A Texture Retrieval Scheme Based on Perceptual Features

Yanhai Gan† Xiaoxu Cai† Jun Liu* Shengke Wang†
* Qingdao Agricultural University, Qingdao, China
† Ocean University of China, Qingdao, China
E-mail: gyh5421@163.com, caixiaoxu90@163.com, liujunqd@163.com, neverme@ouc.edu.cn

***Abstract*—Procedural textures have been widely used as they can be easily generated from various mathematical models. However, the model parameters are not perceptually meaningful or uniform for non-expert users; therefore it is difficult for general users to obtain a desired texture by tuning the parameters. In order to satisfy users' requirement, we propose a novel procedural texture retrieval scheme that can return textures according to commonly used perceptual dimensions. We establish a procedural texture database that includes abundant textures so as to meet the diverse demands of users. All textures in the database are projected into a perceptual space after we construct the mapping model. First, we investigate the salient features of the input texture; then we calculate the Euclidean distance between the input texture and each texture in the database. Experimental results show that our method can effectively retrieve textures that are perceptually consistent with users' input.**

I. Introduction

Procedural models have been widely used in the field of computer graphics for generating procedural textures. However, procedural models are defined by mathematical algorithms [9]; therefore it is difficult to determine the parameters of procedural models to generate desired textures, even for experienced users. For tackling this challenge, we interpret the problem as a retrieval task, i.e., for a given texture, we would like to find procedural textures which are similar to the example in terms of visual perception.

A typical content-based image retrieval (CBIR) consists of two major parts, feature extraction and similarity measurement. First, a set of features are extracted to represent the content of each image; second, a distance measure between features of the query image and each image in the database is calculated so that the most similar image is retrieved. However, the features are typically low-level features, and similarity measurement based on these features always leads to results that may be quite different from human perception.

To describe textures, humans usually use perceptual features, such as "directional", "repetitive" and "structural" [8]. In addition, similar textures substantially have the same salient perceptual features in common among them. Thus, texture retrieval schemes based on perceptual features may yield better retrieval performance. Moreover, similarity measurement based on perceptual features makes the retrieval process more efficient.

In this paper, we investigate two problems as mentioned above. A procedural texture retrieval scheme based on perceptual features is proposed, which may find textures perceptually similar to the example. The highlights of this paper are as follows, (1) representative features are learned by a deep neural network-PCANet efficiently; and the features are used to train mapping models for predicting perceptual scales of the query texture; (2) the prominent perceptual features of the query texture can be determined automatically, which are consistent with human perception; (3) a novel perceptual texture retrieval scheme is proposed, which can find textures that are perceptually similar to the query texture, and the proposed scheme supports a variety of procedural models.

II. Related Works

As more visual information is stored in digital format, image databases are becoming more and more popular. In particular, procedural textures have been widely used in many fields. There exist many texture databases which usually contain abundant textures. A texture database is supposed to provide a simple and timely manner for users to interact with the database. Therefore an efficient CBIR [10] scheme becomes an essential ingredient of a meaningful texture database.

For texture representation, a variety of features have been proposed, e.g., local binary pattern (LBP) [11] and wavelet-like features [3] [7], which are the most popular features used in retrieval systems. Meanwhile, deep neural networks have behaved more powerful in recent years [4]. Recently, a principle component analysis network (PCANet) [1] has been applied in texture classification and object recognition, and achieved great performance success, while requiring much less computation compared to other state-of-the-art deep learning approaches. PCANet simplifies the procedure to learn the convolutional filters, meanwhile, it retains the hierarchical architecture of traditional convolutional neural networks (CNNs) [5] [2]. However, these computational features have no direct relation to human perception even though they are representative and discriminative. What's more, these computational features are always high-dimensional. Similarity measurement based on these computational features might be time consuming, and derive results that depart from human perception. We jointly consider the feature extraction and similarity measurement, and build a mapping from computational features to perceptual features.

