CRAFTING THE OBSERVATION MODEL FOR REGULARIZED IMAGE UP-SAMPLING

Hussein Aly and Eric Dubois

University of Ottawa
School of Information Technology & Engineering (SITE)
Ottawa, ON, K1N 6N5 Canada
{haly, edubois}@site.uottawa.ca

ABSTRACT

Often used as an observation model for image interpolation, is the moving average a correct and accurate model for most circumstances? Are there other options? In this paper we present a novel theoretical analysis of the regularized image up-sampling problem focusing on the data fidelity term. We start with formulation of the physical acquisition processes the image has undergone and develop a generalized design for the correct and accurate data fidelity term for regularized image up-sampling.

1. INTRODUCTION

Image up-sampling to achieve higher perceived resolution is a problem with many potential applications such as video-to-film conversion and law enforcement. A continuous image is acquired by a physical camera to produce a lower-resolution (LR) image(s). The physical camera is modelled as a continuous-space(-time) filter followed by sampling on a lower-density lattice. It is desired to obtain a higher-resolution (HR) version of that image sampled on a denser sampling lattice. The HR image is obtained in principle from the continuous image through a theoretical, not necessarily a physically realizable, camera specifying desired properties of the image. The scenario is shown in Fig.1. The theoretical camera consists of a continuous-space(-time) filter and a denser lattice. Our focus in this paper is to find and design an observation model that can best produce the LR image from the HR image. This observation model is the data fidelity term for the regularized up-sampling process. It should be noted that our problem is the inverse one; we are given the LR image and are trying to obtain the HR image. We study the possibility of obtaining such observation models for any scenario for both cameras. Our results are for some existing physical cameras and arbitrary theoretical cameras. As far as we know, we are the first to perform this study and offer a generalized design of this observation filter for arbitrary scenarios. Image up-sampling and super-resolution is an ill-posed problem that can be solved by combining a data fidelity term with a regularization term. Much research has focused on the regularization term, which might involve different *a priori* constraints, to pick one solution with desirable properties from the infinite number of possible solutions. The data fidelity term used for most image interpolation and super-resolution research [1, 2, 3] is the special case proposed in [4]. Their model supposed that the LR image was obtained from the continuous image by a CCD camera whose aperture is modelled by a rect function [5]. If the HR image is also obtained by a rect aperture, then the modelling observation filter is the discrete moving average.

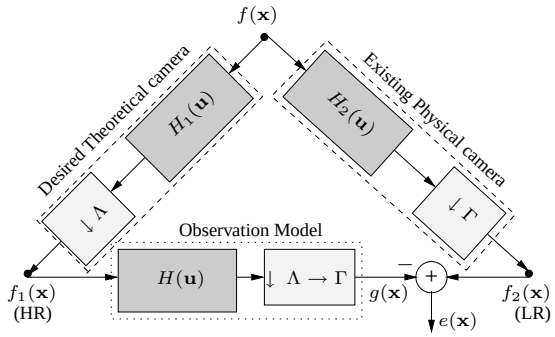


Fig. 1. Scenario for image upsampling.

The motivation in pursuing this study is that an accurate data observation model leads to a better definition of the solution space which is indeed a critical factor for a better quality interpolation [6]. Hence, choices become available for picking nice theoretical cameras and obtaining the corresponding observation model.

2. PROBLEM STATEMENT

Let $f(\mathbf{x})$ be a continuous-space(-time) image that is sampled on two different lattices Λ and Γ. Without loss of generality, assume that $\Gamma \subset \Lambda$, and so $\Lambda^* \subset \Gamma^*$. The superscript * denotes the reciprocal lattice. In cases where neither Λ nor Γ is a subset of the other, then an intermediate lattice

is introduced as in rate conversion by a non-integer factor. The sampling aperture impulse responses for sampling on Λ and Γ are $h_1(\mathbf{x})$ and $h_2(\mathbf{x})$ respectively, yielding the sampled images $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$. We seek a model to relate $f_2(\mathbf{x})$ to $f_1(\mathbf{x})$. The situation is illustrated in Fig.1 where $g(\mathbf{x})$ is the model of $f_2(\mathbf{x})$ and $e(\mathbf{x}) = f_2(\mathbf{x}) - g(\mathbf{x})$ is the modelling error. Here, $g(\mathbf{x})$ is assumed to be obtained by LSI filtering of $f_1(\mathbf{x})$ on Λ, followed by downsampling to Γ.

