Optimization in Multi-Frame Image Super-Resolution

Benedikt Petko[1] and Roger Donaldson[2]

[1] *Department of Mathematics, Imperial College London*
[2] *Department of Mathematics, University of British Columbia*

Abstract

In this paper we will demonstrate the use of optimization methods as a means of achieving image super-resolution. Using multiple low-resolution frames, we will attempt to retrieve subpixel information from a simulated high-resolution signal.

First, we consider a suitable method of detecting frames' mutual shift in terms of precision and robustness to noise. Secondly, we reconstruct the signal by the least squares method and its L1 and TV regularized versions. These optimization procedures are compared with respect to their quality of reconstruction in terms of signal content and resolution.

Contents

1 Introduction **3**
- 1.1 What is super-resolution . . . 3
- 1.2 Research questions . . . 3

2 Camera sensing model **3**

3 Computing frames relative shift **4**
- 3.1 Modeling shifted frames . . . 4
- 3.2 Detecting shifts (signal registration) . . . 5
  - 3.2.1 Spatial method . . . 6
  - 3.2.2 Frequency method . . . 6
  - 3.2.3 Robustness of shift detection against noise . . . 6

4 Image reconstruction methods **7**
- 4.1 Averaging . . . 7
- 4.2 Least squares . . . 7
  - 4.2.1 Inner sum objective . . . 7
  - 4.2.2 Outer sum objective . . . 8
- 4.3 L1 regularized problem . . . 8
  - 4.3.1 Inner sum single constraint . . . 9
  - 4.3.2 Outer sum single constraint . . . 9
  - 4.3.3 Multiple constraints . . . 9
- 4.4 Total variation regularized problem . . . 9
  - 4.4.1 Inner sum single constraint . . . 10
  - 4.4.2 Outer sum single constraint . . . 10
  - 4.4.3 Multiple constraints . . . 10

5 Results **11**
- 5.1 Resolution . . . 11
  - 5.1.1 Signal content . . . 11
  - 5.1.2 Resolution . . . 12
- 5.2 Least squares reconstructions . . . 12
- 5.3 L1 and TV reconstructions . . . 12
  - 5.3.1 L1 reconstruction . . . 14
  - 5.3.2 TV reconstruction . . . 14

6 Further steps **16**

7 Conclusion **16**

A L1 and TV regularized least square problems are SOCP **16**

B SOCP solver **17**
- B.1 Objective function . . . 17
- B.2 Constraints . . . 17

1 Introduction

1.1 What is super-resolution

Super-resolved image is a high resolution (HR) image obtained from one or multiple low resolution (LR) images. This can be understood as recovering subpixel information contained in a picture from multiple shifted frames. Currently, the most commonly used methods in the image processing community are pixel interpolation, statistical methods and optimization.

Specifically, we would like to circumvent the measuring apparatus' limitation which is the size of a pixel. In order to achieve this, we consider a collection of images mutually shifted by a fraction of a pixel. It is assumed that the real image to be reconstructed does not change during the measurement. The shifts are generally random and need to be found before proceeding with an image reconstruction. The process of calculating the shifts is called signal registration.

The real image is continuous. However, depending on the level of detail in the image to be resolved, it is sufficient to have an image of certain degree of fineness. Thus, the image can be approximated as a uniformly spaced, finite grid of colour intensity values. Equivalently, the grid can be represented by a matrix. For an imaging device consisting of an equally spaced array of sensors, the resolution is inversely proportional to the size of a sensor. Using a commercial digital camera, the resolution of one frame is generally multiple times smaller than the required resolution to register a given level of detail. The ratio between LR and HR grid increments is called the super-resolution factor (SRF). A particular model of image acquisition and processing is introduced in the following section.

1.2 Research questions

- What methods to use for signal registration? How do they work in the presence of noise?
- Is there any gain in resolution if we use a more constrained optimization problem?
- What is the relationship among superresolution factor, number of frames, level of noise, quality and bandwidth (i.e. how close can two signals be and still be registered as distinct)?

