Adapting the Sample Size in Texture Synthesis

Roberto Costantini and Sabine Süsstrunk

***Abstract*— Starting from a sample of a given size, texture synthesis algorithms are used to create larger texture images. A good algorithm produces synthesized textures that are pixel-wise different but perceptually indistinguishable from the original image. The sample image should be chosen ensuring that it contains a number of pattern repetitions sufficient to produce valuable synthesis results. Since textures can be characterized by patterns of different dimensions, this must be done in an adaptive way.**

In this article, we propose a method that automatically adapts the sample size for natural textures synthesis, according to the different patterns dimensions. The method is based on the measure of the spatial dependence between the texture pixel values. This measure is used to estimate the size of the smallest texture window that is still perceived as texture by human observers. The sample size is determined from this measure by applying a multiplicative factor that depends on the algorithm used for synthesis. We perform a simple subjective experiment to estimate this factor for three different synthesis algorithms. We show that the measure of spatial dependence based on the correlation between pixels performs well when it is used to adapt the sample size.

I. Introduction

When we look to a set of similar objects, such as "a crowd of people, a bunch of bananas, a shelf of books, a line of cars" (Ariely [1]), we have the impression that they form a single global entity. Using the paradigms of member discrimination and identification, Ariely shows in fact that "the visual system represents the overall statistical, and not individual, properties of sets". A similar phenomenon happens when we look at textures, which indeed are perceived as globally uniform images.

The uniformity derives from the repetition of a given pattern, which is characteristic to the texture. In his early work on texture perception [2], [3] Julesz uses the term *texton* to refer to this pattern. Recent studies claim that textons are the fundamental micro-structures that constitute textures [4], but no globally accepted definition of texton exists, even though some proposals have been formulated [4], [5], [6], [7].

Texture perception is fundamentally influenced by the spatial frequency of the texton repetitions [8]. This can be shown when a texture is zoomed-in on a fixed window frame: at a certain zoom level, the perception of texture is lost, since few repetitions of the pattern are visible.

In this article we consider this fact in the framework of texture synthesis algorithms, where a texture image is taken as sample and used to synthesize larger textures. The sample must contain a number of texton repetitions that allows the algorithm to perform profitably. Here, we propose a method that automatically selects the smallest size of the sample that permits to obtain good synthesis performance. This corresponds to an adaptation to the scale of the texture.

The method consists in two steps. The first is the estimation of the smaller texture crop that is still perceived as texture by observers. We call W^P_{min} the size of this crop, where "P" stands for "perceptive". This is equivalent to find the scale limit factor after which the perception of texture is lost. The second step is the computation of the smallest sample size, called W^S_{min} where "S" stands for "scale". This is obtained by multiplying W^P_{min} by a factor that depends on the algorithm chosen for the synthesis. In the article, we found this factor for three different synthesis algorithms, thus comparing their performance.

W^P_{min} is computed using a measure of the spatial dependence between the pixels of the image, which can give an indication of the texton dimension. We notice, in fact, that completely random textures, which are characterized by independent pixels, can be considered as texture for almost all reasonable scales, as the texton reduces to a single pixel. On the contrary, for natural textures, which show local dependence between pixels and are usually perceived as texture for a limited range of scales. We consider three different measures of spatial dependence and we show that the one based on the simple autocorrelation function permits to obtain good estimation performance.

The article is structured as follows. In Section II we illustrate the methods used to estimate the spatial dependence between pixel values in the texture. In Section III we show the results of the experiment on texture perception that permits to create an ideal observer, which is defined in Section IV. In Section V we present the results of an experiment designed to test the performance of three texture synthesis algorithms and we show how an adapting sample size can be computed from the dependence measures defined in the previous section. Finally, Section VI concludes the paper.

II. Spatial Dependence

This section introduces three measures that estimate the statistical dependence between pixel values of images. The first measure is the Moran's I statistics, generally used in the analysis of geographical data, which is a particular type of spatial data [9]. It is based on the computation of the spatial autocorrelation between adjacent pixels. The second measure is based on the concept of mutual information derived from information theory, and the last measure is based on a test for independence generally used in stock market analysis to verify trends in data [10], [11].

A. Moran's I statistic

Let $\{G(s)\}_{s\in\S}$ be a real-valued stochastic process defined on a certain probability space $(\Omega, \mathcal{F}, P)$, where $S = \{(i,j) \in \mathbb{Z}^2 : i = 0,\dots,N-1; j = 0,\dots,M-1\}$ is a $N \times M$

lattice. In the case of geographical data analysis, where data are usually collected in column vectors, we denote the process $\{G(s)\}_{s\in\S^2}$ as $\mathbf{y} = \{y_i\}_{i=1,\ldots,L}$, where $L = N \times M$ is the cardinality of the set S. Since in $\mathbf{y} = \{y_i\}_{i=1,\ldots,L}$ the spatial contiguity between values is lost, a *spatial weighting matrix* (or *proximity matrix*) is introduced. L being the number of different locations, the proximity matrix is a $L \times L$ matrix denoted as $V = \{v_{ij}\} \in \mathbb{R}^{L\times L}$, where v_{ij} is fixed to zero if two locations i and j are not spatially linked, and $v_{ij} \neq 0$ otherwise. In geographical data analysis, there are different ways to establish the degree of proximity between different locations [12], [13].

Here, we define V such that $v_{ij} = 1$ if locations i and j are at pixel distance d one from each other, and $v_{ij} = 0$ otherwise. We indicate with $V(d)$ this matrix and with $v_{ij}(d)$ its elements. The Moran's I statistic is defined as:

$$I_{MO}(d) = \frac{\sum_{i,j=1}^{L} v_{ij}(d)(y_i - \overline{y})(y_j - \overline{y})}{S^2 \cdot n_d}. \tag{1}$$

where

$$S^2 = \frac{1}{L} \cdot \sum_{i=1}^{L} (y_i - \overline{y})^2 \tag{2}$$

is an estimate of the data variance, y_i is the data value at location i, $\overline{y}$ is the average data value over all locations, and

$$n_d = \sum_{i,j=1}^{L} v_{ij}(d)$$

is the number of locations that belong to the same distance class defined by d.

We notice that this function has exactly the same form as the autocorrelation function, since the numerator is a cross-product (covariance) term, while the denominator is a variance term.

It can be shown [14] that the expected value for Moran's I on L location is:

$$E_L[I(d)] = \frac{-1}{L-1}, \tag{3}$$

and thus it tends to zero as L increases. Moran's I varies on the interval $[-1, 1]$, with 1.0 and -1.0 indicating a perfect positive and negative correlation, respectively.

Fig. **??** show an example of the Moran's I statistic computed on textures (b1) and (b3) of Fig. 10, where d varies from 1 to 20 pixels. We notice that such statistics decay exponentially, according to the typical behavior found in natural images. It is known, in fact, that natural images exhibit a spatial power spectra that decays as $A/(|\omega|^{(2-\eta)})$ where $|\omega|$ is the magnitude of the spatial frequency and η is usually a small number [15], [16]. This implies that the spatial autocorrelation function decays exponentially. Since the Moran's I statistic is based on the computation of the spatial correlation between the image pixels, it also decays exponentially as the spatial distance d increases. For this reason, we express the Moran's I statistics using the following parametric model:

$$I_{MO}(d) = K_{MO} \cdot \exp(-\alpha_{MO} \cdot d), \tag{4}$$

where K_{MO} and α_{MO} are the model's parameters, and d is the pixel distance. The parameter α_{MO} indicates the slope of the Moran's I statistics and thus can be considered as an indicator of the spatial dependence between pixels. This is shown in Fig. 1(a), where the slope of the dashed curve, which corresponds to texture (b3), is more flat than that of the continuous line curve, corresponding to texture (b1). This means that in texture (b3) pixels exhibit a stronger dependence with respect to the other image. Images (b1) and (b3) of Fig. 10 are obtained from the same texture at two different scales (we changed the zoom factor of the digital camera).

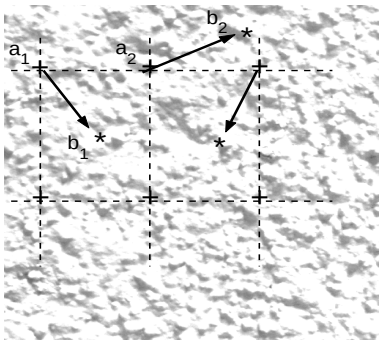

Fig. 2. Collection of the realization of the random variable A and B; a_i is placed on a square grid, while b_i is randomly taken at a distance d from a_i.

