A Comprehensive Study of Stereo Vision Algorithms: Block Matching, Relaxation, Dynamic Programming, and Evaluation

Mohammed-salih Diyari
Université d'Évry Paris-Saclay

***Abstract*—This work investigates classical stereo matching approaches including Block Matching (BM), Relaxation-based global optimization, and Dynamic Programming (DP). Using the Middlebury *Venus* dataset, we systematically evaluate how parameters such as block size, contrast threshold, and disparity search range affect disparity map quality. We additionally implement gradient filtering, hole filling, and left–right consistency validation to refine disparity estimates. Quantitative metrics include Mean Relative Error (MRE) and photometric reprojection error. This document is intentionally written as both a lab report and a future reference guide for revisiting stereo vision concepts.**

CONTENTS

I. INTRODUCTION

Stereo vision estimates depth from two horizontally displaced cameras. The core objective is to compute the disparity map, representing how far each pixel shifts between the left and right views. Disparity estimation is challenging due to noise, textureless regions, occlusions, and ambiguous matches.

This TP involves:

- Implementing local methods (Block Matching)
- Implementing global methods (Relaxation, Dynamic Programming)
- Comparing algorithms using ground truth
- Exploring post-processing techniques
- Performing parameter sweeps and analyzing behavior

—

II. THEORY BACKGROUND

A. Disparity and Depth

Given rectified stereo images, disparity d relates to scene depth Z:

$$Z = \frac{fB}{d}$$

where f is focal length and B is baseline.

Analogy: Hold one finger in front of you. Close one eye, then the other: the amount your finger "jumps" is disparity. Closer objects jump more.

—

B. Matching Cost

Stereo matching relies on comparing pixel neighborhoods:

- SAD: $\sum |I_L - I_R|$
- SSD: $\sum (I_L - I_R)^2$
- NCC: correlation-based, illumination-invariant

—

C. Block Matching (BM)

BM selects the disparity d that minimizes the chosen cost in a small window.

Pros:

- Simple and intuitive

Cons:

- Fails in low texture
- Sensitive to block size and noise

—

D. Relaxation

Relaxation attempts to enforce smoothness:

$$E = C(x, y, d) + \lambda \sum_{\text{neighbors}} |d - d_n|$$

Pixels iteratively update disparity based on neighbors.

—

E. Dynamic Programming (DP)

DP optimizes each scanline:

$$M(x, d) = C(x, d) + \min_{d'}(M(x-1, d') + \lambda|d - d'|)$$

Advantages:

- Globally optimal per row

Drawbacks:

- Horizontal streaks due to missing vertical regularization

—

F. Left–Right Consistency

A correct disparity satisfies:

$$D_L(x, y) = D_R(x - D_L(x, y), y)$$

Pixels failing this check are mismatches or occluded.

—

III. Results and Detailed Analysis

This section includes all final disparity maps and parameter sweeps, each with a detailed explanation tailored for future understanding.

A. Comparison of BM, Relaxation, and DP

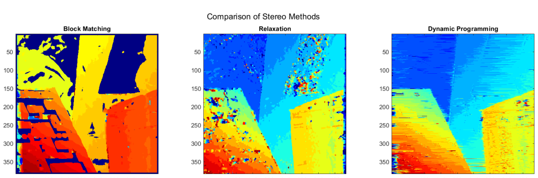

Fig. 1: Comparison between BM, Relaxation, and DP.

Explanation: BM produces blocky surfaces; Relaxation smooths but amplifies noise; DP produces clean slanted surfaces but exhibits horizontal streaking. All three highlight typical trade-offs between local and global stereo methods.

B. Effect of Contrast Threshold

Explanation: A higher contrast threshold rejects unreliable regions. 0.00 $\rightarrow$ very noisy; 0.02 $\rightarrow$ cleaner; 0.05 $\rightarrow$ stable; 0.10 $\rightarrow$ overly sparse. This demonstrates the reliability-vs-coverage trade-off.

C. Effect of Block Size

Explanation: Small blocks = noisy but detailed; Large blocks = smooth but oversmoothed edges. Block size 7–11 is typically optimal for Middlebury-like scenes.

D. Effect of Maximum Disparity Range

Explanation: Too small a range truncates the disparity; too large increases noise. 20 is appropriate for Venus (true disparity max around 18).

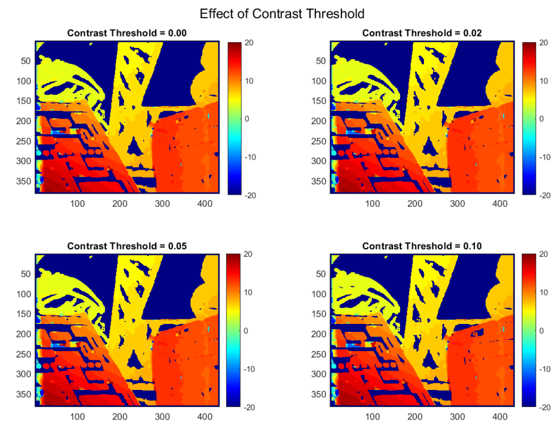

Fig. 2: Contrast threshold influence on disparity estimation.

