Fast Guided Global Interpolation for Depth and Motion

Yu Li, Dongbo Min, Minh N. Do, Jiangbo Lu
Introduction

**Depth upsampling** and **motion interpolation** are often required to generate a dense, high-quality, and high resolution depth map or optical flow field.

Low-res & noisy depth (ToF)  

High-res. color guidance  

High-res. depth

Depth upsampling (color guided)  

Input TOF depth: noisy, low resolution, regularly distributed
Introduction

Depth upsampling and motion interpolation are often required to generate a dense, high-quality, and high resolution depth map or optical flow field.

Color frame and sparse matches from DM

Dense optical flow field

Data density < 1%

Motion interpolation
Input matches: typically reliable, but highly scattered, varying density

[DM: Weinzaepfel et al., "DeepFlow: Large displacement optical flow with deep matching, ICCV 2013."]
Motivation

Existing methods are often tailored to one specific task:

<table>
<thead>
<tr>
<th>Depth upsampling</th>
<th>JBF [Kopf et al. 2007], MRF+nlm [Park et al. 2011], TGV [Ferstl et al. 2013], JGU [Liu et al. 2013], AR [Yang et al. 2014], Data-driven [Kwon et al. 2015], etc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motion interpolation</td>
<td>EpicFlow [Revaud et al. 2015], [Drayer and Brox 2015], [Leordeanu et al. 2013], etc</td>
</tr>
</tbody>
</table>

The common objective for both tasks is to **densify** a set of **sparse data points**, either regularly distributed or scattered, to a full image grid through a 2D **guided interpolation** process.

**Our approach:** Fast Guided Global Interpolation (FGI)

A unified approach that casts the guided interpolation problem into a hierarchical, global optimization framework.
Several Challenges e.g.

- Texture-copy artifacts due to inconsistent structures
- Large occlusions, long-range propagation and extrapolation
- Loss of thin structures, missing motion boundaries
- Complex algorithms and computationally inefficient
Single-scale WLS Method Vs. Our Method

**Depth**
- Color guidance
- WLS
- Our result
- Ground truth

**Optical flow**
- Color guidance
- WLS
- Our result
- Ground truth

[FGS: Min et al., "Fast global image smoothing based on weighted least squares," TIP 2014.]
Our Pipeline: Overview

- A hierarchical (coarse-to-fine), multi-pass guided interpolation framework
- Divide the problems into a sequence of interpolation tasks each with smaller scale factors
- Gradually fill the large gap between the sparse measurement and the dense data
Our Pipeline: Filtering with Alternating Guidances

- From the coarse level \( l = L-1 \), we upsample the signal by a factor of 2 at each level by solving the following weighted least square (WLS) using the recent FGS solver.

- **Guided interp.**:
  \[
  \mathcal{E}(d_*) = (d_* - d_l)^\top M_l (d_* - d_l) + \lambda_1 d_*^\top A_{c_l} d_*
  \]
  - The color image \( c_l \) as the guidance
  - \( A_{c_l} \) is the spatially varying Laplacian matrix defined by \( c_l \)

[FGS: Min et al., "Fast global image smoothing based on weighted least squares," TIP 2014.]

Why FGS? 100 ms for filtering 1MPixels RGB images on 1 CPU core. More details in our paper.
Our Pipeline: Filtering with Alternating Guidances

- Next, another WLS is solved with the output $d_*$ as guidance and bicubic interpolated signal as input.
- **Joint filtering:**
  \[ E(\tilde{d}_l) = (\tilde{d}_l - d_\circ)^\top (\tilde{d}_l - d_\circ) + \lambda_2 \tilde{d}_l^\top A_{d_*} \tilde{d}_l \]
  - The intermediate interpolated map $d_*$ as the guidance
  - $A_{d_*}$ is the spatially varying Laplacian matrix defined by $d_*$
Our Pipeline: Consensus-Based Data Augmentation

- Then, check the **consistency** between the output and the bicubic upsampled data, and pick the most consistent points to add to the data mask map \( \tilde{m}_l \)
  - The bicubic upsampled data is free from texture-copying
  - Proceed in a non-overlapping patch fashion (2x2 patches)
- The entire process is repeated until the finest level \((l = 0)\) is reached.
1D Scanline Illustration

Guidance

Ground truth

One-pass WLS

Ground truth
1D Scanline Illustration
Pipeline Validation on Depth Upsampling

- Single scale WLS
- +Cascaded filtering with alternating guidances - single scale (Sec. 3.1)
- +Hierarchical process
- +Consensus-based data point augmentation (Sec. 3.2)

Depth Upsampling Error (MAD)
Depth Upsampling Results

Average runtime to upsample a $272 \times 344$ depth to $1088 \times 1376$ (in seconds)

<table>
<thead>
<tr>
<th>Method</th>
<th>2X</th>
<th>4X</th>
<th>8X</th>
<th>16X</th>
<th>Average runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRF+nlm</td>
<td>170</td>
<td>420</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>900</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GF</td>
<td>1.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLMF</td>
<td>2.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGI (ours)</td>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1000x faster than AR

Average Depth Upsampling Error on ToF Synthetic Dataset (6 cases)

Close to AR
Our framework also improves other edge-aware smoothing filters, e.g. the guided filter

<table>
<thead>
<tr>
<th>Depth avg. error</th>
<th>2x</th>
<th>4x</th>
<th>8x</th>
<th>16x</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-pass GF</td>
<td>1.31</td>
<td>1.54</td>
<td>2.04</td>
<td>3.12</td>
</tr>
<tr>
<td>GF in our framework</td>
<td>1.06</td>
<td>1.21</td>
<td>1.63</td>
<td>2.59</td>
</tr>
</tbody>
</table>

[GF: He et al., “Guided image filtering,” ECCV 2010.]
Depth Upsampling Results

ToF Synthetic Dataset


Depth

1000x faster than AR

Error map

ToFMark Dataset


650x faster than TGV
Motion Interpolation Results

Performance (EPE) on MPI Sintel training set

<table>
<thead>
<tr>
<th></th>
<th>WLS</th>
<th>EpicFlow-NW</th>
<th>EpicFlow-LA</th>
<th>FGI (ours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean</td>
<td>3.23</td>
<td>3.17</td>
<td>2.65</td>
<td>2.75</td>
</tr>
<tr>
<td>Final</td>
<td>4.68</td>
<td>4.55</td>
<td>4.10</td>
<td>4.14</td>
</tr>
<tr>
<td>Runtime (sec)</td>
<td>0.21</td>
<td>0.80</td>
<td>0.94</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Close to EpicFlow, but over 2x faster

Performance (EPE) on the MPI Sintel testing benchmark

|--------|----------------|-------------|-------------|------------|------------|------------|--------------|----------------|----------------|
Conclusion

• General & versatile technique:
  ▪ Tackle both depth and motion interpolation tasks, and potentially more
  ▪ Generally applicable to other edge-aware smoothing filters, e.g. GF

• Competitive results while running much faster than task-specific state-of-the-art methods

• Simple & effective:
  ▪ No color edge detection & variational minimization in [Revaud et al., CVPR’15]
  ▪ No domain transform filtering for post-smoothing in [Barron & Poole, ECCV’16]

• Further acceleration on GPUs and FPGA, offering a common engine for guided interpolation

Project page (code is available):
http://publish.illinois.edu/visual-modeling-and-analytics/fast-guided-global-interpolation/