2 Camera sensing model

Suppose we want to reconstruct a one dimensional, band-limited continuous signal $u(t)$. By the Nyquist-Shannon sampling theorem, if the spatial frequencies present in $u(t)$ are supported on $[-B, B]$, it is sufficient to sample every $T = \frac{1}{2B}$

or less metres. Therefore, it is sufficient to reconstruct $u[n] := u(nT)$ to underpin all frequencies in the sample.
Let $u[n]$ be the discretized image of desired resolution. However, before the image reaches the user, it is distorted by various effects [1] [2]:

- Movement (translation and rotation)
  This effect can be reversed by precise signal registration and is dealt with in the next part. In one dimension, we only take account of translation:

$$u[n - \alpha_k] \to u[n]$$

- Diffraction in the atmosphere, lens and addition of random electric noise
  This can be modeled as

$$u[n] \to x[n] = \sum_m u[m]h[n - m] + \sigma_n$$

  where h denotes a blurring kernel and σ_n is assumed to be additive white Gaussian noise (AWGN).

- Downsampling
  The high resolution (HR) image $x[n]$ is downsampled to low resolution (LR) image $y[n]$ due to the sensing apparatus' pixel size, which is generally larger than the details to be captured, so that:

$$x[n] \text{ (HR)} \ \to y[n] \text{ (LR)}$$

The ratio between the width of a pixel and T is assumed to be an integer value and is called the superresolution factor (SRF). The blur caused by atmospheric turbulence and diffraction in the lens can be modeled by convolution of the signal with a blurring filter and can be remedied by deconvolution methods. However, in this work we will only focus on reversing the effects of movement, random noise and downsampling.

We will attempt to restore $x[n]$ from multiple measurements (frames) $y^{(k)}[n]$ where $k \in I$, a finite index set. $y^{(k)}$ represents an output of the camera or other imaging device. We identify the finitely supported discrete functions $x[n], y^{(k)}[n]$ with vectors $\mathbf{x}, \mathbf{y}^{(k)}$ respectively. To take into account the width of a pixel, each entry of $\mathbf{y}^{(k)}$ is an average of several neighboring entries of $\mathbf{x}$.

3 Computing frames relative shift

3.1 Modeling shifted frames

It is convenient to work with the circulant matrix when modeling the sampling process. Suppose $\mathbf{u} \in \mathbb{R}^N$ is the noiseless, high-resolution image. We would like

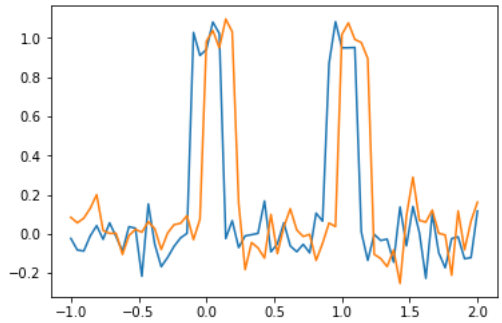


Figure 1: Two shifted frames of a test signal

to reconstruct its noisy, high-resolution representation $\mathbf{x} \in \mathbb{R}^N$ as precisely as possible.

Suppose $C = \frac{1}{n}\begin{pmatrix} 1 & \dots & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 1 & \dots & 0 \\ \vdots & & \ddots & & & \end{pmatrix}$ is a $N \times N$ circulant matrix of the above form, i.e. with $\mathbf{c} = \frac{1}{n}(1 \ \dots \ 1 \ 0 \ \dots \ 0)$ as the first row of C. Let n be the width of a pixel (SRF) and suppose $n|N$ and $\mathbf{c}$ has n 1's. Then we can represent a measurement (a frame) by $\mathbf{y}^{(k)} = R^{(k)}C\mathbf{u}$ where

$$R_{ij}^{(k)} = \begin{cases} 1 & \text{if } j = n(i-1) + \alpha_k \\ 0 & \text{otherwise} \end{cases}$$

with $k = 1 \to m$, α_k are integers in a specified interval and m is the number of frames. R simply takes out every nth entry of $C\mathbf{u}$. In practice, taking m number of frames, we will have data $\{\mathbf{y}^{(k)}, \alpha_k | k \in \{1, \dots, m\}\}$.
Define $\mathbf{y}_u^{(k)}$ the frame upscaled to the high-resolution grid so that $\mathbf{x}$ and $\mathbf{y}_u^{(k)}$ are of the same length and $\mathbf{y}_u^{(k)}$ consists of blocks of SRF-times repeated values.

