

A MEASURE FOR SPATIAL DEPENDENCE IN NATURAL STOCHASTIC TEXTURES

Roberto Costantini^a, Gloria Menegaz^b, and Sabine Süsstrunk^a

^aAudiovisual Communications Laboratory EPFL, Lausanne, Switzerland

^bDepartment of Information Engineering, University of Siena, Italy

e-mail: roberto.costantini@epfl.ch

ABSTRACT

Random textures differ from natural textures because they lack structure. Structure is a concept that is difficult to formalize, but we generally observe that it is associated with spatial dependence between adjacent pixels. Completely random textures, in fact, are characterized by independent pixels, while natural textures show some local dependence. In this paper, we propose a measure for such local dependence. The texture is scanned on a random walk path and the pixels values encountered are collected in form of a time-series. The statistical properties of the time-series are used to characterize the spatial dependence of the texture. We assume that the proposed measure can be used to help model textures, to set the size of operating windows in texture synthesis algorithms, and to compute a simple indicator of texture scale. Moreover, we will show how the measure can be linked to texture perception.

1. INTRODUCTION

When we look at natural textures, we can notice the presence of a certain structure. But what is a “structure?” Generally speaking, the structure of a texture derives from the presence of a randomized and spatially-redundant replica of a pattern. The spatial redundancy acts such that the texture is perceived as a unity, even though it can be formed by isolated elements, while the pattern characteristics define the texture appearance. Recent studies [1] use the term *texton* to refer to this pattern. Textons are the fundamental micro-structures that constitute a texture [1]. Yet, there is no mathematical definition of what a texton is, even though some proposals have been formulated [1, 2, 3, 4].

In this paper, we are interested in finding the minimum number of replica of a single texton or group of textons that permits to form the perception of texture. This problem can be stated in an equivalent way. Given a texture and a smaller window that crops part of it, we aim at estimating the minimum window dimension that ensures the same perception of texture as we have from the whole image. We test if the dimension of this window is linked to the distance between pixels that ensures *spatial independence*.

We propose a method that estimates this spatial independence. By performing a random walk on the texture while maintaining a certain fixed distance between consecutive pixels, we collect the pixel values encountered in a time-series and test the independence of the samples. If samples are dependent, we increase the distance and apply

the method again until independent samples are found. We will show that the estimated distance can be an indicator of texture scale (which is related to the dimension of the texton pattern) and that it is linked to the perception of texture.

2. ESTIMATION OF SPATIAL DEPENDENCY FOR TEXTURES

Several methods exist that establish the dependence between samples of time-series. These methods are generally used in economic trend evaluations to verify the stationarity of stock markets, as for example in [5, 6], or to determine simulation run length, as in [7]. They accept a certain time-series (1D signal) and establish if samples can be considered as independent. Since we are interested in investigating the spatial (2D) dependence in texture, our idea is to extend such 1D methods to 2D. The principle of the test is depicted in Fig. 1. We perform a random walk (RW) on the texture,

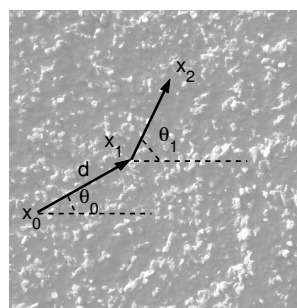


Fig. 1. Random walk scanning of the texture.

ensuring that the consecutive pixels encountered during the walk are at distance d . This is done by randomly choosing an initial position x_0 inside the texture and an initial angle θ_0 in the interval $[0, 2\pi]$. The following pixel x_1 will be selected in the direction indicated by θ_0 , at a distance d from x_0 . The other samples are obtained with the same mechanism, with the difference that angles θ_i at step i are chosen uniformly in the interval $[\theta_0 - \pi/6, \theta_0 + \pi/6]$. This choice of angles prevents the RW to go back and forth too frequently when advancing toward a certain direction. When a border of the image is reached at a certain index j , the next point is chosen by taking a new θ_j uniformly in $[0, 2\pi]$ and ensuring that the new position falls inside the image.

