

Nonparametric Multiscale Energy-Based Model and Its Application in Some Imagery Problems

Max Mignotte

Abstract—This paper investigates the use of a nonparametric regularization energy term for devising an example-based rendering and segmentation technique. We have stated this problem in the multiresolution energy minimization framework and exploited the multiscale structure proposed by Wei and Levoy for the texture synthesis problem. In this nonparametric energy minimization framework, we also propose a computationally efficient coarse-to-fine recursive optimization method to minimize the cost function related to this hierarchical model. In this context, the formulation of our example-based regularization term also allows to directly infer an intuitive dissimilarity measure between two contour shapes. This measure is herein exploited to define an efficient shape descriptor for the contour-based shape recognition and indexing problem.

Index Terms—Nonparametric multiscale energy-based (or multiresolution example-based) model, inpainting, Non-Photorealistic Rendering (NPR), segmentation, contour-based shape recognition, shape indexing.

1 INTRODUCTION

MANY tasks in computer vision, graphics, and image processing can be expressed as global optimization problems. The general issue is to find the global minimum of a cost function (also called a *energy* function) involving the observed data and the hidden variables, which are the representations to be extracted from the original images. Generally, these energy functions (either heuristically designed or guided by a statistical modeling of the interaction between the variables) involve two components. The first one expresses the interaction between the unknown variables and the data, while the second one encodes constraints on the desired solution [1], [2], [3], [4], [5], [6], [7]. This latter component is often a regularization term whose essential role is to rightly constrain the ill-posed nature of the considered inverse problem (it guarantees, to some extent, the existence and the uniqueness of a consistent solution which continuously depends on the data).

Within this energy minimization framework (herein heuristically defined), we first investigate in this paper the use of a nonparametric regularization term, deduced from a recent example-based technique of texture synthesis proposed by Efros and Leung [8], to constrain some inverse (highly nonlinear) problems in computer graphics and image analysis. Combined with an appropriate data-driven term, we show the application of the resulting nonparametric energy-based model to solve some Non-Photorealistic Rendering (NPR) and segmentation problems.

Unfortunately, the introduction of such a nonparametric regularization term leads to an intricate minimization

problem. The specified energy function depends on a very large number of variables and exhibits many local minima. To circumvent this difficulty, hierarchical energy-based models have been proposed in the literature. A classical scheme called the multiresolution approach consists of building a Gaussian pyramid from the full resolution image and considering a set of “similar” energy-based models on the different resolution levels of the pyramid [9], [10], [11], [12]. Instead of considering the minimization problem on the full and original configuration space, the original inverse problem to be solved is decomposed in a sequence of optimization problems of reduced complexity (each one commonly thought of as a “smooth” version of the original problem to be solved). From an algorithmic point of view, the optimization problem can be solved in a coarse-to-fine manner using, as initialization of each optimization problem (at each resolution level), the solution obtained at previous level. The major advantage is to reduce the number of variables. Consequently, it allows us to drastically save computational effort and/or to provide an accelerated convergence toward improved estimate. In addition, hierarchical models are interesting from a modeling point of view. They offer the appealing ability to capture a priori characteristics of the solution image to be estimated within a range of different scales and/or to introduce longer-range dependencies in the neighborhood (on which the model energy is often decomposed as a sum of local interaction functions). They also allow us to integrate the information in a progressive and efficient way.

In the same spirit, we also propose a computationally efficient coarse-to-fine recursive optimization method to minimize the energy function related to our nonparametric hierarchical model. Another contribution of this paper is to directly exploit the example-based regularization term, extracted from Efros and Leung’s model, to infer an intuitive dissimilarity measure between two contour shapes. This measure is herein exploited to define an efficient shape descriptor for the rapid shape retrieval and indexing problem.

• The author is with DIRO, Département d’Informatique et de Recherche Opérationnelle, C.P. 6128, Succ. Centre-ville, Montréal, Canada (Québec), H3C 3J7. E-mail: mignotte@iro.umontreal.ca.

Manuscript received 18 Oct. 2002; revised 19 May 2003; accepted 14 Aug. 2003.

Recommended for acceptance by Y. Amit.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 117624.

The remaining parts of this paper are organized as follows: In Sections 2 and 3, we provide a brief overview of the texture synthesis model introduced by Efros and Leung and its multiscale extension, presented by Wei and Levoy in [13]. We also give some examples and direct applications of this technique for solving the image interpolation and replacement problems. In Section 4, we present the example-based regularization term that will be exploited in our nonparametric energy-based multiresolution model, respectively, in NPR (Section 5) and image segmentation (Section 6). We present the coarse-to-fine recursive energy minimization problem related to this hierarchical model in Section 7. The shape descriptor derived from the example-based regularization term is presented in Section 8. Finally, experimental results and conclusions are presented in Sections 9 and 10.

2 PREVIOUS WORK

Recently, Efros and Leung [8] have proposed a nonparametric 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 semistochastically and then using these first assignments as a seed). In this method, the input texture is modeled as a Markov Random Field (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 of 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 sampling from it with a Gibbs or Metropolis sampler [5]. Due to the exponential growth of the 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 and Leung's algorithm is completely deterministic and avoids explicit conditional

probability construction (like in [14]) and stochastic sampling.¹ By using adequately sized neighborhood, the algorithm produces surprisingly good results in texture synthesis, 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 in that the initial texture contains large scale structure and we have to use large neighborhoods.

To cope with these problems, Wei and Levoy [13] 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 us to use smaller neighborhood and, thus, dramatically increases the speed of the synthesis process. This model is similar to the Scale Causal Multigrid (SCM) model we have proposed in [12], [18] for the sonar image segmentation issue. Nevertheless, the technique introduced in [18] involves a parametric model in which the prior PDF is expressed by a classical Gibbs distribution whose the energy function (i.e., the regularization energy term) consists of a sum of local interaction potentials involving a few parameters (called the clique parameters). Wei and Levoy's method is nonparametric (i.e., example-based) and the synthesized textures can be considerably more complex compared to the ones created with the parametric version of this model [12]. Examples of this synthesis algorithm on complex textures are shown in Fig. 1. The synthesized textures are very similar to the original and seem to come from the same (underlying) stochastic process as the corresponding input textural sample. Results are good both for stochastic textures, i.e., textures without explicit texture elements (texels) or for regular textures (consisting of repeated texels) and do not exhibit "popout" artifacts (i.e., simple repetitions of the original input sample) and discontinuity artifacts.

3 APPLICATIONS OF WEI'S ALGORITHM

This multiscale texture synthesis method can also be used directly as a tool for solving several practical problems in graphics and image processing. In particular, we can cite the textured image interpolation and the image replacement problems.

Image interpolation through texture synthesis can be used for inpainting, retouching, or restoration applications. In image processing, the goal of the inpainting or retouching is to reconstitute the missing or damaged portions of an image in order to revert deteriorations (e.g., cracks, scratches, or scrambled regions on a scanned photograph,

1. This example-based sampling algorithm has inspired some recent research works in computer graphics. We can cite the work of Chen et al. on facial caricature synthesis [15], Hertzmann et al. [16], or the work of Jodoin et al. [17] that efficiently exploits this deterministic sampling technique to generate, after learning from examples, hand-drawn styles of curves for nonphotorealistic rendering and automatic synthesis of freehand sketch drawings with a particular style.

