LOCAL TRILATERAL UPSAMPLING FOR THERMAL IMAGE

Aleksandar S. Cvetkovic † *and Toshikazu Wada* ‡

† Department of Mathematics and Informatics, University of Belgrade,
Kraljice Marije 16, 11120 Belgrade, Serbia
‡ Department of Computer and Communication Sciences, Wakayama University,
930, Sakaedani, Wakayama 640-8510, Japan

ABSTRACT

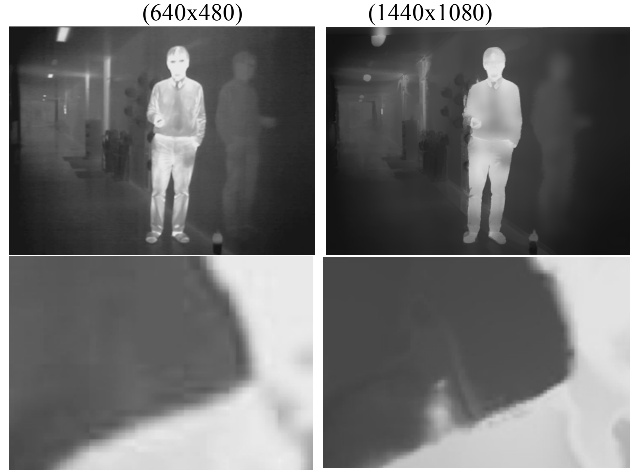
This paper presents an image upsampling method. Joint-bilateral filtering has been successfully applied to this problem that upsample "target" images using high-resolutional "control" images. In this filtering, the kernel is a product of weights representing spatial proximity and color (or intensity) proximity of the "control" image. However, when the "target" image involves textures that are invisible in the "control" image, these textures are damaged by this upsampling. In the case of thermal target image and visible-light control image, we can often find such textures. For solving this problem, we first propose trilateral upsampling, which is a simple extension of joint-bilateral upsampling incorporating color (or intensity) proximity weight in the "target" image. For improving this method so as to increase the spatial resolution, we introduce weight control based on local mutual information between target and control images. We confirmed that resulted images have higher spatial resolution and lower noise.

***Index Terms*—**Image upsampling, Bilateral filter, Trilateral Filter, Thermal image upsampling.

1. INTRODUCTION

Thermal images captured by Long Wave InfraRed (LWIR) cameras are useful for inspecting industrial plants, human health screening in airports, finding missing people or animals from air, night vision, and so on. However, LWIR has a severe limitation on spatial resolution, that is, high-resolutional uncooled LWIR camera cannot be produced by current technology. The other possibility to produce high-resolutional thermal image is upsampling. The upsampling can be done by a simple interpolation and joint-bilateral filtering[6], which produces a high-resolutional image from low resolutional "target" image and high resolutional "control" image. Figure 1 shows original thermal image (a), visible-light control image (geometrically calibrated by homography) (b), interpolation based upsampling (c), and bilateral upsampled image (d). As shown in Figure 1 (c), linear interpolation based upsampling image can have jaggy

(a) Original thermal image (640x480) (b) Visible light image (1440x1080)



(c) Linear-interpolation upsampling (1440x1080) (d) Joint-bilateral filter upsampling (1440x1080)

Figure 1. Two different approaches to upsampling problem.

or blurred edges. On the other hand, joint-bilateral filtered upsampling in Figure 1 (d) have much sharper edges but the tie under the sweater and the reflection on the wall are both blurred. This is because some thermal textures are not visible in the control image. For solving this problem, we propose trilateral upsampling, which uses both visible-light and thermal images as control images. Also, we further improve this method by locally changing the weight of these control images depending on the local mutual information.

2. RELATED WORKS

Bilateral filtering[1] is a well known edge-preserving smoothing, whose kernel (bilateral weight) is the product of

Gaussians in image plane and color space. This filtering is applied not only to edge preserving smoothing but also to many applications, e.g. depth map and saliency map upsampling[2], fusing flash and no-flash images[3], displaying images with high dynamic range[4], and so on. Other applications and extensions can be found in[5].

Bilateral filtering referring two images is called joint-bilateral filtering, where the input image is filtered (upsampled, smoothed) by referring the control image. The control image is used mainly for computing bilateral weights.

It is believed that the joint-bilateral filtering, using visible-light image as a control image, can be used for increasing the spatial resolution and reducing the noise[6].

