The Global Patch Collider

Lina Zhu

June 8 ,2018

Abstract

This paper presents a very effective, fully parallel, global, peer-to-peer communication in computational task-specific algorithmic images and video. Our algorithm Global Patch Collider is based on the detection of a unique collision conditional learning tree structure hash function that uses a set of image points. Contrary to traditional methods, relying on pairing distance calculations, our algorithm isolates unique pairs of pixels in the process to hit the same leaf traversing multiple learning tree structures. The functions that split the intermediate nodes stored in the tree are trained to ensure that only visually similar patches or their geometric or photometric conversion versions fall into the same leaf node. The matching process involves passing through all the pixel position tree structures in the analyzed image. Then we use isolation to calculate the points that match the only ones that collide with each other. Our algorithm is linear and can do a constant amount of time on the number of pixels. Tree traversal on a parallel computing architecture is decoupled for a single image point. We demonstrate our approach by using it to perform some challenging working benchmarks such as optical power flow matching and stereo matching.

1. Introduction

The corresponding estimate is. Estimating how the parts of the task's visual signals correspond to each other is an important and challenging issue for computer vision. Point-to-point correspondence between images or 3D volumes can be used for tasks such as camera pose estimation, multi-view stereo, motion structure, common segmentation, retrieval and compression, etc., and many variations of general communication such as stereo and optical flow estimation problems It has been widely studied in the literature. There are two key challenges to matching visual content across images or volumes. First, robust modelling of luminosity and geometric transformations occurs in real-world data such as occlusion, large displacements, viewpoints, shadows, and lighting changes. Second, perhaps more importantly, the difficulty of performance is inferred in the above model. The latter derives from the computational complexity of performance in potential correspondence and search in a large amount of space is the main obstacle algorithm for real-time development. A popular method of dealing with problems involves detecting points of interest or salient points in the image and then matching the specified hand[7] according to the measured Euclid or understanding[3] descriptors are designed to be unchanged for certain categories of conversions and in some cases can also work across different modalties. For example, in the case of optical flow only by searching, it is used for matching near the pixel position. However, this method cannot detect large movements/displacements. Methods such as[1] overcome this by adaptively sampling the search space and having problems has been proved to be very effective for optical flow and parallax estimation[2]. However, they rely on implicit fields between hypothetical images. When this hypothesis is violated, it goes well and fails. The technique is based on finding nearest-neighbor algorithmic neighbors such as KD-Tree[6] and hash[4] which can be used to search for large-displacement correspondences and have been used to initialize optical flow calculations[8]. However, these methods search for candidates based on appearance similarities and their matching in geometric and photometric measurement scenarios where the inconsistent transitions occur (see Figure 1).

2. Structured light stereo matching

For stereo matching tasks, we collected stereo images with 2200 infrared indoor scenes and Kinect depth sensors based on structured lighting. The reference mode is to use the calibration procedure in[5] to recover. The pattern and the Kinect image are corrected so that parallax matches the horizontal line [5]. We used 1000 frames as the rest of the training set as test tests. The GPC patch size is set to 77. During the training phase, we trained 10 trees to randomly collect 16 layers of each of more than 100 triplets from the training set. In practice we found that choosing the first 7 layers will ensure a good balance between the accuracy of this task and the recall rate, the polarity may greatly re-
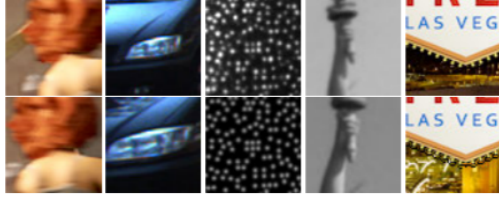
Figure 1. Examples of matched local patches. From left to right: Sintel, Kitti, active stereo, MVS, synthetic.

duce the number of possible match constraints. For each internal node, we generate 1024 random number suggestions and choose the objective definition that maximizes our best solution. Please note that in this one-dimensional matching case, the "sports" learning is simply to train the one-dimensional Gaussian binary classifier as the input difference of each node. The method of colliding with our previous and non-uniqueness further improves accuracy and recalls. We also conducted intensive stereo reconstruction by using our sparse matching to initialize the experiment based on PatchMatch stereo[2], we use our method to display PatchMatch results after 1 iteration and random initialization. Just like us, we can see that our method can produce more complete results and even match PatchMatch with full iteration quality. See Table 1 for quantitative comparisons.

Table 1. PatchMatch based on dense stereo results w/o initialization with global patch collider.

Baseline	1-iter		2-iter	
Random	1.5940	36.81%	1.5698	36.59%
Ours	1.5863	35.26%	1.5642	34.90%

3. Conclusion

This paper proposes a novel algorithm, the Global Patch Collider, for the computation of global point-wise correspondences in images. The proposed method is based on detecting unique collisions between image points using a collection of learned tree structures that act as conditional hash functions. Our algorithm is extremely efficient, fully-parallelizable, task-specific and does not require any pairwise comparison. Experiments on optical flow and stereo matching validates the performance of the proposed method. Future work includes high level applications such as hand tracking, and nonrigid reconstruction of deformable objects.

References

[1] C. Barnes, E. Shechtman, D. B. Goldman, and A. Finkelstein. The generalized patchmatch correspondence algorithm. In *ECCV*, 2010. 1

[2] M. Bleyer, C. Rhemann, and C. Rother. Patchmatch stereostereo matching with slanted support windows. In *BMVC*, 2011. 1, 2

[3] H. Bristow, J. Valmadre, and S. Lucey. Dense semantic correspondence where every pixel is a classifier. In *ICCV*, 2015. 1

[4] Z. Chen, H. Jin, Z. Lin, S. Cohen, and Y. Wu. Large displacement optical flow from nearest neighbor fields. In *CVPR*, 2013. 1

[5] S. R. Fanello, C. Rhemann, V. Tankovich, A. Kowdle, S. Orts Escolano, D. Kim, and S. Izadi. Hyperdepth: Learning depth from structured light without matching. In *CVPR*, 2016. 1

[6] K. He and J. Sun. Computing nearest-neighbor fields via propagation-assisted kd-trees. In *CVPR*, 2012. 1

[7] D. G. Lowe. Object recognition from local scale-invariant features. In *ICCV*, 1999. 1

[8] L. Xu, J. Jia, and Y. Matsushita. Motion detail preserving optical flow estimation. *PAMI*, 2012. 1