Fast Bilateral-Space Stereo for Synthetic Defocus

Team PIX-IT

Mentor TA : Meher Shashwat Nigam

Team Members:

- Apoorva Thirupati (2019121012 - DD CS)
- Balavarun P(2020701012 - MS CS)
- Mohd Omama (2020701006 - MS CS)
- Jhanvi Shingala (2020701008 - MS CS)

GitHub Repo Link

A General Taxonomy of Stereo Algorithms

[3] A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms

A vast majority of stereo algorithms use a subset of following approaches:

- Stereo Matching
- Cost Aggregation
- Disparity Optimization

<u>Local Methods:</u>
Stereo Matching + Cost Aggregation

<u>Global Methods:</u>
Stereo Matching + Disparity Optimization

<u>**This Paper**:</u>
Stereo Matching + Disparity Optimization in Bilateral Space

High Level Overview

- Get disparity map and confidence from Stereo Matching

- Develop technique to make Bilateral Filtering fast. This technique helps us to go from Pixel to Bilateral Space.

- Solve for depth in Bilateral Space

Fast Bilateral Grid

- The normal bilateral grid formation equation - timetaking & slow

$$A_{i,j} = \exp\left(-\frac{\|[x_i, y_i] - [x_j, y_j]\|^2}{2\sigma_{xy}^2} - \frac{\|[r_i, g_i, b_i] - [r_j, g_j, b_j]\|^2}{2\sigma_{rgb}^2}\right)$$

- Faster way to calculate bilateral grid, A, represented as product of sparse matrices:

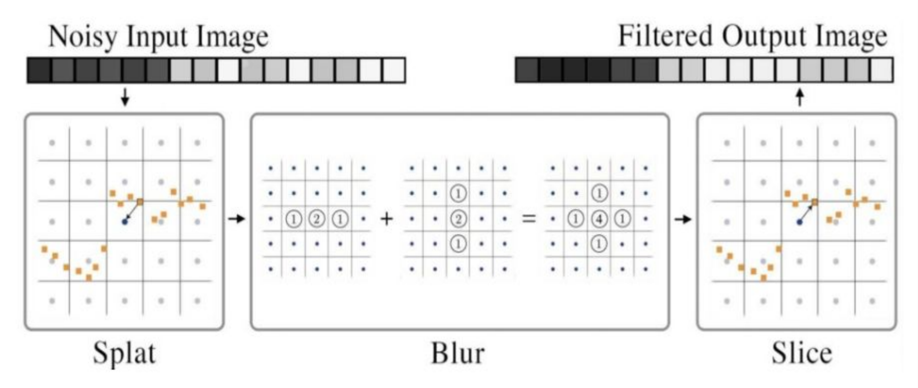
$$A \approx S^{\mathrm{T}} \bar{B} S$$

- Where S is the “splat”, multiplication by B is the “blur” and by S transpose is the “slice”.

Fast Bilateral Grid

Splat Matrix

- Image pixels mapped to a 5-dimensional space (x, y, Y, Cr, Cb)
- Hashing applied to get unique coordinates which represent vertices
- Creating a sparse splat matrix to go from pixels to vertices



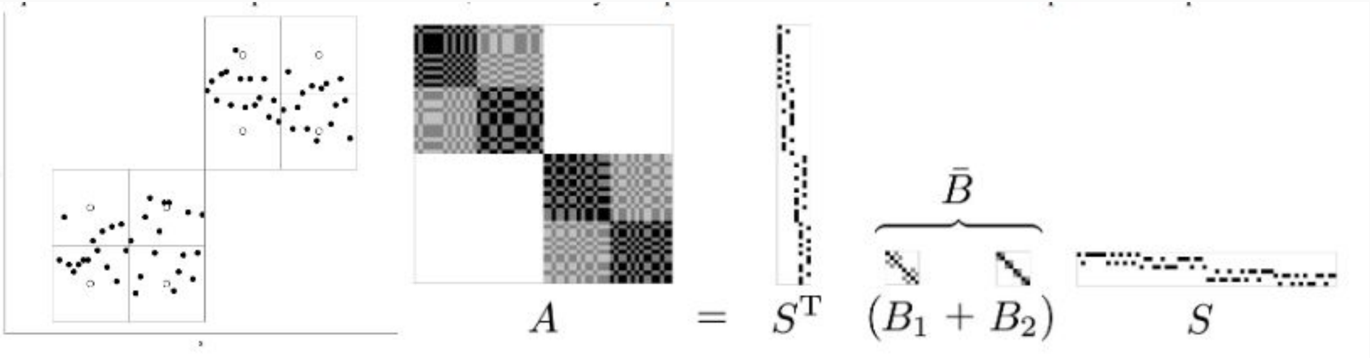
Slice

- Transpose of splat

Fast Bilateral Grid

Blur Matrix

- Actually it is a product of several sparse matrices, one for each dimension. Instead of multiplication, our approach simplifies it to addition.
- This is faster & easier