A drawback of bilateral filtering is its high computational cost. For solving this problem, many acceleration methods have been proposed. Approximation using the fixed size Gaussian filter in extended high dimensional space[7], decomposition to linear kernels and parallel execution[8], approximation by fixed size Gaussian filters applied to 1D pixel arrays that is transformed according to their geodesic distances[9], and so on.

3. TRILATERAL UPSAMPLING

Let $\Omega = [0,1]^2$ be the unit square in 2D plane. In this square, we define the rectangular grid $\Omega_{n,m}$,

$$\Omega_{n,m} = \left\{\left(\frac{i}{n}, \frac{j}{m}\right) \middle| i = 0, \cdots, n,\ j = 0, \cdots, m\right\} \subset \Omega. \tag{1}$$

The problem scaling an image given on a grid $\Omega_{n,m}$ to that on grid Ω_{n_u,m_u} is usually called *upsampling* when the values $n_u > n$ and $m_u > m$.

We measure the distance on the grid $\Omega_{n,m}$ as

$$\|\boldsymbol{p} - \boldsymbol{q}\|^2 = n^2(p_x - q_x)^2 + m^2(p_y - q_y)^2, \tag{2}$$

where $\boldsymbol{p} = (p_x, p_y)^{\mathrm{T}}$, $\boldsymbol{q} = (q_x, q_y)^{\mathrm{T}}$ and $\boldsymbol{p}, \boldsymbol{q} \in \Omega$.

Suppose that two images $T: \Omega \mapsto \mathbb{R}^3$ and $V: \Omega \mapsto \mathbb{R}^3$ represent thermal and visible-light images of the same scene.

In these settings, while creating upsampled copy of the target image T we would like to use information provided by the control image V. The solution is usually represented in the form of joint-bilateral filtering [7] as

$$\hat{T}(\boldsymbol{p}) = \frac{1}{W(\boldsymbol{p})} \sum_{q \in N(\boldsymbol{p})} G_{\sigma_s}(\|\boldsymbol{p} - \boldsymbol{q}\|) G_{\sigma_c}(\|V(\boldsymbol{p}) - V(\boldsymbol{q})\|) \times \mathcal{I}(T)(\boldsymbol{q}),\ \boldsymbol{p} \in \Omega_{n_u,m_u}. \tag{3}$$

In this equation, we denote Gaussian kernel by G_σ, i.e.,

$$G_\sigma(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right), \tag{4}$$

and $\mathcal{I}$ represents some interpolation with the geometric transformation that corresponds pixels in V and T, i.e., $V(\boldsymbol{p})$ and $\mathcal{I}(T)(\boldsymbol{p})$ are the corresponding pixel values.

As usual, the weighted pixel values are summed in the filter window $N(\boldsymbol{p})$ centered at point $\boldsymbol{p}$.

Finally, the result is normalized by

$$W(\boldsymbol{p}) = \sum_{q \in N(\boldsymbol{p})} G_{\sigma_s}(\|\boldsymbol{p} - \boldsymbol{q}\|) G_{\sigma_c}(\|V(\boldsymbol{p}) - V(\boldsymbol{q})\|). \tag{5}$$

An example of bilateral upsampling is shown in Figure1. Original image obtained by the LWIR camera is shown in (a), the image of the scene obtained by the visible-light camera is (b). Upsampled image by Equation (3) is (d), where window size= 32×32, $\sigma_s = 25$, $\sigma_c = 10$. As mentioned in the introduction, we can easily see that some thermal image edges are not preserved in (d). To be more specific, we can check easily that the tie under the sweater and silhouette on the wall are not preserved in the detail. This means that Equation (3) is acting as a simple Gaussian filter on that area.

We can spot also that some details from visible-light image can be imprinted in the upsampled image. For example, the light sources on the celling are changed according to their shape on the visible-light image. Also there are imprinted details on the facial region.

We can improve the above results by extending the control image of the filtering. That is, we can use the visible-light control image as well as thermal control image. Our upsampling is given by

$$\hat{T}(\boldsymbol{p}) = \frac{1}{W(\boldsymbol{p})} \sum_{q \in N(\boldsymbol{p})} G_{\sigma_s}(\|\boldsymbol{p} - \boldsymbol{q}\|) G_{\sigma_c}(\|V(\boldsymbol{p}) - V(\boldsymbol{q})\|) \times G_{\sigma_o}(\|\mathcal{I}(T)(\boldsymbol{p}) - \mathcal{I}(T)(\boldsymbol{q})\|) \mathcal{I}(T)(\boldsymbol{q}), \tag{6}$$

Figure 2. Result of trilateral filtering using Equation (6). Window size=32×32, $\sigma_s = 25$, $\sigma_o = 10$, $\sigma_c = 10$.

for $\boldsymbol{p} \in \Omega_{n_u,m_u}$, where all the notation is the same as in Equation (3). Note that now we are modifying weights in the kernel with information from both images, which are captured by LWIR and visible-light cameras. Kernels G_{σ_c} and G_{σ_o} are responsible for this modification. As can be seen there are three Gaussian kernels involved in this filtering, so we call this procedure *trilateral filtering*.