B. Mutual Information

A value close to zero of the Moran's I statistic does not indicate independence, but decorrelation. The two concepts are not equivalent in general, since two random variables (r.v.) can be decorrelated, but not independent. In order to have a better indicator of independence, we propose a measure based on the notion of mutual information between random variables. Mutual information is preferred to correlation because it is sensitive to more general dependencies [17]. Moreover, the idea of using a measure taken form an information-theoretic framework to test independence has already been applied to design statistical test for independence in economic-trend evaluations, obtaining good performance [18], [19], [20].

Let us consider two discrete r.v. A and B defined on the same set $J \subseteq \mathbb{Z}$. Their mutual information, denoted as $I(A; B)$, is defined as:

$$I(A;B) = -\sum_{a,b\in J} P(A=a, B=b) \cdot \log \frac{P(A=a)P(B=b)}{P(A=a, B=b)}. \tag{5}$$

If A and B are independent, then $P(A = a)P(B = b) = P(A = a, B = b)$ for all a and $b \in J$ and thus the mutual information reduces to zero. Otherwise, it has always a positive value, since it can also be defined as the difference between the entropy of A and the conditional entropy of A knowing B:

$$I(A;B) = H(A) - H(A \mid B). \tag{6}$$

This quantity is always positive.

Mutual information is used to test the dependence between pixel values at a certain distance from each other. Fig. 2 shows the method used to estimate the spatial dependence in texture images. A certain number of image locations is chosen, on a regular grid. Let us call these locations as a_n with $n =$

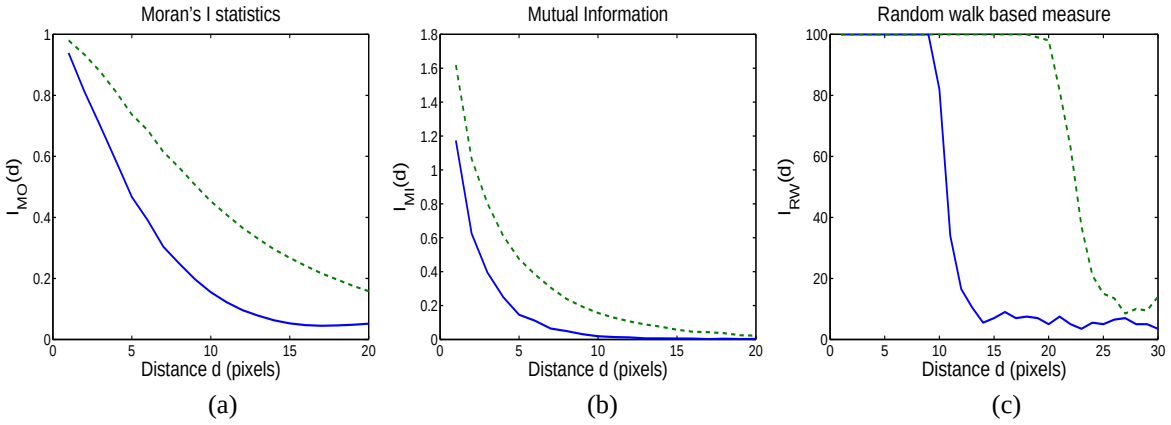

Fig. 1. Spatial dependence curves for texture (b1) (continuous line) and (b3) (dashed line) of Fig. 10: (a) Morans' I statistics; (b) Mutual information; (c) Random walk based measure.

$1, \ldots, N$. From each location a_i, a new location b_i is randomly chosen at distance d. The pixel values at the locations so far obtained can be viewed as the realizations of two random variables $A^d = \{G(a_i)\}_{i=1,\ldots,N}$ and $B = \{G(b_i)\}_{i=1,\ldots,N}$, where G indicates the stochastic process that generates the texture image. The mutual information between them can thus be computed according to Eq. (5) and constitutes a measure of the spatial dependence between the texture pixels. We indicate with $I_{MI}(d)$ this measure: $I_{MI}(d) = I(A^d; B^d)$.
Fig. 1(b) depicts for each spatial distance d an estimate of this measure, computed for textures (b1) and (b3) of Fig. 10, respectively. To compute the mutual information from a finite number of realizations of random variables A and B, we used the estimator of Moddemeijer [21], which is not based on a plug-in method (thus, it is not biased [22]) and can also be applied to dependent bivariate measures.

We notice that the mutual information decreases as the spatial distance increases. This is what we expected, since random variable realizations tend to become more independent as their distance increases. Analogously to the Moran's statistic, we have chosen to model this decay with an exponential law:

$$I_{MI}(d) = K_{MI} \cdot \exp(-\alpha_{MI} \cdot d) \tag{7}$$

where K_{MI} and α_{MI} are the model's parameters. This model turns out to be reasonable for many 2-D natural phenomena, where adjacent location exhibit spatial dependence. Examples can be found in the modelling of coupled map lattices with an Ising-like transition [23], of chaotic neural networks [24], and in the characterization of dynamical systems [25].

C. Random walk and run-up test for dependence

In [11], we proposed a method to establish the spatial dependence between pixel values in a texture. A texture image is scanned with a random path and pixel values are collected in a time-series. The random path ensures that successive samples of the time-series correspond to pixels at a fixed distance in the image. The dependence of the time-series samples is evaluated for different values of such distance, and a measure of dependence is thus obtained.

Fig. 3 depicts an example of a random walk in a finite dimension texture. Starting from the texture (2D signal), a

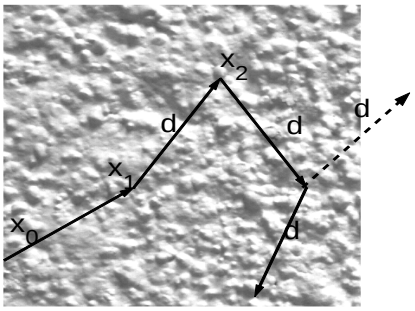

Fig. 3. Random walk scanning of a finite texture at distance d.

time-series (1D signal) is obtained. We indicate it as Y^d, since it contained the texture pixel values encountered along the random walk that are at an Euclidean distance d of the other.

The properties of the time-series Y^d obtained from the random walk are analyzed to study the spatial dependence between pixel values in a texture. One expects that samples of a time-series Y^d are more independent than those of a time-series $Y^{d'}$ if $d > d'$ and that this difference is characteristic to given texture.

Statistical dependence is analyzed using a *run-up test* [10], [11], which accepts or rejects the null hypothesis of sample independence. We indicate its output as:

$$T(Y^d) = \begin{cases} 1 \text{ if samples are dependent} \\ 0 \text{ otherwise} \end{cases} \tag{8}$$

where Y^d is the time-series that is tested.
Being a statistical test, $T(Y^d)$ is a random variable where, if samples are independent, the probability that its value is zero corresponds to the sensitivity the test [26].

In our implementation of the method, we consider finite length time-series, indicated with Y^d_K, where K is the number of samples[1]. We define the measure of spatial dependence as

$$I_{RW}(d) = E[T(Y^d_K)], \tag{9}$$

[1]In [11], we showed that a value of K that ensures reasonable results for a independence test is 6000

where the expectation is taken over the entire collection of time-series we can collect on the texture.
In practice, we only estimate this expectation using a finite number N of time-series of lengths K:

$$\hat{I}_{RW}(d) = \frac{1}{N}\sum_{i=1}^{N} T(\{Y^d\}_i), \tag{10}$$

where $\{Y^d\}_i$ indicates the time-series obtained from the i-th random walk.

Fig. 1(c) reports the measure computed for texture (b1) and (b3) of Fig.10. We notice that when distance d is small (5 to 10 pixels), pixels are very likely dependent, while when the pixel distance begins to increase, they tend to become more independent.
We model this as follows:

$$I_{RW}(d) = \frac{1}{\exp(\alpha_{RW}(d - \mu_{RW})) + 1}, \tag{11}$$

where μ_{RW} and α_{RW} are the model parameters. We use this particular function because it describes quite well the typical behavior of $I_{RW}(d)$.

III. Texture Perception Experiment

This experiment is designed to find the minimum window size that captures the characteristic of a given texture from a subjective point of view. This information will be used to compute the smallest sample image for the synthesis algorithms. The main idea of the experiment is depicted in Fig. 4.