3. DESIGN OF THE OBSERVATION MODEL

Assume that $f(\mathbf{x})$ has a continuous-space Fourier transform $F(\mathbf{u})$. Then $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ are related to $f(\mathbf{x})$ in the frequency domain by

$$\begin{aligned} F_1(\mathbf{u}) &= \frac{1}{d(\Lambda)} \sum_{\mathbf{r}\in\Lambda^*} F(\mathbf{u}+\mathbf{r}) H_1(\mathbf{u}+\mathbf{r}) \\ F_2(\mathbf{u}) &= \frac{1}{d(\Gamma)} \sum_{\mathbf{s}\in\Gamma^*} F(\mathbf{u}+\mathbf{s}) H_2(\mathbf{u}+\mathbf{s}) \end{aligned} \tag{1}$$

where $d(\cdot)$ is the unit-cell area of its argument lattice. Assuming that $\frac{d(\Gamma)}{d(\Lambda)} = \frac{d(\Lambda^*)}{d(\Gamma^*)} = K$, then $\Gamma^* = \bigcup_{k=1}^{K} (\Lambda^* + \mathbf{d}_k)$ where $\mathbf{d}_k \in \Gamma^*, k = 1, .., K$ are the coset representatives of Λ^* in Γ^* [7]. It follows that

$$G(\mathbf{u}) = \frac{1}{K} \sum_{k=1}^{K} F_1(\mathbf{u}+\mathbf{d}_k) H(\mathbf{u}+\mathbf{d}_k). \tag{2}$$

Since $H(\mathbf{u}+\mathbf{r}) = H(\mathbf{u})$ for any $\mathbf{r} \in \Lambda^*$ and substituting from (1) into (2), it follows that (2) can be written

$$G(\mathbf{u}) = \frac{1}{d(\Gamma)} \sum_{\mathbf{s}\in\Gamma^*} F(\mathbf{u}+\mathbf{s}) H_1(\mathbf{u}+\mathbf{s}) H(\mathbf{u}+\mathbf{s}). \tag{3}$$

Thus, $g(\mathbf{x})$ can be obtained in one step by filtering $f(\mathbf{x})$ with the LSI continuous-space(-time) filter whose frequency response is $H_1(\mathbf{u})H(\mathbf{u})$ and sampling on Γ. It should be noted that $H_1(\mathbf{u})$ is aperiodic while $H(\mathbf{u})$ is periodic, coming from a digital filter. We can also write the Fourier transform of the error $e(\mathbf{x})$ as

$$\begin{aligned} E(\mathbf{u}) = &\frac{1}{d(\Gamma)} \sum_{\mathbf{s}\in\Gamma^*} F(\mathbf{u}+\mathbf{s}) \\ &\big(H_2(\mathbf{u}+\mathbf{s}) - H_1(\mathbf{u}+\mathbf{s}) H(\mathbf{u}+\mathbf{s})\big). \end{aligned} \tag{4}$$

Thus, $e(\mathbf{x})$ can be obtained by filtering $f(\mathbf{x})$ with the LSI continuous-space(-time) filter whose frequency response is $H_2(\mathbf{u}) - H_1(\mathbf{u})H(\mathbf{u})$ and sampling on Γ.

If we assume that $f(\mathbf{x})$ is a realization of a continuous-space(-time) stationary random field with power spectral density (PSD) $S_f(\mathbf{u})$, then the error $e(\mathbf{x})$ is a realization of a discrete-space stationary random field with PSD

$$\begin{aligned} S_e(\mathbf{u}) = &\frac{1}{d(\Gamma)} \sum_{\mathbf{s}\in\Gamma^*} S_f(\mathbf{u}+\mathbf{s}) \\ &|H_2(\mathbf{u}+\mathbf{s}) - H_1(\mathbf{u}+\mathbf{s}) H(\mathbf{u}+\mathbf{s})|^2. \end{aligned} \tag{5}$$

