Spatially-Varying Super-Resolution for HDTV

Chih-Tsung Shen, Hung-Hsun Liu, Ming-Sui Lee,
Yi-Ping Hung and Soo-Chang Pei
National Taiwan University

Submitted on Oct. 5, 2012.
Accepted on Dec. 24, 2012.

Outline

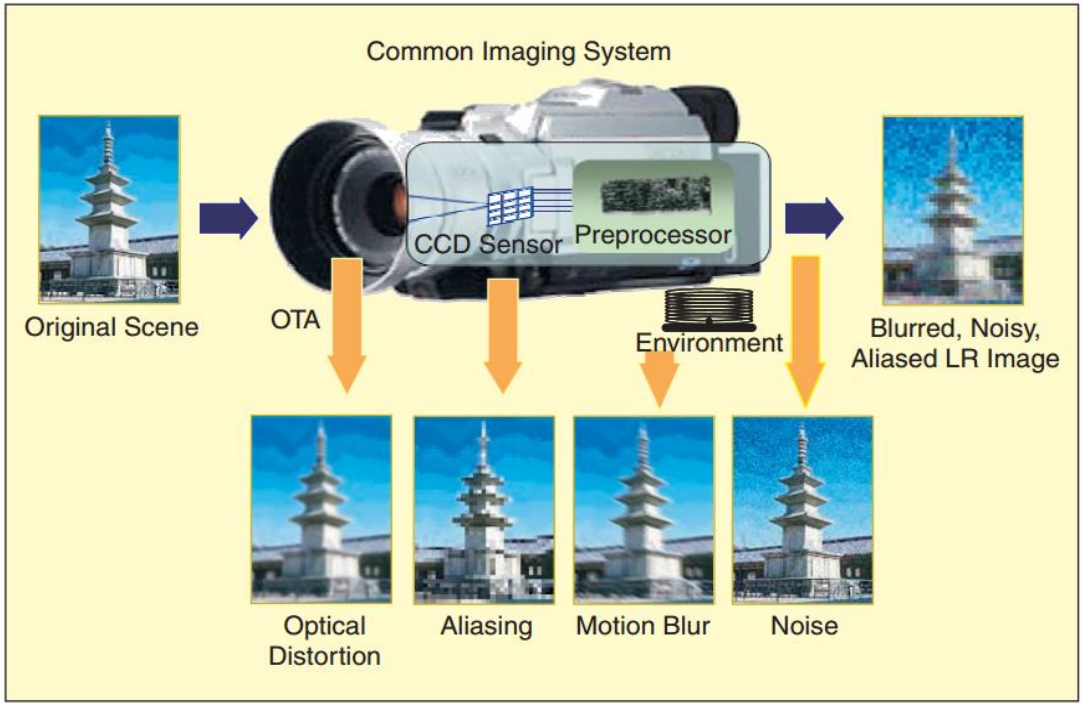
- Problem formulation
- Previous works
- Our System
  - Blur identification
  - L12TTV deblurring model
  - Pixel selection
- Experimental Results
- Conclusions

Problem Formulation

- According to the image formation process, we have

$$g(x,y) = D\{B(x,y) \otimes l(x,y)\} + n(x,y)$$

$g(x,y)$: observed image
$B(x,y)$: blur kernel**s**
$l(x,y)$: latent image
$n(x,y)$: noise
$D\{\}$: down-sample
$\otimes$: convolution



