FAST AND EFFICIENT RESAMPLING FOR MULTI-FRAME SUPER-RESOLUTION

Benoît Vandame

Canon Research Centre France,
Rue de la touche Lambert, F-35510 Cesson Sévigné, FRANCE, benoit.vandame@canon.crf.fr

ABSTRACT

Multi-frame super-resolution aims at producing a single high-resolution image from a set of low resolution images. The aliasing of the input low resolution images can be recovered thanks to multi-frame acquisition of the scene observed with faint camera motion. On another hand, the super-resolution process must be efficient and light to be embedded into products. This article describes a straight super-resolution scheme which emphasises the role of re-sampling. A new family of interpolation functions is described in details and then benchmarked with grey-scaled as well as colour filter array images. The proposed algorithm has the same complexity than basic image re-sampling algorithms and results show a significant gain in resolution with no ringing artefacts.

***Index Terms*—** super-resolution, image warping

1. INTRODUCTION

Multi-frame super-resolution aims at producing a single high-resolution image from a set of low resolution images. The input set is aliased with arbitrary faint camera motion. These two conditions together make super-resolution possible: the aliasing ensures that high-frequencies are not lost but mirrored due to the under-sampling of the sensor; the faint camera motion makes each frame containing new information about the scene[6].

This article deals with signal reconstruction of higher resolution grid from samples on non-uniform grid. An efficient and light solution is desired in order to be embedded into products. Accordingly, a basic and straight super-resolution scheme is proposed where re-sampling plays a major role, and de-blurring is just a post-processing sharpening filter. A new family of interpolation function dedicated for super-resolution is described and benchmarked on grey level images as well as colour filter array images. Results with significant increase in resolution with no ringing artefacts are obtained with a very low complexity re-sampling algorithm.

The article is organized as follows: first the proposed super-resolution scheme is depicted; a review of basic interpolation functions is given; the new re-sampling algorithm dedicated to super-resolution is described in details; and finally results with real images are shown.

2. SIMPLE SUPER-RESOLUTION SCHEME

The super-resolution model considered in this article is formulated by independent steps: 1/ image registration; 2/ image resampling & fusion; and 3/ image de-blurring. This basic scheme offers straight computation with no iteration between the super-resolved image and the low resolution frames. The 3 stages are briefly described below:

Image registration

This stage can be performed by various approaches from a dense optical flow to a global geometric distortion. For simplicity an affine distortion with 6 free parameters is considered (covering: translation, rotation, zooming and shear). They are computed by least square estimation from points of interest extracted from one image set as reference and associated with the corresponding set of points from other images of the sequence.

The geometric transformation is either forward (from one given image to the reference image) or backward (from the reference image to the given image). The affine transforms are further zoomed by the super-resolution zooming factor such as the affine transforms associate a coordinate from a reference image with a coordinate of the super-resolution grid.

Image re-sampling & fusion

The input images are re-sampled one by one according to their zoomed affine distortion model. Image re-sampling is performed by interpolation. While the input images are re-sampled and accumulated into a co-added image, a weight image having the size of the super-resolution grid records the number of accumulated input pixels. All images being re-sampled, the co-added image is then divided by the weight-map.

Image de-blurring

The co-added image is blurred according to the zoomed blur function of the input images. The initial blur function is modelled knowing the characteristics of the camera. This model is scaled to match the super-resolution grid.

For a fast de-blurring of the co-added image, the inverse of the scaled blur function is estimated for instance by common algorithms such as: van-Citter de-convolution[1] or Richardson Lucy de-convolution[4]. The co-added image is then convolved by the inverse estimate. This "direct" strategy offers goods results with reasonable noise because the noise of the co-added image is low due to number of input images.

3. COMMON IMAGE RESAMPLING

Image warping is solving the re-sampling issue of having non-integer transformed coordinates by the means of interpolation. The distortion-models (backward or forward) deliver real (non-integer) coordinates for most integer input coordinates. The interpolation allows placing (forward case) or reading (backward case) a distorted pixel value with a non-integer coordinate into a sampled grid. Unfortunately the interpolation has a visible impact on the warped image. The most common impact is that the high frequencies of the warped image are lowered down, thus providing ringing effect after the de-blurring stage.

