BAYESIAN RENDERING WITH NON-PARAMETRIC MULTISCALE PRIOR MODEL

Max Mignotte

DIRO, Département d'Informatique et de Recherche Opérationnelle, CP 6128,
Succ. Centre-Ville, P.O. 6128, Montréal (Québec), H3C 3J7.

HTTP : WWW.IRO.UMONTREAL.CA/∼MIGNOTTE/
E-MAIL : MIGNOTTE@IRO.UMONTREAL.CA

ABSTRACT

This paper investigates the use of the Bayesian inference for devising an example-based rendering procedure. As prior model of this Bayesian inference, we exploit the multiscale non-parametric model recently proposed by Wei et al. *for texture synthesis. This model appears to be interesting to also capture* some *characteristics of a rendering style from an artistic illustration example. Obtained results, with a prior model capturing the rendering style of drawing samples or trained with synthetic and real input textures, are presented. Our results indicate that the proposed method allows to simulate automatic synthesis of various illustration style. More generally, the proposed scheme is able to re-render an input image in the style of an other image allowing, in this way, to create a very broad range of artistic and visual effects.*

1. INTRODUCTION

In the past few years, there has been a great deal of research work in creating artistic styles by computer. This field of research is known as Non-Photorealistic Rendering (NPR). These methods refer to any image processing techniques which can transform an image into a specific artistic styles (or more precisely that *suggest* an artistic style) such as ink painting, pointillist style, engraving, stylized halftoning, charcoal drawing, etc., or, more generally, in a style other than realism. All of the methods proposed by the computer graphics community are generally specifically tailored to a particular rendering style and/or are purely algorithmic techniques (see [1] for a excellent review of existing NPR techniques).

Recently, Efros *et al.* have proposed an interesting non parametric model to synthesize a texture of arbitrary size from a given texture sample. Contrary to classical parametric models (often based on Markov Random Field (MRF) Models), their approach, improved by a multiscale sampling technique introduced by Wei *et al.* [2], is simple, intuitive and works surprisingly well even for a wide variety of complex textures. The synthesized textures (cf. Fig. 1) remain visually similar to the original and seem to come from the same (underlying) generative process as the original input sample. This model allows to efficiently capture the local and global visual characteristics of a complex input texture and thus seems also well suited to capture *some* characteristics of a particular rendering style from an artistic illustration example.

Subject to finding a probabilistic formulation for the Wei's model and to express it as a prior knowledge, Bayesian statistical theory is a convenient tool to take this *a priori* knowledge into consideration for a rendering procedure. Applied with success, to image segmentation, restoration [3], surface reconstruction [4], visual motion analysis [5], etc., the Bayesian methodology allows to efficiently combine a prior model that expresses the prior characteristics of the solution to be estimated (i.e., the image to be re-rendered in our application) with a statistical model describing the interactions between this solution and an input observation (or a drawing sample). This Bayesian methodology appears thus to be well suited to re-render an input image in the style of an other image.

This paper is organized as follows. In Section 2, we provide a brief overview of the model introduced by Efros *et al.* and we find a probabilistic formulation to their sampling procedure. Section 3 and 4 exploit this model as prior model in a Bayesian rendering procedure and present the coarse-to-fine recursive energy minimization problem related to this technique respectively. Finally, obtained results, with a prior model capturing the output rendering style of drawing samples or learned on texture samples, are presented.

2. PREVIOUS WORK AND PRIOR MODEL

Recently Efros *al.* [6] have proposed a non parametric model to synthesize an output image from a given input texture sample. Their approach works surprisingly well for a wide range of textures. Their algorithm consists of the following steps ;

- The output image is initialized with a random noise whose histogram is equalized with respect to the one of the input sample.
- For each pixel in the output image, in scan-line order, the already-synthesized values in neighborhood of current pixel z of a specific (fixed) size ($N(z)$) is considered and is compared with all possible neighborhoods $N(x)$ from input sample. The value of the input pixel x with the most similar $N(x)$ is then assigned to z.

