Total Variation Based Interpolation.

F. Guichard[1], *F. Malgouyres*[2]

(1) Inrets-Dart, 94114 Arcueil cedex, France.

(2) CMLA, ENS Cachan; 61, avenue du Président Wilson; 94 235 Cachan cedex France.

`e-mail : fguichar@ceremade.dauphine.fr` or malgouy@cmla.ens-cachan.fr

ABSTRACT

We propose a reversible interpolation method for signals or images, in the sense that the original image can be deduced from its interpolation by a sub-sampling. This imposes some constraints on the Fourier coefficients of the interpolated image. Now, there still exist many possible interpolations that satisfy these constraints. The zero-padding method is one of them, but gives very oscillatory images.

We propose to choose among all possibilities, the one which is the most "regular". We justify the total variation of the image as a good candidate as measure of regularity. This yields to minimise a regularisation functional defined in space domain, with a constraint defined in the frequency domain.

1 Introduction.

Many different approaches have been designed to perform image interpolation. Most of them assume that the observed image is a quantised version of a continuous image expressed in a given base of functions. These functions can be defined on the pixel or on a fix set of pixels such as spline interpolation methods (see [5, 13] and references within), oriented interpolation methods (see [12, 2, 4]), or globally as frequencies based method (e.g. zero padding, frequencies extrapolation, ...) [9, 3, 1, 7]. At last, some have tried a regularisation approach of the interpolation using quadratic regularisation functionals [6].

Frequency based methods are very efficient to catch large scale behaviour such as lines. However, generally stated with linear operations, these methods are often rejected due to their difficulties to balance between blurring and aliasing effects especially at boundaries, [12]. As we shall see, we believe that a frequency based method is justified by a reversibility condition. Secondly, we prove that within this condition, it is possible to perform such a balance. We propose to do so by minimising total variation norm, which is considered as a good (perceptually related) regularity measure [10].

In order to simplify the statements, we consider 1D functions. All the propositions can be extended without difficulties in dimension 2. We consider also a signal (or an image) sometimes as a function, and sometimes as a set of values. The signal, as a function, is supposed to be defined on $[0,1[$, and periodic. The corresponding signal of size N is linked to the function by setting $u_i = u(i/N)$. The Fourier transform of u, $\tilde{u}$ is also of size N. Frequencies are going from $-N/2+1$ to $N/2$ integer-wise. We will note these ranges $]-N/2, N/2]$ for commodity. Note that the Fourier transform $\tilde{u}$ is also periodic of period N, so that one can also consider the frequencies to be in $[0, N[$. But, one has to be careful, that the frequency N for a signal of size N corresponds to the frequency 0, whereas it is not true anymore for a signal of larger size (e.g. interpolated signal) since the periodicity is different.

In addition, we will consider here only the interpolation of factor 2, but once again all the propositions can be extended in the case of other integer factors. We always denote by u the original signal of size N, and by w a two times larger interpolation.

2 Reversibility condition of the interpolation.

We want the interpolation method to be reversible, so u shall be deduced from w by a sub-sampling. We are looking for a w such that

$$u = QS(w) \tag{1}$$

Where S is a translation invariant linear smoothing, and Q is a sub-sampling operation. That is that $Q(v)$ is deduced from v by taking one point every two ($Q(v)_i = v_{2i}$). As it is well-known, Q introduces aliasing [9], so one has :

Proposition 1 *Let u and w be two signals, of size respectively N and $2N$, linked by the reversibility condition that is by the relation (1), then one has*

$$2\,\tilde{u}_i = \tilde{s}_i \tilde{w}_i + \tilde{s}_{i+N} \tilde{w}_{i+N} \tag{2}$$

where $\tilde{u}$, $\tilde{w}$ and $\tilde{s}$ are the Fourier transforms of u, w, and of the convolution kernel of S.

