

GenAnalysis: Joint Shape Analysis by Learning Man-Made Shape Generators with Deformation Regularizations

ANONYMOUS AUTHOR(S)
SUBMISSION ID: 0178

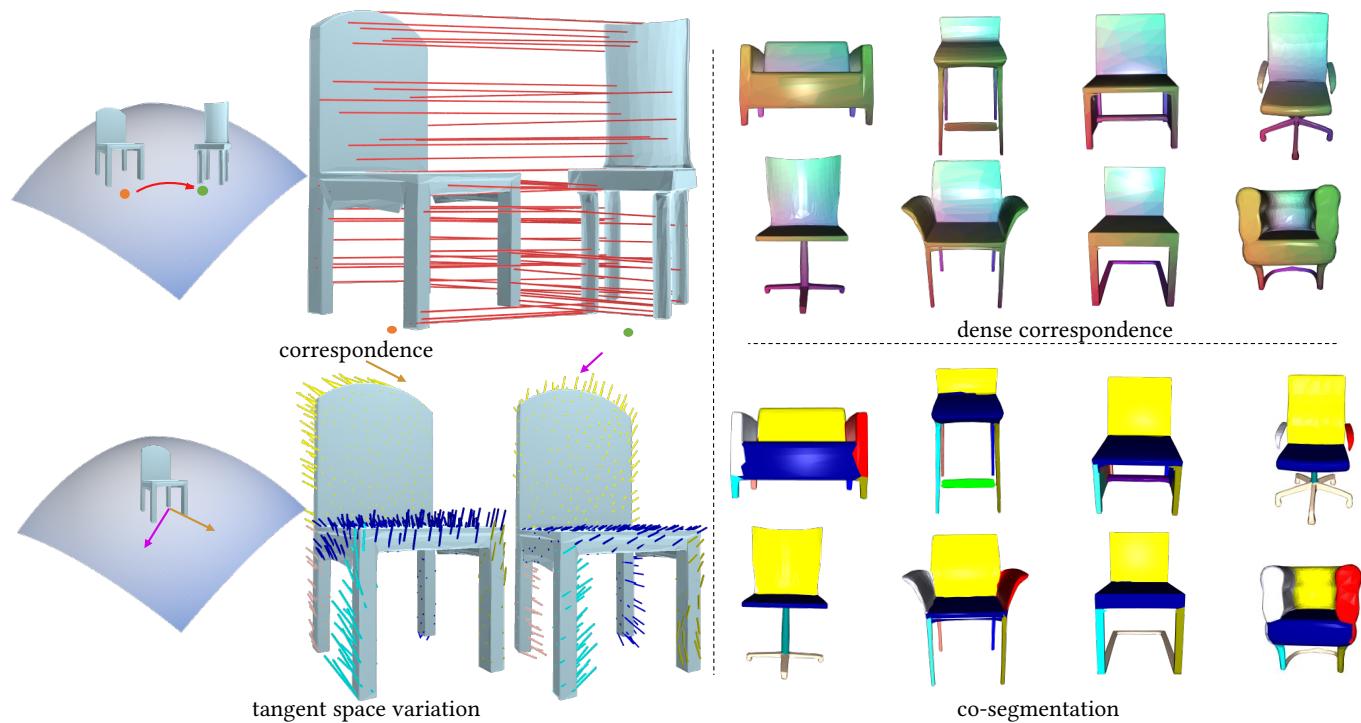


Fig. 1. (Top left) Correspondence from composing correspondences along computed intermediate shapes between the source and target shapes.(Bottom left) Vector fields of shape variation moving in different tangent direction. Arrows of different colors indicate variations in different directions. Different colors in vector fields visualize different piecewise affine components in vector fields. (Top right) Dense correspondence results by performing consistent shape matching on pairwise inputs. The colors-coded shapes visualize dense correspondences where points with similar color are in correspondence. (Bottom right) Unsupervised consistent shape segmentation using our methods in the chair category of ShapeNetPart [Yi et al. 2016]. The colored parts visualize segmentation consistency.

We introduce GenAnalysis, an implicit shape generation framework that allows joint analysis of a collection of man-made shapes. GenAnalysis innovates in learning an implicit shape generator to reconstruct a continuous shape space from the input shape collection. It offers interpolations between pairs of input shapes for correspondence computation. It also allows us to understand the shape variations of each shape in the context of neighboring shapes. Such variations provide segmentation cues. A key idea of GenAnalysis is to enforce an as-affine-as-possible (AAAP) deformation regularization loss among adjacent synthetic shapes of the generator. This loss forces the

generator to learn the underlying piece-wise affine part structures. We show how to extract data-driven segmentation cues by recovering piece-wise affine vector fields in the tangent space of each shape and how to use this generator to compute consistent inter-shape correspondences. These correspondences are then used to aggregate single-shape segmentation cues into consistent segmentations. Experimental results on benchmark datasets show that GenAnalysis achieves state-of-the-art results on shape segmentation and shape matching.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2024 Association for Computing Machinery.
0730-0301/2024/11-ART \$15.00
<https://doi.org/10.1145/nnnnnnnn.nnnnnnn>

CCS Concepts: • Computing methodologies → Shape analysis; • Neural networks;

Additional Key Words and Phrases: shape correspondence, shape co-segmentation, neural network regularization

ACM Reference Format:

Anonymous Author(s). 2024. GenAnalysis: Joint Shape Analysis by Learning Man-Made Shape Generators with Deformation Regularizations. *ACM Trans.*

115 *Graph.* 1, 1 (November 2024), 19 pages. <https://doi.org/10.1145/nnnnnnn>
 116
 117
 118

119 1 INTRODUCTION

120 Shape analysis, where the main goals are shape segmentation and
 121 shape matching, is a fundamental research area in geometry pro-
 122 cessing and related fields. Although early approaches focus on
 123 single-shape segmentation [Chen et al. 2009; Kaick et al. 2015] and
 124 matching a pair of shapes in isolation [Kim et al. 2011; Lipman and
 125 Funkhouser 2009; Melzi et al. 2019; Ovsjanikov et al. 2012a], recent
 126 approaches have shown great promise by jointly analyzing a col-
 127 lection of shapes. In particular, joint shape matching [Deng et al.
 128 2021; Huang and Guibas 2013a; Huang et al. 2012a, 2020; Nguyen
 129 et al. 2011; Zheng et al. 2021] allows one to compute dense corre-
 130 spondences between dissimilar shapes by composing correspon-
 131 dences along a path of similar shape pairs, and joint shape segmen-
 132 tation [Chen et al. 2019; Golovinskiy and Funkhouser 2009b; Huang
 133 et al. 2011, 2019, 2014a; Sidi et al. 2011a; Tulsiani et al. 2017; Yang
 134 and Chen 2021] shows the potential to enforce the consistency of
 135 shape segmentations to reduce segmentation errors in single-shape
 136 segmentations. The performance of joint shape analysis depends on
 137 how to define and compute high-quality correspondences across the
 138 shape collection. This becomes particularly challenging for hetero-
 139 geneous man-made shape collections that exhibit large inter-shape
 140 deformations, which make shape matching difficult, and structural
 141 variations, where only partial correspondences are possible.
 142

143 In this paper, we introduce *GenAnalysis*, a novel framework for
 144 joint shape analysis that allows both shape matching and consis-
 145 tent shape segmentation. The key idea of GenAnalysis is to learn
 146 an implicit shape generator from a shape collection with geom-
 147 etric regularizations [Huang et al. 2021; Yang et al. 2024, 2023] that
 148 forces the generator to capture the common underlying geometric
 149 structures among shapes. We introduce a novel approach to com-
 150 pute dense correspondences between adjacent synthetic man-made
 151 shapes, turning an implicit man-made shape generator into a ex-
 152 plicit mesh generator locally. Correspondence computation employs
 153 a new as-affine-as-possible (AAAP) deformation model tailored for
 154 man-made shapes. GenAnalysis uses a novel regularization loss that
 155 employs a robust norm on the AAAP model to promote piece-wise
 156 affine deformation structures of the implicit shape generator.

157 GenAnalysis presents two contributions in using the shape gen-
 158 erator to compute segmentation cues and inter-shape correspon-
 159 dences. Segmentation cues are based on extracting piece-wise affine
 160 structures in the vector fields derived from the tangent space of
 161 the shape generator at each shape. This local approach is made
 162 possible by enforcing the AAAP regularization loss described above,
 163 in which vector fields of each shape reveal inter-shape variations.
 164 GenAnalysis computes inter-shape correspondences by generating
 165 intermediate shapes between a shape pair and using them to prop-
 166 agate correspondences. We show how to address potential partial
 167 similarity, where the source and target shapes are structurally dif-
 168 ferent. These pairwise matches are fed into an off-the-shelf map
 169 synchronization approach to obtain consistent correspondences
 170 across the shape collection.

171 GenAnalysis performs consistent shape segmentation by using
 172 consistent inter-shape correspondences to aggregate the segmen-
 173 tation cues of each shape. For computational efficiency, we perform
 174 consistent shape segmentation using over-segments computed on
 175 each shape.

176 We have evaluated the performance of GenAnalysis on ShapeNet
 177 for the task of inter-shape correspondence computation and consis-
 178 tent shape segmentation. In both tasks, GenAnalysis outperforms
 179 state-of-the-art techniques by salient margins, justifying the power
 180 of performing shape analysis by learning shape generators. For
 181 example, GenAnalysis outperforms state-of-the-art unsupervised
 182 shape segmentation approach [Chen et al. 2024] by 3.4% in mean
 183 IOU score. An ablation study justifies the importance of each com-
 184 ponent of GenAnalysis.

187 2 RELATED WORK

188 We discuss related work in three groups, namely, neural genera-
 189 tive models for man-made shapes (Section 2.1), data-driven shape
 190 segmentation 2.2, and inter-shape correspondences (Section 2.3).

193 2.1 Neural Generative Models for Man-Made Shapes

194 Recent advances in learning generative shape models have shown
 195 impressive results in synthesizing novel shapes. These approaches
 196 have primarily focused on developing network architectures under
 197 different 3D representations [Achlioptas et al. 2018; Chen and Zhang
 198 2019; Li et al. 2017; Mo et al. 2019; Park et al. 2019; Sinha et al. 2016;
 199 Wang et al. 2018; Wu et al. 2016]. A common scheme is to align
 200 the distribution of the training shapes with that of the synthetic
 201 shapes [Arjovsky et al. 2017; Bouritsas et al. 2019; Goodfellow et al.
 202 2020; Litany et al. 2018; Nakayama et al. 2023; Park et al. 2019;
 203 Ranjan et al. 2018; Song and Ermon 2019; Song et al. 2021; Tan et al.
 204 2018; Yang et al. 2018; Zadeh et al. 2019]. However, the performance
 205 of these approaches is limited to sparse 3D data, which is a lasting
 206 challenge in 3D deep learning.

207 An effective approach to address the data sparsity issue is to
 208 develop regularization losses to train shape generative priors that
 209 enforce prior knowledge about the underlying shape space [Atz-
 210 mon et al. 2021; Gropp et al. 2020; Huang et al. 2021; Liu et al.
 211 2022; Muralikrishnan et al. 2022; Yang et al. 2024]. Similar to Gen-
 212 Corres [Yang et al. 2024], GenAnalysis enforces geometry-aware
 213 regularization terms to train the shape generator. However, the
 214 novelty of GenAnalysis lies in a novel deformation model that is
 215 suitable for man-made shapes that exhibit heterogeneous shape
 216 variations, using the shape generator to obtain structurally similar
 217 shapes, and deriving segmentation cues by analyzing the tangent
 218 spaces of the shape generator.