The similarity of the two considered neighborhoods is computed according to the L_2 norm. In order to obtain good synthesis results, the size of the neighborhood must be large enough to capture the scale of the texture elements of the input sample. Although this searching process is deterministic, the random aspect of each synthesized texture is ensured by the initial random noise (first causing the boundary pixels to be assigned semi-stochastically and then using these first assignments as a seed). In this method, the input texture is modeled as a MRF. This assumes that the Probability Distribution Function (PDF) of brightness values for a pixel depends only of the brightness values of its spatial neighbors. Based on this MRF model, a stochastic sampling would consist in constructing an approximation of $P(z/N(z))$ for a given $N(z)$ (e.g., by using histogramming technique over the input textural sample) and then sample from it with a Gibbs or Metropolis sampler [3]. Due to the exponential growth of amount of data required to empirically estimate a d-dimensional PDF when d (number of neighbors) increases, this estimation technique is not possible in the case of an input sample of small size. Unlike MRF-based stochastic sampling, Efros's algorithm is completely deterministic and avoids explicit conditional probability construction and stochastic sampling. By using adequately sized neighborhood, the algorithm produces surprisingly good results but remains very slow since a full search of the input image is required to synthesize every pixel. This computational load is all the more important that the initial texture contains large scale structure and we have to use large neighborhoods.

To cope with these problems, Wei *et al.* [2] have proposed a multiscale synthesis scheme involving a pyramidal structure of synthesized images and a multiresolution (Gaussian) pyramid for the set of input images. From coarse to fine, each level of the pyramid is synthesized as in Efros's algorithm, except that the considered neighborhood uses spatial adjacent sites and sites belonging to the coarser scale. This method, combined with a tree-structured vector quantization to accelerate the search for the nearest neighbor, allows to use smaller neighborhood and thus dramatically increases the speed of the synthesis process. This model is similar to the Scale Causal Multigrid model we have proposed in [7] [8] for the sonar image segmentation issue. Nevertheless, the technique introduced in [7] involved a parametric model in which the conditional PDF was expressed within an energy function which consists of a sum of local interaction potentials involving a few parameters (called clique parameters). Wei's model is non-parametric and offers the opportunity to capture very efficiently the larger scale characteristics of a given texture sample. Example of this synthesis algorithm on a complex texture is shown in Fig. 1. The synthesized texture is very similar to the original and seems to come from the same (underlying) stochastic process as the input textural sample.

FIG. 1 – *Wei et al.'s texture synthesis algorithm. The resulting texture (right) is synthesized at twice the size of the original input sample (left) (see [2] for more examples).*

Let us now find a probabilistic formulation to the Wei's sampling procedure. To this end, let us consider, at full resolution, a couple of random fields (X, Z), with $Z = \{Z_s, s \in S\}$ and $X = \{X_s, s \in S\}$, the set of variables associated to the N pixels of the input and output texture images respectively and located on a lattice S of N sites s. Each of the X_s and Z_s take their value in $\Lambda = \{0, \ldots, 255\}$ (i.e., the set of grey levels). A particular configuration for Z (the synthesized texture) is given as $\{Z_1 = z_1, \ldots, Z_n = z_n\}$ which is abbreviated to $z (= \{z_s\})$ for convenience. All possible output images z are contained within the configuration space $\Omega = \Lambda^N$. Let $P_{Z/X}(z/x)$ the distribution measure on Ω and let us define,

$$P_{Z/X}(z/x) = \frac{1}{Z_g} \exp\Big\{-\sum_{s \in S} \big(\min_{p \in S} D(N(x_p), N(z_s))\big)\Big\}, \quad (1)$$

where $D(.)$ is the L_2 distance and Z_g is a normalizing constant (depending on x only). From distribution (1), we can synthesize a texture or a particular realization of z by a Stochastic Relaxation (SR) procedure [9]. This is done by starting with any image and iteratively updating pixels in the image with respect to the Local Conditional PDF (LCPDF),

$$P_{Z_s/N(Z_s)}(z_s/N(z_s)) = \frac{1}{Z_l} \exp\big\{-\min_{p \in S} D(N(x_p), N(z_s))\big\},$$

where Z_l is the local partition function and dependence on X is omitted for convenience. A well-known SR algorithm is the Gibbs sampler [9] proposed by Geman *et al.* An alternative approach to this stochastic (and thus computationally expensive) procedure is the Iterative Conditional Modes (ICM) introduced by Besag [3]. Using (1), this method consists in finding for each site, and until convergence is achieved, the value z_s that maximizes $P(z_s/N(z_s))$. From an algorithmic point of view, this procedure is similar to the Wei *et al.* procedure, except that the ICM process has to be repeated until convergence is achieved (and generally more than one pass). This shows that Eq. (1) provides an interesting probabilistic formulation of a non parametric prior model. This prior model will be exploited in the following, via a Bayesian inference, in our rendering procedure.