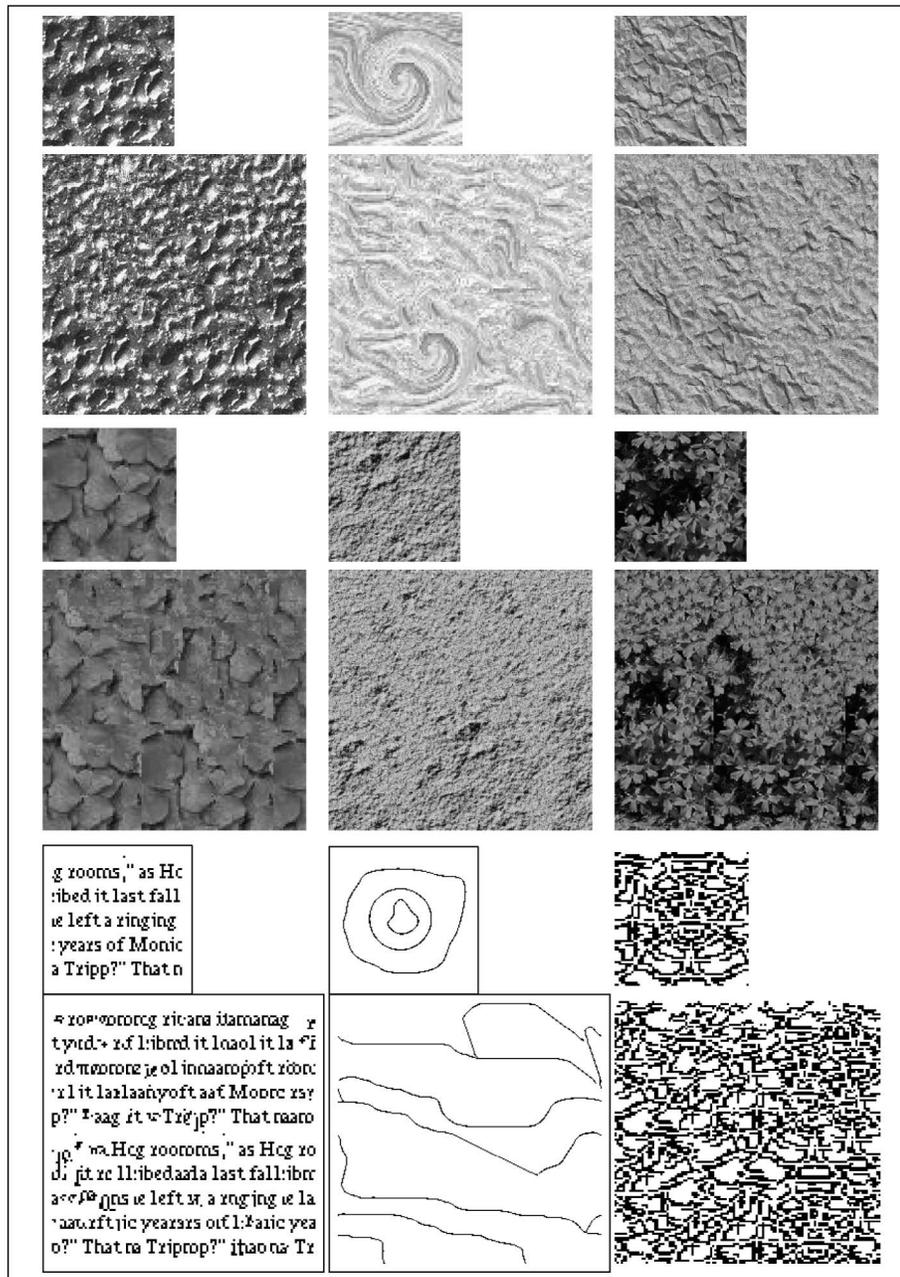


Fig. 1. Wei and Levoy's texture synthesis algorithm. The resulting texture is synthesized at twice the size of the original input sample (see [13] for additional and color examples).

dust spots in film, blocks of missing data in wireless image transmission, etc.) or to add or remove elements (e.g., removal of stamped date, superimposed text, subtitles, or publicity on the image, or entire undesirable objects such as microphones, wires in special effects, etc.) (see [19] for an excellent review of existing image inpainting techniques). Contrary to the classical denoising algorithm, the information to be recovered is not hidden by the (assumed) additive noise but missing. The exploitable and available information is mainly contained in the regions surrounding the area to be inpainted and given by the different textures present in the nondegraded portions of the image.

In order to solve this problem, a simplistic approach would consist of filling in the regions to be inpainted with a patch

extracted somewhere else in the image and exhibiting a similar texture. This technique would produce visible discontinuities at the boundaries and visible seams or repetitive features. In order to overcome these problems, a texture synthesis-based inpainting technique can be directly exploited. To this end, a preprocessing step would consist of preliminary filling of the areas to be inpainted with the average color of the pixels surrounding it. To do that, we can first decompose the image into multiple resolutions by building a color Gaussian pyramid (missing pixels corresponding to the region to be recovered are not considered as being null but are not taken into account in the averaging process). Due to this multiresolution process, the missing regions disappear in the higher levels of the pyramid. Second,

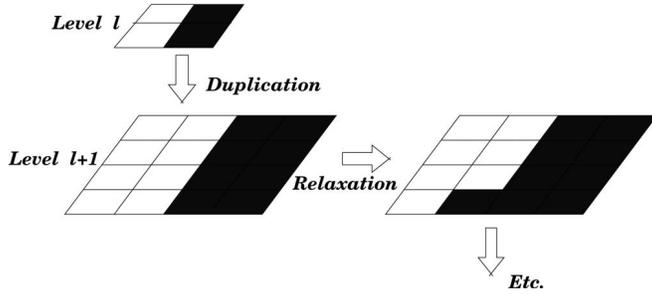


Fig. 2. Duplication and “coarse-to-fine” minimization strategy.

a coarse-to-fine duplication (cf. Fig. 2) of the pixel values is realized for the missing pixels at each resolution level. We then use the multiscale texture synthesis process in the regions to be inpainted.

Figs. 3b and 3c show, respectively, a real image with a synthetic deterioration and, finally, the restored image. In this application, for the first coarse-to-fine sampling, we use a raster scan synthesis ordering, with a causal and spatial neighborhood (using the RGB channels) for the fine level with a symmetric spatial neighborhood in the coarse level. A random order with a symmetric neighborhood in the fine and coarse level is then used for the other iterations of the sampling algorithm (three or four in our application). For each coarse-to-fine pass of the sampling, we rebuild the Gaussian pyramid from the estimated image at full resolution. This strategy allows us to efficiently remove discontinuity artifacts created at the boundaries of the areas to be recovered. The same procedure can be used for image replacement (see Figs. 3d and 3e). More precisely, one portion of the textured image can be substituted with another texture. In this case, the preprocessing step consists of choosing the areas to be replaced and filling in these areas with the average color of the desired texture.

These examples show that this nonparametric sampling technique offers the appealing ability to model and synthesize complex textures. We propose in the following sections to deduce, from this model, a nonparametric (regularization) term that will then be exploited to constrain a rendering and segmentation procedure.

4 EXAMPLE-BASED REGULARIZATION TERM

Let us express Efros and Leung’s sampling technique as a closely related global cost function minimization problem. To this end, let us now consider, at full resolution, a couple of variables (x, z) , with $z = \{z_s, s \in S\}$ and $x = \{x_s, s \in S_x\}$, the set of variables associated to the N and N_x pixels of the output and input texture images, respectively, and located on a lattice S and S_x of sites s . Each of the x_s and z_s take their value in $\Lambda = \{0, \dots, 255\}$ (i.e., the set of gray levels). All possible output images z (i.e., the synthesized textures) are contained within the configuration space $\Omega = \Lambda^N$. Let $U_r(x, z)$ the following energy function,

$$U_r(x, z) = \sum_{s \in S} \min_{p \in S_x} D(N(x_p), N(z_s)), \quad (1)$$

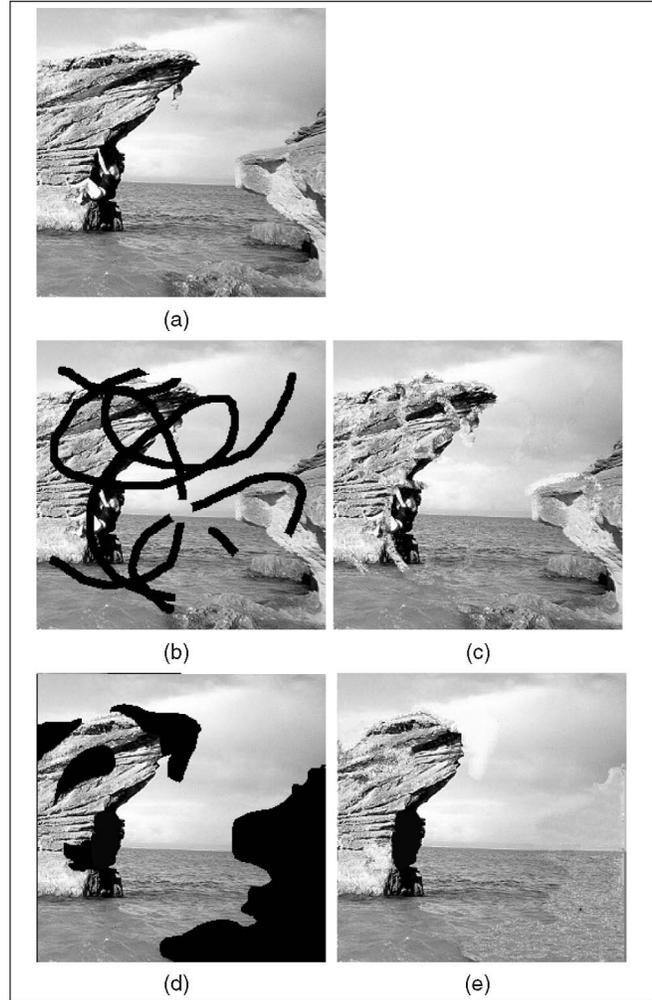


Fig. 3. Texture synthesis-based inpainting restoration example: (a) real image with a synthetic deterioration, (b) real image with a synthetic deterioration, and (c) restoration result. Texture replacement example: (d) real image with areas to be replaced and (e) replacement result.

where $D(\cdot)$ is the L_2 distance and $N(x_p)$ designates the neighborhood of pixel of the output image x at site p (i.e., a spatial neighborhood in the case of Efros and Leung’s sampling procedure and a spatial and scale causal neighborhood for the Wei and Levoy’s algorithm). From the energy function defined by (1), we can find a particular configuration of z that minimizes $U_r(x, z)$ by a global optimization procedure such as a simulated annealing algorithm [1].