It is important to understand the impact of image warping because it is followed by de-convolution. Input images are assumed to be aliased; image warping approximates a *sinc* function which tends to remove the frequencies beyond the Shannon sampling frequency. The input aliasing is lowered down by image warping before it could be unfold by re-sampling and fusion. In other words: image warping prevents fusion of images to recover the aliased frequencies.

3.1. Interpolation kernel

One considers the common interpolation function $I(l)$ based on Lanczos windowing:

$$I(l) = \frac{\sin \pi l/n}{\pi l/n} \cdot \frac{\sin \pi l}{l}$$

Where n is the order of the Lanczos function. The distortion model associates an integer coordinate (x,y) with a non-integer transformed coordinate (X,Y). The transformed coordinate is decomposed into integer part and fractional part $(X,Y) = (\lfloor X \rfloor + \{X\}, \lfloor Y \rfloor + \{Y\})$ where: $\lfloor . \rfloor$ is the ceiling function (the integer part); and $\{\}$ is the fractional function (the non-integer part). The integer part of the coordinate corresponds to basic integer translation, whereas the non-integer part is performed by interpolation. The 1D Lanczos interpolation function $I(l - \{X\})$ is defined for $l =]-n, n]$. The $2n$ values define the 1D interpolation kernel $I_{\{X\}}$. Equivalently one gets 1D interpolation kernel $I_{\{Y\}}$ for the interpolation in the Y direction. The complete 2D interpolation kernel for the fractional shift $\{X,Y\}$ is given by $I_{\{X,Y\}} = I_{\{X\}} \cdot I_{\{Y\}}{}^{T}$.

From a computation point of view, 1D interpolation kernels are pre-computed for $\{X\}$ = 0 to 1 by step of 0.01 or 0.001 depending on the precision required. This pre-computation allows saving many sinus computation per warped coordinate.

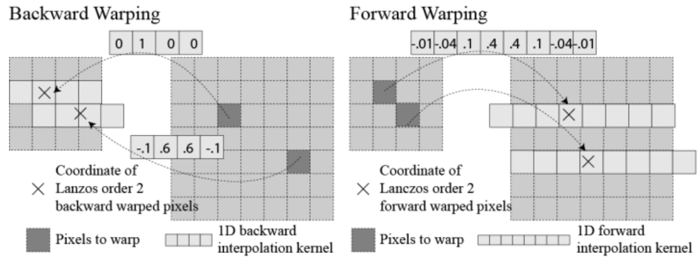

Fig. 1. Detail of Lanczos interpolation for image zooming: backward and forward case.

3.2. Interpolation kernel and zooming

Zooming is just a special type of distortion. Zooming and warping are performed jointly with the interpolation function being scaled by zoomed factor s. The interpolation function ensures that the zoomed image is properly sampled according to the Shannon sampling frequency of the input image.

Image warping and zooming operates slightly differently in case of forward and backward cases. Figure 1 illustrates both cases of image zooming by a factor of 2 with Lanczos interpolation (only 1D interpolation kernels are illustrated). Backward and Forward cases are processed as follows:

- For the backward warping, the s^2N super-resolution pixels are warped for each input image made of *N* pixels. Backward image warping is done independently to the zoom factor *s*. The smoothing is naturally performed by the s^2N warped pixels. The 1D interpolation kernel is computed by sampling the function $I_{Back}(l, \{X\}) = I(l - \{X\})$ for $|l| < n$, l being an integer.
- For the forward Lanczos warping, the 1D interpolation kernel is computed by sampling the interpolation function $I_{Forw}(l, \{X\}, s) = I\left(\frac{l - \{X\}}{s}\right)$ for $|l| < sn$, l being integer. The 1D forward interpolation kernel size is equal to $\lfloor 2ns \rfloor + 1$. Forward warping is applied on the *N* pixels of each input image. Fewer pixels are warped compared to a backward image warping, but the 2D forward interpolation kernels are larger. It results in the same number of calculus.

After image warping, the re-sampled image has a frequency cut equal to $1/2s$. This frequency cut is given by the interpolation kernels which model a scaled *sinc* function. The frequencies of the input images beyond ½ have been filtered out. The frequency cut-off prevents the de-blurring stage to recover the high-frequencies; thus super-resolution cannot be obtained by bi-cubic or Lanczos warping. Several solutions are proposed to adapt the interpolation [7]. To solve this issue, an interpolation function with controlled frequency-cut is proposed.

