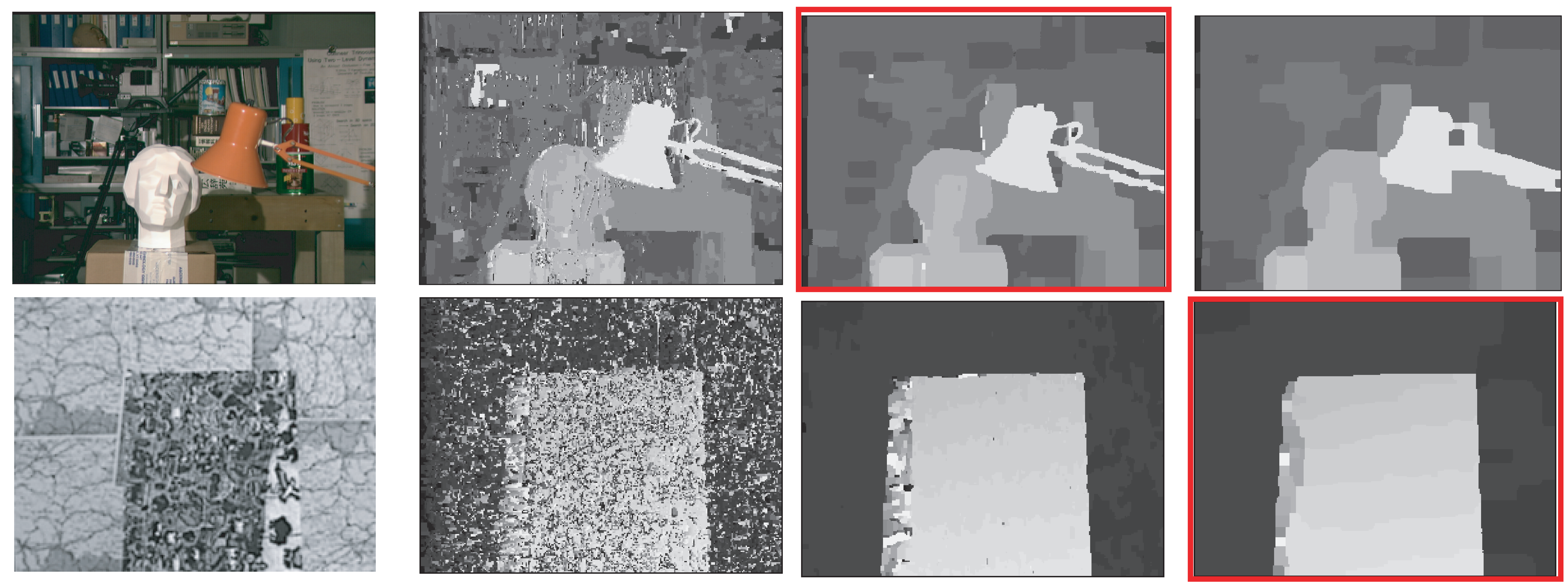
Parameter Estimation for MRF Stereo

Li Zhang **Steve Seitz**
University of Washington, Seattle, WA

Problem

Different stereo pairs require different MRF parameters for better performance. For example, regularization λ=10 gives the best result for the Tsukuba pair, but λ=40 is better for the map pair. ***How to set optimal paramters of MRF stereo algorithms for each stereo pair?*** Specifically, design a wrapper that can be used with existing stereo code.



λ=1 λ=10 λ=40

Introduction

Many stereo algorithms compute a disparity map by minimizing a Markov Random Field (MRF) energy:

