An Image Interpolation Scheme for Repetitive Structures

Hiêp Luong, Alessandro Ledda, and Wilfried Philips

Ghent University, Department of Telecommunications and Information Processing, B-9000 Ghent, Belgium

Abstract. In this paper we present a novel method for interpolating images with repetitive structures. Unlike other conventional interpolation methods, the unknown pixel value is not estimated based on its local surrounding neighbourhood, but on the whole image. In particularly, we exploit the repetitive character of the image. A great advantage of our proposed approach is that we have more information at our disposal, which leads to a better reconstruction of the interpolated image. Results show the effectiveness of our proposed method and its superiority at very large magnifications to other traditional interpolation methods.

1 Introduction

Many applications nowadays rely on digital image interpolation. Some examples are simple spatial magnification of images or video sequences (e.g. printing low resolution documents on high resolution (HR) printer devices, digital zoom in digital cameras or displaying Standard Definition video material on High Definition television (HDTV)), geometric transformation and registration (e.g. affine transformations or computer-assisted alignment in modern X-ray imaging systems), demosaicing (reconstruction of colour images from CCD samples), etc.

Many interpolation methods already have been proposed in the literature, but all suffer from one or more artefacts. Linear or non-adaptive interpolation methods deal with aliasing (e.g. jagged edges in the up scaling process), blurring and/or ringing effects. Well-known and popular linear interpolation methods are nearest neighbour, bilinear, bicubic and interpolation with higher order (piecewise) polynomials, B-splines, truncated or windowed sinc functions, etc. [7,10].

Non-linear or adaptive interpolation methods incorporate a priori knowledge about images. Dependent on this knowledge, the interpolation methods could be classified in different categories. The edge-directed based techniques follow a philosophy that no interpolation across the edges in the image is allowed or that interpolation has to be performed along the edges. This rule is employed for instance in Allebach's EDI method, Li's NEDI method and Muresan's AQua method [1,8,12]. The restoration-based techniques tackle unwanted interpolation artefacts like aliasing, blurring and ringing. Examples are methods based on isophote smoothing, level curve mapping and mathematical morphology [11,9,6]. Some other adaptive techniques exploit the self-similarity property of an image, e.g. iterated function systems [5,15]. Another class of adaptive interpolation

A. Campilho and M. Kamel (Eds.): ICIAR 2006, LNCS 4141, pp. 104–115, 2006.

methods is the training-based approach, which maps blocks of the low resolution image into predefined high resolution blocks [4].

Adaptive methods still suffer from artefacts: their results often look segmented, yield important visual degradation in fine textured areas or random pixels are created in smooth areas [9]. When we use very big enlargements (i.e. linear magnification factors of 8 and more), then all these artefacts become more visible and hence more annoying.

In §2 we will motivate the exploitation of repetitive structures in image interpolation. In §3 we give an image interpolation scheme for repetitive structures and in §4 we show some experiments and results of our proposed method compared to other interpolation techniques.

2 Repetitive Structures

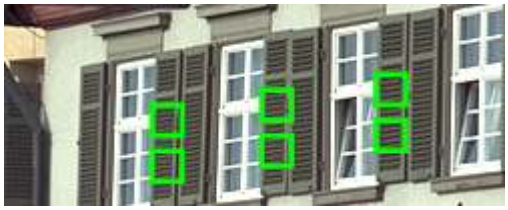
Fractal-based interpolation methods suppose that many things in nature possess fractalness, i.e. scale invariance [5]. This means that parts of the image repeat themselves on an ever-diminishing scale, hence the term self-similarity. This self-similarity property is exploited for image compression and interpolation by mapping the similar parts at different scales. Due to the recursive application of these mappings at the decoder stage, the notion of *iterated function systems* (IFS) is introduced.