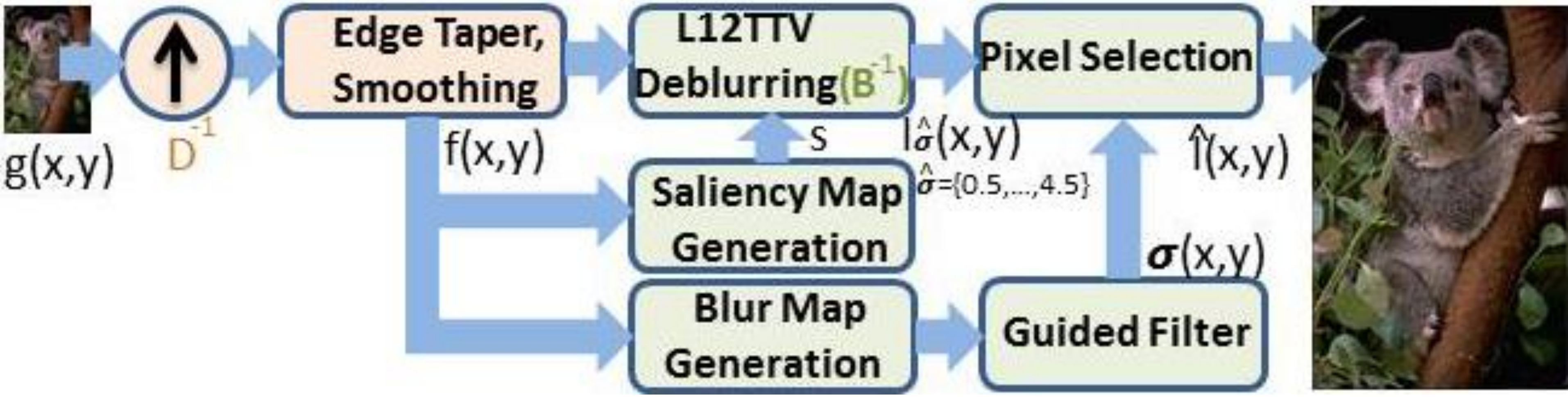
▲ *2. Common image acquisition system.*

IEEE Signal Processing Magazine, 2003

Previous Works

- Multiframe Super-Resolution [Tsai&Huang 84]
- Fast and Robust Multiframe + L1BTV [Farsiu et al. 04]
- Bandlet Super-Resolution [Mallat 06]
- Edge Statistics [Fattal 07]
- Fast Image/Video Upsampling [Shan et al. 08]
- Sparse Representation [Yang et al. 08]
- Self Similarity [Glanser&Irani 09]
- Local Self Example [Freedman&Fattal 11]

Our System: Overview

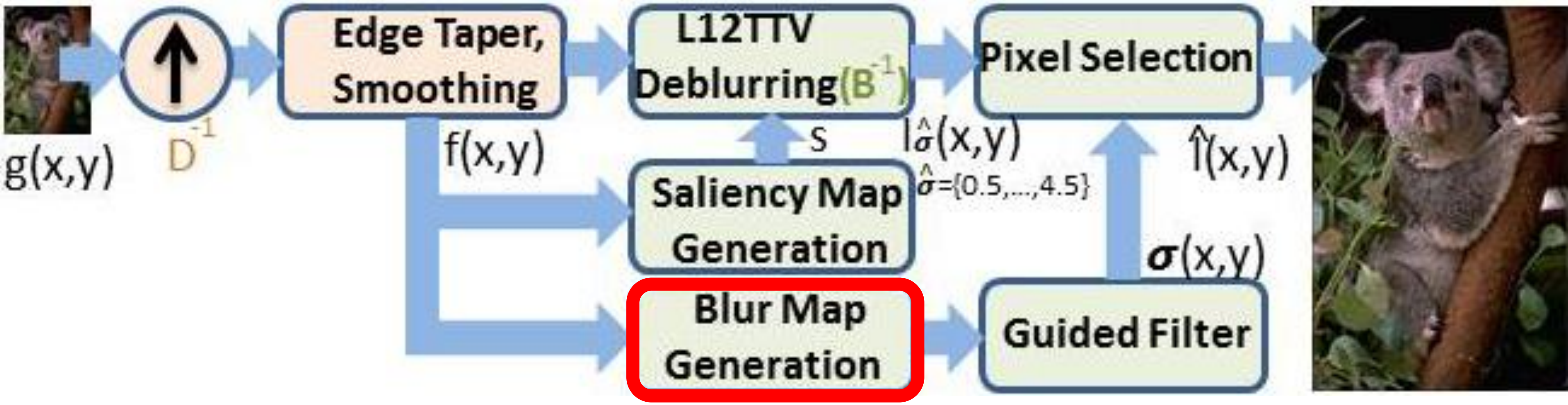


- What we have to do is to inverse the process!

