Dead leaves model for measuring texture quality on a digital camera

Frédéric Cao[1], Frédéric Guichard, Hervé Hornung
DxO Labs, 3 rue Nationale, 92100 Boulogne-Billancourt FRANCE

ABSTRACT

We describe the procedure to evaluate the image quality of a camera in terms of texture preservation. We use a stochastic model coming from stochastic geometry, and known as the dead leaves model. It intrinsically reproduces occlusions phenomena, producing edges at any scale and any orientation with a possibly low level of contrast. An advantage of this synthetic model is that it provides a ground truth in terms of image statistics. In particular, its power spectrum is a power law, as many natural textures. Therefore, we can define a texture MTF as the ratio of the Fourier transform of the camera picture by the Fourier transform of the original target and we fully describe the procedure to compute it. We will compare the results with the traditional MTF (computed on a slanted edge as defined in the ISO 12233 standard) and will show that the texture MTF is indeed more appropriate for describing fine detail rendering. This is true in particular for camera phones that have to apply high level of denoising and sharpening.

Keywords: Texture, sharpness, image quality evaluation, MTF, dead leaves model.

1. INTRODUCTION

Sharpness is one of the main image quality attributes[1,2]. However, qualifying sharpness is not an easy task, since the concept actually covers several close but distinct notions. The sharpness of an image qualifies how acute the boundaries of objects are. It is often mistakenly confused with resolution which is the size of the smallest details that a camera can discriminate. In this paper, we focus on the acuteness of images. The ISO standard 12233 describes a protocol to measure the resolution and the spatial frequency response (SFR) of a camera[4]. However, it does not specify how to compute the sharpness of an image. The I3A Camera Phone Image Quality group is trying to define objective notions of sharpness which are consistent which subjective image quality. The work described in this article has been partly achieved in this workgroup. The first observation is that digital imaging made the traditional concept of sharpness obsolete. Indeed, in traditional photography, sharpness or lack of sharpness is related to the quality of the optics. In digital photography, optical blur is usually corrected by digital processing as sharpen filters. It is also well known that these types of filters increase image noise. Therefore, sharpen filters are often applied selectively on images, particularly on parts which are likely to be edges, which are in general parts with a high enough local contrast. Parts of the image, with more subtle variations may receive different processing. As a consequence, it seems necessary to introduce two concepts of sharpness. The first one is related to edges with "high enough" contrast. The second one is related to the quality of textures and is the object of this article. In a previous article, we discuss the requirements for an objective measurement of texture quality in digital images. We push the analysis further and completely describe the protocol and the algorithm to perform the measurement. The outline of the paper is as follows. In Sect. 2, we remind the main properties of the model that will be used for building a measurement target. In Sect. 3 we describe the target, created to make the measurements and how to shoot it. In Sect. 4, we describe the algorithm to compute the objective measurement, the texture MTF. In Sect. 5, we relate texture MTF and SFR, and study the behavior of the texture MTF in presence of noise.

---

[1] fcao@dxo.com
Frédéric Cao, Frédéric Guichard and Hervé Hornung, "Dead leaves model for measuring texture quality on a digital camera", Proc. SPIE 7537, 75370E (2010); doi:10.1117/12.838902

2. THE DEAD LEAVES MODEL

The dead leaves model was proposed in Mathematical Morphology by G. Matheron[5], as a model of stochastic images that can reproduce occlusions. The generation of the dead leaves model is very intuitive. Consider a set of planar shapes (the dead leaves) that are thrown on the image plane in a temporal series. The most important property of the model is that one shape can be (and probably will be) occluded by the shapes that fall after it. The temporal series is supposed to last long enough so that each point of the whole image plane has been covered by at least one leaf. Leaves can be different in shape, size, orientation or color.

Several studies showed that some elementary statistics, as gradient distribution, wavelet coefficients, co-occurrence values or power spectrum were very close to the ones of natural images[6,7,8,9]. In particular, the power spectrum of dead leaves images turns out to be very close to a power function. The power depends on the type of leaves that are used to generate the image.