4. WEIGHTED IMAGE RESAMPLING

In super-resolution, image warping should adapt to the number of input images. For instance, one considers the super-resolution image computed from 4 low resolution images which are observed with half pixel shift in both axes. Those 4 images are combined into a super-resolved image by just copying pixels in an interleaved manner. Common naïve Lanczos image warping produces a smooth version compared to that interleaved version.

Resampling the multi-frames into the super-resolution grid refers to non-uniform interpolation, also known as scattered interpolation. This is a major stage for multi-frame super-resolution as described in [3][5].

4.1. Weighted interpolation kernel

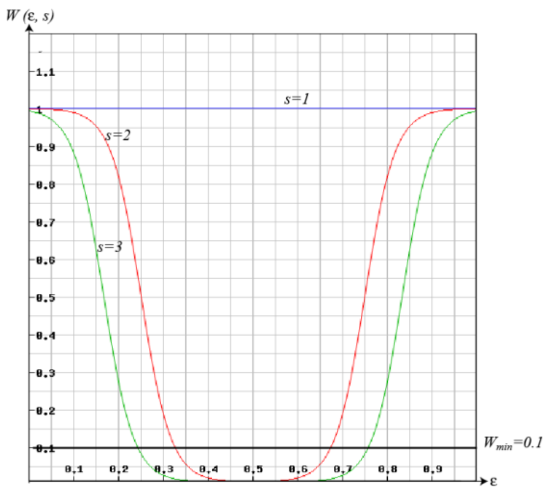

Fig. 2. Illustration of the weighting function as a function of the fractional part and the zoom-factor.

The proposed solution is to associate a weight to an interpolation kernel. The weight is function of 2 parameters: a fractional part $\varepsilon \in [0,1]$ and a zoom-factor s. The proposed weighing is performed in 1D as follows:

$$W(\varepsilon, s) = \max\left(\frac{1}{1+e^{k\left(\varepsilon - \frac{1}{2s}\right)}} + \frac{1}{1+e^{-k\left(\varepsilon - 1 + \frac{1}{2s}\right)}}, W_{\min} \right)$$

The weighing is just the sum of 2 sigmoid functions: the first sigmoid is providing a strong weight for fractional part ε below $\frac{1}{2s}$, the second sigmoid is providing a strong weight for ε beyond $1-\frac{1}{2s}$. The weighting is converging to zero for the fractional part between the 2 inflexion points. The parameter k controls the slope steepness around the inflexion points. $W_{\min}$ controls the minimum weight. It prevents the $W(\varepsilon, s)$ to converge to 0. Indeed, co-added pixels might receive no contribution due to null weight, this case occurs when the arbitrary camera motion is not random. Figure 2 illustrates the weighting function for various zoom-factors (1, 2 and 3), with *k=30*, $W_{\min} = 0$. A common choice of $W_{\min}$ is illustrated by the plain horizontal line. The weighted warping is identical to the standard warping when the zoom factor is 1.

4.2. Application of the weighted interpolation kernel

Applying the weighted image warping is quite different from the backward and forward case:

- For the weighted backward warping, a 1D interpolation kernel is computed by sampling the weighted interpolation function
$$I_{BackWeight}(l, \{X\}, s) = I(l - \{X\}) \cdot W(\{l - \{X\}\}, s)$$
for $|l| < n$, l being an integer. The interpolation function can be simplified because $\{l - \{X\}\}$ is equal to $\{X\}$. It results into a sampling of the weighted interpolation function $I(l - \{X\}) \cdot W(\{X\}, s)$. In other words the weighting is only function of $\{X\}$ and s, it is independent of l the indices within the 1D interpolation kernel.
- For the weighted forward warping, a 1D interpolation kernel is computed by sampling the interpolation function
$$I_{ForwWeight}(l, \{X\}, s) = I\left(\frac{l - \{X\}}{s}\right) \cdot W\left(\left\{\frac{l - \{X\}}{s}\right\}, s\right)$$
for $|l| < sn$, l being an integer. No simplification can be performed.

