Super-Resolution With Significant Illumination Change

Wen-Yi Zhao
Sarnoff Corporation
201 Washington Road, Princeton, NJ 08540, USA

Astract

In this paper, we study the impact of illumination change on super-resolution. Based on analysis and experimental results, we show that large illumination change, not a global/parametric one, poses a significant challenge to existing super-resolution algorithms. Existing robust techniques of handling outliers do not offer a good solution either. We propose a two-step framework to super-resolve the object shape at high-resolution first and then synthesize high-resolution images. For applications where appropriate illumination models are not available, we propose a pre-processing framework. Extensive experimental results are presented based on synthetic and real examples.

1 Introduction

The basic principle of super-resolution is to enhance the spatial image resolution by combining high-frequency spatial information spread across a temporal image sequence. The majority of super-resolution algorithms formulate the problem as a signal reconstruction problem from multiple samples. Under ideal case there is hardly any difference among these algorithms in that perfect reconstruction can be expected. However, many practical issues exist within these algorithms and they can be categorized into: 1) Geometric issues where accurate sub-pixel alignment could be difficult to compute; 2) Photometric issues such as the existence of occluding objects; 3) Other issues.

In this paper, we focus on a pure photometric issue of super-resolution: given a sequence of images undergoing *significant* lighting change, how do super-resolution algorithms behave? Methods exist to handle global/parametric illumination change and are accurately referred to as *photometric registration* [1]. In practice, the illumination change is both significant and complicated. A systematic study of this issue has been lacking though Bayesian methods have been suggested to address small illumination change [2, 3]. It is obvious that the image samples under such condition violate the simple imaging model adopted in existing super-resolution literature. We demonstrate that existing robust techniques that treat pixel intensity variation due to illumination change as outliers do not offer a good solution. From a theoretical point of view, the per-

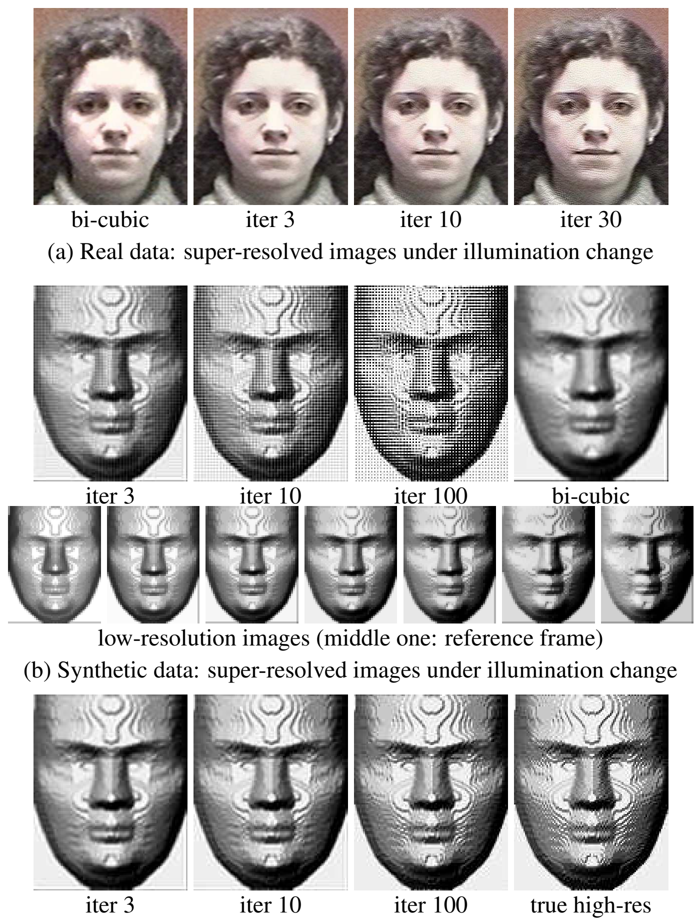


Figure 1: Illustration of the gridding artifact for real data (a) and synthetic data (b-c). For real data, (a) plots the super-resolved high-resolution image at different iterations (refer to Fig. 2(a) for low-resolution images used). For synthetic data, (b) top row plots the super-resolved high-resolution images at different iterations and the bi-cubic interpolated image, and bottom row plots low-resolution images used (the middle one is the reference image that needs to be super-resolved); (c) super-resolved image at different iterations under the same illumination and the true high-resolution image. Notice that under the same illumination, the reconstruction after 100 iterations is almost perfect.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

