SUPER RESOLUTION APPROACHES FOR PHOTOMETRICALLY DIVERSE IMAGE SEQUENCES

Murat Gevrekci and Bahadir K. Gunturk

Louisiana State University
Department of Electrical and Computer Engineering
Baton Rouge, LA 70803
lgevre1@lsu.edu, bahadir@ece.lsu.edu

ABSTRACT

Super resolution (SR) is a well-known technique to increase the quality of an image using multiple overlapping pictures of a scene. SR requires accurate registration of the images, both geometrically and photometrically. Most of the SR articles in the literature have considered geometric registration only, assuming that images are captured under the same photometric conditions. This is not necessarily true as external illumination conditions and/or camera parameters (such as exposure time, aperture size and white balancing) may vary for different input images. Therefore, photometric modeling is a necessary task for super resolution. In this paper, we investigate super-resolution image reconstruction when there is photometric variation among input images.

Index Terms— Super resolution, photometric registration, high dynamic range imaging

1. INTRODUCTION

Detailed visual descriptions are demanded in a variety of commercial and military applications, including surveillance systems, medical imaging, and aerial photography. Imaging devices has limitations in terms of, for example, spatial resolution, dynamic range, and noise characteristics. Researchers are working to improve sensor characteristics by exploring new materials, manufacturing processes, and technologies. In addition to the research in sensor technology, image processing ideas are also explored to improve image quality. One of these image processing ideas is super-resolution image reconstruction, where multiple images are combined to improve spatial resolution. Super resolution (SR) algorithms exploit information diversity among overlapping images through subpixel image registration. Subpixel accurate registration allows to obtain frequency components that are unavailable in individual images. The idea of SR image reconstruction has been investigated extensively, and commercial products are becoming available. For detailed literature surveys on SR, we refer the readers to other sources [1, 2].

In this paper, we focus on a new issue in SR: How to do SR when some of the input images are photometrically different than the others. Other than a few recent papers, almost all SR algorithms in the literature assume that input images are captured under the same photometric conditions. This is not necessarily true in general. External illumination conditions may not be identical for each image. Images may be captured using different cameras that have different radiometric response curves and settings (such as exposure time and ISO settings). Even if the same camera is used for all images, camera parameters (exposure time, aperture size, white balancing, gain, etc.) may differ from one image to another. Therefore, an SR algorithm should include a photometric model as well as a geometric model and incorporate these models in the reconstruction.

In photometric modeling, one should take into account the camera response function (CRF) in addition to the camera settings. The CRF, which is the mapping between the irradiance at a pixel to the output intensity, is not necessarily linear. Charges created at a pixel site due to incoming photons may exceed the holding capacity of that site. When the amount of charge at a pixel site approaches the saturation level, the response may deviate from a linear response. In addition to unavoidable physics-originated nonlinearity of a sensor, camera manufacturers may also introduce intentional nonlinearity to CRF to improve contrast and visual quality.

The saturation of CRF and finite number of bits (typically eight bits per channel) to represent a pixel intensity limit the resolution and the extend of the dynamic range that can be captured by a digital camera. Because a real scene typically has much wider dynamic range than a camera can capture, an image captures information from only a limited portion of a scene. By changing exposure rate, it is possible to get information from different parts of a scene. In high-dynamic-range (HDR) imaging research, multiple low-dynamic-range (LDR) images (that are captured with different exposure rates) are combined to produce a HDR image [3, 4].

Despite the likelihood of photometric variations among images of a scene, there are few SR papers addressing reconstruction with such image sets. In [2], photometric changes were modeled as global gain and offset parameters among image intensities. This is a successful model when photometric changes are small. When photometric changes are large, nonlinearity of CRF should be taken into consideration. In [5], we included a nonlinear CRF model in the imaging process, and proposed an SR algorithm. The algorithm produces high-spatial-resolution and high-dynamic range images based on the maximum a posteriori probability estimation technique. The algorithm presented in [5] is actually one of the approaches that can be taken when there is photometric diversity among input images. This will be explained later in the paper.

