# Fast Synthesis of Dynamic Colour Textures

Jiří Filip Michal Haindl
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
182 08 Prague 8, Czech Republic
{filipj,haindl}@utia.cas.cz

Dmitry Chetverikov
Computer and Automation Research Institute
Hungarian Academy of Sciences
H-1111 Budapest, Hungary
csetverikov@sztaki.hu

## Abstract

*Textural appearance of many real word materials is not static but shows progress in time. If such a progress is spatially and temporally homogeneous these materials can be represented by means of* dynamic texture *(DT). DT modelling is a challenging problem which can add new quality into computer graphics applications. We propose a novel hybrid method for colour DTs modelling. The method is based on eigen-analysis of DT images and subsequent preprocessing and modelling of temporal interpolation eigencoefficients using a causal auto-regressive model. The proposed method shows good performance for most of the tested DTs, which depends mainly on the properties of the original sequence. Moreover, this method compresses significantly the original data and enables extremely fast synthesis of artificial sequence, which can be easily performed by means of contemporary graphics hardware.*

## 1 Introduction

Dynamic or temporal textures (DT) can be defined as spatially repetitive motion patterns exhibiting stationary temporal properties and have also indeterminable spatial and temporal contents. Water surface, fire or straw in the wind can serve as typical DT examples. A video sequence as basic representation of DTs has finite duration. This feature limits the use of DTs in virtual reality systems of any kind so temporally unconstrained modelling of DT is a challenging problem concerning such research areas as computer vision, pattern recognition and computer graphics.

According to the application area we can classify [1] published works on DT to recognition, representation or synthesis, respectively. The DT synthesis is obviously the most difficult one and there are only few papers on this topic available, which can be roughly divided in three major categories.

The first category of DT synthesis approaches use statistical auto-regressive models. Szummer in [11] used spatio-temporal causal auto-regressive model for DT modelling. His implementation has high computational demands and assumes stationarity of the input DT sequence. In the work of Soatto [10] there is proposed a DT model based on auto-regressive moving average process. The model is applied on responses of dimensionality reduction filter based on SVD. However, the model use time consuming iterative gradient descent method for parameters estimation and, similarly to the previous approach it is restricted to mono-spectral DTs only.

The second category of methods is based on synthesis of parametric transformation of original data. In [12] there is presented a generative mono spectral DT model based on moving object structure modelling and their trajectories by means of dictionary containing Gabor bases for particle elements and Fourier bases for wave elements. The synthesis of short DT sequence using this method takes several minutes. Similar problem occurs in the approach of Joseph [4] based on a combination of spatial steerable pyramid and temporal wavelet transformation.

Finally, the last category of DT synthesis methods use video editing techniques. Schödl et al. [9] suggested method generating "video textures", i.e., DTs of arbitrary length. The method is based on searching for transition points where the video can be looped back on itself additionally with blending and morphing techniques. This method shows good results for many colour video sequences, not necessarily textures, however works only if a pair of similar-looking frames can be found, which can be problem when observing fluids such as fire, smoke etc. The further extension of this idea was done by Kwatra et al. [5] who represented DT by several 3D spatio-temporal texture patches with mutually optimal spatio-temporal seams. The seams are computed in temporal neighbourhood of estimated transition points. This method enables fast synthesis of new DT, with relatively high storage requirements of the computed patches. Wei [13] used tree-structured vector quantization for DT synthesis. Although such a synthesis is visually almost perfect, the algorithm require tens of second per one frame. Modelling of DTs using pixel-wise polynomials was mentioned in [6].

The main goal of this paper is to propose straightforward colour DT modelling method with low computational demands enabling extremely fast synthesis of arbitrarily long DT sequence. The method, illustrated on Fig.1, is based on combination of input data dimensionality reduction using eigen-analysis (Section 2) and modelling of resulted temporal coefficients (Section 3) by means of autoregressive model. Section 4 describes synthesis of new DT sequence while Section 5 illustrates and discusses results of the method with the conclusions in Section 6.