In this paper, we use a dead leaves model with the following characteristics:

- Leaves are disks
- The radius of disks is a random variable following a power law with a power equal to -3. In practice, a maximal and an minimal leaf size are two parameters of the model.
- The position of the center follows a uniform Poisson process in the image domain.
- The image is in gray levels (no color); gray levels follow a random variable, uniform in a given interval (which is a parameter o the model)

This particular model has several nice properties:

- It is statistically rotation invariant (because leaves are disks)
- It is statistically shift invariant (because positions follow a Poisson process)
- It contains sharp edges (by occlusion)
- It contains different levels of contrast (gray levels are random)
- It is scale invariant. This means that the statistics of the target does not depend on the viewing distance. In other terms, the target is fractal.
- The power spectrum of the image is a power function. The exponent is approximately -1.93.

The fourth point was proved in [5]. It is a far from trivial consequence of the particular distribution of the size of leaves (power -3). Because of radial symmetry, the Fourier transform of the dead leaves target also has a radial symmetry (up to slight differences due to the fact that pixels form horizontal and vertical lines). The profile of the modulus of the FFT is displayed in log-log scale on Fig.1. Except in very low and very high frequencies, it is very close to a straight line (in log scale). Therefore, it is a power function. In practice, the target is printed such that the highest frequencies of a camera will not reach the highest frequencies of the target, so that we remain in the linear part of the log-power spectrum. For very low frequencies, the power law might slightly overestimate the power spectrum of the dead leaves target.

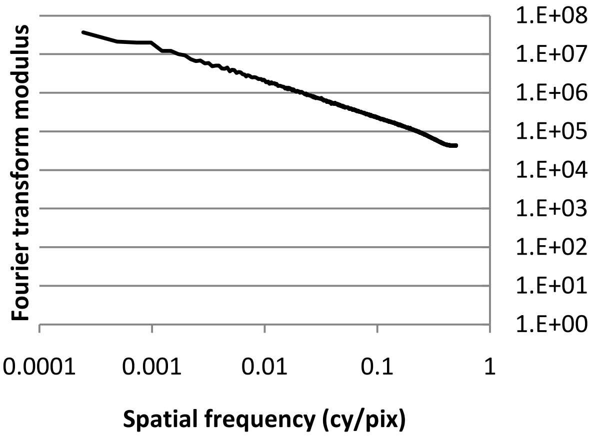


Figure 1. Fourier transform modulus of a dead leaves target in log-log coordinates. Except in very high and very low frequencies, it is very close to a straight line.

3. THE DEAD LEAVES TARGET

The target we propose is shown on Fig. 2.

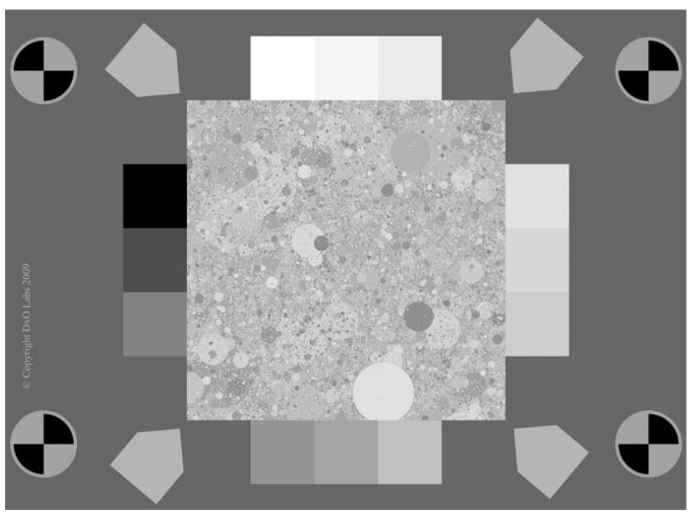


Figure 2. The dead leaves target used for our experiment. See text for a full description.

This target can either be a reflection or transmission target. In the case of a reflection target, the background is a 18% uniform gray. Markers can be used to automatically detect the target. The square around the main central part have precisely chosen reflectance in order to measure the camera OECF. The patterns in the corner have horizontal and vertical edges ± 5° and 45°±5° for SFR measurement. The central part is a disk dead leaves pattern as described in Sect.2. Contrast is chosen low enough. In practice, the contrast of the target in Fig. 2 is 3:1.