The corresponding mean square error (MSE) is

$$\begin{aligned} \sigma_e^2 &= \int_{P_\Gamma^*} S_e(\mathbf{u}) d\mathbf{u} \\ &= \frac{1}{d(\Gamma)} \sum_{\mathbf{s}\in\Gamma^*} \int_{P_\Gamma^*} S_f(\mathbf{u}+\mathbf{s}) \\ &\quad |H_2(\mathbf{u}+\mathbf{s}) - H_1(\mathbf{u}+\mathbf{s}) H(\mathbf{u}+\mathbf{s})|^2 d\mathbf{u} \end{aligned} \tag{6}$$

where P_Γ^* is a unit cell of Γ^*. Thus the best modelling filter $H_0(\mathbf{u})$ in the MSE sense satisfies

$$\mathbf{h_0} = \arg\min_{\mathbf{h}} \sigma_e^2 \tag{7}$$

where $H(\mathbf{u}) = \Sigma_{\mathbf{x}\in\Lambda} h[\mathbf{x}] \exp(-j2\pi\mathbf{u}\cdot\mathbf{x})$ which is a finite sum for FIR filters, and real for zero-phase filters. We can assume that the integrals in (6) are non-negligible for only a few $\mathbf{s} \in \Gamma^*$ in the vicinity of $\mathbf{0}$. This is reasonable because the aperiodic frequency response of the continuous filters $H_1(\mathbf{u})$ and $H_2(\mathbf{u})$ will decay rapidly in the neighborhood of P_Λ^*, P_Γ^* respectively. This is dictated by the physical aperture pre-filter characteristics in cutting down aliasing. Fixing some number N of independent coefficients for $H(\mathbf{u})$, σ_e^2 is just a real function of the N unknowns, and our objective function (6) can be minimized with a general optimization routine, or optimized analytically since it can be written as a quadratic form in $h[\mathbf{n}]$.

4. ANALYTIC OPTIMIZATION

While the general formulation above is applicable to 2 or 3 dimensions, we now consider the specific case of spatial ($2D$) up-sampling. Assume that Λ is a lattice that admits quadrantal symmetry [8]. Suppose $h[\mathbf{x}]$ has quadrantal symmetry, and the independent coefficients are in the quadrant $\mathcal{Q}$, so that $H(\mathbf{u}) = \sum_{\mathbf{x}\in\mathcal{Q}} h'[\mathbf{x}] \cos(2\pi u x) \cos(2\pi v y)$. If H_1, H_2 are real, then expanding the terms of (6) yields

$$\begin{aligned} \sigma_e^2 = &\frac{1}{d(\Gamma)} \sum_{\mathbf{s}\in\Gamma^*} \int_{P_\Gamma^*} S_f(\mathbf{u}+\mathbf{s}) H_2^2(\mathbf{u}+\mathbf{s}) d\mathbf{u} \\ &- \frac{2}{d(\Gamma)} \sum_{\mathbf{x}\in\mathcal{Q}} h'[\mathbf{x}] \sum_{\mathbf{s}\in\Gamma^*} \int_{P_\Gamma^*} S_f(\mathbf{u}+\mathbf{s}) H_1(\mathbf{u}+\mathbf{s}) \\ &H_2(\mathbf{u}+\mathbf{s}) \cos(2\pi(u+s_x)x) \cos(2\pi(v+s_y)y) d\mathbf{u} \\ &+ \frac{1}{d(\Gamma)} \sum_{\mathbf{x}\in\mathcal{Q}} \sum_{\check{\mathbf{x}}\in\mathcal{Q}} h'[\mathbf{x}] h'[\check{\mathbf{x}}] \sum_{\mathbf{s}\in\Gamma^*} \int_{P_\Gamma^*} S_f(\mathbf{u}+\mathbf{s}) \\ &H_1^2(\mathbf{u}+\mathbf{s}) \cos(2\pi(u+s_x)x) \cos(2\pi(v+s_y)y) \\ &\cos(2\pi(u+s_x)\check{x}) \cos(2\pi(v+s_y)\check{y}) d\mathbf{u} \end{aligned} \tag{8}$$

where, $\mathbf{s} = [s_x, s_y]^{\mathrm{T}}$. Putting $h'[\mathbf{x}]$ into a lexicographic vector $\mathbf{h}'[\mathbf{x}]$, (8) can be simply written as

$$\sigma_e^2 = \mathbf{h'}^{\mathrm{T}} D\mathbf{h'} + \mathbf{b}^{\mathrm{T}}\mathbf{h'} + c \tag{9}$$

where the elements of $\mathbf{b}$ and D can be determined by numerical integration. Then, σ_e^2 is easily minimized with

$$\boxed{\mathbf{h'}_0 = -\frac{1}{2}D^{-1}\mathbf{b}.} \tag{10}$$

5. PSD MODEL OF THE IMAGE

Several models for the PSD of the continuous image $S_f(\mathbf{u})$ exist. Since $S_f(\mathbf{u}) \triangleq F.T\{R_f(\mathbf{x})\}$ then we used the basic model for the autocorrelation of continuous images defined by $R_f(x,y) \triangleq \sigma^2 \exp(-\alpha_1|x|)\exp(-\alpha_2|y|)$, where $\sigma^2 = R_f(0,0)$ is the autocorrelation with zero lag and α_1, α_2 are parameters. We derive the PSD to be