2. PHOTOMETRIC MODELING

For a complete SR algorithm, spatial and photometric processes of an imaging system should be modeled. Spatial processes (spatial motion, sampling, point spread function) have been investigated relatively well; we focus on only the photometric side in this paper. As mentioned earlier, in the context of SR, two photometric models have been used. The first one is the affine model used in [2], and the second one is the nonlinear model used in [5]. In this section, we review and compare these two models.

This work was supported by the National Science Foundation under Grant No 0528785.

2.1. Affine Photometric Model

Suppose that N images of a static scene are captured and these images are geometrically registered. Let $\mathbf{q}$ be the irradiance of the scene, and $\mathbf{z}_i$ be the ith measured image. According to the affine model, the relation between the irradiance and the image is as follows:

$$\mathbf{z}_i = a_i \mathbf{q} + b_i, \; i = 1, ..., N, \tag{1}$$

where the gain (a_i) and offset (b_i) parameters can model a variety of things, including global external illumination changes and camera parameters such as gain, exposure rate, aperture size, and white balancing. Then, the ith and the jth images are related to each other as follows:

$$\mathbf{z}_j = a_j \mathbf{q} + b_j = a_j \left(\frac{\mathbf{z}_i - b_i}{a_i} \right) + b_j = \frac{a_j}{a_i} \mathbf{z}_i + \frac{a_i b_j - a_j b_i}{a_i}. \tag{2}$$

Defining $\alpha_{ji} \equiv \frac{a_j}{a_i}$ and $\beta_{ji} \equiv \frac{a_i b_j - a_j b_i}{a_i}$, we can in short write (2) as

$$\mathbf{z}_j = \alpha_{ji} \mathbf{z}_i + \beta_{ji}. \tag{3}$$

The affine relation given in (3) is used in [2] to model photometric changes among the images to be used in SR reconstruction. In [2], the images are first geometrically registered to the reference image to be enhanced. After geometric registration the relative gain and offset terms with respect to the reference image are calculated with least squares estimation. Each image is photometrically corrected using the gain and offset terms. This is followed by SR reconstruction.

2.2. Nonlinear Photometric Model

A typical image sensor has a nonlinear response to amount of light it receives. According to the nonlinear photometric model, an image $\mathbf{z}_i$ is related to the irradiance $\mathbf{q}$ of the scene as follows:

$$\mathbf{z}_i = f\left(a_i \mathbf{q} + b_i\right), \tag{4}$$

where $f(\cdot)$ is the camera response function (CRF), and a_i and b_i are again the gain and offset parameters as in (1). Then, two images are related to each other as follows:

$$\mathbf{z}_j = f\left(\frac{a_j}{a_i} f^{-1}\left(\mathbf{z}_i\right) + \frac{a_i b_j - a_j b_i}{a_i} \right) = f\left(\alpha_{ji} f^{-1}\left(\mathbf{z}_i\right) + \beta_{ji} \right). \tag{5}$$

The function $f\left(\alpha_{ji} f^{-1}(\cdot) + \beta_{ji}\right)$ is known as the intensity mapping function (IMF). Although IMF can be constructed using CRF and exposure ratios, it is not necessary to estimate camera parameters to find IMF. IMF can be extracted directly from the histograms of the images [6].

CRF can also be estimated without finding IMF. In [7] a parametric CRF model is proposed; and these parameters are estimated iteratively. [8] used a polynomial model instead of a parametric model. In [4], a nonparametric CRF estimation technique with a regularization term is presented. Another nonparametric CRF estimation method is proposed in [9], which also includes modeling of noise characteristics.

2.3. Comparison of Photometric Models

Here, we provide an example to compare affine and nonlinear photometric models. In Figure 1(a, b, c, d), we provide four images captured with a hand-held digital camera. One of the images is set as the reference image (Figure 1(d)) and the others are converted to it photometrically using the affine and nonlinear models. (Before photometric conversion, images were registered geometrically.) The residual images computed using the affine model (Figure 1(e, f, g)) and the nonlinear model (Figure 1((i, j, k)) are displayed. The affine model parameters are estimated using the least squares technique and are shown in Figure 1(h). The nonlinear IMFs are estimated using the method in [10]. The estimated mappings are shown in Figure 1(l). As seen from the residual images, the nonlinear model works better than the affine model. The affine model performs better when the exposure ratios are close; the model becomes more and more insufficient as the exposure ratios differ more.