The important property of this method is its robustness. When $V(\boldsymbol{p})$ is uniform, $G_{\sigma_c}(\|V(\boldsymbol{p}) - V(\boldsymbol{q})\|)$ will be constant and Equation (6) works as a simple bilateral filter and interpolation based upsampling. Compared with this, Equation (3) works as a blurring filter in this case. Also, even if $V(\boldsymbol{p})$ is not correlated with $\mathcal{J}(T)(\boldsymbol{p})$, smoothing over thermal texture will not occur by Equation(6).

Result of this filter is shown in Figure 2. Compared with Figure 1(d), we can easily notice that details such as tie under the sweater, silhouette on the wall are the same as those presented in the original image.

4. LOCAL TRILATERAL UPSAMPLING

Trilateral upsampling can avoid the damage caused by the uncorrelated image regions between target and control images. However, since this method uses globally fixed weights, it also suppresses positive aspect of the joint-bilateral upsampling, which produces high resolutional image. For combining the positive aspect of both, it is clear that we should control parameters G_{σ_c} and G_{σ_o} depending on the position $\boldsymbol{p}$. When thermal and visible-light images have strong relationship locally, then the color proximity G_{σ_c} of visible-light image should have bigger influence, otherwise intensity proximity G_{σ_o} of thermal image should be dominant.

For evaluating the relationship between two images, we employ mutual information. Given two discrete random variables X and Y, with probabilities P_X and P_Y respectively, and with joint probability $P_{X,Y}$, mutual information μ of the variables X and Y can be given as

$$\mu_{X,Y} = \sum_{x \in R(X)} \sum_{y \in R(Y)} P_{X,Y}(x,y) \log \frac{P_{X,Y}(x,y)}{P_X(x) P_Y(y)}, \tag{7}$$

where $R(X)$ and $R(Y)$ represent the domains of the random variables X and Y, respectively.

Figure 4 shows an example of local mutual information map between thermal and visible-light images.

For red, green and blue channels in the visible-light image, we interpret the values of the intensities in the given window $N(\boldsymbol{p})$ as a random variable Yr, Yg and Yb. We also interpret the intensity values of the thermal image in the given window $N(\boldsymbol{p})$ as random variable X. For color channel c, we compute the mutual information $\mu_{Yc,X}(\boldsymbol{p})$ between variables Yc and X at pixel $\boldsymbol{p}$. Based on the mutual information, standard deviations of the color channels and thermal image, at the pixel $\boldsymbol{p}$, are defined as

$$\sigma_{o,r}(\boldsymbol{p}) = \sigma\sqrt{1 + (\mu_{Y_r,X}(\boldsymbol{p}))^2}\,,\ \sigma_{c,r}(\boldsymbol{p}) = \sigma\sqrt{\frac{1+(\mu_{Y_r,X}(\boldsymbol{p}))^2}{\mu_{Y_r,X}(\boldsymbol{p})}}\,, \tag{8}$$

Figure 3. Local mutual information by Equation (7).

$$\sigma_{o,g}(\boldsymbol{p}) = \sigma\sqrt{1 + (\mu_{Y_g,X}(\boldsymbol{p}))^2}\,,\ \sigma_{c,g}(\boldsymbol{p}) = \sigma\sqrt{\frac{1+(\mu_{Y_g,X}(\boldsymbol{p}))^2}{\mu_{Y_g,X}(\boldsymbol{p})}}\,, \tag{9}$$

$$\sigma_{o,b}(\boldsymbol{p}) = \sigma\sqrt{1 + (\mu_{Y_b,X}(\boldsymbol{p}))^2}\,,\ \sigma_{c,b}(\boldsymbol{p}) = \sigma\sqrt{\frac{1+(\mu_{Y_b,X}(\boldsymbol{p}))^2}{\mu_{Y_b,X}(\boldsymbol{p})}}\,, \tag{10}$$

$$\sigma_o = \sqrt{\frac{1}{\frac{1}{{\sigma_{o,r}}^2} + \frac{1}{{\sigma_{o,g}}^2} + \frac{1}{{\sigma_{o,b}}^2}}}\,, \tag{11}$$

where $\sigma > 0$ is the hyper parameter of this method.

