An Application of the Least Squares Plane Fitting Interpolation Process to Image Reconstruction and Enhancement

Gabriel Scarmana, Australia

Key words: Image enhancement, Interpolation, Least squares.

SUMMARY

This work applies a least squares plane fitting (LSP) method as an alternative way of interpolating irregularly spaced pixel intensity values that are suitable for image reconstruction of a static scene via super-resolution (SR). SR is a term used within the computer vision and image processing community to describe the process of reconstructing a high resolution image from a sequence of several shifted images covering the same scene.

The accuracy attainable by this process is estimated via tests where the simulation parameters are controlled and where the reconstructed high resolution image can be compared with its original. In these tests the original image is scanned randomly so as to create a sequence of low-resolution and JPEG compressed shifted images. The comparison is based on the r.m.s.e. of the differences between the reconstructed image and the original.

An Application of the Least Squares Plane Fitting Interpolation Process to Image Reconstruction and Enhancement

Gabriel Scarmana, Australia

1. INTRODUCTION

The spatial resolution that represents the number of pixels per unit area in an image is the major aspect in determining image quality. With the development of image processing applications there has been a great requirement for high-resolution (HR) images because HR images not only give the viewer a more pleasing picture, but also offer additional details that are important for the analysis in many applications (Russ, 2011). These applications include, but are not limited to, enhancing the imagery for remote sensing, aerial and satellite imagery, medical imaging, surveillance systems and forensic sciences.

HR images mainly depend on sensor manufacturing technology that attempts to increase the number of pixels per unit area by reducing the pixel size. However, the cost for high-precision optics and sensors may be unsuitable in some applications. In addition, there exist several optical limitations on the pixel size that make it imprudent to further reduce the pixel size beyond those limitations. A major limitation of reducing the pixel size is that pixel sensitivity is significantly reduced as the pixel size is reduced (Gonzalez and Woods, 2009).

To avert these drawbacks, a resolution enhancement approach may be implemented. A particular process that has received much attention in the last two decade is super-resolution (SR) image reconstruction or simply multi-frame resolution enhancement. The basic idea behind SR is the fusion or registration, at a sub-pixel level, of a sequence of low-resolution noisy blurred images depicting the same scene so as to produce a HR image or sequence. A comprehensive review on SR can be found in Nasrollahi and Moeslund (2014).

Images generated by present imaging devices (i.e. digital/video cameras and smart phones) are generally compressed in a lossy manner, such as by the JPEG protocol so as to reduce the storage requirements and speed of transmission over digital links. Lossy compression means that data is lost during compression so the quality after decoding is less than the original picture. Lossy compression protocols also introduce several distortions which can complicate the super-resolution problem. For example, most compression algorithms divide the original image into blocks which are processed independently, thus creating problems of continuity between blocks after decompression. In general, SR reconstruction can be divided into three distinct steps (Baker and Kanade, 2002):

1. Estimation of the motion fields (or shifts) among the different low-resolution images at a sub-pixel level (sometimes referred to as image-to-image registration or image matching),
2. Projecting or mapping the pixels of the low-resolution images onto a higher resolution grid using the shifts detected by the registration process, and,
3. Interpolating or solving sets of equations derived from the geometric relationships existing between low-resolution pixels and high-resolution pixels.

With regard to item 3 above, the data to be interpolated is randomly distributed, because of the random, sub-pixel shifts of each frame. There exists a number of interpolation methods suited to carry out this process. Nearest neighbour interpolation provides a simple and fast solution to this problem but it is less accurate than other methods and not capable to suppress noise. Popular interpolation techniques used in image processing such as bilinear interpolation and cubic convolution are difficult to implement especially when dealing with non-uniformly sampled data.

The alternative process considered here relates to the LSP method of interpolation (see section 3). An important feature of LSP (Gilman et al, 2008) is that it is of fast implementation and gives an estimation of the error at each interpolated point, thus providing a measure of confidence in the resulting enhanced image (Li and Heap, 2008). LSP is ideal for real-time applications as it does not involve complex computations as per standard Kriging or IDW (Inverse Distance Weighted) interpolation processes. LSP also works particularly well in the presence of a dense number of sparse/close points contributing to the computation of the point of interest, and creates sharper images than Splines variants methods which usually produce overly smoothed images (Patil et al., 2007).