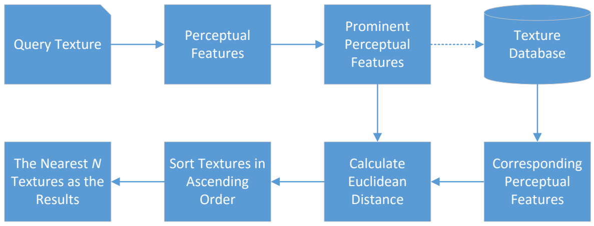

Fig. 1. The architecture of the proposed retrieval scheme.

III. The Proposed Texture Retrieval Scheme

The overall architecture of the texture retrieval scheme is illustrated in Fig. 1. The retrieval scheme consists of two steps. First, we investigate the prominent perceptual features of the query texture. The prominent perceptual features indicate several visual properties particularly perceived by humans at the pre-attentive level of texture perception.

Second, the prominent feature space is constructed. The dimensions of the space are determined by the number of the prominent perceptual features of the query texture. Then, we search the nearest neighbours to the query texture by using the Euclidean distance in the prominent feature space.

A. Extraction of computational features

In our retrieval scheme, we employ a deep neural network-PCANet to learn features from the training textures. PCANet is a variant of traditional deep convolutional neural networks. It has achieved great success in texture classification and object detection. Different from CNNs, PCANet fascinates by its simplified structure and efficient training procedure. The network consists of cascaded feature extraction stages and a non-linear output stage. For each feature extraction stage, principle component analysis (PCA) is applied to calculate the convolutional filters. Several feature extraction stages are concatenated to form a deep network, in which the outputs of preceding extraction stage are propagated as the inputs of the posterior stage. The non-linear output stage is implemented by a binary hashing process and histogram statistics. Benefiting from the unsupervised training process and the efficiency of PCA, PCANet learns features more quickly than conventional deep networks. Taking into consideration of all of the above mentioned aspects, we have PCANet employed in our method to extract computational features for each texture in our database.

B. Mapping Computational Features to Perceptual Features

Given a query texture, the key issue is how to predict the perceptual features. A set of perceptual features that contains 12 perceptual properties [1] defined in [8] is used in our scheme.

Since the computational features of each texture in our database have been extracted, we can train the mapping models from computational features to perceptual features. Because the perceptual features are psychometric scales used to measure the extent of certain visual perceptual properties belonging to a sample, we believe that the scales for each perceptual feature of a sample can also be regarded as a class label. That is to say texture samples can be classified according to their perceptual scales. The task of mapping is then transformed to classification tasks. For one certain perceptual feature, we can classify the texture samples to nine clusters according to their perceptual scales. Support vector machine (SVM) is employed for classification with computational features as the training data and perceptual scales as the class labels. A total of 12 classifiers are trained and each classifier is responsible for a perceptual feature mapping task. Afterwards, we can predict the perceptual scales for each input texture. The overall structure of the mapping model is illustrated in Fig. 2.

C. Identifying Prominent Perceptual Features

Prominent perceptual features of the query texture are defined as perceptual features whose scales are larger than a threshold. We assume that each texture possesses no more than five prominent perceptual features. For each texture, if there are more than five perceptual features whose scales are larger than the threshold, we increase the threshold by one at a time until the number of perceptual features whose scales are above the threshold is no more than five. In this way, prominent perceptual features are determined and coincide with human perception.

IV. Experiments and Results

A. Datasets

The training data we used is the same as in [6] [13], which contains 450 surface textures of size 512×512 pixels and corresponding perceptual scales. The surface textures in the dataset were generated by twenty three representative procedural models. The details of these procedural models can be found in [6]. We cropped each texture into 4 non-overlapping sub-images of size 256×256 pixels, which resulted in a total of 1800 texture samples. Because samples in our database were isotropic textures, we believed that the sub-images had the same perceptual scales as the original texture. The cropping operation increased training samples, which improved the accuracy of the mapping model, meanwhile, the training process was accelerated by the reduced dimensions.