An alternative approach to this stochastic (and, thus, computationally expensive) procedure is the Iterative Conditional Modes (ICM) introduced by Besag [5]. Using (1), this method consists of finding for each site s and, until convergence is achieved, the value z_s that minimizes $\min_{p \in S_x} D(N(x_p), N(z_s))$. From an algorithmic point of view, this procedure is similar to Efros and Leung’s procedure, except that the ICM process has to be repeated until convergence is achieved (and, generally, more than one pass). This shows that (1) provides an interesting nonparametric regularization energy term that will be exploited in the following in our rendering and segmentation energy-based procedure.

5 RENDERING MODEL

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 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. Most 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 [20] for an excellent review of existing NPR techniques). Let us mention, nevertheless, the framework for processing images by example, called image analogies, recently proposed by Hertzmann et al. and involving two phases: a design phase in which a pair of images, with one image purported to be a filtered version of the other, is presented as training data; and an application phase in which the learned filter is applied to some new target in order to create an analogous filtered result [21].

In our energy minimization framework, the example-based regularization energy term given by (1) allows us to efficiently capture the local and global visual characteristics of a complex input texture (see Fig. 1) and also seems well-suited to capture some characteristics of a particular rendering style from an artistic illustration example. Combined with the appropriate data-driven term, this cost function optimization strategy, with two energy terms, appears thus to be well-suited to rerender an input image in the style of an other image (for example, an artistic illustration sample). It is the basic idea of the rendering method we propose.

In this context, let z (the set of hidden variables) now represents the image to be synthesized with rendering effect and y (the observation variables) represent the input image to be rerendered. In our energy-based rendering model, we define the following global energy function to be minimized, at full resolution,

$$U(x, y, z) = \sum_{s \in S} \min_{p \in S_x} \{D(N(x_p), N(y_s)) + \beta D(N(x_p), N(z_s))\}. \quad (2)$$

The first energy term is the data-driven or likelihood energy term, which tends to give a solution z visually and locally similar to the data y . The second term (regularization term) is a contextual example-based energy term (given by (1)) which tends to build z with neighborhood configurations similar to the ones existing in the texture sample x . β is the factor that provides a relative weighting between the two energy terms. For $\beta \rightarrow \infty$, we find an energy function whose the optimization leads to a procedure closely related to the multiscale texture sampling technique proposed by Wei and Levoy. For $\beta = 0$, the solution image z becomes the real input image whose the histogram has been stretched according to the one of x . The resulting global energy function $U(x, y, z)$ is also in a suitable form to exploit an efficient search technique allowing to considerably decrease the computational requirement of

the deterministic optimization algorithm used in our application (see Section 7). We will see in the following section that this global energy form is also well-suited to find a segmentation map of a noisy image corrupted by a standard Gaussian degradation model.

6 SEGMENTATION MODEL

The goal of the segmentation is to assign, to each pixel of a noisy image y , a label indicating to which class the pixel belongs. The set of labels z , i.e., the segmentation to be estimated, is called the label field and now takes its values in $\Lambda' = \{e_0, \dots, e_n\}$, representing the set of n possible classes.

By nature, an image segmentation is an ill-posed problem, i.e., the data in themselves are insufficient to unambiguously define the segmentation. In order to rightly constrain the nature of this problem, some particular knowledge about the scene is necessary. A simple a priori knowledge may express the fact that nearby pixels are fairly likely to belong to the same class. In a probabilistic framework, such regularities are well-captured by MRF [5]. Thanks to the Hammersley-Clifford theorem, this prior model can be parametrically described by a Gibbs distribution based on spatially local interactions, called the clique potential β_{st} . For example, the standard Ising-type prior, expressing the probability of the segmentation, for a configuration $z \in \Omega' (= \Lambda'^N)$, is given by,

$$P(z/\beta_{st}) = \frac{1}{Z(\beta_{st})} \exp\left(-\sum_{\langle s,t \rangle} \beta_{st} [1 - \delta(z_s, z_t)]\right), \quad (3)$$

where $\langle s, t \rangle$ indicates that the summation is over neighboring pairs (traditionally, the four or eight nearest neighbors are used) and δ is the delta Kronecker function. Such (parametric) priors are typically very simple and lead, in a Bayesian segmentation, to simply imposing homogeneous regions in the solution. Unfortunately, this commonly used approach does not allow preserving lines and fine or complex structures which sometimes need to be recovered in the segmentation map. To improve segmentation results, some more complex MRF models, involving larger neighborhood structure, have been proposed. Nevertheless, these methods remain limited and also require complex parameter estimation procedures, often based on approximated estimators. An alternative approach we propose is to once again exploit the example-based regularization energy term defined in (1) and trained from a sample having the same spatial characteristics of the solution to be estimated. By this way, it allows to efficiently restrict the segmented image to some subspace defined by the set of all possible neighborhoods taken on the training sample. In this nonparametric modeling framework, x represents now the training label field taking its values in Λ' . More precisely, let us consider that x and z are now two images quantized into n gray levels representing the gray-level mean of the n possible regions or classes existing in y .

The second component of a Bayesian segmentation model is the likelihood distribution, or the class conditional probabilities $P(y/z)$, expressing the probability of the set of pixel taking each of the observable values, given its class

membership field. A commonly used assumption is to consider that the data y are independent conditional on the labeling process z . If it is assumed, one gets $P(y/z) = \prod_{s \in S} P(y_s/z_s)$ (i.e., the noise is assumed to not be correlated with the data). We shall stick to this assumption throughout. To describe the luminance y within each region, we then consider the standard degradation model expressed by the Gaussian distribution $\mathcal{N}(\mu_{z_s}, \sigma_{z_s})$, where μ_{z_s} and σ_{z_s} are the mean and variance associated with the class z at pixel s . We thus let

$$P(y_s/z_s) = \frac{1}{\sqrt{2\pi\sigma_{z_s}^2}} \exp\left(-\frac{(y_s - \mu_{z_s})^2}{2\sigma_{z_s}^2}\right). \quad (4)$$

Using this likelihood and, for a classical Markovian segmentation, the Ising-type prior expressed by (3) (commonly used in a Markovian segmentation), we can express the most probable MAP segmentation \hat{z}_{MAP} knowing y (and x): $\hat{z}_{\text{MAP}} = \arg \max_{z \in \Omega} P(z/y)$ or, equivalently, the corresponding posterior energy to be minimized,

$$U(x, y, z) = \underbrace{\sum_{s \in S} \Psi_s(y_s, z_s)}_{U_d(y, z)} + \underbrace{\beta \sum_{st} \{1 - \delta(z_s, z_t)\}}_{U_r(x, z)}, \quad (5)$$

where $\Psi_s(y_s, z_s) = -\ln P(y_s/z_s)$. The first term ($U_d(y, z)$) is referred to as the data driven term and tends to give a solution similar to the data. The second one ($U_r(x, z)$) is referred to as the regularization term and encourages simply homogeneous regions in the solution.

In order to recover complex configurations of the labeling process that cannot be expressed by the simple Ising-type energy term, we now propose the following heuristic nonparametric global energy to be minimized, similar to the one used (at full resolution) in our rendering procedure,

$$U(x, y, z) = \sum_{s \in S} \min_{p \in S_x} \{D(N(x_p), N(y_s)) + \beta D(N(x_p), N(z_s))\}. \quad (6)$$

In this global energy function, the second term is the example-based regularization energy term given by (1) whose role, in this segmentation procedure, is to constrain the segmentation problem by recovering from x , complex neighborhood configurations that cannot be expressed mathematically. The first term behaves as a data-driven term which can be easily compared to the one expressed by a classical Markovian segmentation. β is the factor that provides a relative weighting between the two energy terms.