3. BAYESIAN INFERENCE

Let now Z represents the set of hidden variables (the image to be synthesized with rendering effect in our application), and Y, the observation variables (the input image to be re-rendered). The general Bayesian inference defined by Besag [3] consists of the successive stages outlined below ;

- Construct of a prior PDF ; $P_Z(z)$ which represents our initial prior knowledge on the solution.
- Combine the observation variables with the abovementioned prior model through a conditional PDF ; $P_{Y/Z}(y/z)$.
- Compute the posterior density $P_{Z/Y}(z/y)$ from the prior and the conditional probability by Bayes theorem ; $P_{Z/Y}(z/y) \propto P_Z(z)\ P_{Y/Z}(y/z)$.

Finally, choose a criterion based on this posterior probability, in order to define the "best" z given y. In Bayesian analysis, all kinds of inference are made from $P_{Z/Y}(z/y)$. Finding the Maximum A Posteriori (MAP) estimate i.e., the occurrence of z that is the most likely, is one of the most frequently used choice of inference. For $P_{Z/Y}(z/y) \propto \exp(-U(z, y))$ and for some "energy" function U, this MAP estimation problem is equivalent to an energy function minimization problem ; $\hat{z}_{\text{MAP}} = \arg\min_{z \in \Omega}\ U(z, y)$. The energy function $U(z, y)$ involves two components. One arising from $P_Z(z)$, which encodes constraints on the desired solution. It is the *a priori* energy term. In our Bayesian rendering model, we can efficiently exploit the prior model expressed by Eq. (1). The other one, stemming from $P_{Y/Z}(y/z)$, expresses the interaction between the hidden variables and the observed data. In our application, we can define, for this (so-called) likelihood energy term, a likelihood measure of visual difference between the observation and the image z to be synthesized by constraints. More precisely, we use the following difference metric between pixel values of z and y,

$$\begin{aligned} P_{Y/Z}(y/z) &= \frac{1}{Z'} \exp\Big\{ -\varrho \sum_{s \in S} (y_s - z_s)^2 \Big\}, \\ &= \frac{1}{Z'} \exp\Big\{ -\beta \sum_{s \in S} \big(D(N(y_s), N(z_s)\big) \Big\}, \qquad (2) \\ &\approx \frac{1}{Z'} \exp\Big\{ -\beta \sum_{s \in S} \min_{p \in S} \big(D(N(y_s), N(x_p)\big) \Big\}. \end{aligned}$$

with $\beta = \varrho/|N(.)|$, $|N(.)|$ designates the number of neighbors and Z' is a normalizing constant. Using this likelihood and the prior model given by Eq. (1), we can express the most probable MAP configuration $\hat{z}_{\text{MAP}}$ knowing y and x ; $\hat{z}_{\text{MAP}} = \arg\max_{z \in \Omega} P_{Z/Y}(z/y)$. Equivalently, the corresponding posterior energy to be minimized, at full resolution, is,

$$U(x, y, z) = \sum_{s \in S} \min_{p \in S} \big\{ D(N(x_p), N(z_s)) + \beta\, D(N(x_p), N(y_s)) \big\}.$$

4. COARSE-TO-FINE ESTIMATION

Due to the multiscale structure of the model and the scale causal specification of the neighborhood, we have in fact to deal with the following coarse-to-fine recursive energy minimization problem [8],

$$U^l(x, y, z) = \sum_{s \in S^l} \min_{p \in S^l} \big\{ D(N(x_p), N(z_s)) + \beta\, D(N(x_p), N(y_s)) \big\}, \qquad (3)$$

where S^l define the set of pixels at resolution level l. Each of the associated energy can be efficiently solved with an ICM algorithm [3]. Like for other hierarchical approaches, the advantages of this method are twofold. From a modeling point of view, this hierarchical strategy offers an appealing ability to capture *a priori* characteristics of the (underlying generative) stochastic process within a range of different scales. From an algorithmic point of view, contextual information is propagated in a more efficient way [8]. Besides multiscale optimization technique has shown to exhibit fast convergence property and robustness against local *minima* for highly non-linear combinational problems (estimation results are nearly comparable to those obtained by stochastic optimization procedures) [5]. β is the factor that provides a relative weighting between the two energy terms (for $\beta \to \infty$, we find the optimization problem related to the multiscale sampling technique proposed by Wei *et al.* in [2]).

5. EXPERIMENTAL RESULTS

In a preliminary step, we stretch the histogram of the input image y in order to get the same minimal and maximal grey level values of the input texture sample x. Then, the input image and the texture sample are decomposed into multiple resolutions by building two Gaussian pyramids. Each level or image of these pyramids is considered as being toroidal. This assumption allows to handle neighborhoods near the image boundaries. We convert the input image y and the input sample x in grey level values for the Gaussian pyramid constructions. Consequently, only grey level values are used for the searching process. Color can be simply recovered by copying the R, G and B channels at full resolution for each selected pixel by the searching process. We use three levels of pyramid ($L = 3$). For the first coarse-to fine sampling process, we use a 7×7 spatial and causal neighborhood in the fine level with a symmetric 5×5 spatial neighborhood in the coarse level. A symmetric 5×5 neighborhood in the fine and coarse levels are then used for the other passes of the algorithm.

