

Accelerated Gravitational Point Set Alignment with Altered Physical Laws*

—Supplementary Material—

Vladislav Golyanik¹
¹MPI for Informatics

Christian Theobalt¹
²University of Kaiserslautern

Didier Stricker^{2,3}
³DFKI

This supplement provides additional insights on the Barnes-Hut Rigid Gravitational Approach (BH-RGA). More details on 2^D -tree generation can be found in Sec. I. Next, Sec. II explains how the reprojection of the scene flow is visualised in Fig. 4 of the main draft. The detailed statistics of the resolved initial misalignments in the *clean-500* and *N500-U100* experiments is provided in Sec. III. Sec. IV elaborates on the differences in parameters of GA [7] and the proposed approach. Original reference numbers to the tables and figures from the main paper are preserved, and Roman numerals refer to the tables and figures introduced in this supplementary material.

I. Building a 2^D -Tree

In this section, we summarise the rules for building a 2^D -tree. The recursive algorithm was initially proposed by Barnes and Hut [3] for the N-body simulation problem [1], and we adapt it with minor high-level modifications as described in the main matter.

A BH-tree is initialised as a root node with 2^D empty external nodes v_l^j , $j \in \{1, \dots, 2^D\}$, and the depth of the tree l equals to 1. Always starting from the root, a new particle p_k is added to the tree following three rules, and every insertion results in a new leaf, *i.e.*, an occupied external node. Every node corresponds to some subdivided area of space (the root node stands for the entire space). Every insertion begins with localisation of p_k in the target space, *i.e.*, a hierarchical determination of the quadrants (2D), octants (3D), *etc.* which p_k belongs to, as well as the corresponding nodes, until an empty external node is reached. Localisation is valid on all levels of the tree, and it is valid to say that p_k is localised at node v_l^j at tree level l . The insertion rules for the localised p_k are:

- 1) [p_k is localised in an empty external v_l] If an external node v_l is empty, add p_k to v_l and done. The depth of the tree remains the same.

* supported by the ERC Consolidator Grant 4DReply (770784) and the BMBF projects DYNAMICS (01IW15003) and VIDETE (01IW18002).

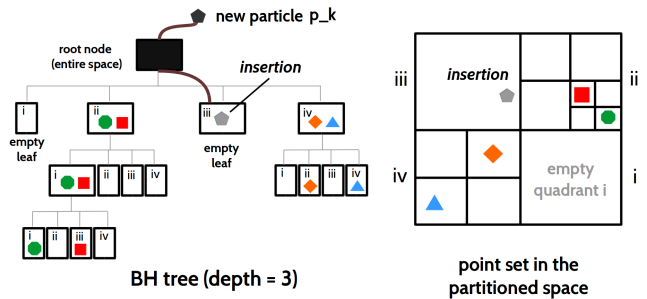


Figure 1: Insertion rule for building a 2^D -tree for an empty external node.

- 2) [p_k is localised in an internal (=non-empty) v_l] If an internal node v_l is encountered, update its mass and the centre of mass. Next, while relocalising p_k , visit and update all internal nodes hierarchically descending until an external node is reached. Follow rule 1) or 3) depending on whether the external node is empty or occupied.
- 3) [p_k is localised in a non-empty external v_l] If an external node v_l is occupied, declare v_l first to be a new internal node, — introduce centre of mass and a mass of the node — and then split v_l into 2^D new nodes. The depth of the tree grows by one. After the splitting, there will be one new occupied external node and $2^D - 1$ new empty external nodes. Add p_k to one of 2^D external nodes v_{l+1}^j following either rule 1) or 3).

Fig. 1 exemplifies insertion rule 1). Fig. 2 in the main paper illustrates rules 2) and 3).

