Human Vision and Electronic Imaging XV, B. E. Rogowitzand T. N. Pappas, eds.,
Proc. SPIE Vol. 7527, (San Jose, CA), Jan. 2010.

Subband analysis and synthesis of real-world textures for objective and subjective determination of roughness

René Van Egmond,[*a] Thrasyvoulos N. Pappas,[b] Huib de Ridder[a]

[a] Faculty of Industrial Design Engineering, Delft University of Technology, Landbergstraat 15, 2628 CE Delft, The Netherlands

[b] Electrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208, USA

ABSTRACT

In a previous study we investigated the roughness of real world textures taken from the CUReT database. We showed that people could systematically judge the subjective roughness of these textures. However, we did not determine which objective factors relate to these perceptual judgments of roughness. In the present study we take the first step in this direction using a subband decomposition of the CUReT textures. This subband decomposition is used to predict the subjective roughness judgments of the previous study. We also generated synthetic textures with uniformly distributed white noise of the same variance in each subband, and conducted a perceptual experiment to determine the perceived roughness of both the original and synthesized texture images. The participants were asked to rank-order the images based on the degree of perceived roughness. It was found that the synthesis method produces images that are similar in roughness to the original ones except for a small but systematic deviation.

Keywords: CUReT database, texture, roughness, subband decomposition, texture synthesis

1. INTRODUCTION

In our daily life, visual, auditory, or tactile textures provide information about events and objects in the environment. For example, the visual appearance of a road may indicate whether we are about to drive over a smooth or a bumpy surface. The granularity of sandpaper is a good predictor of how it will feel when we touch it. Textures also provide strong clues about the material composition of objects in the environment. For example, we can recognize flour from its visual appearance, or cotton from the way it feels and sounds. The primary focus of this paper is on the perception of visual textures. Many different attributes can be used to describe our experience of visual textures, for example, roughness, sharpness, and pleasantness. However, little is known about what image features may be good predictors of such attributes. In Van Egmond, *et al.*[1] the subjective roughness and pleasantness of images from the CUReT database[2] were determined. It was found that the roughness judgments were systematic and could differentiate between images. The pleasantness judgments, on the other hand, were less systematic and seemed to cluster subjects into two groups, those who found rough images pleasing and those who found rough textures unpleasing. This study takes the next step in trying to find an objective predictor of subjective visual roughness. The proposed predictor is based on a subband decomposition and analysis of the variance of the subband coefficients. In addition, to test the hypothesis that visual roughness can be determined from the variance of the subband coefficients, we synthesize images with the same subband variance and conduct subjective tests to determine if these images evoke the same experience of roughness.

*r.vanegmond@tudelft.nl, phone: +31152788324

1.1. Visual roughness research

Roughness is an attribute that has strong tactile roots. Bergmann, Tiest, and Kappers[3] compare the haptic and visual perception of roughness, while Ho *et al.*[4,5] attempt to relate roughness to surface attributes, direction of illumination, and viewpoint. Another study of how surface parameters affect visual perception of roughness is reported by Padilla, *et al.*[6] However, there has been relatively little work in deriving metrics that can directly estimate visual roughness from images; for example, Sarkar and Chaudhuri[7] examine the relation between visual perception of roughness and fractal dimension.

1.2. Statistical descriptions of visual textures: Deriving a measure for uniformly distributed textures.

Motivated by the texture analysis/synthesis literature,[8,9] we base our texture analysis on multiscale frequency decomposition. Such decompositions have been widely used for acoustic and visual signal analysis and compression, as they provide good approximations of early acoustic and visual processing in mammals. A number of decompositions have been used, including steerable filters,[10] the Gabor transform,[11,12] as well as separable subband and wavelet decompositions. The latter are simpler to implement and have been widely used in image compression and quality evaluation.[13] The main drawback of such separable decompositions is that they cannot discriminate between the two diagonal directions . On the other hand, the filters can be designed to have fairly sharp cutoffs, and hence very little aliasing and interference between the different subbands. In this paper we adopt the generalized quadrature mirror filter bank (GQMF)[14] that was used in the development of the perceptual image coder.[15] (We use an 8x8 GQMF decomposition.) An added advantage of using this filter bank is that direct measurements of subband sensitivity thresholds have been obtained (for certain viewing conditions: 53.6 cycles per degree).

Our goal is to derive a metric of texture roughness. Our first hypothesis was that perceived texture roughness correlates with the distribution of energy in the subbands. As we will see below, our subjective data supports this hypothesis. The question then was how the energy (i.e., variance) in each subband affects the perceived roughness. Another question was whether our objective metric should take into consideration the sensitivity of each subband, that is, whether the variances should be normalized by the subband sensitivity thresholds, or equivalently, expressed as just noticeable distortion (JND) ratios. In the following, we will refer to such normalized variances as perceptually weighted variances.