3.2 Detecting shifts (signal registration)

Before reconstructing the image, it is essential to reverse the motion of the frames which occured during their acquisition.

Suppose we have two frames $y^{(1)}[n], y^{(2)}[n]$ mutually shifted by n_0 and their discrete Fourier transforms $Y^{(1)}[\omega], Y^{(2)}[\omega]$. Here n and n_0 represent increments on the high-resolution grid and we assume that the domain has N points, i.e. $|D| = N$. Then $y^{(1)}[n - n_0] = y^{(2)}[n]$, assuming the high-resolution image did not change between these two measurements. In practice, due to noise we only have $y^{(1)}[n - n_0] \approx y^{(2)}[n]$, therefore there might not be an exact solution for

an integer n_0.

There are two prominent algorithms in the literature that compute the mutual shift n_0 [1].

3.2.1 Spatial method

Although $y^{(1)}[n - n_0] \approx y^{(2)}[n]$ only approximately, we can always find the closest integer n_0 by solving the problem

$$\min_{n_0} \sum_{n \in D} |y^{(1)}[n - n_0] - y^{(2)}[n]|^2.$$

Disregarding the presence of noise, this can be solved by taking the maximum of the lagged autocorrelation function:

$$R[\tau] = \sum_{n \in D} y^{(1)}[n - \tau] y^{(2)}[n]$$

At the maximum $\tau = n_0$, $R[n_0]$ is simply the energy of $y^{(1)}[n]$ and $y^{(2)}[n]$.

3.2.2 Frequency method

By a well-known identity of the Fourier transforms:

$$\mathrm{DFT}\{y^{(1)}[n - n_0]\} = e^{-2\pi n_0 \frac{\omega}{N}} Y^{(1)}[\omega]$$

Then $y^{(1)}[n - n_0] \approx y^{(2)}[n]$ implies that

$$e^{-2\pi n_0 \frac{\omega}{N}} Y^{(1)}[\omega] \approx Y^{(2)}$$

Thus we can find an approximate integer solution for n_0 by solving

$$\min_{n_0} \sum_{\omega \in B} |e^{-i2\pi \frac{\omega}{N} n_0} Y^{(1)}[\omega] - Y^{(2)}[\omega]|^2$$

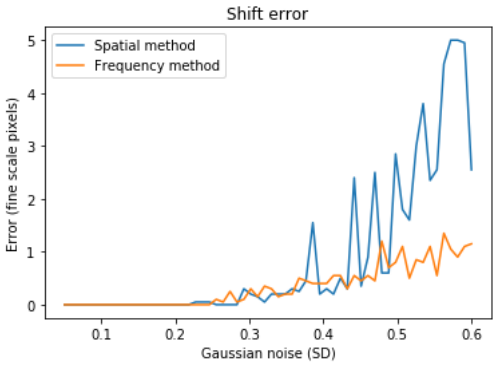
for some bandwidth B which is tricky to choose. The bandwidth is chosen less than the Nyquist rate because of the overlaps of periodic continuation of the FT of the image in the frequency domain. It seems like the frequency approach will be hard to implement if FT has fat tails/high frequencies (it will be hard to establish the true period/bandwidth).

3.2.3 Robustness of shift detection against noise

As a test image, we used two peaks of width 0.2 each on the domain $[-1, 2]$ and tested whether the computed shift corresponded to the true shift.

There seems to be an offset of the error from noise level around $\sigma = 0.4$.

The following graph shows errors where the whole spectrum was considered.



This shows significantly higher robustness of the frequency method against noise. The reason is that the Fourier spectrum of Gaussian noise is constant.

4 Image reconstruction methods

We have several approaches to exploiting multiple measurements. In the following discussion, assume we are given a set of frames $\{\mathbf{y}^{(k)}|k \in I \subset \{1,\ldots,m\}\}$ with $\mathbf{y}^{(k)}$ aligned and upscaled (converted to the fine scale).

4.1 Averaging

The very simplest method to obtain a reconstruction is to take the average of the frames aligned accordingly:

$$\mathbf{x} = \sum_k \mathbf{y}^{(k)}$$

This averages out the noise while still retaining some of the resolution by offsetting the frames. However, the averaging acts much like a smoothing filter applied to the high-resolution test signal. For example, from Fig. 2 we see that the reconstruction did little to preserve edges originally present in the test signal.