The target shall be uniformly illuminated. If possible, the size of the whole target in the field shall not exceed 30% of the image field so as to limit the influence of light and color shading, whatever it is due to non uniform illumination or lens and sensor characteristics. The dimensions of the texture part shall be larger than 250 pixels, and preferably larger than 500 pixels. For low resolution cameras, it is preferred that the texture part is larger, even though the accuracy of the measurement may suffer from field non-uniformity.

Since digital processing is usually dependent on the noise level in the input raw image, measurements can be made at different levels of illumination. The measurement can be performed on the three channels independently. The result also depends on the illuminant. If possible, white balance should be adjusted to the ambient illuminant.

The image shall then be linearized by inverting the OECF on each color channel. This step will also remove a color hue due to an imperfect white balance.The OECF patches should contain enough pixels so that the average value can be accurately estimated despite noise, in particular in low light conditions.

The dead leaves MTF can then be computed on each channel independently.

4. TEXTURE MTF COMPUTATION

4.1 Definition

Let us first state the definition.

Let I be a $N \times N$ gray level (single channel) image of the dead leave target. Denote by $\hat{I}(m,n)$ be the discrete Fourier transform on I, $-\frac{N}{2} < m, n \leq \frac{N}{2}$. The 2D texture MTF is

$$K(m,n) = \frac{\left|\hat{I}(m,n)\right|}{\left|\hat{T}(m,n)\right|},$$

where

$$\hat{T}(m,n) = \frac{c(N)}{(m^2+n^2)^{\eta/2}}, \text{ with } c(N) = \frac{\mathrm{var}(I)}{\sum_{-\frac{N}{2}+1,\ldots,\frac{N}{2}} \frac{1}{(m^2+n^2)^{\frac{\eta}{2}}}} N^4.$$

Since two-dimensional MTF is not easy to handle, we usually assume that $K(m,n)$ has a radial symmetry. The one-dimensional texture MTF is the average of $K(m,n)$ over all directions.

4.2 Interpretation

Let us now explain the definition. If the image is obtained from the original target by a convolution with a blur kernel, then the MTF of this kernel can be computed as the ratio of the Fourier transform of the shot and the Fourier transform of the original target. That is to say, if we denote by $\hat{I}$ the Fourier transform of the image, by $\hat{k}$ the Fourier transform of the convolution kernel and $\hat{T}$ the Fourier transform of the original target, we then have

$$\left|\hat{k}\right| = \frac{\left|\hat{I}\right|}{\left|\hat{T}\right|}.$$

Even though, digital processing is non linear, we can still make the analogy and define the texture MTF as this ratio. In this equation, we obviously need to know the Fourier transform of the original target. As described in Sect. 2, the Fourier transform of a dead leaves pattern is very close to a power function. The approximation fails for very low and very high frequencies. For practical purposes, it does not really matters. Indeed, we can harmlessly assume that the texture MTF is continuous for low frequencies. For high frequencies, the target can be printed with a high enough resolution so that an image shot by a usual camera will never attain such high frequencies. Therefore, we can assume that

$$\hat{T}(m,n) = \frac{c(N)}{(m^2+n^2)^{\eta/2}}.$$

The constant $c(N)$ only depends on the crop size.

It can be proved that

$$c(N) = \frac{\mathrm{var}(I)}{\sum_{-\frac{N}{2}+1,\dots,\frac{N}{2}} \frac{1}{(m^2+n^2)^{\frac{\eta}{2}}}} N^4.$$

The proof is given in annex.

5. EXPERIMENTS

5.1 Acutance

A requirement of image quality evaluation is to be able to compare images or cameras. Therefore, measurements should be limited to a single number. We first defined texture MTF in two dimensions, and then considered the problem in one dimension by assuming that the MTF is radial symmetric. This result is comparable to the SFR computed on a slanted edge. In order to summarize the MTF in a single number, a contrast sensitivity function (CSF) can be used to weigh the different spatial frequencies, leading to a single acutance value. The CSF models human eye sensitivity to spatial frequencies, expressed in cycles per degree.

The CSF we used is explicitly defined by

$$CSF(\nu) = a \cdot \nu^c \cdot e^{-b\nu},$$

Where $\mathrm{b} = 0.2$, $\mathrm{c} = 0.8$, and a is chosen such that $\int CSF(\nu)d\nu = 1$.

