2010 Fifth IEEE International Symposium on Electronic Design, Test & Applications

Least-squares optimal interpolation for fast image super-resolution

Andrew Gilman, Donald G. Bailey and Stephen Marsland
School of Engineering and Advanced Technology
Massey University
Palmerston North, New Zealand
e-mail: a.gilman@massey.ac.nz, d.g.bailey@massey.ac.nz, s.r.marsland@massey.ac.nz

***Abstract*—Image super-resolution is generally regarded as consisting of three steps – image registration, fusion, and deblurring. This paper presents a novel technique for resampling a non-uniformly sampled image onto a uniform grid that can be used for fusion of translated input images. The proposed method can be very fast, as it can be implemented as a finite impulse response filter of low order (10th order results in good performance). The technique is based on optimising the resampling filter coefficients using a simple image model in a least squares fashion. The method is tested experimentally on a range of images and shown to have similar results to that of a least-squares optimal filter. Further experimental comparisons are made against a number of methods commonly used in image super-resolution that show that the proposed method is superior to these.**

Keywords-super-resolution; image reconstruction; image fusion; least squares; non-uniform interpolation; resampling

I. INTRODUCTION

Image super-resolution is the process of combining multiple images of the same scene, generally captured from different viewpoints, to reconstruct the higher frequency information that has been lost due to aliasing. It is generally regarded as consisting of three steps – image registration, fusion, and deblurring [1]. This paper presents a novel technique for resampling a non-uniformly sampled image onto a uniform grid that can be used for fusion of translated input images once relative translations have been estimated by means of image registration. The method can also be generalised and utilised for any problem that requires resampling in two (or more) dimensions.

Image interpolation is a relatively well established part of image processing; however, most of the traditional image interpolation methods are specifically designed to interpolate uniformly sampled data [2]. Adaptation of these methods to work with non-uniformly sampled data, as required in image super-resolution is somewhat problematic. As an alternative, researchers have either employed already existing non-uniform interpolation techniques (see [3, 4] for examples) or formulated the problem as a discrete inverse problem, where traditional interpolation can be applied (see [5-7] for examples). The approach we propose in this article is of the non-uniform interpolation kind.

In previous work [8] we investigated the concept of a least-squares optimal resampling filter. We showed that filters optimised on one image can interpolate another with fairly good results. We also showed that a filter optimised on a synthetic image consisting of just a black disk on a white background works well at resampling normal images. This led to the hypothesis that step-edges are a major contributor to the optimisation of filter weights. This work was extended further in [9]: considering that to a first approximation, most images can be assumed to be piece-wise constant (consisting of relatively flat areas separated by step-edges), we optimised the coefficients of a 1D filter on an image consisting of randomly positioned step-edges. The use of this step-edge model resulted in a piece-wise cubic interpolation kernel. Visual comparison of this kernel to optimal kernels of three normal images showed high degrees of similarity. However, to simplify the analysis, only the one-dimensional resampling problem was considered. In this paper, we extend our previous work to two dimensions, and propose a new two-dimensional image model to optimise 2D resampling filters.

II. OPTIMAL INTERPOLATION

The problem considered here is that of resampling a non-uniformly sampled image formed by "stacking" a number of low-resolution input images, which are related to each other by pure translation. The new sampling grid is assumed to be of the same sampling period as the input images, but, there is nothing to stop us from repeating this resampling process with a shifted output sampling grid. By shifting this output grid by half a pixel in the x direction, half a pixel in the y direction, and both together and interleaving the images resampled on each of those grids (in the right order) will result in an output image with doubled sampling rate. The only constraint on constructing the final output image in this way is that the sampling rate can only be increased by an integer factor.

We will re-use the notation from [9], extending it to two dimensions by replacing scalar variables x, q, and p with two-element vectors $\mathbf{x}$, $\mathbf{q}$, and $\mathbf{p}$. The resampled image $f_O(\mathbf{q})$ is then given from the discrete convolution of the non-uniformly sampled input image, f_I, with the corresponding interpolation kernel, H:

$$\hat{f}_O(\mathbf{q}) = \sum_{\mathbf{p}\in P_{\mathbf{q},w}} f_I(\mathbf{p}) H(\mathbf{q}-\mathbf{p}), \quad \mathbf{q}\in Q \ , \tag{1}$$