II. Scene Flow Visualisation by Reprojection

We can interpret the result of rigid point set alignment in the experiment with the *SINTEL* dataset [5] as an RGB-D based scene flow. The objective of RGB-D scene flow estimation is, given two RGB-D measurements \mathcal{D}_t and \mathcal{D}_{t+1} , to reconstruct a 3D displacement field $\rho(\mathcal{D}_t): \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3$ warping \mathcal{D}_t to \mathcal{D}_{t+1} . We presume, without loss of generality, that each three-dimensional point in $\mathcal{D}_{[\cdot]}$ has a single intensity value. In our matrix notation, \mathbf{Y} and \mathbf{X} correspond

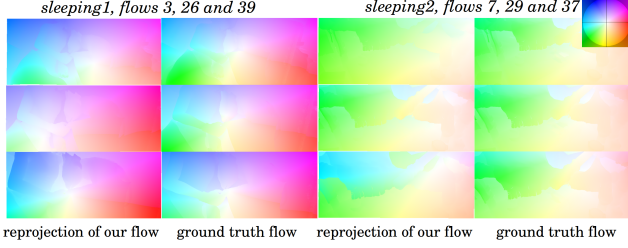


Figure II: Examples of reprojected 3D displacements obtained by BH-RGA, compared to ground truth optical flows, for the *sleeping1* and *sleeping2* sequences from the SINTEL collection [5]. The used *Middlebury* optical flow encoding [2] is given on the top right.

to \mathcal{D}_t and \mathcal{D}_{t+1} , respectively. Recall that in the *sleeping1* and *sleeping2* sequences [5], scene transformations are rigid or nearly rigid, and point sets are parametrised on a grid induced by $\mathcal{D}_{[\cdot]}$. The scene flow $\rho(\mathcal{D}_t)$ can then be computed as the difference between the transformed and the initial template in absolute coordinates:

$$\rho(\mathcal{D}_t) = \mathbf{F}_{3 \times M} = [\mathbf{f}_i] = (\mathbf{R} - \mathbf{I})\mathbf{Y} + \mathbf{T}, \quad (1)$$

with $\mathbf{T}_{3 \times M} = [\mathbf{t} \ \mathbf{t} \ \dots \ \mathbf{t}]$ and $i \in \{1, \dots, M\}$. M equals to the total number of pixels in the RGB-D image.

Suppose the focal lengths $\{f_x, f_y\}$ and the principal point (c_x, c_y) of the camera are known (*SINTEL* [5] provides the intrinsic camera parameters). In the following, we use the operator $\pi^{-1}: \mathbb{Z}^2 \times \mathbb{R} \rightarrow \mathbb{R}^3$, which maps a 2D image point (x, y) to the 3D space given its depth value:

$$\pi^{-1}(x, y, z) = \left(z \frac{(x - c_x)}{f_x}, z \frac{(y - c_y)}{f_y}, z \right). \quad (2)$$

The optical flow for visualisation $\rho_{2D}: \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is obtained for every $(x, y) \in \Omega$ (image domain) by the *scene flow projection operator* $\pi_\rho: \mathbb{Z}^2 \times \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}^2$ [8]:

$$\pi_\rho(x, y, \mathbf{P}_z, \mathbf{f}_i) = \begin{bmatrix} f_x \frac{\mathbf{f}_{i,x} + \mathbf{P}_x}{\mathbf{f}_{i,z} + \mathbf{P}_z} + (c_x - x) \\ f_y \frac{\mathbf{f}_{i,y} + \mathbf{P}_y}{\mathbf{f}_{i,z} + \mathbf{P}_z} + (c_y - y) \end{bmatrix}, \quad (3)$$

where $\mathbf{P} = \pi^{-1}(x, y, \mathbf{P}_z)$, $\mathbf{P}_z = \mathcal{D}_t(x, y)$ and the consecutive index i in \mathbf{f}_i is uniquely determined by (x, y) coordinates on the grid using the row-major rule. From (3), we see that ρ_{2D} can be alternatively calculated by projecting every transformed \mathbf{y}_i to the image plane and subtracting from the projected value the respective point coordinates (x, y) on the initial 2D image grid.