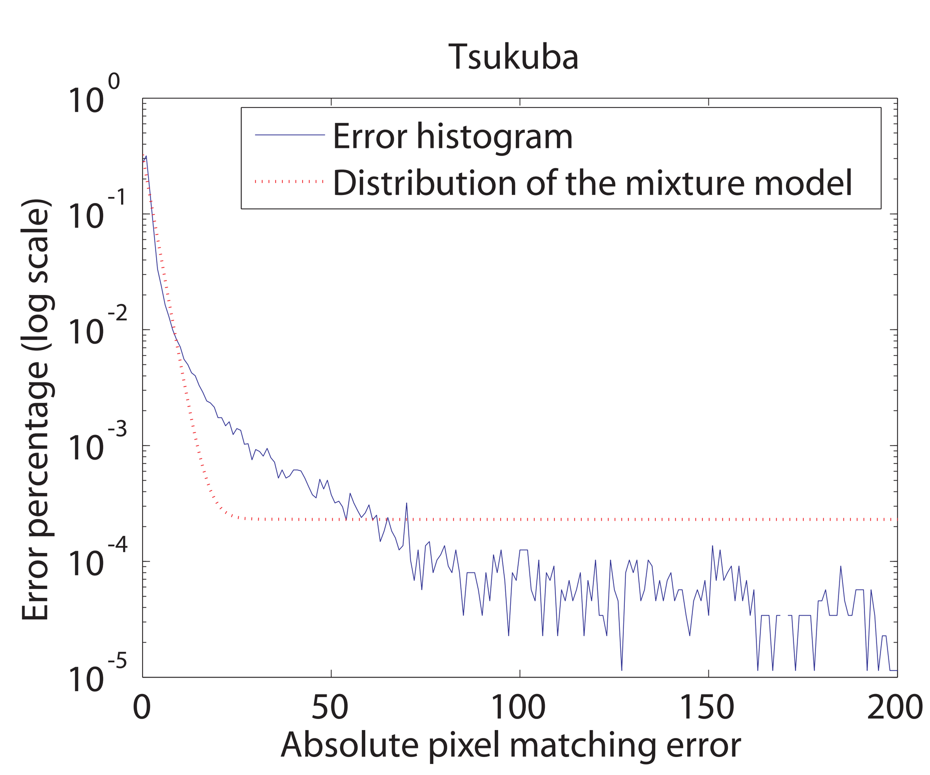
$$\sum_{i\in\mathcal{I}} U(d_i) + \lambda \sum_{(i,j)\in\mathcal{G}} V(d_i, d_j)$$

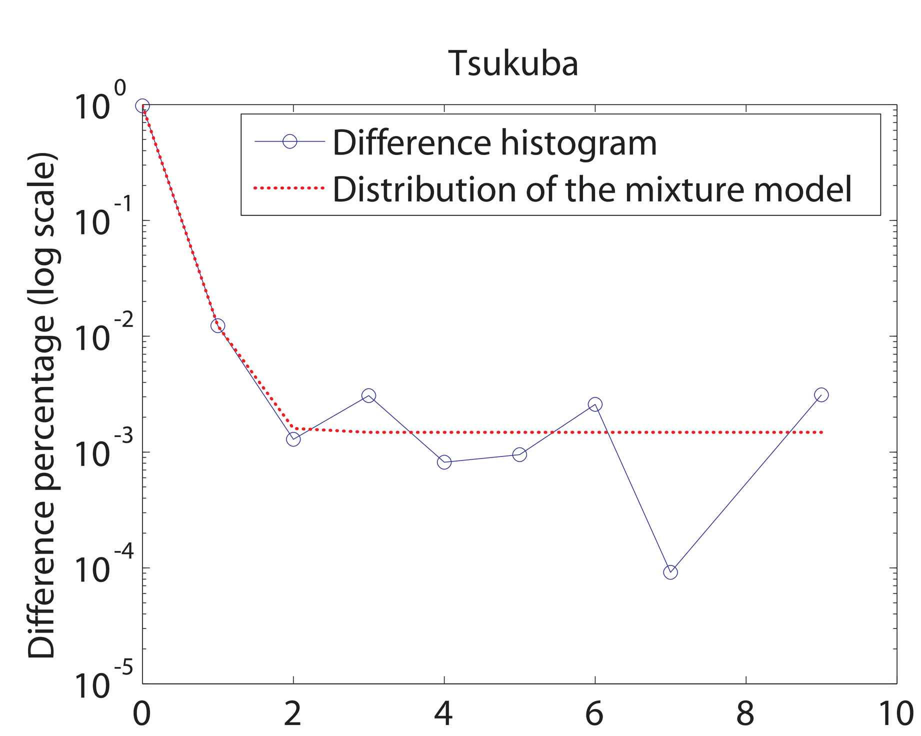
We focus on truncated linear for U and V terms:

$$\begin{aligned} V(d_i,d_j) &= \min(|d_i - d_j|, \tau) \\ U(d_i) &= \min(|I(x_i,y_i) - J(x_i - d_i, y_i)|, \sigma) \end{aligned}$$

Parameters to estimate:

- σ -- based on expected error of corresponding pixels
- τ -- based on expected disparity differences between neighbors
- λ -- based on expected smoothness, dependent on local intensity gradients





Approach

We derive MRF stereo parameters by fitting distribution models to the histograms of *pixel matching errors* and *neighboring disparity differences.*
We use *mixtures of exponential distribution and uniform outlier* for both histograms.

Initialize σ, τ, λ

Estimate disparity map using an existing stereo matcher

Compute pixel matching errors and neighboring disparity differences

Fit mixture models to the histograms of matching errors and disparity differences

Compute σ, τ, λ from mixture model parameters

We iterate between estimating parameters and stereo matching until convergence.