$$g(x,y) = D\{B(x,y) \otimes l(x,y)\} + n(x,y)$$

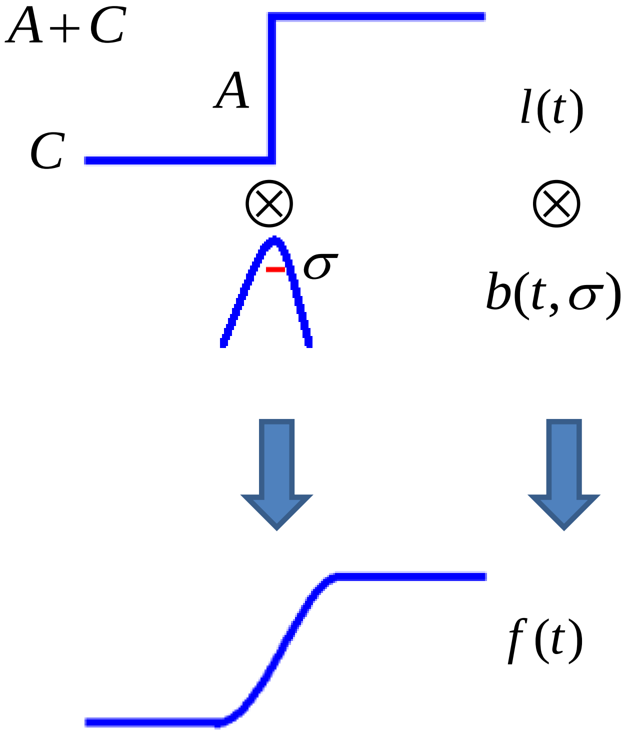
$$\hat{l}(x,y) = B^{-1}\{D^{-1}\{g(x,y) - n(x,y)\}\}$$

Our System: Blur Identification



Our System: Blur Identification

- Different to the previous works to deal with the single blur kernel (PSF), we deal with **spatially-varying** blur condition.



- Let's model the blurred edges.

$$l(t) = A \cdot u(t) + C$$

$$b(t,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} \cdot \exp(-\frac{t^2}{2\sigma^2})$$

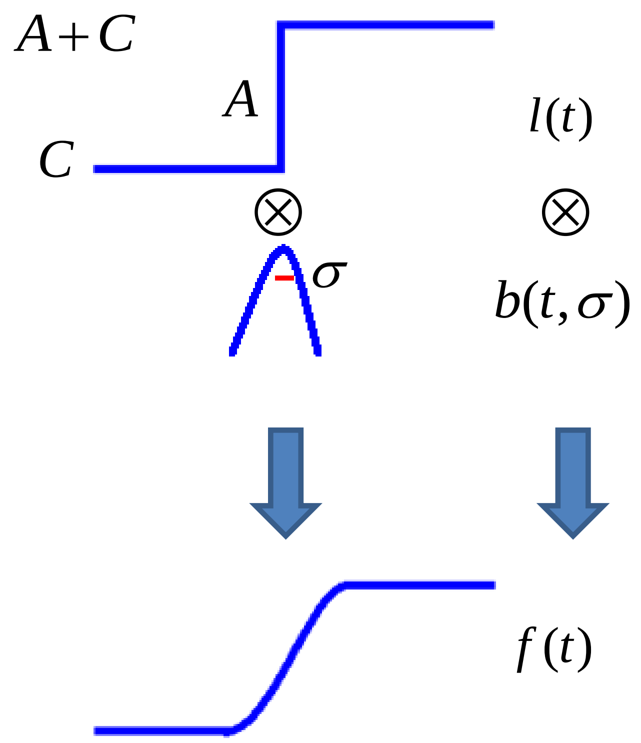
$$f(t) = l(t) \otimes b(t,\sigma)$$

Our System: Blur Identification

- Take 1st derivatives:

$$f(t) = l(t) \otimes b(t,\sigma)$$

$$\frac{\partial}{\partial t} f(t) = \frac{\partial}{\partial t}(l(t) \otimes b(t,\sigma)) = A \cdot b(t,\sigma)$$



- Edge Properties: $\max |\exp(0)| = 1$

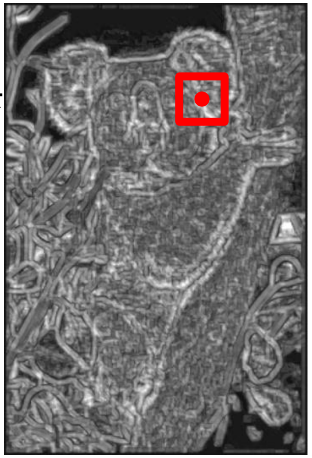
Our System: Blur Identification

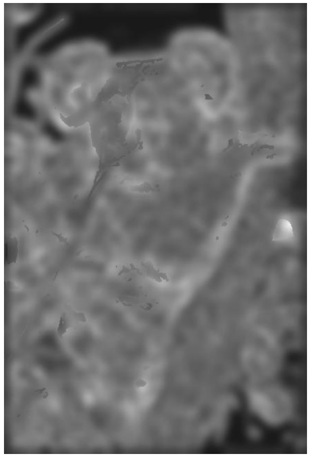
- We can adopt the local contrast prior:

$$\frac{\max\limits_{(x,y)\in(x',y')}\sqrt{\left(\frac{\partial}{\partial x}f(x,y)\right)^2+\left(\frac{\partial}{\partial y}f(x,y)\right)^2}}{\max\limits_{(x,y)\in(x',y')}f(x,y)-\min\limits_{(x,y)\in(x',y')}f(x,y)}\approx\frac{A\cdot\max|b(t,\sigma)|}{A+C-C}=\frac{1}{\sqrt{2\pi}\sigma(x,y)}$$

- We can obtain our blur map:

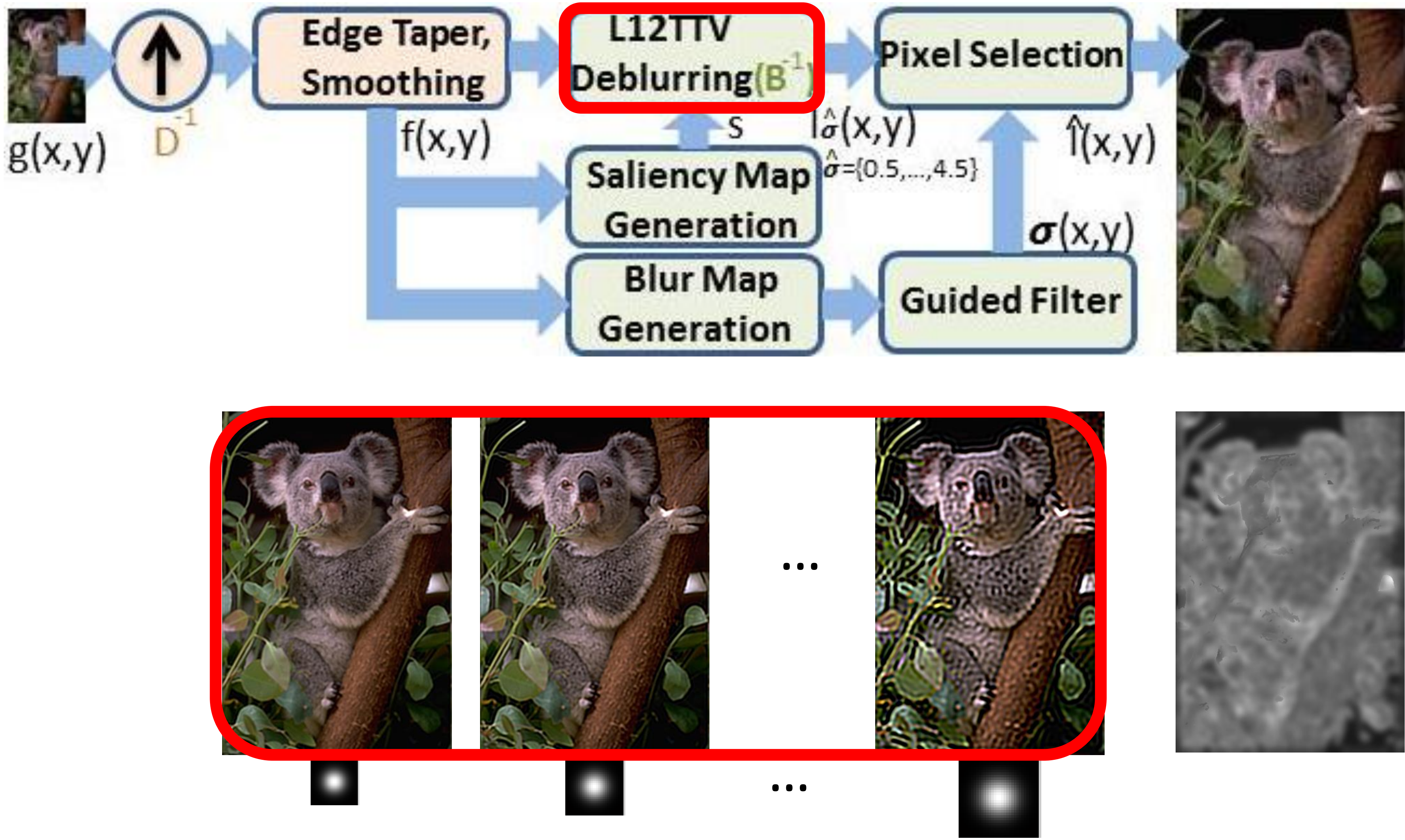
$$\sigma(x,y)\approx\frac{1}{\sqrt{2\pi}}\frac{\max\limits_{(x,y)\in(x',y')}f(x,y)-\min\limits_{(x,y)\in(x',y')}f(x,y)}{\max\limits_{(x,y)\in(x',y')}\sqrt{\left(\frac{\partial}{\partial x}f(x,y)\right)^2+\left(\frac{\partial}{\partial y}f(x,y)\right)^2}}$$