For *sleeping1* and *sleeping2* [5], the average endpoint error (AEPE) between the reprojected 3D flow fields parametrised by the recovered $\{\mathbf{R}, \mathbf{t}\}$ and the ground truth optical flow amounts to 2.242 and 0.914, respectively (see Fig. II for visualisations). Both sequences are fifty frames long. In our optical flow notation, EPE is defined per displacement vector as $\|\rho_{2D,u} - u_{GT}, \rho_{2D,v} - v_{GT}\|$, where (u_{GT}, v_{GT}) is the ground truth flow vector. The optical flow ρ_{2D} is visualised with the Middlebury colour scheme [2].

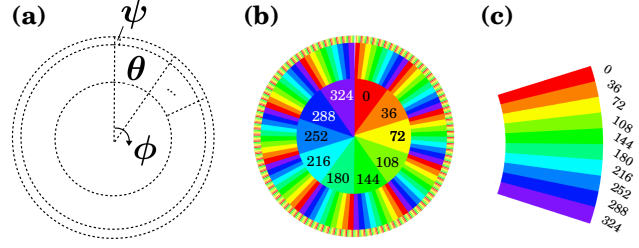


Figure III: Colour scheme for the statistics of the resolved initial misalignments. One line in the disk (from the disk centre to the disk boundary) denotes a successfully aligned pair of point sets differing in the global orientation by Euler angles ϕ , θ and ψ encoded by colours, around the x -, y - and z -axes, respectively. (a): Disk area subdivision for ϕ , θ and ψ . (b): Colouring for all tested combinations of ϕ , θ and ψ . The numbers in the ϕ area denote ϕ angle in degrees. (c): Angle sectors in the θ and ψ areas; recall that $\psi \in \{0^\circ, 36^\circ, 72^\circ, 108^\circ, 144^\circ\}$.

III. Varying Angles of Initial Misalignment

Whereas Tables 1 and 2 of the main paper report the summary of the rotation resolution capability of the tested methods in multiple settings, we would also like to analyse statistics of the resolved initial misalignments in detail. We define a colour scheme for all possible initial misalignments on a disk, see Fig. III. The disk is subdivided into three areas indicating ϕ , θ and ψ Euler angles used to transform the template point set in the *clean-500*, *N500-U50* and *N500-U100* experiments. All angles are sampled with the angular step of 36° . We discard duplicated states and obtain 500 input point set pairs with different initial misalignment in total, *i.e.*, $\phi, \theta \in \{0^\circ, 36^\circ, 72^\circ, 108^\circ, 144^\circ, 180^\circ, 216^\circ, 252^\circ, 288^\circ, 324^\circ\}$, and $\psi \in \{0^\circ, 36^\circ, 72^\circ, 108^\circ, 144^\circ\}$.

Thus, one line on the statistics disk stands for one resolved initial misalignment. In total, up to 500 lines at different angles can be present in the disk. Every line is subdivided into three segments with three colours, and each colour indicates ϕ , θ and ψ of the resolved initial misalignment. Fig. III-(b) shows at the same time the colour coding if all initial misalignments are resolved (this should be the desired case for a *global* point set alignment method).

Fig. IV shows the colour-coded statistics of the resolved initial misalignments for the *clean-500* and *N500-U100* experiments. We see that ICP [4], CPD (6 DoF) [10], GMR (7 DoF) [9] and GA [7] considerably worsen their accuracy on the *N500-U100* dataset compared to *clean-500* (the results of GMR (7 DoF) and GA and not shown for *N500-U100*). Uniform noise affects the accuracy of the tested methods in various ways. The success rate of LM-ICP [6] drops by $\sim 60\%$, and the success rate of CPD (6 DoF) halves on *N500-U100*. GMR (7 DoF) does not resolve point sets differing even by a small rotation, and GA resolves only five cases. BH-RGA and CPD (7 DoF) are the only meth-