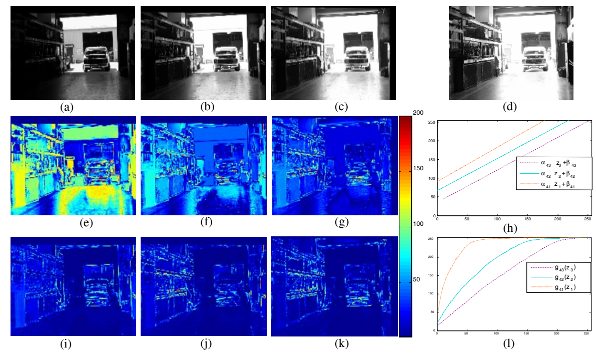

Fig. 1. Comparison of affine and nonlinear photometric conversion. (a)-(d) are the images captured with different exposure rates. All camera parameters other than the exposure rates are fixed. The images are geometrically registered. The relative exposure rates are $1/16$, $1/4$, $1/2$, and 1, respectively. Image $\mathbf{z}_4$ is set as the reference image and other images are photometrically registered to it. (e)-(g) are the residuals using the affine model. (h) The photometric mappings for (e)-(g). (i)-(k) are the residuals using the nonlinear model. (l) The photometric mappings for (i)-(k).

A super-resolution algorithm requires an accurate modeling of the imaging process. The restored image should be consistent with the observations given the imaging model. A typical iterative SR algorithm starts with an initial estimate, calculates an observation using the imaging model, finds the residual between the calculated and real observations, and projects the residual back on the initial estimate. When the imaging model is not accurate or registration parameters are not estimated correctly, the algorithm would fail. In this section, we conclude that nonlinear photometric models should be a part of SR algorithms when there is a possibility of photometric diversity among input images.

3. SR UNDER PHOTOMETRIC DIVERSITY

When all input images are not photometrically identical, there are two possible ways to enhance a reference image: (i) spatial resolution enhancement and (ii) spatial resolution and dynamic range enhancement. In (i), only spatial resolution of the reference image is improved. This requires photometric mapping of all input data to the reference image. In (ii), both spatial resolution and dynamic range of the reference image are improved. This can be considered as a combination of high-dynamic-range imaging and super-resolution image restoration.

3.1. Spatial Resolution Enhancement

In spatial resolution enhancement, all input images are converted to the tonal range of reference image. After photometric registration, a traditional SR reconstruction algorithm can be applied. However, this is not a straightforward process when the intensity mapping is

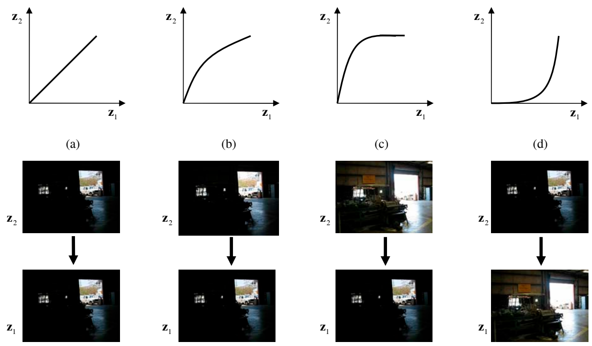

Fig. 2. Various photometric conversion scenarios. First row illustrates possible photometric conversion functions. Correspondingly, second row shows the input, and last row shows the reference image which produces the tonal conversion functions in first row.

nonlinear. Refer to Figure 2 that shows various intensity mapping functions (IMFs). Suppose that $\mathbf{z}_1$ is the reference image to be enhanced. Input image $\mathbf{z}_2$ is aimed to be photometrically converted onto $\mathbf{z}_1$ in all cases. There are four cases in Figure 2:

• In case (a), the input image $\mathbf{z}_2$ is photometrically same as the reference image; so there is no photometric registration necessary.