The complexity of computing the 1D weighted interpolation kernels is not an issue because they are preliminarily pre-computed as for standard image warping. The weighted image warping needs to be performed with a weight-map having the size of the super-resolved grid. In case of backward warping the weight of $W(\{X\}, s) \cdot W(\{Y\}, s)$ is added to the weight-map at the location of the super-resolved pixel (x, y). In case of the forward warping the 2D interpolation kernel is added into the weight-map at location $(\lfloor X \rfloor, \lfloor Y \rfloor)$. Figure 3 illustrates the standard Lanczos interpolation function order 4 multiplied by weighted function $W(\varepsilon, s)$ for various sampling factor *s*.

4.3. Fourier Analysis

Fourier analysis shows that weighed interpolation kernel does not produce a frequency cut-off. A sharp weighted (Door weighing) *sinc* interpolation function $I_p(l)$ is equal:

$$I_p(l) = \frac{\sin(\pi l)}{\pi l} \left[\Delta(l) \otimes \Pi_{1/2s}(l) \right]$$

Where $\Pi_{1/2s}(l)$ is a Door function equal to 1 for $|l| \leq 1/2s$, and $\Delta(l)$ is the Dirac comb function.

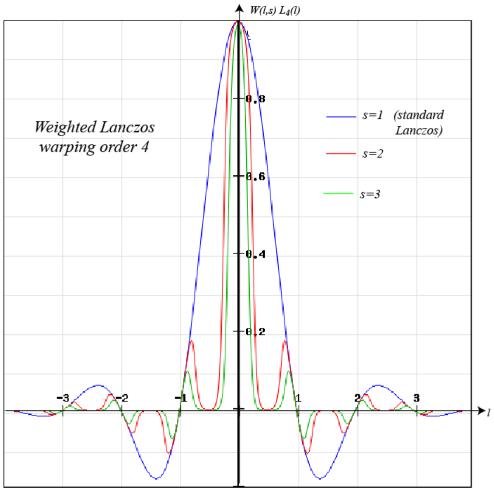

Fig. 3. Visualization of the weighted Lanczos interpolation functions for various zoom factors.

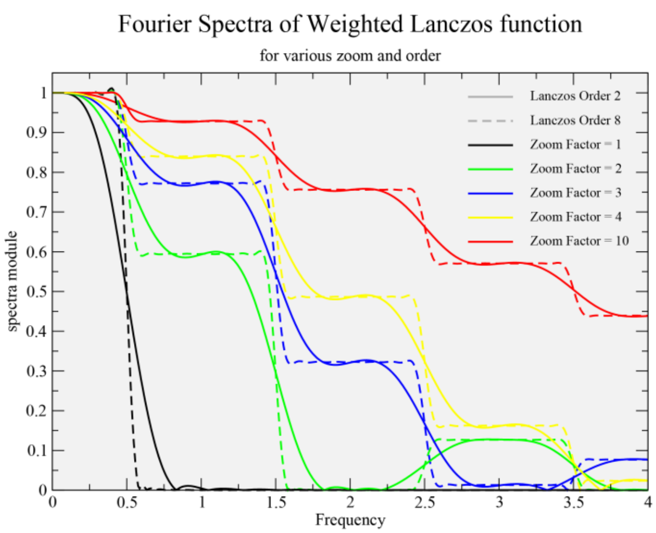

Fig. 4. Fourier transform of the weighted Lanczos interpolation function for various zoom factor.

The Fourier transform $i_p(l)$ is trivial and equal to:

$$i_p(\upsilon) = \Pi(\upsilon) \otimes \left[\Delta(\upsilon) \frac{\sin(\pi\upsilon/2s)}{\pi\upsilon} \right]$$

The $i_p(\upsilon)$ is a sum of translated Door functions with decreasing amplitude. Figure 4 illustrates the Fourier transform of weighted Lanczos for various super-resolution factor s (the weighting is given by the two sigmoid functions with parameter k=30), one notices the various steps defined by the sum of translated Door functions.