7 COARSE-TO-FINE OPTIMIZATION

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 [12],

$$U^l(x, y, z) = \sum_{s \in S^l} \min_{p \in S_x^l} \{D(N(x_p), N(y_s)) + \beta D(N(x_p), N(z_s))\}, \quad (7)$$

where S^l and S_x^l define the set of pixels, at resolution level l , of the output and input images, respectively.

In this multiresolution optimization framework, the image z to be estimated, at each resolution level, leads to a global energy function minimization problem requiring a set of estimations, given a query vector in the (d -)dimensional space (and under the l_2 norm), of the nearest neighbor. In this context, the main drawback of this set of exact estimations of the nearest neighbor remains its high-computational requirement (for reasonably large dimensions, brute-force search, which compares the query point to each data point, often remains the most efficient in practice). In order to overcome this problem and to reduce the computational load, we propose computing an approximated estimation of each nearest neighbor. Due to the presence of $N(x_p)$ in the data-driven and regularization energy term, this energy form is suitable for such a technique. In this context, an interesting practical data structure (based on KD-tree) and efficient search method, called Approximate Nearest Neighbor (ANN) have been recently proposed [22] and can be used, at each resolution level l , in conjunction with a deterministic optimization algorithm such as an ICM algorithm [5] to minimize U^l . More precisely, the ANN search technique can be efficiently used to accelerate the search of the conditional mode of the ICM procedure (i.e., the configuration corresponding to the maximum of $\min_{p \in S_x^l} \{ \cdot \}$). The deterministic iterative minimization at level $(l-1)$ is then initialized by the interpolation of the result previously obtained at coarser level l [12] (see Fig. 2).

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 solution image z to be estimated within a range of different scales. Consequently, the solution is better constrained. From an algorithmic point of view, contextual information is propagated in a more efficient way [12]. Besides, multiscale optimization technique has been shown to exhibit fast convergence property and robustness against local *minima* for highly nonlinear combinatorial problems (estimation results are nearly comparable to those obtained by stochastic optimization procedures) [2], [12]. In (7), β is the factor that provides a relative weighting between the two energy terms (for $\beta \rightarrow \infty$, we find the optimization problem that is closely related to the multiscale sampling technique proposed by Wei and Levoy in [13]).

8 CONTOUR-BASED SHAPE RECOGNITION

8.1 Introduction

Along with color and texture, widely used in the content-based image retrieval community, shape information remains an important, although relatively underexploited, cue for pattern recognition and retrieval purposes in large image database (see [23] for an extensive survey of shape matching in computer vision). Nevertheless, even when color and texture are absent, as in line drawing, the human visual

system can recognize objects quite well and can also make efficiently coarse distinctions very quickly. This demonstrates the highly discriminative and descriptive power that is only conveyed by the information modeled by a single shape contour. One shape-based approach, commonly used in an automatic image database retrieval system, is the deformable template-based framework [23], [24], [25]. This approach of shape matching through global deformations is now well known and gives quite good results. A drawback of this technique is its computational cost. Such an approach could not be viable on a database of several thousand objects. While a final stage of deformable matching may be necessary to make very fine distinctions, it is interesting to do fast early pruning based on much cruder measures in order to produce an accurate short list of candidate matches suitable for a more careful (and also more time consuming) (possibly shape-based) matching engine.

In that prospect, the example-based regularization term, defined in (1) can once again be exploited. To this end, let us notice that, using this constraint term and given the initial textural sample x , an output synthesized image z is a configuration for which the distance $\text{Dist}(x||z)$, at full resolution, is low with,

$$\text{Dist}(x||z) = \sum_{s \in S} \min_{p \in S_x} D(N(x_p), N(z_s)), \quad (8)$$

This multiscale measure involving a spatial and scale causal neighborhood offers the opportunity to capture very efficiently the larger scale characteristics of a given binary textural image (such as line drawings of a contour shape; see, for example, the bottom right synthesized texture of Fig. 1). Of course, a shape contour is fundamentally different from a textural sample. Shapes generally contain detailed and nonrepeating variations. Nevertheless, the distance $\text{Dist}(x||z)$ also seems well-suited to capturing some local geometric characteristics of a shape contour and could be used to infer and define an efficient shape descriptor [26].

8.2 Shape Similarity Distance

The query shape describes only one of the possible instances of the considered shape class. In order to take into account the variability of each possible shape related to the object class to be detected in the database, we do not directly exploit the discrete set of points, sampled from the internal or external contours on the shape. Instead, we exploit an edge potential field or a distance map. This field is determined by the positions of the edge points in the contour images. For a pixel (i, j) in the contour image I , we define its edge potential by,

$$\Phi_I(i, j) = \exp\left\{-\frac{1}{\rho} (\delta_i^2 + \delta_j^2)^{\frac{1}{2}}\right\}, \quad (9)$$

where (δ_i, δ_j) is the displacement to the nearest edge point in the image and ρ is a smoothing factor which controls the degree of smoothness of the potential field and, consequently,

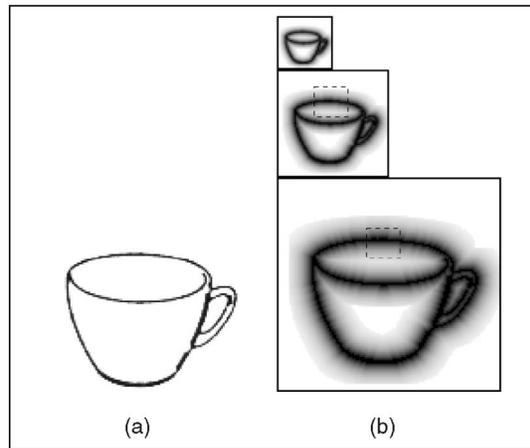


Fig. 4. (a) Original line drawing stemming from the Snodgrass and Vanderwart database. (b) Multiscale edge potential map (with two levels of pyramid and $\rho = 4$) and example of multiscale piece of contour extracted from this edge potential map.

the intrinsic variability of the object class to be detected.² We compute this potential field both for the (internal and external) contours of the query shape and for the contours of the candidate shape. The query and candidate images are then decomposed into multiple resolutions by building two Gaussian pyramids in order to get the multiscale structure that will be used in our matching process. An example of this multiscale potential edge field is shown in Fig. 4. This multiscale 2D distance map can also be viewed as a hierarchical Chamfer map [27], currently exploited in the shape-based (binary) object detection techniques using distance transforms [28].

Finally, in order to get a more robust shape dissimilarity measure, the distance $\text{Dist}(q, c)$ between a query (q) and a candidate (c) shape is defined as,

$$\text{Dist}(q, c) = \max\{\text{Dist}(q||c), \text{Dist}(c||q)\}, \quad (10)$$

$$\text{with } \text{Dist}(q||c) = \sum_{s \in C_c} \min_{p \in C_q} D(N(\Phi_{q_p}), N(\Phi_{c_s})), \quad (11)$$

where the first summation is over the set of contour points (C_c) of the candidate shape c . The search of the min is over which are not previously selected. $N(\Phi_{q_p})$ and $N(\Phi_{c_s})$ designates, respectively, the (multiscale) set of edge potential values contained in the (symmetric and causal) neighborhood of contour point p and s of the query and candidate image. The shape similarity distance $\text{Dist}(q, c)$ can be also interpreted as a reconstruction error measure of the contour of the query shape using multiscale contour elements of the candidate shape (like a reconstruction given by a puzzle made up of multiscale pieces) (see Fig. 5). In this way, we expect to obtain low distance values for similar shapes and high distance values for distinct ones.

2. We can easily complete this edge potential field $\Phi_I(i, j)$ by adding to it a directional component in order to obtain a directional edge potential field, $\varphi_I(i, j) = \Phi_I(i, j) |\cos \beta(i, j)|$, where $\beta(i, j)$ is the angle between the tangent of the nearest edge and the tangent direction of the contour at (i, j) . This potential is similar to the one proposed in [24] for the globally deformable template-based matching approach.