Square crops of different size are cut from a given texture and displayed within a fixed-size (reference) window. Subjects have to state if the images displayed on the reference window are perceived as texture or not. In Fig. 4 we report, for instance, two crops of dimension W_1 and W_2, respectively. The image corresponding to the first crop is shrunk to fit the reference window W, while the other crop is enlarged. The first crop is an image still perceived as texture, while this is not true for the second crop. We expect, in fact, that when the scale is large, observers fail to perceive the image as a texture, because the number of texton repetitions within the window is small. By analyzing observers' responses, we can identify the scale limit that allows to have texture perception. We expect that this limit varies according texture structure and, consequently, of the spatial dependence between pixels.
We define the minimum window size that ensure texture perception as the smallest window that ensures a percentage of 75% of positive answers. Details of the experiments are given in Appendix I.

Results and Discussion

The experimental results are reported in Fig. 5. The abscissa indicates the ratio between the dimension of the original crop and its resized version when fitted to the reference window. The ratio is between the visual angle defined by the two images, defined as follows:

$$r = \frac{\arctan[W_i \cdot W_d/(2 \cdot N_x \cdot D)]}{\arctan[W \cdot W_d/(2 \cdot N_x \cdot D)]} \tag{12}$$

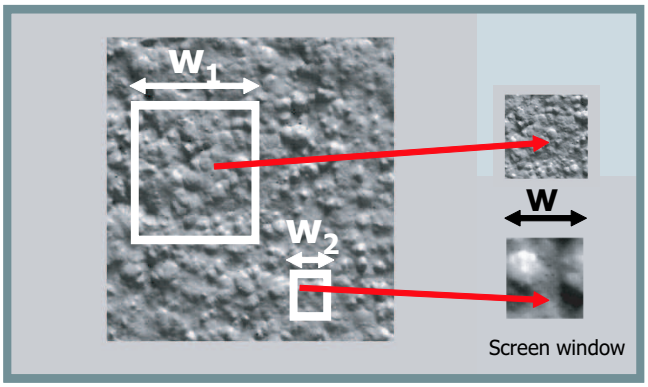

Fig. 4. Schematic representation of the hypothesis tested in the subjective experiment; two crops of different dimensions are considered (W_1 and W_2). Both are displayed in a fixed-size window. Observers are asked to state if the image displayed in the reference window W can be still considered as a texture.

where W_d is the reference screen width (cm), N_x is its horizontal resolution (pixels), D is the distance from the screen (cm), W_i is the crop dimension, and W is the dimension of the reference screen window. This value can be thought as a "spatial frequency" indicator. The results show that the percentage of positive responses is clearly sensitive to the dimension of the crops considered: the bigger the crop is, the better the image is perceived as texture. Fixing the dimension of the cropped window to a ratio equal to 2, for instance, the curves in Fig. 5(a) have bigger values than those of Fig. 5(b) and (c), for each scale considered. This is because the texton element in the first texture is bigger than that in the two other texture images.

It is clear that the observer responses are sensitive to texture scale. In the next we discuss how to link the results of the test with an estimation of the scale given by the spatial dependence measure of Section II.

IV. The Ideal Observer

The *ideal observer* is a function that links the measure of spatial dependence defined in Section II to the results obtained in the experiment of the previous section. This function estimates the probability that a crop taken from a given texture is perceived as texture itself. Since this corresponds to the task in the experiment described in the previous section, the ideal observer imitates observers' responses to this experiment. We indicate this function as $I_{\text{obs}}(r, \theta_1, \theta_2)$, where r is the ratio of Eq.(12), θ_1 and θ_2 are parameters that characterize the texture according to the measure of spatial dependence used. Mathematically, the ideal observer is defined as following:

$$I_{\text{obs}}(L, \theta_1, \theta_2) = \exp[-K_1\theta_1(L - K_2\theta_2^C)], \tag{13}$$

where $K1$, K_2, and C are given constants and the parameters are defined as:

$$\theta_1 = \begin{cases} \alpha_{MO} \text{ for Moran's I statistics} \\ \alpha_{MI} \text{ for mutual information} \\ \alpha_{RW} \text{ for RW-based measure} \end{cases} \tag{14}$$

$$\theta_2 = \begin{cases} \alpha_{MO}^{-1} \text{ for Moran's I statistics} \\ \alpha_{MI}^{-1} \text{ for mutual information} \\ \mu_{RW} \text{ for RW-based measure} \end{cases}$$

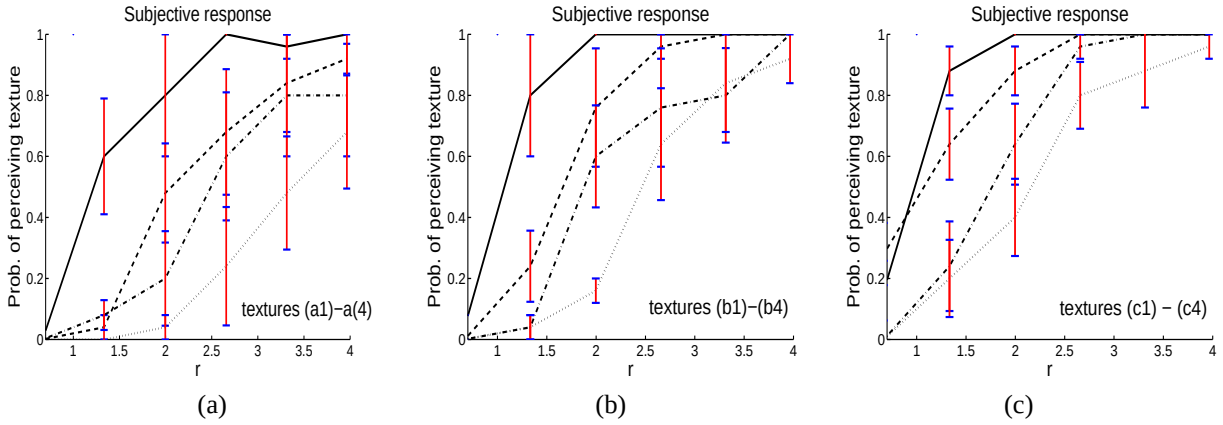

Fig. 5. Results of the experiment performed on textures of Fig. 10: (a) Results for the textures (a1)-(a4) (continuous, dashed, point-dashed, and point line, respectively); (b) Results for the textures (b1)-(b4); (c) Results for the texture (c1)-(c4);

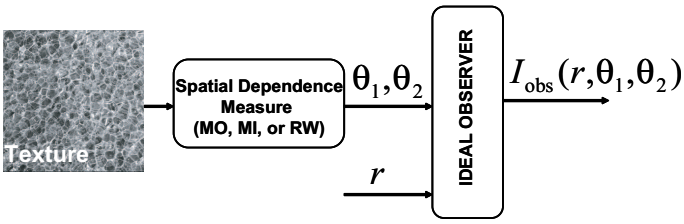

Fig. 6. Scheme illustrating the use of the ideal observer and the measure of spatial dependence; two parameters are extracted from the texture and they are used to predict observers' responses in term of probability of texture perception for a given ratio r.

Fig. 6 illustrate the concept of the ideal observer.
For each texture, the spatial dependence is computed using one of the three methods defined in Section II. According to Eq.(14), the parameters θ_1 and θ_2 are computed. Using Eq.(13), we compute the probability that a square crop of a given ratio r is perceived as texture itself by a human observer. When using the reference screen, it is possible to pass from the value of r to pixel values, using Eq.(12).

The parametric model of Eq. (13) imitates the typical shape of subject's responses found from the behavioral test in the previous experiment. The parameter θ_1 controls the slope of the curve and is directly linked to the α parameters of the spatial dependence models. The parameter θ_2 is inversely proportional to α_{MO} and α_{MI}, and directly proportional to μ_{RW}. The reason for this can be seen in Tab. I. Here we report the results obtained for the measures of the spatial dependence for the 12 textures used in the perceptual experiment. The images correspond to four increasingly zoomed pictures of three different textures respectively. Spatial dependence changes as the zoom factor increases, and this can be noticed in the decrease of α and the increase of μ_{RW}. In fact, when the zoom factor has a large value, the texture exhibits greater spatial dependence and the slope of the spatial dependence measure increases, since the transition between dependence and independence is larger. Analogously, μ_{RW} increases. The choice of θ_1 and θ_2 follows this behavior, since θ_1 decreases with scale, while θ_2 increases.
In the formula of the ideal observer, we notice that parameters K_{MO} and K_{MI} of Eq. (4) and Eq. (7) are not considered. From Tab. I we see in fact that those parameters are almost constant for all textures and are not sensitive to scale variation. Using a sign test and a one-sample z-test, we verified that these parameters do not change significantly: they can be considered as constant and combined with the model constants.