The relation (2) constrains the possible functions w, and in the same way indicates where is the freedom in the choice of the interpolation. For each frequency of u : $\tilde{u}_k$, we can choose its repartition within two frequencies of w : $\tilde{w}_k$ and $\tilde{w}_{k+N}$. Shannon's theorem [11], gives the exact inverse of the quantisation in the case of an original signal w having a Fourier transform supported inside $]-N/2, N/2[$ (the aliased term is zero). This yields to a particular choice of repartition : the zero-padding; which makes the choice to distribute everything on the lowest frequency. It is defined by the following relations : $\tilde{w}_i = 2\,\tilde{u}_i/\tilde{s}_i, \forall i \in]-N/2, N/2[$, $\tilde{w}_{N/2} = w_{-N/2} = u_{N/2}$, and, $\tilde{w}_i = 0$ otherwise. Now in the general case, as shown in the experiments section, this method yields oscillatory images due to Gibbs effect (see figure 3). At this point, there is no mathematical consideration that can say how this repartition has to be done.

Let us call $\mathcal{W}_{u,s}$ the set of functions w that satisfies the constraint (2). It is easy to see that this is an affine subspace. ($\forall w_1, w_2$ in $\mathcal{W}_{u,s}$, one has $(1-\alpha)w_1 + \alpha w_2 \in \mathcal{W}_{u,s}$).

We propose to choose, among functions of $\mathcal{W}_{u,s}$, the one which is the most "regular". Of course, we then have to precise what we mean by "regular"...

3 Choice of a measure of regularity of the images.

We propose to define such a measure among norms L^α, $\alpha > 0$ of the gradient. A particular α will be chosen later. We define, for a periodic signal of size N

$$E_\alpha(w) = N^{\alpha-1} \sum_{k=0}^{N-1} (|w_{k+1} - w_k|^\alpha) \qquad (3)$$

It is the Riemann approximation of $\int_0^1 |\nabla w|^\alpha$ in the case of a w defined on the continuous domain $[0, 1[$. The normalisation by $N^{\alpha-1}$ of the energy makes comparable the energy of a signal and its zoomed versions. Let us now state our aim.

P. *Given a signal u of size N, we are looking for a signal w or size $2N$ such that it minimises $E_\alpha(w)$, among the functions of $\mathcal{W}_{u,s}$.*

We assume that the function $\tilde{s}$ is known or chosen. In general case, in order to avoid excessive deblurring, we will assume that s is normalised such that $\sum_{i=-N+1}^{N} s(i) = 1$, and that $\tilde{s}(i)$ is not null for all $i \in]-N/2, N/2]$. Under these conditions, we are sure that there always exists at least one function w that satisfies the constraint, that is $\mathcal{W}_{u,s} \neq \emptyset$.

Before investigating different power α, let us first consider the classical H^1 norm, that is $\alpha = 2$.

Proposition 2 *With $\alpha = 2$, the function that achieves the minimum of the energy (3) up to the constraint (2) is unique, and is given by :*

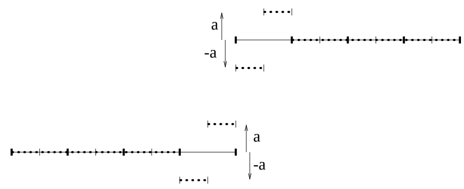

Figure 1: The dot signal has a lower H^1 norm than the continuous one, and they are both sub-sampled in a heavy-side function, in the case where S is a mean on two pixels. H^1 is not a good measure of regularity.