• In case (b), the IMF is nonlinear; however, there is no saturation. Therefore, the intensities of $\mathbf{z}_2$ can be mapped onto the range of $\mathbf{z}_1$ using the IMF without loss of information.

• In case (c), there is bright saturation in $\mathbf{z}_2$. The IMF is not a one-to-one mapping. The problematic region is where the slope of the IMF is zero or close to zero. For saturated regions, there is no information in $\mathbf{z}_2$ corresponding to $\mathbf{z}_1$. Therefore, photometric registration is not possible from $\mathbf{z}_2$ to $\mathbf{z}_1$. Even for the small-slope regions, noise in $\mathbf{z}_2$ would be amplified and reconstruction would be affected negatively.

• In case (d), there are regions of small slope and large slope. Large-slope regions are not issue because mapping from $\mathbf{z}_2$ to $\mathbf{z}_1$ would not create any problem. The problem is still with the small-slope regions (dark saturation regions in $\mathbf{z}_2$), where quantization and noise floor are effective.

One solution to this saturation problem is to use a certainty function associated with each image. The certainty function should weight the contribution of each pixel in a photometrically registered image based on the reliability of conversion. If a pixel is saturated or close to saturation, then the certainty function should be close to zero. If a pixel is from a reliable region, then the certainty function should be close to one.

We now put these ideas in SR reconstruction. Let $\mathbf{x}$ be the (unknown) high-resolution version of a reference image $\mathbf{z}_r$, and define $g_{ri}(\mathbf{z}_i)$ as the IMF that takes $\mathbf{z}_i$ and converts it to the photometric range of $\mathbf{z}_r$, (therefore, $\mathbf{x}$). Referring to equation (5), $g_{ri}(\mathbf{z}_i)$ includes the CRF $f(\cdot)$, and gain α_{ri} and offset β_{ri} parameters:

$$g_{ri}(\mathbf{z}_i) \equiv f\left(\alpha_{ri} f^{-1}(\mathbf{z}_i) + \beta_{ri}\right). \tag{6}$$

We also need to model spatial processes. Define $\mathbf{H}_i$ as the linear mapping that takes a high-resolution image and produces a low-resolution image. In this case, $\mathbf{H}_i$ is applied on $\mathbf{x}$ to produce the photometrically converted ith observation $g_{ri}(\mathbf{z}_i)$. $\mathbf{H}_i$ includes motion (of the camera or the objects in the scene), blur (caused by the point spread function of the sensor elements and the optical system), and downsampling. (Details of $\mathbf{H}_i$ modeling can be found in the special issue of the IEEE Signal Processing Magazine [1] and the references therein.)

We need to find $\mathbf{x}$ that produces $g_{ri}(\mathbf{z}_i)$ when $\mathbf{H}_i$ is applied to it, for all i. The least squares solution to this problem would minimize the following cost function:

$$C(\mathbf{x}) = \sum_i \left\| g_{ri}(\mathbf{z}_i) - \mathbf{H}_i \mathbf{x} \right\|^2. \tag{7}$$

As explained earlier, the problem associated with the saturation of the IMF can be solved using a certainty function, $w(\mathbf{z}_i)$. We formulate our equations using a generic weight function $w(\mathbf{z}_i)$. Our specific choice will be given in the experimental results section. We now define a diagonal matrix $\mathbf{W}_i$ whose diagonal is $w(\mathbf{z}_i)$, and use this matrix as multiplicative term in equation (7) for constructing the weighted least squares cost function. This new cost function is

$$C(\mathbf{x}) = \frac{1}{2} \sum_i \left(g_{ri}(\mathbf{z}_i) - \mathbf{H}_i \mathbf{x}\right)^T \mathbf{W}_i \left(g_{ri}(\mathbf{z}_i) - \mathbf{H}_i \mathbf{x}\right). \tag{8}$$