The constant K_1, K_2, and C of the model are estimated by minimizing the distance between human and calculated responses of the experimental results. This estimation is done for each spatial dependence model using a least square non-linear regression technique. The estimated values are given in Tab. II, together with their respective standard deviation. These have been obtained by a leave-one-out validation technique, where we fitted the model N times, each time using N-1 data, thus finding N-1 parameter estimates and then computing the variance of such estimates. From the ideal observer, we reproduced the responses of the experiment obtained by human observers. A rank test for statistical equivalence was used to compute the percentage of statistically equivalent responses on the entire set of textures. For each algorithm a percentage of 98% of statistical equivalent responses was obtained.

The minimum size of the window that ensures texture perception is defined as the size that ensures a probability equal to 0.75 to perceive texture: $I_{OBS}(r_{min}) = 0.75$. Using the reference screen, we calculate the size of this window in pixels, defined as W^P_{min} (see Section III. This value will be used to estimate the size of the smallest texture sample in texture synthesis.

V. Texture Synthesis

Starting from a texture sample, texture synthesis algorithms generate a larger texture that should be visually indistinguishable from the original sample. The result of the synthesis depends on the algorithm used and on the size of the sample image, which has to be big enough to contain the texture characteristics.

In this section, we describe the result of a subjective test that aims at finding the minimum sample size that permits to obtain visually indistinguishable synthesized textures. The main idea of the test is depicted in Fig. 7. Texture samples are

	MO		MI		RW	
texture	$\hat{\alpha}_{MO}$	$\hat{K}_{MO}$	$\hat{\alpha}_{MI}$	$\hat{K}_{MI}$	$\hat{\alpha}_{RW}$	$\hat{\mu}_{RW}$
a1	0.1215	1.1764	0.3515	1.8715	0.0394	48.5249
a2	0.0849	1.1898	0.2350	1.9944	0.0440	64.7567
a3	0.0668	1.1715	0.1913	1.9350	0.0284	85.1662
a4	0.0510	1.1143	0.1556	1.9126	0.0025	40.1285
b1	0.1852	1.1629	0.5203	1.9331	0.0469	33.8353
b2	0.1291	1.2242	0.3458	2.0299	0.0388	48.2233
b3	0.0972	1.1601	0.2815	2.0261	0.0336	58.4203
b4	0.0752	1.1418	0.2169	1.8450	0.0216	78.8393
c1	0.2357	1.2852	0.6158	1.8965	0.0824	25.4028
c2	0.1720	1.2628	0.4686	1.9059	0.0632	34.2284
c3	0.1388	1.2200	0.3742	1.8255	0.0531	43.9571
c4	0.1112	1.1897	0.3087	1.8695	0.0449	52.3728

TABLE I

ESTIMATES OF THE TEXTURE PARAMETERS FOR TEXTURES OF FIG. 10 OBTAINED USING THE THREE MEASURES OF SPATIAL DEPENDENCE PROPOSED.

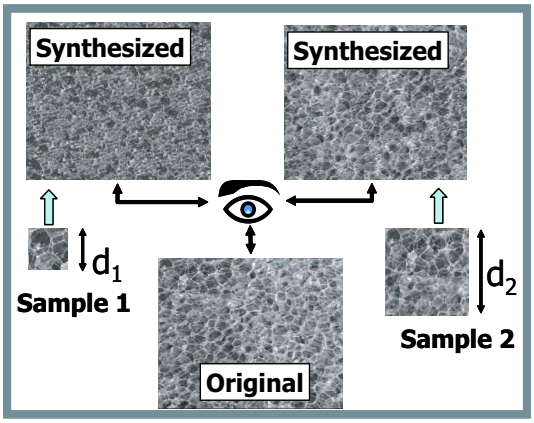

Fig. 7. Schematic representation of the subjective web-experiment for texture synthesis.

	MO	MI	RW
$\hat{K}_1$	$23.9418 \pm 3\%$	$8.5405 \pm 2\%$	$66.0958 \pm 4\%$
$\hat{K}_2$	$0.2811 \pm 4\%$	$0.6208 \pm 3\%$	$0.0540 \pm 10\%$
$\hat{C}$	$0.8464 \pm 2\%$	$0.9138 \pm 2\%$	$0.8852 \pm 3\%$

TABLE II

ESTIMATES OF THE IDEAL-OBSERVER MODEL PARAMETERS OBTAINED USING THE THREE MEASURE OF SPATIAL DEPENDENCE PROPOSED (INDICATED AS MO, MI, AND RW); THE VARIANCE OF THE ESTIMATE IS INDICATED IN PERCENTAGE OF THE PARAMETER VALUE.

taken from a given *original* texture. The samples correspond to a region of the original texture. The samples are then used to synthesize a texture having the same dimension as the original. For instance, Fig. 7 shows two samples of different size d_1 and d_2, together with the synthetic textures obtained from each sample respectively. The experiment consists of asking observers if they perceive them similar to the original images. Different textures and samples sizes have been used and, as a result, the minimum sample size was defined as the one that ensures a percentage of positive responses of 75%.

In this Section we show the results obtained in the subjective test and the link that exists between the minimum sample size and the minimum size window that is estimated by the ideal observer of Section IV. We show that a direct proportion can be established between the two window sizes. This proportion can be seen as a *score* that permits to evaluate a given algorithm, since it indicates how much texture information is needed by the algorithm to produce good results with respect to the information that observers need to have to perceive it as texture. Also, we show for a certain number of the texture considered, the score can be used to find in an adaptive way the size of a minimum sample size by using a measure of the spatial dependence defined in Section II.

A. Texture Synthesis Algorithms

Synthesis algorithms are mainly divided into two categories: model based and patch-based methods.

In the first category, textures are synthesized using a model derived from a sample texture representative of a texture class. Heeger and Bergen [27] proposed to use Laplacian and steerable pyramids to analyze textures, reconstructing similar textures by matching the same histogram of the filtered coefficients through an iterative process. A similar method was used by Portilla and Simoncelli [28], [29]. They achieved impressive results for the synthesis of stochastic textures, but the method failed for more structured textures. A different model was derived from the Markov random field theory. The work of Popat [30] is the first attempt to model a texture as a Markov random field, finding the conditional probability function and synthesizing a texture by sampling from this probability. Faster methods were proposed by Paget [31], [32], [33], who introduced the top-down multiresolution approach, where the texture characteristics are combined into a synthetic texture going from low to high frequency content progressively. A convincing method was then proposed by Zhu et alt. [34], [35] to connect the filter bank approach with Markov random field models, obtaining a model called FRAME. This constitutes a Markov random field with a richer vocabulary and thus with stronger ability to model textures. The main problem, which was partly addressed later [36], [37], remains the speed of the synthesis.

It is exactly to increase speed that the patch-based approach to texture synthesis was proposed. In this approach, a given patch of the sample texture is cropped and replicated over space in order to recreate bigger textures. The first methods were proposed by Xu et al. [38], [39]. The main problem of such an approach is to find the correct superposition of patches that ensures the smallest error and thus the least annoying effect to observers. A first convincing solution was published by Ashikhmin [40]. He proposed an effective measure of the visual error in patch alinement. Other algorithms used this measure to target other problems of patched-based synthesis algorithm, such as increase in speed ([41], [42], [43], [44]), or accuracy ([45], [46]). Generally, patch-based methods permit to obtain good synthesis performance even for structured textures, and are considered to be better methods than the model-based one. However, it is not possible to use them for different applications than synthesis, as for instance for texture segregation, recognition, and so on.

B. Testing Texture Synthesis with a Web-Based Visual Experiment

We test three synthesis algorithms: the model-based algorithm of Simoncelli and Portilla [29], and two recent patch-based algorithms from Ashikhmin [40] and Zelinka [41], respectively. The algorithm of Zelinka is based on the same principle of Ashikmin's algorithm, but uses a faster method to match the texture patches in a suboptimal way. This method permits to obtain good quality results with a significant increase in synthesis speed.