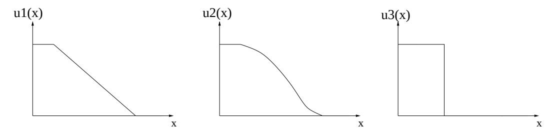

Figure 2: These three signals differ by the way they go from 1 to 0. From Left to Right, the maximum of the gradient (slope) of the signal increase. The norm E_α with α larger than 1, will prefer the left signal. With α lower than 1, it will prefer the right one. And, all of them are equivalent with $\alpha = 1$.

for $k \in]-N/2, N/2] \setminus \{0\}$, and $t \in \{0, 1\}$

$$\tilde{w}_{k+tN} = 2 \frac{\tilde{s}_{k+tN}}{|exp(i2\pi\frac{k+tN}{2N}) - 1|^2 \phi(k)} \tilde{u}_k \ , \qquad (4)$$

where

$$\phi(k) = \frac{\tilde{s}_k^2}{|exp(i2\pi\frac{k}{2N}) - 1|^2} + \frac{\tilde{s}_{k+N}^2}{|exp(i2\pi\frac{k+N}{2N}) - 1|^2}$$

and $\tilde{w}_0 = 2\,u_0$, and $\tilde{w}_N = 0$.

Proposition 3 *If one has for all $|k| > N/2$, $\tilde{s}_k = 0$, and $\tilde{s}_{N/2} = \tilde{s}_{-N/2} \neq 0$, then with $\alpha = 2$, the function that achieves the minimum of the energy (3) up to the constraint (2) corresponds to the zero-paddings.*

These propositions show that we can directly derive the solution of the problem stated with the quadratic norm of the gradient ($\alpha = 2$). Surprisingly, in the introduction we have criticised the zero-padding method because it yields toward Gibbs effects and irregular images. And here we show that the zero-padding interpolation sometimes minimises the quadratic norm of the gradient. Moreover, looking at the proposition 2 one can see that we just duplicate the spectra (modulo a weight), and that we can not recover lost frequencies (where $\tilde{s} = 0$)(see figure 4). Therefore, we conclude that this norm is not a good measure of regularity, (see figure 1). We are led to choose another power α.

In order to understand the differences between the possible powers, let us consider an example. Let u be a function from $\mathbb{R}$ into $\mathbb{R}$ such that $u(x) = 1$ for $x \leq 0$, $u(x) = 0$ for $x \geq 1$. For x between 0 and 1, we will consider different shapes (see figure 2 for u_i's definitions).

	$E_\alpha(u_1)$	$E_\alpha(u_2)$	$E_\alpha(u_3)$
$\alpha > 1$	1	> 1	$+\infty$
$\alpha = 1$	1	1	1
$\alpha < 1$	1	> 0	0

The preceding tabular indicates that a minimisation of E_α will yield different shapes of functions depending on whether α is larger, equal, or lower than 1. A α larger than 1 yields a function as flat as possible. Conversely a α lower than 1 prefers a function as straight as possible. At last, with $\alpha = 1$, any monotone way to go from 1 to 0 gives the same value. Therefore, the minimisation of E_1 will not decide on the shape of the signal.

With respect to the zoom problem, we believe that we do not have to decide whether or not a signal (or an image) has to be smooth or unsmooth. This should be driven somehow by the existing low frequencies, and not by an a-priori norm. According to us, this justifies by itself the choice of $\alpha = 1$, that is the total variation of the image. With such a norm, we do not enforce a particular shape on the signal (smooth or unsmooth). We minimise in fact the amount of variation of the signal, that is its oscillations [10]. In other words, we prefer a monotonous signal to oscillating one.

We also remark that all the functions w that satisfy the reversibility condition have a total variation larger than the function u. Indeed,

Proposition 4 *For any $\alpha \geq 1$, and any signal u.*

$$E_\alpha(u) \leq ||s||_1^\alpha min_{w \in \mathcal{W}_{u,s}} E_\alpha(w)$$

with $||s||_1$ the l^1 norm of the filter s.

We believe this proposition justifies our approach of the interpolation (seen as a regularisation problem). Moreover, this means that the quantisation process smoothes a signal (in the sense that it makes E_α decrease), which is an argument in favour of E_α for $\alpha \geq 1$.

Now, despite the convexity of the total variation, in general the problem does not have a unique solution. However, there is no local-minima, and the set of solutions is convex. And in addition, as shown by the following proposition, solutions are not far from each others.