Details

Mixture model for pixel matching errors

$$P(e(d_i)|d_i,\gamma_i) = \begin{cases} \zeta e^{-\mu|e(d_i)|}, & \gamma_i = 1. \\ \frac{1}{N}, & \gamma_i = 0. \end{cases}$$

$$P(\gamma_i = 1) = \alpha$$

Mixture model for neighboring disparity differences

$$P(\Delta d_g|\theta_g) = \begin{cases} \eta e^{-\nu|\Delta d_g|}, & \theta_g = 1. \\ \frac{1}{L}, & \theta_g = 0. \end{cases}$$

$$P(\theta_g = 1) = \beta$$

Intensity gradient cues

$$P(\Delta I_g|\theta_g) = \begin{cases} \xi e^{-\kappa|\Delta I_g|}, & \theta_g = 1. \\ \frac{1}{K}, & \theta_g = 0. \end{cases}$$

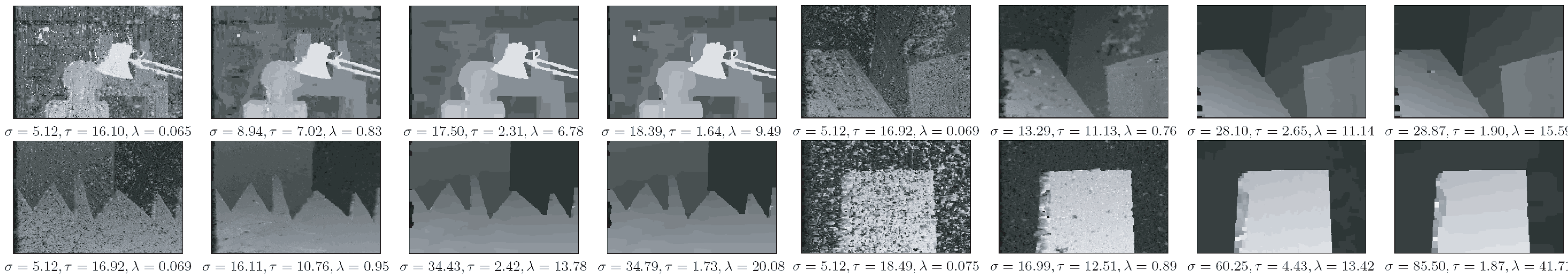
From mixture models to MRF parameters: $\sigma = \frac{t_d}{s_d}$, $\tau = \frac{t_p}{s_p}$, $\lambda = \frac{s_p}{s_d}$, where

$$s_d = \frac{\alpha\zeta\mu}{\alpha\zeta + (1-\alpha)\frac{1}{N}} \qquad t_d = \log(1 + \frac{\alpha\zeta N}{1-\alpha}) \qquad s_p = \frac{\beta\eta\nu}{\beta\eta + (1-\beta)\frac{1}{L}} \qquad t_p = \log(1 + \frac{\beta\eta L}{1-\beta})$$

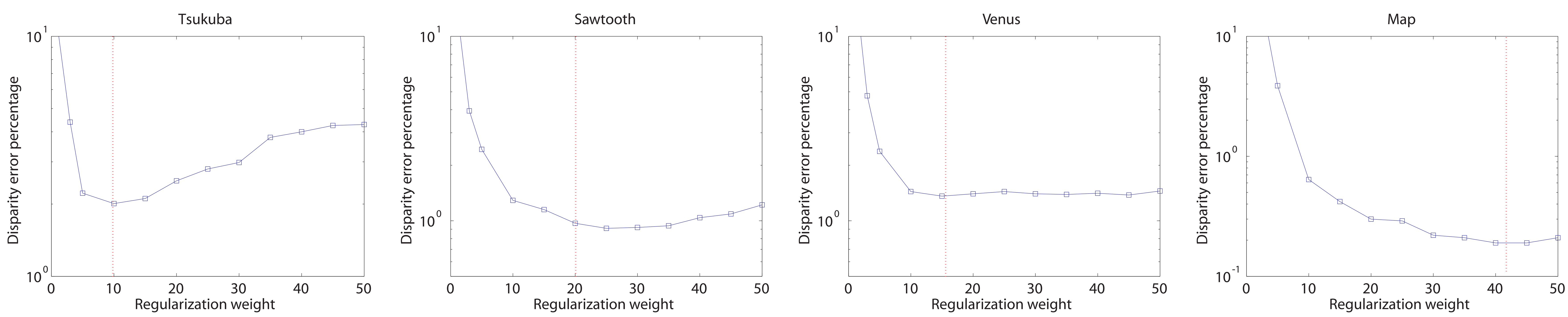
To include the intensity gradient cue, $s_p = \frac{\beta\xi\eta\nu e^{-\kappa|\Delta I_g|}}{\beta\xi\eta e^{-\kappa|\Delta I_g|} + \frac{1-\beta}{KL}}$ $t_p = \log(1 + \frac{\beta\xi\eta K L e^{-\kappa|\Delta I_g|}}{1-\beta})$

Results

Stereo matching with Belief Propagation



Convergence on the four Middlebury benchmarks.