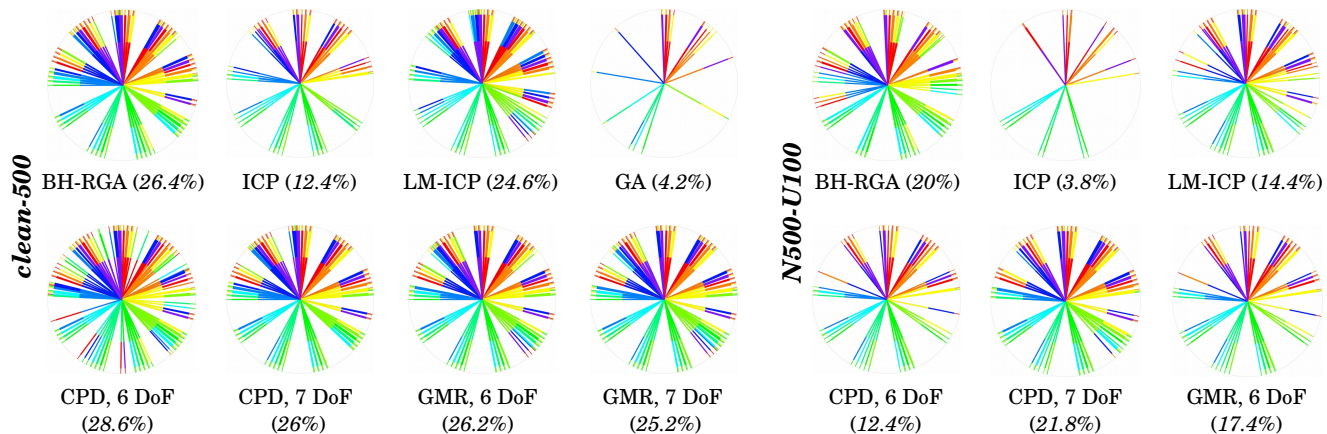


Figure IV: Colour-coded statistics of the resolved initial misalignments in the *clean-500* and *N500-U100* experiments for ICP [4], LM-ICP [6], CPD [10], GMR [9], GA [7] and BH-RGA (ours). The numbers in the brackets indicate the portion of the successfully resolved initial configurations in %.