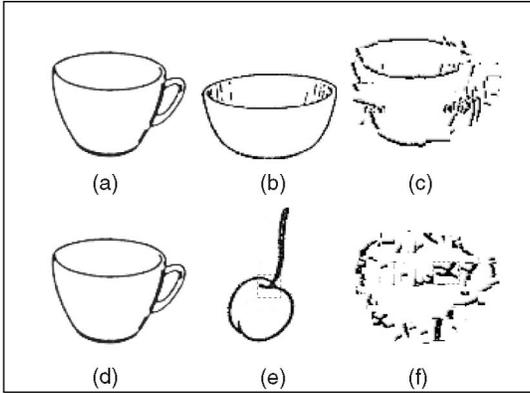


Fig. 5. Examples of reconstructions between a given query shape and two candidate contour shapes ($\rho = 1$ in this example). (a) and (d) Query shape q . (b) and (e) Candidate shapes c . (c) Good reconstruction result and, consequently, low distance value ($\text{Dist}(c\|q) = 6.8 \times 10^5$) and good similarity between the two shapes (a) and (b). (f) Less good reconstruction result and, consequently, higher distance value ($\text{Dist}(c\|q) = 9.65 \times 10^5$) and low similarity between the two shapes (d) and (e). The images in (e) and (f) show an example of a multiscale piece of contour taken from the candidate shape to reconstruct a part of the query shape.

9 EXPERIMENTAL RESULTS

We present, in this section, experimental results of the rendering, segmentation, and contour-based shape recognition procedure presented in Sections 5, 6, and 8.

9.1 Rendering

In a preliminary step, we stretch the histogram of the input image y in order to get the same minimal and maximal gray 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 us to handle neighborhoods near the image boundaries. We convert the input image y and the input sample x in gray-level values for the Gaussian pyramid constructions. Consequently, only gray-level values are used for the searching process. Color is 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 for the constraint term and a 5×5 symmetric spatial neighborhood for the data-driven term in the fine level. A 5×5 symmetric spatial neighborhood is used in the coarse level for the constraint and data-driven term. For the other passes of the algorithm, we use a 5×5 symmetric spatial neighborhood in the fine and coarse levels for the constraint and data-driven term (cf. Fig. 6 and Table 1).

In order to eliminate the difference of contrast between x and y , we use, in (7), instead of $D(N(x_p), N(y_s))$, the following measure,

$$D(\gamma \bar{N}(x_p) + (1 - \gamma)N(x_p), N(y_s)), \quad (12)$$

where $\bar{N}(x_s)$ designates the mean of pixel values contained in the neighborhood of x_s and γ is a parameter ($\in [0, 1]$) allowing us to normalize the local variance between x and y . We recall that we use an ANN search technique [22] to accelerate the search for the conditional mode of the ICM

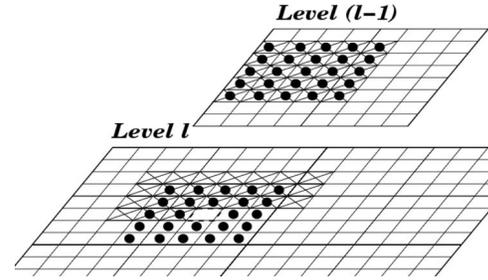


Fig. 6. Two consecutive levels of the pyramidal data structure and neighborhood used for the regularization term, respectively, for the first pass of the ICM (black crosses) and for the other iterations (black circles).

procedure. Finally, for each coarse-to-fine pass of the multiscale ICM, we rebuild the hierarchy of estimation $\hat{z}^1, \dots, \hat{z}^L$ from \hat{z}^0 with a Gaussian pyramid.

Several examples of rendering results are shown in Figs. 7 and 8. 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 example-based regularization energy term allows to efficiently capture the local and global visual characteristics of the line drawing texture, the pointillist texture, the charcoal pen style, or the ink painting or painting effect of the input source drawing (see also [29] and <http://www.iro.umontreal.ca/~mignotte/ICPR02/> for additional and color rendering results). We also present the results of our Rendering method when the regularization term is trained with an input textural sample (cf. Fig. 9). Obtained results show that the proposed method allows us to constrain an input image toward an image with the same local and global characteristics of the input texture sample. β allows us to control the weighting between the intrinsic visual characteristics of the

TABLE 1
Multiscale Local Rendering Procedure

Multiscale Rendering Procedure	
y	Input image (to be rendered)
z	Image to be synthesized (with rendering effect)
x	Input sample
$\hat{z}^0, \dots, \hat{z}^L$	A pyramid of synthesized images ($\hat{z}^0 = \hat{z}$)
β	Rendering weighting factor
Initialization:	
• Stretch the histogram of y according to the one of x	
• Initialize $\hat{z}^0, \dots, \hat{z}^L$ with a random noise (at each level)	
• $G_y \leftarrow$ Build Gaussian Pyramid from y	
• $G_x \leftarrow$ Build Gaussian Pyramid from x	
• Stretch the histogram of \hat{z}^i according to the one of $x^i \forall i \in [0, \dots, L]$	
– iterations $\leftarrow 0$	
while (iterations < ITERMAX) do	
foreach level l , from lower to higher resolution of G_z do	
• ICM optimization [5] with ANN search [22] on $U^l(x, y, z)$ = $\sum_{s \in S^l} \min_{p \in S_x^l} \{D(N(x_p), N(y_s)) + \beta D(N(x_p), N(z_s))\}$ (Eq. (7))	
• Duplication (cf. Fig. 2)	
$G_z \leftarrow$ Build Gaussian Pyramid from z^0	
iterations \leftarrow iterations + 1	



Fig. 7. Drawing image with, respectively, the following rendering styles: (a) hatching, (b) pointillist, (c) charcoal, (d) mixture of pointillist and cross-hatching, and (e) etching (e). Real photograph and rendering results ($\beta = 1$) based on each drawing example. (see also <http://www.iro.umontreal.ca/~mignotte/ICPR02/> for additional and color rendering results).

input image and the characteristics of the texture sample. The sampling process takes about 20 minutes per image (for a nonoptimized code) on a 1.2 Ghz PC workstation under Linux (for a 512×512 size image). In addition to a rendering or a stylized depiction procedure, the proposed technique may eventually be applicable to a much broader application, such as textural morphing (i.e., creation of image sequence transforming a real image progressively into a texture or vice versa), texture mixing algorithm, or in order to create a very broad range of visual effects.

9.2 Segmentation

We use one level of the pyramid ($L = 1$) and a 5×5 symmetric spatial neighborhood is used in the fine and coarse levels and for the data-driven and regularization energy term. For the first iteration of the ICM algorithm, we set $\beta = 0$, thus considering a segmentation result without example-based regularization term.

In order to get an unsupervised segmentation model, parameters of the noise model can be estimated either by conducting, alternatively, the segmentation and estimation (by using Maximum Likelihood (ML) estimator of the

complete data on the segmented image). In this strategy, initial parameters can be estimated by using ML estimator on a rough segmentation obtained by a K-mean clustering procedure on the set of gray value levels of the noisy image. An alternative approach consists of having a two-step process [12]. First, a parameter estimation step involving, for example, the Expectation Maximization [30] method or the Stochastic Expectation Maximization (SEM) [31] algorithm. Then, a second step devoted to the segmentation itself based on the values of estimated parameters.

Figs. 11 and 12 show two images (respectively, (a) shows a noisy image of block letters and (b) shows an image made up with line segments) corrupted by an additive synthetic white Gaussian noise (with parameters $\mu_{e_0} = 55$, $\sigma_{e_0}^2 = 500$, $\mu_{e_1} = 100$, $\sigma_{e_1}^2 = 500$) and their corresponding ground truth segmentations. Figs. 11 and 12 show (c) the resulting segmentations obtained with a Pott-type prior-based Markovian model and (d) our example-based method trained with the sample shown in Fig 10 with $\beta = 1$ (the segmentation process takes 5 minutes on a 1.2 Ghz PC workstation). In order to be able to only compare the segmentation model, the noise parameters of the likelihood