If indeed the perceived texture roughness depends only on subband variance, then any texture with the same variance should lead to the same perceived roughness. The simplest such texture that one can generate is one that consists of white noise of the given variance in each subband. This ignores higher order subband statistics, such as skewness and kurtosis (third and fourth order statistics, respectively), as well as pixel correlations within and across subbands. Our hypothesis then would that such simple synthesized textures would have the same perceived roughness as the original textures. Note that, since both the original and synthesized textures are perceived by the observer, the perceptual weighting should only be used in the calculation of perceived roughness, not in the texture synthesis.

2. Subjective visual roughness and variances in subbands

In Van Egmond, *et al.*[1] the subjective roughness of 49 images of the CUReT database[2] was determined. The selected set corresponds to lightning and viewing condition 122: the polar angle of the viewing direction was .88, the azimuthal angle of the viewing direction was -2.38, the polar angle of the illumination direction was .88, and the azimuthal angle of the illumination direction was -.76. The subjective roughness was obtained by conducting a pairwise comparison of circular cutouts from the grayscale component of the 49 CUReT textures. The circular shape and the conversion to grayscale were used in order to limit the possibility of identification or even recognition. Therefore, it was suggested that only the structural aspects of these images were evaluated by the participants. However, no analysis was conducted to determine the main objective predictors of subjective roughness. Here, we first analyze these images with a basic statistic, i.e., the variance of the distribution of energy in each subband (i.e., the standard deviation, SDEV). To determine if the SDEV would be a predictive measure, the first step is to determine if there is a systematic relation among the textures using the distribution of energy in the subband decomposition of the CUReT textures.

2.1. Associations between subjective roughness judgments on CUReT textures and perceptually weighted variances of energy in subbands

In order to establish if there is a systematic association between the subjective roughness judgments and the perceptually weighted variances in the subbands of the textures, correlations between the 49 textures were determined on the basis of the perceptually weighted variances in the 64 subbands. The resulting correlations (N=1176) varied between .15 and .998. Ninety percent of the correlations were higher than .50 and 75% was higher than .75. This indicates that the textures are highly correlated on the basis of the perceptually weighted variances in their subbands. We choose one specific example to explain this relation. In Figure 1 the standard deviations of the distributed energy determined for the 64 subbands of CUReT image 47 are shown as a function of the standard deviations of the distributed energy determined for the 64 subbands of CUReT image 17. These textures were chosen because CUReT image 47 was judged as the roughest texture and CUReT image 17 was judged as the smoothest texture (see, Table 1, Van Egmond *et al.*[1]). It can be seen that for each subband there is a systematic increase in the perceptually weighted variance between image 17 and 47. This is confirmed by conducting a linear regression analysis with SDEV of both textures as input. The relation can be expressed by: Texture_47 = -0.71 + 8.52*Texture_17. The explained variance is very high (R^2=.97). In addition it can be seen that especially the lower subbands contribute the most to the relation between these two textures. This observed relation between the lower subbands is very similar across textures. Moreover, taking CUReT image 17 as reference the slope of the regression line systematically increases with the judged roughness of most of the other images. Thus, it may be concluded that an increase of the perceptually weighted variances of the energy distribution in the subbands may be a good predictor for the subjective roughness judgments.

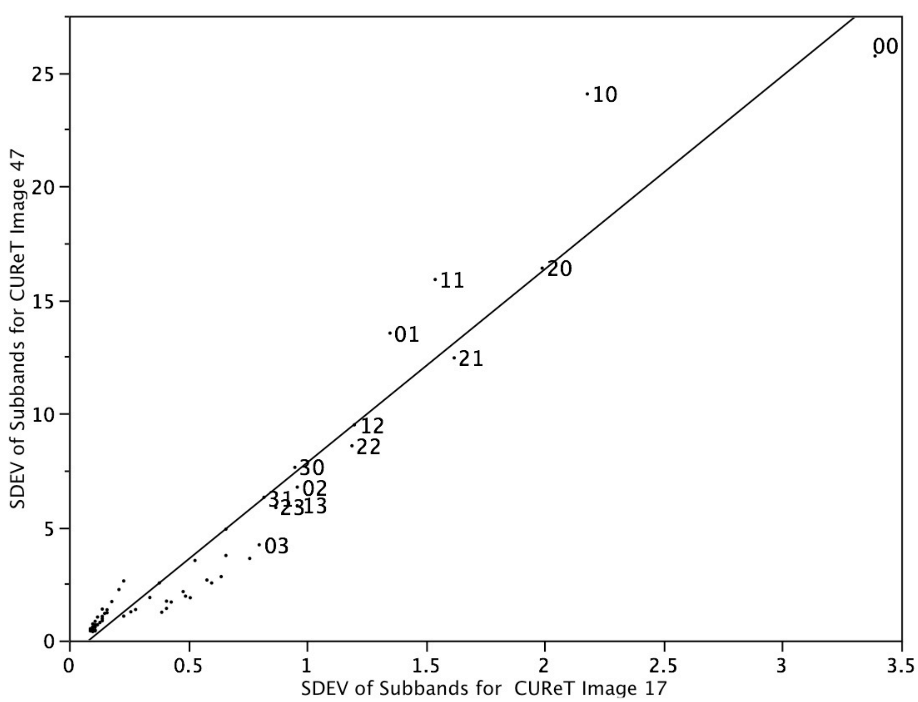

Figure 1. The perceptually weighted variances (SDEV) of 64 subbands for CUReT image 47 (roughest picture, Table 1, van Egmond *et al.*,[1]) as a function of the perceptually weighted variances of 64 subbands for CUReT image (smoothest picture, Table 1, van Egmond *et al.*,[1]). The most relevant subbands have been indicated by their number.