2. IMAGE MATCHING: SUB-PIXEL MOTION ESTIMATION

There exists a number of methods for determining the sub-pixel shifts (i.e. in x and y and rotations) existing between two images depicting the same scene (LeMoigne et al, 2011). The method used here was devised by Sicairos et al. (2008) and is based on DCT (Discrete Fourier Transforms) and normalized cross-correlation registration techniques.

This registration technique was preferred because it allows images to be registered or matched without using control points in the registration procedure and, depending on the image entropy, can achieve accuracies of up to 0.01 of a pixel. For two given images *A* and *B* the registration yields the normalized root-mean-squared error (n.r.m.s.e.) between *A* and *B*, their global phase angle and the row (x) and column (y) shifts between the two images respectively. In this work, all registration results for n.r.m.s.e. values less that 0.1 was be considered suitable. The best image features required for an effective image to image registration are gentle changes between two areas of grey values because these areas are generally unchanged by aliasing (Schowengerdt, 2007). Hence, before the low-resolution images of the same scene are registered, they are blurred (Russ, 2011). The purpose of the blurring filter is to smooth or reduce:

1. sharp edges and small details
2. abrupt changes of pixel intensity values and
3. the distortions created by the compression process.

The registration process determines the x- and y-shifts between any two images. By reiterating the procedure for an additional reference image, a second estimation for the relative positions is made. The replication of this process for all images in the sequence defines a better estimate of the relative shifts (Hardie, 2007). The numerical measure that was used to determine the 'best' possible value of the image shifts was the median as the median would be a tangible computed value and not an averaged one. In addition, the median is not biased by outliers and/or skewed data (Gelfand, 2010).

3. IMAGE RECONSTRUCTION

Once all the low-resolution images have been registered to a sub-pixel level, they are projected or mapped onto a uniformly spaced high-resolution grid. In the idealized super-resolution set-up of Figure 1 the images (b)-(d) are taken with sub-pixel shifts of half a pixel in the horizontal, vertical and diagonal directions in relation to image (a). Their pixels can then be inserted in the correct location so as to generate a higher resolution image with a magnification factor (p) equal to 4, that is, the image contains 4 times more pixels than any of the low-resolution images.