Unlike IFS, we exploit the similarity of small patches in the same scale, i.e. spatially. In order to avoid confusion, we will use the term repetitivity. Another class of upscaling methods which also takes advantage of repetitivity, is called *super resolution* (SR) reconstruction. SR is a signal processing technique that obtains a HR image from multiple noisy and blurred low resolution images (e.g. from a video sequence). Contrary to image interpolation, SR uses multiple source images instead of a single source image. It is well known that SR produces superior results to conventional interpolation methods [3].

It is often assumed that true motion is needed for SR, however many registration methods do not yield true motion: their results are optimal to some proposed cost criterion, which are not necessarily equal to true motion. With this in mind, we can hypothetically assume that repetitive structures could serve as multiple noisy observations of the same structure (after proper registration). Results of our experiments in §4 will confirm that this hypothesis holds for real situations. The concept of repetitive structures has already succesfully been used for image denoising [2]. Besides repetitivity in texture, we can also find this recurrent property in other parts of the image, some examples are illustrated in figure 1.

Our method is found perfectly suitable to some applications: some examples are interpolation of scanned text images (availability of multiple repeated characters regardless of their font and the scan orientation) and gigantic satellite images (long roads and a lot of texture provide a huge amount of training data). A special application is SR in surveillance: when very few low resolution images are available, missing data could be filled up with the use of repetitive

structures. Because of the close relationship with SR, we can also denote our method as an intra-frame SR technique. In that way, surveillance applications could use a combination of inter- and intra-frame SR.



(a) Repetition in different objects.

(b) Repetition along edges.

(c) Repetition in uniform areas.

Fig. 1. Examples of repetitive structures in images

3 Proposed Reconstruction Scheme

We propose a simple interpolation method which exploits this repetitive behaviour. Our interpolation method is based on our camera model as shown in figure 2. Our scheme is quite straightforward and consists of three consecutive steps:

1. Matching repetitive structures and subpixel registration of these structures on the HR grid.
2. Robust data fusion.
3. Restoration (i.e. denoising and deblurring).

In the rest of this paper we will simply treat each R,G,B-channel of colour images separately. For an enhanced colour treatment we refer to [14]. The following sections will discuss each component carefully.

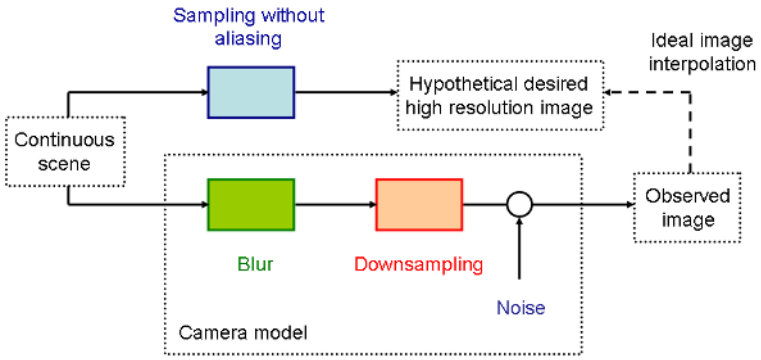


Fig. 2. Observation model of the image acquisition

3.1 Matching and Registration of Repetitive Structures

For the sake of simplicity, we define small rectangular windows B as basic structure elements. Two criterions are used in our algorithm to find matching windows or blocks across the whole image, namely the *zero-mean normalized cross correlation* (CC) and the *mean absolute differences* (MAD):

$$E_{\mathrm{CC}} = \frac{\sum_{\mathbf{x}\in\Omega}(B(m(\mathbf{x})) - \overline{B})(B_{\mathrm{ref}}(\mathbf{x}) - \overline{B_{\mathrm{ref}}})}{\sqrt{\sum_{\mathbf{x}\in\Omega}(B(m(\mathbf{x})) - \overline{B})^2 \sum_{\mathbf{x}\in\Omega}(B_{\mathrm{ref}}(\mathbf{x}) - \overline{B_{\mathrm{ref}}})^2}} \tag{1}$$