2.2. Perceptually weighted variances in subbands related to subjective roughness

The analysis in Section 2.1 suggested that there is a systematic relation between perceptually weighted variances and subjective roughness. In order to test this, two analyses were performed: a multi-dimensional scaling analysis and a correlational analysis. To reduce the dimensionality of the data (a 49 by 64 matrix) the data were analyzed using multi-dimensional scaling. The data were analyzed using the PROXSAL module in SPSS 18. In the first step of this analysis the similarity distances between the textures were determined using a Euclidean distance measure. In the second step, a scaling model was chosen that transformed distances with an interval proximity transformation with a 2D solution. This 2D solution had a normalized raw stress of .0028 and Dispersion Accounted For of .997. This indicates that the solution represents the data very well.

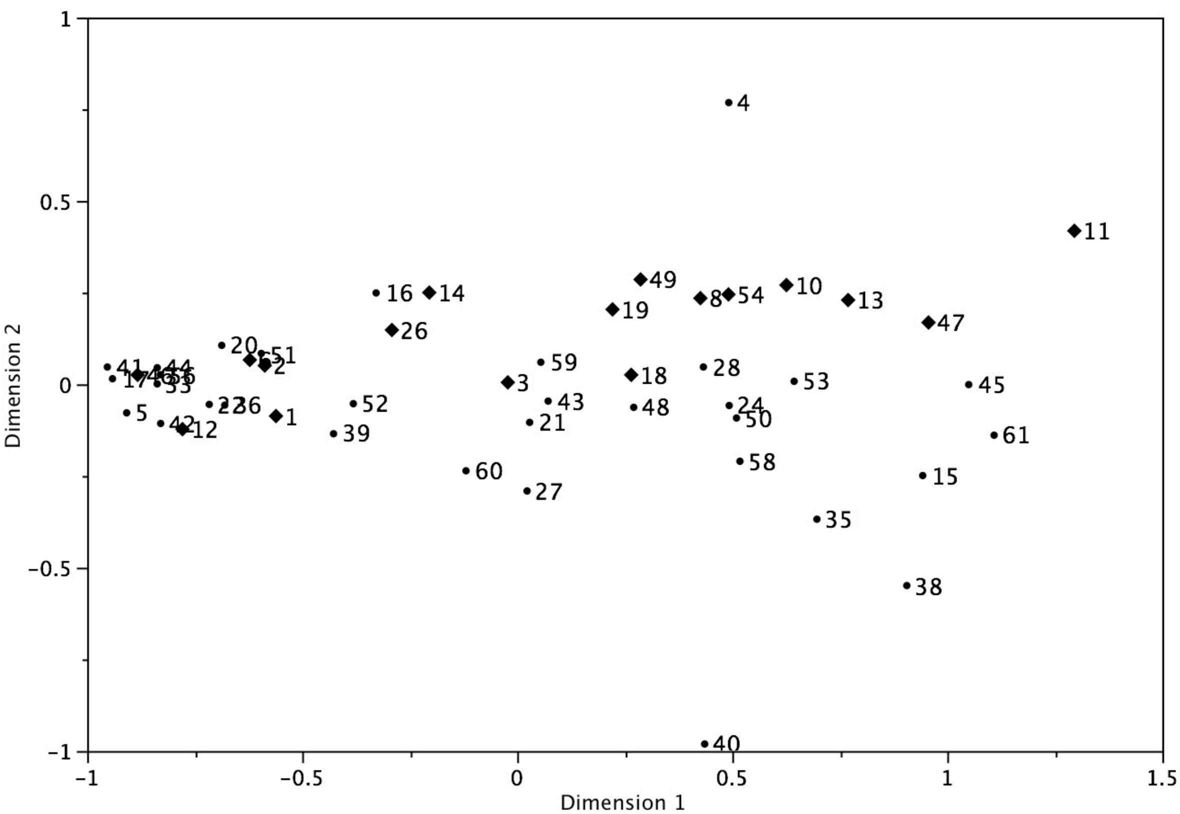

Figure 2. 2D scaling solution of the perceptually weighted variance of the energy in the subbands of CUReT textures. The numbers are the CUReT numbers. Images indicated with a diamond were selected for Experiment 1.