The following proposed methods are optimization-based, i.e. the reconstruction is a solution to an optimization problem.

4.2 Least squares

We can distinguish two least squares methods with two distinct objective functions, differing in the position of the sum over the measurements.

4.2.1 Inner sum objective

In this case, the objective function is the norm of the sum of the difference between the modeled and observed frame. Given the data $\{\mathbf{y}^{(k)}|k \in I \subset$

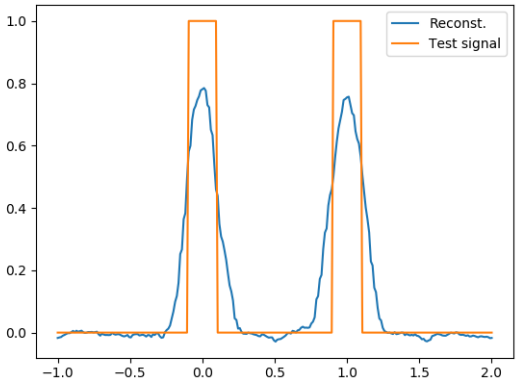


Figure 2: A typical reconstruction of image by averaging frames

$\{1, \ldots, m\}\}$, solve the problem

$$\min_{\mathbf{x}} \ \| \sum_k A^{(k)}\mathbf{x} - \mathbf{y}^{(k)} \|_{L^2}^2.$$

4.2.2 Outer sum objective

An alternative problem is

$$\min_{\mathbf{x}} \ \sum_k \| A^{(k)}\mathbf{x} - \mathbf{y}^{(k)} \|_{L^2}^2.$$

In both cases the problem arises that some of $A^{(k)}$ or $\sum_k A^{(k)}$ might be ill-conditioned and so $\mathbf{x}$ will have infinitely many solutions. Python picks the shortest one. The ill-posedness is dealt with by dropping some of the small eigenvalue components. However, dropping the small eigenvalues means disregarding high frequency components of the image.

A way to deal with the ill-posedness of the least squares problem is to introduce a regularization term. According to literature, TV and L1 regularization work well for preserving resolution. The reason is that real world images are sparse, which means that there are few discontinuities and pixels are clustered in monochromatic patches.

4.3 L1 regularized problem

We considered three types of optimization problems with various regularization terms, listed in order of computational complexity.

$A^{(k)}\mathbf{x} = R^{(k)}C\mathbf{x}$ represents the shifted, down-sampled real image. Assume we are given the data $\{\mathbf{y}^{(k)} | k \in I \subset \{1, \dots, m\}\}$.

4.3.1 Inner sum single constraint

Solve the regularized problem for some λ:

$$\min \|\mathbf{x}\|_1 + \lambda \| \sum_k A^{(k)}\mathbf{x} - \mathbf{y}^{(k)}\|_2^2$$

It turns out that this is equivalent to solving the problem with a single inequality constraint for some ϵ [4]:

$$\min \|\mathbf{x}\|_1 \text{ subject to } \| \sum_k A^{(k)}\mathbf{x} - \mathbf{y}^{(k)}\|_2^2 \leq \epsilon^2$$

4.3.2 Outer sum single constraint

Solve the regularized problem for some λ:

$$\min \|\mathbf{x}\|_1 + \lambda \sum_k \|A^{(k)}\mathbf{x} - \mathbf{y}^{(k)}\|_2^2$$

This is equivalent to the problem with a single inequality constraint:

$$\min \|\mathbf{x}\|_1 \text{ subject to } \sum_k \|A^{(k)}\mathbf{x} - \mathbf{y}^{(k)}\|_{L^2}^2 \leq \epsilon^2$$

4.3.3 Multiple constraints

Solve the regularized problem for some set of λ_k:

$$\min \|\mathbf{x}\|_1 + \sum_k \lambda_k \|A^{(k)}\mathbf{x} - \mathbf{y}^{(k)}\|_2^2$$

This is equivalent to the problem with multiple inequality constraints for some ϵ:

$$\min \|\mathbf{x}\|_1 \text{ subject to } \|A^{(k)}\mathbf{u} - \mathbf{y}^{(k)}\|_2^2 \leq \epsilon_k^2 \text{ for } k \in I.$$

4.4 Total variation regularized problem

For a discrete signal x, the total variation (TV) is defined as $\|x\|_{TV} := \|Dx\|_1$ where

$$D = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & & \\ \vdots & & \ddots & & \\ -1 & 0 & \dots & 0 & 1 \end{pmatrix}$$

is the discrete derivative operator.