$$E_{\mathrm{MAD}} = \frac{1}{\kappa(\Omega)} \sum_{\mathbf{x}\in\Omega} |B(m(\mathbf{x})) - B_{\mathrm{ref}}(\mathbf{x})| \tag{2}$$

where Ω contains all the pixels of the reference window B_{ref} and $\kappa(\Omega)$ is the cardinality (i.e. the number of pixels) of Ω. $\overline{B}$ and $\overline{B_{\mathrm{ref}}}$ are denoted as the mean values of respectively B and B_{ref}. The transformation of the coordinates is characterized by the mapping function m. To simplify the registration problem and in particularly to save computation time, we assume that we are only dealing with pure translational motions of B. The main motive to use the CC and MAD criterions is because they are somewhat complementary: CC emphasizes the similarity of the structural or geometrical content of the windows, while MAD underlines the similarity of the luminance and colour information. A matched window is accepted if the two measures E_{CC} and E_{MAD} satisfy to the respective thresholds τ_{CC} and τ_{MAD}, more specifically: $E_{\mathrm{CC}} > \tau_{\mathrm{CC}}$ and $E_{\mathrm{MAD}} < \tau_{\mathrm{MAD}}$. Since we only want to have positive correlation, E_{CC} must lay between 0.0 and 1.0 and τ_{MAD} denotes the maximum mean absolute pixel difference between two windows. The choice of E_{CC} and τ_{MAD} depends heavily on the noise content of the image (e.g. due to additive noise or due to quantization noise in DCT-based compressed images such as JPEG): the higher the noise variance, the lower E_{CC}

and the higher τ_{MAD} must be chosen in order to match the same repetitive structures. Our implementation uses an simple exhaustive search in order to find the matching windows, but more intelligent (pattern-based) search algorithms could reduce the computation time enormously.

Common ways to achieve subpixel registration in the spatial domain, is to interpolate either the image data or the correlation data. In order to save computation time we only resample the reference window B_{ref} on a higher resolution. In this way we represent the downsampling operator in the camera model in figure 2 as a simple decimation operator as illustrated in figure 3. We estimate the subpixel shifts by minimizing the MAD criterion in equation 2. As a simplification of the optimization problem, we use the HR grid as the discrete search space for the subpixel shifts. After the subpixel registration, the pixel values of B are mapped onto the HR grid.

Most existing techniques use linear methods to resample B_{ref}. However, these interpolation methods typically suffer from blurring, staircasing and/or ringing. These artefacts not only degrade the visual quality but also affect the registration accuracy. Therefore we adopt a fast non-linear restoration-based interpolation technique based on level curve mapping, which suffers less from these artefacts [9].

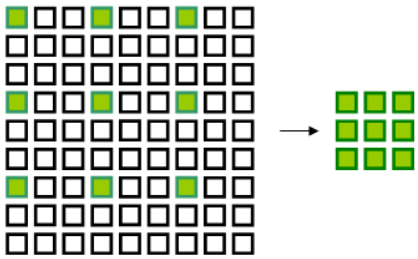


Fig. 3. The $3:1$ decimation operator maps a $3M \times 3N$ image to an $M \times N$ image

3.2 Data Fusion

In this step we determine an initial pixel value for every pixel of the HR grid. In the previous registration step we already obtained zero or more observations for these pixels.

Several observations are available. Starting from the maximum likelihood principle, it can be shown that minimizing the norm of the residuals is equivalent to median estimation [3]. A residual is the difference between an observed pixel value and the predicted pixel value. The median is very robust to outliers, such as noise and errors due to misregistration. For this reason we adopt the median estimate of all observations for every pixel in the HR grid for which we have at least one observation.

No observation is available. These unknown pixel values are initialised with the values of the interpolated reference windows B_{ref}. We do not need additional computations since this image is already constructed for the registration step (see §3.1).

Restoring the original pixels. These original pixel values are simply mapped back on the HR grid, because we know that these pixel values contain the least noise.