Since dimensions of the matrices are large, we wanted to avoid matrix inversion and apply the gradient descent technique to find $\mathbf{x}$ that minimizes this cost function. Starting with an initial estimate $\mathbf{x}^{(0)}$, each iteration updates $\mathbf{x}^{(0)}$ in the direction of the negative gradient of $C(\mathbf{x})$:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \gamma \sum_i \mathbf{H}_i^T \mathbf{W}_i \left(g_{ri}(\mathbf{z}_i) - \mathbf{H}_i \mathbf{x}^{(k)}\right), \tag{9}$$

where γ is the step size at the kth iteration. We found γ using the exact line search that minimizes $C\left(\mathbf{x}^{(k)} + \gamma \mathbf{\Phi}\right)$ at each step, where $\mathbf{\Phi}$ is the negative gradient of C.

3.2. Spatial Resolution and Dynamic Range Enhancement

Here, the goal is to produce a high-resolution and high-dynamic range image. This requires formulating the image acquisition from the unknown high-resolution irradiance $\mathbf{q}$ to each observation $\mathbf{z}_i$. Adding the spatial processes (geometric warping, blurring with the PSF, and downsampling) to equation (4), the imaging process can be formulated as

$$\mathbf{z}_i = f\left(a_i \mathbf{H}_i \mathbf{q} + b_i\right), \tag{10}$$

where $\mathbf{H}_i$ is the linear mapping (including warping, blurring, and downsampling operations) from a high-spatial-resolution irradiance signal to a low-spatial-resolution irradiance signal. $f(\cdot)$, a_i, and b_i are the CRF, gain, and offset terms as in (4).

This time the weighted least squares estimate of $\mathbf{q}$ minimizes the following cost function:

$$C(\mathbf{q}) = \frac{1}{2} \sum_i \left(\frac{f^{-1}(\mathbf{z}_i) - b_i}{a_i} - \mathbf{H}_i \mathbf{q}\right)^T \mathbf{W}_i \left(\frac{f^{-1}(\mathbf{z}_i) - b_i}{a_i} - \mathbf{H}_i \mathbf{q}\right) \tag{11}$$

This cost function is basically analogous to the cost function in equation (8). Starting with an initial estimate for $\mathbf{q}$, the rest of algorithms work similar to the previous one. The only difference is that intensity-to-intensity mapping $g_{ri}(\cdot)$ in (8) is replaced with intensity-to-irradiance mapping $\frac{f^{-1}(\cdot) - b_i}{a_i}$. Unlike the intensity-to-intensity mapping, intensity-to-irradiance mapping requires explicit estimation of the CRF, gain and offset parameters.

In [5], we investigated this spatial and dynamic range enhancement idea in more detail. We left out the details in this paper.

4. GEOMETRIC AND PHOTOMETRIC REGISTRATION

SR requires accurate geometric and photometric registration. If the actual CRF and the exposure rates are unknown, the images must be geometrically registered before these parameters can be estimated. On the other hand, geometric registration is problematic when images are not photometrically registered. There are three possible approaches to this problem: (i) Images are first geometrically registered using an algorithm that is insensitive to photometric changes. This is followed by photometric registration. (ii) Images are first photometrically registered using an algorithm that is insensitive to geometric misalignments. This is followed by geometric registration. (iii) Geometric and photometric registration parameters are estimated jointly.

We take the third in experiments. For geometric registration, we used a feature-based algorithm, which requires robust exposure-insensitive feature extraction and matching. In our experiments, feature points are first extracted using the Harris corner detector [11]. These feature points are matched using normalized cross correlation, which is insensitive to contrast changes. The RANSAC method is then used to eliminate the outliers and estimate the homographies. After geometric registration comes photometric registration. In our experiments, we used the method in [7].