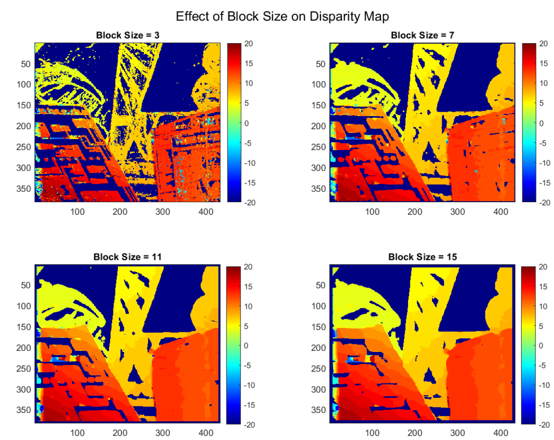

Fig. 3: Block size effect on BM performance.

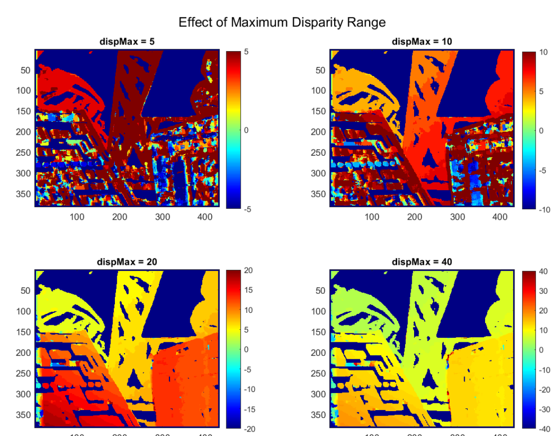

Fig. 4: Effect of search range on disparity accuracy.

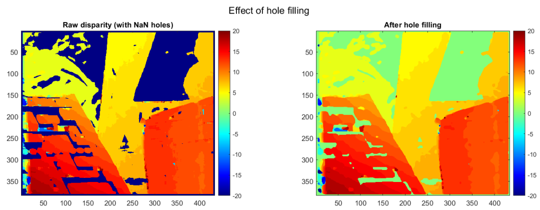

Fig. 5: Hole filling before and after.

E. Hole Filling Analysis

Explanation: Holes (NaNs) arise from occlusions or filtering. Hole filling ensures continuity, though it may hide actual uncertainty.

F. Evaluation Against Ground Truth

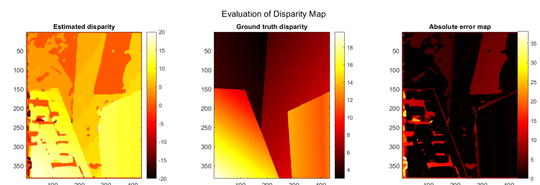

Fig. 6: Estimated disparity, ground truth, and absolute error.

Explanation: The error map directly indicates where the algorithm fails: edges, occlusions, and low-texture regions. This visualization is crucial for understanding algorithmic failure modes.

G. Global Evaluation (Direction Corrected)

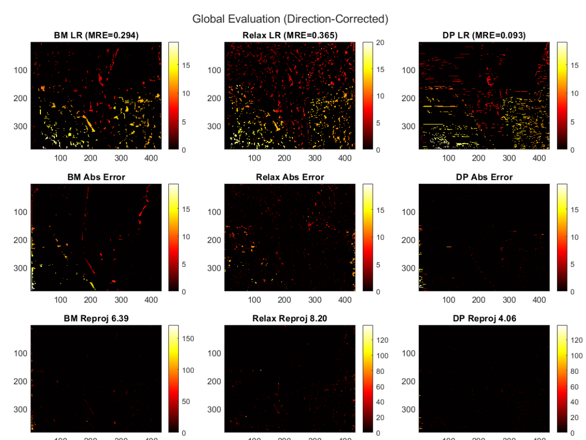

Fig. 7: Combined evaluation: MRE, reprojection error, LR consistency.

Explanation: This figure reveals:

- BM performs unexpectedly well in reprojection error but poorly in consistency.
- Relaxation performs worst under all quantitative metrics.
- DP achieves lowest reprojection error but still fails structurally.

It also demonstrates why disparity *direction* correctness is essential.

H. Left–Right Consistency Check

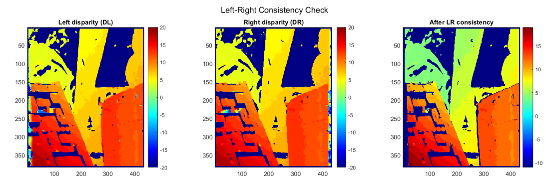

Fig. 8: LR consistency validation.

Explanation: LR consistency identifies invalid matches. Even visually plausible maps may fail this check. It is one of the strongest sanity-check tools in stereo vision.

—

IV. Conclusion

This TP provided hands-on experience with classical stereo methods. Key insights include:

- BM is simple but highly sensitive to parameters.
- Relaxation smooths disparity but is fragile to noise.
- DP achieves good alignment but introduces structural artifacts.
- LR consistency and reprojection error are essential evaluation tools.
- Correct disparity direction is absolutely critical.

Future improvements include Semi-Global Matching (SGM), Census transform costs, and modern neural stereo networks.

—