220 2.2 Data-Driven Shape Segmentation

221 A common paradigm among non-deep learning-based joint shape
 222 segmentation approaches [Golovinskiy and Funkhouser 2009a; Hu
 223 et al. 2012; Huang et al. 2011, 2019, 2014a; Sidi et al. 2011b; Wang
 224 et al. 2013, 2014] is to promote segmentation of each shape to be
 225 consistent among a shape collection, using point-wise correspon-
 226 dences [Huang et al. 2011], primitive-level correspondences [Hu et al.
 227

229 2012; Huang et al. 2011; Sidi et al. 2011b], or functional maps [Huang
 230 et al. 2019, 2014a; Ovsjanikov et al. 2012b; Wang et al. 2013, 2014].
 231 However, all these approaches only exhibit limited performance on
 232 heterogeneous shape collection, where computing consistent corre-
 233 spondences between highly dissimilar shapes is very challenging.
 234

235 Deep learning based methods are introduced to handle joint shape
 236 segmentation for its improved generalization ability on heteroge-
 237 neous shape collection. Most of the methods for learning consistent
 238 segmentation are supervised learning[Kalogerakis et al. 2017; Qi
 239 et al. 2017a,b; Zhao et al. 2019]. To alleviate the need for 3D anno-
 240 tation, unsupervised learning methods learn deformable part tem-
 241 plates to fit input shapes for shape segmentation. These approaches
 242 use a variety of part abstractions, including cuboids [Sun et al. 2019;
 243 Tulsiani et al. 2017; Yang and Chen 2021], superquadrics [Paschalidou
 244 et al. 2019], convex polygons [Chen et al. 2020; Deng et al. 2020],
 245 sphere templates [Paschalidou et al. 2021], MLPs [Chen et al. 2024,
 246 2019; Niu et al. 2022], and part point clouds [Huang et al. 2015]. Part
 247 deformations include explicit deformation fields [Huang et al. 2015],
 248 and implicit neural fields [Chen et al. 2024, 2019; Niu et al. 2022;
 249 Shuai et al. 2023]. Deformable part templates can also be encoded
 250 explicitly [Huang et al. 2015] or implicitly using neural network
 251 branches [Chen et al. 2024, 2019; Paschalidou et al. 2021; Shuai et al.
 252 2023; Tertikas et al. 2023]. However, these approaches still require
 253 some prior knowledge of the underlying part structure, e.g., the
 254 number of parts and part shapes. They exhibit limited performance
 255 in shapes with complex part variations. In contrast to GenAnalysis
 256 that learns a generative model of entire shapes, these approaches
 257 do not explicitly model correlations among parts.

258 GenAnalysis advances the state-of-the-art in unsupervised data-
 259 driven shape segmentation in two ways. First, as we will discuss
 260 next, GenAnalysis presents improves shape correspondences be-
 261 tween shapes with large structural shape variations to aggregate
 262 shape segmentation cues. Second, GenAnalysis enforces the prior
 263 knowledge about shape parts using a regularization loss. The regu-
 264 larization loss is enforced on adjacent synthetic shapes and tolerates
 265 large inter-shape variations among the input shape collection. We
 266 do not enforce specific knowledge about the number of parts, part
 267 shapes, and part configurations.

268 2.3 Inter-Shape Correspondences

269 Computing correspondences between geometric shapes is a long-
 270 standing problem in computer graphics. It is beyond the scope of
 271 this paper to provide a comprehensive review. We refer to [Sahilli-
 272 oglu 2020; Tam et al. 2013; van Kaick et al. 2010] for surveys on
 273 this topic. Most approaches [Boscaini et al. 2016; Kim et al. 2011;
 274 Melzi et al. 2019] deal with deformable shapes in which dense cor-
 275 respondences are well defined. Such approaches do not apply to
 276 man-made shapes where only sparse feature correspondences or
 277 part-level correspondences make sense. This issue is addressed in
 278 the functional map representation [Huang et al. 2014b; Ovsjanikov
 279 et al. 2012a], which provides a unified framework for encoding pixel-
 280 wise and part-wise correspondences. However, functional maps still
 281 require initial pairwise correspondences as input, which is the focus
 282 of GenAnalysis.

283 Another influential idea in correspondence computation is to
 284 compute correspondences through intermediate shapes, i.e., consis-
 285 tent shape matching [Deprelle et al. 2019; Huang and Guibas 2013a;
 286 Huang et al. 2012b; Kim et al. 2012]. The performance of consistent
 287 shape matching depends on the graph of similar shape pairs and
 288 the associated pairwise correspondences. GenAnalysis employs a
 289 shape generator to find structurally similar shapes and significantly
 290 improve pairwise correspondences.

291 More recently, neural field-based template learning has shown
 292 superior performance in finding dense correspondence for generic
 293 man-made shapes with diverse structural variation [Deng et al. 2021;
 294 Genova et al. 2019; Kim et al. 2023; Zheng et al. 2021]. Similar to
 295 co-segmentation, a common theme is to use the learned templates
 296 to derive inter shape correspondences, c.f., [Deng et al. 2021; Kim
 297 et al. 2023; Yang et al. 2024; Zheng et al. 2021]. However, these ap-
 298 proaches only exhibit limited performance on heterogeneous shape
 299 collections, as it is hard to precisely reconstruct the input shapes
 300 due to expressivity limitation of the templates. In contrast, Gen-
 301 Analysis can handle heterogeneous shape collections. We show that
 302 consistent shape matching on input correspondences derived from
 303 the shape generator significantly outperforms these approaches.

3 PROBLEM STATEMENT AND APPROACH OVERVIEW

304 This section presents the problem statement (Section 3.1) and an
 305 overview of GenAnalysis (Section 3.2).

306 3.1 Problem Statement

307 The input to GenAnalysis is a collection of man-made shapes $\mathcal{S} =$
 308 $\{S\} \subset \bar{\mathcal{S}}$ where $\bar{\mathcal{S}}$ denotes the ambient shape space. GenAnalysis
 309 aims at learning an implicit shape generator, which enables us to
 310 perform two fundamental shape analysis tasks:

- 311 • *Shape Segmentation.* For each input shape $S \in \mathcal{S}$, we want
 312 to decompose it into shape parts based on underlying shape
 313 variations provided by \mathcal{S} . We also want these segmenta-
 314 tions to be consistent between shape pairs that shared part
 315 substructures.
- 316 • *Shape Correspondence.* Given a source shape $S \in \mathcal{S}$ and a
 317 target shape $S' \in \mathcal{S}$ and a point p on S , we want to compute
 318 its correspondence on S' . The output consists of a flag that
 319 indicates whether this correspondence is well defined (due
 320 to partial similarities) and, if so, the corresponding point.

321 We also assume that we have another shape collection $\mathcal{S}_{\text{train}} =$
 322 $\{S\} \subset \bar{\mathcal{S}}$ that offers the data prior of the underlying shape space.

323 3.2 Approach Overview

324 GenAnalysis proceeds in four stages (See Figure 2 for an illustration).

325 **3.2.1 Shape Generator Learning.** The first stage learns an implicit
 326 shape generator from $\mathcal{S}_{\text{train}}$. We accomplish this by enforcing a
 327 novel as-affine-as-possible (AAAP) deformation prior on the dense
 328 correspondences computed between nearby shapes in the genera-
 329 tor’s latent space. The resulting deformation regularization loss
 330 significantly improves the quality of the shape generator, and the

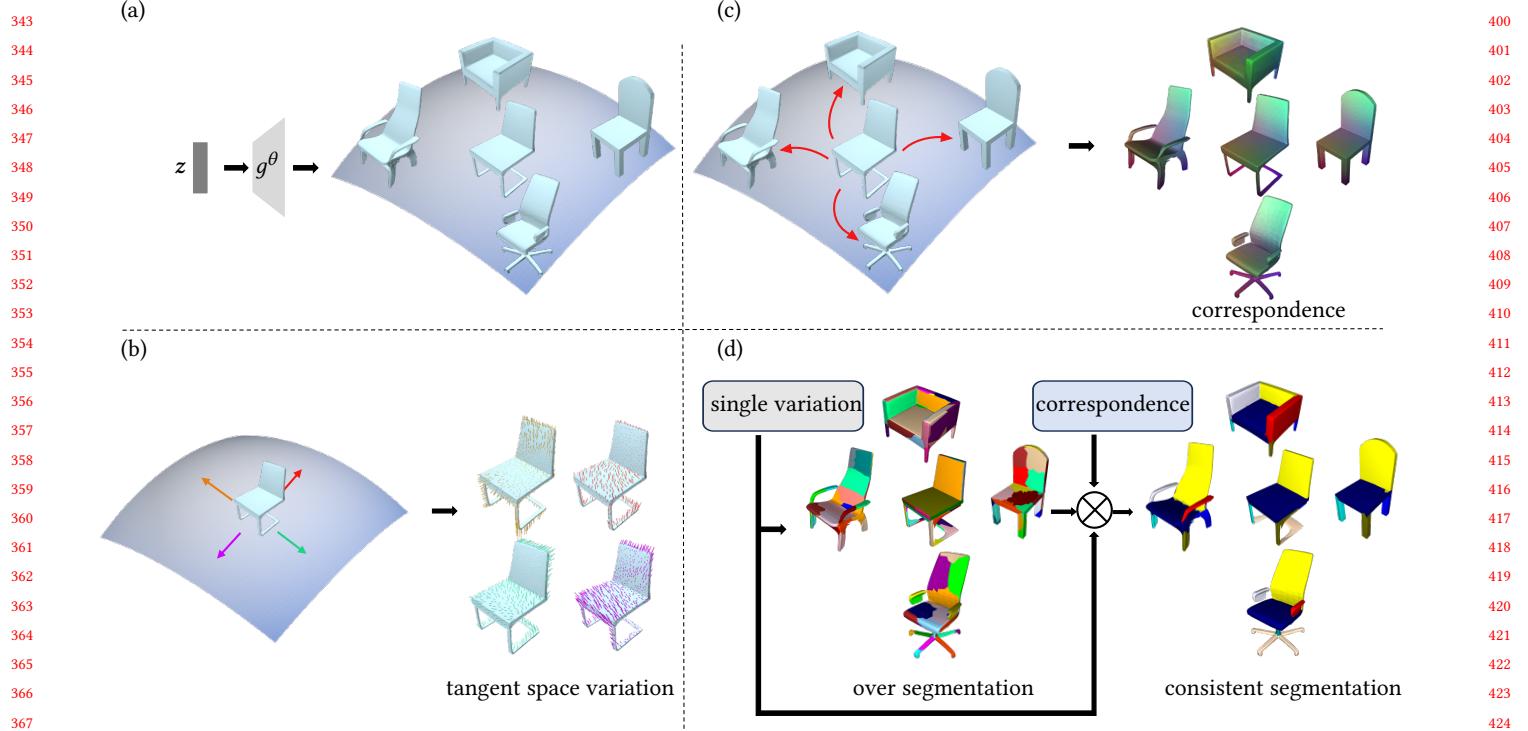


Fig. 2. The pipeline of GenAnalysis, which consists of four stages. (a) The first stage learns an implicit shape generator to fit the input shapes by combining a data loss and an as-affine-as possible deformation loss. (b) The second stage extracts single-shape segmentation cues by extracting piece-wise affine structures in vector fields of each shape derived from the tangent space of each shape. The colored vector field in each shape indicates a variation in the corresponding colored tangent direction. (c) The third stage computes consistent shape correspondences from pairwise shape correspondences obtained by composing correspondences along intermediate shapes defined by the generator. The colors-coded shapes visualize dense correspondences where points with similar color are in correspondence. (d) The last stage performs consistent segmentation using the consistent correspondences obtained in stage three to integrate single-shape segmentation cues obtained in stage two.

interpolations reveal structural variations useful for downstream tasks.

3.2.2 Shape Variation Analysis. The second stage computes single-shape segmentation cues by analyzing the tangent spaces of the learned shape manifold. This is done by solving a generalized eigenvalue problem to extract per-shape vector fields that exhibit piecewise affine structures. This is only possible because we enforce the AAAP deformation loss during training. Then, local fitting problems are solved for all vector fields to obtain the underlying infinitesimal affine transformation on each shape. These affine transformations then produce a distance function on each shape where the distance between points correlates to if they belong to the same semantic part. Finally, GenAnalysis clusters the pairwise distance into over-segments along with a single distance matrix on the over-segments by aggregating the point-based distance matrix. As we will subsequently explain, these over-segments and the associated distance matrix provide single shape segmentation cues that can be further clustered by Genanalysis to produce semantic segments of each shape,

3.2.3 Consistent Shape Correspondences. The third stage of GenAnalysis computes consistent shape correspondences that are used to aggregate single-shape segmentation cues from the second stage. We use the map synchronization algorithm of [Huang et al. 2014b] and improve the input quality with the dense correspondence generated by Genanalysis. In this context, the topology of the graph and the individual correspondence quality become the key, c.f., [Huang and Guibas 2013b]. We show that unlike matching pairs of shapes directly, which leads to poor results under large shape variations, GenAnalysis obtains matches by propagating matches on the intermediate shapes interpolated from the learned deformation model. Thanks to our AAAP deformation regularization, this guarantees that intermediate matches are obtained between structurally similar shapes, which offers more accurate dense correspondences and leads to a meaningful similarity score for each shape pair. We feed the similarity scores into [Heath et al. 2010] to identify a strongly connected graph of shape pairs with high similarity scores. This becomes the input to the map synchronization algorithm.