As in the case of the L1 regularized problem, we have three ways to reconstruct the original image, using various regularization terms or equivalently different constraints.

4.4.1 Inner sum single constraint

Solve the regularized problem for some λ:

$$\min \|\mathbf{x}\|_{TV} + \lambda \| \sum_k A^{(k)}\mathbf{x} - \mathbf{y}^{(k)}\|_2^2$$

This is equivalent to

$$\min \|\mathbf{x}\|_{TV} \text{ subject to } \| \sum_k A^{(k)}\mathbf{x} - \mathbf{y}^{(k)}\|_2^2 \leq \epsilon^2$$

.

4.4.2 Outer sum single constraint

Solve the regularized problem for some λ:

$$\min \|\mathbf{x}\|_{TV} + \lambda \sum_k \|A^{(k)}\mathbf{x} - \mathbf{y}^{(k)}\|_2^2$$

This is equivalent to

$$\min \|\mathbf{x}\|_{TV} \text{ subject to } \sum_k \|A^{(k)}\mathbf{x} - \mathbf{y}^{(k)}\|_2^2 \leq \epsilon^2$$

.

4.4.3 Multiple constraints

Solve the regularized problem for some set of λ_k:

$$\min \|\mathbf{x}\|_{TV} + \sum_k \lambda_k \|A^{(k)}\mathbf{x} - \mathbf{y}^{(k)}\|_2^2$$

This is equivalent to

$$\min \|\mathbf{x}\|_{TV} \text{ subject to } \|A^{(k)}\mathbf{x} - \mathbf{y}^{(k)}\|_2^2 \leq \epsilon_k^2 \text{ for } k \in I.$$

Remark: In total, we thus have 2 different least squares problems and 6 different regularized least squares problems, which can be solved to yield an image reconstruction. All of the above are convex optimization problems since every norm is a convex function.

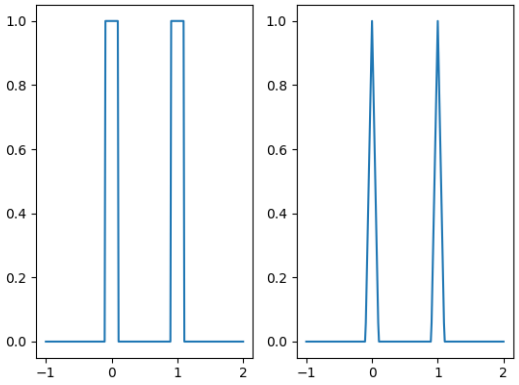


Figure 3: Step-like (left) and spike (right) test signals

5 Results

Software Python 3.6 and 2.7 and the open source package CVXOPT [3] for convex optimization tasks were used to simulate the proposed signal reconstruction methods. The following simulation parameters were used for all experiments unless otherwise specified. The domain consisted of 256 uniformly spaced points, SRF was set to 16 and the number of frames was set to 36. The test signal were either two step-like peaks or two spikes of height 1 centered at 0 and 1, each of width 0.2. The Gaussian noise was randomly generated with mean 0 and standard deviation 0.3.

As we will show graphically, reconstructions share similar features regardless of the constraint type chosen, although they might differ in bandwidth and artifacts.

5.1 Resolution

A user can perceive the quality of an image in two ways: the signal content in the reconstruction, and the ability to distinguish fine features, for instance discerning two sources of light being close together.

