AUTOMATIC DETECTION OF DIGITAL ZOOMS

Jérôme BUZZI and Frédéric GUICHARD*

DOLabs
3, rue Nationale - 92100 Boulogne - FRANCE
jbuzzi@dolabs.com, fguichard@dolabs.com

ABSTRACT

We propose and evaluate a method to determine whether a given digital image is the result of a digital zoom in or more generally of a linear interpolation. This information is important to some post-processing. It also sheds some light on the actual resolution of some digital cameras.

1. INTRODUCTION

The analysis and optimization of digital zooms has received some attention (see, e.g., [4, 3, 9, 15]). However it seems that its detection has only be considered within movies (see, e.g., [6, 11]) and not for still images as in this paper. Yet such a detection provides a useful input to some image processing or can be used to compress or index images in a database. It can also give some information on the inner workings of digital cameras as we shall see.

In this paper, we propose and evaluate a method to determine whether a given digital image is the result of a digital zoom in or more generally of a linear interpolation or a JPEG compression [10].

This paper is organized as follows. We recall the basics of linear zooms, then we draw theoretically the consequences in Fourier space of zooms and JPEG compressions and present a practical method. Then we apply this to digital zooms, JPEG compression, digital cameras using natural images. Finally we conclude on the possibilities and limitations of this method

Basics of zooms and linear interpolations

In this discussion, an ideal image is a real function f defined over the plane. That is, we ignore boundary effects and round-off errors. We also pretend to have only one channel (in practice one examines the red, blue and green channels separately).

The discretization of f at resolution $\delta x \times \delta y$ is represented mathematically by convolution with the "Dirac comb", i.e., the distribution (see, e.g., [12] for background): $C_{\delta x,\delta y} = \sum_{k,l} \delta_{(k\delta x, l\delta y)}$, where k and l are integers. The corresponding discrete image is:

$$u = C_{\delta x,\delta y}.f$$

This formalism is convenient for dealing with multiple scales of discretization.

A digital zoom in (or just "zoom" in this paper) is a transformation which takes a discrete image at some precision $\delta x \times \delta y$ and produce a discrete image v which is an approximation to the discretization of the (ideal) image f at some better precision $\delta x' \times \delta y'$. That is,

$$v = C_{(\delta x', \delta y')}.F$$

where F is temptative reconstruction of f using only u. Following e.g. [14], in this paper we consider a linear zoom as a zoom where the reconstruction is linear and translation-invariant. Linear zooms are by far the most used and studied zooms (however see [7, 9] for arguments in favor of non-linear zooms). Hence F has the form:

$$F = u * \varphi$$

for some function φ defined on the plane with integral equal to 1. We shall also assume that the interpolation is separable, i.e., $\varphi(x,y) = \varphi_1(x)\varphi_2(y)$. Then it is enough to consider the one-dimensional case and we make this restriction henceforth. Since, little assumptions are made on φ, the given definition of linear zoom is compatible with generalized interpolations (as described in [2, 14]) and therefore includes most classical zooms.

2. REDUNDANCY IN FOURIER SPACE

We compute the Fourier transform $\hat{v}$ of v in terms of those of u and f. We have:

$$\hat{u} = \hat{C}_\delta * \hat{f} = \frac{2\pi}{\delta} C_{1/\delta} * \hat{f}$$

i.e., $\hat{u}$ is the $1/\delta$-periodization of $\hat{f}$. Of course, u being real, $\hat{u}(-\omega) = \hat{u}(\omega)^*$. We also have:

$$\hat{v} = \hat{C}_{\delta'} * \hat{F} = \frac{2\pi}{\delta'} C_{1/\delta'} * \hat{F}$$

*From C.N.R.S. U.M.R. 7640 & Ecole polytechnique, Palaiseau, FRANCE.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

and $\hat{F} = \hat{u}.\hat{\varphi}$. We obtain:

$$\hat{v}(\cdot) = \frac{2\pi}{\delta'} \sum_{k'} \hat{u}(\cdot + k'/\delta')\hat{\varphi}(\cdot + k'/\delta') \tag{1}$$