In Figure 2 the 2D scaling solution is shown. It can be readily seen that the textures are systematically distributed along the first dimension (x-axis). On the second dimension (y-axis) the textures are somewhat differentiated but not that strong. This differentiation appears to be the strongest for the textures in the right half of the panel (right from 0.4 on the x-axis). A visual analysis of the textures showed that the roughness of the textures appeared to be increasing from left to right. Thus, the first dimension appears to capture the subjective roughness. A visual analysis of the textures positioned on the second dimension was more difficult to explain. However, it could be suggested that textures positioned low on Dimension 2 show higher order regularities than textures that are positioned high on Dimension 2. In other words, if these textures are differentiated on this dimension they do not show a uniformly distributed texture. In order to determine if Dimension 1 was associated with subjective roughness, correlations between the two dimensions and the subjective roughness judgments (see, Table 1, Van Egmond *et al.*[1]) were determined. The correlation between Dimension 1 and the subjective roughness judgments was very high (r = -.88; note that in Table 1 of Van Egmond, *et al.*[1] the rougher a texture was judged the more negative the scale value was; this explains the negative correlation) and a very low correlation between Dimension 2 and the subjective scale values was found (r = -.05).

In conclusion, we have shown that there is a systematic relation between the subjective roughness judgments and the distribution of energy in subbands. This raises the question that if there is such a strong relation then it may be possible to synthesize images based on the distribution of energy in the subbands that are similar in roughness to the original textures.

3. EXPERIMENT

Seventeen images were selected from the set of 49 images used in Van Egmond *et al.,*[1] They were selected such that they were equally distributed on Dimension 1 of Figure 2 and were not differentiated by Dimension 2. The textures are indicated with diamonds in Figure 2. A set of seventeen synthetic images were also generated, containing uniformly distributed white noise in each subband with variance equal to the variance of the corresponding subband of the corresponding original texture image. The synthesized images had the same mean values as the original images. In four of the synthesized images (1, 2, 6, and 12), the base (lowest horizontal and vertical frequency) band was set to zero, because the original image contained very low frequency variations that, when evenly distributed over the subband, resulted in funny artifacts that did not resemble the original texture.

In Figure 3 sixteen original textures of the CUReT database and the sixteen corresponding synthesized textures are shown. In each panel the top-row displays the original and the bottom-row the synthesized versions. Below each texture the CUReT numbers are presented with extension “o” and “s” representing the original and synthesized textures respectively. It can be seen that the synthesized textures capture the roughness aspects of the original textures very nicely. To test this similarity in roughness a rank-order task with the 34 images was conducted. Participants had to rank-order the textures on the basis of their roughness.

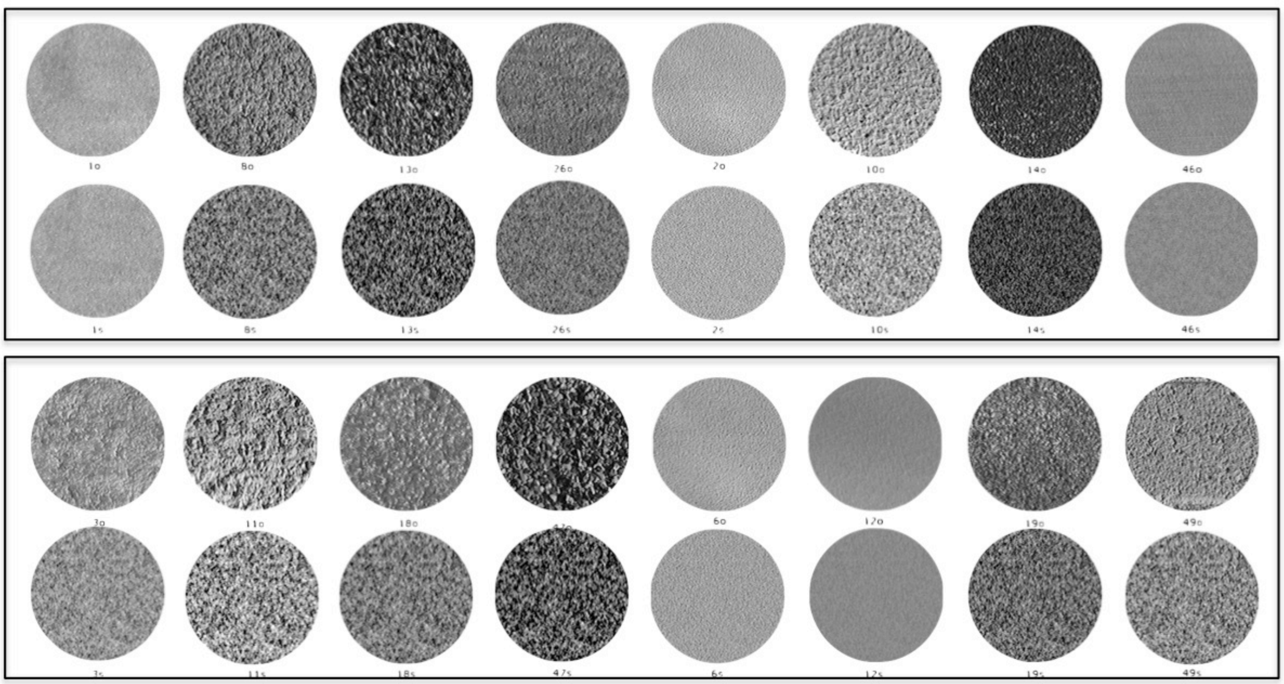

Figure 3. 16 original textures of the CUReT database and 16 synthesized textures created using the energy distributions in the subbands and uniformly distributed white noise. In each panel the top row presents the original textures and the bottom row the synthesized textures. Note that the 17th texture (CUReT number 54) is left out only because of graphical displaying issues.

