Multi-Scale Geometric Consistency Guided and Planar Prior Assisted Multi-View Stereo

Presented by Group 3

Elliot Huang, Ji-Ze Jang, Kshitij Karnawat (Scribe), Rishikesh Jadhav, Sathvik Inteti

Motivation

PatchMatch MVS allows for *fast* global search for the best match in corresponding images and can be implemented in parallel on GPUs.

However, this depends on a stable visual similarity measure, which cannot capture the discrimination in low-textured areas.

Undesirable trade-off:

efficient propagation $\Longleftrightarrow$ accurate multi-view matching cost evaluation

Question: how to design a more robust view selection method based on the checkerboard pattern?

Figure source: https://openaccess.thecvf.com/content_CVPR_2019/papers/Xu_Multi-Scale_Geometric_Consistency_Guided_Multi-View_Stereo_CVPR_2019_paper.pdf (Figure 2(a), 2(b))

Key Prior works

According to the prior works - **MVS methods** can be categorized in 4 groups :

1. **Voxel-based methods** - Limitations arise from the predefined voxel grid resolution.
2. **Surface evolution based methods** - These methods depend on a good initial solution, which can be challenging to obtain.
3. **Patch-based methods** - Completeness issues occur due to the reliance on matched keypoints.
4. **Depth map based methods** - These methods require depth estimation for all images and subsequent fusion into a 3D scene representation, which can be computationally intensive.

The authors propose an innovative depth map-based method comprising two components:

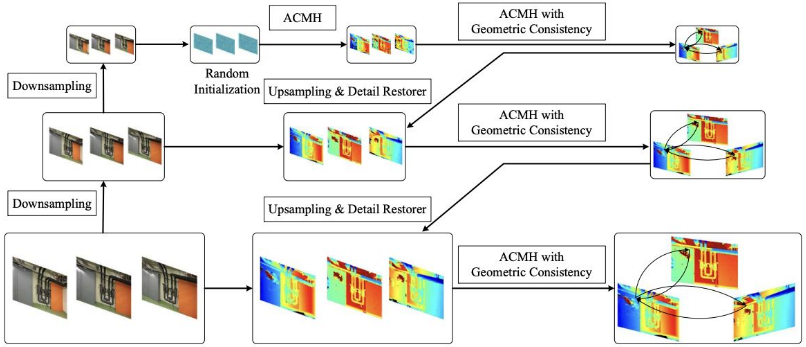
A. **Adaptive Checkerboard Sampling and Multi-Hypothesis Joint View Selection (ACMH):** This component introduces a novel approach to depth map estimation.
B. **ACMH with the Proposed Multi-Scale Geometric Consistency Guidance Scheme (ACMM):** This enhancement integrates a multi-scale geometric consistency guidance scheme into the ACMH method.

Overview / Highlights

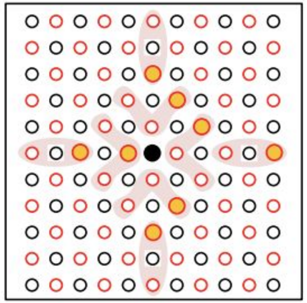
Adaptive checkerboard sampling scheme

Multi-scale patch matching scheme to deal with ambiguities in low-textured areas

Novel planar prior to tackle depth estimation in low-textured areas



ACMH



(c)

Random Initialization - randomly generate hypotheses (depth and normal)

Adaptive Checkerboard Sampling - Four V's, Four strips

- Sample good hypotheses according to cost (calculated using plane-induced homography)

MH Joint View Selection - select views that have a lower costs for the 8 hypotheses

Assign weight of each selected view s.t. most important view of previous iteration influences selection of views for current iteration

Refinement - two new hypotheses. 6 hypotheses tested again. minimum cost hypothesis chosen as estimate for p

