AN APPLICATION OF MATTING LAPLACIAN MATRIX TO RANGE IMAGE SUPER-RESOLUTION

YUXIANG YANG
Department of Automation
University of Science and Technology of China
Hefei, P. R. China
yangyux@mail.ustc.edu.cn

ZENGFU WANG
Institute of Intelligent Machines
Chinese Academy of Sciences
Department of Automation
University of Science and Technology of China
Hefei, P. R. China
zfwang@ustc.edu.cn

Received 25 December 2011
Accepted 21 March 2012

This paper describes a successful application of Matting Laplacian Matrix to the problem of generating high-resolution range images. The Matting Laplacian Matrix in this paper exploits the fact that discontinuities in range and coloring tend to co-align, which enables us to generate high-resolution range image by integrating regular camera image into the range data. Using one registered and potentially high-resolution camera image as reference, we iteratively refine the input low-resolution range image, in terms of both spatial resolution and depth precision. We show that by using such a Matting Laplacian Matrix, we can get high-quality high-resolution range images.

Keywords: Range measuring technology; range image super-resolution; Matting Laplacian Matrix; steepest descent method.

1. Introduction

There are a variety of range measuring technologies to acquire 3D information about our world. For example, laser range scanners can provide extremely accurate and dense 3D measurement over a large working volume [Batle et al., 1998; Besl, 1989; Jarvis, 1983; Poussart and Launardeau, 1989; Salvi et al., 2004]. But, most of these high-quality scanners measure a single point at a time and it limits their applications to static environments only.

Recently some new sensors (Canestavision; Z-cam; Swiss Ranger SR-2) have been developed...
to overcome this limitation. These sensors measure time delay between transmission of a light pulse and detection of the reflected signal on an entire frame once by using extremely faster shutter. Though this technology is promising, in the current generation, these sensors are very expensive and very limited in terms of resolution. For example the Canasta EP DevKit sensors can provide range images only up to $64 \times 64$.

To overcome this problem, some range image super-resolution approaches have been proposed, which roughly fall into two categories:

1. **Combining range and image data**: This kind of range image super-resolution methods has primarily been accomplished by using a high-resolution camera image taken from the same location. The core idea is to enforce simple statistical between range image and camera image, and smoothness of geometry in areas of largely uniform color. Regularization has taken the form of a MRF [Diebel and Thrun, 2006], bilateral filtering of the cost volume [Yang et al., 2007], and bilateral filtering in the image plane [Kopf et al., 2007]. Although these methods are often computationally very efficient, they frequently suffer from artifacts that are due to the heuristic nature of the enforced statistical model.

2. **Super-resolution from range data only**: This kind of range image super-resolution methods enhances the resolution by combining only range recordings of a static scene that were taken from slightly displaced viewpoints. The core idea is the observation model between low-resolution range image and high-resolution range image. Kil et al. [2006] were among the first to explore such an idea for laser triangulation scanners. Rosenbush et al. [2007] used this idea to Flash LADAR range data. Rajagopalan et al. [2008] proposed a Markov Random Field-based resolution enhancement method for PMD camera data. Schuon et al. [2008, 2009] and Edeler et al. [2009] used this idea to TOF camera data. Though this strategy only needs low resolution range data to enhance the resolution, the information in the range data is limited.

In this paper, we present a novel range image super-resolution approach. This method combines the advantages of those two categories. Specifically, we use the range observation model between low resolution range image and high resolution range image to construct a minimization problem; then we apply Matting Laplacian Matrix to exploit the information between range image and camera image and construct a regularization term. Figure 1 gives the framework of our proposed method. In the experiments, we show that by using such a Matting Laplacian Matrix and the observation model, we can get high-quality high-resolution range images.

The paper is organized as follows: Sec. 2 presents the observation model used in this paper. In Sec. 3 we show how to use the Matting Laplacian Matrix to exploit the information in the camera image data. Section 4 describes our proposed method and the algorithm details. The experimental results are reported in Sec. 5, followed by a conclusion in Sec. 6.