3.1. Method

Participants had to rank-order 34 images (17 original textures from the CUReT database and 17 synthesized images) on the basis of their perceived roughness.

3.1.1. Participants

Twenty participants (average age *M*=21.4 years) volunteered in this experiment. All were students from the Delft University of Technology and had normal or corrected-to-normal vision.

3.1.2. Stimuli

Seventeen textures from the CUReT database were used. The textures were selected on the basis of their distribution along Dimension 1 of Figure 2 that represented roughness. The lighting and viewing conditions were as follows: the polar angle of the viewing direction was .88, the azimuthal angle of the viewing direction was -2.38, the polar angle of the illumination direction was .88, and the azimuthal angle of the illumination direction was -.76. These were the same conditions as in Van Egmond, *et al.*[1] The original textures were converted to grey-scale and a circular region on the texture was selected. After the subband analysis of the original textures, the synthesized images were generated using the variance of the energy in the subbands and uniformly distributed white noise. The images were printed on laminated paper so that the subjects could rank-order them on a table.

3.1.3. Procedure

The images were placed in random order on a table. A participant was asked to rank-order them on the basis of the perceived roughness. The images were marked and the experimental leader wrote down the rank-order after a participant completed the task.

3.2. Results

The data were analyzed in four steps. First, a 2 dimensional scaling solution of the rank-order data was calculated. Second, these dimensions were then correlated with the dimensions of the 2D scaling solution of the perceptually weighted variances in the subbands for the original and synthesized textures. Third, a 1D scaling solution of the rank-order data was determined in order to establish if there was a systematic difference in judgment of the original and synthesized textures. Fourth, an overall statistical measure of the perceptually weighted subbands was determined and correlated with the roughness judgments represented by the 1D scaling solution.

A multi-dimensional scaling solution was determined for the rank-order data to reduce the dimensionality of the data matrix using PROXSCAL, SPSS 18. The similarity distances between the textures were determined using a Euclidean distance measure. A scaling model was employed that transformed the distances with an ordinal untied proximity transformation using a 2D solution. The 2D solution has a value of Normalized Raw Stress of .0004 and a Dispersion Accounted For of .9996. Although this shows that the 2D solution represents the ranking data very well, the 1D solution showed even a better fit. The 1D solution has a value of Normalized Raw Stress of .0003 and a Dispersion Accounted For of .9997. This means that the 1D solution is the best representative of the roughness ranking judgments. In order to display the actual images of the original and synthesized textures legible, we chose to represent these in the 2-dimensional space. In Figure 4 the original and the synthesized textures are represented on the two dimensions. It can be readily seen that the first dimension is the most important one. A visual analysis shows that the textures appear to become rougher from left to right. The second dimension shows differentiation only at the outer points of Dimension 1. The horseshoe form confirms that in fact the 1D solution is the proper one. The original and synthesized smoother textures (presented on the left side in Figure 4) appear to be situated closer together than the rougher textures. For example, the distances on Dimension 1 between the original and synthesized textures 13 and 18 are relatively large compared to other textures. In addition, one can distinguish visually three main clusters on the outer left side (smooth), on the outer right side (rough) and in the middle (medium rough) of Dimension 1.

In Section 2.2 it was shown that the perceptually weighted variance of the energy in the subbands of the 49 original CUReT textures was highly associated with subjective roughness (note that the subjective roughness measure was obtained in Van Egmond *et al.*[1] by a different paradigm, i.e., pairwise comparison). Therefore, a 2D scaling solution was calculated for the perceptually weighted variance of the energy in the subbands of the 17 original and 17 synthesized textures. The method of scaling was the same as described in Section 2.2. The 2D solution had a dispersion accounted for of .998 and a Normalized Raw Stress of .002. The values on Dimension 1 of the subjective roughness solution correlated highly with those on Dimension 1 of the 2D solution of the weighted variances (r = .93) but low with those on Dimension 2 of the 2D solution of the weighted variances (r = .16). As expected, the values on Dimension 2 of the 2D subjective roughness solution hardly correlated with those on Dimensions 1 and 2 of the 2D solution of the weighted variances (respectively, r = -.004 and r = .21). This indicates that, in line with what we found for the original 49 CUReT images, the perceptually weighted variances of energy in the subbands again explain the subjective roughness judgments.