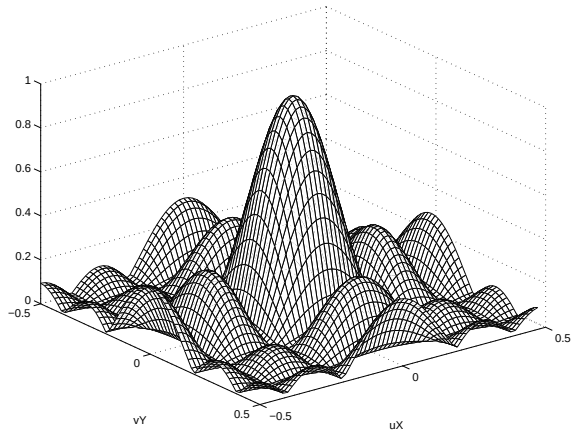
$$S_f(u,v) = 4\sigma^2 \frac{\alpha_2}{4\pi^2 v^2 + \alpha_2^2} \frac{\alpha_1}{4\pi^2 u^2 + \alpha_1^2}. \tag{11}$$

In our simulations we also used the Welch-modified Periodogram to estimate the PSD from a very high resolution image.

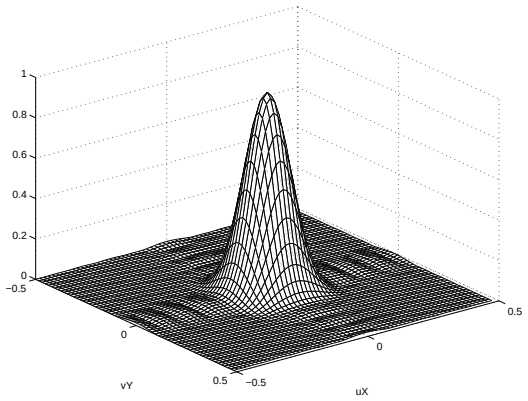
6. SIMULATIONS AND EXPERIMENTS

Most physical apertures have an impulse response that is approximated with a $2D$ Gaussian, rect, or circ function. Hence, our experiments used scenarios with the Gaussian and rect functions for H_2. We will name each scenario according to the aperture used for H_1 and H_2 respectively and the value of K; in other words if H_1 is a Gaussian, H_2 is a rect function, and $K = 25$, then we call it Gauss-Rect($\downarrow$ 25). Our sample scenarios in this paper are Gauss-Gauss($\downarrow$25), Rect-Rect($\downarrow$25), Rect-Gauss($\downarrow$25), and Gauss-Rect($\downarrow$25). For each scenario we obtained an FIR modelling filter $h[\mathbf{x}]$ described in the first row in Table 1.

In order to verify the results obtained we ran a simulation of each scenario and used the modelling filter obtained to measure the actual modelling error $e(\mathbf{x})$ on the images. We started with a very high resolution image simulating the continuous signal $f(\mathbf{x})$. We simulated H_1, H_2 by the appropriate digitally designed Gaussian or moving average with the appropriate cut-off angular frequency. The very-high-resolution image is filtered by H_1, H_2 and then downsampled by different large factors, in order to minimize the error between the digital simulation and the reality of the continuous spectrum analysis, simulating the sampling on both lattices Λ, Γ respectively. We chose Λ, Γ to be rectangular, where $\Lambda = \mathrm{LAT}(\mathrm{diag}(X,Y))$. The images obtained are used to simulate $f_1(\mathbf{x}), f_2(\mathbf{x})$. We measured



(a) Gauss-Rect ($\downarrow$ 25).



(b) Rect-Gauss ($\downarrow$ 25).