5.1.1 Signal content

The quality of an image depends on the amount of noise and artifacts. Define the signal-to-noise ratio $SNR := \frac{\text{maximum of the signal}}{\text{maximum of the noise}}$. Define the quality of

reconstructed image $\mathbf{x}$ in terms of the original image $\mathbf{u}$ as

$$Q := 1 - \frac{\|\mathbf{x} - \mathbf{u}\|_2^2}{\|\mathbf{x}\|_2^2}.$$

Q measures what proportion of the reconstructed image is the true signal in terms of the L^2 norm: if $Q = 1$, we have a pure signal, if $Q \approx 0$, most of the image is noise.
Alternatively,

$$Q := 1 - \frac{\|\mathbf{x} - \mathbf{u}\|_1}{\|\mathbf{x}\|_1}.$$

5.1.2 Resolution

The main point of multi-frame image restoration is to retrieve subpixel information, that is to reconstruct the high-resolution grid. We define the resolution of a reconstruction of the given test signal as

$$\alpha = 1 - (\text{the lowest value between the two test peaks}).$$

α is the factor of confidence that the two test peaks were resolved and it holds that $0 \leq \alpha \leq 1$. Thus, $\alpha = 0$ means the two peaks were not resolved at all, whereas $\alpha = 1$ means they were completely resolved.

5.2 Least squares reconstructions

The least squares reconstructions were computed using the numpy.linalg.lstsq() function. Running simulations shows that the outer and inner sum objectives yield qualitatively very similar reconstructions.

Fig 5 shows a comparison of the both least squares methods with 256 points, SRF set to 16 and 50 frames with respect to the L2-based quality factor. This shows that the level of artifacts is similar for both inner and outer sum objective functions. Moreover, the signal content in the reconstruction diminishes even for low noise levels. Therefore, we can conclude that least squares methods are not robust to noise. This can be explained by the ill-posedness of the least squares problem as mentioned previously. A small distortion of the measurement by noise leads to a substantially different least squares output.

5.3 L1 and TV reconstructions

The ill-posedness of the least squares problem motivates the use of regularization terms. As a result, these methods turn out to be very robust to noise.

The L1 and TV regularized problems described in the previous section can be converted into a second-order cone program (SOCP). See Appendix A for

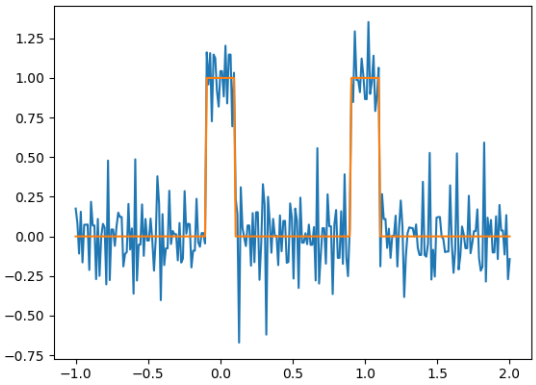


Figure 4: A typical least squares reconstruction (blue) with noise level $\sigma = 0.05$

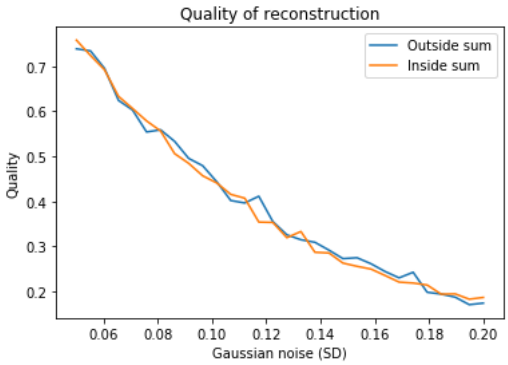


Figure 5: Comparison of the two least squares versions in terms of signal content

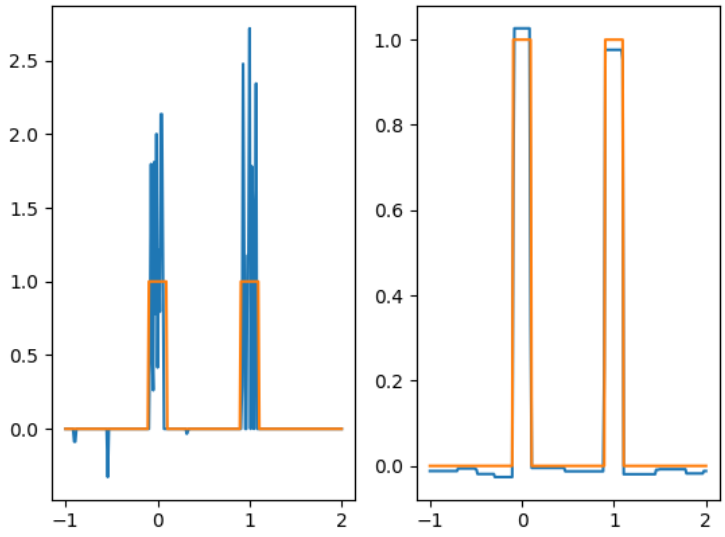


Figure 6: Comparison: Typical reconstruction (blue) using L1 (left) and TV (right) regularization with a test signal (orange)