978-0-7695-3978-2/10 $26.00 © 2010 IEEE
DOI 10.1109/DELTA.2010.59

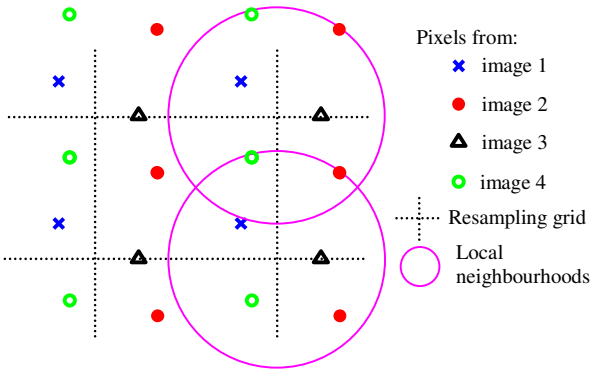

Figure 1. "Stacked" non-uniform image showing pixels originating from four individual input images with an output sampling grid overlaid on top.

where P is the set of discrete locations of the input samples, Q is the set of locations that correspond to the output sampling grid, and $P_{\mathbf{q},w}$ defines the region of support used for the interpolation – the subset of P that includes all input pixels within the distance w of output grid location $\mathbf{q}$:

$$P_{\mathbf{q},w} \triangleq \{\mathbf{p} \mid \mathbf{p} \in P, \|\mathbf{p}-\mathbf{q}\| < w\}, \tag{2}$$

where $\|\cdot\|$ denotes the Euclidean distance.

We exploit the inherent periodicity in the non-uniform input image. For any point on the output resampling grid, the locations of input pixels within some small neighbourhood around it are exactly the same (as demonstrated in Fig. 1). Consequently, the weights associated with the input pixels in a neighbourhood are the same for each neighbourhood (since the weights are a function of the relative position of the input pixel to the output pixel, i.e. $\mathbf{q}-\mathbf{p}$).

The collection of output pixels and the neighbourhoods of input pixels around them offer multiple observations and the set of weights can be optimised in a least squares fashion to minimise the sum of square errors across all observations:

$$\min_{H} \sum_{\mathbf{q}\in Q} \left(f_O(\mathbf{q}) - \sum_{\mathbf{p}\in P_{\mathbf{q},w}} f_I(\mathbf{p}) H(\mathbf{q}-\mathbf{p}) \right)^2, \tag{3}$$

subject to the "partition of unity" constraint

$$\sum_{\mathbf{p}\in P_{\mathbf{q},w}} H(\mathbf{q}-\mathbf{p}) = 1 \quad \forall \mathbf{q} \in Q \tag{4}$$

to make sure that there is no overall gain in flat regions of the image [2, 10].

Solving this minimisation problem produces a set of weights that are least squares optimal and result in minimum mean square error across the whole image. However, this is impractical for obvious reasons – it requires knowing the ideal output $f_O(\mathbf{q})$ to derive the least squares optimal approximation of the output $\hat{f}_O(\mathbf{q})$. Instead, the weights can be optimised on an image model and then used to resample the input image.

Following the approach of [9], we redefine the origin for the filter to be at an integer location on the output grid by defining the set $P' = P_{\mathbf{q},w} - \mathbf{q}$, as the local neighbourhood centered on the output pixel $\mathbf{q}$. This produces the same pixel coordinates for any output location $\mathbf{q}$ in the output grid. We introduce the notation $M(\mathbf{p};\boldsymbol{\alpha})$ as the area-sampled value of the model centred on $\boldsymbol{\alpha}$ at $\mathbf{p}$. Because of centering, $f_O(\mathbf{q}) = M(0;\boldsymbol{\alpha})$. In (3), the objective function is a sum over all output pixels Q – this is what provides the multiple observations. We assume that the model is located somewhere within the region of support, but at an unknown location. Hence, we want to optimise the filter coefficients over all possible model locations and we achieve this by replacing the sum with an integral over all possible model locations, $\boldsymbol{\alpha}$, which is a double integral as the pixel grid is 2D:

$$\min_{H} \iint \left(M(\mathbf{0};\boldsymbol{\alpha}) - \sum_{\mathbf{p}\in P'} M(\mathbf{p};\boldsymbol{\alpha}) H(-\mathbf{p}) \right)^2 d\boldsymbol{\alpha}, \tag{5}$$

where the integrals range over the support region.