In a nutshell, the HR grid consists of three classes: the original pixels (OR), the unknown pixels (UN) and the fused pixels (FU). The last mentioned class will provide the extra information which gives us better interpolation results compared to conventional upscaling techniques.

3.3 Denoising and Deblurring

We assume that the blur in the camera model in figure 2 is characterized by a shift-invariant *point spread function* (PSF). The inverse problem becomes highly unstable in the presence of noise. This can be solved by imposing some prior knowledge about the image. Typically we will try to force spatial smoothness in the desired HR solution. This is usually implemented as a penalty factor in the generalized minimization cost function:

$$\widehat{I}(\mathbf{x}) = \arg\min_{I(\mathbf{x})} \left[\rho_{\text{R}}(I(\mathbf{x})) + \lambda \rho_{\text{D}}(H * I(\mathbf{x}) - I(\mathbf{x}, 0))\right] \tag{3}$$

where H denotes the PSF-kernel (typically Gaussian blur, which is characterized by its standard deviation σ_{blur}) and λ is the regularization parameter between the two terms, respectively called the regularization term ρ_{R} and the data fidelity term ρ_{D}. Image $I(\mathbf{x}, 0)$ is the HR image obtained in §3.2.

The minimization problem of equation 3 could be transformed to the following partial differential equation (PDE) which produces iteratively diffused images $I(\mathbf{x}, t)$ starting from the inialisation image $I(\mathbf{x}, 0)$:

$$\frac{\partial I(\mathbf{x}, t)}{\partial t} = \rho'_{\text{R}}(I(\mathbf{x}, t)) + \lambda \rho'_{\text{D}}(H * I(\mathbf{x}, t) - I(\mathbf{x}, 0)) \tag{4}$$

where the chain rule is applied on both ρ-terms: $\rho'(I(t)) = \frac{\partial \rho(I(t))}{\partial I(t)} \cdot \frac{\partial I(t)}{\partial t}$.

The use of the so-called edge-stopping functions in the regularization term is very popular, because it suppresses the noise better while retaining important edge information [13]. Therefore we apply one of the most successful edge-preserving regularization terms proposed for image denoising, namely the total variation (TV):

$$\rho_{\text{R}}(I(\mathbf{x}, t)) = |\nabla I(\mathbf{x}, t)| \tag{5}$$

Since the interpolation of repetitive structures is closely related to super resolution problems, we could assume that the outliers (due to the inaccuracy of the image registration, blur, additive noise and other kinds of error which are

not explicitly modeled in the fused images) are better modeled by a Laplacian probability density function (PDF) rather than a Gaussian PDF according to [3]. The maximum likelihood estimate of data in the presence of Laplacian noise is obtained through the L_1-norm minimization. That is why we will use the L_1-norm function for the data fidelity term:

$$\rho_{\mathrm{D}}(H * I(\mathbf{x}, t) - I(\mathbf{x}, 0)) = |H * I(\mathbf{x}, t) - I(\mathbf{x}, 0)| \tag{6}$$

These ρ-functions are very easy to implement and are very computationally efficient. Other robust ρ-functions could also be employed [13]. We now adapt the PDE in equation 4 locally to the several classes of pixels on the HR grid:

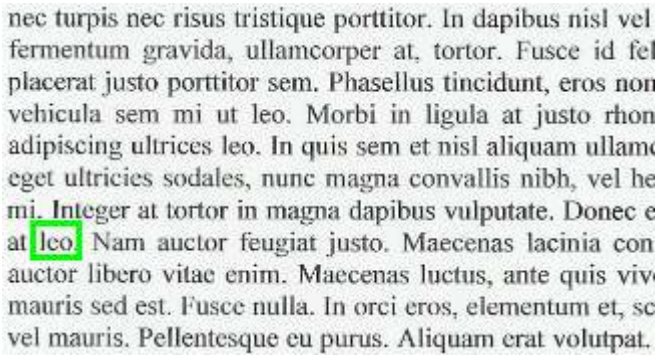
- Class OR: since these pixels contain the least noise as discussed in §3.2, very little regularization has to be applied. This means that these pixels depend mainly on the data fidelity term and thus λ is set to λ_{MAX}.
- Class UN: these pixels are most likely noise and depend only on the regularization term ($\lambda = 0$).
- Class FU: these pixels contain noise and relevant information. The more observations we have, the more robust the initial estimation will be and hence the lesser noise the pixel will contain. That is why we apply equation 4 with λ proportional to the number of available observations α: in our implementation we use $\lambda = \min(\lambda_{\mathrm{MAX}}, \frac{\alpha}{10}\lambda_{\mathrm{MAX}})$. Where 10 is the buffer length of the observations.

Finally the PDE of equation 4 is iteratively applied to update the blurred and noisy image in the restoration process.

4 Experiments and Results

As a first realistic experiment we have both printed and scanned one A4 paper containing the *Lorem ipsum* text with the HP PSC 2175 machine at 75 dpi as shown in figure 4. The Lorem ipsum text is very popular as default model text, but additionally it has a more-or-less normal distribution of letters [16].

The basic structure element was chosen to be a 18×12 rectangular window. We have enlarged the *region of interest* (ROI) with a linear magnification factor of 8 and compared our result with the popular cubic B-spline interpolation technique in figure 4. The parameters for our method are $\sigma_{\mathrm{blur}} = 4.0$, $\tau_{\mathrm{CC}} = 0.6$, $\tau_{\mathrm{MAD}} = 40.0$, $\lambda_{\mathrm{MAX}} = 100$ and 100 iterations for the restoration process. The parameter selection was based on trial and error, i.e. to produce the visually most appealing results. Also the detection rate of the word *leo* was perfectly, this means a recall and a precision of both 100% (on 13 samples). We can clearly see in figure 4 that our method outperforms traditional interpolation techniques: the letters are much better readable and reconstructed, noise and JPEG-artefacts are heavily reduced and much less blurring, staircasing and ringing artefacts are created.



(a)

(b) (c) (d)

Fig. 4. Interpolation of the Lorem ipsum text (8× enlargment): (a) a part of the original scanned text image (at 75 dpi), (b) nearest neighbour interpolation of the ROI of (a), (c) cubic B-spline interpolation, (d) our proposed method (partition of the HR grid: 1.5625% (OR), 14.0625% (FU) and 84.375% (UN))

(a) (b) (c)

Fig. 5. Experiment with the scanned Lorem ipsum text (2× enlargment): (a) original scanned text at 150 dpi, (b) cubic B-spline interpolation of the ROI of figure 4a at 75 dpi, (c) our proposed method applied on the ROI of figure 4a at 75 dpi (partition of the HR grid: 25% (OR), 50% (FU) and 25% (UN))

We have scanned the same text at 150 dpi as shown in figure 5. If we visually inspect the 2× linear enlargment of our method and the cubic B-spline interpolation to this ground truth data, we can conclude that our method manage to reconstruct the characters much better. The images in figure 5 are 4× enlarged with the nearest neighbour interpolation in order to achieve a better visibility.

As a second experiment we have interpolated a real image. In figure 6 we show a part of the original image with a 8× nearest neighbour interpolation of the

(a) A part of the original image.

(b) Nearest neighbour of the region of interest of (a).

Fig. 6. An example with a real image

region of interest. As basic structure elements we use 5×5 windows and we have enlarged the image with a linear magnification factor of 8. For the matching step we have used the following threshold parameters: $\tau_{\text{CC}} = 0.9$ and $\tau_{\text{MAD}} = 9.0$. For the denoising and deblurring we have applied the PDE within 100 iterations and with $\sigma_{\text{blur}} = 4.0$ and $\lambda_{\text{MAX}} = 10$. With these parameters we obtained the following partition of the different classes: 1.5625% (OR), 30.9091% (FU) and 67.5284% (UN).