The result of this scan is a time-series formed by the values of the pixels encountered along the RW. We indicate it with X_d , where d is the distance (in pixels) used in the RW. In our experiment, we set the length of the time-series to 6000 samples. In Fig. 2(a), the first 1000 samples of the time-series obtained from the RW scan of the texture in Fig. 1 with a distance $d = 10$ are reported. Once

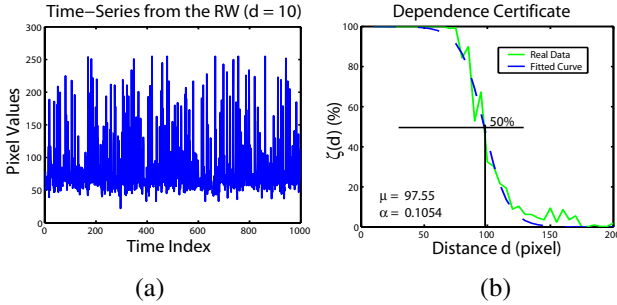


Fig. 2. Result of the run-test: (a) first 1000 samples of the time-series obtained by the RW scan of the texture in Fig. 1 for a distance $d = 10$; (b) global result of the run-up test for the texture in Fig. 1.

the time-series X_d is created, we use a simple yet robust statistical test called *run-up test* for checking sample independence [7]. The null hypothesis H_0 is that the samples are independent. The test output will be TRUE (or 1) if H_0 is rejected, i.e., if the samples are dependent, and FALSE (0) otherwise.

Statistical tests have a certain *power*. This is the probability that the test does not accept a false hypothesis, and it corresponds to the complementary event of a Type II error [8]. For this reason, we perform the test a certain number of times, and we define a *dependence certificate* $\zeta(d)$ as:

$$\zeta(d) = \frac{\text{n. of times the test is positive}}{\text{total number of tests performed}} \quad (1)$$

This certificate represents an indication of the probability that at distance d pixels are dependent, and provides a method to describe the spatial dependence behavior of a texture. In Fig. 2(b), the certificate $\zeta(d)$ obtained for the texture in Fig. 1 is drawn with a continuous line, expressed in percentages from 0 to 100%. When distance d is small (10 to 30 pixels), the pixels are very likely dependent, while when the pixel distance begins to increase, they tend to become more independent. The dashed line represents the fitting of $\zeta(d)$, according to the function:

$$f(d) = \frac{100}{e^{\alpha(d-\mu)} + 1} \quad (2)$$

where μ and α are some parameters. We use this particular function since it describes quite well the typical behavior of $\zeta(d)$ and has the property that at $d = \mu$, its value is 50%. We estimate the parameters μ and α by fitting the real data to $\zeta(d)$ using MSE minimization.

Briefly, our method has two main outputs: the dependence certificate, which is a function of pixel distance and

describes globally the spatial dependence of the texture, and the SDMD (Spatial Dependency Mean Distance), which is a scalar that represents the distance between pixels at which they are both equally likely to be dependent or independent ($\zeta(\mu) = 50\%$, $d = \mu$).

2.1. Run-up Test

The run-up test is a test that verifies the dependence between samples of a given time-series. It is based on the statistics of the number of *run-up* that appear in the sequence. By definition [7], a run-up of length L is a sequence of increasing values of length L : $x_i < x_{i+1} < \dots < x_{i+L-1} (> x_L)$. In a time-series formed by iid samples, the probability that a run-up has length k is:

$$\Pr(L = k) = \frac{k}{(k+1)!} \quad (3)$$

In order to verify that the samples of the time-series are independent, the run-up test counts the number of run-up of length from 1 to 6 and then compares their relative frequency with the theoretical one given by Eq. (3). The comparison is done using the Pearson χ^2 -test, allowing a 5% Type I error percentage. We refer to [7] for details.