We further enlarged the database by generating more textures by varying parameters for the 23 procedural models. In the end, 31150 additional textures were generated. The appearance of these textures varied in visual properties, e,g., contrast, repetitiveness, roughness and so on. Each texture was rendered using Luxrender [12] under the same area lighting conditions and diffuses reflectance as in [6]. Some examples of the surface textures are shown in Fig. 3. We cropped each of the 31150 new produced textures into four non-overlapping sub-images as mentioned before, which produced 124600 textures. It was impossible to obtain the perceptual scales for each texture in such a large database by psychophysical experiment,

[1] The features are, in order, contrast, repetition, granularity, randomness, roughness, feature density, directionality, structural complexity, coarseness, regularity, locally orientation and uniformity.

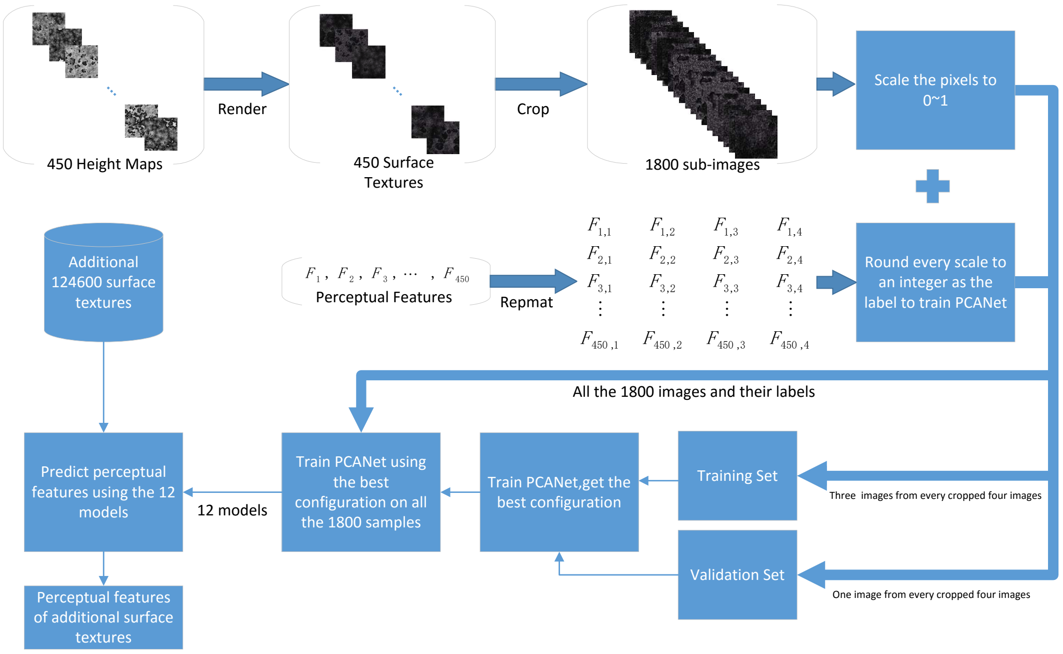

Fig. 2. Procedure for predicting the perceptual features of additional surface textures.

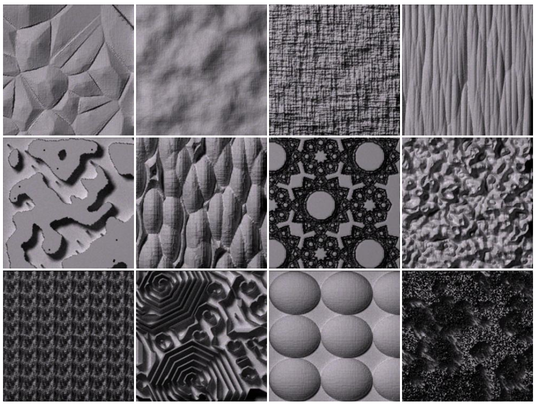

Fig. 3. Ten examples of surface textures.

so the perceptual scales of the new 124600 samples were approximately predicted by the mapping models. Finally our database extended to 126400 procedural textures, and each of them owned its perceptual scales.