Formulation of Optimization Problem in Bilateral Space

Revisiting Problem in Pixel Space

- Objective in pixel space is a combination of data and smoothness terms:

$$\underset{\mathbf{x}}{\text{minimize}} \; \frac{\lambda}{2} \sum_{i,j} \hat{W}_{i,j} \left(x_i - x_j\right)^2 + \sum_i c_i (x_i - t_i)^2 \qquad \text{--(1)}$$

  - ti, ci are target and confidence values obtained from stereo matching
  - W matrix represents measure of smoothness

- In our case, W_hat is bistochastic version of Bilateral Matrix W:

$$W_{i,j} = \exp\left(- \frac{\|[[p_i^x, p_i^y] - [p_j^x, p_j^y]\|^2}{2\sigma_{xy}^2} - \frac{(p_i^l - p_j^l)^2}{2\sigma_l^2} - \frac{\|[p_i^u, p_i^v] - [p_j^u, p_j^v]\|^2}{2\sigma_{uv}^2} \right) \qquad \text{--(2)}$$

From Pixel to Bilateral Space

- Approximating W using Splat, Blur, and Slice Matrices

$$W = S^{\mathrm{T}} \bar{B} S$$ --(3)

- Bistochastization of W. (Pseudo code in paper)

$$\hat{W} = S^{\mathrm{T}} D_{\mathbf{m}}^{-1} D_{\mathbf{n}} \bar{B} D_{\mathbf{n}} D_{\mathbf{m}}^{-1} S \qquad S S^{\mathrm{T}} = D_{\mathbf{m}}$$ --(4)

- Variable Substitution. Going from pixel to vertices (Bilateral Space)

$$\mathbf{x} = S^{\mathrm{T}} \mathbf{y}$$ --(5)

Final Problem in Bilateral Space

- Substituting (4) and (5) in (1), our problem in Bilateral Space becomes:

$$\underset{\mathbf{y}}{\text{minimize}} \quad \frac{1}{2}\mathbf{y}^{\mathrm{T}}A\mathbf{y} - \mathbf{b}^{\mathrm{T}}\mathbf{y} + c$$

$$A = \lambda(D_{\mathbf{m}} - D_{\mathbf{n}}\bar{B}D_{\mathbf{n}}) + \mathrm{diag}(S\mathbf{c}) \qquad \mathbf{b} = S(\mathbf{c} \circ \mathbf{t}) \qquad c = \frac{1}{2}(\mathbf{c} \circ \mathbf{t})^{\mathrm{T}}\mathbf{t}$$

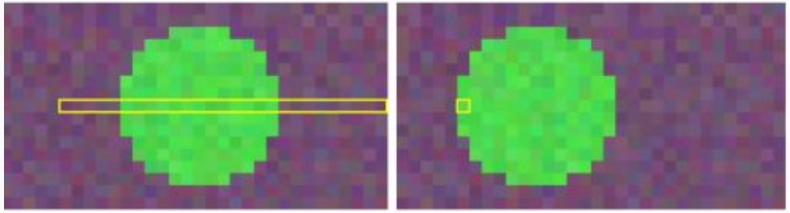
Advantages:

- y << x, Hence fast
- Edge aware depth
- Useful for defocus applications

Stereo Matching (Scratching The Surface)