3.2.4 Consistent Segmentation. The last stage of GenAnalysis performs a consistent segmentation. We formulate a spectral-based consistent segmentation approach that operates on a generalized

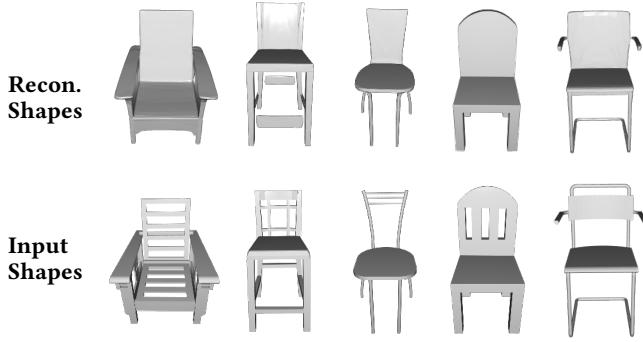


Fig. 3. **Loss of nonessential detail with regularization.** The AAAP regularization loss and the data loss remove some non essential details in the reconstructed shapes which violate the piece-wise affine assumption.

adjacency matrix among over-segments of all input shapes. The diagonal blocks store single-shape segmentation cues obtained in stage two. The off-diagonal blocks encode correspondences between over-segments converted from point-wise correspondences obtained in stage three. Spectral clustering outputs consistent groups of over-segments across the input shapes as the consistent segmentation output.

4 APPROACH

This procedure describes four stages of GenAnalysis in detail (from Section 4.1 to Section 4.4).

4.1 Shape Generator Learning

The first stage of GenAnalysis learns an implicit shape generator $g^\theta(\mathbf{x}, \mathbf{z})$ that takes a 3D position \mathbf{x} and a latent code $\mathbf{z} \in \mathcal{Z} \cong \mathbb{R}^d$ ($d = 256$ in our experiments) as input and outputs the signed distance function (SDF) value of the underlying 3D shape with the latent code \mathbf{z} . With \mathbf{z}_S we denote the latent code of S . We learn g by combining a data loss, and latent code loss, and a regularization loss:

$$\begin{aligned} \min_{\theta} \quad & \frac{1}{|S|} \sum_{S \in \mathcal{S}_{\text{train}}} d^2(S, g^\theta(\cdot, \mathbf{z}_S)) + \lambda_{\text{KL}} \text{KL}(\{\mathbf{z}_S\}, \mathcal{N}(\mathbf{0}, I_d)) \\ & + \lambda_d \mathbb{E}_{\mathbf{z} \sim \mathcal{N}(\mathbf{0}, I_d)} r(\theta, \mathbf{z}). \end{aligned} \quad (1)$$

The first term $d(\cdot, \cdot)$ in Eq.(1) aligns the generator with the input shapes. Given an input shape S , let $\mathcal{P}_S = \{(\mathbf{p}, d)\}$ collect samples \mathbf{p} and their corresponding SDF values d using the DeepSDF [Park et al. 2019] strategy. As our goal is to recover the underlying piece-wise affine deformations between the input shapes, we use a loss that encourages that the reconstruction from the generator smooths out geometric details of the input shapes whose variations violate the piece-wise affine assumption:

$$d^2(S, g^\theta(\cdot, \mathbf{z}_S)) = \sum_{(\mathbf{p}, d) \in \mathcal{P}_S} \phi(|g^\theta(\mathbf{p}, \mathbf{z}_S) - d|) \quad (2)$$

where $\phi(t) = t^4/\alpha^2$ encourages small penalty when $t \leq \alpha$.

Figure 3 shows the effects of this data loss, where the reconstructed shapes are close to the input shapes in Euclidean distance, while the shape details are smoothed out. We perform both joint-shape matching and joint-shape segmentation on the reconstructed shapes. Correspondence and segmentation results of the input shapes are derived from the nearest-neighbor correspondences between the input shapes and the reconstructed shapes.

The second term $\text{KL}(\cdot, \cdot)$ aligns the empirical latent distribution defined by the training shapes and the prior Gaussian distribution under the KL divergence measure. This enables us to define regularizations on latent codes sampled from the Gaussian distribution.

The third term $r(\cdot, \cdot)$, which is a key contribution of GenAnalysis, enforces as-affine-as-possible (AAAP) deformations among adjacent synthetic shapes (See Figure 4 for effects). This model is motivated from the assumption that inter-shape variations among man-made shapes are largely piece-wise affine, i.e., part variations are affine. Note that this assumption also models part additions and deletions, in which the scaling along one axis becomes zero. When enforcing this prior to train a shape generative model, it can recover the underlying interpolations between shapes that have small AAAP deformations. This AAAP deformation regularization is enforced between neighboring synthetic shapes (in \mathcal{Z}). Under the implicit representation, it can tolerate complex part variations, e.g., between a swivel chair base to a four-leg chair base.

4.1.1 AAAP regularization. Our formulation takes inspiration from GenCorres [Yang et al. 2024], which enforces an as-rigid-as-possible deformation prior on implicit deformable shape generators. The contribution of GenAnalysis is a formulation that applies to man-made shapes and enforces AAAP deformation under a robust norm to promote piece-wise affine structures.

Specifically, we first discretize $g^\theta(\mathbf{x}, \mathbf{z}) = 0$ using a mesh with $n = 2000$ vertices $\mathbf{p}_i^\theta(\mathbf{z})$, $1 \leq i \leq n$, e.g., via Marching Cube [Lorensen and Cline 1987]. To define the AAAP deformation prior, our goal is to determine the corresponding location $\mathbf{p}_i^\theta(\mathbf{z}^c + \epsilon \mathbf{v})$ on the neighboring implicit surface $g^\theta(\mathbf{x}, \mathbf{z} + \epsilon \mathbf{v}) = 0$, where \mathbf{v} is the direction of the perturbation and $\epsilon = 10^{-3}$ is an infinitesimal value. Approximate $\mathbf{p}_i^\theta(\mathbf{z} + \epsilon \mathbf{v}) \approx \mathbf{p}_i^\theta(\mathbf{z}) + \epsilon \mathbf{d}_i^\theta(\mathbf{z})$. The technical challenge for correspondence computation is that the implicit surface representation only provides one constraint on $\mathbf{d}_i^\theta(\mathbf{z})$:

$$\frac{\partial g^\theta}{\partial \mathbf{x}} (\mathbf{p}_i^\theta(\mathbf{z}), \mathbf{z})^T \mathbf{d}_i^\theta(\mathbf{z}) + \frac{\partial g^\theta}{\partial \mathbf{z}} (\mathbf{p}_i^\theta(\mathbf{z}), \mathbf{z})^T \mathbf{v} = 0 \quad (3)$$

To address the uniqueness problem in $\mathbf{d}_i^\theta(\mathbf{z}^c)$, GenCorres [Yang et al. 2024] solves a linearly constrained quadratic optimization problem to find $\mathbf{d}_i^\theta(\mathbf{z}^c)$ jointly. We adopt this idea and introduce a deformation model that applies to man-made shapes that exhibit approximate affine deformations in a piece-wise manner.

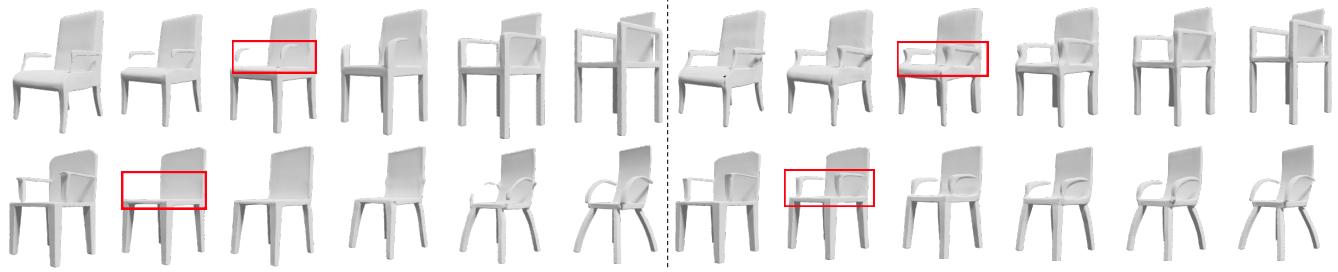


Fig. 4. AAAP regularization leads to more meaningful interpolations. (Left) Shape interpolations from the shape generator without AAAP regularization. (Right) Shape interpolations from the shape generator with AAAP regularization. The chair arm is structurally preserved with AAAP regularization.

Specifically, we associate each vertex $\mathbf{p}_i^\theta(z)$ with a local transformation $I_3 + A_i \in \mathbb{R}^{3 \times 3}$. For the sake of optimization, we reparameterize A_i as

$$s_i I_3 + \mathbf{c}_i \times + \begin{pmatrix} \frac{a_i^1}{\sqrt{2}} - \frac{a_i^2}{\sqrt{6}} & 0 & 0 \\ 0 & \frac{2a_i^2}{\sqrt{6}} & 0 \\ 0 & 0 & -\frac{a_i^1}{\sqrt{6}} - \frac{a_i^2}{\sqrt{6}} \end{pmatrix} + \begin{pmatrix} 0 & a_i^3 & a_i^4 \\ a_i^3 & 0 & a_i^5 \\ a_i^4 & a_i^5 & 0 \end{pmatrix} \quad (4)$$

Denote $\mathbf{a}_i = (a_i^1, a_i^2, a_i^3, a_i^4, a_i^5)^T$, it is clear that \mathbf{c}_i , s_i , and \mathbf{a}_i describe the rotation, similarity, and remaining affine component, in which \mathbf{a}_i forms a orthogonal basis. As we shall see immediately, a nice property of this parametrization is that it is easy to regularize A_i by penalizing L^2 norms of s_i and \mathbf{a}_i , while maintaining a quadratic total objective function.

Let \mathcal{N}_i collect the neighbors of the vertex i and itself. Let $\mathbf{p}^\theta(z) \in \mathbb{R}^{3n}$ and $\mathbf{p}^\theta(z) + \epsilon \mathbf{d}^\theta(z) \in \mathbb{R}^{3n}$ collect $\mathbf{p}_i^\theta(z)$ and $\mathbf{p}_i^\theta(z) + \epsilon \mathbf{d}_k^\theta(z)$, respectively. We model the deformation energy between them as

$$e(\mathbf{p}^\theta(z), \mathbf{d}^\theta(z)) := \min_{\{A_i\}} \sum_{i=1}^n \left(\sum_{j \in \mathcal{N}_i} \|A_i(\mathbf{p}_i^\theta(z) - \mathbf{p}_j^\theta(z))\right. \\ \left. - (\mathbf{d}_i^\theta(z) - \mathbf{d}_j^\theta(z))\|^2 + (\mu_r s_i^2 + \mu_s \|\mathbf{a}_i\|^2) \right) \quad (5)$$

where $\mu_r = 1$ and $\mu_s = 1$ are tradeoff parameters.

Introduce $\mathbf{y} \in \mathbb{R}^{9n}$ where $\mathbf{y}_i = (s_i; \mathbf{c}_i; \mathbf{a}_i)$, $1 \leq i \leq n$. It is clear that $e(\mathbf{p}^\theta(z), \mathbf{d}^\theta(z))$ is quadratic in \mathbf{y} and $\mathbf{d}^\theta(z)$. Therefore, the resulting $e(\mathbf{p}^\theta(z), \mathbf{d}^\theta(z))$ and the optimal solution to \mathbf{y} are

$$e(\mathbf{p}^\theta(z), \mathbf{d}^\theta(z)) = \mathbf{d}^\theta(z)^T L^\theta(z) \mathbf{d}^\theta(z), \quad \mathbf{y}^* = B^\theta(z) \mathbf{d}^\theta(z) \quad (6)$$

where both $L^\theta(z) \in \mathbb{R}^{3n \times 3n}$ and $B^\theta(z) \in \mathbb{R}^{9n \times 3n}$ are sparse matrices (see Appendix A).

Combing Eq.(6) and Eq.(3), we arrive at the following linearly constrained quadratic optimization problem for solving $\mathbf{d}^\theta(z)$:

$$\mathbf{d}^\theta(z) := \lim_{\mu \rightarrow 0} \operatorname{argmin}_{\mathbf{d}} \mathbf{d}^T L^\theta(z) \mathbf{d} + \mu \|\mathbf{d}\|^2 \\ \text{s.t. } C^\theta(z) \mathbf{d} = -F^\theta(z) \mathbf{v} \quad (7)$$

where $C^\theta(z) \mathbf{d} = -F^\theta(z) \mathbf{v}$ is the matrix representation of Eq.(3). $\mu \|\mathbf{d}\|^2$ avoids degenerate solutions.

It is easy to check the optimal solution to (7) is given by

$$\mathbf{d}^\theta(z) = M^\theta(z) \mathbf{v}. \quad (8)$$

$$M^\theta(z) = -(I_{3n}, 0) \begin{pmatrix} L^\theta(z) & C^\theta(z)^T \\ C^\theta(z) & 0 \end{pmatrix}^+ F^\theta(z) \quad (9)$$

where A^+ denotes the pseudo inverse of A . Note that unlike the formulation of GenCorres [Yang et al. 2024] that requires computing the pseudo inverse of a dense matrix, Eq.(9) only needs the pseudo inverse of a sparse matrix, which can be computed efficiently by LU pre-factorization.