fect approach to study and address this issue is to introduce a complete imaging model that connects the high-resolution shape and images at both high-resolution and low-resolution. Then solving this issue becomes a problem of super-resolving shape from low-resolution images where shape-from-shading [9, 10] offers an attractive solution. When appropriate illumination models are not available, we propose to extend the photometric registration approach to pre-process images to adjust illumination. The main goal of this paper is to provide an insightful study on this challenging issue and offer a suite of algorithms suitable for different scenarios respcetively.

The remainder of this paper is organized as follows: We study the impact of illumination change upon super-resolution in Section 2. Section 3 presents a shape-from-shading framework to recover the high-resolution image by first estimating the object shape at high-resolution. We then propose a pre-processing framework to reduce the impact of illumination change in Section 4. Experiments based on real and synthetic examples are presented in Section 5. Finally, we conclude our paper in Section 6.

2 The Impact of Illumination

Before we formally study the impact of illumination change, we want to demonstrate an interesting gridding artifact when existing super-resolution algorithms are applied to images with significant lighting change (Fig. 1). The algorithm used here is the BP (back-projection) method [4]. Two sets of data are shown here: while the first set of real data (Fig. 1(a)) is used to *confirm* the existence of the gridding artifact, the second set of synthetic data (Fig. 1(b-d)) is used to *analysize* this artifact.

It is clear from Fig. 1(a) that gridding artifact becomes worse along with the increase of iterations. Figure 1(b) shows a sequence of low-resolution images synthesized from a 3D face model using the Lambertian reflectance model [9] and estimated high-resolution images. If there were no lighting change, the image reconstruction would have been perfect (Fig. 1(c)). Instead, the more iterations the algorithm runs, the worse the gridding artifact is (Fig. 1(b)). We will explain this phenomenon after we incorporate illumination into super-resolution.

2.1 Basics of super-resolution

The basic assumption of any reconstruction algorithms is the modeling of high-res–to–low-res image transformation. Let us denote I_h as the true high-resolution image, g_k as the low-resolution images of the k-th frame, and $\tilde{g}_k$ as the simulated version of g_k from I_h. Then the standard *imaging model* is as follows

$$\tilde{g}_k = \{[I_h]^{B_k} \cdot h\} \downarrow s, \tag{1}$$

where $\downarrow s$ denotes a down-sampling operator by a factor s, $[\cdot]^{B_k}$ denotes a backward-warping process and h is a blurring kernel.

Based on this model, a least square problem can be formulated and either ML estimation or MAP estimation could be sought. Without loss of generality, we choose the BP method [4] as our baseline algorithm. To be specific, the iterative update of I_h in the BP method is

$$I_h^{(n+1)} = I_h^{(n)} + \frac{1}{K}\sum\{[(g_k - \{[I_h^{(n)}]^{B_k} \cdot h\} \downarrow s) \uparrow s]^{F_k} \cdot p\}, \tag{2}$$

where $I_h^{(n)}$ is the recovered high-resolution image at the n-th iteration, p is a back-projection kernel, $\uparrow s$ denotes a up-sampling operator by a factor s, $[\cdot]^{F_k}$ denotes a forward-warping process, and $\tilde{g}_k^{(n)} = \{[I_h^{(n)}]^{B_k} \cdot h\} \downarrow s$.

2.2 Why gridding artifact?

To explain the interesting behavior (Fig. 1) of super-resolution algorithms, we first describe the physical imaging process used to generate these images. We borrow the general form of *image irradiance equation* [9]

$$I[x, y] = R(p[x, y], q[x, y]), \tag{3}$$

where $(p[x, y], q[x, y])$ are the partial derivatives of the object surface. The form of the irradiance map R depends on the chosen reflectance model.

