Accelerating the Distribution Estimation for the Weighted Median/Mode Filters

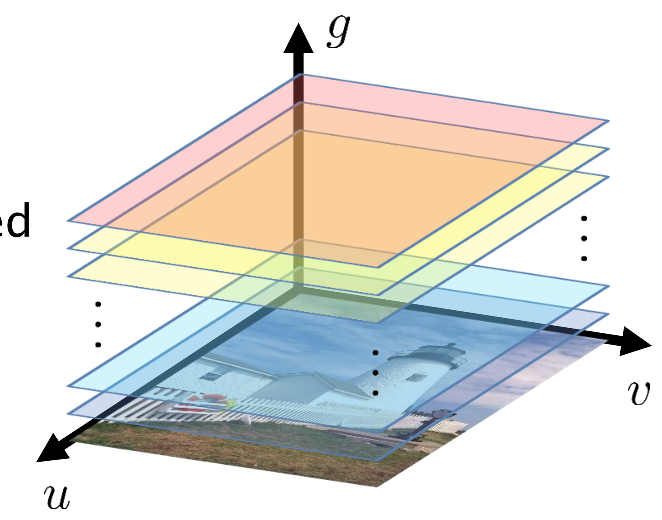
Lu Sheng *King Ngi Ngan*, Fellow, IEEE *Tak-Wai Hui*
Department of Electronic Engineering, the Chinese University of Hong Kong
{lsheng, knngan, twhui}@ee.cuhk.edu.hk

ABSTRACT

Various image filters for computer vision applications require the properties of the local statistics of the input image, which are always defined by the local distribution or histogram. But the huge expense of computing the distribution hampers the popularity of these filters in real-time or interactive systems. In this paper, we present an efficient and practical method to estimate the local weighted distribution for the weighted median/mode filters based on the kernel density estimation with a new separable kernel defined by a weighted combination of a series of probabilistic generative models. It reduces the number of filtering operations in previous constant-time algorithms [1,2], which is also adaptive to the structure of the input image. The proposed accelerated weighted median/mode filters are effective and efficient for a variety of applications, and have comparable performance against the current state-of-the-art counterparts and cost only a fraction of their execution times.

MOTIVATION

- Conventional weighted distribution estimation
  $h(\mathbf{x},g) = \frac{1}{\mathrm{Z}(\mathbf{x})}\sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} w(\mathbf{x},\mathbf{y})\phi_{\mathbf{x}}(g,f_{\mathbf{y}})$
- The median of the distribution needs to be tracked
  $\mathcal{C}(\mathbf{x},\hat{g}) = \frac{1}{\mathrm{Z}(\mathbf{x})}\sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} w(\mathbf{x},\mathbf{y})\cdot\int_{-\infty}^{\hat{g}}\phi_{\mathbf{x}}(g,f_{\mathbf{y}})\,dg$
- The mode requires the quantity
  $\frac{\partial h(\mathbf{x},g)}{\partial g}\Big|_{g=\hat{g}} = \frac{1}{\mathrm{Z}(\mathbf{x})}\sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} w(\mathbf{x},\mathbf{y})\cdot\frac{\partial\phi_{\mathbf{x}}(g,f_{\mathbf{y}})}{\partial g}\Big|_{g=\hat{g}}$



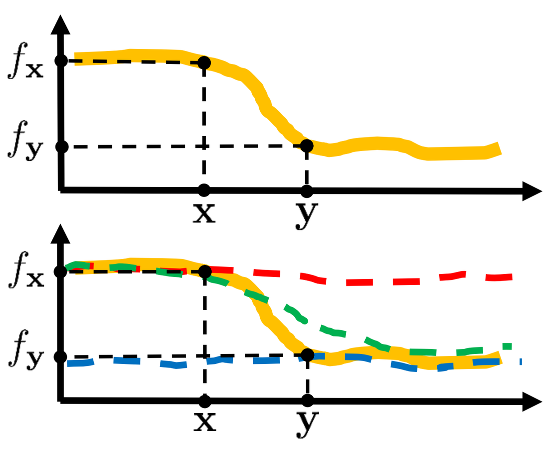
$\mathbf{x}=(u,v)$	pixel location
$\phi_{\mathbf{x}}(g,f_{\mathbf{y}})$	data kernel
$\mathcal{N}(\mathbf{x})$	local window
$w(\mathbf{x},\mathbf{y})$	weight at $\mathbf{x}$ & $\mathbf{y}$
$\mathrm{Z}(\mathbf{x})$	normalization factor