3. TESTS AND RESULTS

In order to analyze the proposed measure of spatial dependence proposed, we have tested it on textures whose spatial dependence was under control. In Fig. 3, we show two samples of such textures, which are obtained from a random field defined from its 2D Fourier transform:

$$X(f_1, f_2) = e^{\frac{f_1^2 + f_2^2}{2r^2}} e^{-j\Phi_1} e^{-j\Phi_2} \quad (4)$$

where f_1 and f_2 are the spatial frequencies (from -0.5 to 0.5), Φ_1 and Φ_2 are random phases chosen in the interval $[0, 2\pi]$, and r is a parameter. The textures are obtained by considering the absolute value of the inverse transform:

$$X(s_1, s_2) = |F^{-1}\{X(f_1, f_2)\}| \quad (5)$$

where s_1 and s_2 are the spatial coordinates of the pixels.

The parameter r permits to change the range of spatial dependence between pixels, as illustrated in Fig. 3. This can be regarded as a change of scale. Twenty synthetic textures

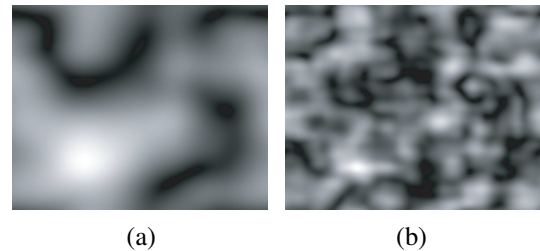


Fig. 3. Synthesized textures: (a) $r = 0.4$; (b) $r = 0.7$.

for each value of r were created, with r ranging from 0.4

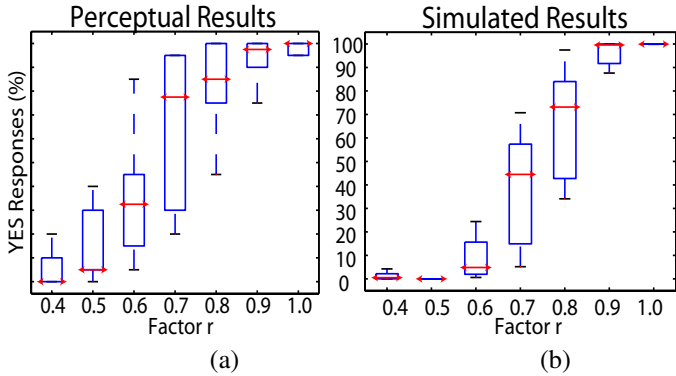


Fig. 4. Results from the psychovisual test on synthetic textures perception (Experiment 1). (a) Data obtained with subjects; (b) Simulated data.

to 1.0 with a step of 0.1, thus obtaining a total set of 140 synthetic textures. For each of them, the SDMD (μ) and the parameter α of Eq. (2) have been estimated. For each scale r , their average values among the twenty textures are reported in Tab. 1. As it can be seen, the average estimate of SDMD follows the change of scale in a coherent way, since smaller scales (low r -values) have a bigger μ values than larger scales.

| r | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |
|----------------|------|------|------|------|------|------|------|
| SDMD (μ) | 52.4 | 33.0 | 22.9 | 16.8 | 13.5 | 9.7 | 8.9 |
| α | 0.17 | 0.27 | 0.38 | 0.48 | 0.98 | 0.69 | 0.16 |

Table 1. Measured parameters for the 7 synthesized textures.

3.1. Synthetic Textures: Psychovisual Test

When does a certain spatial configuration becomes a texture from a perceptual point of view? This is a problem we would like to study and interpret in the framework illustrated above. In order to do this, we conducted a psychovisual test on texture perception (Experiment 1). We used 10 subjects (6 naive and 4 experienced subjects with respect to image processing) and provided them with a general definition of texture. We defined a texture as an image having at least two out of the three following characteristics: randomness, quasi-periodicity, and visual redundancy.

We used the 140 synthesized textures described above as test images. Each texture was shown on a computer screen within a window of dimension 126×126 pixels viewed at a normal viewing distance of 50 cm from the monitor. Subjects had to answer if the image could be considered a texture, according to the stated definition. The results are reported in Fig. 4(a).