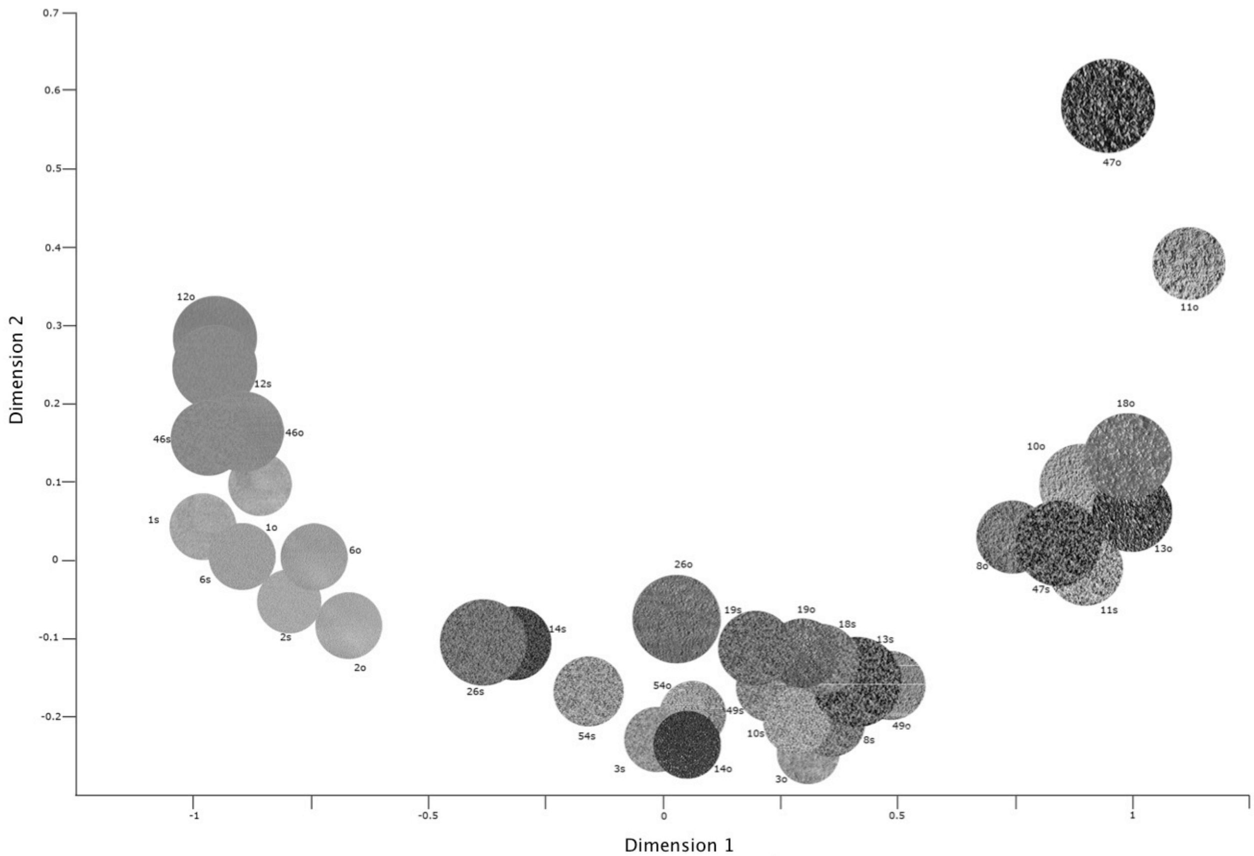

Figure 4. 2 Dimensional scaling solution of ranking data on roughness on synthesized and original (CUReT) textures. The numbers indicate the CUReT number, the "o" the original texture, the "s" the synthesized texture.

Because only one dimension of the 2D scaling solution of the weighted variances appears to be an important measure in explaining the subjective roughness measure—which also is best represented on one dimension—an aggregated measure of the weighted variances was used to predict subjective roughness. The formula for objective roughness as function of the variance of the k^{th} subband σ_k^2 and the corresponding just noticeable distortion threshold t_k (where K is the number of subbands) is:

$$roughness = \left(\frac{1}{K} \sum_{k=1}^{K} w_k \sigma_k^2 \right)^{\frac{1}{2}} \quad where \ \ w_k = \frac{1}{t_k^2}$$

In Figure 5 objective roughness as a function of subjective roughness (1D scaling solution) is presented for the original (left frame) and the synthesized (right frame) textures separately. It can be seen that there is a strong association between the objective roughness measure and the subjective roughness judgments for the original (r = .94) and the synthesized textures (r = .95). In both pictures, the textures 11, 18, 54, and 47 tend to deviate from the overall trend. In general it can be concluded that the objective roughness measure is a good predictor for subjective roughness judgments for both original and synthesized images.