Proposition 5 *Let w_1 and w_2 achieving the minimum of E_1 under the constraint, then $w_1(k+1) - w_1(k)$ and $w_2(k+1) - w_2(k)$ (resp. ∇w_1 and ∇w_2) have the same sign (in 1-D), (resp. direction (in N-D)) .*

4 Minimising the total variation in the spatial domain with a reversibility condition in frequency domain.

In this section we propose a simple schemes to compute a solution at the problem **P** which is now stated as follow :

$$min \sum |\nabla w|, \text{ under the constraint that } w \in \mathcal{W}_{u,s}$$

Figure 3: Up-left : Part of Lenna image zoomed x8 by duplication. Up-Right : Zoom by the zero-padding method. Down-left : Zoom yields by the minimisation of the H^1 semi-norm. Down-right : Zoom yields by the minimisation of the total variation of the image.

We begin with a function w_0 of $\mathcal{W}_{u,s}$, e.g. the zero-padding interpolation of u. We then define a "constrained" gradient descent, by projecting the gradient of the energy on the space $\mathcal{W}_{0,s}$.

The gradient of the total variation is given by the function $k(w) = div(Dw/|Dw|)$. Since, $\mathcal{W}_{0,s}$ is a vectorial subspace, we can consider the orthogonal projection $P_{\mathcal{W}_{0,s}}$ on this subspace. The "constrained" gradient descent is then given by :

$$\frac{\partial w}{\partial t} = P_{\mathcal{W}_{0,s}}(div(\frac{Dw}{|Dw|})) \quad (5)$$

So, given w_n, we compute $\frac{\partial w_n}{\partial t}$, and then compute the optimal step s such that $\sum | \nabla(w_n + s\frac{\partial w_n}{\partial t}) |$ is minimum. We then compute $w_{n+1} = w_n + s\frac{\partial w_n}{\partial t}$, and iterate.

5 Experiments.

We display in figure 3 different zooms of factor 8 of a part of the famous Lenna image. We choose for the linear smoothing s, the average on the square 8x8 pixels corresponding to one pixel of the original image. Moreover we compute the gradient using a simple finite difference, but the result presented here also holds for other kinds of schemes. Since the average does not satisfy fully the condition of Proposition 3, the zero-padding interpolation differs from the H^1 norm minimisation. In the Up-right image, we see that the zero-padding yields a very oscillatory image. Due to the chosen smoothing

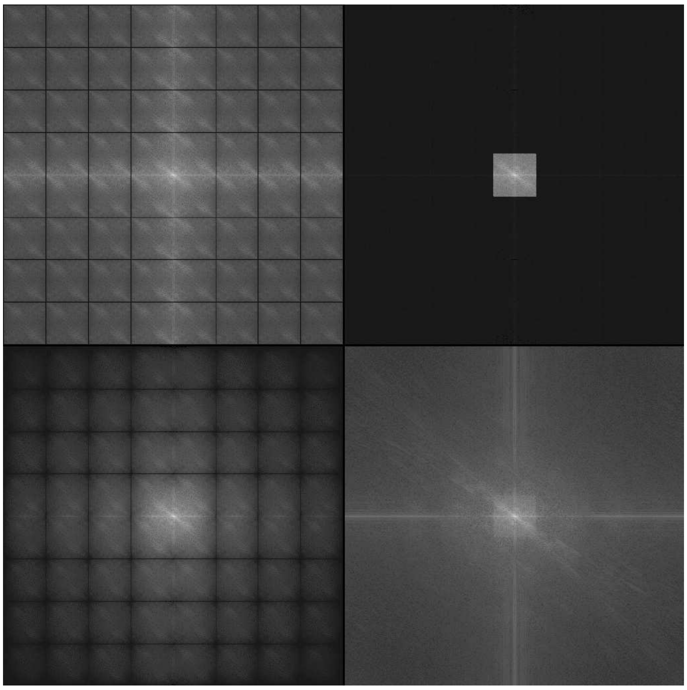

Figure 4: Spectre of images in figure 3 which are zoom x8 by: Up-left : duplication. Up-Right : Zero-padding method. Down-left : H^1 minimisation. Down-right : Total variation minimisation.