Extending the step-edge model used in [9] to two dimensions is more problematic. The logical extension is to have edges of every orientation, although with a rectangular pixel structure, this makes it very difficult to calculate a closed form solution. Simply using horizontal and vertical edges results in the least squares solution of (5) being rank deficient [11]. A compromise that gives a solution that is relatively easy to determine is to use a 2D rectangular pulse model: a pulse with dimensions of one in *x*-direction and one in *y*-direction (the size of a pixel), consisting of two horizontal and two vertical step-edges with corners at the intersections of these edges (see Fig. 2). Area-sampling can be represented by a convolution of the model with another rectangular pulse (assuming square pixels with 100% fill factor) and point-sampling the result. This model has the advantage of separability, simplifying the calculation. The area-sampled rectangular pulse can be expressed as

$$M(\mathbf{p};\boldsymbol{\alpha}) = M_{1D}(p_x;\alpha_x) M_{1D}(p_y;\alpha_y), \tag{6}$$

where $\boldsymbol{\alpha} = (\alpha_x, \alpha_y)$, $\mathbf{p} = (p_x, p_y)$, and

$$M_{1D}(p;\alpha) = \begin{cases} 1-\alpha+p & \text{if } \alpha-1 < p \le \alpha \\ 1+\alpha-p & \text{if } \alpha < p \le \alpha+1 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

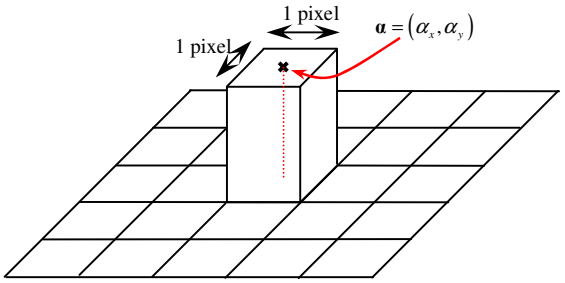

Figure 2. 2D rectangular pulse model, centred on $\boldsymbol{\alpha}$.

We solve (5) by equating the derivative of the objective function to zero, after substituting the "partition of unity" constraint (4) into it. Given a finite region of support of the filter, it can be assumed that R input pixels fall within the region of support. Numbering each of these input pixels with an index $i \in \{0,1,...,R-1\}$ in the order of increasing Euclidean distance from the origin, such that $\mathbf{p}_0$ is the location of the closest input pixel and $\mathbf{p}_{R-1}$ is the location of the farthest input pixel from the origin, the objective function can be rewritten to simplify the notation as

$$\iint \left(M(\mathbf{0};\boldsymbol{\alpha}) - \sum_{i=0}^{R-1} M_i h_i \right)^2 d\boldsymbol{\alpha}, \tag{8}$$

where $M_i = M(\mathbf{p}_i;\boldsymbol{\alpha})$ is the value of the input sample at $\mathbf{p}_i$ and $h_i = H(-\mathbf{p}_i)$ is the weight associated with input i. The "partition of unity" constraint is then

$$h_0 = 1 - \sum_{i=1}^{R-1} h_i . \tag{9}$$

Substituting this into (8) results in the objective function to be minimised

$$\iint \left(\bar{M}(\mathbf{0};\boldsymbol{\alpha}) - \sum_{i=0}^{R-1} \bar{M}_i h_i \right)^2 d\boldsymbol{\alpha}, \tag{10}$$

where $\bar{M}(0;\boldsymbol{\alpha}) = M(0;\boldsymbol{\alpha}) - M(\mathbf{p}_0;\boldsymbol{\alpha})$ and $\bar{M}_i = M_i - M_0$.

The minimisation of (10) can be computed as a linear system of equations by the standard techniques of differentiating with respect to each filter coefficient and equating to zero and swapping the order of differentiation and integration. The resulting equations are

$$\begin{cases} -2\iint \left(\bar{M}(\mathbf{0};\boldsymbol{\alpha}) - \sum_{i=0}^{R-1} \bar{M}_i h_i \right) \bar{M}_1 d\boldsymbol{\alpha} = 0 \\ \vdots \\ -2\iint \left(\bar{M}(\mathbf{0};\boldsymbol{\alpha}) - \sum_{i=0}^{R-1} \bar{M}_i h_i \right) \bar{M}_j d\boldsymbol{\alpha} = 0 \\ \vdots \\ -2\iint \left(\bar{M}(\mathbf{0};\boldsymbol{\alpha}) - \sum_{i=0}^{R-1} \bar{M}_i h_i \right) \bar{M}_{R-1} d\boldsymbol{\alpha} = 0 \end{cases} \tag{11}$$