- It needs many samples g to output a good estimation
- Each sample requires a non-trivial filtering operation to output the probability at its location
- The distribution estimation and related median or mode seeking are time-consuming

PROPOSED METHOD

Novel Kernel Definition

- Assume the input image is modeled by several *models*
  - Each model is governed by $p(\eta_{\mathbf{x}}|l), l\in\mathcal{L}$
  - They represent the distinct local structures
- Two pixels are similar in a model if they both have high probabilities that agree with this model as



$$\kappa^l(f_{\mathbf{x}},f_{\mathbf{y}}) = p_{\mathbf{x}}(f_{\mathbf{x}}|l)p_{\mathbf{y}}(f_{\mathbf{y}}|l) = \int_{\eta_{\mathbf{x}}\in\mathcal{H}_{\mathbf{x}}} p(f_{\mathbf{x}}|\eta_{\mathbf{x}})p(\eta_{\mathbf{x}}|l)d\eta_{\mathbf{x}}\times\int_{\eta_{\mathbf{y}}\in\mathcal{H}_{\mathbf{y}}} p(f_{\mathbf{y}}|\eta_{\mathbf{y}})p(\eta_{\mathbf{y}}|l)d\eta_{\mathbf{y}}$$

- The overall kernel is defined as a weighted combination

$$\kappa(f_{\mathbf{x}},f_{\mathbf{y}}) = \sum_{l=1}^{L}\kappa^l(f_{\mathbf{x}},f_{\mathbf{y}})p_{\mathbf{x},\mathbf{y}}(l) = \sum_{l=1}^{L}p_{\mathbf{x}}(f_{\mathbf{x}}|l)p_{\mathbf{y}}(f_{\mathbf{y}}|l)\underline{p_{\mathbf{x},\mathbf{y}}(l)}$$ [Model compatibility prior]

Probability Distribution Approximation

$$\tilde{h}(\mathbf{x},g) \propto \sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} w(\mathbf{x},\mathbf{y})\sum_{l=1}^{L}p_{\mathbf{x}}(g|l)p_{\mathbf{y}}(f_{\mathbf{y}}|l)p_{\mathbf{x},\mathbf{y}}(l) = \sum_{l=1}^{L}p_{\mathbf{x}}(g|l)\cdot\psi_{\mathbf{x}}(l)$$

$$\psi_{\mathbf{x}}(l) = \sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} w(\mathbf{x},\mathbf{y})p_{\mathbf{y}}(f_{\mathbf{y}}|l)p_{\mathbf{x},\mathbf{y}}(l)$$

A joint filtering with the guidance of
- the joint image/feature map
- and the compatibility prior

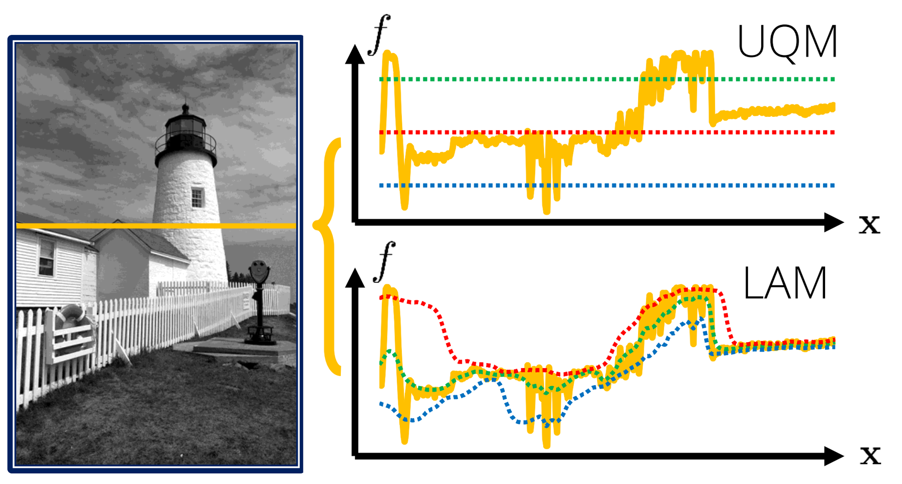
- A mixture of L densities
- Pre-compute $\psi_{\mathbf{x}}(l)$ requires only L filtering operations, each of which is independent of g