B. Retrieval Results

First, we evaluate the performance of the mapping models. We used 1800 textures and their perceptual scales to train the mapping models. The validation set contained 450 surface textures randomly chosen from the 1800 textures. The rest of the textures were used as the training set. The parameters of PCANet were chosen by validation. We had demonstrated that two-stage PCANet was in general sufficient to achieve good performance. Thus, in our experiments, two-stage PCANet architecture was employed. After validation, in PCANet, the filter size was $k_1 = k_2 = 5$, the number of filters $L_1 = L_2 = 8$, and block size 64×64. Since we had found the best configuration, we trained PCANet and the 12 classifiers using all the 1800 samples including the training and validation set. Table I summarizes the classification accuracies.

Classification results based on PCANet features suggest that we can predict perceptual scales of each texture in our database accurately.

TABLE I
COMPARISON OF ACCURACY(%) AMONG LBP, GABOR, GABOR+LBP, AND PCANET FEATURES FOR DIFFERENT PERCEPTUAL FEATURE PREDICTION TASK.

Perceptual Features	Accuracy(%)			
	LBP	Gabor	Gabor+LBP	PCANet
contrast	87.99	91.91	92.72	97.33
repetitive	85.70	86.44	89.38	97.45
granular	82.11	85.95	87.99	97.10
random	87.58	87.99	90.77	97.15
rough	87.09	87.83	91.42	97.45
feature density	87.01	90.28	92.16	97.51
direction	87.66	90.93	93.14	97.98
structural complexity	84.86	86.44	89.95	96.28
coarse	88.07	88.89	90.93	95.92
regular	85.29	88.15	91.42	97.82
oriented	86.93	90.11	92.97	97.38
uniform	86.52	89.46	91.09	96.82
average	86.40	88.70	91.16	97.18

Second, we evaluated the performance of the retrieval scheme. Given a procedural texture, we predicted its perceptual scales and investigated its prominent perceptual features. After constructing the prominent perceptual feature space, we calculated the Euclidean distance between the query and each texture in our database and took the top N closest textures as the retrieval results. Fig. 4 illustrates some of the results

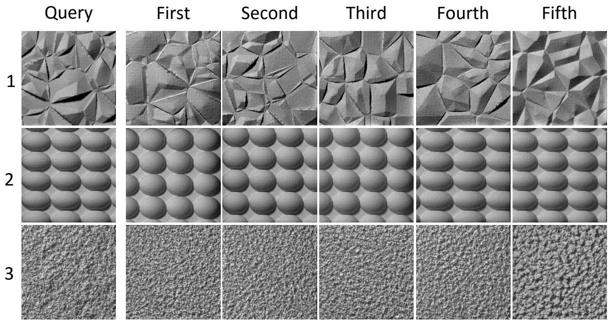

Fig. 4. Results of the retrieval scheme based on perceptual features. In the same row, texture on the left is the query and textures on the right are corresponding results arrayed in ascending order by their distances to the query.

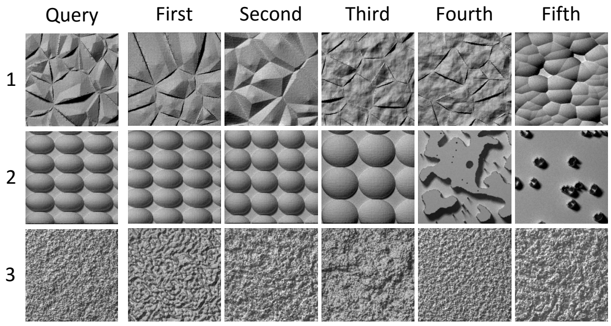

Fig. 5. Retrieval results by using Gabor features. In the same row, texture on the left is the query and textures on the right are corresponding results arrayed in ascending order by their distances to the query.

of our retrieval scheme. We made $N = 5$ here to implement the retrieval process. It is obvious that most of the results are perceptually consistent with the queries. We also compared the results to that achieved by using Gabor features. The retrieval results by using Gabor features are shown in Fig. 5, in which the same textures are used as the queries. We can see that the retrieved textures by using Gabor features have similar pixel distributions over space, but they depart from human perception. Only random textures can be retrieved precisely by using Gabor features. These results demonstrate the effectiveness of our method. In addition, the predicted perceptual scales of the query textures are shown in Table II.