## 2 Dynamic Texture Eigen-Analysis

RGB pixels of individual images from the DT sequence are arranged into normalised column vectors forming matrix $\mathbf{C}$ $(n \times t)$ where $n$ is number of pixel values $n = 3MN$ depending on the image resolution $M \times N$, $t$ is a number of DT colour learning frames $(t \ll n)$ and $\mu_{\mathbf{C}}$ is a mean image of the sequence. From the matrix $\mathbf{C}$ a covariance matrix $\mathbf{A}$ $(t \times t)$ is created including spatial and spectral correlation of the DT sequence according to $\mathbf{A} = \mathbf{C}^T\mathbf{C}$. The resulted matrix $\mathbf{A}$ is decomposed using singular value decomposition $\mathbf{A} = \mathbf{UDU}^T$ [2] where $\mathbf{U}$ is orthogonal matrix of eigen-vectors and $\mathbf{D}$ is diagonal matrix of corresponding eigen-numbers sorted in ascending order. From matrix $\mathbf{U}$ only a number $k$ of eigen-vectors fulfilling $k < t \ll n$ is preserved in the matrix $\tilde{\mathbf{U}}$ corresponding to eigen-numbers bearing the most of the information (see Tab.1). Using $\hat{\mathbf{U}} = \mathbf{C}\tilde{\mathbf{U}}\tilde{\mathbf{D}}$, where $\tilde{\mathbf{D}} = diag\{\sigma_1^{-\frac{1}{2}}, \ldots, \sigma_k^{-\frac{1}{2}}\}$, we obtain the matrix of eigen-images $\hat{\mathbf{U}}$ ordered into $k$ columns of the length $n$. Finally the course of temporal mixing coefficients of individual eigen-images $\hat{\mathbf{U}}$ for all frames from the original DT sequence is computed using $\mathbf{M} = \hat{\mathbf{U}}^T\mathbf{C}$. Only the matrix $\mathbf{M}$ $(k \times t)$ is a subject of further processing and modelling as it is explained in the following sections.

## 3 Temporal Features Analysis

The matrix $\mathbf{M}$ contains time behaviour of eigen-image temporal coefficients of the original DT sequence length. However, the data of such a length (typically 250 frames) are insufficient to learn our statistical model. Moreover the data often show spatial discontinuity between successive images in DT sequences of very fast processes, and consequently between corresponding temporal coefficients in $\mathbf{M}$. For this purpose we performed the interpolation of individual temporal coefficients separately by means of cubic splines [8]. This simple technique enables us to obtain arbitrarily smooth interpolation between columns of $\mathbf{M}$, i.e., individual frames of the original DT. As a result of this preprocessing step generating $s$ additional frames between each pair of original ones we obtain the enlarged matrix $\hat{\mathbf{M}}$ of size $k \times L$ where $L = s(t-1) + t$.

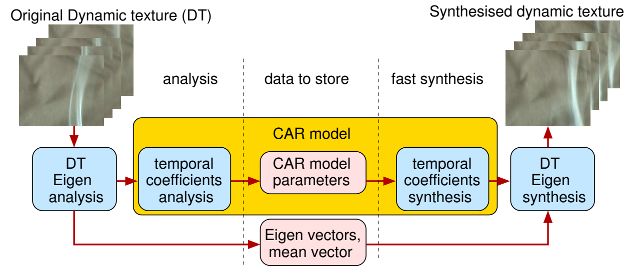


**Figure 1. Scheme of the proposed dynamic texture hybrid model.**

Modelling of all smoothed temporal coefficients of the matrix $\hat{\mathbf{M}}$ was done simultaneously by means of Gaussian noise driven causal auto-regressive (CAR) model. The advantage of the CAR models is that they can be solved, under several additional and acceptable assumptions, analytically. Let the matrix $\hat{\mathbf{M}}$ represents a finite rectangular two-dimensional underlying lattice $I$ $(k \times L)$. Let us denote a $r$ as index in time axis of matrix $\hat{\mathbf{M}}$. The CAR random field is a family of random variables with a Gaussian probability density on the set of all possible realizations $\mathbf{Y}$ of the lattice $I$, subject to several conditions [3]. The CAR model can be expressed as a stationary causal uncorrelated noise driven autoregressive process:

$$\mathbf{Y}_r = \Gamma_{(r-1)}\mathbf{X}_r + \mathbf{E}_r \ , \tag{1}$$

where $\Gamma$ $(k \times kh)$ is the parameter matrix with $h$ denoting the length of the causal neighbourhood $I_r^c$, i.e., how many positions in history is taken into account for creating of design data vector during each shift of $r$ on lattice $I$, $\mathbf{E}_r$ is a Gaussian white noise vector with zero mean and a constant but unknown covariance matrix $\Sigma$ $(k \times k)$ and $\mathbf{X}_r$ is a corresponding vector of $\mathbf{Y}_{r-h}$ collecting previously generated data from $I_r^c$.