Rewriting this in matrix notation results in

$$\mathbf{A}\mathbf{h} = \mathbf{b}, \tag{12}$$

where $\mathbf{h} = [h_1 \quad \cdots \quad h_{R-1}]$, $\mathbf{A}$ is a $(R-1)\times(R-1)$ matrix with element in column i and row j being of the form

$$A_{j,i} = \iint \bar{M}_i \bar{M}_j d\boldsymbol{\alpha}, \tag{13}$$

and $\mathbf{b}$ is a $(R-1)\times 1$ vector with element in row j being of the form

$$b_j = \iint \bar{M}_i \bar{M}(0;\boldsymbol{\alpha}) d\boldsymbol{\alpha} . \tag{14}$$

The linear system described by (12) can be solved using any standard method, such as Gaussian elimination. The resulting filter can be applied to the input images to produce the output high-resolution image sampled on a uniform grid.

The step-edge model expressed by (6) and (7) is one example. The resulting expressions for (13) and (14), when this model is in use, can be found in [11]. Other models can also be used, based on different assumptions about the images or using different image features.

III. Method

We test the proposed method using simulated low-resolution images, where the ground-truth output high-resolution image is readily available and a quantitative comparison in the form of root-mean-square-error (RMSE) can be made. We employ the same 8-bit grayscale test images used in [9], shown in Fig. 3. These scenes provide a variety of compositions of high-frequency detail, such as step-edges, texture and relatively low-frequency flat areas, and should indicate how well the proposed method can interpolate different types of scenes.

Four low-resolution images, shifted by various sub-pixel amounts, are used as the input. The offsets are assumed to be known *a priori*. For a given test image, the quality of the reconstruction will depend on the image content, but also on the locations of the input pixels, i.e. the sub-pixel offset of each input image. It can, for example, happen that the input pixels are positioned exactly over the location of the output grid, making the resampling trivial. Therefore, to measure

the performance over a wide range of input image offsets, the high-resolution image was reconstructed one thousand times using random combinations of offsets. This number was sufficient to obtain an estimate of the underlying error distribution. The same one thousand combinations were used for each interpolation method.

The proposed step-edge model-based filter is first tested against the least-squares optimal filter and then compared to a number of existing methods. We implemented the inverse distance weighting method (IDW) first attributed to Shepard [12]; a method by Lertrattanapanich and Bose [3] based on Delaunay triangulation [13]; and a local polynomial expansion method similar to that of Pham et al. [4].

The Delaunay triangulation-based method employed the bivariate cubic polynomial described in [3] (labelled here as Tcub) and also a simple planar fit to each triangle (labelled as Tlin). The local polynomial expansion method employed both the bivariate quadratic (labelled as Quadratic) and cubic (labelled as Cubic) polynomials. We chose to test the 10^{th} order model-based filter (labelled as Model10), as it seemed to provide a good trade-off between computational complexity and performance. It also had an average extent of the region of support of 2 pixels, which is similar to the regions of support employed by IDW and local polynomial expansion methods.

Figure 3. Test images 'bird', 'cat' and 'face'.

In addition, another experiment was performed to produce images for visual comparison. Sixteen low-resolution images were combined to produce a high-resolution image with four times the sampling rate in each direction. The resampled image was deblurred using MATLAB's Image Processing Toolbox `deconvblind()` command.

IV. Experimental Results

The model-based filter follows closely the performance of the optimal filter (see Fig. 4). The performance of both