The arrows inside the boxes indicates the mean value of YES response percentage. Each box has lines at the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of a box to show the extent of the rest of the data. We notice that for r -values at the extremes of the range, the responses are quite confident

(low variance), while at the center (at $r = 0.7$ for instance), subjects are quite uncertain (high variance). This behavior could result from the variability of images and from subject uncertainty within certain scales, where the image is perceived as “almost” a texture.

To see how our measure of spatial dependency can be linked to texture perception, we compared the SDMD and α parameters measured by our method with those that can be estimated from the psychovisual experiment. The parameter estimation was done by inverting Eq. (2): we computed for each scale r the parameters μ and α of Eq. (2) such that, at a distance $d = 126$ pixels (which is the size of the window used in our experiment), would permit to have $f(d)$ equal to the percentage of YES responses normalized to one.

We found that the ratio between the parameters estimated from the psychovisual data and those measured by our method was approximately linear for μ and constant for α , with respect to increasing r . Using the measured values and this relationship, we simulated the responses of observers starting from the measured values. The results are shown in Fig. 4(b). We obtained a similar behavior to the one that characterizes subject responses, i.e. the increase of uncertainty in the middle range of the scale.

3.2. Natural Textures

The experiments conducted on synthetic textures revealed that texture perception and scale are linked, since texture, as defined in the experiment, was perceived at certain scales, but not at others. Here we test if this could be applied to natural textures as well.

We consider two experiments: one where no definition of texture is given and one where the same definition used for synthetic textures is given. For the first experiment (Experiment 2) we used 13 observers (8 naive and 5 expert observers) and a collection of 11 textures, grouped in three categories: textures belonging to a single recognizable (wood, tree bark, cobble surface) or non recognizable object (parts of different walls), and texture formed by an ensemble of the same type of objects (such as beans, sunflower seeds, and pasta).

Images, original size 800×1200 pixel, were displayed at 10 different sized crops varying from 100 to 600 pixels for each image on a 100×100 pixel screen window. This is equivalent to 60 artificially-zoomed different images for each single texture, for a total number of 660 test images. The sequence of crop positions and images used during the test was randomized and made the same for all subjects. The subjects’ task was to state if they perceived the crops to be texture or not, as in the test for synthetic textures. Results are given in Fig. 5(a).

A common behavior and some strong differences can be seen at the same time. The similar behavior consists in the increase of YES response percentages when the scale becomes larger; this happens in all the three texture categories. The differences are visible in the absolute value of the responses: non recognizable images (continuous line) are in fact almost always perceived as textures (average over all scales: 85%), while images made of repetitions of the same object (dash-point line) are seldom considered as texture (average over all scales: 30%). This is because subjects

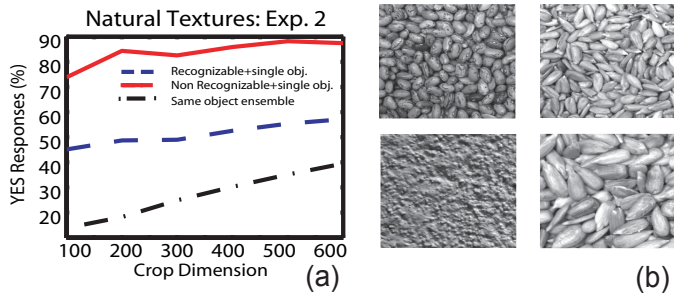


Fig. 5. (a) Results from the psychovisual experiment on natural textures where no definition of texture was given (Experiment 2); (b) Textures used in Experiment 3.

| Texture | SDMD (μ) | P-SDMD | α | P- α |
|---------|----------------|--------|----------|-------------|
| (a) | 23 | 184 | 0.15 | 0.015 |
| (a) | 49 | 337 | 0.01 | 0.014 |
| (a) | 64 | 388 | 0.07 | 0.011 |
| (a) | 84 | 513 | 0.05 | 0.032 |
| (b) | 12 | 152 | 0.71 | 0.025 |
| (b) | 19 | 250 | 0.93 | 0.016 |
| (b) | 23 | 307 | 0.87 | 0.016 |
| (b) | 29 | 380 | 0.68 | 0.016 |
| (c) | 9 | 134 | 1.50 | 0.033 |
| (c) | 13 | 159 | 1.34 | 0.015 |
| (c) | 17 | 265 | 0.61 | 0.020 |
| (c) | 20 | 317 | 0.68 | 0.012 |