Gaussian Models for the Kernel

- Uniformly Quantized Models (UQM)
- Locally Adaptive Models (LAM)
  - Hierarchically segment [3] pixels according to different local structures
  - Model parameters are generated as

$$\theta^l_{\mathbf{y}} = 1_{[\mathbf{y}\in\mathcal{S}_l]},\ \mathrm{W}^l_{\mathbf{x}} = \sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})}\theta^l_{\mathbf{y}}$$
$$\mu^l_{\mathbf{x}} = \frac{1}{\mathrm{W}^l_{\mathbf{x}}}\sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})}\theta^l_{\mathbf{y}}f_{\mathbf{y}}$$
$$\Sigma^l_{\mathbf{x}} = \frac{1}{\mathrm{W}^l_{\mathbf{x}}}\sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})}\theta^l_{\mathbf{y}}f_{\mathbf{y}}f_{\mathbf{y}}^\top - \mu^l_{\mathbf{x}}\mu^{l\top}_{\mathbf{x}}$$

$\mathcal{S}_l$ is the set of pixels in the l^{th} model



- Model Compatibility Prior

$$p_{\mathbf{x},\mathbf{y}}(l) = \exp\left(-\frac{1}{2}\left(\mu^l_{\mathbf{x}}-\mu^l_{\mathbf{y}}\right)^\top\Sigma_n^{-1}(\mu^l_{\mathbf{x}}-\mu^l_{\mathbf{y}})\right)$$

- Kernel for the l^{th} Model [data noise variance]

$$\kappa^l(f_{\mathbf{x}},f_{\mathbf{y}}) = N(f_{\mathbf{x}}|\mu^l_{\mathbf{x}},\Sigma_n+\Sigma^l_{\mathbf{x}})N(f_{\mathbf{y}}|\mu^l_{\mathbf{y}},\Sigma_n+\Sigma^l_{\mathbf{y}})$$

- Filters to calculate $\psi_{\mathbf{x}}(l)$
  - Guided image filter [4]
  - Domain transform filter [5]

ACCELERATING THE WEIGHTED FILTERS

Weighted Median Filter

- Find the median of $\tilde{h}(\mathbf{x},g)$
  - Estimate the cumulative probability at each model's mean as $\tilde{\mathcal{C}}(\mathbf{x},\mu^l_{\mathbf{x}}), \forall\, l\in\mathcal{L}$
  - Find two adjacent cumulative probabilities $\tilde{\mathcal{C}}(\mathbf{x},\mu^k_{\mathbf{x}})\le 0.5$ and $\tilde{\mathcal{C}}(\mathbf{x},\mu^{k+1}_{\mathbf{x}}) > 0.5$
  - Linearly interpolate these two cumulative probabilities as

$$g^{\mathrm{med}}_{\mathbf{x}} \approx \frac{0.5-\tilde{\mathcal{C}}(\mathbf{x},\mu^k_{\mathbf{x}})}{\tilde{\mathcal{C}}(\mathbf{x},\mu^{k+1}_{\mathbf{x}})-\tilde{\mathcal{C}}(\mathbf{x},\mu^k_{\mathbf{x}})}(\mu^{k+1}_{\mathbf{x}}-\mu^k_{\mathbf{x}})+\mu^k_{\mathbf{x}}$$

Weighted Mode Filter

- Find the global mode of $\tilde{h}(\mathbf{x},g)$
  - Fixed-point iteration with respect to g letting $\partial\tilde{h}(\mathbf{x},g)/\partial g = 0$, as

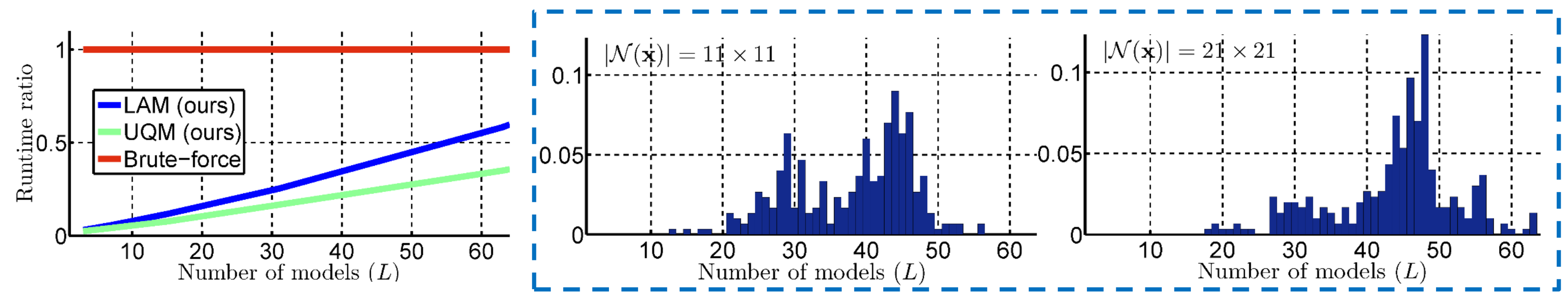
$$g^{n+1}_{\mathbf{x}} = \left(\sum_{l=1}^{L}\mathcal{B}^l_{\mathbf{x}}(g^n_{\mathbf{x}})\left(\Sigma_n+\Sigma^l_{\mathbf{x}}\right)^{-1}\right)^{-1}\left(\sum_{l=1}^{L}\mathcal{B}^l_{\mathbf{x}}(g^n_{\mathbf{x}})\left(\Sigma_n+\Sigma^l_{\mathbf{x}}\right)^{-1}\mu^l_{\mathbf{x}}\right)$$