Fig. 2. Magnitude of the frequency response for the observation model $H(\mathbf{u})$ for two scenarios.

the actual modelling error $\|e(\mathbf{x})\|_2$ and computed the peak-signal-to-noise ratio (PSNR). For sake of comparison we also chose some reasonable filters to be the modelling filter and measured the actual error. For all the Gaussian filters used as modelling filters, we optimized the variance that maximizes PSNR because we found that results changes drastically with non-optimal variance! All the PSNR measures for all scenarios with all filters are shown in Table 1. Our result for the Rect-Rect scenario agrees with that of [4]. For the Gauss-Gauss scenario our method designed a Gaussian-like observation filter with the optimized variance. We obtained totally new observation models for the Rect-Gauss and Gauss-Rect scenarios.

Another verification we used was to obtain a modelling filter using the images f_1, f_2 and solving a numerical optimization problem, directly from the images. Here we used

Modelling Filter	**Gauss-Gauss (↓25)**	**Rect-Rect (↓25)**	**Gauss-Rect (↓25)**	**Rect-Gauss (↓25)**
Resulting Optimal Modelling filter	Optimized Gaussian-like: Variance is critical	Moving Average: Phase correction is critical	see Fig.2(a): Moving Average can work but not optimal	see Fig.2(b): Neither Gaussian nor moving average
Simulated Optimization	61.96	63.22	53.48	58.35
Our Model with PSD model: (11)	53.51	58.39	52.20	55.15
Moving Average	32.31	63.22	50.42	31.77
Maximally flat $\omega_{cut} = \pi/4$	34.45	37.22	38.03	29.00
Maximally flat $\omega_{cut} = \pi/8$	38.07	33.03	32.76	38.97
Optimized Gaussian[a]	53.69[b]	40.90[c]	40.43[d]	38.42 [e]

[a] Standard deviation of optimized Gaussian [b] 2.1772 pixels [c] 1.5635 pixels [d] 1.497 pixels [e] 1.874 pixels

Table 1. Performance of modelling filters for different scenarios :$\|e(\mathbf{x})\|_2$, MSE expressed as PSNR (dB)

the least squares method because the problem is overdetermined. The simulated modelling filter obtained was also used to compute the modelling error which appears in the first row of results in Table 1 and is named *simulated optimization*. This helped us in obtaining an upper bound on the PSNR that can be obtained from the simulation experiment. However, it can't be practically used since f_1 is not available, and it results in a modelling filter that is biased towards the specific features of the underlying image.

7. CONCLUSION

We were able to formulate a generalized design for the observation model and succeeded in obtaining the optimal observation model for important scenarios in image up-sampling and super-resolution. Furthermore, our method allows us to solve for any desired scenarios; standards conversion posing known physical specifications or up-sampling to images acquired with nice a theoretical aperture specifying desired properties of the image! With our optimal modelling filter tightening the relation between the HR and LR, we can cut down noise amplification during the interpolation process. This leads to relaxing the smoothness (regularization) constraint(s) and helps in obtaining less blurred (over smoothed) results, if a suitable regularizer is selected. We used the bounded total variational regularizer, along with our designed observation model, and the results will appear in a future paper.

8. REFERENCES

[1] B.C. Tom and A.K. Katsaggelos, "Resolution enhancement of monochrome and color video using motion compensation," *IEEE Trans. Image Process.*, vol. 10, pp. 278–287, Feb. 2001.

[2] D. Rajan and S. Chaudhuri, "Generation of super-resolution images from blurred observations using Markov random fields," *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, vol. 3, pp. 1837–1840, 2001.

[3] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, pp. 1167–1183, Sept. 2002.

[4] R. R. Schulz and R. L. Stevenson, "A Bayesian approach to image expansion for improved definition," *IEEE Trans. Image Process.*, vol. 3, pp. 233–242, May 1994.

[5] J. P. Allebach, "Image scanning, sampling, and interpolation," in *Handbook of Image and Video Processing*, A. Bovik, Ed., chapter 7.1, pp. 629–643. Academic Press, San Diego, CA, 2000.

[6] S. Borman and R. L. Stevenson, "Super-resolution for image sequences - a review," *Proc. IEEE Int. Symp. Circuits and Systems*, pp. 374–378, 1998.

[7] E. Dubois, "Video sampling and interpolation," in *Handbook of Image and Video Processing*, A. Bovik, Ed., chapter 7.2, pp. 645–654. Academic Press, San Diego, CA, 2000.

[8] S. Coulombe and E. Dubois, "Linear phase and symmetries for multidimensional FIR filters over lattices," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 45, pp. 473–481, Apr. 1998.