Parameter estimations (2),(4) of the CAR model using the Bayesian method and the normal-Wishart parameter prior can be found analytically. The CAR model parameter matrix [3] is

$$\hat{\Gamma}_{r-1}^T = \mathbf{V}_{\mathbf{XX}(r-1)}^{-1}\mathbf{V}_{\mathbf{XY}(r-1)} \tag{2}$$

where the used notion is $\tilde{\mathbf{V}}_{\Delta\Omega(r-1)} = \sum_{i=1}^{r-1} \delta_i \omega_i^T$ ,

$$\tilde{\mathbf{V}}_{r-1} = \begin{pmatrix} \tilde{\mathbf{V}}_{\mathbf{YY}(r-1)} & \tilde{\mathbf{V}}_{\mathbf{XY}(r-1)}^T \\ \tilde{\mathbf{V}}_{\mathbf{XY}(r-1)} & \tilde{\mathbf{V}}_{\mathbf{XX}(r-1)} \end{pmatrix} \ , \tag{3}$$

where $\mathbf{V}_{r-1} = \tilde{\mathbf{V}}_{r-1} + \mathbf{V}_0$ matrix $\mathbf{V}_0$ is the matrix derived from the normal-Wishart parameter prior [3].
The Gaussian noise variance matrix is defined as

$$\hat{\Sigma}_{r-1} = \Lambda_{(r-1)}/\beta(r) \ , \tag{4}$$

where $\beta(r)$ represents number of model movements in $\hat{\mathbf{M}}$ $\beta(r) = \beta(0) + r - 1$ , $\beta(0) > 1$ and

$$\Lambda_{(r)} = \mathbf{V}_{\mathbf{YY}(r)} - \mathbf{V}^T_{\mathbf{XY}(r)} \mathbf{V}^{-1}_{\mathbf{XX}(r)} \mathbf{V}_{\mathbf{XY}(r)} \ . \qquad (5)$$

## 4 Dynamic Texture Synthesis

The CAR model synthesis is very simple, new temporal mixing coefficients of individual eigen-images can be directly generated from the model equation (1) using the model parametric matrix $\hat{\Gamma}$ $(k \times hk)$. and a multivariate Gaussian generator with estimated noise variance $\hat{\Sigma}$ $(k \times k)$. A new DT frame $\hat{\mathbf{C}}_r$ (vector of size $n$) is obtained as a linear interpolation of $k$ individual eigen-images $\hat{\mathbf{U}}$ according to synthesised temporal coefficients $\mathbf{Y}_r = [y_{r,1}, \ldots, y_{r,k}]$ synthesised from the CAR model

$$\hat{\mathbf{C}}_r = [\hat{\Gamma}_{(r-1)} \mathbf{X}_r]^T \hat{\mathbf{U}} + \mu_{\mathbf{C}} = \mathbf{Y}_r \hat{\mathbf{U}} + \mu_{\mathbf{C}} \ . \qquad (6)$$

Both the synthesis of new temporal coefficients and the following eigen-images interpolation can be done even faster using features of contemporary graphics hardware. Moreover, this technique enables significant compression of the original DT data, typically of the ratio between $\frac{1}{5}$ and $\frac{1}{10}$, depending on the length and the character of DT sequence. After several thousand of synthesised steps the DT frames can converge to a mean image of the original sequence or become unstable. To solve this problem the original estimated model parameters $\hat{\Gamma}, \hat{\Sigma}$ can be iteratively reloaded for the infinite DT synthesis.

## 5 Results and Discussion

As a source of test data we used the dynamic texture data sets from DynTex texture database [7] and one data set from MIT [11]. Each DT sample is typically represented by 250 RGB images. For testing purposes we down-sampled these data to resolution 200×150. As a test sequences were chosen *fire*, *boil*, *escalator*, *smoke* and *straw*.

Robust and reliable similarity comparison between two static textures is still unsolved problem up to now. Moreover, when we switch to the dynamic textures the complexity of comparison between original and synthesised DT sequence increase even more. We proposed two statistical DT similarity measures: either original DT images with synthesised ones (**A**) or the original underlying temporal coefficients with those synthesised (**B**).