![Fig. 1. Framework of our proposed method.](image)
2. The Observation Model

We cast super-resolution as the problem of inverting the formation process of low-resolution range image of a high-resolution 3D scene. This formation process can be described as range image observation model as follows:

\[ g = DHf + v. \] (1)

Here, \( f \) (of size \([r^2N^2 \times 1]\)) and \( v \) (of size \([N^2 \times 1]\)) are mutually independent and the noise is an additive white Gaussian noise with variance \( \sigma^2 \).

The observed low-resolution range image \( G \) (of size \([rN \times rN]\)) represents a blurring filter, and \( D \) (of size \([N^2 \times r^2N^2]\)) represents the downsampling operator.

Assuming that the pixels of range image are mutually independent and the noise is an additive white Gaussian noise with variance \( \sigma^2 \), we can get:

\[ p(g | f) = \left( \frac{1}{2\pi\sigma^2} \right)^{\frac{r^2N^2}{2}} \times \exp \left\{ -\frac{1}{2\sigma^2} \| g - DHf \|_2^2 \right\}. \] (2)

Then, the ML estimate of \( f \) under the Bayesian framework can be described as:

\[ f^* = \arg \max_f p(g | f). \] (3)

So, to extract the high-resolution original scene from the observed low-resolution range image, we need to solve the following minimization problem:

\[ f^* = \arg \min_f \| g - DHf \|_2^2. \] (4)

To obtain a high-quality high-resolution range image (the original scene), regularization term \( \gamma(f) \) should be added to this minimization problem. In the paper, we apply Matting Laplacian Matrix to exploit the information between range image and camera image to construct such a regularization term.

3. The Matting Laplacian Matrix

The resolution of the observed range image is very limited. For example the Canasta EP DevKit sensors can provide range images of only up to \(64 \times 64\). Just by using the information during the observed low-resolution range image, we cannot get a high-quality high-resolution range image. Providentally it is simple to get high-resolution camera image of the same scene, considering that the range discontinuities in a scene often co-occur with color or brightness changes within the associated camera image of the same scene. So the information in the high-resolution camera image can help us to get high-quality high-resolution range image of the same scene.

Since discontinuities in range image \( f \) can account for the discontinuities in camera image \( I \) of the same scene, this assumption allows us to write \( f \) as a linear function of \( I \):

\[ f_i \approx a_i I_i + b, \forall i \in w, \] (5)

where \( w \) is a small image window. This linear relation is similar to the one prior used in Levin et al. [2006]. The goal in this paper will be to find \( f, a \) and \( b \) minimizing the cost function:

\[ J(f, a, b) = \sum_{j \in l} \left( \sum_{i \in w_j} (f_i - a_i I_i - b_j)^2 + \varepsilon a_j^2 \right), \] (6)

where \( w_j \) is a small image window around pixel \( j \).

First we define \( J(f) \) as:

\[ J(f) = \arg \min_{a, b} J(f, a, b). \] (7)

Then using the theorem proposed in Levin et al. [2006], we can get:

\[ J(f) = f^T L f, \] (8)

where \( L \) is an \(r^2N^2 \times r^2N^2\) Matting Laplacian Matrix, whose \((i, j)\)th entry is:

\[ \sum_{k \in w_k} \left( \delta_{ij} - \frac{1}{|w_k|} \left( 1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{|w_k| + \sigma_k^2} \right) \right). \] (9)

Here \( \delta_{ij} \) is the Kronecker delta, \( \mu_k \) and \( \sigma_k^2 \) are the mean and variance of the intensities in the window \( w_k \) around \( k \), and \(|w_k|\) is the number of pixels in this window.
4. Final Optimization Problem

We use $J(f)$ as the regularization term $\gamma(f)$ and add it to the minimization problem.

$$f^* = \arg\min_f (\alpha\|g - DHf\|_2^2 + f^T Lf).$$

(10)

Define:

$$L(f) = (\alpha\|g - DHf\|_2^2 + f^T Lf).$$

(11)

Then, we use the steepest descent method to solve this optimization problem as:

$$f_{r+1} = f_r + kd.$$  

(12)

Here, $d$ is the descent gradient, $k$ is the step size.