a more detailed discussion. A SOCP can be solved using CVXOPT's cvxopt.solvers.socp() function. Fig. 6 shows general appearance of these reconstructions.

5.3.1 L1 reconstruction

L1 regularization is particularly suitable for preserving slopes, spikes and impulses, but tends to disrupt monochromatic patches (see Fig. 6). The socp() function's input were

$$G_0 = \begin{pmatrix} I & -I \\ -I & -I \end{pmatrix}, h_0 =$$

5.3.2 TV reconstruction

Total variation regularization works well for preserving edges and monochromatic patches. Fig. 7 shows how the value of α (factor of resolution) changes depending on values of noise and the gap between the test peaks.

The collected data shows that in presence of little or no noise, multiple constraints allow to reconstruct an image with higher resolution (see example in Fig. 8). Moreover, outer sum constraint yields better resolution for all tested

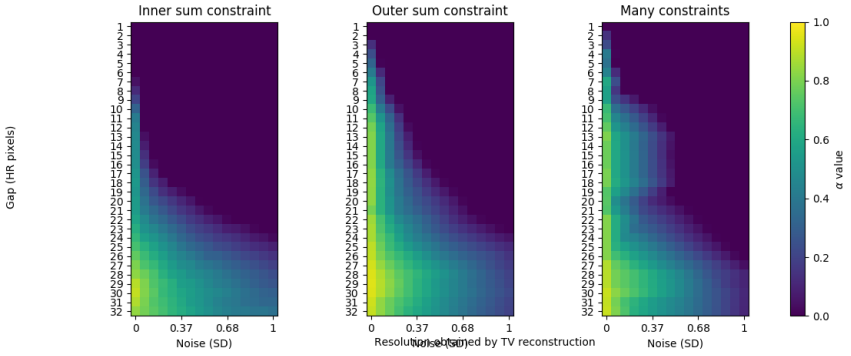


Figure 7: α values with varying noise and bandwidth of the test signal

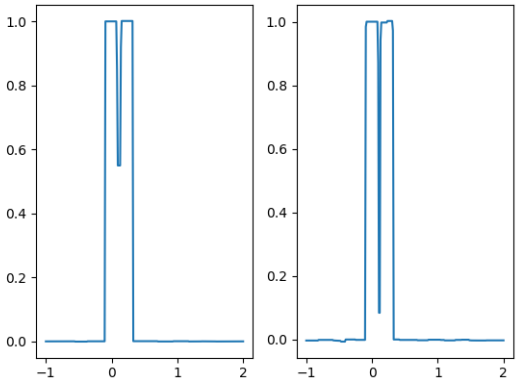


Figure 8: Inner sum single constraint (left) and many constraints (right) methods

values of noise and bandwidth. Therefore we can conclude that the use of a stricter constraint is justified.

6 Further steps

We modeled the degrading process as a noisy Gaussian channel. By Shannon's Third Fundamental Theorem (Noisy Channel Coding Theorem), there might be an upper limit on how much resolution (information) can be obtained given the level of noise. In other words, the atmosphere and the apparatus are a noisy channel which transmit the image, therefore we can never reconstruct it perfectly. We will further inquire what the information-theoretical bounds might be.

Moreover, we only dealt with 1D signals. For a practical application, it is necessary to develop a procedure to reconstruct higher dimensional images. In more than one dimension, rotation of frames should be taken account of besides translation. Also with increasing dimensions and samples the computational power becomes an issue, so the algorithm and code ought to be optimized if possible.