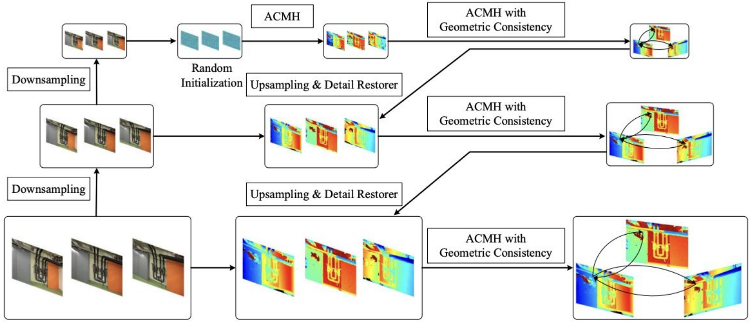


Figure source: https://openaccess.thecvf.com/content_CVPR_2019/papers/Xu_Multi-Scale_Geometric_Consistency_Guided_Multi-View_Stereo_CVPR_2019_paper.pdf (Figure 1(c), 3)

ACMH with Geometric Consistency Guidance Scheme (ACMM)

Geometric Consistency:

- Depth maps obtained by ACMH are noisy due to ambiguities and occlusions.
- Perform geometric consistency at the coarsest scale to optimize these initial depth maps.
- Conduct geometric consistency guidance twice to refine depth maps at each scale in our experiments.

Detail Restorer (why is it needed?):

- Geometric consistency leads to blurred details,
- Lost details cause loss in depth information at coarser scales,
- Upscaling adds extra error in details

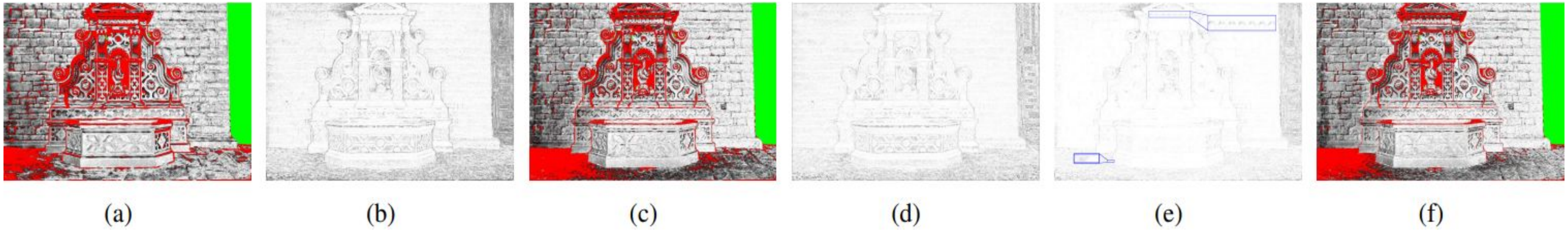


Figure source: https://openaccess.thecvf.com/content_CVPR_2019/papers/Xu_Multi-Scale_Geometric_Consistency_Guided_Multi-View_Stereo_CVPR_2019_paper.pdf (Figure 4)