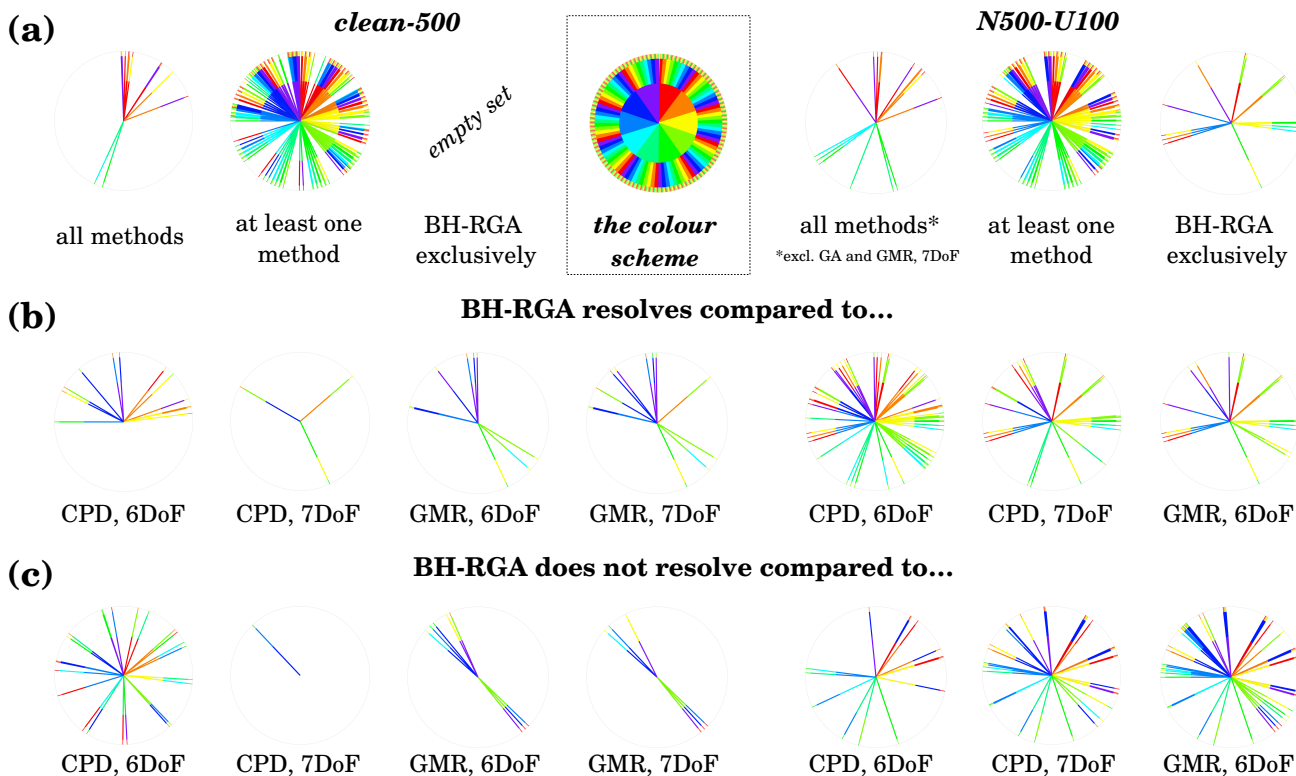


Figure V: Colour-coded **comparative** statistics of the resolved initial misalignments in the *clean-500* and *N500-U100* experiments for the evaluated approaches. The colour scheme is given in the top row, middle. **(a):** Joint statistics for all methods. **(b), (c):** Statistics for BH-RGA *versus* CPD and GMR.

ods with the decreased success rate by as little as 6.4% and 4.2%, respectively. Whereas BH-RGA, CPD (6 DoF), CPD (7 DoF), GMR (6 DoF), GMR (7 DoF) demonstrate comparable success rates on *clean-500*, only BH-RGA and CPD (7 DoF) show comparable top performance on *N500-U100*. We observe that in the case of CPD, allowing for scaling

helps to cope with large amounts of noise. BH-RGA, thanks to the globally multiply-linked alignment policy, achieves high accuracy with 6 DoF (the scaling is not included).

It is also insightful to perform a comparative analysis of resolved initial misalignments. Fig. V provides comparative statistics of successfully registered configurations in

<i>parameter</i>	<i>interpretation</i>
G	gravitational constant
$m^{\mathbf{x}_j}$	default mass of template points
$m^{\mathbf{y}_i}$	default mass of reference points
ε	softening length
η	strength of energy dissipation
Δt	forward integration step
$\vec{v}_{\mathbf{X}}^0$	<i>initial template's velocity (optional)</i>

Table I: Overview of GA [7] parameters.

<i>parameter</i>	<i>interpretation</i>
ϵ	Huber loss threshold
γ	distance threshold of 2^D -tree
$m_{\mathbf{y}_i}$	<i>template point masses (optional)</i>
$m_{\mathbf{x}_j}$	<i>reference point masses (optional)</i>

Table II: Overview of BH-RGA parameters.

	<i>clean-500</i>	<i>N500-U100</i>
<i>no Huber loss</i>	less resolved cases	slightly higher RMSE (in the fourth digit after the comma)
$\epsilon = 0.01$	<i>used in all experiments (Tables 1–3 of the paper)</i>	
$\epsilon = 0.1$	same number of resolved cases, $\sim 15\%$ higher RMSE	

Table III: Influence of ϵ on the experimental results.

the *clean-500* and *N500-U100* experiments and helps us to understand the strengths of the evaluated approaches better. For both experiments, we report in Fig. V-(a) the cases which are successfully resolved by all methods, cases which at least one method is able to resolve, and cases which only BH-RGA resolves. It is noticeable that the number of cases which BH-RGA can resolve exclusively increases from zero in *clean-500* to fourteen in *N500-U100*.