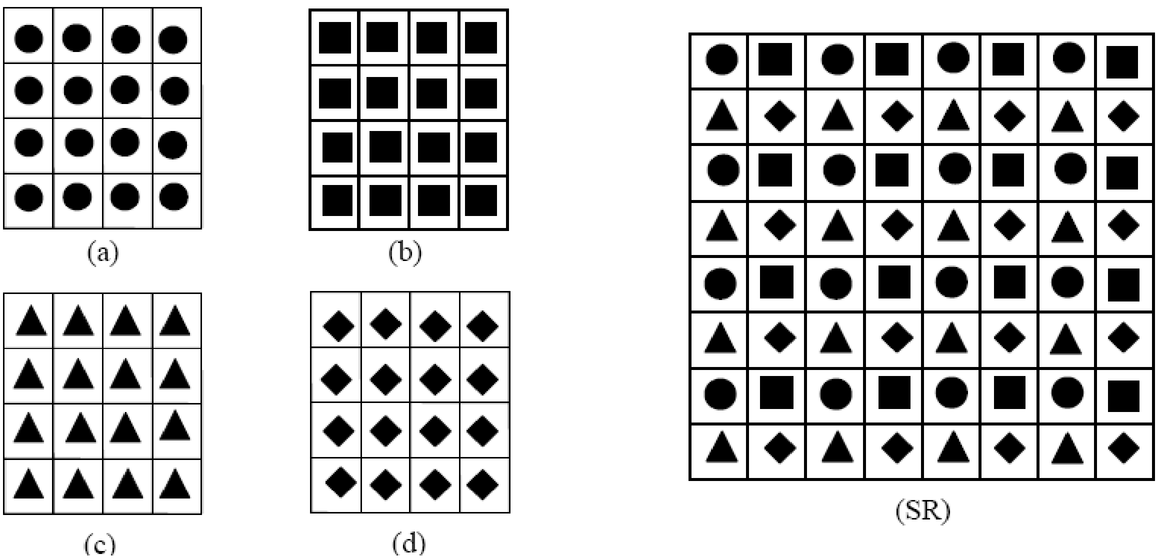

Figure 1 - An idealized super-resolution setup. Images (b)-(d) are taken with sub-pixel shifts of half a pixel in the horizontal, vertical and diagonal directions in relation to image (a). Their pixels can then be interleaved to generate a double resolution image in each coordinate direction.

As previously cited, the estimation or interpolation process which determines the pixel brightness which would exist at the intersections of an ordered grid using randomly spaced pixel locations (representing the pixels of the low-resolution images) is carried out using LSP. Traditionally, the method of least squares regression allows to find a two-variable linear equation y = mx + b that provides the best possible fit for the data points. In a two dimensional sense the best fit is assumed in terms of minimizing the squared vertical errors, that is finding the values of m and b that minimize the function:

$$F(m, b) = \sum(y_i - mx_i - b)^2 \tag{1}$$

The solution can be found with matrices because the system $\partial F/\partial m = 0$ and $\partial F/\partial b = 0$ is a linear system of equations. In multiple linear regression, the process is extended to find the equation of a plane z = ax + by + c that minimizes the vertical distances between the points (x_i, y_i, z_i) and the plane. To achieve this, the values of a, b, and c that minimises the equation below must be established:

$$G(a, b, c) = \sum(z_i - ax_i - by_i - c)^2 \tag{2}$$

by solving the system $\partial G/\partial a = 0$, $\partial G/\partial b = 0$, and $\partial G/\partial c = 0$. As the system $\partial G/\partial a = 0$, $\partial G/\partial b = 0$, and $\partial G/\partial c = 0$ is linear, it is possible to solve it in matrix form. The matrix equation for a, b and c is:

$$\begin{bmatrix} \sum x_i^2 & \sum x_i y_i & \sum x_i \\ \sum x_i y_i & \sum y_i^2 & \sum y_i \\ \sum x_i & \sum y_i & n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum x_i z_i \\ \sum y_i z_i \\ \sum z_i \end{bmatrix} \tag{3}$$

When the matrix on the left is invertible (determinant not equal to zero) then there is a unique solution set (a, b, c).

Based on the above explanation and on Figure 2 below, *n* points in the vicinity of the new sample position are selected. A least squares plane is fit through those *n* points and the value of this plane at the required position is used as the new sample value. The number of data points (i.e. 12 scattered pixel values in Figure 2) used to compute the new position and the new pixel value (the black triangle in Figure 2) is fitted to adjust between noise suppression and resolution improvement.

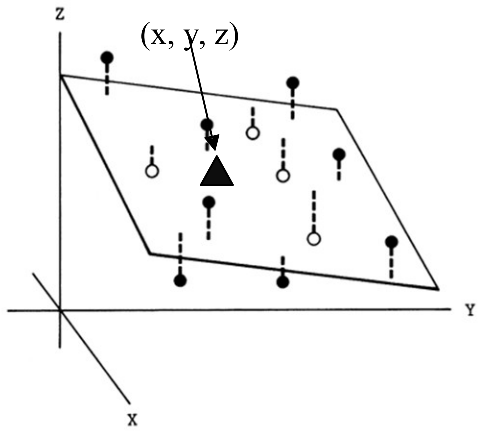

Fig. 2 – The **z** value (pixel intensity value) of the black triangle at a required grid intersection (**x, y**) coordinates is the value calculated on the least square plane determined by twelve adjacent pixel intensity values.

In order to improve the computational time it is possible to set bounds to the *n* points that contribute to the calculation of the interpolated value. For the particular application developed in this work, it was estimated that the most appropriate search radius was 10 surrounding pixels closest to the point of interest. This was determined based on testing this estimation process on the reconstruction of a significant number of high-resolution grey-scale images of varied entropy values.