Finally, we can describe our filtering as follows

$$\begin{aligned} \hat{T}(\boldsymbol{p}) = \frac{1}{W(\boldsymbol{p})} \sum_{q \in N(\boldsymbol{p})} & G_{\sigma_s}(\|\boldsymbol{p} - \boldsymbol{q}\|) \\ & \times G_{\sigma_o}(\|\mathcal{J}(T)(\boldsymbol{p}) - \mathcal{J}(T)(\boldsymbol{q})\|) \\ & \times \tilde{G}_{\sigma_{c,r}\sigma_{c,g}\sigma_{c,b}}\big(V(\boldsymbol{p}) - V(\boldsymbol{q})\big)\mathcal{J}(T)(\boldsymbol{q}), \end{aligned} \tag{12}$$

for $\boldsymbol{p} \in \Omega_{n_u,m_u}$, where

$$\log \tilde{G}_{\sigma_{c,r}\sigma_{c,g}\sigma_{c,b}}(x) = -\frac{1}{2}\left(\frac{x_r^2}{\sigma_{c,r}^2} + \frac{x_g^2}{\sigma_{c,g}^2} + \frac{x_b^2}{\sigma_{c,b}^2}\right), \tag{13}$$

for $\boldsymbol{x} = (x_r, x_g, x_b)^T$.

It is clear that when the mutual information is high the standard deviation of the control image remains bounded, while the standard deviation of the thermal image becomes bigger. The result of local mutual information is shown in Figure 3.

Figure 4 shows the result of local trilateral upsampling. Compared with other upsampling methods, it can clearly be seen that noise level is quite small but also jags are reduced. Figure 5 shows the log histogram of horizontal derivative. The sharper peak of this histogram represents that the noise level is low. The wider support represents high-frequency component including textures and jags. Linear interpolation produces wider peak and narrower support. This means high noise level and high frequency component is blurred. Trilateral and local trilateral produce low noise level images. The support of local trilateral result is bit narrower than that of trilateral. This is because jags in trilateral result are relaxed by local parameter control.

Figure 4. Result of local trilateral filtering using Equation (12). Window size=32×32, $\sigma_s = 25$, $\sigma_o = 10$, $\sigma_c = 10$.

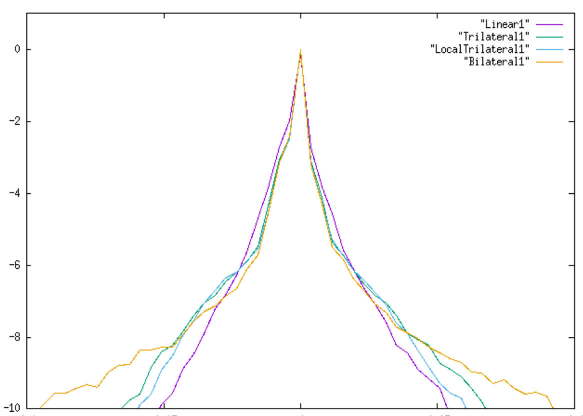
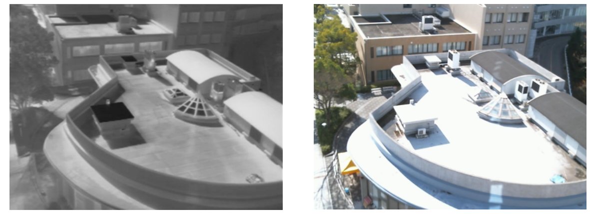


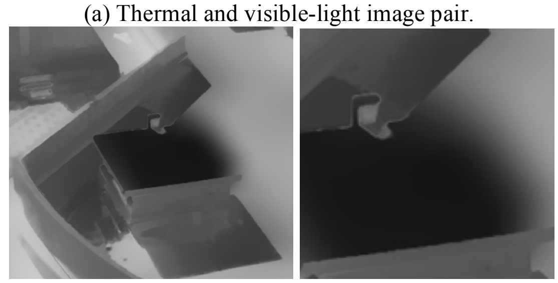
Figure 5. Log histogram of horizontal derivatives.

Other examples are shown in Figure 6.