ACMM Pipeline

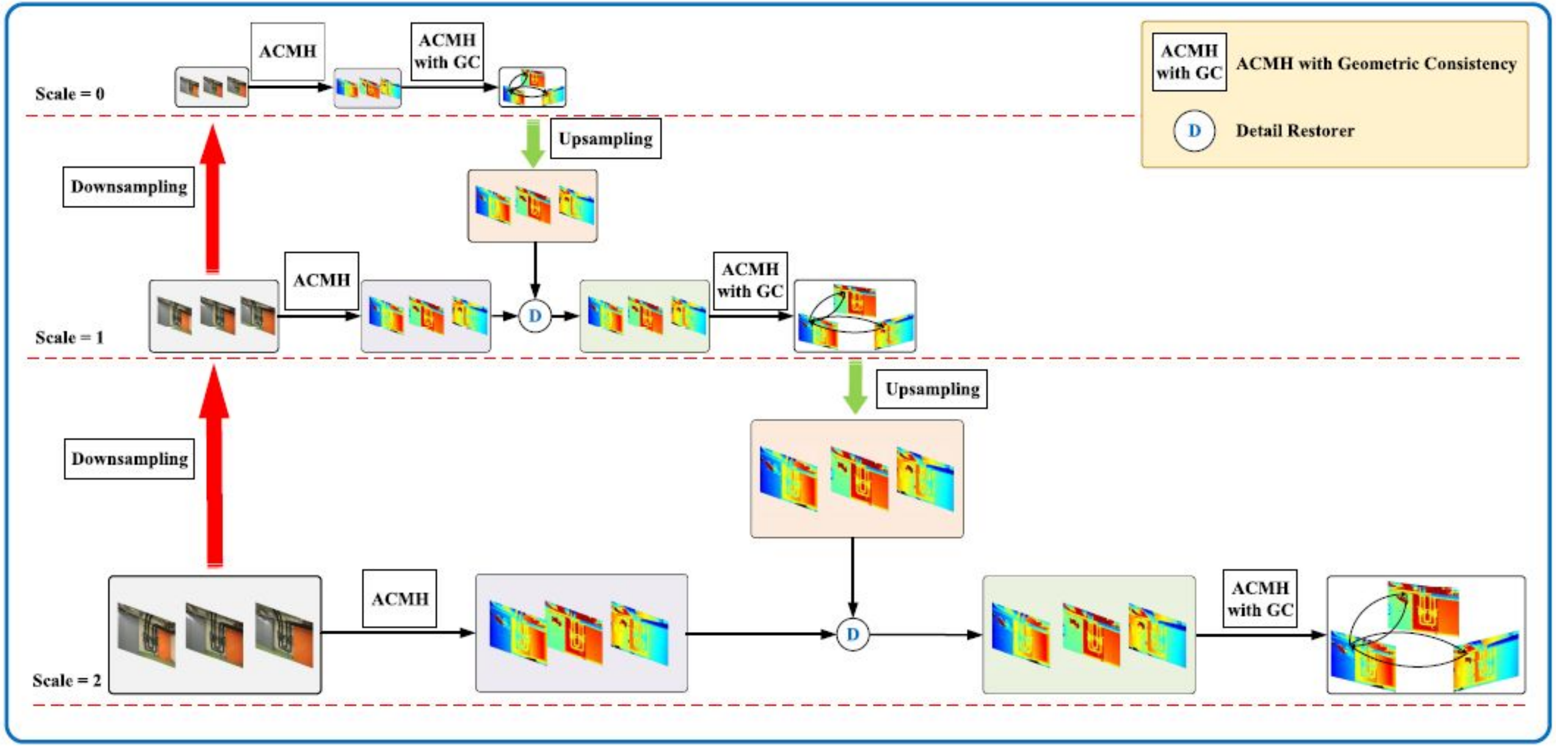


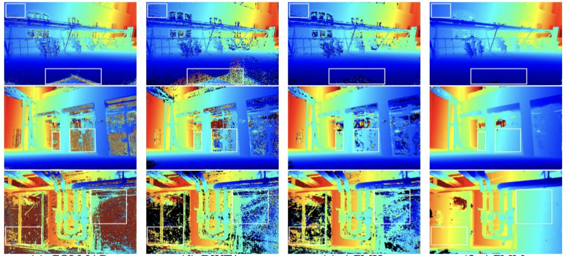
Figure source: https://ieeexplore.ieee.org/document/9863705 (Figure 6)

Experiments & Conclusion

MVS datasets used: Strecha dataset and ETH3D benchmark

Depth Map Evaluation

- ACMH surpasses DWTA and is on par with COLMAP without geometric consistency
- ACMM outperformed other datasets, especially in challenging datasets (e.g., indoor, poorly textures)
- ACMH is the second-best overall



(c) COLMAP (d) DWTA (e) ACMH (f) ACMM

Point Cloud Evaluation

- Accuracy: all methods tested (i.e., PMVS, Gipuma, LTVRE, COLMAP, ACMH and ACMM) show similar accuracy
- Completeness:
  - ACMH is competitive, with same completeness with ACMM in outdoor scenes
  - ACMM has more advantages in challenging areas (e.g., indoor scenes)

Runtime Performance

- ACMH is 6x faster than COLMAP, also faster than Gipuma which uses checkerboard propagation
- ACMM is slower than ACMH, but no more than twice as slow, which means still 3x faster and COLMAP

Conclusion

- Both methods achieve good efficiency and state-of-the-art performance, with ACMM offering exceptionally detailed estimation in challenging datasets

Figure source: https://openaccess.thecvf.com/content_CVPR_2019/papers/Xu_Multi-Scale_Geometric_Consistency_Guided_Multi-View_Stereo_CVPR_2019_paper.pdf (Figure 5(c), 5(d), 5(e), 5(f))