- Guided Image Filtering

Our System: L12TTV Deblurring



Our System: L12TTV Deblurring

- We can firstly interpolate the input image to desired size. In this paper, we adopt the **bicubic** interpolation with a smoothing operation. However, the **edge-directed interpolation** methods can also perform well.

$$
\begin{aligned}
& f(x, y) \\
& = D^{-1}\left\{g(x, y) - n(x, y)\right\} + \varepsilon(x, y) \\
& = B(x, y) \otimes l(x, y) + \varepsilon(x, y)
\end{aligned}
$$

2D Gaussian-like blurring kernel B(x,y) with its blur level σ(x,y):

$$
B(x, y) = \frac{1}{2\pi\sigma(x, y)^2} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma(x, y)^2}\right)
$$

Our System: L12TTV Deblurring

- We propose a L12TTV deblurring model with discrete B(x,y) (or $\sigma(x,y)=0.5,1.0,\ldots,4.5$).

$$\min_{l(x,y)} \gamma \sum_{(x,y)} |B(x,y)\otimes l(x,y) - f(x,y)| + \frac{\mu}{2}\sum_{(x,y)} |B(x,y)\otimes l(x,y) - f(x,y)|^2$$
$$+ s\sum_{(x,y)} |\nabla l(x,y)| + (1-s)\sum_{(x,y)} |\nabla l(x,y)|^2$$

$\sum_{(x,y)} |B(x,y)\otimes l(x,y) - f(x,y)|$: L1 fidelity, L1 data term

$\sum_{(x,y)} |B(x,y)\otimes l(x,y) - f(x,y)|^2$: L2 fidelity, L2 data term

$\sum_{(x,y)} |\nabla l(x,y)|$: Total variation

$\sum_{(x,y)} |\nabla l(x,y)|^2$: Tikhonov-like regularizer

Our System: L12TTV Deblurring

- It is not easy to solve the L1 fidelity and TV term. As a result, we approximate L1 fidelity to v(x,y) and TV term to w(x,y), respectively.

$$\min_{l,v,w} \frac{\gamma\alpha}{2}\sum_{(x,y)} |B(x,y)\otimes l(x,y) - f(x,y) - v(x,y)|^2$$

$$+\gamma\sum_{(x,y)} |v(x,y)| + \frac{\mu}{2}\sum_{(x,y)} |B(x,y)\otimes l(x,y) - f(x,y)|^2$$

$$+s\sum_{(x,y)} |w(x,y)| + \frac{s\beta}{2}\sum_{(x,y)} |w(x,y) - \nabla l(x,y)|^2$$

$$+(1-s)\sum_{(x,y)} |\nabla l(x,y)|^2$$

Equation

Our System: L12TTV Deblurring

- We adopt the alternating minimization to solve equation (★). Given l(x,y), we can have v(x,y) and w(x,y):

$$\min_{v,w} \frac{\gamma\alpha}{2} \sum_{(x,y)} | B(x,y) \otimes l(x,y) - f(x,y) - v(x,y) |^2$$
$$+\gamma \sum_{(x,y)} | v(x,y) | + s \sum_{(x,y)} | w(x,y) |$$
$$+\frac{s\beta}{2} \sum_{(x,y)} | w(x,y) - \nabla l(x,y) |^2$$

- Since v(x,y) and w(x,y) are independent, we have:

$$v(x,y) = \max\left\{ | B(x,y) \otimes l(x,y) - f(x,y) | - \frac{1}{\alpha}, 0 \right\} \cdot \frac{B(x,y) \otimes l(x,y) - f(x,y)}{| B(x,y) \otimes l(x,y) - f(x,y) |}$$

$$w(x,y) = \max\left\{ | \nabla l(x,y) | - \frac{1}{\beta}, 0 \right\} \cdot \frac{\nabla l(x,y)}{| \nabla l(x,y) |}$$