We synthesized texture images starting from samples of different sizes. We expect that, starting from a small sized sample does, the algorithm produces poor visual results, since the few textural information is contained.

The experiment can be found on Web [2], where people can perform it by following the instructions and finally stating if synthesized images look similar to original ones. Web experiments obtained wide attention in the last years [47], because they have the advantage of being accessible on the web, and can be accessed by many observers. The experiment set-up is described in Appendix II.

Gathering and fitting the data

42 observers participated to the experiment: 10 were familiar with image processing and 32 were naive observers. For each sample size, texture class, and algorithm used, we computed the percentage of people that considered the synthetic textures similar to the original ones.

In Fig. 8 we display this percentage (called *performance*) as a function of the size of the sample. The stimulus intensity corresponds to the size of the sample texture used for the synthesis. The values have been fitted using a Weibull psychometric function [48], [49] (dotted line), in order to find the sample size value that ensures a performance of 75%. We define this as the threshold value between the perception of similar and different texture pairs. We used the Psignifit toolbox for Matlab to compute the threshold.

The results

The threshold values obtained for Brodatz and MT textures are reported in Tab. III and Tab. IV, respectively. For a more direct interpretation of the data, the threshold is indicated in pixels and not in degrees of visual angle. The symbols "POR", "ASH", and "ZEL" indicate the synthesis algorithm of Portilla, Ashkhmin, and Zelinka, respectively.

Table III reports the results corresponding to 11 Brodatz textures out of the 13 used in the experiment. The fitting of the data did not work for two textures, indicated in Fig. 9, where observers always perceived the synthetic texture as different from the original one. Table IV reports the threshold values obtained considering the MT texture. The symbol "/" indicates that we did not considered the texture because even using the full 350×350 pixel image as sample for the synthesis, the result obtained was poor. This specially happens when using the model-based synthesis algorithm, as This was expected, since when applied to more regular textures model based algorithms perform less efficiently than patch-based algorithms.

[2]It can be found at http://ivrgwww.epfl.ch/SynthWebExp/main.html

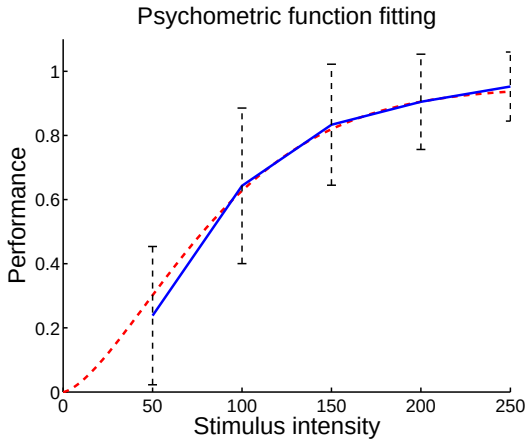

Fig. 8. Psychometric fitting of observers' responses. The data corresponds to the responses given by observers on a Brodatz texture. Dotted line: Weibull psychometric function; continuous line: linear interpolation between data; vertical dotted lines: confidence interval for the responses.

Texture n.	ASH	JMP	POR
1	88	51	98
2	120	221	286
3	47	149	30
4	52	94	39
5	53	53	30
6	56	34	34
7	123	145	121
8	81	170	61
9	50	49	35
10	53	48	23
11	128	164	257

TABLE III

THRESHOLD SIZES FOR THE BRODATZ TEXTURE CONSIDERED; THE VALUES CORRESPONDS TO THE SAMPLE SIZE THAT ENSURES THAT ABOUT 75% OF OBSERVERS CONSIDERED THE SYNTHETIC AND ORIGINAL TEXTURES AS SIMILAR.

Algorithm score

Using the minimum-sample window, we can define a score for a given algorithm. We indicate the minimum window size obtained from the ideal observer as $(W^P_{min})_i$, where $i \in \{$"MO", "MI", "RW"$\}$ is the index indicating the spatial dependence measure used and $(W^S_{min})_j$ the minimum size for the synthesis algorithms, where $j \in \{$"POR","ASH","ZEL"$\}$ is the algorithm used for synthesis. For each combination of dependence measure and synthesis algorithm used, we can thus defined a ratio between the two windows:

$$\gamma_i^j = (W^S_{min})_j / (W^P_{min})_i. \quad (15)$$

This ratio is a gain or *score* obtained using the synthesis algorithm with respect to human performance. In Tab. V we indicate the (average) values of the ratios obtained from Brodatz and MT texture, respectively and we also show the standard deviation between scores when using a given algorithm. The smaller the score, the less information the synthesis algorithm needs.

Texture n.	ASH	JMP	POR
1	237	265	883
2	223	257	217
3	224	/	/
4	90	60	/
5	268	/	/
6	207	/	/
7	194	/	/
8	463	/	/
9	463	364	/
10	252	397	/
11	539	/	/
12	217	181	150
13	342	370	/
14	264	345	/
15	320	293	/
16	136	189	/
17	326	/	/
18	297	324	/
19	423	297	/
20	88	77	78
21	76	52	80
22	111	303	73
23	114	593	150

TABLE IV

RESULTS OBTAINED FROM THE WEB-EXPERIMENT FOR THE MT TEXTURES; THE VALUES ARE INDICATED IN PIXELS

	Brodatz Textures			**MT Textures**		
γ_i^j	**ASH**	**ZEL**	**POR**	**ASH**	**ZEL**	**POR**
MO	0.47	0.72	0.78	0.53	0.74	1.01
MI	0.44	0.64	0.88	0.50	0.72	1.13
RW	0.45	0.63	0.89	0.44	0.64	1.20
std.	20%	50%	30%	30%	50%	90%

TABLE V

SYNTHESIS ALGORITHMS SCORE FOR BRODATZ AND MT TEXTURES; THE AVERAGE SCORE VALUE IS INDICATED TOGETHER WITH ITS STANDARD DEVIATION PER ALGORITHM USED.

The ratio can be also seen as a multiplicative factor applied to $(W_{min}^P)_i$ to adapt the size of the sample used for synthesis of the particular texture considered.

Discussion

Synthesis algorithms are compared using Table V. A first comparison is made between the "ASH" and the "JMP" algorithms. Tab. V shows that the performance of "JMP' are slightly worse than those of "ASH", where the score has a smaller value both for Brodatz and MT texture collections. This indicates that the a bigger sample size is necessary to synthesize textures using "JMP" algorithm. This can be explained by the presence of visible artifacts in some synthetic texture obtained using the "JMP" algorithm, as shown in Fig. 11.

In Fig. 11(a) we report the result of the "ASH" algorithm on a foliage texture, while in Fig. 11(b) the result obtained using the algorithm "JMP' is given. We notice the presence of certain artifacts in the texture. The "JMP" algorithm, in fact, is based on a faster but suboptimal search procedure for the best match between texture patches. Since observers were allowed to look at the original-synthetic texture pairs for as long as they want, those artifacts were easily observed and caused a large

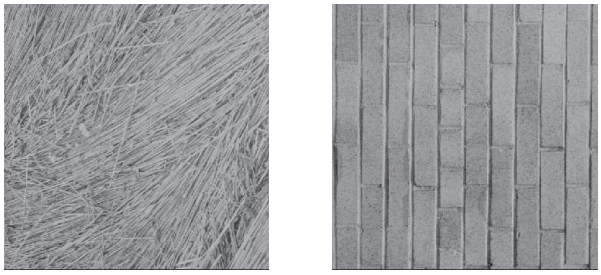

Fig. 9. Brodatz textures for which the data obtained from the web-experiment did not fit the psychometric function.

percentage of negative responses when the similarity between texture pairs were tested.

For the Brodatz collection, the results obtained using the model-based method of Portilla and Simoncelli ("POR") are better than those obtained by the two other methods, while they are worse when considering the MT collection. This is because the Brodatz textures that we considered were less structured than those belonging to the MT collection. This confirms the fact the model-based algorithms perform less well than patch-based algorithm for structured textures. As an example, we show in Fig. 11(a) the result obtained by "POR" algorithm on the foliage and pasta textures, obtained using a 350×350 sample size.

