Fast BTF Texture Modelling

Abstract—This paper presents a novel fast model-based algorithm for realistic multispectral BTF texture modelling potentially capable of direct implementation inside the graphical card processing unit. The algorithm starts with range map estimation of the BTF texture followed by the spectral and spatial factorization of an input multispectral texture image. Single orthogonal monospectral band-limited factors are independently modelled by their dedicated Gaussian Markov random field models (GMRF). We estimate an optimal contextual neighbourhood and parameters for each GMRF. Finally single synthesized band-limited factors are collapsed into the fine resolution monospectral images and using the inverse Karhunen-Loeve transformation we obtain the smooth multispectral texture. Both multispectral and range information is combined in a bump mapping filter of the rendering hardware. The presented model offers huge BTF texture compression ration which cannot be achieved by any other sampling-based BTF texture synthesis method.

I. INTRODUCTION

Photo realism in virtual or mixed reality scenes cannot be accomplished without nature-like colour textures covering visualized scene objects. These textures can be either smooth or rough (also refered as the bidirectional texture function - BTF). The rough textures which have very rough surfaces do not obey Lambertian law and their reflectance is illumination and view angle dependent as it is illustrated in Fig. 1. Textures can be either digitised natural textures or textures synthesized from an appropriate mathematical model. The former simplistic option suffers among others with extreme memory requirements for storage of a large number of digitised cross-sectioned slices through different material samples. Sampling solution become unmanageble for rough textures which require to store thousands of different illumination and view angle samples for every texture.

Several intelligent sampling methods [1], [2], [3], [4], [5] were proposed with the aim to diminish these huge memory requirements. The method [1] constructs texture in coarse-to-fine fashion, preserving conditional distribution of filter outputs over multiple scales, while another multiscale method [4] uses histograms of filter responses. The quilting method [3] is based on the overlapping tiling and subsequent minimum error boundary cut. Similarly the algorithm [5] uses regular tiling combined with a deterministic chaos transformation. All these methods are based on some sort of original small texture sampling and the best of them produce very realistic synthetic textures. However these methods require to store thousands images for every combination of viewing and illumination angle of the original target texture sample, they often produce visible seams, some of them are computationally demanding and they cannot generate textures unseen (unstored) by the algorithm.

Synthetic textures are far more flexible, extremely compressed (few parameters have to be stored only), they may be evaluated directly in procedural form and can be designed to meet certain constraints or properties, so that they can be used to fill an infinite texture space without visible discontinuities.

Several monospectral smooth texture modelling approaches were published, e.g., [6], [7],[8], among them also few colour models, e.g., [9], [10], [11], and some survey articles are available [12],[13] as well. Previous colour texture modelling methods either mapped colour to gray tones [14] sacrificing considerable amount of image information or used 3D models [10],[11] with corresponding problems of robust nonlinear estimation of large amount of model parameters.

Modelling multispectral images requires three dimensional models however if we are willing to sacrifice some information a 3D model can be approximated with a set of much simpler 2D models without compromising its visual realism. Among such possible models the Gaussian Markov random fields are appropriate for texture modelling not only because they do not suffer with some problems of alternative options (see [9],[12],[13],[15] for details) but they are also easy to synthesize and still flexible enough to imitate a large set of natural and artificial textures. While the random field based models quite successfully represent high frequencies present in natural textures low frequencies are much more difficult for them. One possibility how to overcome this drawback is to use a multiscale random field model.

Multiple resolution decomposition (MRD) such as Gaussian/Laplacian pyramids, wavelet pyramids or subband pyramids [16] present efficient method for the spatial information
comparing. The hierarchy of resolutions provides a transition between pixel-level features and region or global features and hence to model a large variety of possible textures. Unfortunately Markov random fields in general, and Gaussian Markov random fields in particular are not invariant to multiple resolution decomposition (MRD) even for simple MRD like subsampling and the lower-resolution images generally lose their Markovianity. Fortunately we can avoid computationally demanding approximations [17] of a non-Markov multigrid random field by Markov random fields because there is no need to transfer information between single spatial factors hence it is sufficient to analyze and synthesize each resolution component independently.

We propose a novel algorithm for efficient rough texture modelling which combines an estimated range map with synthetic multiscale Markov random field based generated smooth texture. The texture visual appearance during changes of viewing and illumination conditions simulated using the bump mapping technique. The obvious advantage of this solution is the possibility to use hardware support of bump map techniques in contemporary visualization hardware.