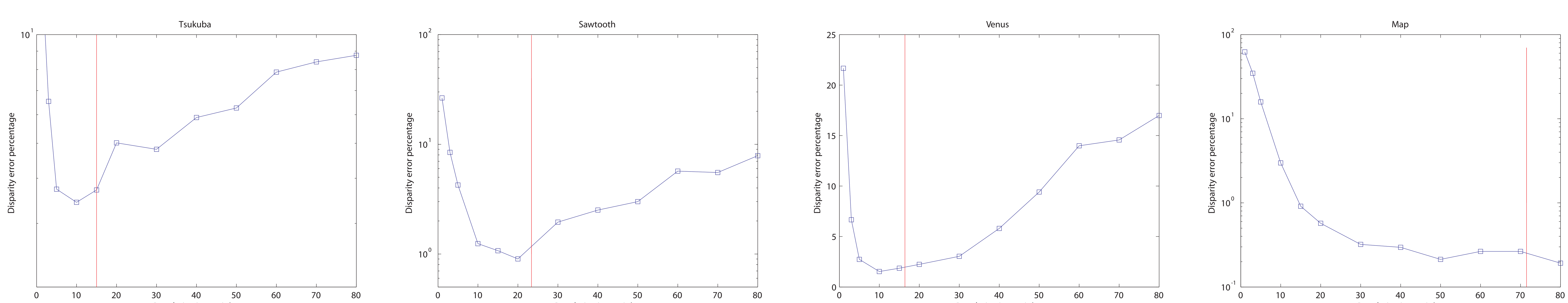


Graphs of error rate with respect to ground truth as a function of regularization weight λ while fixing (σ,τ). The vertical dotted lines are our estimates for λ.

Parameter setting	Overall ranking	Tsukuba			Sawtooth			Venus			Map	
		all	untex.	disc.	all	untex.	disc.	all	untex.	disc.	all	disc.
Adaptive+Grad	5	1.87 16	0.67 10	7.13 7	0.83 13	0.32 19	3.48 7	1.53 15	0.92 8	10.37 17	0.20 3	2.20 1
Fixed+Grad 2	11	1.84 14	1.05 13	9.87 17	0.87 13	0.28 17	5.78 15	1.22 12	1.05 9	13.39 21	0.20 3	2.49 3
Adaptive	13	2.12 17	1.36 17	10.76 17	0.97 13	0.31 19	6.79 16	1.33 13	1.13 10	14.65 24	0.18 2	2.20 1
Fixed	19	1.84 14	1.33 16	10.02 17	1.24 18	0.32 19	7.18 19	1.34 13	1.18 11	15.17 24	0.38 12	3.47 7
Fixed+Grad 1	31	7.68 32	5.76 28	11.79 17	5.92 35	0.30 18	13.12 25	6.90 34	3.20 25	14.11 23	20.19 36	35.41 36

Performance comparison of fixed and adaptive stereo solvers. ***Our method improves a baseline stereo matcher by 6 slots on the Middlebury rankings***. Fixed: (σ,τ,λ) = (10,2,10). Adaptive: estimated (σ,τ,λ). Fixed+Grad1: estimated (α,μ,β,ν), fixed κ = 1. Fixed+Grad 2: estimated (α,μ,β,ν), fixed κ = 0.01. Adaptive+Grad: estimated (α,μ,β,ν,κ).

Stereo matching with Graph Cuts (Potts model)



Graphs of error rate with respect to ground truth as a function of regularization weight λ while fixing σ. The vertical dotted lines are our estimation for λ.

Related work

> Prior work on MRF parameter estimation in the image processing literature not easily extended to the stereo problem.
> Cheng & Caelli 04 proposed a more restricted approach for stereo MRF parameter estimation, using MCMC.
> Statistical learning for improving vision algorithms [Freeman et al.] require training images.

Contribution

> A novel approach estimates optimal MRF parameters for each stereo pair
> Boosts the proformance by 6 slots on the Middlebury rankings
> Works as a "wrapper" around existing stereo code