Next, we plot the cases which BH-RGA resolves, but another tested method (in this case, either CPD or GMR) does not, and vice versa. If on *clean-500*, CPD (6 DoF) resolves more cases which BH-RGA cannot resolve, the situation inverts on *N500-U100*, and we observe that BH-RGA resolves comparably more cases which CPD (6 DoF) cannot resolve. This positions BH-RGA as a method which can tackle many cases which are difficult for CPD (6 DoF) on *N500-U100*. Recall that CPD (6 DoF) is the method with the highest success rate on *clean-500*, followed by BH-RGA. Next, there are no cases which BH-RGA cannot resolve but either ICP, LM-ICP or GA can. Thus, the corresponding visualisations are trivial and, hence, not plotted. The input sets which CPD (6 DoF), CPD (7 DoF), GMR (6 DoF) and BH-RGA can resolve in *clean-500*, result in the intersection with a high degree of overlapping. Last but not least, the sets of configurations which BH-RGA, CPD (7 DoF) or GMR (6 DoF) can exclusively resolve, become more *disjoint* in *N500-U100* compared to the noiseless case.

IV. Parameters in GA [7] and BH-RGA

BH-RGA reduces the number of compulsory parameters by the factor of three compared to GA [7]. Tables I and II provide an overview of GA and BH-RGA parameters. For GA, if we choose default unit masses of template and reference points, we have to adjust the gravitational constant G . The latter, in turn, influences the choice of the softening length ε . The strength of the energy dissipation η depends on all other parameters. It seems reasonable to fix masses and G ; although it turns out practically that for a new dataset, other parameters have to be selected nevertheless.

In BH-RGA, we have to set the Huber loss threshold ϵ and the distance threshold of the 2^D -tree γ . We designate both parameters compulsory, though we have learned from the experiments that there exist values which work well for many datasets. Table III elaborates on the influence of ϵ on the outcome of the quantitative experiments. With no Huber loss or different ϵ values, both the number of resolved cases and RMSE slightly vary. Likewise, BH-RGA will probably converge for a wide range of γ values, with a difference in the speed and accuracy.

References

- [1] Sverre J. Aarseth. *Gravitational N-body Simulations: Tools and Algorithms*. Cambridge University Press, 2003. 1
- [2] Simon Baker, Daniel Scharstein, J. P. Lewis, Stefan Roth, Michael J. Black, and Richard Szeliski. A database and evaluation methodology for optical flow. *International Journal of Computer Vision (IJCV)*, 92(1), 2011. 2
- [3] Josh Barnes and Piet Hut. A hierarchical $o(n \log n)$ force-calculation algorithm. *Nature*, 324:446–449, 1986. 1
- [4] Paul J. Besl and Neil D. McKay. A method for registration of 3-d shapes. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 14(2):239–256, 1992. 2, 3
- [5] Daniel Jonas Butler, Jonas Wulff, Garrett B. Stanley, and Michael J. Black. A naturalistic open source movie for optical flow evaluation. In *European Conference on Computer Vision (ECCV)*, pages 611–625, 2012. 1, 2
- [6] Andrew W. Fitzgibbon. Robust registration of 2D and 3D point sets. In *British Machine Vision Conference (BMVC)*, pages 662–670, 2001. 2, 3
- [7] Vladislav Golyanik, Sk A. Ali, and Didier Stricker. Gravitational approach for point set registration. In *Computer Vision and Pattern Recognition (CVPR)*, 2016. 1, 2, 3, 4
- [8] Vladislav Golyanik, Kihwan Kim, Robert Maier, Matthias Nießner, Didier Stricker, and Jan Kautz. Multiframe scene flow with piecewise rigid motion. In *International Conference on 3D Vision (3DV)*, 2017. 2
- [9] Bing Jian and Baba C. Vemuri. Robust point set registration using gaussian mixture models. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 33(8):1633–1645, 2011. 2, 3
- [10] Andriy Myronenko and Xubo Song. Point-set registration: Coherent point drift. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2010. 2, 3