For (**A**) we computed pixel-wise mean value and variance both resulting into corresponding mean and variance images of original and synthesised DT (see Fig.2). Additionally we compared averaged means and variances of 24 Gabor features (6 orientations and 4 scales, filter variance $\sigma = 1.2$ pixels) through all DT images in CIE Lab space of original and synthesised sequence as it is shown in


**Figure 2. The mean (rows 1,2) and variance (rows 3,4) images comparison of original (rows 1,3) and synthesised DT (rows 2,4).**

Tab.2. The table shows differences of averaged Gabor features (means and variances) between the first and the second part of the original DT ($O$), between original DT and results of its eigen-analysis ($E$), and between original and synthesised DT ($S$). From the table we can see that the average difference between Gabor coefficients caused by dimensionality reduction was significantly higher than additional difference caused by synthesis of the resulted temporal coefficients. For (**B**) we computed averaged mean value

**Table 1. Difference of the averaged mean and variance for all original temporal coefficients of DT ($E$) and their synthesis ($S$).**

| | model params. | | | $D_{\mu(\mathbf{M})}$ | | $D_{\sigma(\mathbf{M})}$ | |
|---|---|---|---|---|---|---|---|
| DT | $k$ | $h$ | $s$ | $\frac{1}{2}$-$\frac{2}{2}E$ | $E$-$S$ | $\frac{1}{2}$-$\frac{2}{2}E$ | $E$-$S$ |
| *fire* | 18 | 2 | 10 | 3.91 | 40.82 | $1.1 \cdot 10^7$ | $1.5 \cdot 10^7$ |
| *boil* | 25 | 1 | 22 | 2.23 | 22.56 | $4.8 \cdot 10^4$ | $3.3 \cdot 10^5$ |
| *escal.* | 25 | 1 | 5 | 25.92 | 14.73 | $1.1 \cdot 10^6$ | $1.1 \cdot 10^6$ |
| *smoke* | 25 | 1 | 7 | 9.08 | 46.45 | $7.1 \cdot 10^5$ | $5.9 \cdot 10^5$ |
| *straw* | 17 | 2 | 34 | 3.35 | 6.26 | $1.5 \cdot 10^5$ | $3.4 \cdot 10^5$ |

and variance through all temporal coefficients with results figured in Tab.1. The table shows that average mean and variance values of spline interpolation of PCA temporal coefficients $\hat{\mathbf{M}}$ ($E$) and their synthesis ($S$) are similar. The only significant difference for *boil* DT was probably caused by fitting of AR model to the different temporal frequencies than those presented in the original DT which are actually very fast causing clear spatial discontinuities between individual frames. This table includes also optimal numbers of eigen-components $k$, length of model history $h$ and interpolated extra frames $s$ for individual DTs. The comparison of original DT frames with those being synthesised is shown in Fig.3. Distance between frames of individual DTs is chosen to show the most of the DT's dynamics and the 5th synthesised image shows always the 300. synthesised frame.

Analysis time of the original DT was ∼3 minutes. Synthesis of a new DT sequence is very fast ∼60 frames/s using SW implementation on the PC Athlon 2GHz.

**Table 2. Averaged Gabor features statistics comparing original DT frames ($O$), eigen-analysed ($E$) and synthesised ($S$) frames.**

| | $D_{\mu(\mathbf{C})}$ | | | $D_{\sigma(\mathbf{C})}$ | | |
|---|---|---|---|---|---|---|
| DT | $\frac{1}{2}$-$\frac{2}{2}O$ | $O$-$E$ | $O$-$S$ | $\frac{1}{2}$-$\frac{2}{2}O$ | $O$-$E$ | $O$-$S$ |
| *fire* | 0.86 | 1.53 | 1.60 | 1.29 | 2.22 | 2.63 |
| *boil* | 0.45 | 4.32 | 3.23 | 0.46 | 4.66 | 3.37 |
| *escalator* | 0.57 | 1.22 | 1.58 | 0.59 | 0.93 | 1.52 |
| *smoke* | 1.37 | 2.05 | 2.47 | 1.25 | 2.01 | 2.08 |
| *straw* | 0.40 | 1.91 | 2.16 | 0.33 | 1.73 | 1.98 |

## 6 Conclusions

A novel method for fast synthesis of dynamic multi-spectral textures was proposed. The method is based on probabilistic modelling of temporal coefficients resulted from input data dimensionality reduction step. The proposed approach enables extremely fast synthesis of arbitrary number of multi-spectral frames, which can be efficiently performed by contemporary graphical hardware. Moreover, the proposed analytical estimation of model parameters is very fast and this approach enables DT synthesis in variable temporal scales.


## Acknowledgements

This research was supported by the EC project FP6-507752 MUSCLE, grants No. 1M0572 (DAR), A2075302 and 1ET400750407. The authors wish to thank to M. Huiskes of the Centre for Mathematics and Computer Science, Amsterdam for providing us DynTex database.