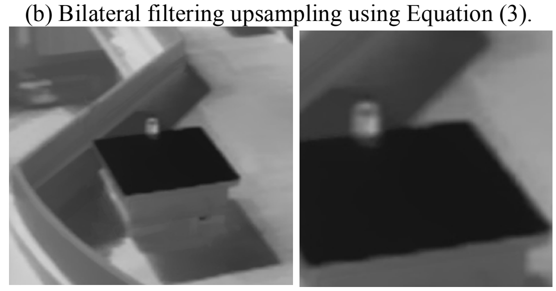
5. CONCLUSIONS



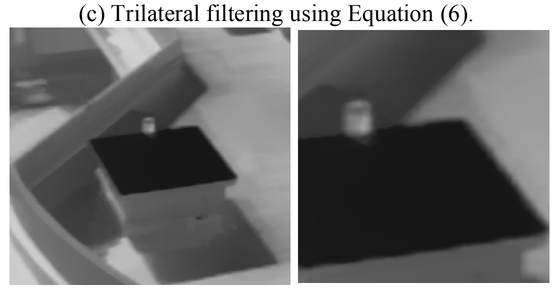
(a) Thermal and visible-light image pair.



(b) Bilateral filtering upsampling using Equation (3).



(c) Trilateral filtering using Equation (6).



(d) Local-trilateral filtering using Equation (8).

Figure 6. Other example comparing different upsampling methods.

In this paper, we presented the followings.

- Bilateral upsampling imprints control image on the filtered image. This breakdown may cause every image pairs having inconsistent edges or textures.
- Trilateral upsampling that uses spatial proximity weight, color similarity weight of control image, and intensity similarity weight of the target image. This setting prevents the breakdown.
- We further introduce local parameter control for improving spatial resolution. The local trilateral upsampling improves the noise and spatial resolution in the upsampled images.

Current local trilateral upsampling changes σ_o and σ_c depending on μ, but the mapping is defined in an ad hoc manner. The optimization of this mapping and the acceleration of this method will be done in the future works.

12. REFERENCES

[1] A C. Tomasi, R. Manduchi, “Bilateral Filtering for Gray and Color Images,” In Proc. of ICCV, Bombay, India, 1998

[2] K. Wang, L. Lin, J. Lu, C. Li, and K. Shi. “PISA: Pixelwise Image Saliency by Aggregating Complementary Appearance Contrast Measures with Edge-Preserving Coherence,” in IEEE Trans. Image Processing, vol. 24, no. 10, pp. 3019-3033, 2015.

[3] G. Petschnigg, M. Agrawala, H. Hoppe, R. Szeliski, M. Cohen, K. Toyama, “Digital Photography with Flash and No-Flash Image Pairs,” in Proc. of ACM SIGGRAPH, Vol 4, 664-672, 2004.

[4] F. Durand, J. Dorsey, “Fast Bilateal Filtering for the Display of High-Dynamic-Range Images,” in Proc. of ACM SIGGRAPH 2002, Vol 21, Issue 3, pp. 257-266, 2002.

[5] S. Paris, P. Kornprobst, J. Tumblin, F. Durand, “Bilateral Filtering: Theory and Applications,” Foundations and Treands in Computer Graphics and Vision, Vol 4, 1-73, 2009.

[6] J. Kopf, M.F. Cohen, D, Lischinski, M. Uyttendaele, “Joint Bilateral Upsampling,” in Proc. of ACM SIGGRAPH 2007, Vol 26, Issue 3, article 96, 2007.

[7] S. Paris and F. Durand, “A Fast Approximation of the Bilateral Filter using a Signal Processing Approach,” in Proc. of ECCV’06 , 2006

[8] Q. Yang, K.-H. Tan, and N. Ahuja, “Real-time O(1) bilateral filtering,” in Proc. of IEEE CVPR 2009, pp.557-564, 2009

[9] E. S.L. Gastal and M. M. Oliveira, “Domain Transform for Edge-Aware Image and Video Processing,” in Proc. of ACM SIGGRAPH 2011, Vol. 30, Issue 4, Article No. 69, 2011

[10] R. Szeliski, S.B. Kang, C.L. Zitnick, W.T. Freeman, “Automatic Estimation and Removal of Noise from a Single Image,” IEEE Trans. PAMI, Vol 30, No 2, pp. 299-314, 2008.

[11] H. Jinggang and D. B. Mumford, “Statistics of natural images and models,” In Proc. of CVPR’99 , pp. 541-547, Fort Collins, Colorado, 1999