In order to eliminate the difference of contrast between x and y, we use, instead of $D(N(x_p), N(y_s))$ in Eq. 3, the following measure,

$$D\big(\gamma \overline{N}(x_p) + (1 - \gamma) N(x_p), N(y_s)\big),$$

where $\overline{N}(x_s)$ designates the mean of pixel values contained in the neighborhood of x_s and γ is a parameter ($\in [0, 1]$) allowing to normalize the local variance between x and y. Finally, we use an Approximate-Nearest-Neighbor search technique [10] to accelerate the search of the conditional mode of the ICM procedure, i.e., the configuration corresponding to the maximum of the LCPDF. Finally, for each coarse-to-fine pass of the multiscale ICM we rebuild the hierarchy of estimation $\hat{z}^1, \ldots, \hat{z}^L$ from $\hat{z}^0$ with a Gaussian pyramid.

Several examples of Bayesian rendering results are shown in Figs 2 and 3. In these examples, we try to transfer some artistic rendering styles from a drawing image to an input real photograph. In these cases, the multiscale prior model used in our Bayesian procedure allows to efficiently capture the local and global visual characteristics of the line drawing texture, the halftoning texture, the charcoal pen style or the ink painting effect respectively of the input source drawing. We also present the results of our Bayesian Rendering method when the prior model is trained with an input textural sample (cf. Figs 4). Obtained results show that the proposed method allow to constraint an input image towards an image with the same local and global characteristics of the input texture sample. β allows to control the weighting between the intrinsic visual characteristics of the input image and the characteristics of the texture sample. The sampling process takes about 2 hours per image on a 400Mhz PC workstation. In addition to a rendering or a stylized depiction procedure, the proposed technique may eventually be applicable to a much broader of application, such as *textural morphing*, texture mixing algorithm or in order to create a very broad range of visual effects.

6. CONCLUSION

In this paper, we have presented a statistical method for the rendering problem. We have stated this problem in the Bayesian framework and exploited the multiscale non parametric model proposed by Wey *et al.* as prior model of this Bayesian inference. This prior model has shown itself to be well suited to efficiently capture the local and global visual characteristics of a complex input texture or to capture *some* characteristics of a rendering style given by an artistic illustration example. Combined with a likelihood model, the resulting Bayesian rendering procedure allows to constraint a given image with the visual characteristics or the (artistic representation) style of an other image, yielding a very broad range of interesting visual effects.

[1] † cf. Http :www.iro.umontreal.ca/∼mignotte/ICPR02/ for additional examples.