Calculating the derivation of (11):

$$\frac{\partial L(f)}{\partial f} = \alpha[2H^T D^T DHf - 2H^T D^T g] + 2Lf.$$  

(13)

Thus, $d$ is written as:

$$d = -(\alpha H^T D^T DHf - \alpha H^T D^T g + Lf).$$

(14)

Then order $\frac{\partial L(f)}{\partial k} = 0$, we can obtain $k$ as:

$$k = \frac{d^T d}{\alpha d^T H^T D^T DHd + d^T Ld}.$$  

(15)

So, using the steepest descent method we can solve the proposed optimization problem and get the final high-resolution range image after iterations.

The proposed method can be briefly described by the following steps:

1. Apply bilinear interpolation to the given range low-resolution image $g$ and obtain the initial estimate $f_0$.
2. Begin the iteration loop index $l = 0$, apply the steepest descent method to calculate $f_{l+1}$ based on (12)–(14), iterate until $\|f_{l+1} - f_l\| < \eta$ or the iteration count exceeds a maximum.
3. Use the final high-resolution range image to reconstruct the three-dimensional scene image.

5. Experiments and Results

The experimental system consists of a Canesta EP DevKit camera and FLEA digital camera, which is the same as the system used in Yang et al. [2007]. The EP DevKit camera can produce range images with size up to
64 \times 64 of the objects in its view, and the FLEA camera can produce color image with resolution up to 1024 \times 768 (640 \times 640 used in this paper).

The method proposed in Yang et al. [2007] have got a lot of attention and performs excellent within this research community. So, we mainly compare the results of our method with those of Yang et al. [2007].

Figure 2 gives part of the super-resolution results of different methods to show the visual effect of our method clearly.

Figure 3 is the experimental results of super-resolution range and contains four columns: Fig. 3(a) shows the high-resolution camera images; Fig. 3(b) shows the low-resolution input range images; Fig. 3(c) shows the super-resolution results of Yang et al. [2007]; Fig. 3(d) shows the super-resolution results of our method.

Figure 4 shows the three-dimensional images reconstructed from the high-resolution range images of different methods. To show the visual effect of our method clearly, Fig. 4 gives a
Fig. 4. Reconstructed three-dimensional pictures: (a) Result of Yang et al. [2007]; (b) result of our method.
different angle of view of the reconstructed three-dimensional images.

From Figs. 2 and 3 we can see that the visual effect of our method is better, especially the super-resolution edge of our method is much smoother. From Fig. 4 we can see that our reconstructed three-dimensional picture has fewer mistakes and has more similarity with the real 3D scene.

6. Conclusion

In this paper, we present a novel range image super-resolution method. This method combines the advantages of those two main categories. We apply the observation model between low-resolution range image and high-resolution range image to obtain the information in the range data; then we apply Matting Laplacian Matrix to exploit the information between range image and camera image. In the experiments, we show that by using such a Matting Laplacian Matrix and the observation model, we can get high-quality high-resolution range images. But it is found recently that our method can only be applied successfully to those simple scenes. In the future, we will make great efforts to improve our method to more complex scenes.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (No. 60875026) and the Key Laboratory of Biomimetic Sensing and Advanced Robot Technology, Anhui, China. The authors would like to thank them for their help.

References


Biography

Yuxiang Yang was born in 1987, Taizhou, Zhejiang, P. R. China. He received the B.S. degree in department of automation from the University of Science and Technology of China in 2008, and has been studying for a Ph.D. degree in pattern recognition of professional at University of Science and Technology of China since 2008.

He is now a graduate student at the Joint Laboratory of Intelligent Science and Technology at the University of Science and Technology of China. His research interests include image sparse representation, image processing, image super-resolution and stereo vision.

Zengfu Wang was born in 1960, Hefei, Anhui, P. R. China. He received the B.S. degree from the University of Science and Technology of China in 1982, the M.S. degree in communications and electronic engineering from Nanjing Electronic Engineering Research Center, and the Ph.D. degree in control engineering from Osaka University, Japan, respectively.

He is now a professor and Director at the Institute of Intelligent Machines, Chinese Academy of Sciences, and a professor at the University of Science and Technology of China. His research interests include stereo vision, biometrics, affective computing and intelligent robotics.