The face shape in Fig. 1(b-c) is quite complex and different lighting angles yield images with significantly different appearances (bottom row of Fig. 1(b)). First, we should not expect perfect reconstruction from existing algorithms that adopt the the standard imaging model (Eq. 1). Next, from Eq. 2 it is clear that the key component of the BP algorithm is the *updating term*, i.e., the sum of zoomed, warped, and filtered version of the individual *projection error*: $(g_k - \tilde{g}_k^{(n)})$. At each iteration, the algorithm updates previous estimate $I_h^{(n)}$ with this term. Ideally, this term brings new information only due to the difference between previous estimate and the true high-resolution image. However, we know that *extra* difference due to lighting change (Eq. 3) could be significant and varies depending upon pixel location. This varying difference plus the discrete computational scheme result in local edges, or *grids*. Once these artifacts are introduced into the iteration (Eq. 2), they will become more evident and introduce more artifacts through back-projection.

In practice, we may see less severe gridding artifact (Fig. 1(a)). For example, real-world surfaces may be less sensitive to lighting change than that predicted by theoretical models. It is important to understand why such phenomenon happen and distinguish it from similar artifacts such as ringings due to the lack of samples [5].

2.3 Image-only operation

To address this significant issue of illumination, the first thought is to address it in the *image space* where image-only operations are allowed. Under this constraint, a new *imaging model* should be adopted in super-resolution

$$\tilde{g}_k^L = L_k^{-1}(\{[I_h]^{B_k} \cdot h\} \downarrow s) = L_k^{-1}(\tilde{g}_k) \tag{4}$$

where L_k^{-1} is the inverse function of $L_k(\cdot)$ that represents the mapping from the illumination condition of g_k to that of $\tilde{g}_k$. When mapping L is of parametric form, it can be solved via photometric registration [1]. For general illumination change, it could be difficult to recover L *perfectly*.

3 A Shape-from-Shading Framework

From the above analysis of illumination issue, it is clear that generally perfect reconstruction can not be achieved in the *image space* alone. In stead, we need to operate in the *shape space* where different high-resolutions images are generated from the *common* object shape. In this Section, we present a shape-from-shading framework that recovers the object shape first and then its images, both at high-resolution. Given an illumination model [9] and various illumination angles, the synthesis of respective images from a known shape is straight-forward.

To recover the high-resolution object shape, we propose a two-step approach. In the first step, we propose an multi-image based least square algorithm to estimate a low-resolution version of the high-resolution object shape from low-resolution images. This algorithm appears to be similar to the photometric stereo approach [10]. However, the assumption that all images correspond to one common shape is not true in our case.

In the second step, we propose an algorithm that refines the high-resolution shape interpolated from the estimated low-resolution shape.

3.1 Estimating object shape at high-resolution

Comparing this *shape super-resolution* problem with the *image super-resolution* problem, we notice that it is a more complex problem due to the additional mapping from the object shape to images at high-resolution. In summary, image super-resolution is based on the following quadratic cost function of I_h:

$$E = \sum ||g_k - \{[I_h]^{B_k} \cdot h\} \downarrow s||^2; \quad (5)$$

while the cost function for shape super-resolution is

$$E_s = \sum ||g_k - \{[R(p_h[x,y], q_h[x,y])]^{B_k} \cdot h\} \downarrow s||^2. \quad (6)$$

If we look carefully into Eq. 2 that is used to solve Eq. 5, it is just a special version of Gauss-Newton method with additional operations such as warping, filtering, subsampling and up-sampling. We notice that all these operations are linear, hence the essence of Gauss-Newton method keeps unchanged. For the optimization problem (Eq. 6) we can apply similar special Gauss-Newton method. However, for cost function that has higher-than quadratic order, we need to apply an iterative strategy that first linearizes the error item e_k ($E_s = \sum e_k^T e_k$)

$$e_k = (g_k - \{[R(p_h[x,y], q_h[x,y])]^{B_k} \cdot h\} \downarrow s), \quad (7)$$

and then starts Gauss-Newton updating. For more stable solution, the celebrated Levenberg-Marquardt algorithm should be used. In summary, we propose the following local algorithm that also enforces surface integrability [7].

Shape Super-Resolution Algorithm

1. Linearization iteration starts.
2. At each iteration, process all pixels via a moving window centered at pixel $[x, y]$
   - Re-arrange the (p_h, q_h)'s within the neighbor to form a vector X.
   - Linearize the error item e_k at X_0, the previous estimate of X: $A_k[X_0]X - b_k[X_0]$.
   - Run linear updating (to be described next) to update X.
   - Obtain only $(p_h[x,y], q_h[x,y])$ from vector X.
   - Move to the next pixel to compute (p_h, q_h).
3. Linearization iteration ends.