Our System: L12TTV Deblurring

- Since v(x,y) and w(x,y) are calculated, we can obtain l(x,y) from equation (★) .

$$\min_{l,v,w} \frac{\gamma\alpha}{2} \sum_{(x,y)} | B(x,y) \otimes l(x,y) - f(x,y) - v(x,y) |^2$$

$$+ \frac{\mu}{2} | B(x,y) \otimes l(x,y) - f(x,y) |^2$$

$$+ \frac{s\beta}{2} \sum_{(x,y)} | w(x,y) - \nabla l(x,y) |^2$$

$$+ (1-s) \sum_{(x,y)} | \nabla l(x,y) |^2$$

- Since B(x,y) is a Gaussian-like blur kernel, we can solve l(x,y) using **2D-FFTs**.

$$l(x,y) = F^{-1} \left\{ \frac{s\beta F^* \{\nabla\} F\{w\} + kF^* \{B\} F\{f\} + \gamma\alpha F^* \{B\} F\{v\}}{(s\beta + 2 - 2s) F^* \{\nabla\} F\{\nabla\} + kF^* \{B\} F\{B\}} \right\}$$

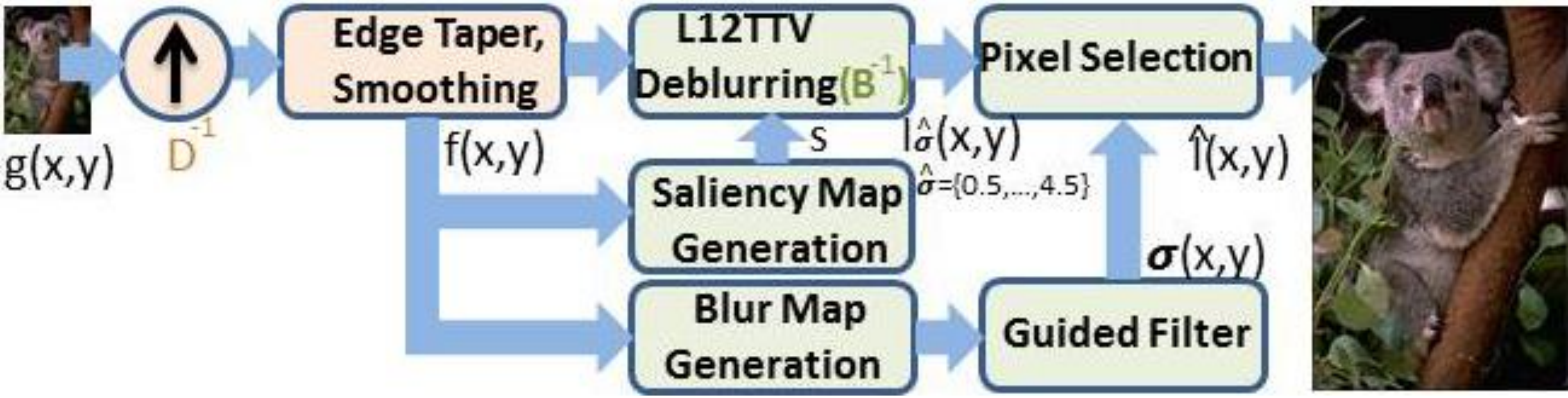
Our System: L12TTV Deblurring

- s is the weighting parameter. We let it equal to the proportion of the Salient pixels inside image.