TABLE II
PREDICTED PERCEPTUAL FEATURES OF INPUT SURFACE TEXTURES. SCALES IN BOLD CORRESPOND TO PROMINENT PERCEPTUAL FEATURES OF GIVEN TEXTURES.

Perceptual Features	Scales		
	Texture 1	Texture 2	Texture 3
contrast	3	**7**	5
repetitive	4	**8**	4
granular	3	3	2
random	**6**	2	5
rough	**5**	5	5
feature density	**5**	5	3
direction	4	5	**7**
structural complexity	4	3	3
coarse	**6**	**6**	5
regular	3	**9**	4
oriented	4	3	**7**
uniform	4	4	5

V. CONCLUSION

In this paper, we propose a method to retrieval textures that are in accordance with human perception. In order to achieve this purpose, we implement our retrieval scheme by estimating prominent perceptual features. Because psychometric perceptual features are expensive to obtain, a mapping model is constructed to estimate the perceptual scales for each surface texture in our database. By searching the nearest neighbours in the prominent feature space, retrieval results that are perceptually consistent with users' input are obtained.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation Of China (NSFC) under Grant No. 61271405, 61401413 and 61301241, the Ph.D. Program Foundation Of Ministry Of Education Of China under Grant No. 20120132110018, the International Science & Technology Cooperation Program of China (ISTCP) under Grant No. 2014DFA10410, the Fundamental Research Projects of Qingdao Science and Technology Plan under Grant No. 12-1-4-1-(8)-jch, the Shandong Science and Technology Development Plan Projects under Grant No. 2012GHY11524, the Natural Science Foundation of Shandong under Grant No. ZR2014FQ023.

REFERENCES

[1] Chan, T.H., Jia, K., Gao, S., Lu, J., Zeng, Z., Ma, Y.: Pcanet: A simple deep learning baseline for image classification? arXiv preprint arXiv:1404.3606 (2014)
[2] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. pp. 1097–1105 (2012)
[3] Laine, A., Fan, J.: Texture classification by wavelet packet signatures. Pattern Analysis and Machine Intelligence, IEEE Transactions on 15(11), 1186–1191 (1993)
[4] LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature pp. 436–444 (2015)
[5] LeCun, Y., Bengio, Y.: Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks 3361(10) (1995)
[6] Liu, J., Dong, J., Cai, X., Qi, L., Chantler, M.: Visual perception of procedural textures: Identifying perceptual dimensions and predicting generation models. PloS one 10(6) (2015)
[7] Manjunath, B.S., Ma, W.Y.: Texture features for browsing and retrieval of image data. Pattern Analysis and Machine Intelligence, IEEE Transactions on 18(8), 837–842 (1996)
[8] Rao, A.R., Lohse, G.L.: Towards a texture naming system: identifying relevant dimensions of texture. In: Visualization, 1993. Visualization'93, Proceedings., IEEE Conference on. pp. 220–227. IEEE (1993)
[9] Ross, B.J., Zhu, H.: Procedural texture evolution using multi-objective optimization. New Generation Computing 22(3), 271–293 (2004)
[10] Smeulders, A.W., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. Pattern Analysis and Machine Intelligence, IEEE Transactions on 22(12), 1349–1380 (2000)
[11] Suruliandi, A., Meena, K., Reena Rose, R.: Local binary pattern and its derivatives for face recognition. Computer Vision, IET 6(5), 480–488 (2012)
[12] Team, L.: Luxrender v0. 8 (2012)
[13] Zheng, Y., Zhong, G., Liu, J., Cai, X., Dong, J.: Visual texture perception with feature learning models and deep architectures. In: Pattern Recognition, pp. 401–410. Springer (2014)