The rest of the paper is organized as follows. After a description of range map estimation technique, we discuss the underlaying multiscale Markovian smooth texture model together with its parameter estimation and synthesis solutions. Results are reported in Section 3, followed by a conclusion and discussion about future work in Section 4.

II. RANGE MAP ESTIMATION

There are several methods for determining a surface range map. The most accurate range map can be estimated by direct measurement of the observed surface using corresponding range cameras (e.g. laser or structured light based). However this method requires special hardware and measurement methodology. Hence several alternative approaches for range map estimation from surface images were developed. One of them is a Photometric Stereo which estimates surface range map from at least three images obtained for different position of illumination source while the camera position is fixed. Its obvious disadvantage is the requirement to obtain at least three mutually registered images. An alternative option is to use a method from the Shape from Shading group.

The range map estimation of texture images within this paper was performed by shape from shading algorithm published by Frankot&Chellappa [18]. This method exploits the fact that image intensity \( Y_r \) of each pixel observed in texture image is given according to reflectance map \( R \):

\[
Y_r = R(z_{r1}, z_{r2}, L, C, \rho) \tag{1}
\]

where \( z_{r1} \) and \( z_{r2} \) are surface slopes in directions \( r_1 \) and \( r_2 \), \( L \) is vector from surface to illumination source, \( C \) is vector from surface to the camera and \( \rho \) is albedo of observed material. The advantage of this method is fast range map estimation from single surface image illuminated by single light source from known position. In comparison with other methods, as for example Horn&Brooks [19], this algorithm includes constrain which enforces integrability of surface slopes (2). These slopes are then updated in each iteration step.

\[
\frac{\partial^2 Z_r}{\partial r_1 \partial r_2} = \frac{\partial^2 Z_r}{\partial r_2 \partial r_1} \tag{2}
\]

where \( Z_r \) is the unknown surface height in the location \( r = (r_1, r_2) \). This integrability condition is performed in the Fourier coefficient representation as follows:

\[
C(u, v) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} Z_r \cdot e^{j(ur_1 + vr_2)} dr_1 dr_2
\]

\[
Z_r = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} C(u, v) \cdot e^{-j(ur_1 + vr_2)} dudv
\]

Based on these assumptions the reconstruction of surface height is performed in the Fourier space with coefficients obtained by the above equation. Frankot&Chellappa proved [18] that the algorithm maps closed convex sets into closed convex sets in each iteration step and thus the estimated surface slopes are integrable. Finally the inverse transformation leads to a surface range map.

III. SMOOTH TEXTURE MODEL

Modelling general multispectral (e.g., colour) texture images requires three dimensional models. If a 3D data space can be factorized then these data can be modelled using a set of less-dimensional 2D random field models, otherwise it is necessary to use some 3D random field model. Although full 3D models allows unrestricted spatial-spectral correlation modelling its main drawback is large amount of parameters to be estimated and in the case of Markov models (MRF) also the necessity to estimate all these parameters simultaneously [10]. The factorization alternative is attractive because it allows using simpler 2D data models with less parameters (one third in the three-spectral case of colour textures).

Spectral factorization using the Karhunen-Loeve expansion transforms the original centered data space \( \tilde{Y} \) defined on the rectangular \( M \times N \) finite toroidal lattice \( I \) into a new data space with K-L coordinate axes \( \tilde{Y} \). This new basis vectors are the eigenvectors of the \( d \times d \) second-order statistical moments matrix \( \Phi = E(\tilde{Y}_{r,\bullet} \tilde{Y}_{r,\bullet}^T) \) where the multiindex \( r \) has two components \( r = [r_1, r_2] \), the first component is row and and the second one column index, respectively. The projection of random vector \( \tilde{Y}_{r,\bullet} \) (the notation \( \bullet \) has the meaning of all possible values of the corresponding index) onto the K-L coordinate system uses the transformation matrix \( T = [u_1^T, \ldots, u_d^T]^T \) which has single rows \( u_j \) that are eigenvectors of the matrix \( \Phi \).

\[
\tilde{Y}_{r,\bullet} = T\tilde{Y}_{r,\bullet} \tag{3}
\]