Eq.(8) offers an explicit parameterization of $g^\theta(x, z') = 0$ for latent codes z' in the local neighborhood of z . We proceed to define the structure-preserving regularization term using this explicit parameterization. Let $\{A_i\}$ be the latent transformations decoded from Eq.(6). We define the structure-preserving regularization term $r(\theta, z)$ to enforce that $\mathbf{d}^\theta(z)$ admits a piece-wise affine structure via

$$r(\theta, z) := \int_{\mathbf{v} \in \mathcal{B}^d} \sum_{i=1}^n \sum_{j \in \mathcal{N}_i} r_{ij}^\alpha(z, \mathbf{v}) d\mathbf{v}, \quad (10)$$

$$r_{ij}(z, \mathbf{v}) := \|A_i(\mathbf{p}_i^\theta(z) - \mathbf{p}_j^\theta(z)) - (\mathbf{d}_i^\theta(z) - \mathbf{d}_j^\theta(z))\|.$$

where we set $\alpha = 1$ to promote a heavy-tailed distribution in $r_{jk}^\alpha(z, \mathbf{v})$, which corresponds to a piece-wise affine deformation.

4.2 Shape Variation Analysis

The second stage of GenAnalysis extracts single-shape segmentation cues by analyzing variations of each shape derived from the learned shape generator in the first stage. For each shape S , let $\mathcal{P}_S = \{\mathbf{p}_i\}$ be n samples on the surface of the corresponding reconstruction $g^\theta(\cdot, z_S)$. Our goal is to decompose $\mathcal{P}_S = \mathcal{P}_S^1 \cup \dots \cup \mathcal{P}_S^m$ into m over-segments ($m = 60$ in our experiments) and compute a distance matrix $D_O \in \mathbb{R}^{m \times m}$ that encodes the affinity scores between the over-segments.

To do this, we first calculate a distance matrix $D_P \in \mathbb{R}^{n \times n}$ among all sample points. D_P summarizes piece-wise affine structures in vector-fields $\mathbf{u}_l = M^\theta(z_S) \mathbf{v}_l$, $1 \leq l \leq L$ of \mathcal{P}_S ($L = 20$ in our implements), where $M^\theta(z_S)$ is defined in Eq.(9). Motivated by the modal analysis framework described in ([Huang et al. 2009]), we compute \mathbf{v}_l , $1 \leq l \leq L$ as the smallest eigenvectors of

$$H^\theta(z_S) = M^\theta(z_S)^T L^\theta(z_S) M^\theta(z_S).$$



Fig. 5. (Left) Leading vector fields of example shapes. (Right) Visualizations of distance functions of samples colored in gold.

Intuitively, \mathbf{u}_l are those vector fields with small AAAP deformation energies, which possess affine deformation structures in parts. Let λ_l be the eigenvalue of \mathbf{u}_l . Similar to [Huang et al. 2009], we weight each \mathbf{u}_l using $w_l = \frac{\lambda_1(H^\theta(z_S))}{\lambda_l(H^\theta(z_S))}$, so that vector fields with small deformation energies have larger weights. Figure 5(Left) shows vector fields of an example shape.

For each sample \mathbf{p}_i and each vector field $\mathbf{u}_l = (\mathbf{u}_{l1}; \dots; \mathbf{u}_{ln})$, we fit an affine transformation A_{li}, \mathbf{b}_{li} to $\mathbf{u}_{lj}, j \in \mathcal{N}_i$:

$$A_{li}, \mathbf{b}_{li} = \underset{A, \mathbf{b}}{\operatorname{argmin}} \sum_{j \in \mathcal{N}_i} \|A\mathbf{p}_j + \mathbf{b} - \mathbf{u}_{lj}\|^2. \quad (11)$$

For any sample \mathbf{p}_j , the residual $\epsilon_{lij} = \|A_{li}\mathbf{p}_j + \mathbf{b}_{li} - \mathbf{u}_{lj}\|$ reveals whether \mathbf{p}_j and \mathbf{p}_i belong to the same underlying semantic part or not. Specifically, when ϵ_{lij} is small, \mathbf{p}_j and \mathbf{p}_i are probably in the same part. In contrast, they are likely in different parts when ϵ_{lij} is large. Therefore, we define

$$D_P(i, j) = \sum_{l=1}^L w_l \epsilon_{lij}^2.$$

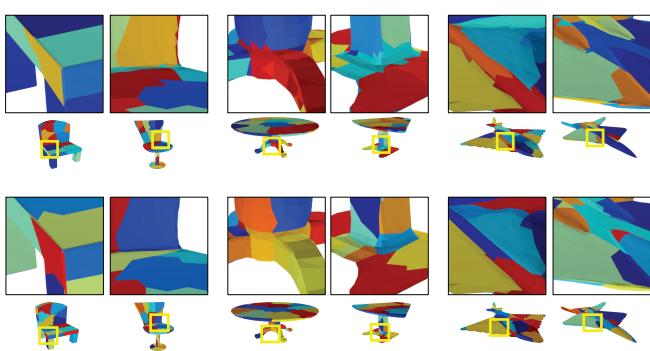


Fig. 6. Comparison between over-segments derived from different strategies. (Top) Over-segments derived from angles between adjacent faces [Golovinskiy and Funkhouser 2008]. (Bottom) Over-segments using our approach, which offer more meaningful segment boundaries.

Figure 5(Right) visualize distances of $D_P(i, j)$ when fixing i while varying j . We can see that the resulting distance field is indeed part-aware.

To compute the over-segments $O_S = \{v_S^j | 1 \leq j \leq m\}$ of S , we convert D_P into an affinity matrix A_P where $A_P(i, j) = \exp(-D_P(i, j)/2\sigma^2)$ and apply NormalizedCut [Shi and Malik 2000] using A_P to obtain the over-segments O_S . The distance $D_O(s, s')$ between over-segments s and s' is given by the mean value of $D_P(i, j)$ where $i \in s$ and $j \in s'$. In our experiments, we set σ as the median of the 5% percentile of the distances of each point in D_P . The default number of over-segments is set as $m = 60$. As shown in Figure 6, we can see that the resulting over-segments are better than those derived from angles between adjacent faces [Golovinskiy and Funkhouser 2008].

4.3 Consistent Matching

The third stage of GenAnalysis computes consistent inter-shape correspondences by performing map synchronization. Map synchronization takes point-wise correspondences between pairs of shapes along a graph $\mathcal{G} = (\mathcal{S}, \mathcal{E})$ of similar shape pairs as input and outputs consistent correspondences along \mathcal{G} . As studied in [Chen et al. 2014; Huang and Guibas 2013b], map synchronization offers a perfect result if the noise level in the input is below a ratio of the algebraic connectivity of the graph. Therefore, instead of improving the map synchronization procedure, we focus on improving the input to map synchronization, which is a dominant factor of the output quality.

To this end, we introduce a novel approach for computing dense correspondences between samples \mathcal{P}_S on $S := \{x | g^\theta(\cdot, z_S)\}$ and $S' := \{x | g^\theta(\cdot, z_{S'})\}$. The idea is to use the generator $g^\theta(\cdot, z)$ to obtain K intermediate shapes $S^k := \{x | g^\theta(x, z^k) = 0\}$, where $z^k = z_S + \frac{k}{(K+1)}(z_{S'} - z_S)$. We propagate \mathcal{P}_S to S' via $S^k, 1 \leq k \leq K$. Each propagation step begins with applying (9) to obtain a displacement vector field to the current samples positions. As (9) is based on a linear approximation of the implicit surface constraint, applying the displacement vector field does not lie exactly on the next implicit

surface. Therefore, we introduce a projection step to snap the samples onto the surface. Instead of doing so for each point in isolation, we solve an optimization problem to project them consistently.

Specifically, consider a point cloud $P = \{\mathbf{p}_i, 1 \leq i \leq n\}$ in the vicinity of an implicit surface $g^\theta(\mathbf{x}, z) = 0$. Our goal is to optimize a displacement \mathbf{d}_i to each \mathbf{p}_i so that $\mathbf{p}_i + \mathbf{d}_i$ is closer to the surface. Similar to Eq. (7), we formulate the following constrained optimization problem to solve the displacement vector $\mathbf{d} = (\mathbf{d}_1; \dots; \mathbf{d}_n)$

$$\begin{aligned} \mathbf{d}^* := \lim_{\mu \rightarrow 0} \operatorname{argmin}_{\mathbf{d}} \quad & \mathbf{d}^T L(P) \mathbf{d} + \mu \|\mathbf{d}\|^2 \\ \text{s.t.} \quad & \frac{\partial g^\theta(\mathbf{p}_i, z)}{\partial \mathbf{x}} \mathbf{d}_i = -g^\theta(\mathbf{p}_i, z), \quad 1 \leq i \leq n, \end{aligned} \quad (12)$$

where $L(P)$ is defined similarly as $L^\theta(z)$ in Eq.(6) (See Appendix. A). In our implementation, we alternate between one step of propagation and one step of projection.

The output of this procedure gives for each point $\mathbf{p} \in \mathcal{P}_S$ its corresponding point on S' . When S and S' are structurally similar, such correspondences are meaningful. When S and S' are only partially similar, some of these correspondences are not well defined. We identify these correspondences by calculating the distortion of the neighborhood of each point. In contrast to computing the distortion between S and S' directly, we find that it is more stable to accumulate the distortion during the propagation procedure. For each point \mathbf{p}_i , consider its neighborhood \mathbf{p}_j^k and neighborhood \mathbf{p}_j^{k+1} at step k and step $k+1$ (after one step of propagation and projection). We first calculate the underlying affine transformation A_{ik} between them by adopting Eq. (5):

$$A_{ik} = \operatorname{argmin}_A \sum_{j \in \mathcal{N}_i} \|A(\mathbf{p}_j^k - \mathbf{p}_i^k) - (\mathbf{p}_j^{k+1} - \mathbf{p}_i^{k+1})\|^2 + (\mu_r s^2 + \mu_c \|\mathbf{a}\|^2). \quad (13)$$

We then define the distortion as

$$e_{ik} := \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} \|A(\mathbf{p}_j^k - \mathbf{p}_i^k) - (\mathbf{p}_j^{k+1} - \mathbf{p}_i^{k+1})\|.$$

The accumulated distortion is then defined as

$$e_i = \sum_k e_{ik}.$$

The similarity weight of the i -th sample point is then given by

$$w(\mathbf{p}_i) = \exp(-e_i^2/2\sigma^2)$$

where σ is the median of e_i . Figure 7 illustrates the similarity weights of the source shapes with respect to the target shapes. We can see that the distortions reveal structural similarities and differences.

We define the deformation distance from S to S' as the average of e_i . We feed this edge distance into [Heath et al. 2010] to optimize a strongly connected graph whose edges have small deformation distances. We then apply [Huang et al. 2014a] to this graph of pairwise shape correspondences to obtain consistent shape correspondences.

4.4 Consistent Segmentation

The last stage of GenAnalysis performs consistent segmentation by integrating single-shape segmentation cues formulated in stage II using consistent correspondences obtained in stage III. This is



Fig. 7. Color-map of similarity weights $w(\mathbf{p}_i)$ between the source shape (left) and the target shape (right) defined by the local distortions e_i of the correspondences computed using our approach. Here, σ is the median of e_i . We can see that the similarity weights characterize the structural similarities.

achieved by performing spectral clustering on an affinity matrix $A \in \mathbb{R}^{|S|m \times |S|m}$, an $|S| \times |S|$ block matrix. Each diagonal block A_i encodes the segmentation cue of $S_i \in \mathcal{S}$:

$$A_i(j, k) = \exp\left(-\frac{D_{O_i}(j, k)}{2\sigma^2}\right), \quad 1 \leq j, k \leq m. \quad (14)$$

We set $A_{ij} = 0, \forall (i, j) \notin \mathcal{E}$. Each non-empty off diagonal block $A_{ij}, (i, j) \in \mathcal{E}$ encodes the correspondences between over-segments of S_i and over-segments of S_j . Given two over-segments $s_i \in O_i$ and $s_j \in O_j$, we define their affinity score as

$$A_{ij}(s_i, s_j) = \lambda \frac{|o(s_i, s_j)|}{\max(|s_i|, |s_j|)} \cdot \underset{\mathbf{p}_i, \mathbf{p}_j \in o(s_i, s_j)}{\operatorname{mean}} \frac{(w(\mathbf{p}_i) + w(\mathbf{p}_j))}{2} \quad (15)$$

where $o(s_i, s_j)$ collects point pairs from s_i and s_j that are in correspondences; the second term in (15) calculates the mean of similarity scores of the correspondences in $o(s_i, s_j)$; λ is a hyper-parameter that balances single shape segmentation cues and consistency in segmentations. As the quality of our inter-shape correspondences is high, we set $\lambda = 2$ in our experiments.

Given A , we employ a variant of spectral clustering to compute clusters of over-segments across all input shapes, which lead to consistent segmentation. Please refer to Appendix B for details of this spectral clustering procedure.