The score for the "ASH" algorithm are quite similar both for Brodatz and MT collection, and are close to 0.5. This algorithm shows a constant performance over a larger collection of texture, regardless of the method of spatial dependence used to characterize the texture structure. This does not happens with the other algorithms, which have a smaller score for Brodatz than for MT texture collection. In the case of "JMP" algorithm this is due to the creation of artifacts, while in the case of "SAP" this is due to the poor performance on structured images.

Table V also shows that the results do not vary according to the method use to measure the spatial dependence between pixels values. This can be seen by inspecting the columns of Table V. The values do not change too much when different algorithms and different test texture are used. In fact, the ideal observer should not be influenced by the particular measure of spatial dependence, since the models has been obtained from the single psychophysical experiment of Sec. III. The measure based on the mutual information (MI) seems more conservative with respect to the other measures (bigger values of the score) for the Brodatz, but this does not happens for MT collection, where the MO gives higher score. Dependence measures can be considered as comparable and the choice between them can be made according to the computational complexity issues. For this reason, the measure based on the Moran's I statistics reveals to be the most indicated, since it permits to obtain the best compromise in term of quality of results and computational complexity.

From Tab. V we notice that the standard deviation of the score has a relevant value, globally bigger for the MT than for the Brodatz textures. This reflects the uncertainty and arbitrariness of certain responses obtained in the web experi-

ment, especially when more structured textures were presented (MT textures). The "JMP" method, where the presence of visible artifacts clearly distracted the observers, exhibits a higher score variability than the "ASH" method. In the case of the Brodatz collection it has the higher standard deviation. In the case of MT collection, the "POR" method has the higher standard deviation. We already noticed that this method is in fact less adapted to strongly structured textures.

To take into account the variability of the score, we considered a modified score, which is more conservative. It is obtained by adding 50% of their value to the scores of Tab. V. In the case of the "ASH" method this is equivalent to an overestimation of the real score, while in the case of the "POR" algorithm to an underestimation.

Adapting the sample size

The algorithm score is used to adapt the size of the samples used to synthesize a given texture. The way this id done is resumed as follows:

- We consider a measure of spatial dependence and compute the parameters θ_1 and θ_2;
- We choose a range of values for r, according to the possible sample size we consider;
- We compute W^P_{min} using the ideal observer's model of Eq.(13);
- We choose an algorithm and use the score γ to compute $W^S_{min} = \gamma \cdot W^P_{min}$.

Fig. 12 reports, for instance, the results obtained using two images taken randomly from the Brodatz and MT collection (BT n.4 and MT n.20). Here we notice that in the case of the Brodatz texture, the window estimated by the ideal observers had a size of 100 pixels. Using a score factor of 0.5, as suggested by Tab. V, we obtained a minimum sample size of 50 pixels, which is similar to the one obtained in the web-experiment (52 pixels). We see that the synthesized texture looks similar to the original one. The same is done for texture MT, where the estimated window size is 180 pixel and we use a ratio almost equal to one, as suggested by Tab. V.

VI. Conclusions

We have proposed a method to estimate the smallest sample size for texture synthesis. This method is based on the construction of an ideal observer and the measure of spatial dependence between texture pixel values. Using the results obtained from a web-experiment on texture synthesis quality assessment, we defined a score for the synthesis algorithm and show how to compute the smallest sample size using this score. We show that the score can predict algorithms' performance when texture with different structure are used. The measure based on simple autocorrelation function give results comparable to that obtained by more complex measures. It can thus be considered as a good compromise between performance and computational cost.

Appendix I

We define a texture as an image having at least two out of the three following characteristics:

- quasi-periodicity;
- randomness;
- visual redundancy.

To perform the experiment, we used the three textures shown in Fig. 10, indicated as (a1-a4), (b1-b4), and (c1-c4), respectively. They consist of 12 images, derived from 3 textures at 4 different scales obtained with an optical zoom of a digital camera. The size of the images is 800×1200 pixels. We performed the experiment on a CRT monitor with 4:3 aspect ratio, a resolution of 1280×1024 pixels and a diagonal width of 45 cm (called *reference* screen). Subjects were at a distance of 50 cm from the screen in a dim light conditions. For each texture, we randomly selected 5 different crops of dimension varying from 100 to 600 pixels. Each image was displayed on a 150×150 pixel reference screen window. When resizing the test images within this window, we checked that the image quality was not degraded by the interpolation. Four naive and one experienced observers were tested. The task was to state if the crop displayed on the screen window can be considered as texture according to the given definition.

Appendix II

We used 36 original texture images, composed of 13 Brodatz textures (indicated here as BT) and 23 textures that we collected (indicated here as MT). The original size of the Brodatz textures was 512×512 pixels. Images were resized to 350×350 pixels to make them more suitable for an average size computer display. MT textures, whose size was 800×1200, were cropped to 800×800 and then resized to a 350×350 pixel size. The 36 images of size 350×350 pixel were considered as original textures.

From each original texture, 7 sample images were created, with dimension varying form 50×50 to 300×300 pixels, with an interval of 50 pixels[3]. This was simply done by cutting out a square block from the texture starting from its top-left corner location and taking a block of the specified dimensions.

We obtain a total number of 252 sample textures. For each sample, one 512×512 synthetic texture was generated for each algorithm considered, resulting in 756 synthesized textures. Each of them was finally cropped to fit a 350×350 window, thus ensuring that synthetic and original images have the same size.

The original and the synthesized images were then displayed on the computer screen during the web-experiment and the observer's task was to state if they could be considered as similar. The subject were instructed about the meaning of the word "similar" in the introduction. We discarded the texture pairs that gave obvious responses, thus reducing the number of comparisons to 399. The images appeared in random order, and we ensured that they had the same size for all observers, independently of screen resolution or dimension. This was done using a simple experiment in the preliminary part of the

[3]Window sizes chosen when using Portilla's synthesis algorithm were taken form crops of size 64, 128, 192, 256, and 320 pixels respectively; in fact, algorithmic constrains imposed a size multiple of 32 pixel.

test, where we asked users to measure a red square displayed on their screen. This measure was used to remotely adjust the image dimension on the screen. We assumed that all observers performed the experiment with a fixed distance from the screen of 50 cm.

ACKNOWLEDGMENTS

We would like to thank Prof. Herzog for the helpful discussions about texture and minimum window size we had. Moreover, we would like to thank all the participants to the web-experiment.

REFERENCES

[1] D. Ariely, "Seeing sets: Representation by statistical properties," *Psychological Science*, vol. 12, no. 12, pp. 157–162, March 2001.

[2] B. Julesz and J. Bergen, "Texton, the fundamental elements in preattentive vision and perception of textures," *The Bell System Tech. Journal*, vol. 62, no. 6, pp. 1619–1645, July-August 1990.

[3] B. Julesz, "Texton gradients: The texton theory revisited," *Giol. Cybern.*, vol. 54, pp. 245–251, 1986.

[4] S.-C. Zhu, C. Guo, Y. Wu, and Y. Wang, "What are textons," *Proc. of 7th European Conf. on Computer Vision*, May-June 2002.

[5] B. Olshausen and D. Field, "Sparse coding with an overcomplete basis set: A strategy employed by v1?" *Vis. Res.*, vol. 37, no. 23, pp. 3311–3325, 1997.

[6] T. Leung and J. Malik, "Recognizing surface using three-dimensional textons," *Proc. of 7th ICCV*, 1999.

[7] F. B. and J. N., "Transformed component analysis: joint estimation of spatial transforms and image components," *Proc. of Int'l Conf. on Comp. Vis., Corfu, Greece*, 1999.

[8] F. Kingdom and D. Keeble, "On the mechanism for scale invariance in orientation-defined textures," *Vision Res.*, vol. 39, no. 8, pp. 1477–1489, 1999.

[9] N. A. Cressie, *Statistics for Spatial Data*. John Wiley and Sons, Inc., 1993, ch. 6, p. 427.

[10] E. Chen and W. Kelton, "Determining simulation run length with the runs test," *Modelling Practice and Theory*, vol. 11, pp. 237–250, 2003.

[11] R. Costantini, G. Menegaz, and S. Susstrunk, "A measure for spatial dependence in natural stochastic textures," *IEEE International Conference on Image Processing*, pp. 1525–1528, 2004.

[12] L. Anselin and A. Getis, "Spatial statistical analysis and geographical information systems," *Annals of Regional Science*, vol. 22, pp. 509–536, 1992.