Components of the transformed vector \( \tilde{Y}_{r,\bullet} \) (3) are mutually uncorrelated and if \( \tilde{Y}_{r,\bullet} \) are Gaussian they are also independent hence each transformed monospectral factor can be modelled independently of remaining spectral factors. Texture
modelling does not require computationally demanding MRD approximations (e.g. [17]) because it does not need to propagate information between different data resolution levels. It is sufficient to analyze and subsequently generate single spatial frequency bands without assuming a knowledge of some MRF-type global multi-grid model. Input multispectral image is factorized using (3) into $d$ monospectral images $Y_{s,i}$ for $i = 1, \ldots, d$. These components are further decomposed into a multi-resolution grid and each resolution data are independently modelled by their dedicated GMRF. Each one generates a single spatial frequency band of the texture. An analysed texture is decomposed into multiple resolutions factors using Laplacian pyramid and the intermediary Gaussian pyramid $\tilde{Y}_{s,i}$ which is a sequence of images in which each one is a low-pass down-sampled version of its predecessor. The Gaussian pyramid for a reduction factor $n$ is

$$\tilde{Y}_{s,i}^{(k)} = \downarrow^n Y_{s,i}^{(k-1)} \otimes u \quad k = 1, 2, \ldots ,$$

where $\tilde{Y}_{s,i}^{(0)} = Y_{s,i}$ ,

$\downarrow^n$ denotes down-sampling with reduction factor $n$ and $\otimes$ is the convolution operation. The Laplacian pyramid $\tilde{Y}_{r,i}^{(k)}$ contains band-pass components and provides a good approximation to the Laplacian of the Gaussian kernel. It can be constructed by differencing single Gaussian pyramid layers:

$$\tilde{Y}_{r,i}^{(k)} = \tilde{Y}_{r,i}^{(k-1)} \uparrow^n (\tilde{Y}_{s,i}^{(k+1)}) \quad k = 0, 1, \ldots ,$$

where $\uparrow^n$ is the up-sampling with an expanding factor $n$.

Single orthogonal monospectral components are thus decomposed into a multi-resolution grid and each resolution data are independently modeled by their dedicated Gaussian Markov random field (GMRF) as follows. The Markov random field (MRF) is a family of random variables with a joint probability density on the set of all possible realisations $Y$ of the lattice $I$, subject to following conditions:

$$p(Y_{s,i}) > 0, \quad \forall Y ,$$

and

$$p(Y_{r,i} | Y_{s,i} \forall s \in I \setminus \{r\}) = p(Y_{r,i} | Y_{s,i} \forall s \in I_{r,i}) ,$$

and $I_{r,i}$ is a contextual support set of the $i$-th monospectral field.

If the local conditional density of the MRF model (8) is Gaussian, we obtain the continuous Gaussian Markov random field model (GMRF):

$$p(Y_{r,i} | Y_{s,i} \forall s \in I_{r,i}) =$$

$$(2\pi\sigma_i^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_i^2}(Y_{r,i} - \bar{\mu}_{r,i})^2\right\} ,$$

where the mean value is

$$E\{Y_{r,i} | Y_{s,i} \forall s \in I \setminus \{r\}\} = \bar{\mu}_{r,i} =$$

$$\mu_{r,i} + \sum_{s \in I_{r,i}} a_{s,i} (Y_{r-s,i} - \mu_{r-s,i})$$

and $\sigma_i, a_{s,i} \forall s \in I_{r,i}$ are unknown parameters. The 2D MRF model can be expressed as a stationary non-causal correlated noise driven 2D autoregressive process:

$$\tilde{Y}_{r,i} = \sum_{s \in I_{r,i}} a_{s,i} \tilde{Y}_{r-s,i} + e_{r,i}$$

where the noise $e_{r,i}$ is random variable with zero mean. The $e_{r,i}$ noise variables are mutually correlated

$$R_{e_i} = E\{e_{r,i}e_{r-s,i}\}$$

$$= \begin{cases} \sigma_i^2 & \text{if } s = (0,0), \\ -\sigma_i^2 a_{s,i} & \text{if } s \in I_{r,i}, \\ 0 & \text{otherwise}. \end{cases}$$

Correlation functions have the symmetry property $E\{e_{r,i}e_{r+s,i}\} = E\{e_{r,i}e_{r-s,i}\}$ hence the neighbourhood support set and their associated coefficients have to be symmetric, i.e. $s \in I_{r,i} \Rightarrow -s \in I_{r,i}$ and $a_{s,i} = a_{-s,i}.$

A. Parameter Estimation

The selection of an appropriate GMRF model support is important to obtain good results in modelling of a given random field. If the contextual neighbourhood is too small it can not capture all details of the random field. Inclusion of the unnecessary neighbours on the other hand add to the computational burden and can potentially degrade the performance of the model as an additional source of noise. We use hierarchical neighbourhood system $I_{r,i},$ e.g., the first-order neighbourhood is $I_{r,i} = \{r - (0,1), r + (0,1), r - (1,0), r + (1,0)\}$, etc. An optimal neighbourhood is detected using the correlation method [9] favouring neighbours locations corresponding to large correlations over those with small correlations.