On a digital image, frequencies are expressed in cycles/pixels. Therefore, acutance is relative to viewing conditions. In what follows, the viewing conditions are as follows: 29.7" screen, resolution 2560x1600, and viewing distance in 34".

The acutance is computed by

$$A = \int MTF(\nu) \cdot CSF(\nu)\, d\nu.$$

5.2 Comparison SFR vs. texture MTF

Raw images

The texture MTF aims to measure the spatial response for a low contrasted pattern. The difference with SFR on a high contrasted edge is that digital processing may behave very differently on these two structures. However, on a RAW image, no digital processing has been applied. Therefore, texture MTF and SFR should only be characteristic of lens performance and be very close to each other. However, on a RAW image, each plane is sampled at half Nyquist frequency. Standard ISO 12233 still permits to estimate the SFR at Nyquist rate. For texture MTF, we assume that all the channels have the same blur. By applying white balance scales directly estimated on the image, we reconstruct a single plane gray level image at the sensor resolution. The texture MTF is computed on this image. The result is displayed on Fig. 3. The MTF curves (SFR and texture) are close. For low frequencies, the texture MTF is a bit under estimated, as could be expected, since the power law is not exact for lower frequencies. For very high frequencies, the texture MTF is a bit higher for two reasons: the three channels do not have exactly the same amount of blur (it might be due to a slight longitudinal chromatic aberration) and the white balance may not be completely accurate, yielding some slight oscillations up to Nyquist frequency.

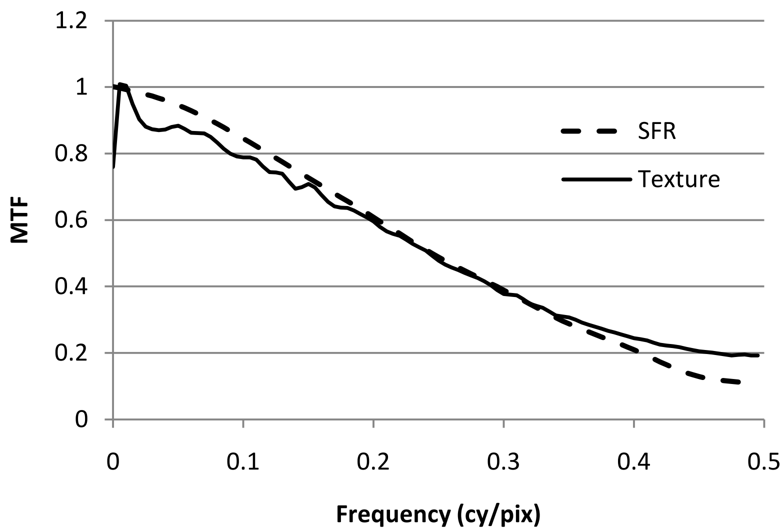
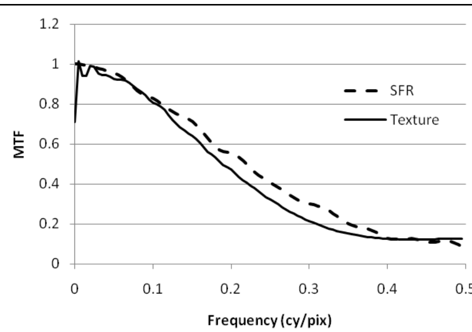
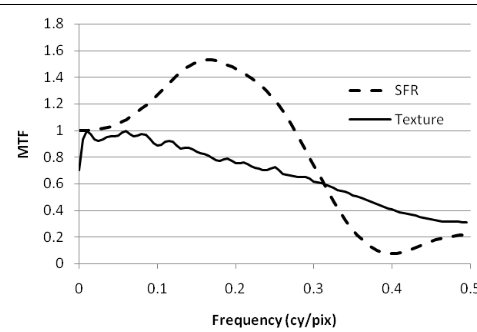