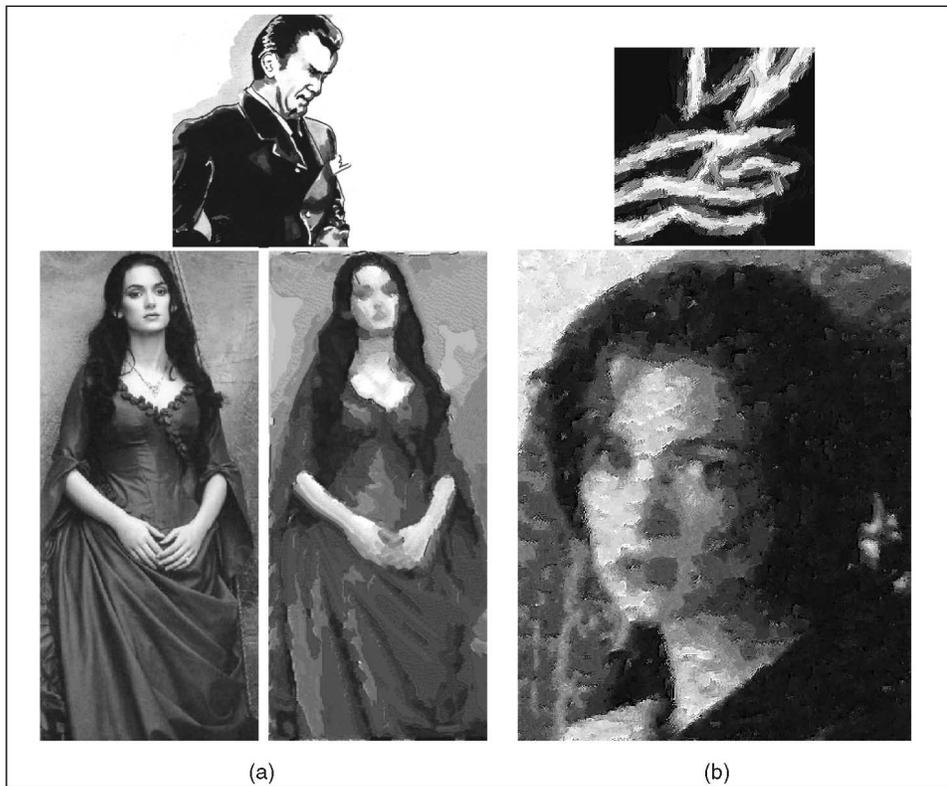


Fig. 8. (a) Drawing image with the ink painting style and rendering result ($\beta=1$). (b) Drawing image with the painting style and rendering result ($\beta=1$). (see also [http:// www.iro.umontreal.ca/~mignotte/ICPR02/](http://www.iro.umontreal.ca/~mignotte/ICPR02/), for additional and color rendering results).



Fig. 9. Texture sample (respectively: fire, water, clouds, satin) and rendering results ($\beta = 9$) based on each one.

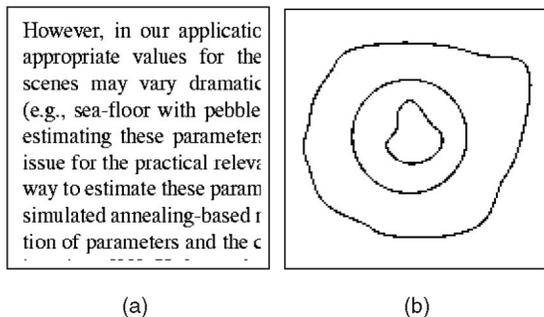


Fig. 10. Training samples used by the nonparametric example-based regularization energy term, respectively, for (a) the block letter segmentation problem and (b) the thin line segmentation problem.

are the optimal parameters in the two cases. For the local a priori model of the Markovian segmentation, we adopt a standard isotropic Potts model with the 8 connectivity spatial neighborhood (see (3)). In this comparison, we use $\beta_{st} = 1$ for the horizontal, vertical, and diagonal cliques. We use the stochastic procedure called simulated annealing [1] with 10,000 iterations (taking 10 minutes on a 1.2 Ghz PC) to minimize the energy function related to this model. We can notice that the results obtained by the Markovian segmentation are poor compared to the ones obtained by our model due to the unappropriated regularization term of Potts that tends to only impose homogeneous regions in the solution. By comparisons, synthetic textures, based on realizations of the nonparametric model itself (without data-driven term), show that the considered nonparametric model allows to efficiently and easily capture complex structure such as block letters or thin line segments (see the bottom left and the bottom middle texture of the Fig. 1).

In order to test the ability of our example-based method to generalize the training data, we present two segmentation results obtained for two noisy images exhibiting block letters with a different font (see Figs. 13b and 13f) and the resulting

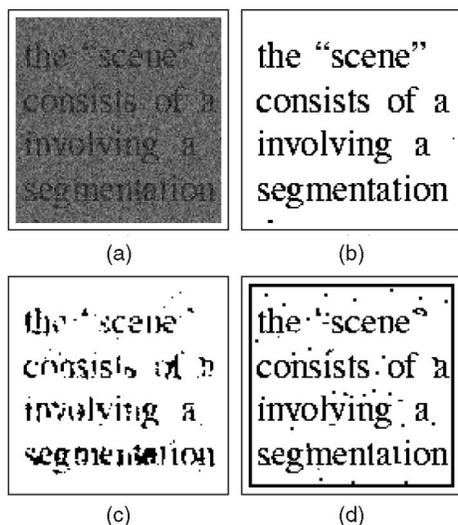


Fig. 11. (a) Image of block letters corrupted by an additive synthetic white Gaussian noise (with parameters $\mu_{e_0} = 55$, $\sigma_{e_0}^2 = 500$, $\mu_{e_1} = 100$, $\sigma_{e_1}^2 = 500$) and (b) corresponding ground truth segmentation, resulting segmentations obtained with (c) a Pott-type prior-based Markovian model compared to (d) our example-based method.

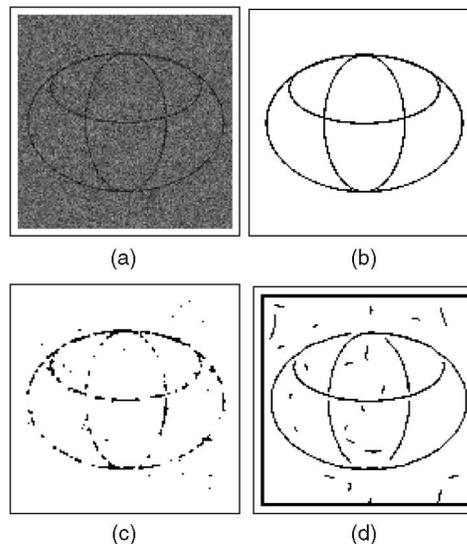


Fig. 12. Image made up of line segments corrupted by an additive synthetic white Gaussian noise (with parameters $\mu_{e_0} = 55$, $\sigma_{e_0}^2 = 500$, $\mu_{e_1} = 100$, $\sigma_{e_1}^2 = 500$), (b) corresponding ground truth segmentation, resulting segmentations obtained with (c) a Pott-type prior-based Markovian model compared to (d) our example-based method.

segmentations obtained with our example-based method compared to a Pott-type regularization energy term based on a Markovian model. Obtained segmentations with our method are of lower quality compared to the one obtained for a noisy image with a similar font, but remain better than a Markovian segmentation based on a Pott-type prior.

9.3 Shape Descriptor

In this application, we use two levels of pyramid and a 17×17 spatial and symmetric neighborhood in the fine level with a symmetric 17×17 spatial neighborhood in the coarse level. We set $\rho = 4$. Finally, we also use an ANN search technique [22] to accelerate the search of each nearest neighbor in (11). In order to reduce the computational cost, the first summation and the search of the min of (11) is over a subset of (regularly spaced) contour points on the query and the candidate shapes. The algorithmic procedure allowing to compute $\text{Dist}(q||c)$ is summarized in Table 2. A similarity measure takes about 75 ms to perform a single matching between two shapes on a 1.2 Ghz PC workstation. Due to the local characteristic of this measure, an implementation on parallel machines remains possible and could greatly reduce the computational time.