5 EXPERIMENTAL RESULTS

We demonstrate the effectiveness of GenAnalysis and provide both quantitative and qualitative results. We begin with the experimental setup in Section 5.1. We then evaluate the shape correspondence and shape segmentation results of GenAnalysis in Section 5.2 and Section 5.3, respectively. Section 5.4 presents an ablation study.

5.1 Experimental Setup

5.1.1 Dataset. We train and evaluate GenAnalysis and baselines on ShapeNet[Chang et al. 2015] dataset. We follow the training and testing split of ShapeNetPart[Yi et al. 2016]. To evaluate shape correspondences, we conduct various surrogate tasks including part label transfer and keypoint transfer due to the lack of labeled

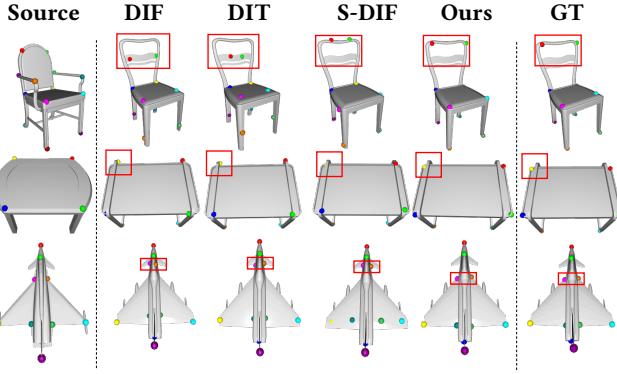


Fig. 8. **Keypoint transfer result on ShapeNet.** We transfer keypoint labels from source shapes to target shapes and compare our result with DIF[Deng et al. 2021], DIT[Zheng et al. 2021] and Semantic DIF[Kim et al. 2023].

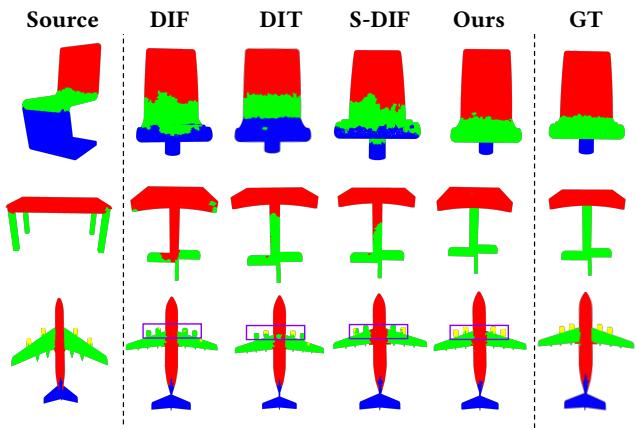


Fig. 9. **Label transfer result on ShapeNet.** We transfer ShapeNetPart[Yi et al. 2016] labels from source shapes to target shapes and compare our result with DIF[Deng et al. 2021], DIT[Zheng et al. 2021] and Semantic DIF[Kim et al. 2023].



Fig. 10. **Texture transfer result on ShapeNet.** We transfer texture from source shapes to target shapes with different structure on ShapeNet objects using corespondences generated by GenAnalysis.

datasets providing ground truth dense correspondence for direct evaluation. We use labels from ShapeNetPart[Yi et al. 2016] and KeypointNet[You et al. 2020] for the part label transfer and key point

	mIOU	Chair	Table	Airplane	
AtlasNetV2[Deprelle et al. 2019]	67.1	59.6	56.8		970
SIF[Genova et al. 2019]	61.5	62.7	54.3		971
DIT[Zheng et al. 2021]	79.6	68.7	64.4		972
DIF[Deng et al. 2021]	80.4	68.6	71.9		973
SemanticDIF[Kim et al. 2023]	80.6	69.5	71.8		974
Ours-NR	80.3	70.0	70.7		975
Ours	82.6	73.0	73.3		976

Table 1. **Label transfer result on three categories of ShapeNetPart.** The report numbers are mean IOU. Higher is better. Ours-NR means GenAnalysis without AAAP regularization.

	PCK	Chair	Table	Airplane	
AtlasNetV2[Deprelle et al. 2019]	16.6/37.1	24.5/45.3	25.7/42.4		983
SIF[Genova et al. 2019]	20.1/40.7	28.6/47.2	28.1/46.7		984
DIT[Zheng et al. 2021]	24.6/45.3	38.9/54.2	31.7/52.0		985
DIF[Deng et al. 2021]	32.9/52.5	40.5/61.4	36.9/54.7		986
SemanticDIF[Kim et al. 2023]	25.9/44.5	26.9/47.9	20.0/31.1		987
Ours-NR	28.9/49.7	37.3/57.2	35.7/61.4		988
Ours	34.9/58.6	431/64.2	40.5/67.8		989

Table 2. **Keypoint transfer result on three categories of ShapeNetPart.** We report the PCK scores with threshold of 0.01/0.02. Higher is better. Ours-NR means GenAnalysis without AAAP regularization.

transfer tasks, respectively. We evaluate shape correspondence on three popular categories, i.e., chair, table, and airplane. For shape co-segmentation, we also use labels from ShapeNetPart[Yi et al. 2016] for evaluation. We report the performance in the chair, table, and airplane categories, as do most baseline approaches, and also report the mean performance throughout the entire dataset that has 16 categories.

5.1.2 Baseline approaches. For shape correspondence, we compare our result with the template learning methods, including DIT[Zheng et al. 2021], DIF[Deng et al. 2021] and Semantic DIF[Kim et al. 2023] since they achieve the state-of-the-art performance in finding dense correspondence on ShapeNet. We also include AtlasNetV2[Deprelle et al. 2019] and SIF[Genova et al. 2019] as additional baselines.

We compare our shape co-segmentation results with BAE-Net [Chen et al. 2019], RIM-Net [Niu et al. 2022] and DAE-Net [Chen et al. 2024], which performs unsupervised shape co-segmentation using branched auto-encoding implicit network. We also compare against shape abstraction methods including VP[Tulsiani et al. 2017], HA[Sun et al. 2019] and CA[Yang and Chen 2021]. There are some other works on structure learning that could also perform unsupervised shape co-segmentation. However, they split each input shape into an excessive set of small parts[Chen et al. 2020; Deng et al. 2020], so they are not compared in our evaluation.

5.2 Analysis of Shape Correspondences

5.2.1 Part label transfer. We transfer part segment labels from the source shape to the target shape using our correspondence results. Specifically, for each of the three categories, we select 5 labeled

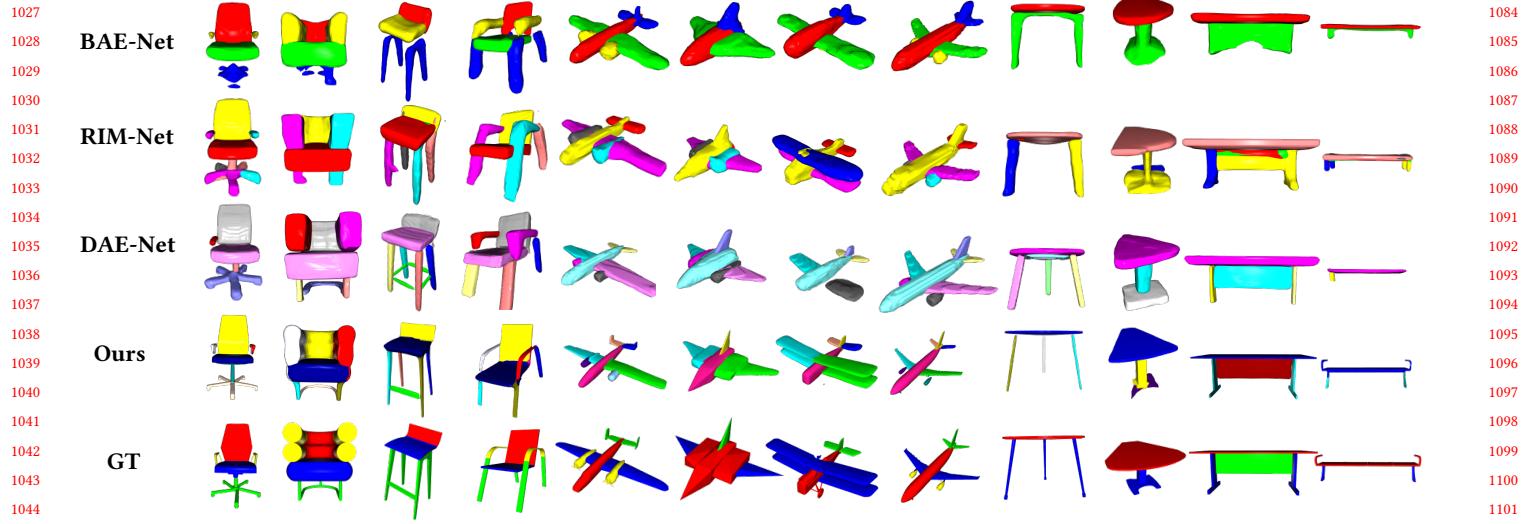


Fig. 11. Qualitative evaluation of shape co-segmentation quality on ShapeNet. We compare co-segmentation results with BAE-Net[Chen et al. 2019], RIM-Net[Niu et al. 2022] and DAE-Net[Chen et al. 2024]. The colored parts visualize segmentation consistency across different shapes in the same category

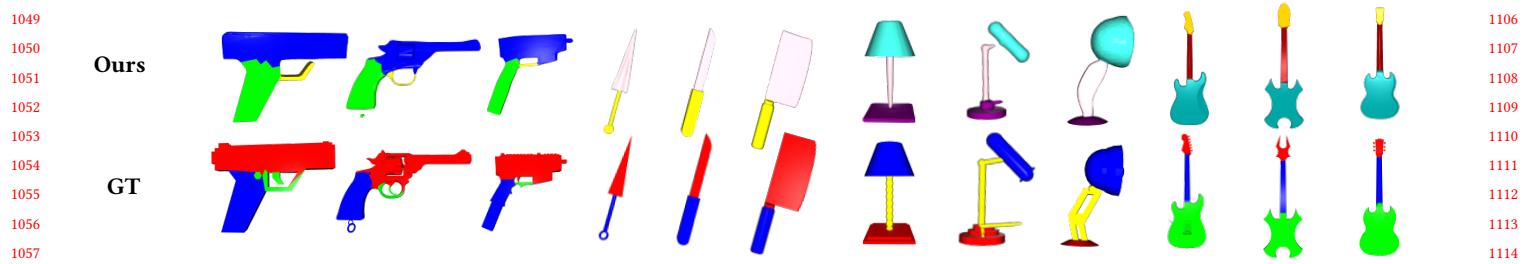


Fig. 12. More Qualitative evaluation of shape co-segmentation quality on pistol , knife, lamp and guitar category in ShapeNet The colored parts visualize segmentation consistency across different shapes in the same category

shapes as source shape, and transfer their labels to target shapes and compare with ground truth part labels to record the best results. This task can be viewed as 5-shot 3D shape segmentation using 5 samples as training data. Table 2 presents quantitative results on the 5-shot part label transfer task, measured by average per-part IOU. Our method achieves the best performance in all three categories compared to all the baseline methods. Figure 9 illustrates the qualitative results of part label transfer. AtlasNetV2 and SIF often generate inconsistent correspondences among small regions such as the engine region of airplane and arm region of chair due to the limitation of their shape representations. Template-based learning method [Deng et al. 2021; Kim et al. 2023; Zheng et al. 2021] suffers from finding accurate correspondences among structurally distinct and less common shapes due to the limited flexibility of the shared template to represent those shapes. On the other hand, GenAnalysis performs much better on those less common shapes with large geometric variations, such as a high-base chair, table. More qualitative visualization results can be found in Figure 22 in the Appendix.

5.2.2 Keypoint transfer. Table 1 presents the quantitative result in keypoint transfer task. We transfer keypoints from the source shape to the target shape using our correspondence results and measure performance by the percentage of correct keypoints (PCK) [Yi et al. 2017]. We compute the geodesic distance between the transferred keypoints and the ground truth points and report the PCK score under a distance threshold of 0.01/0.02. Table 1 shows that our method outperforms other baselines. It should be noted that the relevant improvement in keypoint transfer is higher than that in segmentation transfer. This can be understood as the fact that GenAnalysis optimizes point-wise correspondences, while segmentation transfer accuracy does not measure the quality of the correspondence in each segment. Figure 8 shows the qualitative keypoint transfer results. Our method shows better results in finding the correspondences between structurally distinct shapes, while template-based learning methods [Deng et al. 2021; Kim et al. 2023; Zheng et al. 2021] fail to find a consistent scale to fit the template and produce poor correspondence results. More qualitative visualization results can also be found in Figure 23 in the Appendix.

5.2.3 *Texture transfer.* We can also transfer mesh attributes such as texture from the source shape to the target shape using the dense correspondence generated by GenAnalysis. Figure 10 shows texture transfer result between shapes in ShapeNet. Our method preserves complex patterns in texture and transfers them to semantically consistent regions.