4. TESTS AND RESULTS

A series of tests was conducted so as to establish the achievable accuracy of the selected image registration routine and to validate the reconstruction of a higher resolution image. Firstly, the performance of the image registration was measured by estimating the known shifts between two synthetic images. This was carried out for ten different image pairs and each time 50 random shifts were computed per pair of images.

By way of consistency, the size of the images was in all cases 400^2. The shifts were estimated once with the use of a Gaussian blurring filter (18 pixel radius) and once without it. Table 1 shows the results in terms of the absolute errors. The results indicate that the registration process performs better by adding a Gaussian filter. In this instance, registering with a Gaussian filter produced registrations almost 42 times more exact.

In a second test, the mapping or reconstruction process described in the previous section was tested by constructing a high-resolution image using 40 (90^2 pixels) low-resolution images obtained by randomly scanning an image of a gray scale scene. The process used by the scanner to create the low-resolution (or down-sampled) images was unknown, and so were the relative shifts existing among the low-resolution images.

Filter	Mean (x-axis)	Variance (x -axis)	Mean (y-axis)	Variance (y-axis)
Filterless	2.69 x 10-2	2.77 x 10-4	2.56 x 10-2	2.03 x 10-4
Gaussian filter	6.13 x 10-4	1.66 x 10-7	3.88 x 10-4	1.41 x 10-7

Table1 – Accuracy results for the registration process (with and without a Guassian filter) used to determine sub-pixel shifts between two given images depicting the same scene.

In this instance the 'true' image of the scene shown in Figure 3 was known prior to the enhancement as it was scanned to be (450^2 pixels) and thus the accuracy of the enhancement and some of the factors that influence the process could be investigated and quantified. The aim was to recover a high resolution image with five times more pixels in each coordinate direction (i.e. x and y) than any of the low-resolution images so as to be comparable with the true image.

Also, each of the 40 low-resolution images was jpeg compressed using the same compression ratio (10:1). No rotations were applied when scanning the low-resolution images. Correlation obviously exists amongst a sensor and orientation parameters such as tilts, rotations and affinity/obliquity of the sensor (as those found in digital and video cameras) and, in a controlled experiment where the intention was to ascertain a resolution enhancement process in itself, it was considered unwise to include such intricacy.

Figure 3(a) shows the quality of the enhancement by combining via LSP the 40 low-resolution images using the sub-pixel shift calculated through the registration technique of section 2. The accuracy of the reconstructed image was determined by subtracting it from the true initial image. The difference produced a grey scale R.M.S.E. (Root Mean Square Error) of 12 pixel intensity values with maximum and minimum errors ranging between -16 and +18 pixel intensity values. In this specific example, the final composite in Figure 3(b) was obtained using only 25 low-resolution images as no further improvement was observed in the R.M.S.E. with more than this number of images.

In order to subtract the higher resolution composite from the original 'true' scanned image, the enhanced composite was aligned to said original image by way of the same registration process given in section 2. Although this experiment relates to a sequence of grey scale images, the same process can be applied when using colour. Colour images can be considered as three separate images containing red, green and blue channels (RGB). Each of these channels can be enhanced independently and then fused to produce a colour image with enhanced resolution.

(a) (b)

Figure 3 – A sequence of 40 compressed low-resolution images (one example shown on the left) was used to reconstruct the higher resolution image shown in (b).

5. NUMBER OF LOW-RESOLUTION IMAGES REQUIRED

When image noise is low and knowledge of sub-pixel shift values is accurate, resolution improvement is limited primarily by the number of low-resolution shifted images. In several controlled tests conducted by the author and similar to the one presented in section 4 the number of images required to solve for a given magnification factor (p) could be approximated to $4p^2$. By way of example, it would require 16 images (50^2 pixels) to reconstruct a composite of 200^2 pixels (p=4). However, at noise levels characteristics of conventional video and digital image sensors, additional low resolution images may be needed so as to compensate for noise reduction. Also, in order to minimise the influence of noise it is important that the distribution of the shifts between the low- resolution images be as complete as possible.