7 Conclusion

Some prior knowledge of the characteristics of the signal helps when choosing a suitable reconstruction method. For a high bandwidth signal with little noise, we recommend using TV or L1 reconstruction with multiple constraints. However, if significant amount of noise is present, it is advisable to use one of the single constraint programs. Moreover, for a signal with edges to be resolved and monochromatic areas, total variation methods yield results with lower level of artifacts than L1 methods. On the other hand, if the signal contains slopes and spikes, L1 reconstruction performs better in preserving such features.

A L1 and TV regularized least square problems are SOCP

In the L1 case, the objective function $\|\mathbf{x}\|_1$ can be substituted with the objective function $\mathbf{1}^T \begin{pmatrix} \mathbf{x} \\ \mathbf{v} \end{pmatrix}$ and the constraint $-\mathbf{v} \preceq \mathbf{x} \preceq \mathbf{v}$, optimizing with respect to x, v where v is an arbitrary vector of the same length as x. In matrix form, this is then the problem

$$\min_{\mathbf{x},\mathbf{v}} \mathbf{1}^T \begin{pmatrix} \mathbf{x} \\ \mathbf{v} \end{pmatrix} \text{ subject to } \begin{pmatrix} I & -I \\ -I & -I \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{v} \end{pmatrix} \preceq \mathbf{0}$$

and corresponding inequality conditions of choice on the frames (inner or outer sum constraint or multiple constraints).

Similarly, we can convert the TV objective to

$$\min_{\mathbf{v},\mathbf{x}} \mathbf{1}^T \begin{pmatrix} \mathbf{x} \\ \mathbf{v} \end{pmatrix} \text{ subject to } \begin{pmatrix} D & -I \\ -D & -I \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{v} \end{pmatrix} \preceq \mathbf{0}$$

B SOCP solver

The following inputs for cvxopt.solvers.socp() can be used to replicate the experiments in this paper. The function takes as the input a vector c, which defines the objective function for the SOCP, and matrices G_0, h_0 and G_i, h_i for $i = 1 \to m$, which represent various inequality constraints. Their exact functioning in the algorithm is described in detail in CVXOPT's documentation [7].

B.1 Objective function

Always set $c = \mathbf{1}$ (a vector of ones of dimension $2N$). This imposes the objective function:

$$\min \mathbf{1}^T \begin{pmatrix} \mathbf{x} \\ \mathbf{v} \end{pmatrix}.$$

Set $h_0 = \mathbf{0}$ (a zero vector of dimension $2N$).
For an L1 regularized problem, set

$$G_0 = \begin{pmatrix} I & -I \\ -I & -I \end{pmatrix}.$$

This imposes the inequality constraint

$$-\mathbf{v} \preceq \mathbf{x} \preceq \mathbf{v}.$$

For a TV regularized problem, set

$$G_0 = \begin{pmatrix} D & -I \\ -D & -I \end{pmatrix}$$

which imposes the inequality constraint

$$-\mathbf{v} \preceq D\mathbf{x} \preceq \mathbf{v}.$$

B.2 Constraints

To introduce the inner sum constraint

References

[1] Yang J, Huang T. Image Super-Resolution: Historical Overview and Future Challenges. In: Milanfar P. (ed.) *Super-Resolution Imaging*. Boca Raton, Florida: CRC Press, Taylor & Francis Group: 2010. p. 3-35.

[2] Gonzalez RC, Woods RE. *Digital Image Processing*. Second Edition. Upper Saddle River, New Jersey: Prentice Hall; 2002.

[3] Andersen M, Dahl J, Vandenberghe L. *CVXOPT*. (Version 1.1.9) [Code] Available from: http://cvxopt.org/. 2017.

[4] Boyd S, Vandenberghe L. *Convex Optimization*. New York: Cambridge University Press; 2004.

[5] Candès EJ, Fernandez-Granda C. Towards a Mathematical Theory of Super-resolution. *Communications on Pure and Applied Mathematics* 2014; 67(6): 906-956.

[6] Ivanov V, Kirichuk V, Kosykh V. Formation of a high-resolution image from a series of mutually shifted images using optimal linear prediction. *Optoelectronics, Instrumentation and Data Processing* 2009; 45(2): 91-99.

[7] Andersen M, Dahl J, Vandenberghe L. *CVXOPT User's Guide*. Available from: http://cvxopt.org/userguide/coneprog.html#second-order-cone-programming. 2017