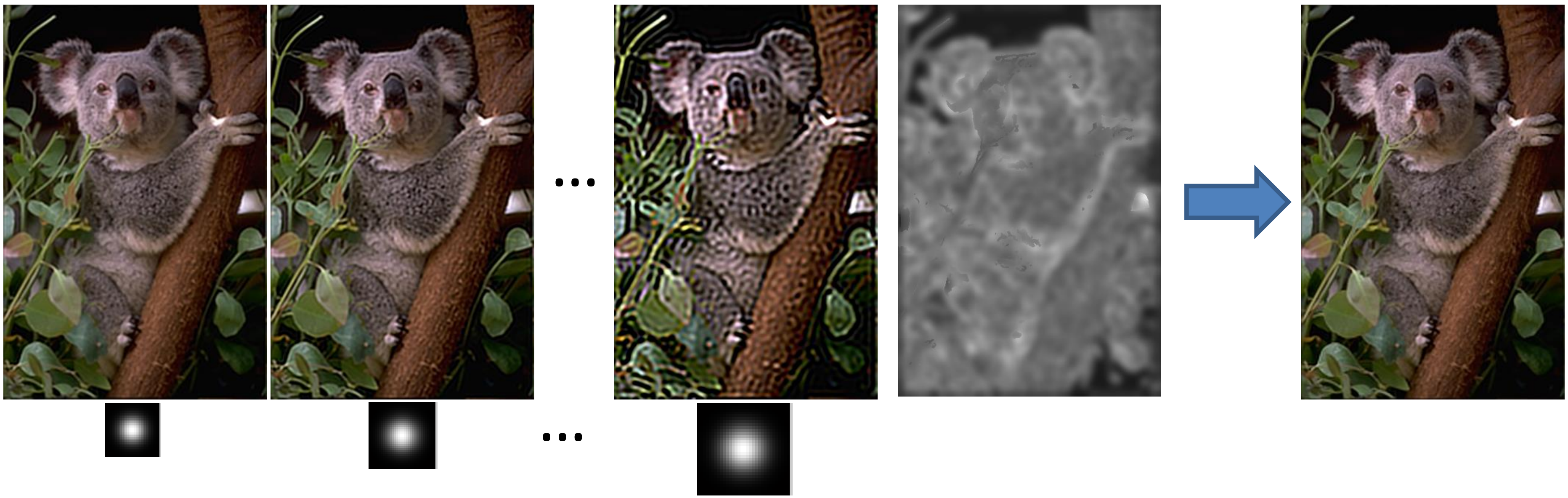
$$Sal(x,y)=\sqrt{\sum_{I,M} Lm(x,y)^2}$$

- If Sal(x,y) is high, we may meet the texture areas so as to deblur stronger (on TV term).