The reconstruction of a higher resolution image with less than the minimum quantity of coarse images is feasible, but it should not be estimated to always accomplish optimal results, especially for higher magnification factors. High magnification factors require a greater numbers of low resolution images, meaning that these low-resolution images must be relatively close to one another, that is, relatively small shifts. The accuracy of detecting those offsets will clearly influence the accuracy of the enhanced image composite as the ambiguity in an offset's calculation may be of the same magnitude as the shift itself (Robinson and Milanfar, 2006).

6. CONCLUSIONS

In this paper, the problem of restoring a high-resolution image from a sequence of low-resolution and compressed images was investigated in terms of applying a Least Squares Plane (LSP) fitting interpolation process. The registration or matching methodology and subsequent use of the enhancement process may lead to a general approach to the problem of generating a higher resolution image from compressed sequences of slightly shifted/under-sampled images. The application of the enhancement process has been demonstrated in tests which demonstrated that a significant gain in spatial resolution can be obtained. Tests also showed that the LSP method of interpolation produced suitable results with regard to processing time, visual perspective and accuracy of the image reconstruction.

Improvements to the proposed process are being undertaken to increase the accuracy achievable for larger image magnification factors (i.e. greater than 5), while adapting this device independent process to a generalized scheme whereby both sensor and object are unstable, the illumination is non-uniform and the low-resolution images are of different spatial resolutions.

REFERENCES

Baker S. and Kanade T. 2002. Limits on Super-Resolution and How to Break Them. IEEE Transactions on Pattern Analysis and Machine Intelligence. Volume 24. Issue 9.

Bailey D. and Marsland S. 2008. Interpolation Models for Image Super-resolution. 4th IEEE International Symposium on Electronic Design, Test & Applications. 23-25 January. Hong Kong, SAR, China.

Eastman R., Nataniahu N. and LeMoigne J. 2011. Image Registration for Remote Sensing. Cambridge University Press.

Gelfand A. E. 2010. Handbook of Spatial Statistics. Publisher: Boca Raton, CRC, 607 pages. Gilman

Gonzalez R. E., Woods S. and Eddins L. 2009. Digital Image Processing Using Matlab. Second Edition. Gatesmark Publishing. A Division of Gatesmark LLC. 782 pages.

Guizar-Sicairos M., Thuman S. and Fienup J R. 2008. Efficient sub-pixel image registration algorithms. Optics Letters, Vol. 33, Issue 2, pp. 156-158.

Hardie R. 2007. A Fast Image Super-Resolution Algorithm Using an Adaptive Wiener Filter. IEEE Transaction on Image Processing, Vol. 16, No. 12, Dec. pp 1-5.

Kamal Nasrollahi and Thomas B. Moeslund. 2014. Super-resolution: A comprehensive survey Machine Vision & Applications Vo. 26, Issue 6.pp 1423-1468.

Li J. and Heap D. A. 2008. A review of Spatial Interpolation Methods for Environmental Scientists. Record 2008/23. Geocat No. 68229. Geosciences Australia. Australian Government.

Patil V.and Bormane D. 2007. Interpolation for Super-Resolution Imaging. Innovations and Advanced Techniques in Computer and Information Sciences and Engineering. pp. 483-489. Springer Netherlands.

Robinson M. D. and Milanfar P. 2006. Statistical performance analysis of super-resolution, IEEE Transactions on Image Processing, vol. 15, no. 6, pp. 1413–1428.

Russ C. J. 2011. The Image Processing Handbook. Publisher: CRC. Boca Raton. 885 pages.

Schowengerdt R. A. 2007. Remote Sensing: Models and Methods for Image Processing. Elsevier. 515 pages.

BIOGRAPHICAL NOTES

Gabriel Scarmana completed his PhD in 2004 from the University of Newcastle (Australia) from the school Civil, Surveying and Environmental Engineering. He maintains an active research interest in the areas of image processing and mobile mapping technology.

CONTACTS

Gabriel Scarmana

Institution

University of Southern Queensland

Toowoomba

AUSTRALIA