Observe that φ being smooth, $\hat{\varphi}$ is a rapidly decreasing function. Hence, one expects the dominant terms in eq. (1) to be:

$$\hat{v}(\omega) \approx \frac{2\pi}{\delta'} \Big(\hat{u}(\omega)\hat{\varphi}(\omega) + \hat{u}(\omega + 1/\delta')\hat{\varphi}(\omega + 1/\delta') + \\ + \hat{u}(\omega - 1/\delta')\hat{\varphi}(\omega - 1/\delta') \Big). \tag{2}$$

3. SELF-CORRELATIONS IN FOURIER SPACE

We now discuss self-correlations introduced in the Fourier transform by zooms and compressions and how to measure them.

3.1. Theoretical analysis for zooms

We first consider a zoomed image v and analyze the correlation spectrum of the Fourier transform in the following model. We consider u to be a random signal and make the (brash) assumption that $\hat{u}(\omega)$ and $\hat{u}(\omega')$ are independent except for $\omega - \omega'$ a multiple of $1/\delta$.

We study the correlation function $\frac{\delta'^2}{4\pi^2} <\hat{v}(\omega)\hat{v}(\omega+\Delta)^*>$ for $0 < \Delta < 1/\delta'$ taking into account the periodicity of $\hat{v}$. This correlation function becomes, using the approximation (2):

$$\sum_{\epsilon_1,\epsilon_2 \in \{-1,0,+1\}} <\hat{u}(\omega+\epsilon_1/\delta')\hat{u}(\omega+\Delta+\epsilon_2/\delta')^*> \times \\ \times \hat{\phi}(\omega + \epsilon_1/\delta')\hat{\phi}(\omega + \Delta + \epsilon_2/\delta')^* \tag{3}$$

Because of the periodicity of $\hat{u}$, the ϵ_1, ϵ_2-term above is non-zero for:

$$\omega + \epsilon_1/\delta' = \omega + \Delta + \epsilon_2/\delta' \bmod 1/\delta.$$

Hence, they occur for: $\Delta = (\epsilon_2 - \epsilon_1)/\delta' + k/\delta$, for some integer k. Because of the rapid decrease of ϕ, the non-trivial dominant terms are those for which $\epsilon_1 = 0$ (to maximize the first factor $\hat{\phi}$) and $\epsilon_2 = 1$, i.e., $\Delta = 1/\delta' - k/\delta$, for some $0 < k < \lceil \delta/\delta' \rceil$. The last factor in the corresponding term in (3) is: $\phi(\omega + \Delta - 1/\delta')^* = \phi(\omega - k/\delta)^*$. Hence the dominant non-trivial effect is obtained for $k = 1$ and the correlation function is maximum at:

$$\Delta_* = 1/\delta' - 1/\delta =: M - N \tag{4}$$

if M, resp. N, is the size in pixels of the zoomed, resp. initial, image.

3.2. Theoretical analysis for JPEG compression

We recall the principle of JPEG compression [10]. The image is cut into 8×8 pixel squares. A discrete cosine transform (DCT) is applied to each square. The resulting coefficients are rounded off to get a more compact representation. We consider the effect of this rounding off for a given coefficient i of the DCT. The contribution of this coefficient to the reconstruction of the image is:

$$w_i := a_i * (C_{8\delta}.(u.b_i))$$

for some functions a_i, b_i. Thus

$$\hat{w} = \hat{a}_i.\hat{b}_i * C_{1/8\delta} * \hat{u}$$

We see that $\hat{w}_i(\omega)$ and $\hat{w}_i(\omega + k/8\delta)$ will be correlated. We conjecture that the rounding off introduces a measure of independance between the different w_i's so that correlations between $\hat{w}(\omega)$ and $\hat{w}_i(\omega+k/8\delta)$ survive[1]. Hence we expect correlations for

$$\Delta_* = \frac{k}{8\delta} = \frac{k}{8}M \qquad k = 0, 1, 2, \ldots, 7.$$

This is observed experimentally in Figure 3 where a natural image has been compressed by a ratio of 10. These peaks grow with the compression ratio (they are barely visible for a ratio below 5).