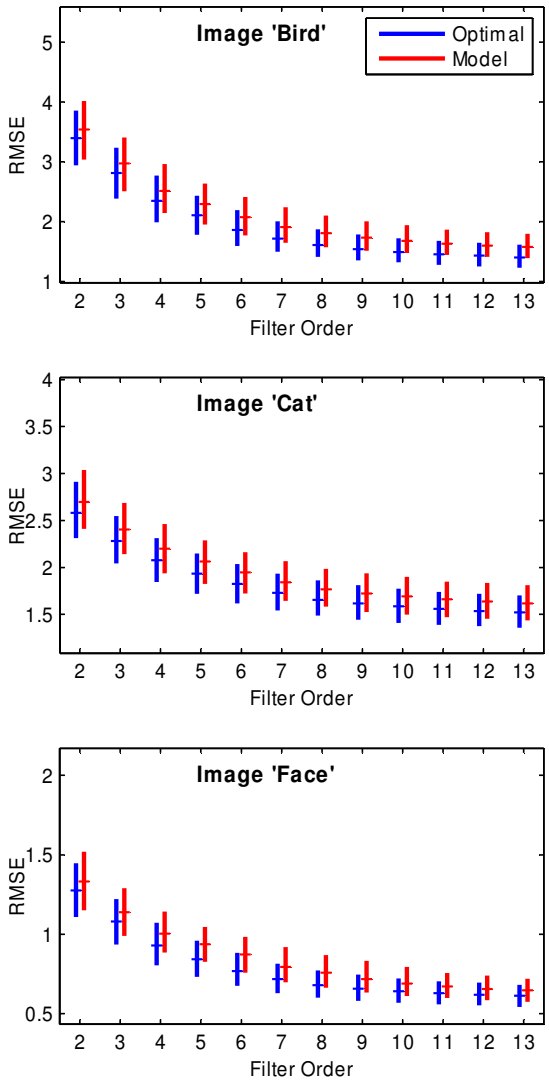

Figure 4. RMSE (in grayscale levels) vs Filter order for optimal and model-based filters. Top and bottom of the vertical lines indicate 3rd and 1st quartiles, respectively of the error distribution. Horizontal dash indicates the median result.

filters improves and the variability of the errors decreases with increasing filter order. However, both filters exhibit asymptotically decreasing RMSE as the filter order is increased.

For image 'bird', the proposed model-based filter resulted in median RMSE (on average across filter orders 2-13) 0.18 grayscale levels higher than optimal (recall that images are 8-bit, hence 256 grayscale levels), for image 'cat' it resulted in 0.11 grayscale levels higher than optimal and for image 'face' 0.065 grayscale levels higher than optimal.

Fig. 5 shows the results of other methods in comparison to the proposed method with filter order 10. IDW method performs the worst, with the Delaunay triangulation methods also having poor performance. It can be observed that the triangulation method with cubic fit performs worse than the local cubic polynomial fit. The triangulation method also has higher variability, most probably due to the fact that surface fit to the triangular patches is sensitive to the shape of the triangles, which is highly dependent on relative offsets of the input images. The proposed method outperformed the other methods for all three test images.

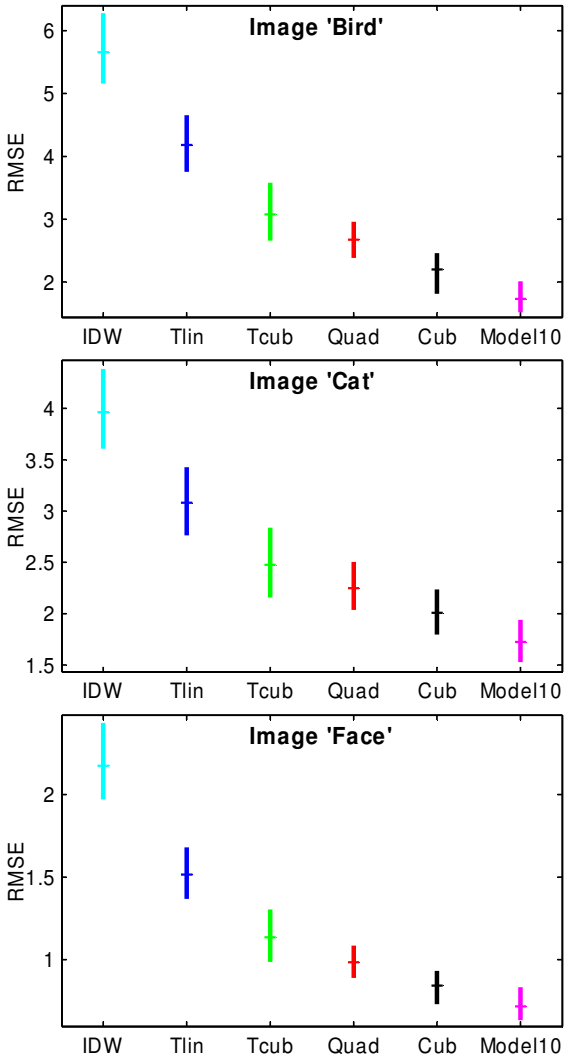

Figure 5. Results of high-resolution image reconstruction using inverse-distance weighting (IDW), Delaunay triangulation with planar fit (Tlin), Delaunay triangulation with cubic fit (Tcub), local polynomial expansion with quadratic (quad) and cubic (cub) fits and the proposed method of order 10. Top and bottom of the vertical lines indicate 3rd and 1st quartiles, respectively. Horizontal dash indicates the median result.