where $\mathcal{B}^l_{\mathbf{x}}(g^n_{\mathbf{x}}) = N(g^n_{\mathbf{x}}|\mu^l_{\mathbf{x}},\Sigma_n+\Sigma^l_{\mathbf{x}})\psi_{\mathbf{x}}(l)$

  - Initialization: $g^0_{\mathbf{x}} = \mu^{m^\star}_{\mathbf{x}}$ where $m^\star = \arg\max_m\sum_{l=1}^{L}\mathcal{B}^l_{\mathbf{x}}(\mu^m_{\mathbf{x}})$

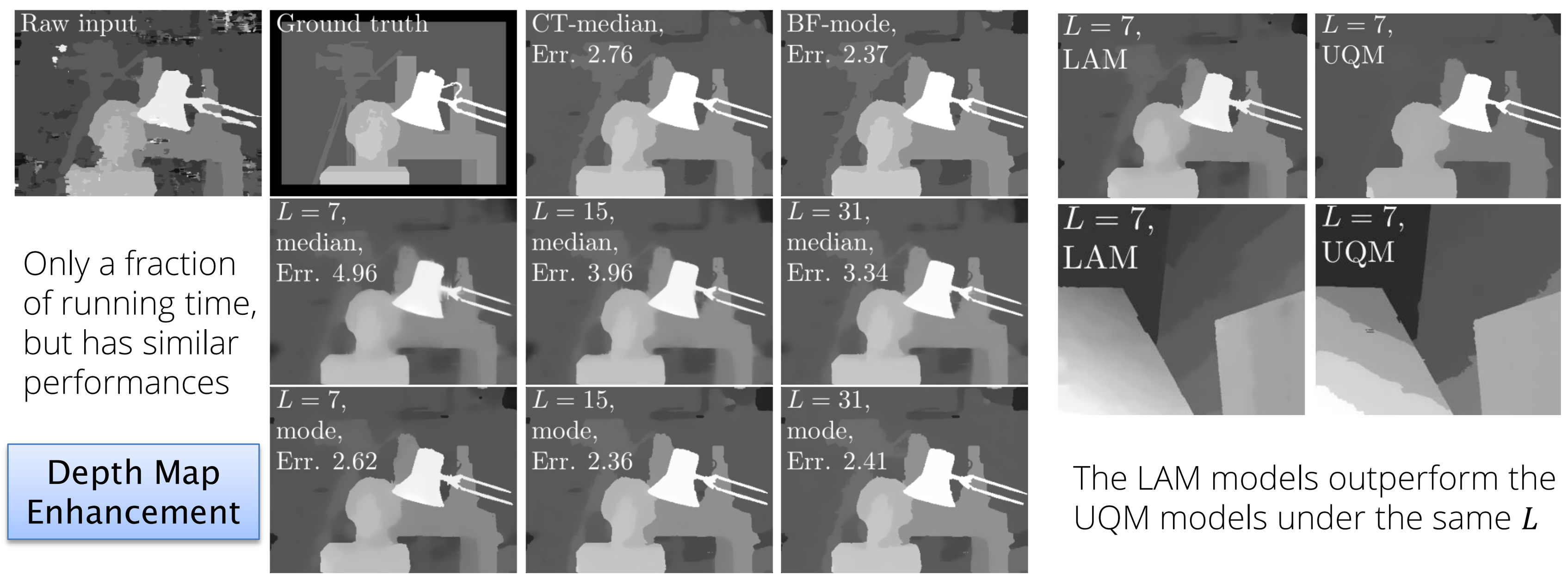
EXPERIMENTAL RESULTS & DISCUSSIONS

Performance Evaluation



- Runtime comparison on the distribution estimation
- Less than 0.5 of that of Brute-force when $L = 50$
- The number of Necessary LAM models in **BSD300** dataset under different window sizes
- 2~3x faster for grayscale image, 6~9x faster for color image, even further speedup for disparity or cartoon style images due to high structure homogeneity