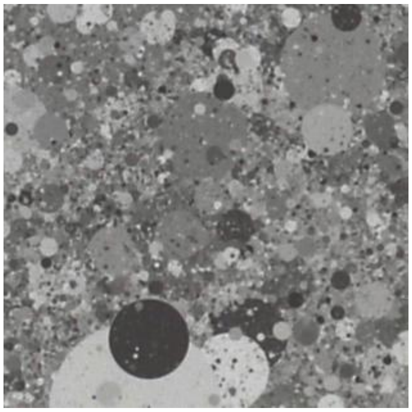
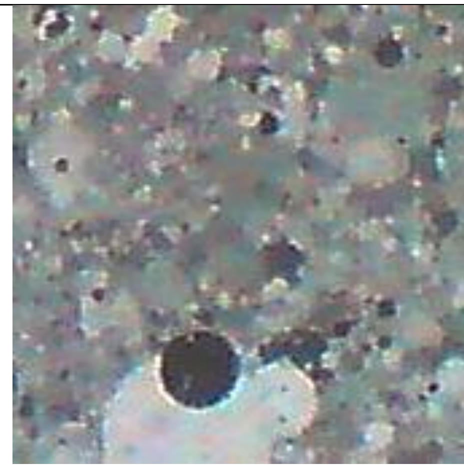
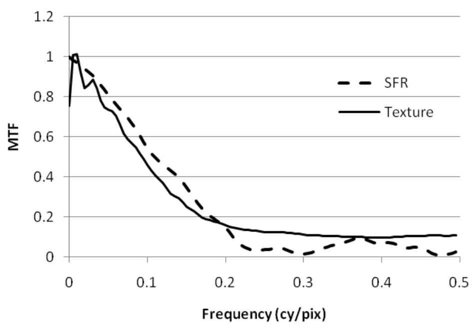
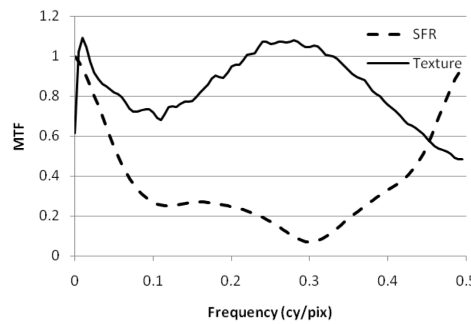
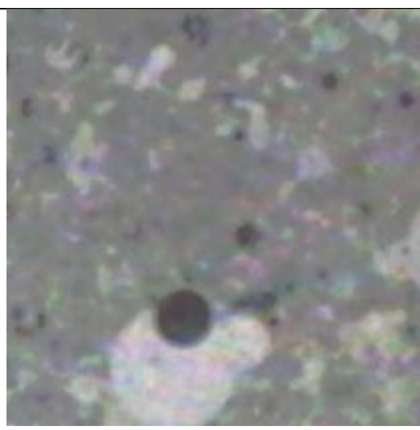
Figure 3. Comparison of SFR and texture MTF on a RAW image (Nikon D3x image)

JPEG images

The purpose of the measurement on the RAW image is only a consistency check between the edge MTF (SFR) and the texture MTF. The actual purpose of the measurement is to evaluate the quality of the final image, so the analysis of JPEG images is more relevant.

Table 1. The four different configurations of texture vs. edge sharpness.

Nikon D3x iso 100	Sony Ericsson K850i iso 100
	

	
LG KU990i iso 400	Fuji Z33WP iso 1600
	
	

There are four combinations of MTF:

- High texture MTF, high SFR: these are usually good cameras, good focus. Quality is good.
- Low texture MTF, high SFR: images have usually been digitally sharpened on edges, denoised in smooth areas. Camera has not been able to detect the dead leaves as relevant signal. Quality is questionable, in particular if the scene contains texture.
- Low texture MTF, low SFR: images are blurry, for instance because of a bad focus. Quality is bad.

- High texture MTF, low SFR: this case may happen if the image is blurry but also very noisy. This is the less likely configuration.

The four cases are illustrated in Table 1 above.

5.3 Influence of noise

The sensitivity of texture MTF using the dead leaves model obviously depends on the contrast of the target compared to the amplitude of noise. It is very difficult to avoid considering noise as texture. Noise is actually a particular kind of texture. For some natural textures, noise also sometimes help by adding a crispy look to the image. However, we want the measurement to get rid of this influence as much as possible.

Therefore, we use SFR as a reference MTF. Whenever texture MTF is larger than SFR, we consider that texture is actually due to noise. On the opposite, SFR is usually larger than texture MTF since edges undergo digital sharpening.