Fig. 6 shows the result of the last part of the experiment. One of the sixteen low-resolution input images is shown (after upsampling by a factor of four using pixel replication) next to the final super-resolved image. The output image contains more detail and it is clear that some aliasing, such as in the streaks in the hair, has been successfully removed.

V. Conclusions

We extended the work presented in [9] to interpolation of two-dimensional non-uniformly sampled images, formed by stacking translated low-resolution images. To perform resampling of real images where the output is unknown, wepropose computing the resampling filter coefficients based on a simple 2D image model, consisting of a rectangular pulse (which is area-sampled to be consistent with the imaging process). Using linear least-squares minimisation to optimise the weights allows for a direct solution for the filter coefficients as a function of the relative translations between the input images.

Figure 6. One low-resolution input image (top) and the final super-resolved image, including deblurring (bottom).

Experimental results show that the resampling filters optimised on the proposed image model perform similarly to the least-squares optimal filters. This demonstrates that the compromises made to the model to enable a simpler closed-form solution have not significantly affected the performance of the interpolation filters.

The model-based filters also outperform a number of commonly used methods for non-uniform image interpolation used in image super-resolution. Visual inspection of the resulting super-resolved image shows noticeable improvement in detail and reduction of aliasing artefacts.

Although the optimum filter was derived from the periodicity of the local neighbourhoods, this constraint is not required for the model based near optimal interpolation. This allows for interpolation of completely scattered data by recalculating the filter weights for each output pixel's local neighbourhood.

Future work will include investigating how the proposed resampling filters perform in the presence of noise and how misregistration errors affect the results.

The optimal linear filter requires the same weights to be used globally. As the filters presented here show near optimal results, the only way an improvement can be achieved is through non-linear means. An extension of this work worth investigating is an adaptive approach by using different models for patches of the image with different characteristics.

REFERENCES

[1] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," IEEE Signal Processing Magazine, vol. 20, no. 3, pp. 21-36, May 2003.

[2] P. Thévenaz, T. Blu, and M. Unser, "Image interpolation and resampling," Handbook of medical imaging, pp. 393-420, Orlando, FL, USA: Academic Press, Inc., 2000.

[3] S. Lertrattanapanich, and N. K. Bose, "High resolution image formation from low resolution frames using delaunay triangulation," IEEE Transactions on Image Processing, vol. 11, no. 12, pp. 1427-1441, December 2002.

[4] T. Q. Pham, L. J. van Vliet, and K. Schutte, "Robust fusion of irregularly sampled data using adaptive normalized convolution," EURASIP Journal on Applied Signal Processing, 2006.

[5] M. Elad, and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," IEEE Transactions on Image Processing, vol. 6, no. 12, pp. 1646-1658, December 1997.

[6] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," IEEE Transactions on Image Processing, vol. 6, no. 12, pp. 1621-1633, December 1997.

[7] C. Vazquez, E. Dubois, and J. Konrad, "Reconstruction of nonuniformly sampled images in spline spaces," IEEE Transactions on Image Processing, vol. 14, no. 6, pp. 713-725, June 2005.

[8] A. Gilman, and D. G. Bailey, "Near optimal non-uniform interpolation for image super-resolution from multiple images," in Image and Vision Computing New Zealand, Great Barrier Island, New Zealand, 2006, pp. 31-35.

[9] A. Gilman, D. G. Bailey, and S. R. Marsland, "Interpolation models for image super-resolution," in 4th IEEE International Symposium on Electronic Design, Test and Applications, Hong Kong, 2008, pp. 55-60.

[10] M. Unser, "Sampling-50 years after Shannon," Proceedings of the IEEE, vol. 88, no. 4, pp. 569-587, 2000.

[11] A. Gilman, "Least-squares optimal interpolation for direct image super-resolution," Ph.D. dissertation, Massey University, Palmerston North, New Zealand, 2009.

[12] D. Shepard, "A two-dimensional interpolation function for irregularly-spaced data," in Proceedings of the 1968 23rd ACM national conference, 1968.

[13] B. N. Delaunay, "Sur la sphère vide," Izvestia Akademia Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk, vol. 7, pp. 793-800, 1934.