Table 2. Measured (SDMD (μ) and α) and perceptual (P-SDMD and P- α) parameters for the 3 natural textures used in Experiment 3, for each of the 4 scales.

recognized the repeating object (bean, or pasta, for example) and did not consider the ensemble as a texture according to their personal interpretation. Recognizable textures that belong to a single object lay exactly in the middle of these two extreme cases (dashed line). They are recognized as texture with approximately 50% of probability. This means that they are as likely considered texture as not.

From these data, it is not possible to make some quantitative analysis, since subject performance is biased by high level phenomena (identification, classification, etc.). Thus a second experiment (Experiment 3) has been conducted where a definition of texture was given. We used the three different textures, shown in Fig. 5(b) (first row and bottom left corner), each at 4 different scales obtained by an optical zoom of the digital camera (one zoomed image is shown in Fig. 5, bottom right corner). We applied the same methodology used in the previous test.

In this case, the collected data show that subject performance follows the typical shape of the Dependence Certificate of Fig. 2(b) and valid parameter estimates can be found. The results are reported in Tab. 2. We indicate with P-SDMD and P- α the values estimated from the psychovisual data curve obtained from observers (P stands for “perceptual”), and with SDMD and α the values obtained by our measure. We notice that in this case, when the scale

changes, the P-SDMD values change accordingly to scale and to the measured parameters. Moreover, the ratios between perceptual and measured parameters are almost constant, having a different value for each particular texture.

This is different from what has been found for synthetic textures, where the ratio between measured and estimated SDMD changed linearly with r . For natural textures, in fact, the different scales have been obtained by changing the zoom factor: this is not equivalent to changing the parameter r in synthetic textures. Similar behavior of ratios could still be found in synthetic texture, when using as r values the square root of those used in the Experiment 1 (see exponent of r in Eq. (4)).

4. CONCLUSIONS

In this paper, we propose a method that estimates a characteristic spatial dependence distance in natural stochastic textures. The distance produces coherent results for synthetic textures where the spatial dependence is adjustable. In addition, we showed that it can be used as a measure to estimate texture scale. In psychovisual tests performed both on synthetic and natural texture, we found that the proposed distance can produce coherent predictions of the perception of texture when a clear definition of texture was given. A lack of instruction resulted in arbitrary interpretation, and thus in unpredictable performance.

5. REFERENCES

- [1] Zhu S.-C., Guo C., Wu Y., and Wang Y., “What are textures,” *Proc. of 7th European Conf. on Computer Vision*, May-June 2002.
- [2] Olshausen B.A. and Field D.J., “Sparse coding with an overcomplete basis set: A strategy employed by v1?,” *Vis. Res.*, vol. 37, no. 23, pp. 3311–3325, 1997.
- [3] T. Leung and J. Malik, “Recognizing surface using three-dimensional textures,” *Proc. of 7th ICCV*, 1999.
- [4] Frey B. and Jovic N., “Transformed component analysis: joint estimation of spatial transforms and image components,” *Proc. of Int’l Conf. on Comp. Vis., Corfu, Greece*, 1999.
- [5] M. Klimek, G. Strandell, and J. Tysk, “Testing weak stationarity of stock returns,” *Theory of Stochastic Processes*, vol. 7, no. 23, pp. 196–202, 2001.
- [6] D. Yu, W. Lu, and R.G. Harrison, “Space time-index plots for probing dynamical nonstationarity,” *Phys. Lett. A*, vol. 250, no. 323, 1998.
- [7] E.J. Chen and W.D. Kelton, “Determining simulation run length with the runs test,” *Modelling Practice and Theory*, vol. 11, pp. 237–250, 2003.
- [8] T.W. Anderson and J.D. Finn, *The New Statistical Analysis of Data*, Springer, 1996.