This shape similarity measure has been tested on the Snodgrass and Vanderwart line drawing database. This data set, frequently used in the psychophysics community for tests with human subjects, contains line drawings of 260 commonly occurring and distinct objects (only one image per object) [32]. In this test, no preprocessing phase is required for normalizing the shape size or for extracting the edges since the images are only line drawings of fixed size. In the case of raw images, a canny edge detector [33] would allow us to extract line features from the images and a pretreatment would be required for normalizing the size of the shapes. In

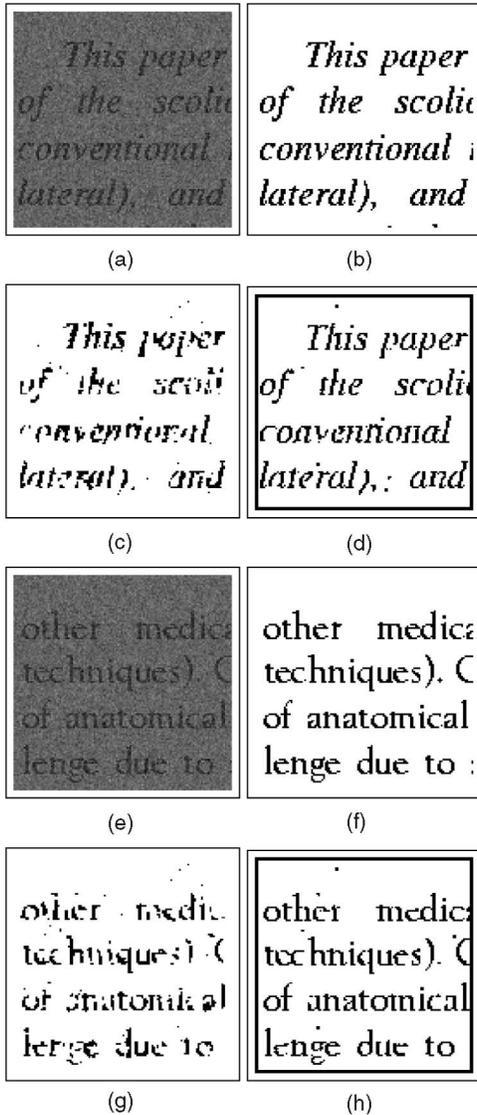


Fig. 13. (a) and (e) Images of block letters, with a different font of the considered training sample font, corrupted by an additive synthetic white Gaussian noise (with parameters $\mu_{e_0} = 55$, $\sigma_{e_0}^2 = 500$, $\mu_{e_1} = 100$, $\sigma_{e_1}^2 = 500$), (b) and (f) corresponding ground truth segmentation, segmentation results obtained with a Pott-type prior based on a Markovian model (c) and (g), and our method based on the training sample shown in Fig. 10a (d) and (h).

our application, we use only an Hotelling transform [34] in order to align the binary shapes with their principal axes. In order to also test the ability of this pruning method to deal with possible deformations and occlusions (i.e., contours partially detected) of the object class to be detected, we create, for each shape, a synthetic distorted and/or warped and/or occulted contour version that is added to the initial data set.

Fig. 14 shows some query shapes and a short list of candidate matches (i.e., the four most similar shapes ranked from 1 to 4) proposed by our pruning method. Our shape similarity measure is quite insensitive to the deformations and the distorted and/or occluded version of each shape is correctly selected in the four closest matches. Moreover, our pruning method allows us to find quite correctly visually similar shapes. Results of the ranking are quite consistent with the human visual system. This strategy of test is

TABLE 2
Multiscale Shape Similarity Distance

Multiscale Shape Similarity Distance	
Dist($q c$)	
q	Binary contour image (query shape)
c	Binary contour image (candidate shape)
Initialization:	
• $\Phi_q \leftarrow$ Compute the edge potential field from q (Eq. (9))	
• $\Phi_c \leftarrow$ Compute the edge potential field from c (Eq. (9))	
• $C_q \leftarrow$ Set of (regularly spaced) contour points of q	
• $C_c \leftarrow$ Set of (regularly spaced) contour points of c	
– $C_t \leftarrow C_q$	
– Dist($q c$) $\leftarrow 0$	
foreach contour point $s \in C_c$ do	
– $\hat{p} = \arg \min_{p \in C_q} D(\underbrace{N(\Phi_{q_p}), N(\Phi_{c_s})}_d)$	
– Dist($q c$) \leftarrow Dist($q c$) + d	
– Remove the contour point \hat{p} from C_q	
– if $C_q = \emptyset \Rightarrow C_q \leftarrow C_t$	

similar to the one adopted by Mori et al. for the evaluation of the discriminative power of their (context-based) shape descriptor and their results can be compared to ours since they use the same Snodgrass database [35]. Our shape descriptor produces a more accurate short list of candidate matches because our similarity measure is more sensitive to the local difference of two contour shapes. But, on the other hand, our shape similarity distance also requires (approximately) four times more computation.

Our method can be efficiently combined in a second stage with a more powerful and more time consuming comparison technique, such as a deformable template-based matching algorithm, to only the short list, in order to refine the search process.

10 CONCLUSION

In this paper, we have presented a example-based procedure for the rendering and segmentation problem. We have stated this problem in the multiresolution energy minimization framework and exploited the scale causal and hierarchical structure proposed in [12] and [13]. This procedure uses, for the modeling part, a multilevel nonparametric regularization term involving a spatial and scale-causal neighborhood and, for the algorithmic part, a multiresolution optimization technique involving a coarse-to-fine sequence of energy function minimization. From a modeling point of view, it offers the appealing ability to capture a priori characteristics of the solution image to be estimated within a range of different scales. From an algorithmic point of view, instead of considering the minimization problem on the full and original configuration space, the original inverse problem to be solved is decomposed in a sequence of approximated optimization problems of reduced complexity allowing to drastically save computational effort and/or to provide an accelerated convergence toward improved estimate.

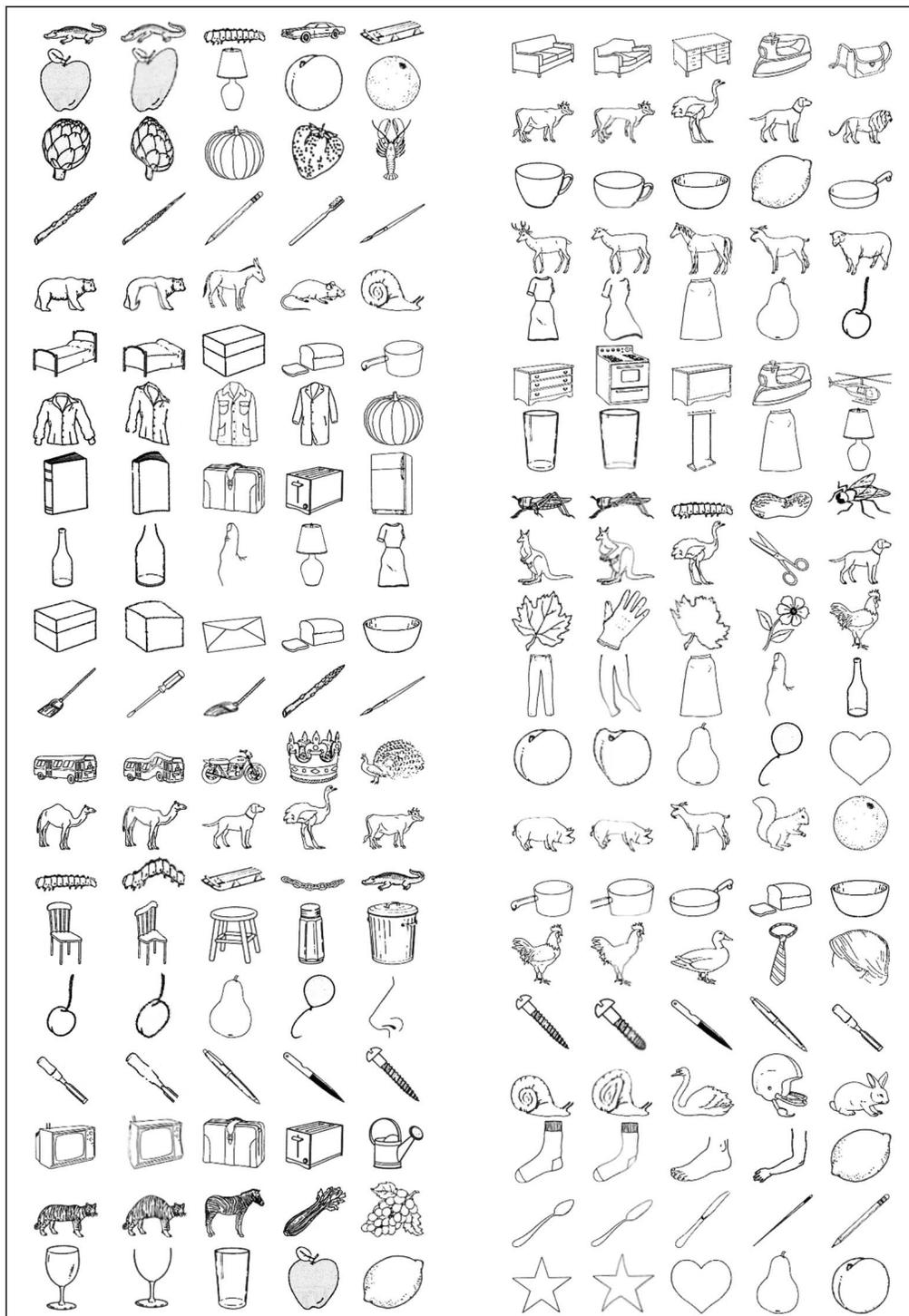


Fig. 14. Retrieval results. The first object of each column is the query shape. The remaining objects of each column show the closest four matches to each query object.