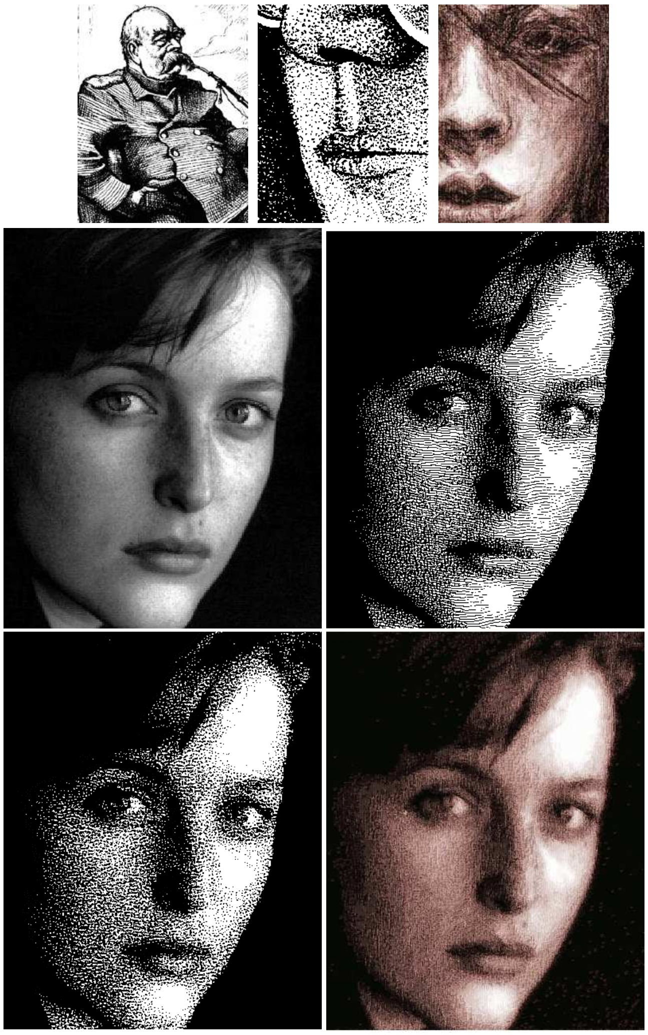

FIG. 2 – *Drawing image with respectively the following rendering styles ; hatching, pointillist, charcoal drawing. Real photograph and Bayesian rendering results ($\beta=1$) based on each drawing example*†.

FIG. 3 – *Drawing image with the* ink painting *style, original image and Bayesian rendering result* †.

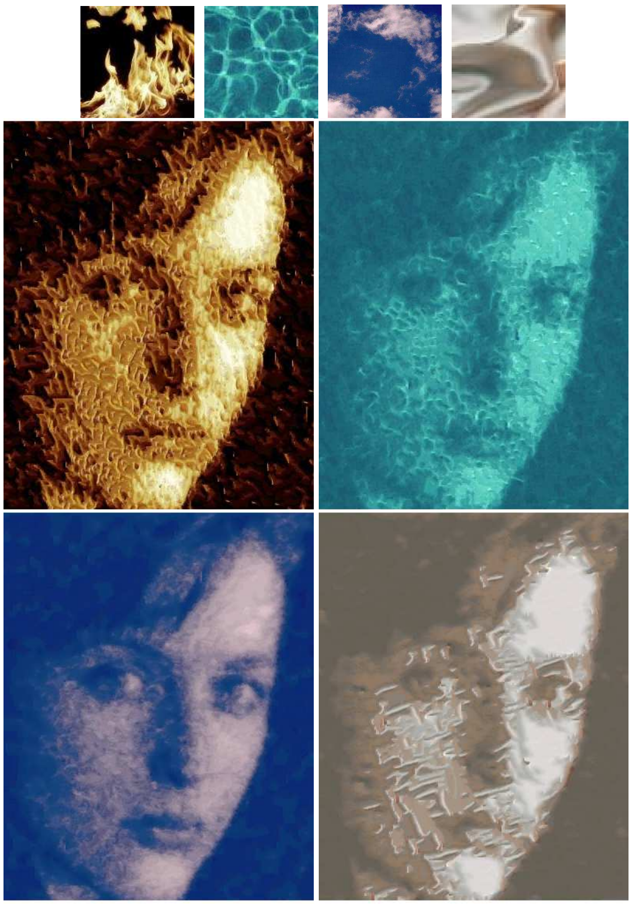

FIG. 4 – *Texture sample (respectively : fire, water, clouds, satin). Real photograph. Bayesian rendering results (β=3) based on each textural sample* †.

7. REFERENCES

[1] C. Reynolds. Stylized depiction in computer graphics, NPR, painterly and toon rendering. http ://www.red3d.com/cwr/npr/.

[2] L.-Y. Wei and M. Levoy. Fast texture synthesis using tree-structured vector quantization. In *Proceedings of SIGGRAPH'2000*, pages 479–488, July 2000.

[3] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, B-48 :259–302, 1986.

[4] P. B. Chou and C. M. Brown. The theory and practice of Bayesian image labeling. *International Journal of Computer Vision*, 4 :185–210, 1990.

[5] F. Heitz, P. Pérez, and P. Bouthemy. Multiscale minimisation of global energy functions in some visual recovery problems. *CVGIP : Image Understanding*, 59(1) :125–134, 1994.

[6] A. A. Efros and T. K. Leung. Texture synthesis by non-parametric sampling. In *7th International Conference on Computer Vision, (ICCV'99)*, pages 1033–1038, 1999.

[7] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy. Unsupervised Hierarchical Markovian segmentation of sonar images. In *4th IEEE International Conference on Image Processing, (ICIP'97)*, page on CD ROM, October 1997.

[8] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy. Sonar image segmentation using a hierarchical MRF model. *IEEE Trans. on Image Processing*, 9(7) :1216–1231, 2000.

[9] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6(6) :721–741, 1984.

[10] S. Arya, D. M. Mount, N.S. Netanyahu, R. Silverman, and A.Y. Wu. An optimal algorithm for aproximate nearest neighbor searching in fixed dimensions. *Journal of the ACM*, 45(6) :891–923, 1998.

---

[1] † cf. Http : www.iro.umontreal.ca/∼mignotte/ICPR02/ for colored images.