Note that such correlation peaks give "zoom factors" of $z = 8/(8-k)$ for $k = 1, 2, \ldots, 7$. i.e., $z \approx 1.14285$, 1.33333, 1.6, 2, 2.66666, 4, 8.

3.3. Method for measuring correlations

Let $V : [1, M] \times [1, N] \to R$ be the digital image to be analyzed for the presence of a horizontal zoom. We extract L lines $V_i : [1, M] \to R$. ($V_i(k) = V(k, l_i)$ for some line numbers $l_1 < l_2 < \ldots < l_L$). We compute the discrete Fourier transforms of each one of these lines $\tilde{V}_i : [1, M] \to R$. We compute the correlation spectrum $\kappa(\Delta)$ as follows. For each frequency $\omega \in [1, M]$ we compute

$$\kappa(\omega, \Delta) = \frac{\sum_i \tilde{V}_i(\omega)\tilde{V}_i(\omega + \Delta)^*}{\left(\sum_i |\tilde{V}_i(\omega)|^2\right)^{1/2} \left(\sum_i |\tilde{V}_i(\omega + \Delta)|^2\right)^{1/2}}$$

$$\text{and} \qquad \kappa(\Delta) = \frac{1}{M} \sum_{\omega=1}^{M} \kappa(\omega, \Delta)$$

One looks for the Δ_* maximizing $\kappa(\Delta)$ away from $\Delta = 0$. If $\kappa(\Delta_*)$ is significantly higher than the background fluctuations (say the variance), then one declares that a zoom of a factor

$$z = \frac{M}{M - \Delta_*}$$

has been detected in accordance to (4).

[1] Of course doing the exact sums of the w_i's destroys these correlations since it gives back u exactly.

4. APPLICATIONS

In this section we present the averaged correlation function $\kappa(z) := \kappa(\Delta_*(z))$ for various images.

Application to known zooms: We first check that the proposed method works for the detection of zoom applied in a controlled way. We apply various zooms to scanned pictures.

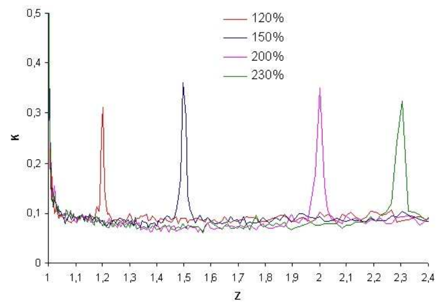

Figure 1. Bicubic zooms of various factors. The unique peak of the correlation function K gives the zoom factor Z.

In figure 1, we have applied a bicubic zoom [8] with varius zoom factors. We see that $\kappa(z)$ presents a very sharp peak exactly at the value of the zoom. The precision here is of the order of 1%.

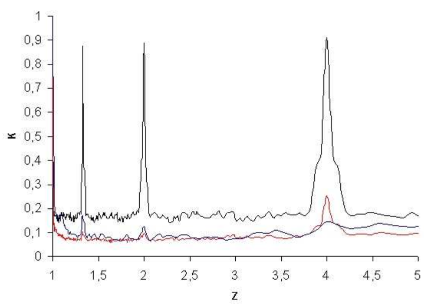

Figure 2. Comparison of the correlation function K for (a) bilinear, (b) bicubic and (c) non-linear zooms (the curves from top to bottom at $z = 4$ are (a), (b),(c)). The increasing quality shows up in decreasing correlations.