Parameter estimation of a MRF model is complicated by the difficulty associated with computing the normalization constant. Fortunately the GMRF model is an exception where the normalization constant is easy obtainable however either Bayesian or ML estimate requires iterative minimization of a nonlinear function. Therefore we use the pseudo-likelihood estimator which is computationally simple although not efficient. The pseudo-likelihood estimate for $a_{s,i}$ parameters has the form

$$\gamma_{a,i} = \left[a_{s,i} \forall s \in I_{r,i}\right] =$$

$$\left[\sum_{r \in I} X_{r,i}^T X_{r,i}^{-1}\right] \sum_{r \in I} X_{r,i}^T Y_{r,i} ,$$

(12)

where $X_{r,i} = [Y_{r-s,i} \forall s \in I_{r,i}]$ and

$$\sigma_i^2 = \frac{1}{MN} \sum_{r=1}^{MN} (Y_{r,i} - \gamma_{a,i} X_{r,i}^T)^2 .$$

(13)

B. Model Synthesis

Several possibilities exist [9] for a finite lattice GMRF synthesis. The most effective synthesis method uses the discrete fast Fourier transformation. GMRF can be generated [8] from $Y_{s,i} = F^{-1}\{\tilde{Y}_{s,i}\} + U_i,$ where $U_i$ the mean vector of the
whole field and \( \hat{Y}_{r,i} \) is generated from the Gaussian generator \( N(0, N_{MSY}(r,i)) \). \( S_Y(r,i) \) is the associated power spectrum [12] and \( N \times M \) is the underlying generated lattice size. Single GMRF models synthesize spatial frequency bands of the texture. Each monospectral fine-resolution component is obtained from the pyramid collapse procedure (inversion process to (4),(5)). Finally the resulting synthesized multispectral texture is obtained from the set of synthesized monospectral images using the inverse K-L transformation:

\[
\hat{Y}_{r} = T^{-1}\hat{Y}_{r,i}
\]  

(14)

IV. RESULTS

We have tested the algorithm on BTF colour textures from the CUReT [20] database, several our and the University of Bonn BTF measurements. Fig.2 shows a corduroy, upholstery and knitwear textures synthesized using the presented method for three illumination angles. For comparison, the included corresponding smooth texture synthesis results using the same sets of models demonstrate clear improvement of the presented approach. All these figures demonstrate also the colour quality of the model. Our synthesized images manifest comparable colour quality with the much more complex 3D models [11]. The multi-scale models demonstrate [9] their clear superiority over their single-scale counterparts while the colour quality is comparable between single-scale and multi-scale alternative models. All presented experiments use the bump map technique approximation for combining range and intensity data. The synthesised intensity image is identical for all viewing angles but differs for illumination angle changes.

V. SUMMARY AND CONCLUSIONS

Our extensive testing results of the algorithm are encouraging. Some synthetic textures reproduce given digitized texture images so that both natural and synthetic texture are visually indiscernible. Even the not so successful results can be used for the preattentive BTF textures applications. The overwhelming amount of original colour tones were reproduced realistically in spite of restricted spectral modelling power of the model. The multi-scale approach is more robust and allows far better results than the single-scale one if the synthesis model is inadequate (lower order model, non stationary texture, etc.). The proposed method allows huge compression ratio (unattainable by alternative intelligent sampling approaches) for transmission or storing texture information while it has still moderate computation complexity. The method does not need any time-consuming numerical optimization like for example the usually employed Markov chain Monte Carlo methods. The replacement of the bump filter technique with the novel graphics cards feature - the displacement mapping is expected to further significantly improve the visual quality of our algorithm results.

The presented method is based on the estimated model and as such it can only approximate realism of the original measurement. However it offers unbeatable data compression ratio (tens of parameters per texture only), easy simulation of even non-existing (previously not measured) BTF textures and fast seamless synthesis of any size texture.

ACKNOWLEDGEMENTS

This research was supported by the EC project no. IST-0000-00000 XXXXX grant No. and partially by the grant .

REFERENCES

Fig. 2. Three examples of synthetic rough (BTF) textures (corduroy, textile, knitwear) using the presented method. The upper rows in each example contain original textures ($\theta_i = 60^\circ$) illuminated with tilt angles 0, 90 (100) and 180 degrees, respectively and the range map obtained from the original. The bottom rows contain the corresponding synthesized rough textures (the rightmost image is one of smooth synthesised results using the same model).