5. EXPERIMENTS AND RESULTS

We captured a data set of 22 images with a hand-held digital camera. The exposure rate was changed manually to introduce photometric diversity into the data set. The resolution enhancement factor is four and the number of iterations was set to two in all experiments. The PSF is taken as a Gaussian window of size [7x7] and of variance 1.7. We use a hat function as in [4] for the certainty function. The results are shown in Figure 3. For the spatial-only enhancement approach, we did experiments when the reference is chosen as an over-exposed image and also when it is chosen as an under-exposed image to show the robustness of the algorithm. For the spatial and dynamic range enhancement approach, we created an initial estimate by applying a standard "HDR from multiple exposures" [4] algorithm. The initial estimate is then updated iteratively as in [5]. As seen, Figure 3(g) and (h) have higher resolution than the corresponding input images. Figure 3(h) is the result of the resolution and dynamic range enhancement algorithm. Notice that both spatial resolution and dynamic range are improved. Also note this approach estimates the irradiance $\mathbf{q}$, which needs to be compressed in dynamic range to display on limited range displays. Displaying HDR images on limited range displays is an active research area. Here, we used a gamma correction to display the image. (The gamma parameter is 0.5.)

6. CONCLUSIONS AND FUTURE WORK

In this paper, we showed how to do SR when the photometric characteristics of the input images are not identical. We showed two possible approaches, one of them enhancing spatial resolution only, and the other enhancing both spatial resolution and the dynamic range. This idea can also be utilized in HDR imaging applications. We demonstrated that nonlinear photometric modeling should be preferred to affine photometric modeling. Because of the limited space, we could not provide further details and discussions on some of the topics, such as weight function, geometric registration, and tonal mapping. These details will be provided in later publications.

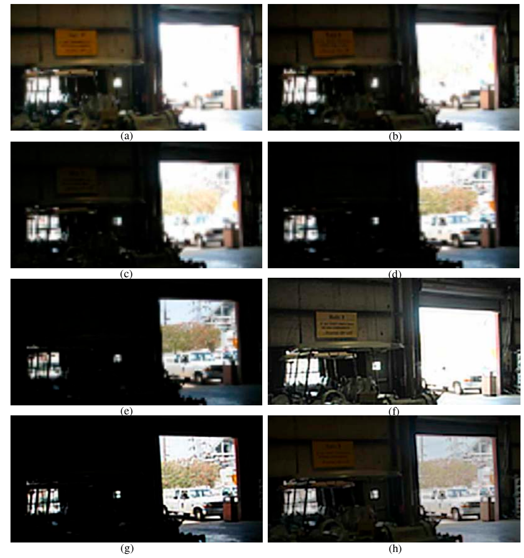

Fig. 3. Cropped regions of observations and SR results. (a)-(e) Some of the bilinearly interpolated input images. (f) SR when (a) is the reference image. (g) SR when (e) is the reference image. (h) SR using the technique presented in Section 3.2.

7. 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] D.Capel, *Image Mosaicing & Super-resolution*. Spring., 2004.

[3] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, in *High dynamic range imaging: acquisition, display and image-based lighting*. Morgan Kaufmann, 2006.

[4] P. E. Debevec and J. Malik, "Modeling and rendering architecture from photographs," in *Proc. of the ACM SIGGRAPH*, 1997, pp. 369–378.

[5] B. K. Gunturk and M. Gevrekci, "High-resolution image reconstruction from differently exposed images," *Signal Processing Letters*, vol. 13, no. 4, pp. 197–200, April 2006.

[6] M. D. Grossberg and S. K. Nayar, "Determining the camera response from images: what is knowable?" *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 11, pp. 1455–1467, November 2003.

[7] S. Mann and R. Mann, "Quantigraphic imaging: estimating the camera response and exposures from differently exposed images," vol. 1, pp. 842–849, 2001.

[8] T. Mitsunaga and S. Nayar, "Radiometric self calibration," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, vol. 1, June 1999, pp. 374–380.

[9] Y. Tsin, V. Ramesh, and T. Kanade, "Statistical calibration of the ccd imaging process," *Proc. Int. Conf. Computer Vision*, vol. 1, pp. 480–487, 2001.

[10] S. Mann, "Comparametric equations with practical applications in quantigraphic image processing," *IEEE Trans. Image Processing*, vol. 9, no. 8, pp. 1389–1406, August 2000.

[11] C. G. Harris and M. Stephens, "A combined corner and edge detector," in *Fourth Alvey Vision Conference*, 1988, pp. 147–151.