Applied as rendering procedure, the proposed scheme is able to re-render an input image in the style of another image allowing to create a very broad range of artistic and visual effects. Applied to image segmentation, the proposed technique allows us to recover complex configurations of the labeling process that cannot be expressed easily in mathematical form (for example, by classical Gibbs distributions). Finally, applied to pattern recognition, the definition of our example-based regularization term allows us to infer an

intuitive dissimilarity measure between two contour shapes. This measure is herein used to define an efficient shape descriptor for the rapid shape retrieval and indexing problem.

ACKNOWLEDGMENTS

The author is grateful to the anonymous reviewers for useful comments and suggestions on earlier drafts of this work.

REFERENCES

- [1] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721-741, 1984.
- [2] F. Heitz, P. Pérez, and P. Bouthemy, "Multiscale Minimisation of Global Energy Functions in Some Visual Recovery Problems," *CVGIP: Image Understanding*, vol. 59, no. 1, pp. 125-134, 1994.
- [3] E. Memin and P. Pérez, "Dense Estimation and Object-Based Segmentation of the Optical Flow with Robust Techniques," *IEEE Trans. Image Processing*, vol. 7, no. 5, pp. 703-719, 1998.
- [4] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Bayesian Inference and Optimization Strategies for Some Detection and Classification Problems in Sonar Imagery," *Proc. SPIE Conf. Non Linear Image Processing, Electronic Imaging*, pp 14-27, Jan. 1999.
- [5] J. Besag, "On the Statistical Analysis of Dirty Pictures," *J. Royal Statistical Soc., B*, vol. 48, pp. 259-302, 1986.
- [6] C. Kervrann and F. Heitz, "Statistical Model-Based Segmentation of Deformable Motion," *Proc. Int'l Conf. Image Processing*, pp. 937-940, 1996.
- [7] P.B. Chou and C.M. Brown, "The Theory and Practice of Bayesian Image Labeling," *Int'l J. Computer Vision*, vol. 4, pp. 185-210, 1990.
- [8] A.A. Efros and T.K. Leung, "Texture Synthesis by Non-Parametric Sampling," *Proc. Int'l Conf. Computer Vision*, pp. 1033-1038, Sept. 1999.
- [9] C. Regazzoni, F. Arduini, and G. Vernazza, "A Multilevel GMRF-Based Approach to Image Segmentation and Restoration," *Signal Processing*, vol. 34, pp. 43-67, 1993.
- [10] S. Krishnamachari and R. Chellappa, "Multiresolution GMRF Models for Image Segmentation," *IEEE Trans. Image Processing*, vol. 6, no. 2, pp. 251-267, 1997.
- [11] L. Wang and J. Liu, "Texture Classification Using Multiresolution," *Pattern Recognition Letters*, vol. 20, no. 2, pp. 171-182, 1999.
- [12] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Sonar Image Segmentation Using a Hierarchical MRF Model," *IEEE Trans. Image Processing*, vol. 9, no. 7, pp. 1216-1231, 2000.
- [13] L.-Y. Wei and M. Levoy, "Fast Texture Synthesis Using Tree-Structured Vector Quantization," *Proc. SIGGRAPH*, pp. 479-488, July 2000.
- [14] K. Popat and R.W. Picard, "Cluster-Based Probability Model and Its Application to Image and Texture Processing," *IEEE Trans. Image Processing*, vol. 6, no. 2, pp. 268-284, 1997.
- [15] H. Chen, Y.Q. Xu, H.Y. Shum, S.C. Zhu, and N.N. Zhen, "Example-Based Facial Sketch Generation with Non-Parametric Sampling," *Proc. Int'l Conf. Computer Vision*, pp. 433-438, July 2001.
- [16] A. Hertzmann, N. Oliver, B. Curless, and S.M. Seitz, "Curve Analogies," *Proc. Eurographics Workshop Rendering*, pp. 233-245, June 2002.
- [17] P. M. Jodoin, E. Epstein, M. Granger-Piché, and V. Ostromoukhov, "Hatching by Example: A Statistical Approach," *Proc. Non-Photorealistic Animation and Rendering*, June 2002.
- [18] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Unsupervised Hierarchical Markovian Segmentation of Sonar Images," *Proc. IEEE Int'l Conf. Image Processing*, Oct. 1997.
- [19] T.F. Chan et al. "Image Inpainting," <http://www.math.ucla.edu/~imagers/htmls/inp.html>, 2003.
- [20] C. Reynolds, "Stylized Depiction in Computer Graphics, NPR (Non Photo-Realistic Rendering), Painterly and Toon Rendering," <http://www.red3d.com/cwr/npr/>, 2003.
- [21] A. Hertzmann, C.E. Jacobs, N. Olivier, B. Curless, and D.H. Salesin, "Image Analogies," *Proc. SIGGRAPH*, pp. 327-340, July 2001.
- [22] 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," *J. ACM*, vol. 45, no. 6, pp. 891-923, 1998.
- [23] R.C. Veltkamp and M. Hagedoorn, "State of the Art in Shape Matching," Technical Report UU-CS-1999-27, Utrecht, 1999.
- [24] A.K. Jain, Y. Zhong, and S. Lakshmanan, "Object Matching Using Deformable Templates," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 3, pp. 267-278, Mar. 1996.
- [25] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Hybrid Genetic Optimization and Statistical Model-Based Approach for the Classification of Shadow Shapes in Sonar Imagery," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 2, pp. 129-141, Feb. 2000.
- [26] M. Mignotte, "A New And Simple Shape Descriptor Based on a Non-Parametric Multiscale Model," *Proc. IEEE Int'l Conf. Image Processing*, vol. 1, pp. 445-448, Sept. 2002.
- [27] P.F.M. Nacken, "Chamfer Metrics in Mathematical Morphology," *J. Math. Imaging and Vision*, vol. 4, pp. 233-253, 1994.
- [28] D.M. Gavrilu and V. Philomin, "Real-Time Object Detection for Smart Vehicles," *Proc. Int'l Conf. Computer Vision*, pp. 87-93, Sept. 1999.
- [29] M. Mignotte, "Bayesian Rendering with Non-Parametric Multiscale Prior Model," *Proc. IEEE Int'l Conf. Pattern Recognition*, vol. 1, pp. 247-250, Aug. 2002.
- [30] A.P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Royal Statistical Soc.*, pp. 1-38, 1976.
- [31] B. Braathen, P. Masson, and W. Pieczynski, "Global and Local Methods of Unsupervised Bayesian Segmentation of Images," *Graphics and Vision*, vol. 2, no. 1, pp. 39-52, 1993.
- [32] J.G. Snodgrass and M. Verderwart, "A Standardized Set of 260 Pictures: Norms for Name Agreement, Familiarity and Visual Complexity," *J. Experimental Psychology: Human Learning and Memory*, vol. 6, pp. 174-215, 1980. <http://www.cog.brown.edu/~tarr/stimuli.html>.
- [33] J. Canny, "A Computational Approach Edge Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679-697, 1986.
- [34] R.C. Gonzales and R.W. Woods, *Digital Image Processing*. Addison-Wesley, 1992.
- [35] G. Mori, S. Belongie, and J. Malik, "Shape Contexts Enable Efficient Retrieval of Similar Shapes," *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 723-730, Dec. 2001.



Max Mignotte received the DEA (postgraduate degree) in digital signal, image, and speech processing from the INPG University, France (Grenoble) in 1993 and the PhD degree in electronics and computer engineering from the University of Bretagne Occidentale (UBO) and the Digital Signal Laboratory (GTS) of the French Naval Academy, France, in 1998. He was an INRIA postdoctoral fellow at the University of Montréal (DIRO), Canada (Québec), from 1998 to 1999. He is currently with DIRO at the Computer Vision & Geometric Modeling Lab as an assistant professor (professeur adjoint) at the University of Montréal. He is also a member of LIO (Laboratoire de Recherche en Imagerie et Orthopédie, Centre de Recherche du CHUM, Hôpital Notre-Dame) and researcher at CHUM. His current research interests include statistical methods and Bayesian inference for image segmentation (with hierarchical Markovian, statistical templates or active contour models), hierarchical models for high-dimensional inverse problems from early vision, parameters estimation, tracking, classification, shape recognition, deconvolution, 3D reconstruction, and restoration problems.

► For more information on this or any other computing topic, please visit our Digital Library at <http://computer.org/publications/dlib>.