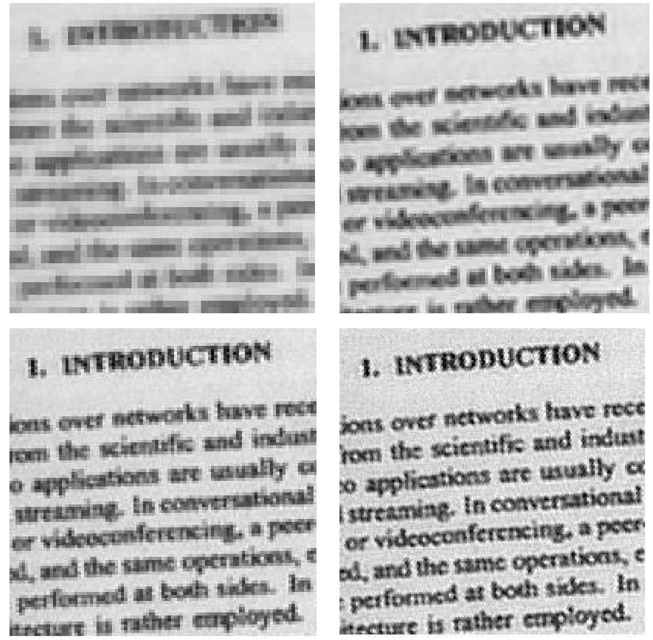

Fig. 5. From right to left, top to bottom: A/ One of the 30 input images; B/ super-resolved image by a factor of 2 with common forward Lanczos warping; and C/ & D/ super-resolution with a forward weighted Lanczos warping with a super-resolution factor 2 and 4 respectively.

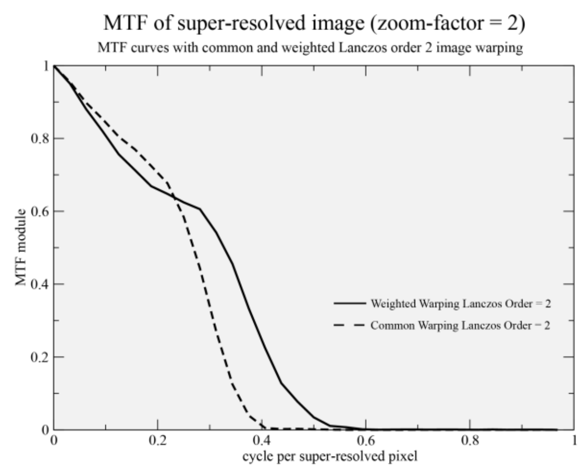

Fig. 6. Modulation Transfer Function of the super-resolved image. Plain line: with weighted image warping; dashed line: with common image warping.

5. APPLICATION

5.1. Application to grey level images

The proposed re-sampling algorithm is applied to a set of grey-scaled images obtained with a micro-eye camera. The aliasing of the input images is important. Figure 5 illustrates the super-resolution images without and with weighted image warping for a super-resolution factor of 2 and 4. One notices the absence of ringing effect with the weighted image warping. The weighted image warping with the super-resolution factor of 2 and 4 looks similar (the variation of contrast might be related to the estimation of the inverse blur function).

Figure 6 illustrates the Modulation Transfer Function (MTF) computed with a slanted edge. The MTF obtained with common Lanczos warping tends to 0 at a frequency of 0.4, which characterizes the ringing effect. For the weighted image warping, the MTF tends to 0 at a frequency of 0.5.