5.3 Analysis of Shape Segmentations

We then evaluate the performance of consistent shape segmentation. We quantitatively evaluate the consistent segmentation result by mean IOU in three main categories of ShapeNet, as commonly reported by other baseline methods. In addition, we report the mean result for all 16 categories. Table 3 shows that our result achieves higher accuracy in chair, table, airplane, and mean in all categories, compared to the baseline methods. The improvement in table is the highest, as table shapes satisfy the piece-wise affine assumption the best among all categories. Although both chair and airplane have shapes with complex part variations, the performance of GenAnalysis still leads to satisfactory improvements. The overall relative improvement of GenAnalysis from DAE-Net is higher than in chair and airplane, because the margins of improvement in other categories are higher than those two categories. We also present qualitative consistent segmentation results in Figure 11. Note that since the ground truth segmentation label is coarse, for example, the four legs of chair are grouped as one category instead of four individual category, so as the two back wings of airplane, many baseline methods especially BAE-Net fail to produce fine-grained segmentation. We also notice that RIM-Net struggles to segment shapes with rare and distinct structures, such as the three-leg table, due to the limited flexibility of the binary tree structure of its network. In contrast, GenAnalysis produces more consistent and fine-grained segmentation than other baseline methods. Furthermore, our method shows much better reconstruction quality compared to all the baseline methods. In this way, we do not need to perform additional post-processing steps used by other baselines such as projecting labels to ground truth shapes for evaluation. This is because instead of finding consistent partition of the shape in the template using a bottleneck network, as other baseline methods do, we analyze the piecewise affine variation of the shape to produce consistent segmentation, which hurts little network reconstruction ability. Figure 12 visualizes additional consistent segmentations. We show that the piecewise affine transformation assumption fits part variation among a wide range of shape category, which supports the powerful generalization ability of our method to more complex shape collections.

Please refer to Figure 15- 21 for additional qualitative results.

5.4 Ablation Study

This section presents an ablation study of GenAnalysis.

5.4.1 *Without regularization.* As show in Table 1 and Table 2, the correspondence quality of GenAnalysis in terms of mIOU in the part label transfer task and PCK scores in Keypoint transfer task decreases without the AAAP regularization loss. This is expected because, without the AAAP regularization, the interpolations from

mIOU	Chair	Table	Airplane	Mean
VP [Tulsiani et al. 2017]	64.7	62.1	37.6	–
HA [Sun et al. 2019]	75.6	67.4	60.2	–
CA [Yang and Chen 2021]	78.3	65.8	58.8	–
BAE-Net [Chen et al. 2019]	56.1	58.4	74.3	56.2
RIM-Net [Niu et al. 2022]	80.2	54.9	76.0	53.6
DAE-Net [Chen et al. 2024]	85.5	75.5	78.0	76.9
Ours-NR	58.1	49.2	52.4	–
Ours-NSV	72.3	69.2	62.4	–
Ours	88.4	82.6	79.1	80.3

Table 3. **Shape co-segmentation result on ShapeNetPart.** We show the results on chair, table, airplane, and the mean over the entire dataset. We report the mean IOU. Higer is better. GenAnalysis-NR means GenAnalysis without AAAP regularization. GenAnalysis-NSV means no structural variation segmentation cue.

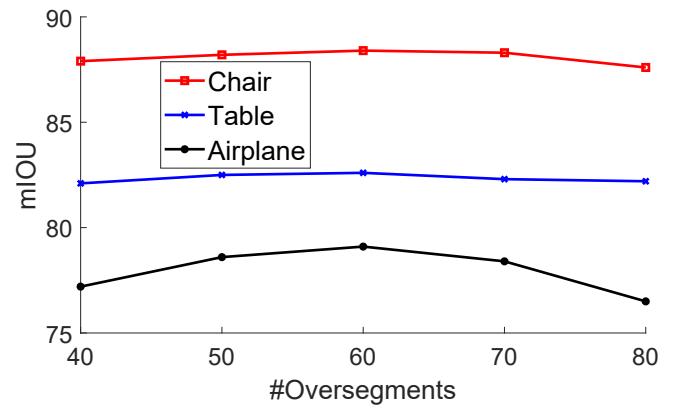


Fig. 13. Mean IOU values when varying the number of over segments. We show three categories, i.e., chair, table, and airplane.

the generator do not recover the underlying deformations between shape pairs. The resulting correspondences easily drift away. Performing map synchronization cannot address this issue, due to poor inputs.

In terms of consistent segmentation, Table 3 shows that performance drops in miou values are much more significant when the AAAP regularization term is withdrawn, that is, by 30.3%, 33.4% and 26.7% on the chair, table, and airplane. In addition to having low-quality shape correspondences in this setting, another important factor is that the shape variation segmentation cues become ineffective in this setting.

5.4.2 *Without structural variation cue.* Next, we drop the structural variation cue on each input shape and use the edge angle formulation in [Golovinskiy and Funkhouser 2008] to generate the over-segments and single shape segmentation cues. In this case, the miou values on the chair, table, and airplane drop by 16.1%, 13.4%, and 16.7%, respectively. On the one hand, this shows the effectiveness of the shape variation segmentation cues. However, the decrements are not glaring. This can be understood as the power of consistent shape segmentation using inter-shape correspondences.

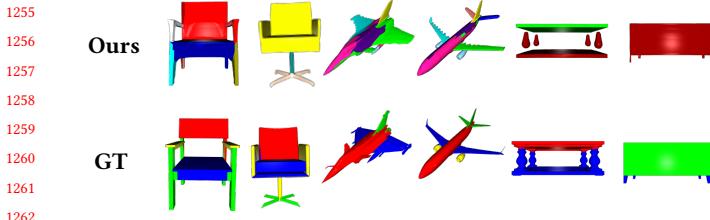


Fig. 14. Unsuccessful co-segmentation results.

5.4.3 *Varying the number of over-segments m .* We proceed to analyze the effects when varying the number of over-segments m per shape. Figure 13 shows the mIOU values of chair, table, and airplane, when varying the number of over-segments m . We can see that the values peak around $m = 60$ and drop slightly when m becomes too small or too large. We can understand this behavior as follows. When m is small, the best shape segmentation offered by the over-segments is of low quality. However, the consistent shape segmentation problem, which employs spectral relaxation, is easier to solve, leading to a good approximate solution. In contrast, when m is large, the best shape segmentation offered by the over-segments is of high quality. Yet, it is difficult to obtain the optimal solution, because of the approximation nature of spectral clustering. However, the variations in mIOU are not significantly, showing the robustness of GenAnalysis.

6 CONCLUSIONS, LIMITATIONS, AND FUTURE WORK

In this paper, we have introduced GenAnalysis, a novel framework for performing joint shape analysis by learning implicit shape generators. The key idea is to enforce an as-affine-as-possible (AAAP) deformation prior among neighboring synthetic shapes by establishing correspondences between them. This allows us to establish inter-shape correspondences, to extract structurally similar shapes, and to understand shape variations. We show how to extract single-shape segmentations by recovering piece-wise affine structures from the vector fields in the tangent space of each shape. We also show how to perform consistent shape segmentation by integrating segmentation cues from single-shapes using consistent shape correspondences derived from the shape generator. For both shape matching and shape segmentation, GenAnalysis has achieved state-of-the-art results on ShapeNetPart.

Since our method relies on spectral clustering to perform consistent segmentation, for some isolated shapes where shape correspondence quality is poor, our method could perform poorly. See the chair example in Figure 14 where the arm of the chair is not distinguished from the back of the chair and the table example. In addition, due to the nature of spectral clustering, we occasionally observe over-segmentations and under-segmentations, as shown in airplane and table in Figure 14.

GenAnalysis opens the door to use shape generators to understand shape variations and establish inter-shape correspondences. There are ample opportunities for future research. On the analysis side, we would like to detect clusters of structurally similar shapes and jointly learn a shape generator of each cluster. This can be

achieved by decoupling the latent space into a geometry latent code and a structure latent code. The structure latent code can model structural variations, avoiding explicit clustering. Another direction is to explore shape generation. We propose to introduce another latent code to model shape details and enforce AAAP when varying geometry and structure latent codes.

REFERENCES

- Panos Achlioptas, Olga Diamanti, Ioannis Mitliagkas, and Leonidas J. Guibas. 2018. Learning Representations and Generative Models for 3D Point Clouds. In *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholm, Sweden, July 10-15, 2018 (Proceedings of Machine Learning Research, Vol. 80)*, Jennifer G. Dy and Andreas Krause (Eds.). PMLR, Stockholm, Sweden, 40–49. <http://proceedings.mlr.press/v80/achlioptas18a.html>
- Martin Arjovsky, Soumith Chintala, and Léon Bottou. 2017. Wasserstein Generative Adversarial Networks. In *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017 (Proceedings of Machine Learning Research, Vol. 70)*, Doina Precup and Yee Whye Teh (Eds.). PMLR, Sydney, NSW, Australia, 214–223. <http://proceedings.mlr.press/v70/arjovsky17a.html>
- Matan Atzmon, David Novotný, Andrea Vedaldi, and Yaron Lipman. 2021. Augmenting Implicit Neural Shape Representations with Explicit Deformation Fields.
- Davide Boscaini, Jonathan Masci, Emanuele Rodolà, and Michael Bronstein. 2016. Learning Shape Correspondence with Anisotropic Convolutional Neural Networks. In *Proceedings of the 30th International Conference on Neural Information Processing Systems (Barcelona, Spain) (NIPS'16)*. Curran Associates Inc., Red Hook, NY, USA, 3197–3205.
- Giorgos Bouritsas, Sergiy Bokhnyak, Stylianos Ploumpis, Michael Bronstein, and Stefanos Zafeiriou. 2019. Neural 3D Morphable Models: Spiral Convolutional Networks for 3D Shape Representation Learning and Generation. In *The IEEE International Conference on Computer Vision (ICCV)*. IEEE Computer Society, Washington, DC, USA, 7213–7222.
- Angel X. Chang, Thomas A. Funkhouser, Leonidas J. Guibas, Pat Hanrahan, Qi-Xing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu. 2015. ShapeNet: An Information-Rich 3D Model Repository.
- Xiaobai Chen, Aleksey Golovinskiy, and Thomas Funkhouser. 2009. A Benchmark for 3D Mesh Segmentation. *ACM Trans. Graph.* 28, 3, Article 73 (jul 2009), 12 pages. <https://doi.org/10.1145/1531326.1531379>
- Yuxin Chen, Leonidas J. Guibas, and Qi-Xing Huang. 2014. Near-Optimal Joint Object Matching via Convex Relaxation. In *Proceedings of the 31th International Conference on Machine Learning, ICML 2014, Beijing, China, 21-26 June 2014 (JMLR Workshop and Conference Proceedings, Vol. 32)*. JMLR.org, JMLR.org, 100–108. <http://proceedings.mlr.press/v32/chend14.html>
- Zhiqin Chen, Qimin Chen, Hang Zhou, and Hao Zhang. 2024. DAE-Net: Deforming Auto-Encoder for fine-grained shape co-segmentation. *ACM Transactions on Graphics (TOG)* 50, 6 (2024), 8.
- Zhiqin Chen, Andrea Tagliasacchi, and Hao Zhang. 2020. Bsp-net: Generating compact meshes via binary space partitioning.. In *CVPR*. Computer Vision Foundation / IEEE, Washington, DC, USA, 45–54.
- Zhiqin Chen, Kangxue Yin, Matthew Fisher, Siddhartha Chaudhuri, and Hao Zhang. 2019. BAE-Net: Branched autoencoder for shape co-segmentation.. In *CVPR*. Computer Vision Foundation / IEEE, Washington, DC, USA, 8490–8499.
- Zhiqin Chen and Hao Zhang. 2019. Learning Implicit Fields for Generative Shape Modeling. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*. Computer Vision Foundation / IEEE, Long Beach, CA, USA, 5939–5948. <https://doi.org/10.1109/CVPR.2019.00609>
- Boyang Deng, Kyle Genova, Soroosh Yazdani, Sofien Bouaziz, Geoffrey E. Hinton, and Andrea Tagliasacchi. 2020. CvxNet: Learnable Convex Decomposition. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*. Computer Vision Foundation / IEEE, Washington, DC, USA, 31–41. <https://doi.org/10.1109/CVPR42600.2020.00011>
- Yu Deng, Jiaolong Yang, and Xin Tong. 2021. Deformed Implicit Field: Modeling 3D Shapes With Learned Dense Correspondence. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*. Computer Vision Foundation / IEEE, Washington, DC, USA, 10286–10296. <https://doi.org/10.1109/CVPR46437.2021.01015>
- Theo Deprelle, Thibault Groueix, Matthew Fisher, Vladimir G. Kim, Bryan C. Russell, and Mathieu Aubry. 2019. Learning elementary structures for 3D shape generation and matching. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d’Alché-Buc, Emily B. Fox, and Roman Garnett (Eds.). Curran