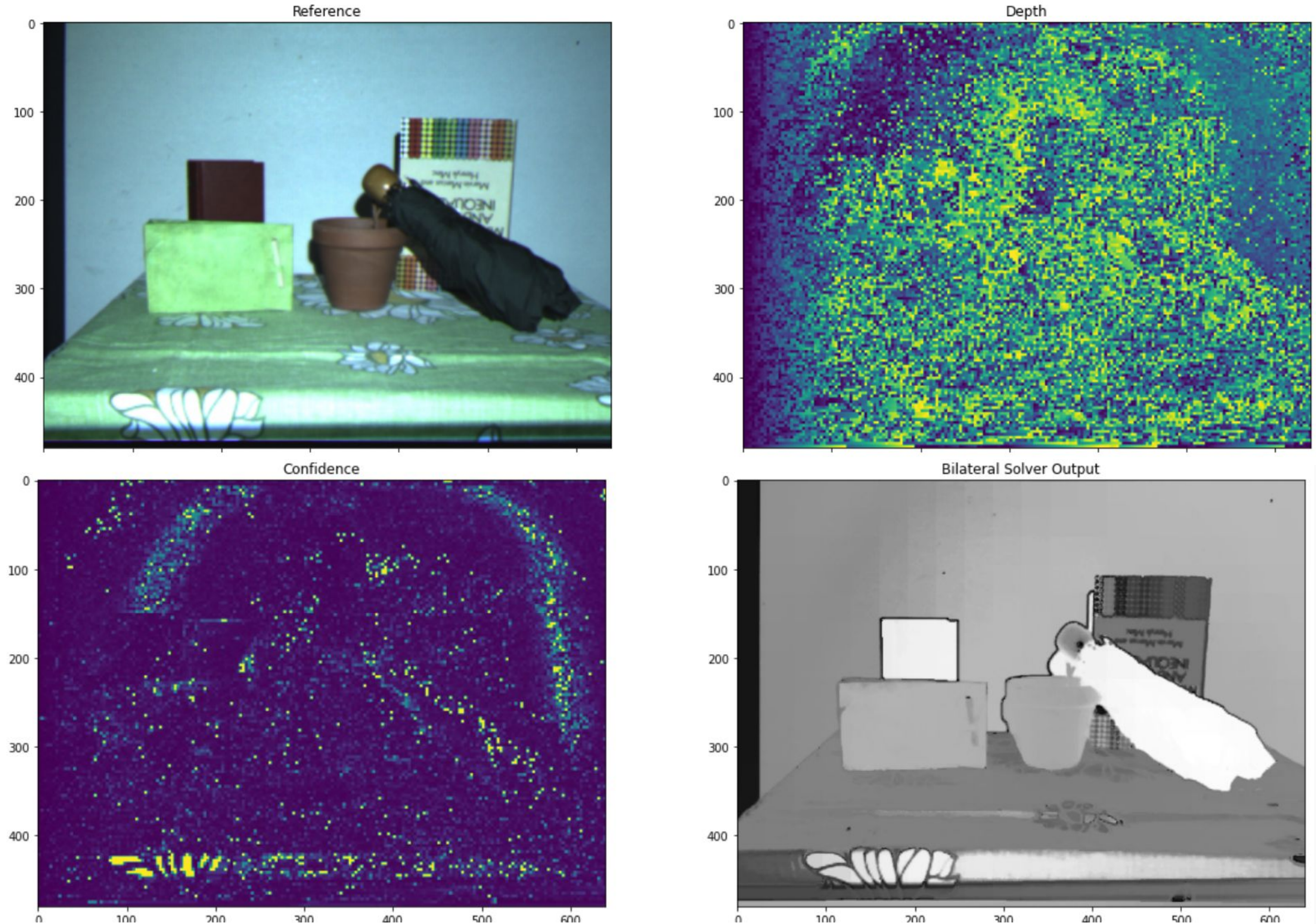
Approaches:

- Per-Pixel
  - SSD (Sum of Squared Difference)
  - SAD (Sum of Absolute Difference)
- Block Matching (Aggregation)
- Using Birchfield-Tomasi Measures



Stereo Matching in Rectified Images

Results


Reference
Depth
Confidence
Bilateral Solver Output
0
100
200
300
400
500
600

Synthetic Defocus

Synthetically Defocused Image

Reference Image

Conclusion

- Pros:
  - Edge aware stereo optimization
  - Considerably fast
  - Ideal for synthetic defocus use case
- Issues :
  - Cannot be used for fine resolution of depth
  - Not ideal when multiple objects present in foreground

- Scopes of Improvement :
  - Different stereo algorithms can be used to generate disparity map
  - Other ways of generating splat matrix
  - Testing on benchmark datasets

References

1. Jonathan T. Barron, Andrew Adams, YiChang Shih, Carlos Hernandez, ***Fast Bilateral Space Stereo for Synthetic Defocus***, CVPR 2015

2. Jonathan T. Barron, Andrew Adams, YiChang Shih, Carlos Hernandez, ***Fast Bilateral Space Stereo for Synthetic Defocus Supplemental Material***, CVPR 2015

3. Daniel Scharstein and Richard Szeliski, ***A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms***, IJCV

4. Jonathan T. Barron and Ben Poole , ***The Fast Bilateral Solver***, ECCV 2016

5. Stan Birchfield and Carlo Tomasi, ***Depth Discontinuities by Pixel-to-Pixel Stereo***, IJCV 1999

Work Distribution

- **Balavarun**:
  - Bilateral Grid
    - Splat, Blur, Slice Matrices
  - Defocus

- **Omama**
  - Bilateral Grid
    - Splat, Blur, Slice Matrices
  - Bilateral Solver
  - Code Integration
  - Final Testing, Parameter Tweaking

- **Apoorva**:
  - Stereo Matching (Per Pixel)
  - Testing Different Approaches of Stereo Matching
  - Slides, Documentation
- **Jhanvi** :
  - Stereo Matching
    - Birchfield Tomasi, Sum of Squared Difference
    - Block-Matching Interval Cost Volume