the H^1 norm gives a more acceptable image, in spite of remaining oscillations and "staircase effect" close to strong edges. At last, the zoom based on total variation (Down-right) does not have Gibbs effect, and removes as well the "staircase effect" on non vertical or horizontal boundaries. Moreover, one can see on figure 4, the modulus of images spectra of figure 3(the contrast is enhanced). One can notice that they all have particular shapes which do not seam natural except the one concerning the total variation. This one tends to prolong spectre's shape. This is a consequence of the way the total variation deals with edges.

6 Conclusion.

As shown in the experiment section, we have defined a frequency interpolation method that does not create Gibbs effect on the image. However, the minimisation of the total variation under the interpolation reversibility constraint does not define a unique interpolation. And, the implemented minimisation process chooses among the possible one's. That means also, that there still exists a freedom that could be used to "ask" something more for the interpolated image, as we have done for the regularity.

Another issue, is the choice of the measure of "regularity". Here, we have chosen the total variation norm, because it measures the amount of oscillation of the gray-level. In 2D, that is for images, the total variation measures also an other kind of oscillation. Indeed, TV is exactly equal to the sum of the length of the level lines of the image. Its minimisation yields the removal of some level lines (reduction of the gray-level oscillation), but also to the minimisation of their length. This means that we minimise also the oscillation of shapes boundaries. And for example, a circle is consider as far more regular than a square. In the same way that we choose the total variation norm because it does not decide between different shapes of monotonous signal, we should reject it because it chooses between convex shapes...

Acknoledgment. The authors would like to thank J.M. Morel and S. Durand for all fruitful discussions on that subject.

References

[1] J. Agbinya, *Interpolation Using the Discrete Cosine Transform.* Electronics Letters, U.K., vol. 28, No 20, pp. 1927-1928. 1992,

[2] J. Allebach, P.W. Wong *Edge directed interpolation.* Proceedings. International conference on image processing. pp. 707-10, 1996.

[3] F. Brouaye, *Sampling and Wavelets: A New Approach to the Zero-Padding Method.* Traitement du Signal, France, vol 9, No 2, pp. 193-200, 1992.

[4] S. Carrato, G. Ramponi, S. Marsi *A simple edge-sensitive image interpolation filter.* Proceedings. International conference on image processing. pp. 711-14, 1996.

[5] N.A. Dodgson, *Quadratic Interpolation for Image Resampling* IEEE Trans. on image processing, vol 6. No. 9. 1997.

[6] Karayiannis and Venetsanopoulos. *Image interpolation based on variational principles* Signal Processing 25, pp 259-288, 1991.

[7] C. Lee T. Komuro. *Electronic image resolution enhancement by frequency-domain extrapolation.* US patent abstract. No 5,666,163, Sept. 9, 1997.

[8] J.A Parker, R. V. kenyon, and D. troxel, *Comparison of interpolation methods for image resampling.* IEEE trans. Med. Imag. MI-2, p 31-39, 1983.

[9] K.W. Pratt; *Digital image processing.* John Wiley and Sons, Inc,1978.

[10] L. Rudin, S. Osher. *Total Variation based restoration of noisy, blurred images,* submtd. SIAM J. Num Analysis, 1992.

[11] C.E. Shannon, *Communication in the presence of noise.*. Proc. I.R.E., 37 (1949), pp 10-21.

[12] S. Thurnhofer, S. Mitra, *Edge-enhancement image zooming.* Opt. Eng., vol 35 no. 7, p 1862-1870, 1996.

[13] M. Unser, A. Aldroubi, M. Eden. *Fast B-Spline Transforms for Continuous Image Representation and Interpolation.* IEEE trans. on pattern Analysis and machine intelligence,vol 13 no.3 pp 277-285, 1991.