Linear Updating

1. Starts iterative updating based on the following equation

$$X^{[n+1]} = X^{[n]} + \frac{1}{K}\sum\{[(A_k^T b_k - A_k^T A_k\{[X^{[n]}]^{B_k} \cdot h\} \downarrow s) \uparrow s]^{F_k} \cdot p\} + \alpha CC^T X^{[n]} \quad (8)$$

2. Iterative updating ends.

Comparing Eq. 8 with Eq. 2, we see that this algorithm is a generalization of the original BP algorithm. More specifically, $A_k^T b_k$ is related to the low-resolution images g_k and X represents the high-resolution shape that serves the same role as the high-resolution image in the original BP algorithm. The extra item $CC^T X^{[n]}$ in Eq. 8 is due to the surface integrability constraint. We expect this algorithm to have a good convergence property provided that an reasonable initial result is estimated in the first step.

4 A Pre-Processing Framework

As suggested in Section 2, in practice it would be a good idea to apply some pre-processing algorithms so illumination variations can be compensated for to some extent. To handle general illumination change, we propose a wavelet-/pyramid-based adjustment. More specifically, we replace the complicated $\tilde{g}_k$ with $[g]^{b_k}$. It is simply the warped version of g towards g_k where b_k is the low-resolution of the backward motion B_k.

Preprocessing framework For each image g_k that needs illumination compensation, do the following:

1. Back-warp the low-resolution image g to image g_k.
2. Apply algorithms to adjust the illumination of g_k so it will be close to that of $[g]^{b_k}$.

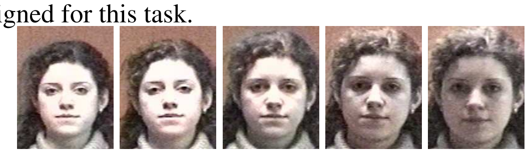
Without complicated operations such as shadow detection and segmentation, the basic task of illumination adjustment algorithms is to adjust the global intensity and keep local structure intact. While many choices are available, wavelet or pyramid is an ideal candidate for the task of signal reconstruction while changing the global intensity [8]. In particular, pyramid/wavelet representations are localized in spatial frequency as well as in space. This in turn helps to reconstruct images with decent global structure and fine local details. In short, here is the simple algorithm to adjust the illumination of g_k to match that of g.

Pyramid-based illumination adjustment

1. Decompose both g_k and $[g]^{b_k}$ into multi-level $[0 \to N]$ Gaussian pyramids.
2. Construct a new $\hat{g}_k$ using pyramids $[0 \to (N-1)]$ from g_k and the top level N from $[g]^{b_k}$.

5 Experiments

We have applied various algorithms to both real and synthetic data. For real data, we test the pre-processing framework and compare several algorithms. An interlaced video was captured using a regular 8MM SONY analog camcoder while the subject is walking underneath overhead lights. We first track the region of interest, the subject's face, and then compute consistent flow [12] for performing super-resolution. As can be seen from Fig 2(a), the subject's face undergoes significant non-rigid motion and illumination change. For comparision (Fig. 2), the best overal super-resolution result is obtained using the pyramid-based illumination adjustment method combined with masking. When correlation-based masking is not applied, pyramid-based BP method significantly reduces the gridding artifact compared to the basic BP method. It is understandable that the failed robust BP method was not designed for this task.



(a) De-interlaced low-resolution samples

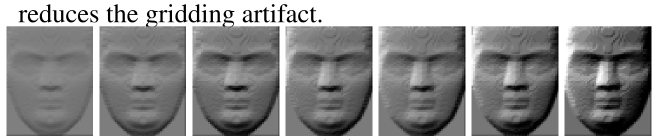
BP method   Pyramid BP   Masking Pyramid   Robust BP

(b) Comparison of super-resolution results

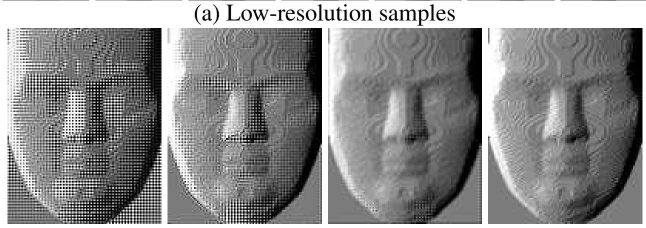
Figure 2: Algorithm comparison. (a) Low-resolution samples of a de-interlaced real image sequence. The *third* column is the reference image needs to be resolved. (b) Result comparison of BP method, pyramid-based illumination adjustmeent method (without and with masking), and robust BP method.