[13] W. Mitchell and A. Bill, "Spatial dependence in regional unemployment in australia," *A Future that Works - economics, employment and the environment Conference*, vol. Working Paper No. 04-11, Dec. 8-10, 2004.

[14] D. Barkley, H. Mark, S. Bao, and B. Kerry, "How functional are economic areas: Tests for spatial association using gis-based analytic techniques," *Papers of Regional Science*, vol. 74, pp. 1–19, 1995.

[15] D. Ruderman, "Origins of scaling in natural images," *Vision Res.*, vol. 37, no. 23, pp. 3385–3398, 1997.

[16] S. A., L. A.B., S. E.P., and Z. S.-C., "On advances in statistical modeling of natural images," *Journal of Math. Imaging and Vision*, vol. 18, pp. 17–33, 2003.

[17] T. Schreiber, "Spatio-temporal structure in coupled map lattices: two-point correlations versus mutual information," *Journal of Physics A: Mathematical and General*, vol. 23, pp. 393–398, 1990.

[18] P. Robinson, "Consistent non-parametric entropy-based testing," *Review of Economic Studies*, vol. 58, pp. 437–453, 1991.

[19] H. Skaug and D. Tjostheim, *Nonparametric tests of serial independence*, 1993.

[20] C. Tsallis, "Generalized entropy-based criterion for consistent testing," *Physical Review E*, vol. 58, pp. 1442–1445, 1998.

[21] R. Moddemeijer, "The variance of the mutual information estimator," 1997, cS-R9701. [Online]. Available: http://www.cs.rug.nl/ rudy/papers/abstracts/RM9705.html

[22] I. Nemenmann, "Entropy and information estimation: an overview," *NIPS'03 Entropy Estimation Workshop*, 2003. [Online]. Available: http://www.menem.com/ ilya/pages/NIPS03/nemenman.pdf

[23] C. O'Hern, D. A. Egolf, and H. Greenside, "Lyapunov spectral analysis of a nonequilibrium ising-like transition," *Phys. Rev. E*, vol. 53, pp. 3374–3386, 1996.

[24] S. Mizutani, T. Sano, T. Uchiyama, and N. Sonehara, "Lyapunov exponents and mutual information of chaotic neural networks," *Neural Networks for Signal Processing VI. Proceedings of IEEE Signal Processing Society Workshop*, pp. 200–209, 1996.

[25] J. P. Crutchfield and D. P. Feldman, "Regularities unseen, randomness observed: The entropy convergence hierarchy." *Chaos*, vol. 15, pp. 25–54, 2003.

[26] T. Anderson and J. Finn, *The New Statistical Analysis of Data*. Springer, 1996.

[27] D. J. Heeger and J. R. Bergen, "Pyramid-based texture analysis/synthesis," *ACM Transactions on Graphics, SIGGRAPH 1995*, pp. 229–238, 1995.

[28] E. P. Simoncelli and J. Portilla, "Texture characterization via joint statistics of wavelet coefficient magnitudes," *5th IEEE Int'l Conf on Image Processing*, pp. 229–238, 1998.

[29] J. Portilla and E. Simoncelli, "A parametric model based on joint statistics of complex wavelet coefficients," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 49–71, 2000.

[30] K. Popat and R. Picard, "Novel cluster-based probability model for texture synthesis, classification, and compression," *Proc. SPIE Visual Communications '93*, 1993.

[31] R. Paget and I. D. Longstaff, "Texture synthesis via a nonparametric markov random field," *Proceedings of DICTA-95, Digital Image Computing: Techniques and Applications*, vol. 1, pp. 547–552, 1995.

[32] ——, "A nonparametric multiscale markov random field model for synthesising natural textures," *Fourth International Symposium on Signal Processing and its Applications*, vol. 2, pp. 744–747, 1996.

[33] ——, "Texture synthesis via a noncausal nonparametric multiscale markov random field," *IEEE Transactions on Image Processing*, vol. 7, no. 6, pp. 925–931, 1998.

[34] S. Zhu, Y. Wu, and D.B.Mumford, "Minimax entropy principle and its applications to texture modeling," *Neural Computation*, vol. 9, pp. 1627–1660, 1997.

[35] ——, "Frame : Filters, random fields and maximum entropy — towards a unified theory for texture modeling," *Int. Journal of Computer Vision,*, vol. 27, no. 2, pp. 1–20, 1998.

[36] S. Zhu, X. Liu, and Y. Wu, "Exploring julesz ensembles by efficient markov chain monte carlo—towards a trichromacy theory of texture," *IEEE Trans. PAMI*, vol. 22, no. 6, pp. 554–569, 2000.

[37] ——, "Equivalence of julesz ensemble and frame models," *Int. Journal of Computer Vision*, vol. 38, no. 3, pp. 247–265, 2000.

[38] Y. Xu, B. Guo, and H. Shum, "Chaos mosaic: Fast and memory efficient texture synthesis," *Technical Report, Microsof Research*, vol. 22, no. 6, 2000. [Online]. Available: ftp://ftp.research.microsoft.com/pub/tr/tr-2000-32.pdf

[39] Y. Q. Xu, S. C. Zhu, B. N. Guo, , and H. Y. Shum, "Asymptotically admissible texture synthesis."

[40] M. Ashikhmin, "Synthesizing natural textures," *ACM Symposium on Interactive 3D Graphics*, pp. 217–226, 2001.

[41] S. Zelinka and M. Garland, "Towards real-time texture synthesis with the jump map," *Proceedings of the Thirteenth Eurographics Workshop on Rendering Techniques*, pp. 99–104, 2002.

[42] X. Tong, J. Zhang, L. Liu, X. Wang, B. Guo, and H.-Y. Shum, "Synthesis of bidirectional texture functions on arbitrary surfaces," *SIGGRAPH 2002*, pp. 665–672, 2002.

[43] A. Nealen and M. Alexa, "Hybrid texture synthesis," *Proceedings of the Eurographics Symposium on Rendering 2003*, pp. 97–105, 2003.

[44] ——, "Fast and high quality overlap repair for patch-based texture synthesis," *Proceedings of CGI 2004*, pp. 582–585, 2004.

[45] A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," *Proceedings of SIGGRAPH 2001*, pp. 341–346, 2001.

[46] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin, "Image analogies," *Proc. SIGGRAPH 2001*, pp. 327–340, 2001.

[47] U.-D. Reips, "Standards for internet-based experimenting," *Exp. Psych.*, vol. 49, no. 4, pp. 243–256, 2002.

[48] F. Wichmann and N. Hill, "The psychometric function. i. fitting, sampling, and goodness of fit," *Percept. Psychophys*, vol. 63.

[49] ——, "The psychometric function. ii. bootstrap-based confidence intervals and sampling," *Percept. Psychophys*, vol. 63.