(a) Cubic B-spline.

(b) IFS [15].

(c) Our proposed method.

Fig. 7. Results of several interpolation methods for the ROI of figure 6a

Figure 7 shows our result compared to a linear interpolation (cubic B-spline) and an IFS method (obtained from commercial software [15]). Significant improvements in visual quality can be noticed in our method: there is a very good reconstruction of the edges and our result contains much less annoying artefacts. The result produced with our method is also better denoised while important edges are preserved.

5 Conclusion

In this paper we have presented a novel interpolation scheme based on the repetitive character of the image. Exploiting repetitivity brings more information at our disposal, which leads to much better estimates of the unknown pixel values. Results show the effectiveness of our proposed interpolation technique and its superiority at very large magnifications to other interpolation methods: edges are reconstructed well and artefacts are heavily reduced. In special applications with text images, we can achieve excellent results: characters could be made much better readable again. This could be very advantageous for optical character recognition (OCR) applications or when the image resolution can not be improved at the sensor because of technological limitations or because of high costs.

References

1. Allebach, J., Wong, P.W.: Edge-Directed Interpolation. Proc. of IEEE International Conference on Image Processing (1996) 707–710
2. Buades, A., Coll, B., Morel, J.: Image Denoising By Non-Local Averaging. Proc. of IEEE International Conference of Acoustics, Speech end Signal Processing **2** (2005) 25–28
3. Farsiu, S., Robinson, M.D., Elad, M., Milanfar, P.: Fast and Robust Multiframe Super Resolution. IEEE Trans. on Image Processing **13** (2004) 1327–1344
4. Freeman, W.T., Jones, T.R., Pasztor, E.C.: Example-Based Super-Resolution. IEEE Computer Graphics and Applications **22** (2002) 56–65
5. Honda, H., Haseyama, M., Kitajima, H.: Fractal Interpolation For Natural Images. Proc. of IEEE International Conference of Image Processing **3** (1999) 657-661
6. Ledda, A., Luong, H.Q., Philips, W., De Witte, V., Kerre, E.E.: Image Interpolation Using Mathematical Morphology. Proc. of 2nd IEEE International Conference On Document Image Analysis For Libraries (2006) (to appear)
7. Lehmann, T., Gönner, C., Spitzer, K.: Survey: Interpolations Methods In Medical Image Processing. IEEE Trans. on Medical Imaging **18** (1999) 1049–1075
8. Li, X., Orchard, M.T.: New Edge-Directed Interpolation. IEEE Trans. on Image Processing **10** (2001) 1521–1527
9. Luong, H.Q., De Smet, P., Philips, W.: Image Interpolation Using Constrained Adaptive Contrast Enhancement Techniques. Proc. of IEEE International Conference of Image Processing **2** (2005) 998–1001
10. Meijering, E.H.W., Niessen, W.J., Viergever, M.A.: Quantitative Evaluation Of Convolution-Based Methods For Medical Image Interpolation. Medical Image Analysis **5** (2001) 111–126

11. Morse, B.S., Schwartzwald, D.: Isophote-Based Interpolation. Proc. of IEEE International Conference on Image Processing (1998) 227–231
12. Muresan, D.: Fast Edge Directed Polynomial Interpolation. Proc. of IEEE International Conference of Image Processing **2** (2005) 990–993
13. Pižurica, A., Vanhamel, I., Sahli, H., Philips, W., Katartzis, A.: A Bayesian Approach To Nonlinear Diffusion Based On A Laplacian Prior For Ideal Image Gradient. Proc. of IEEE Workshop On Statistical Signal Processing (2005)
14. Vrhel, M.: Color Image Resolution Conversion IEEE Trans. on Image Processing **14** (2005) 328–333
15. Genuine Fractals 4: http://www.ononesoftware.com
16. Lorem Ipsum generator: http://www.lipsum.com