For synthetic data, we test extensively the shape-from-shading framework to recover the high-resolution shape and image. The same face shape (Fig. 1) has been used throughout the experiments. We have successfully experimented with two different reflectance models to recover the low-resolution shape: the Lambertian model (Fig. 1) and a linear shading model [11] where intensity value is linear in terms of p and q (Fig. 3(a)).

To recover the high-resolution shape, we have currently tested the linear shading model(Fig. 3). As in Fig. 1, this set of challenging examples are designed to test the limits of super-resolution: 1) the known parametric transformations are 0.25 or 0.5 pixels so the masking algorithm does not work, 2) rich 3D structure creates so much illumination variation that gridding artifact is widespread. Figure 3 plots results from various algorithms after 100 iterations. The proposed shape-from-shading framework almost perfectly recovers the high-resolution image. The results from other methods are not good though the robust BP method reduces the gridding artifact.



(a) Low-resolution samples



Pyramid adjustmeent   Robust BP   Shape-fom-Shading   True high-res

(b) Comparison of super-resolution results

Figure 3: Comparison of super-resolution algorithms at 100 iterations with low-resolution samples synthesized from a linear shading model [11].

6 Conclusion

We have presented a systematic study of how illumination change impacts super-resolution and proposed a suite of algorithms to address this issue. The first one is based on shape-from-shading and can offer perfect solution under ideal case. The second one does not offer perfect solution but does reduce the impact without assuming an appropriate illumination model. Experimental results have been reported with both real and synthetic data.

References

[1] D. Capel and A. Zisserman, "Computer Vision Applied to Super Resolution," *IEEE Signal Processing Magazine,* pp. 75-86, May, 2003.

[2] H. Shekarforoush, M. Berthod, J. Zerubia, and M. Werman, "Subpixel Bayesian Estimation of Albedo and Height," *Int. Journal of Computer Vision,* 19(3), pp. 289-300 1996.

[3] V.N. Smelyanskiy, P. Cheeseman, D. Maluf, and R. Morris, "Bayesian Super-Resolved Surface Reconstruction from Images," *Proc. CVPR,*, Vol I, pp. 375-382, 2000.

[4] Irani, M. and Peleg, S. 1993. Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency. *Journal of Visual Comm. and Image Repre.,* Vol. 4, pp. 324-335.

[5] Baker, S. and Kanade, T. 2000. Limits on Super-Resolution and How to Break Them. In *Proc. Conf. Comp. Vision and Patt. Recog.*

[6] A. Rav-Acha, A. Zomet, and S. Peleg, Robust Super Resolution, 2001. In *Proc. Conf. Comp. Vision and Patt. Recog.* Vol I, pp. 645-650.

[7] Frankot, R.T. and Chellappa, R. 1988. "A Method for Enforcing Integrability in Shape from Shading Algorithms," *IEEE Trans. on PAMI*, Vol. 10, pp. 439-451.

[8] P. Burt and E. Adelson, "The Laplacian Pyramid as A Compact Image Code," *IEEE Trans. on Communications,* Vol 31, pp.583-540, 1983.

[9] B.K.P. Horn, and M.J. Brooks, *Shape from Shading,* MIT Press: Cambridge, MA, 1989.

[10] Woodham, R. 1980. Photometric Method for Determining Surface Orientation from Multiple Images. *Optical Engineering,* Vol. 19, pp. 139-144.

[11] A. Pentland, "Linear Shape form Shading," *Int. Journal of Computer Vision,* Vol. 4(2), pp. 153-162, 1990.

[12] W. Zhao and H. Sawhney, "Is super-resolution with optical flow possible?" In *Proc. European Conf. Comp. Vision,* I:599-613, 2002.