Therefore texture quality is characterized by min(SFR, texture MTF). Noise contribution can be defined as max(0,texture MTF-SFR). The rational of this definition is that the dead leaves target is composed of many edges in every direction and any level of contrast. However, a very good camera should be able to discriminate every single edge and reproduce it as it reproduces an isolated edge as a SFR slanted edge.

6. CONCLUSION

In this paper, we precisely describe the procedure to compute texture MTF from a dead leaves target as well as the protocol to shoot the target. We showed that texture MTF and SFR coincide on RAW images, discrepancy being due to image processing. We discuss the influence of noise on the dead leaves MTF and therefore propose to use the minimum value of SFR and texture MTF to evaluate the texture content.

7. ANNEX: NORMALIZATION OF TEXTURE MTF

The normalization constant of a NxN dead leaves image is

$$c(N) = \frac{\mathrm{var}(I)}{\sum_{-\frac{N}{2}+1,\ldots,\frac{N}{2}} \frac{1}{(m^2+n^2)^{\frac{\eta}{2}}}} N^4.$$

Proof: The Fourier transform of an $N \times N$ image u_{kl} is defined by

$$\hat{I}_{mn} = \sum_{0 \le i,j \le N-1} I_{kl} e^{-\frac{2i\pi(km+ln)}{N}}$$

Parseval equality is

$$\sum_{-\frac{N}{2}+1}^{\frac{N}{2}} \left|\hat{I}_{mn}\right|^2 = N^{\mathbf{2}} \sum_{0,\ldots,N} I_{kl}^2.$$

Applying this to

$$\left|\hat{I}_{mn}\right|^2 = \frac{c(N)}{(m^2+n^2)^{\eta/2}},$$

we get

$$\sum_{mn \neq 0} \frac{c(N)}{(m^2+n^2)^{\frac{\eta}{2}}} + |\hat{I}_{00}|^2 = N^4 E(I^2).$$

Since $\hat{I}_{00} = N^2 E(I)$, we get the results.

REFERENCES

[1] Wang, Z., Bovik, A.C., “A Universal Image Qualiy Index,” IEEE Signal Processing Letters, 9(3) 2002.
[2] Keelan, B., “Handbook of Image Quality: Characterization and Prediction,” Optical Science and Engineering; 2002.
[3] ISO 12333, “Photography – Electronic still-picture cameras – Resolution measurements”.
[4] Industrial International Imaging Association, Camera Phone Image Quality. Ongoing work.
[5] Matheron, G., “Random Sets and Integral Geometry,” John Wiley and Sons, New York 1975.
[6] Lee, A.B. , Mumford D., Huang J., “Occlusion Models for Natural Images: A Statistical Study of a Scale-Invariant Dead Leaves Model,” IJCV 41 (1/2), 35-59, 2001.
[7] Ruderman, D.L.. “The statistics of natural images,” Network, 5(4): 517-548, 1994.
[8] Alvarez, L., Gousseau, Y., Morel, J.M., “The size of objects in natural and artificial images,” Advances in Imaging and Electron Physics, vol 111. Academic Press: San Diego, CA: 167-242, 1999.
[9] Gousseau, Y., Roueff, F., “Modeling Occlusion and Scaling in Natural Images,” SIAM J. of Multiscale Modeling and Simulation, 6(1), 2007: 105-134, 2007.
[10] Artmann, U., “Noise-Reduction versus Spatial Resolution in Digital Images,” Diplomarbeit, University of Applied Sciences Cologne.
[11] Lee, J.S., “Digital Image Smoothing and the Sigma Filter”, CVGIP 24: 255-269, 1983.
[12] Phillips, J., Jin, E., Chen, Y., Clark, J., “Correlating Objective and Subjective Evaluation of Texture Appearance with Applications to Camera Phone Imaging,” Electronic Imaging 2009.
[13] Buzzi, J. Guichard, F. “Uniqueness of Blur Measure”, ICIP 2004.
[14] McElvain J.S., Campbell S.P, Jin E. Miller J., "Texture-based measurement of spatial frequency response using the dead leaves target: extensions, and application to real camera systems", Electronic Imaging 2010.
[15] Philipps J.B., Christoffel D.W, "Validating a texture metric for camera phone images using a texture-based softcopy quality ruler", Electronic Imaging 2010.