Our System: Pixel Selection



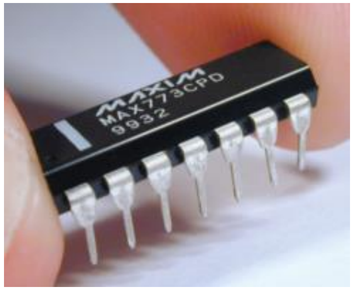
Our System: Pixel Selection

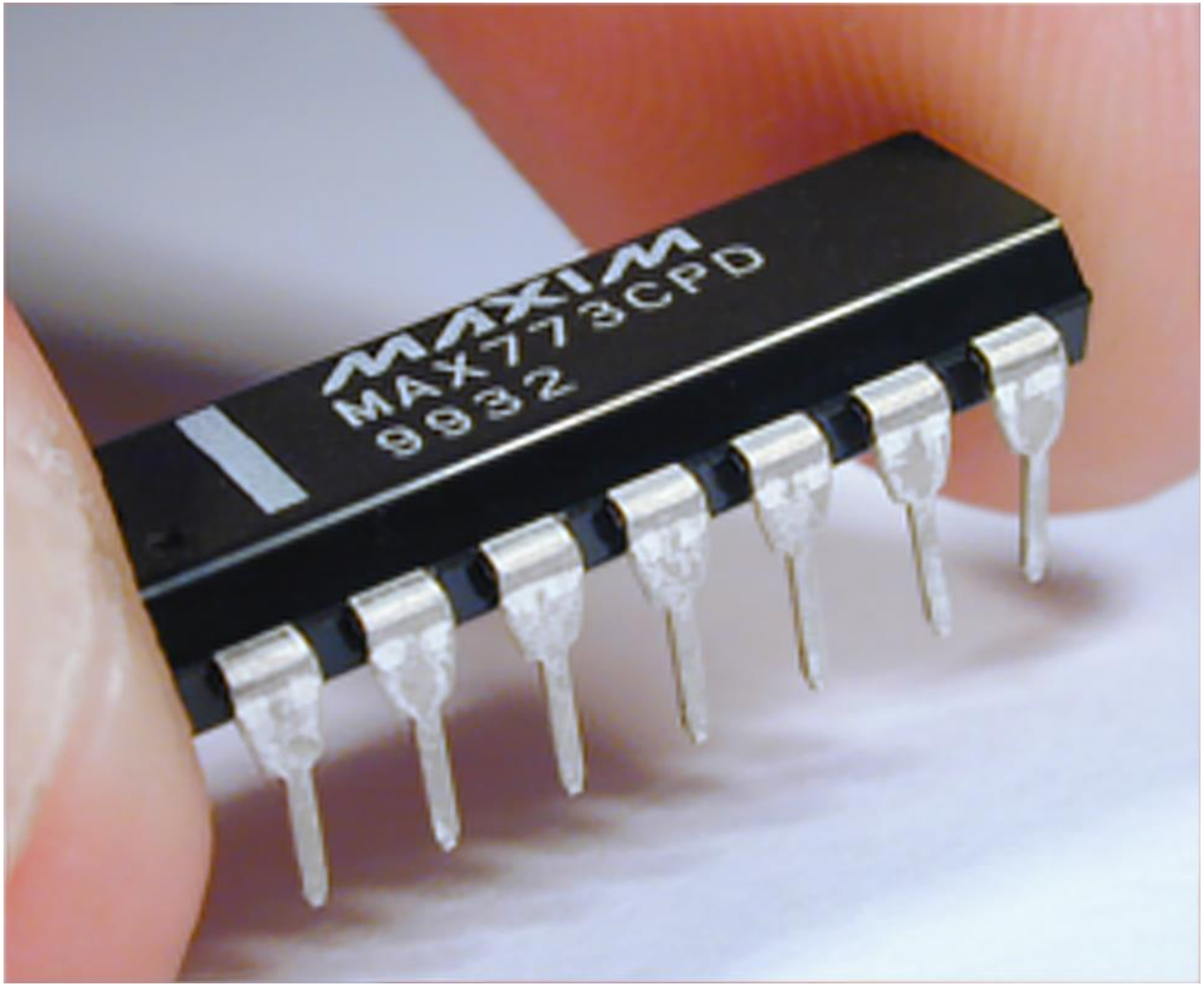


- We use the blur map as the index and reconstruct the latent HD image.

$$\hat{l}(x,y) = \sum_{(x,y)} l_{\hat{\sigma}(x,y)}(x,y)$$

Experimental Results





(a) input (b)ours

Experimental Results

(a) bicubic (b) Fattal's (c) ours

Experimental Results

(a) input (b) bicubic (c) GeniueFractals2012 (d) ours

Experimental Results

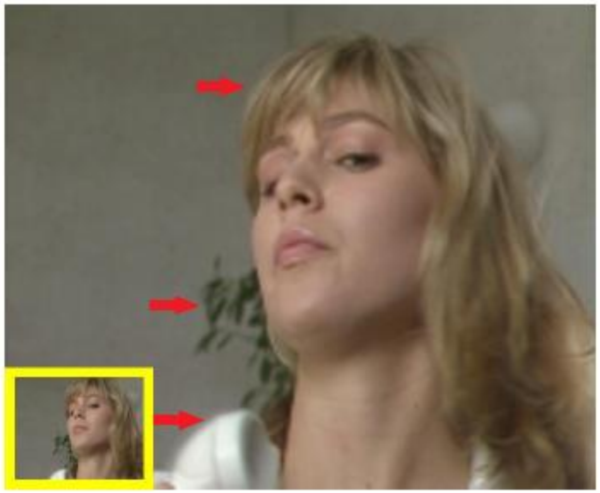
(a) input

(b) bicubic

(c) GeniueFractals2012

(d) ours

Experimental Results



(a) original

(b) Video Enhancer

(c) Shan et al.'s

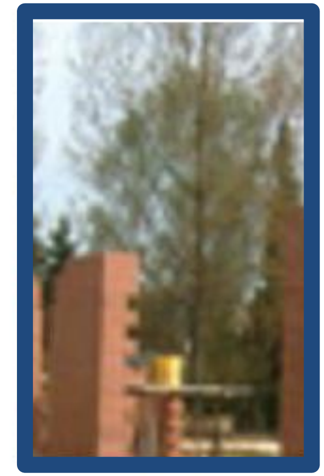
(d) ours

Experimental Results

(a) original

(b) QELabs

(e) bicubic



(f) QELabs

(c) Shan et al.'s

(d) ours

(g)Shan

(h) ours

Experimental Results

Experimental Results

Experimental Results

Experimental Results

Experimental Results

Experimental Results

Experimental Results

Experimental Results

Conclusions

- We propose a **Spatially-Varying** scheme to Super-Resolution for HDTV.
- Different to previous works, we deal with spatially-varying blur condition.
- We propose a **L12TTV** deblurring model.
- **Pixel Selection** ensures our final HD output.

- WeiBo : ShenChihTsung
- Gmail : shenchihtsung@gmail.com