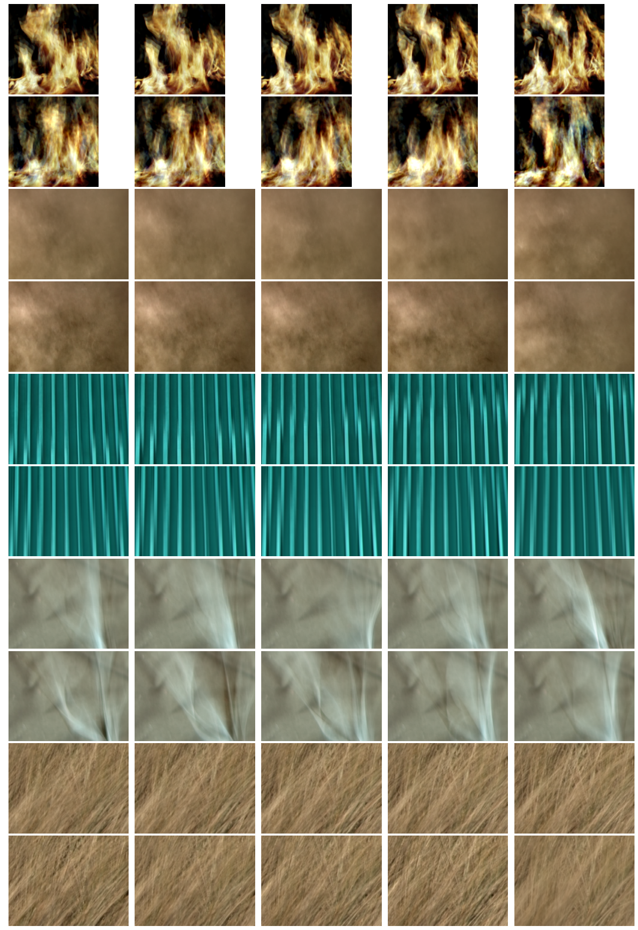

**Figure 3. Examples of frames from original DT (odd rows) and the corresponding synthesised frames using the proposed model (even rows) for five of tested DTs.**

## References


[1] D. Chetverikov and R. Péteri. A brief survey of dynamic texture description and recognition. In *Proceedings 4th Int. Conf. on Computer Recognition Systems (CORES'05)*, pages 17–26. Springer Advances in Soft Computing, 2005.

[2] G. H. Golub and C. F. Van Loan. *Matrix Computations, (3rd ed.)*. Johns Hopkins University Press, Baltimore, MD, 1996.

[3] M. Haindl and S. Šimberová. *Theory & Applications of Image Analysis*, chapter A Multispectral Image Line Reconstruction Method, pages 306–315. World Scientific Publishing Co., Singapore, 1992.

[4] Z. Joseph, R. El-Yaniv, D. Lischinski, and M. Werman. Texture mixing and texture movie synthesis using statistical learning. *IEEE Transactions on Visualization and Computer Graphics*, 7(2):120–135, 2001.

[5] V. Kwatra, A. Schödl, I. Essa, and A. Bobick. Graphcut textures: image and video synthesis using graph cuts. *ACM SIGGRAPH 2003, ACM Press*, 22(2):277–286, July 2003.

[6] T. Malzbender, D. Gelb, and H. Wolters. Polynomial texture maps. In *ACM SIGGRAPH 2001, ACM Press*, pages 519–528. Eurographics Association, Switzerland, 2001.

[7] R. Péteri and M. Huiskes. A comprehensive database of dynamic textures. *URL:http://www.cwi.nl/projects/dyntex/*.

[8] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recepies in C*. Cambridge University Press, 1992.

[9] A. Schödl, R. Szeliski, D. Salesin, and I. Essa. *ACM SIGGRAPH 2000, ACM Press*, pages 489–498, July 2000.

[10] S. Soatto, G. Doretto, and Y. Wu. Dynamic textures. In *ICCV*, pages 439–446, 2001.

[11] M. Szummer and R. Picard. Temporal texture modeling. In *Proceedings of IEEE International Conference on Image Processing (ICIP 96)*, 1996.

[12] Y. Wang and S. Zhu. Analysis and synthesis of textured motion: Particle, wave and cartoon sketch. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, 2004.

[13] L. Wei and M. Levoy. Fast texture synthesis using tree-structured vector quantization. *ACM SIGGRAPH 2000, ACM Press*, pages 479–488, 2000.