- 1369 Associates, New York, NY, USA, 7433–7443. <https://proceedings.neurips.cc/paper/2019/hash/d360a502598a4b64b936683b44a5523a-Abstract.html>
- 1370 Kyle Genova, Forrest Cole, Daniel Vlasic, Aaron Sarna, William T. Freeman, and Thomas A. Funkhouser. 2019. Learning Shape Templates With Structured Implicit Functions. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*. IEEE, Washington, DC, USA, 7153–7163. <https://doi.org/10.1109/ICCV.2019.00725>
- 1371 Aleksey Golovinskiy and Thomas Funkhouser. 2008. Randomized cuts for 3D mesh analysis. In *ACM SIGGRAPH Asia 2008 papers (Singapore) (SIGGRAPH Asia '08)*. ACM, New York, NY, USA, Article 145, 12 pages. <https://doi.org/10.1145/1457515.1409098>
- 1372 Aleksey Golovinskiy and Thomas Funkhouser. 2009a. Consistent Segmentation of 3D Models. *Computers and Graphics (Shape Modeling International 09)* 33, 3 (June 2009), 262–269.
- 1373 Aleksey Golovinskiy and Thomas A. Funkhouser. 2009b. Consistent segmentation of 3D models. *Comput. Graph.* 33, 3 (2009), 262–269. <https://doi.org/10.1016/J.CAG.2009.03.010>
- 1374 Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2020. Generative Adversarial Networks. *Commun. ACM* 63, 11 (oct 2020), 139–144. <https://doi.org/10.1145/3422622>
- 1375 Amos Groppe, Lior Yariv, Niv Haim, Matan Atzmon, and Yaron Lipman. 2020. Implicit Geometric Regularization for Learning Shapes. In *ICML (Proceedings of Machine Learning Research, Vol. 119)*. PMLR, Virtual, 3789–3799.
- 1376 Kyle Heath, Natasha Gelfand, Maks Ovsjanikov, Mridul Aanjaneya, and Leonidas J. Guibas. 2010. Image webs: Computing and exploiting connectivity in image collections. In *The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, 13-18 June 2010*. IEEE Computer Society, Washington, DC, USA, 3432–3439. <https://doi.org/10.1109/CVPR.2010.5539991>
- 1377 Ruizhen Hu, Lubin Fan, and Ligang Liu. 2012. Co-Segmentation of 3D Shapes via Subspace Clustering. *Comput. Graph. Forum* 31, 5 (aug 2012), 1703–1713. <https://doi.org/10.1111/j.1467-8659.2012.03175.x>
- 1378 Haibin Huang, Evangelos Kalogerakis, and Benjamin Marlin. 2015. Analysis and Synthesis of 3D Shape Families via Deep-Learned Generative Models of Surfaces. In *Proceedings of the Eurographics Symposium on Geometry Processing (Graz, Austria) (SGP '15)*. Eurographics Association, Goslar, DEU, 25–38.
- 1379 Qixing Huang and Leonidas Guibas. 2013a. Consistent Shape Maps via Semidefinite Programming. In *Proceedings of the Eleventh Eurographics/ACMSIGGRAPH Symposium on Geometry Processing (Genova, Italy) (SGP '13)*. Eurographics Association, Goslar, DEU, 177–186.
- 1380 Qixing Huang and Leonidas J. Guibas. 2013b. Consistent Shape Maps via Semidefinite Programming. *Comput. Graph. Forum* 32, 5 (2013), 177–186. <https://doi.org/10.1111/CGF.12184>
- 1381 Qixing Huang, Xiangru Huang, Bo Sun, Zaiwei Zhang, Junfeng Jiang, and Chandrajit Bajaj. 2021. ARAPReg: An As-Rigid-As Possible Regularization Loss for Learning Deformable Shape Generators. In *ICCV*. IEEE, Washington, DC, USA, 5795–5805.
- 1382 Qixing Huang, Vladlen Koltun, and Leonidas Guibas. 2011. Joint Shape Segmentation with Linear Programming. *ACM Trans. Graph.* 30, 6 (dec 2011), 1–12. <https://doi.org/10.1145/2070781.2024159>
- 1383 Qixing Huang, Zhenxiao Liang, Haoyun Wang, Simiao Zuo, and Chandrajit Bajaj. 2019. Tensor Maps for Synchronizing Heterogeneous Shape Collections. *ACM Transaction on Graphics* 38 (2019), 106:1–106:18. Issue 4.
- 1384 Qixing Huang, Fan Wang, and Leonidas Guibas. 2014a. Functional Map Networks for Analyzing and Exploring Large Shape Collections. *ACM Trans. Graph.* 33, 4, Article 36 (July 2014), 11 pages. <https://doi.org/10.1145/2601097.2601111>
- 1385 Qixing Huang, Fan Wang, and Leonidas Guibas. 2014b. Functional Map Networks for Analyzing and Exploring Large Shape Collections. *ACM Trans. Graph.* 33, 4, Article 36 (jul 2014), 11 pages. <https://doi.org/10.1145/2601097.2601111>
- 1386 Q. Huang, G. Zhang, L. Gao, S.M. Hu, A. Butscher, and L. Guibas. 2012a. An optimization approach for extracting and encoding consistent maps in a shape collection. *ACM Transactions on Graphics* 31, 6 (2012), 167.
- 1387 Qi-Xing Huang, Martin Wicke, Bart Adams, and Leonidas J. Guibas. 2009. Shape Decomposition using Modal Analysis. *Comput. Graph. Forum* 28, 2 (2009), 407–416.
- 1388 Qi-Xing Huang, Guo-Xin Zhang, Lin Gao, Shi-Min Hu, Adrian Butscher, and Leonidas Guibas. 2012b. An Optimization Approach for Extracting and Encoding Consistent Maps in a Shape Collection. *ACM Trans. Graph.* 31, 6, Article 167 (nov 2012), 11 pages. <https://doi.org/10.1145/2366145.2366186>
- 1389 Ruqi Huang, Jing Ren, Peter Wonka, and Maks Ovsjanikov. 2020. Consistent ZoomOut: Efficient Spectral Map Synchronization. *Comput. Graph. Forum* 39, 5 (2020), 265–278. <https://doi.org/10.1111/CGF.14084>
- 1390 Oliver Van Kaick, Noa Fish, Yanir Kleiman, Shmuel Asafi, and Daniel Cohen-OR. 2015. Shape Segmentation by Approximate Convexity Analysis. *ACM Trans. Graph.* 34, 1, Article 4 (dec 2015), 11 pages. <https://doi.org/10.1145/2611811>
- 1391 Evangelos Kalogerakis, Melinos Averkiou, Subhransu Maji, and Siddhartha Chaudhuri. 2017. Deep convolutional priors for indoor scene synthesis. In *CVPR*. Computer Vision Foundation / IEEE, Washington, DC, USA, 3779–3788.
- 1392 Sihyeon Kim, Juyeon Ko, Minseok Joo, Juhan Cha, Jaewon Lee, and Hyunwoo J. Kim. 2023. Semantic-Aware Implicit Template Learning via Part Deformation Consistency.
- 1393 In *IEEE/CVF International Conference on Computer Vision, ICCV 2023, Paris, France, October 1-6, 2023*. IEEE, Washington, DC, USA, 593–603. <https://doi.org/10.1109/ICCV51070.2023.00061>
- 1394 Vladimir G. Kim, Wilmot Li, Niloy J. Mitra, Stephen DiVerdi, and Thomas Funkhouser. 2012. Exploring Collections of 3D Models Using Fuzzy Correspondences. *ACM Trans. Graph.* 31, 4, Article 54 (jul 2012), 11 pages. <https://doi.org/10.1145/2185520.2185530>
- 1395 Vladimir G. Kim, Yaron Lipman, and Thomas Funkhouser. 2011. Blended Intrinsic Maps. *ACM Trans. Graph.* 30, 4, Article 79 (jul 2011), 12 pages. <https://doi.org/10.1145/2010324.1964974>
- 1396 Jun Li, Kai Xu, Siddhartha Chaudhuri, Ersin Yumer, Hao (Richard) Zhang, and Leonidas J. Guibas. 2017. GRASS: generative recursive autoencoders for shape structures. *ACM Trans. Graph.* 36, 4 (2017), 52:1–52:14.
- 1397 Yaron Lipman and Thomas Funkhouser. 2009. Möbius Voting for Surface Correspondence. In *ACM SIGGRAPH 2009 Papers (New Orleans, Louisiana) (SIGGRAPH '09)*. Association for Computing Machinery, New York, NY, USA, Article 72, 12 pages. <https://doi.org/10.1145/1576246.1531378>
- 1398 Or Litany, Alexander M. Bronstein, Michael M. Bronstein, and Ameesh Makadia. 2018. Deformable Shape Completion With Graph Convolutional Autoencoders. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*. IEEE Computer Society, Salt Lake City, UT, USA, 1886–1895. <https://doi.org/10.1109/CVPR.2018.00202>
- 1399 Hsueh-Ti Derek Liu, Francis Williams, Alec Jacobson, Sanja Fidler, and Or Litany. 2022. Learning Smooth Neural Functions via Lipschitz Regularization. In *ACM SIGGRAPH 2022 Conference Proceedings (Vancouver, BC, Canada) (SIGGRAPH '22)*. Association for Computing Machinery, New York, NY, USA, Article 31, 13 pages. <https://doi.org/10.1145/3528233.3530713>
- 1400 William E. Lorensen and Harvey E. Cline. 1987. Marching Cubes: A High Resolution 3D Surface Construction Algorithm. *SIGGRAPH Comput. Graph.* 21, 4 (Aug. 1987), 163–169. <https://doi.org/10.1145/37402.37422>
- 1401 Simone Melzi, Jing Ren, Emanuele Rodolà, Abhishek Sharma, Peter Wonka, and Maks Ovsjanikov. 2019. ZoomOut: Spectral Upsampling for Efficient Shape Correspondence. *ACM Trans. Graph.* 38, 6, Article 155 (nov 2019), 14 pages. <https://doi.org/10.1145/3355089.3356524>
- 1402 Kaichun Mo, Paul Guerrero, Li Yi, Hao Su, Peter Wonka, Niloy J. Mitra, and Leonidas J. Guibas. 2019. StructureNet: hierarchical graph networks for 3D shape generation. *ACM Trans. Graph.* 38, 6 (2019), 242:1–242:19.
- 1403 Sanjeev Muralikrishnan, Siddhartha Chaudhuri, Noam Aigerman, Vladimir G. Kim, Matthew Fisher, and Niloy J. Mitra. 2022. GLASS: Geometric Latent Augmentation for Shape Spaces. In *CVPR*. IEEE, Washington, DC, USA, 470–479.
- 1404 George Kiyohiro Nakayama, Mikaela Angelina Uy, Jiahui Huang, Shi-Min Hu, Ke Li, and Leonidas J. Guibas. 2023. DiffFacto: Controllable Part-Based 3D Point Cloud Generation with Cross Diffusion. In *IEEE/CVF International Conference on Computer Vision, ICCV 2023, Paris, France, October 1-6, 2023*. IEEE, Washington, DC, USA, 14211–14221. <https://doi.org/10.1109/ICCV51070.2023.01311>
- 1405 Andy Nguyen, Mirela Ben-Chen, Katarzyna Welnicka, Yinyu Ye, and Leonidas J. Guibas. 2011. An Optimization Approach to Improving Collections of Shape Maps. *Comput. Graph. Forum* 30, 5 (2011), 1481–1491. <https://doi.org/10.1111/J.1467-8659.2011.02022.X>
- 1406 Chengjie Niu, Manyi Li, Kai Xu, and Hao Zhang. 2022. RIMNet: Recursive implicit fields for unsupervised learning of hierarchical shape structures.. In *CVPR*. Computer Vision Foundation / IEEE, Washington, DC, USA, 11769–11778.
- 1407 Maks Ovsjanikov, Mirela Ben-Chen, Justin Solomon, Adrian Butscher, and Leonidas Guibas. 2012a. Functional Maps: A Flexible Representation of Maps between Shapes. *ACM Trans. Graph.* 31, 4, Article 30 (jul 2012), 11 pages. <https://doi.org/10.1145/2185520.2185526>
- 1408 Maks Ovsjanikov, Mirela Ben-Chen, Justin Solomon, Adrian Butscher, and Leonidas J. Guibas. 2012b. Functional maps: a flexible representation of maps between shapes. *ACM Trans. Graph.* 31, 4 (2012), 30:1–30:11. <https://doi.org/10.1145/2185520.2185526>
- 1409 Jeong Joon Park, Peter Florence, Julian Straub, Richard A. Newcombe, and Steven Lovegrove. 2019. DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*. Computer Vision Foundation / IEEE, Washington, DC, USA, 165–174. <https://doi.org/10.1109/CVPR.2019.00025>
- 1410 Despoina Paschalidou, Ali Osman Ulusoy, and Andreas Geiger. 2021. Neural Parts: Learning Expressive 3D Shape Abstractions With Invertible Neural Networks. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*. Computer Vision Foundation / IEEE, Washington, DC, USA, 3204–3215. <https://doi.org/10.1109/CVPR46437.2021.00322>
- 1411 Despoina Paschalidou, Ali Osman Ulusoy, and Andreas Geiger. 2019. Superquadrics Revisited: Learning 3D Shape Parsing Beyond Cuboids. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*. Computer Vision Foundation / IEEE, Washington, DC, USA, 10344–10353. <https://doi.org/10.1109/CVPR.2019.01059>
- 1412 Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. 2017a. Pointnet: Deep learning on point sets for 3d classification and segmentation.. In *CVPR*. Computer Vision Foundation / IEEE, Washington, DC, USA, 652–660.
- 1413 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482