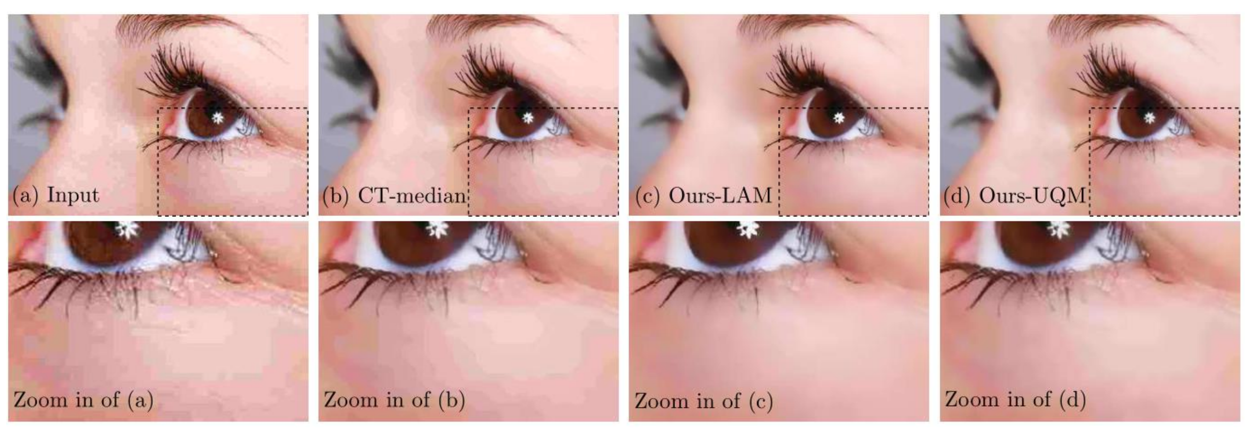
Applications



Only a fraction of running time, but has similar performances

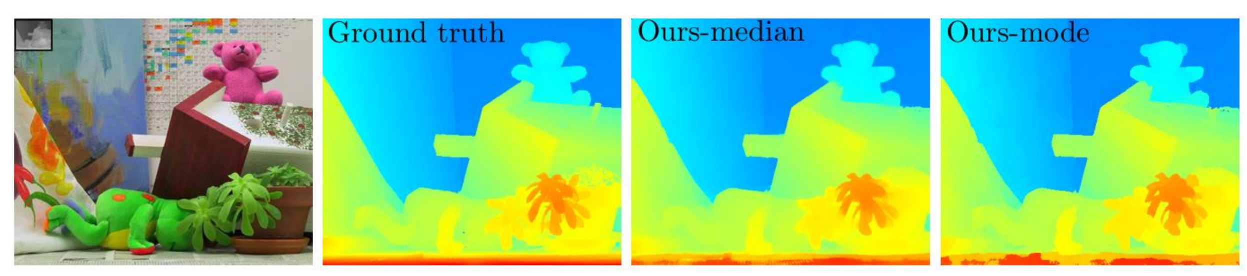
Depth Map Enhancement

The LAM models outperform the UQM models under the same L



1) Our filter exploiting LAM models outputs piecewise smooth result but also preserves the structure
2) Other methods give piecewise constant results

JPEG Artifact Removal



Joint upsample & smooth a 1/8 downsampled noisy depth map

Joint Upsamping & Noise Reduction

Weighted median filter under the LAM models

Detail Enhancement

[1] Z. Ma *et al.*, Constant time weighted median filtering for stereo matching and beyond. In Proc. ICCV. 2013.
[2] D. Min *et al.*, Depth video enhancement based on weighted mode filtering, In IEEE TIP. 2012
[3] E.S.L. Gastal and M.M.Oliveira, Adaptive Manifolds for Real-Time High-Dimensional Filtering. In ACM TOG. 2012
[4] K. He *et al.*, Guided image filtering. In Proc. ECCV. 2010
[5] E.S.L. Gastal and M.M.Oliveira, Domain transform for edge-aware image and video processing. In ACM TOG. 2011