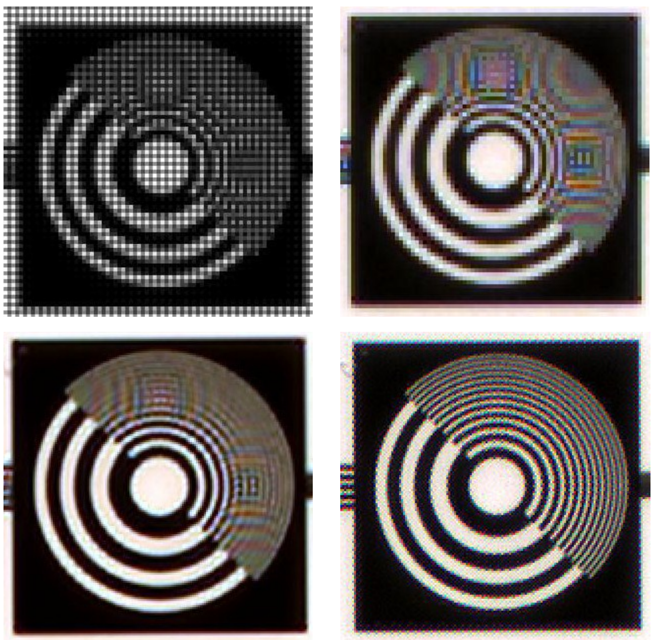

Fig.7. From right to left, top to bottom: A/ One of the 30 input raw Bayer images; B/ one raw image de-mosaiced with Hamilton algorithm; C/ super-resolution applied on the RGB images with common Lanczos interpolation; D/ super-resolution applied on raw images with the proposed weighted Lanczos interpolation.

5.2. Application with Colour Filter Array

Most of the colour images are taken from sensor mounted with a Colour Filter Array (CFA). From the raw image, where one colour is recorded per pixel, to a RGB image a de-mosaicing algorithm interpolates the missing colour components (as illustrated respectively in Fig. 7A and 7B). In a super-resolution context, it is worth combining the raw images straight into the super-resolution grid without recovering the missing colour components. Indeed, de-mosaicing algorithms perform non-linear interpolations which produce artificial aliasing for the missing colour components. Merging the artificial aliasing of various input images into the super-resolution grid produces smooth and unstable results.

Super-resolution of CFA images refers multi-frame de-mosaicing[2]. Figure 7 illustrates the super-resolution by a factor of 2 when applied on RGB images and raw images (Fig. 7C and 7D respectively). The recovery of high resolution from RGB images is not possible because de-mosaicing produces wrong missing components which do not benefit from image re-sampling and fusion. In contrast, the recovery of details is good with super-resolution performed on raw images with a weighted Lanczos interpolation.

It is worth noting that super-resolution from raw image is computed with only via a forward warping (one raw pixel is cumulated into the co-added image following to the 2D interpolation kernel), backward warping is impossible due to the CFA. The Red and Blue pixels are super-resolved by *2s* as they are under-sampled by a factor of 2 on the raw images observed with a Bayer pattern. The jaggy results visible on Fig. 7D are due to “missing” observations: some super-resolution pixels are associated to a null weight (or a minimal weight). These pixels could be spatially interpolated just after the raw images are re-sampled and fused into the co-added image.

6. CONCLUSION

A new family of interpolation functions dedicated to multi-frame super-resolution is described in details and is showing significant resolution improvement. The computation is performed by a low complexity algorithm prone to be embedded into products.

The quality of the super-resolved image is function of the randomness of the relative motions between the input images and the level of aliasing of the input images.

The proposed super-resolution procedure is independent to the image content. The re-sampling weighted interpolation function could be enhanced with a post processing in order to interpolate co-added pixel associated to a null of low level weight. This post processing based on spatial prior knowledge would decrease the jaggy effect observed on super-resolved images.

7. REFERENCES

[1] A. Bennia, S.M. Riad, "Filtering capabilities and convergence of the Van-Cittert deconvolution technique," IMTC-1991. 8th IEEE, pp.319-323, 14-16 May 1991

[2] S. Farsiu, M. Elad, P. Milanfar, "Multiframe demosaicing and super-resolution of color images," Image Processing, IEEE Transactions, vol.15, no.1, pp.141-159, Jan. 2006

[3] A. Gilman, D.G. Bailey, "Near optimal non-uniform interpolation for image superresolution from multiple images" IVCNZ'06, pp.31 -36, Nov. 2006.

[4] W. Richardson, "Bayesian-Based Iterative Method of Image Restoration," J. Opt. Soc. Am. 62, 55-59 (1972).

[5] D. Shepard, "A two-dimensional interpolation for irregularly-spaced data.” ACM '68 Proceedings of the 23rd ACM national conference, January 1968.

[6] P. Vandewalle, S. Süsstrunk and M. Vetterli, “A Frequency Domain Approach to Registration of Aliased Images with Application to Super-Resolution”, EURASIP ID 71459, 2006.

[7] Sung C. Park, Min K. Park, Moon G. Kang “Super-resolution image reconstruction: a technical overview” Signal Processing Magazine IEEE, Vol. 20, No. 3. (May 2003), pp. 21-36.