In figure 2, we compare several zooms by bilinear, bicubic and non-linear interpolation. The non-linear interpolation is a mix of bicubic interpolation in smooth zones and minimum total variation interpolation in other zones, see [1], inspired by [9]. In contrary to the linear zooms, this last zoom generates only little correlations.

Effect of JPEG compression: We compute the correlation function κ of an image compressed by JPEG. The predicted peaks corresponding to 'spurious' zoom factors of $1.14285, 1.33333, 1.6, 2, 2.66666, 4$ and 8 do appear for high enough compression rates starting around 10. The effect is approximately the same in each channel.

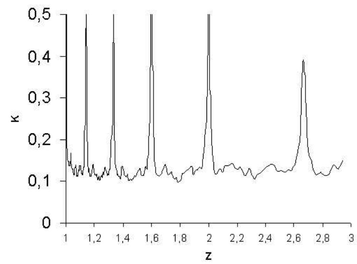

Figure 3. Correlation function induced by a JPEG-compression by a factor of 35 with sharp peaks at multiples of $M/8$ (Δ is normalized to the sampling frequency).

Analysis of digital cameras: We now present some "real life experiments", i.e., we compute the correlation function for images produced by digital cameras.

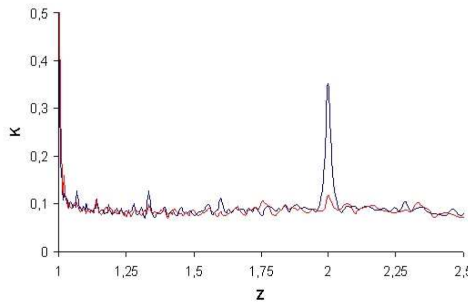

Figure 4. Canon Powershot S40. Correlation functions for two images with the embedded zoom (a) set to a factor of 2, (b) turned off.

In Figure 4, one sees a simple case where there is no other phenomenon (JPEG-compression is essentially turned off). The zoom factor gives a very well-defined peak and can be read without difficulty.

The situation is not always so favorable. In many cameras, one sees extra correlations coming from demosaicing, i.e., the interpolation from the sensor data to produce a regular RGB image. This is illustrated in Figure 5. We see that the red and blue channels are zoomed by a factor of 2 whereas the green channel is not. This is compatible with a Bayer sensor, in which the photoreceptors are organized in 4×4 squares with one red, one blue and two green sensors.

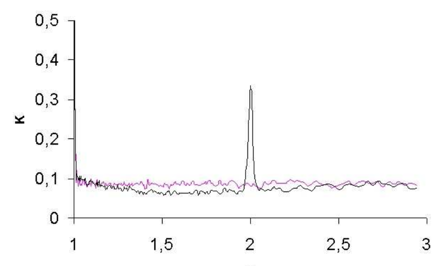

Figure 5. Kodak DC4800-(experiment made with non compressed images) Peak for $z = 2$ in the red and blue correlation function (only the red one is pictured) but not in the green one. This is induced by the demosaicing.

In the Figure 6, there is no trace of demosaicing, maybe because of the Fujifilm "SuperCCD" sensor which is organized in hexagonal cells[2]. On the other hand, one detects peaks compatible with internal JPEG compression.

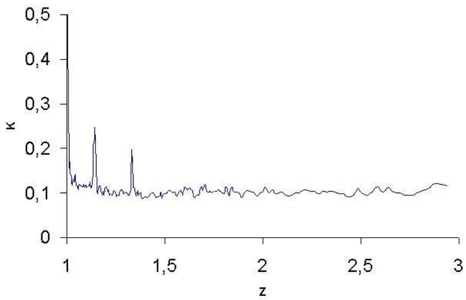

Figure 6. Fuji Digital Still Camera. No peak attributable to demosaicing is visible in the correlation function, but there are peaks coming from internal JPEG compression.