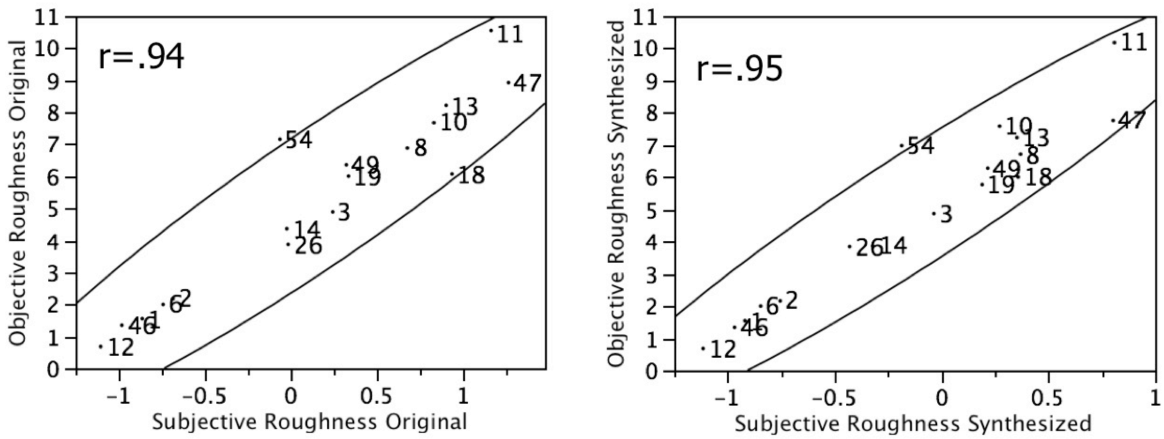

Figure 5. Objective roughness as a function of subjective roughness. In the left frame the values for the original textures are presented; in the right frame the values of the synthesized textures are presented. The density ellipse is the bivariate normal ellipse (p=.95). Numbers indicate the CUReT numbers.

Although for both original and synthesized textures the objective roughness measure showed a strong association with the subjective roughness judgments, a careful look at Figure 4 reveals a systematic difference between original and synthesized textures. To confirm this difference, a regression analysis was conducted with the subjective roughness values of the original textures as a predictor for the subjective roughness values of the synthesized textures (the values from the 1D scaling solution were used). In Figure 6 the subjective roughness values for the synthesized textures are presented as a function of the subjective roughness values for the original textures. It can be clearly seen that there is a strong systematic relation between the values of the original textures and the values of the synthesized textures. This was confirmed by the regression analysis. The regression showed a strong fit: *Synthesized* = -.23 + .78 *Original*, $F(1,15)=457.39$, $p<.0001$; and a high explained variance (R^2=.97). The fact that the slope significantly deviates from one suggests that there is indeed a systematic deviation between the judgments on the original and synthesized textures: the synthesized textures are judged to be less rough than the original ones and this difference increases with the roughness of the original texture. This difference can be explained in part by the fact that the variance of the synthesized images was found to be lower than that of the original images. This was true even though the images were synthesized with the same variances as the original textures. However, imperfections in the filter design account for an energy leakage or cancellation, as replacing the actual subband contents with white noise destroys the perfect aliasing cancellation property of the GQMF subband decomposition. Other factors may also account for this systematic bias, for example, texture synthesis may change the glossiness or other attributes of the textures, which may have an effect on perceived roughness. Indeed, the relation of roughness to other perceptual texture attributes will be examined in future studies.

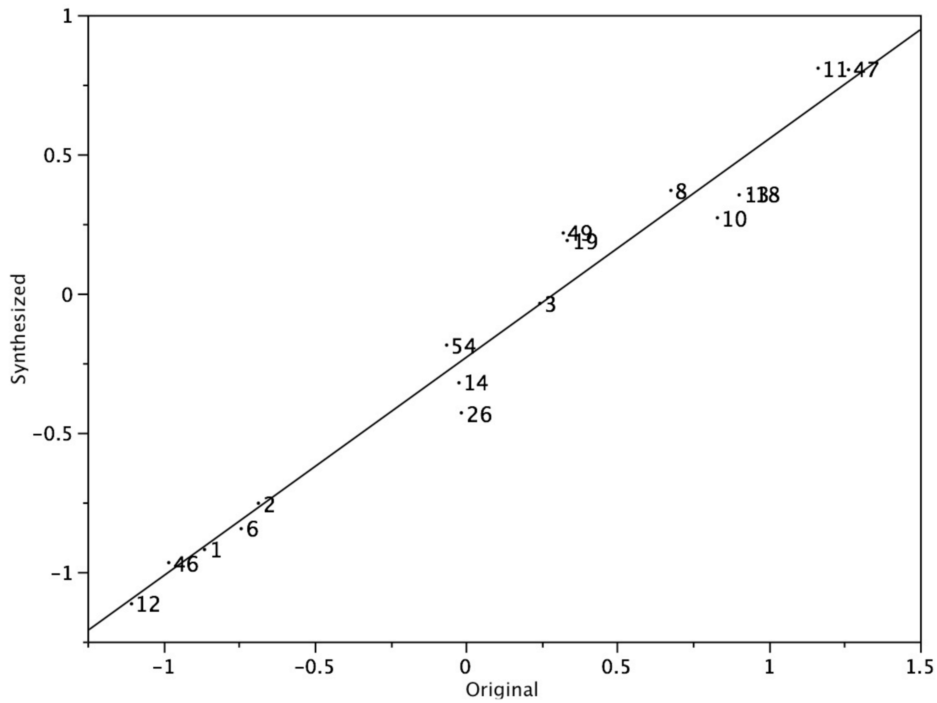

Figure 6. Subjective roughness (1D solution) of synthesized textures as a function of the subjective roughness (1D solution) of the original textures. Numbers identify the CUReT numbers.