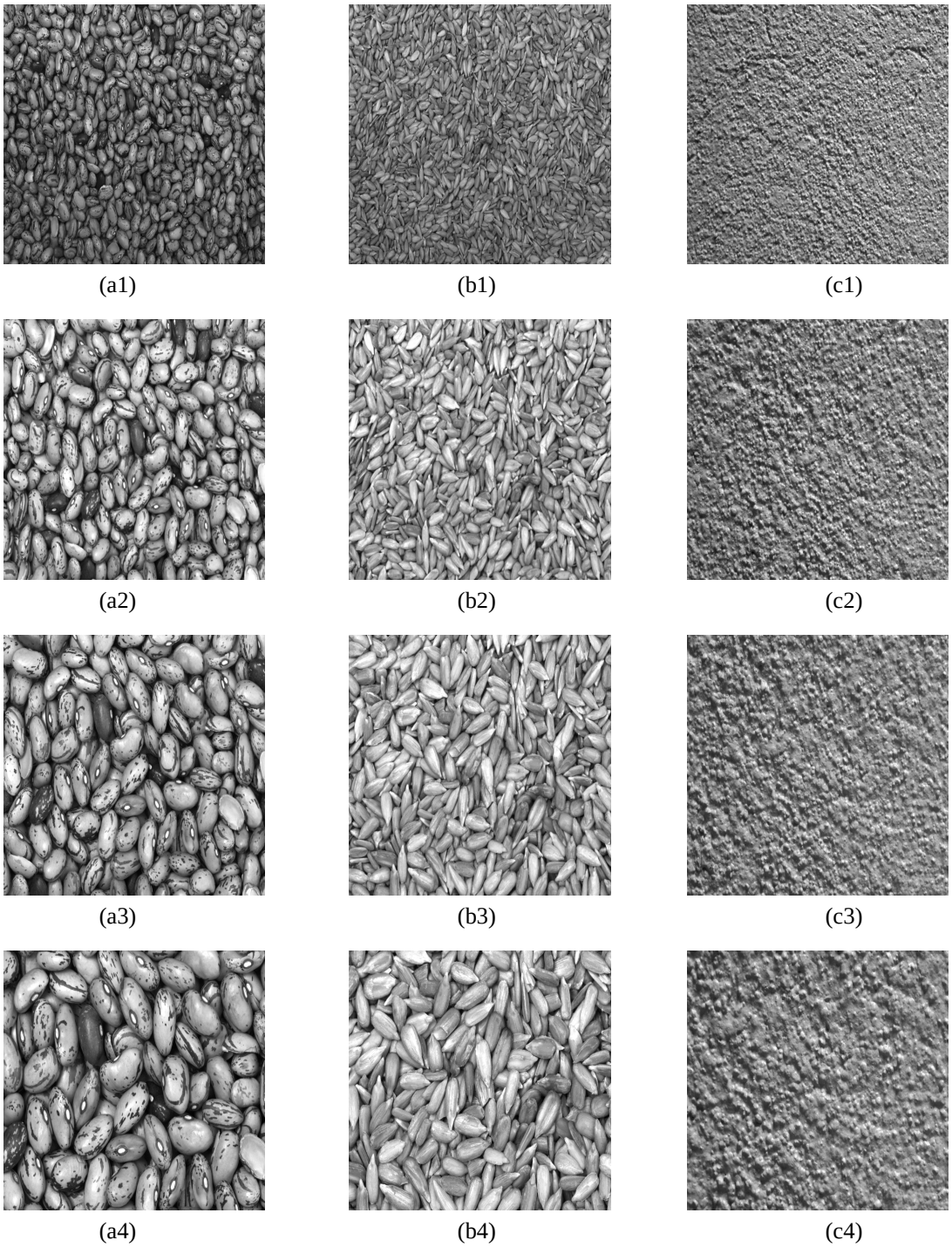

Fig. 10. The other 12 textures belonging to the MT collection; they are composed by four different zoomed textures of: (a1-a4) Beans; (b1-b4) Sunflower seeds; (c1-c4) A wall.

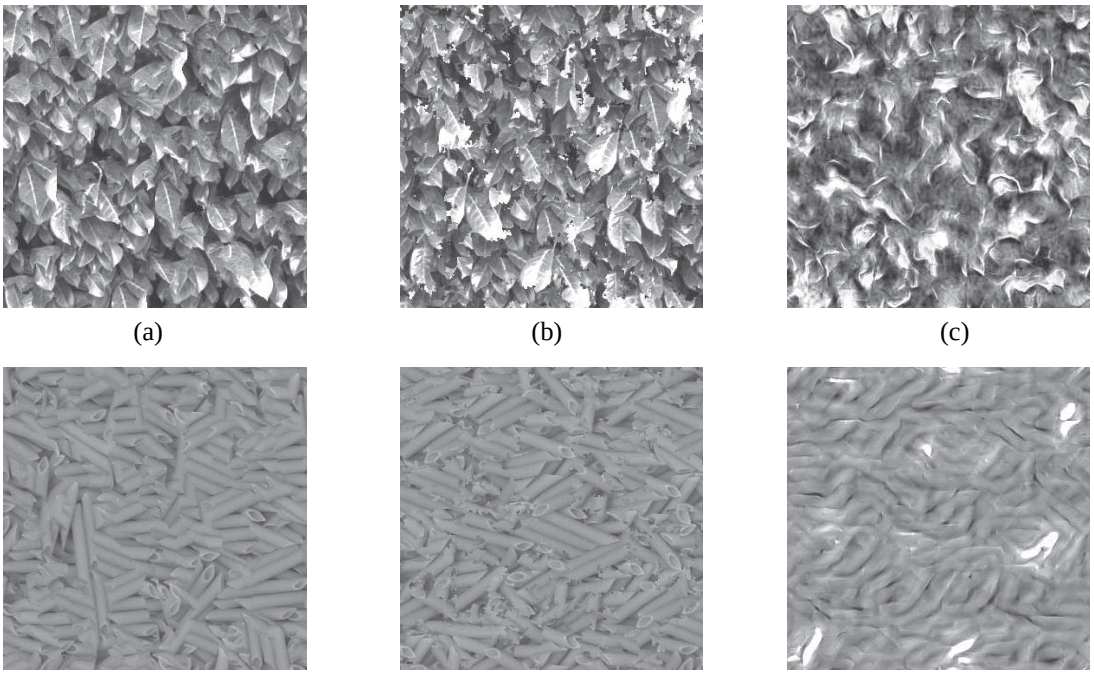

Fig. 11. Texture synthesis examples: (a-c) results obtained using the algorithm "ASH","JMP", and "POR", respectively for the synthesis of texture representing foliage; (d-f) results obtained using the algorithm "ASH","JMP", and "POR", respectively for the synthesis of texture representing pasta ensemble.

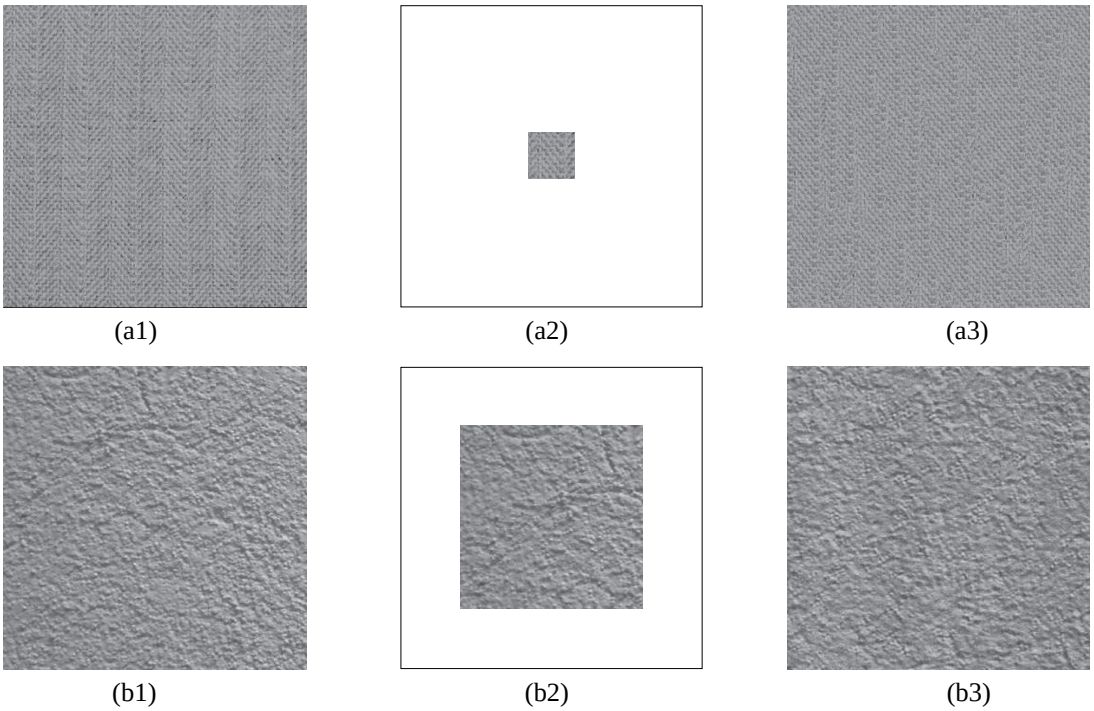

Fig. 12. Examples of synthesis results obtained by adapting the sample window to the texture according to the score; (a1-a3) Results for a texture of the Brodatz collection obtained using the "ASH" algorithm: (a1) original texture; (a2) A 50×50 sample, computed from $W^P_{min} = 100$ using a ratio $\gamma = 1/2$; (a3) Synthetic texture; (b1-b3) result for a texture of the MT collection using the "POR" algorithm: (a1) original texture; (a2) A 192×192 sample, computed from $W^P_{min} = 180$ using a ratio $\gamma = 0.94$; (a3) Synthetic texture.