5. CONCLUSION

We have checked both theoretically under simplifying assumptions and experimentally using natural images and real life devices how numerical zoom-in and JPEG compression introduce self-correlations in the Fourier transform of an image. We have proposed a method to detect a numerical zoom automatically.

Our method still shows some limitations: we have not considered the new standard of compression JPEG-2000 [13]. We also have to point out that combination of successive zooms and/or JPEG compressions can create a number of peaks making sometimes the interpretation difficult.

In addition to the practical utility of detecting eventual numerical zoom (since the zoom factor is usually not recorded by digital cameras), our method shows up a theoretical interest: the amount of self-correlation seems related to the degradation of the image. We have seen how this could give indications on the **quality of interpolations**, compressions or digital cameras.

More precisely, we think that the amount of self-correlations could be used to estimate the information content at least in similar images. One would then obtain a new criterium besides the usual L^2 norm (see for instance [4, 5]) to measure the quality of various algorithms or digital equipments or optimize them.

Acknowledgments. We thank Eric Royer of Vision IQ for writing the relevant software.

6. REFERENCES

[1] A.I. Belamidi, *PDEs applied to image restoration and image zooming*, PhD Dissertation, Univ. Paris IX, 2003

[2] T. Blu, P. Thevenaz, M. Unser, Generalized Interpolation: Higher Quality at No Additional Cost, Proc ICIP-99, 27PO4, 1999

[3] T. Blu, P. Thévenaz, M. Unser, MOMS: maximal order interpolation of minimal support, *IEEE Trans. Image Proc.,* **10** (2001), no. 7, 1069–1080.

[4] T.Blu, M.Unser, Quantitative Fourier Analysis of Approximation Techniques: Part I—Interpolators and Projectors, Part II—Wavelets, *IEEE Trans. Signal Process.,* **47** (1999), no. 10, 2783–2806.

[5] A.M. Bruckstein, M. Elad, R. Kimmel, Down-scaling for better transform compression, *IEEE Trans. Image Proc.* **12** (2003), no. 9, 1132–1144.

[6] C. Dorai, and V. Kobla, Perceived visual motion descriptors from MPEG-2 for content-based HDTV annotation and retrieval, In *Proc. IEEE Third Workshop on Multimedia Signal Processing,* pp. 147-152, Sept., 1999.

[7] F. Guichard, F. Malgouyres, Total variation based interpolation, In *Proc. Europ. signal proc. conf.*, vol. 3, 1741–1744, 1998.

[8] R. G. Keys. Cubic convolution interpolation for digital image processing, *IEEE Trans. Acoustics, Speech, and Signal Proc.,* **29** (1981), no. 6, 1153–1160.

[9] F. Malgouyres, F. Guichard, Edge direction preserving image zooming: a mathematical and numerical analysis, *SIAM J. Num. Anal.* **39** (2002), no. 1, 1–37.

[10] W.B. Pennebaker, J.L. Mitchell, *JPEG still image compression standard,* New York: Van Nostrand Reinhold, 1992.

[11] B. Rousso, S. Peleg, I. Finci, A. Rav-Acha, Universal Mosaicing using Pipe Projection, *Proc. ICCV'98,* Bombay, January 1998, pp. 945-952.

[12] W. Rudin, *Functional analysis,* International Series in Pure and Applied Mathematics, McGraw-Hill, New York, 1991.

[13] D. S. Taubman, M. W. Marcellin, *JPEG 2000: Image Compression Fundamentals, Standards and Practice,* Kluwer International Series in Engineering and Computer Science, 642, 2001.

[14] P. Thévenaz, T. Blu, M. Unser, Interpolation revisited, IEEE Transactions on Medical Imaging, vol. 19, no. 7, pp. 739-758, 2000

[15] M. Unser, A. Aldroubi, M. Eden, Enlargement or reduction of digital images with minimum loss of information, *IEEE Trans. Image Process* **4** (1995), 247–258.

[2] Note that the interpolation in this case is not separable.