4. Discussion

There are two main findings in this study. First, it has been established that subband analysis produces a strong and simple predictor for subjective roughness measures on uniformly distributed textures. This predictor has been based on the distribution of energy in these subbands. Furthermore, this has been confirmed by using two subjective experimental paradigms. In Van Egmond, *et al,*[1] a pairwise comparison task has been employed, and in the present study, a rank-ordering task has been employed. Both experimental paradigms yield the same relationship between subjective judgments and subband analysis. Second, the subjective roughness of the synthesized textures is highly similar to that of the roughness of the original textures. This shows that by using the variation of energy in the subbands and uniform white noise as a source, the experienced roughness can be recreated. This may have important implications for synthesizing textures in virtual environments. However, there appears to be a systematic deviation between the synthesized textures and the original textures. The cause of this difference is a topic of further investigation. Furthermore, note that only uniformly distributed textures have been used, without any directional or other higher order cues. However, it may be interesting to investigate if these higher order aspects can be captured in a simple additive model such as the one we proposed in this paper. In conclusion, it has been shown that by using very simple measures one can effectively predict and generate textures. Our next step will be to investigate if a similarly simple measure may be a good predictor for the structural aspects in textures.

ACKNOWLEDGMENTS

We gratefully acknowledge the work of Kiran Sagoo for conducting the experiments and creating some of the figures. This research was partly made possible by a visiting grant of the Delft University of Technology.

REFERENCES

[1] R. Van Egmond, P. Lemmens, T. N. Pappas, H. De Ridder, "Roughness in sound and vision," in *Human Vision and Electronic Imaging XIV* (B. E. Rogowitz and T. N. Pappas, eds.), SPIE Proc. vol. 7240, (San Jose, CA), pp. 72400K-1-12, Jan. 2009.

[2] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real-world surfaces," *ACM Transactions on Graphics*, vol. 18, pp. 1-34, Jan. 1999.

[3] W. M. Bergmann Tiest and A. M. L. Kappers, "Haptic and visual perception of roughness," *Acta Psychologica,* vol. 124, no. 2, pp. 177-189, 2007.

[4] Y.-X. Ho, M. S. Landy, and L. T. Maloney, "How direction of illumination affects visually perceived surface roughness," *Journal of Vision*, vol. 6, pp. 634-648, 2006.

[5] Y.-X. Ho, L.T. Maloney, and M.S. Landy, "The effect of viewpoint on perceived visual roughness," *Journal of Vision*, vol. 7, no. 1, pp. 1-16, 2007.

[6] S. Padilla, O. Drbohlav, P. R. Green, A. Spence, and M. J. Chantler, "Perceived roughness of 1/f[beta] noise surfaces," *Vision Research*, vol. 48, no. 17, pp. 1791-1797, 2008.

[7] N. Sarkar and B. B. Chaudhuri, "An efficient approach to estimate fractal dimension of textural images," *Pattern Recognition*, vol. 15, no. 9, pp. 1035-1041, 1992.

[8] D. J. Heeger and J. R. Bergen, "Pyramid-based texture analysis/synthesis," in *Proc. Int. Conf. Image Processing (ICIP-95)*, vol. III, (Washington, DC), pp. 648-651, Oct. 1995.

[9] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int. J. Computer Vision* , vol. 40, pp. 49-71, Oct. 2000.

[10] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in *Proc. ICIP-95, vol. III*, (Washington, DC), pp. 444-447, Oct. 1995.

[11] J. G. Daugman and D. M. Kammen, "Pure orientation filtering: A scale invariant image-processing tool for perception research and data compression," *Behavior Research Methods, Instruments, and Computers*, vol. 18, no. 6, pp. 559-564, 1986.

[12] M. Porat and Y. Y. Zeevi, "Localized texture processing in vision: Analysis and synthesis in Gaborian space," *IEEE Trans. Biomed. Eng.,* vol. 36, no. 1, pp. 115-129, 1989.

[13] T. N. Pappas, R. J. Safranek, and J. Chen, "Perceptual criteria for image quality evaluation," in *Handbook of Image and Video Processing* (A. C. Bovik, ed.), pp. 939-959, Academic Press, second ed., 2005.

[14] R. V. Cox, "The design of uniformly and nonuniformly spaced pseudoquadrature mirror filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 1090-1096, Oct. 1986.

[15] R. J. Safranek and J. D. Johnston, "A perceptually tuned sub-band image coder with image dependent quantization and post-quantization data compression," in *Proc. ICASSP-89*, vol. 3, (Glasgow, Scotland), pp. 1945-1948, May 1989.