- Charles R Qi, Li Yi, Hao Su, and Leonidas J Guibas. 2017b. Pointnet++: Deep hierarchical feature learning on point sets in a metric space.. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2017*. Curran Associates, New York, NY, USA, 5099–5108.
- Anurag Ranjan, Timo Bolkart, Soubhik Sanyal, and Michael J Black. 2018. Generating 3D faces using convolutional mesh autoencoders. In *Proceedings of the European Conference on Computer Vision (ECCV)*. Springer Science+Business Media, Berlin/Heidelberg, Germany, 704–720.
- Yusuf Sahilioglu. 2020. Recent advances in shape correspondence. *Vis. Comput.* 36, 8 (2020), 1705–1721. <https://doi.org/10.1007/s00371-019-01760-0>
- Jianbo Shi and Jitendra Malik. 2000. Normalized Cuts and Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (August 2000), 888–905. Issue 8. <https://doi.org/10.1109/34.868688>
- Qingyao Shuai, Chi Zhang, Kaizhi Yang, and Xuejin Chen. 2023. DPF-Net: Combining Explicit Shape Priors in Deformable Primitive Field for Unsupervised Structural Reconstruction of 3D Objects. In *IEEE/CVF International Conference on Computer Vision, ICCV 2023, Paris, France, October 1–6, 2023*. IEEE, Washington, DC, USA, 14275–14283. <https://doi.org/10.1109/ICCV51070.2023.01317>
- Oana Sidi, Oliver van Kaick, Yanir Kleiman, Hao Zhang, and Daniel Cohen-Or. 2011a. Unsupervised Co-Segmentation of a Set of Shapes via Descriptor-Space Spectral Clustering. *ACM Trans. Graph.* 30, 6 (dec 2011), 1–10. <https://doi.org/10.1145/2070781.2024160>
- Oana Sidi, Oliver van Kaick, Yanir Kleiman, Hao Zhang, and Daniel Cohen-Or. 2011b. Unsupervised co-segmentation of a set of shapes via descriptor-space spectral clustering. *ACM Trans. Graph.* 30, 6 (2011), 126.
- Ayan Sinha, Jing Bai, and Karthik Ramani. 2016. Deep Learning 3D Shape Surfaces Using Geometry Images. In *ECCV (6) (Lecture Notes in Computer Science, Vol. 9910)*. Springer Science+Business Media, Berlin/Heidelberg, Germany, 223–240.
- Yang Song and Stefano Ermon. 2019. Generative Modeling by Estimating Gradients of the Data Distribution. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8–14, 2019, Vancouver, BC, Canada*, Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d’Alché-Buc, Emily B. Fox, and Roman Garnett (Eds.). Curran Associates Inc., Red Hook, NY, USA, 11895–11907. <https://proceedings.neurips.cc/paper/2019/hash/3001ef257407d5a371a96dc947c7d93-Abstract.html>
- Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. 2021. Score-Based Generative Modeling through Stochastic Differential Equations.
- Chun-Yu Sun, Qian-Fang Zou, Xin Tong, and Yang Liu. 2019. Learning adaptive hierarchical cuboid abstractions of 3d shape collections. *ACM Transactions on Graphics (TOG)* 38, 6 (2019), 13.
- Gary K. L. Tam, Zhi-Quan Cheng, Yu-Kun Lai, Frank C. Langbein, Yonghuai Liu, A. David Marshall, Ralph R. Martin, Xianfang Sun, and Paul L. Rosin. 2013. Registration of 3D Point Clouds and Meshes: A Survey from Rigid to Nonrigid. *IEEE Trans. Vis. Comput. Graph.* 19, 7 (2013), 1199–1217. <https://doi.org/10.1109/TVCG.2012.310>
- Qingyang Tan, Lin Gao, Yu-Kun Lai, and Shihong Xia. 2018. Variational Autoencoders for Deforming 3D Mesh Models. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18–22, 2018*. IEEE Computer Society, Salt Lake City, UT, USA, 5841–5850. <https://doi.org/10.1109/CVPR.2018.00612>
- Konstantinos Tertikas, Despoina Paschalidou, Boxiao Pan, Jeong Joon Park, Mikaela Angelina Uy, Ioannis Z. Emiris, Yannis Avrithis, and Leonidas J. Guibas. 2023. Generating Part-Aware Editable 3D Shapes without 3D Supervision. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2023, Vancouver, BC, Canada, June 17–24, 2023*. IEEE, Washington, DC, USA, 4466–4478. <https://doi.org/10.1109/CVPR52729.2023.00434>
- Shubham Tulsiani, Hao Su, Leonidas J Guibas, Alexei A Efros, and Jitendra Malik. 2017. Learning shape abstractions by assembling volumetric primitives.. In *CVPR*. Computer Vision Foundation / IEEE, Washington, DC, USA, 2635–2643.
- Oliver van Kaick, Hao Zhang, Ghassan Hamarneh, and Daniel Cohen-Or. 2010. A Survey on Shape Correspondence. In *31st Annual Conference of the European Association for Computer Graphics, Eurographics 2010 - State of the Art Reports, Norrköping, Sweden, May 3–7, 2010*, Helwig Hauser and Erik Reinhard (Eds.). Eurographics Association, Eindhoven, Netherlands, 61–82. <https://doi.org/10.2312/egst.20101062>
- Fan Wang, Qixing Huang, and Leonidas J. Guibas. 2013. Image Co-segmentation via Consistent Functional Maps. In *IEEE International Conference on Computer Vision, ICCV 2013, Sydney, Australia, December 1–8, 2013*. Computer Vision Foundation / IEEE, Washington, DC, USA, 849–856.
- Fan Wang, Qixing Huang, Maks Ovsjanikov, and Leonidas J. Guibas. 2014. Unsupervised Multi-class Joint Image Segmentation. In *2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014, Columbus, OH, USA, June 23–28, 2014*. Computer Vision Foundation / IEEE, Washington, DC, USA, 3142–3149.
- Nanyang Wang, Yinda Zhang, Zhuwen Li, Yanwei Fu, Wei Liu, and Yu-Gang Jiang. 2018. Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images. In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part XI*. Springer Science+Business Media, Berlin/Heidelberg, Germany, 55–71. https://doi.org/10.1007/978-3-030-01252-6_4
- Jiajun Wu, Chengkai Zhang, Tianfan Xue, William T. Freeman, and Joshua B. Tenenbaum. 2016. Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling. In *Proceedings of the 30th International Conference on Neural Information Processing Systems (Barcelona, Spain) (NIPS’16)*. Curran Associates Inc., Red Hook, NY, USA, 82–90.
- Haitao Yang, Xiangru Huang, Bo Sun, Chandrajit Bajaj, and Qixing Huang. 2024. GenCorres: Consistent Shape Matching via Coupled Implicit-Explicit Shape Generative Models. In *12th International Conference on Learning Representations, ICLR 2024, Vienna, Austria, May 7–11, 2024, Conference Track Proceedings*. OpenReview.net, Online, 1–20.
- Haitao Yang, Bo Sun, Liyan Chen, Amy Pavel, and Qixing Huang. 2023. GeoLatent: A Geometric Approach to Latent Space Design for Deformable Shape Generators. *ACM Trans. Graph.* 42, 6, Article 242 (dec 2023), 20 pages. <https://doi.org/10.1145/3618371>
- Kaizhi Yang and Xuejin Chen. 2021. Unsupervised learning for cuboid shape abstraction via joint segmentation from point clouds. *ACM Transactions on Graphics (TOG)* 49, 152 (2021), 11.
- Zhenpei Yang, Lihang Liu, and Qixing Huang. 2018. Learning Generative Neural Networks for 3D Colorization. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2–7, 2018*. AAAI, New York, USA, 2580–2587. <https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17379>
- Li Yi, Vladimir G. Kim, Duygu Ceylan, I-Chao Shen, Mengyan Yan, Hao Su, Cewu Lu, Qixing Huang, Alla Sheffer, and Leonidas Guibas. 2016. A Scalable Active Framework for Region Annotation in 3D Shape Collections. *ACM Trans. Graph.* 35, 6, Article 210 (dec 2016), 12 pages. <https://doi.org/10.1145/2980179.2980238>
- Li Yi, Hao Su, Xingwen Guo, and Leonidas J. Guibas. 2017. SyncSpecCNN: Synchronized Spectral CNN for 3D Shape Segmentation. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21–26, 2017*. IEEE Computer Society, Washington, DC, USA, 6584–6592. <https://doi.org/10.1109/CVPR.2017.697>
- Yang You, Yujing Lou, Chengkun Li, Zhoujun Cheng, Liangwei Li, Lizhuang Ma, Cewu Lu, and Weiming Wang. 2020. KeypointNet: A Large-Scale 3D Keypoint Dataset Aggregated From Numerous Human Annotations. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13–19, 2020*. Computer Vision Foundation / IEEE, Washington, DC, USA, 13644–13653. <https://doi.org/10.1109/CVPR42600.2020.01366>
- Amir Zadeh, Yao-Chong Lim, Paul Pu Liang, and Louis-Philippe Morency. 2019. Variational Auto-Decoder.
- Yongheng Zhao, Tolga Birdal, Haowen Deng, and Federico Tombari. 2019. 3d point capsule networks. In *CVPR*. Computer Vision Foundation / IEEE, Washington, DC, USA, 1009–1018.
- Zerong Zheng, Tao Yu, Qionghai Dai, and Yebin Liu. 2021. Deep Implicit Templates for 3D Shape Representation. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19–25, 2021*. Computer Vision Foundation / IEEE, Washington, DC, USA, 1429–1439. <https://doi.org/10.1109/CVPR46437.2021.00148>

A EXPRESSION OF THE QUADRATIC ENERGY

We can express $e(\mathbf{g}^\theta(\mathbf{z}), \mathbf{d}^\theta(\mathbf{z}))$

$$:= \min_{\mathbf{y}} \left(\begin{array}{c} \mathbf{d}^\theta(\mathbf{z}) \\ \mathbf{y} \end{array} \right)^T \left(\begin{array}{cc} K & E^\theta(\mathbf{z}) \\ E^\theta(\mathbf{z})^T & G^\theta(\mathbf{z}) \end{array} \right) \left(\begin{array}{c} \mathbf{d}^\theta(\mathbf{z}) \\ \mathbf{y} \end{array} \right). \quad (16)$$

Here $K = \text{diag}(\{|N_i|\}) \otimes I_3$; $E^\theta(\mathbf{z})$ and $G^\theta(\mathbf{z})$ are $n \times n$ sparse and diagonal block matrices, whose expressions are deferred to Appendix A.1.

As discussed in Appendix A.2, (5) is a quadratic form:

$$e(\mathbf{g}^\theta(\mathbf{z}), \mathbf{d}^\theta(\mathbf{z})) = \mathbf{d}^\theta(\mathbf{z})^T \mathbf{L}^{-\theta}(\mathbf{z}) \mathbf{d}^\theta(\mathbf{z}). \quad (17)$$

where

$$\mathbf{L}^\theta(\mathbf{z}) := K - E^\theta(\mathbf{z}) G^\theta(\mathbf{z})^{-1} E^\theta(\mathbf{z})^T. \quad (18)$$

Moreover, the optimal transformations are given by

$$\mathbf{y}^* = -G^\theta(\mathbf{z})^{-1} E^\theta(\mathbf{z})^T \mathbf{d}^\theta(\mathbf{z}). \quad (19)$$

1597 A.1 Expressions of Matrices in (16)

1598 We can parameterize the elements of A_i as follows:

$$1600 \quad \text{vec}(A_i) = \bar{J} \cdot \begin{pmatrix} s_i \\ c_i \\ a_i \end{pmatrix} = \bar{J} \mathbf{y}_i$$

1603 where

$$1605 \quad \bar{J} := \begin{pmatrix} 1 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & 0 & 0 & 0 \\ 1606 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1607 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1608 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1609 & 1 & 0 & 0 & 0 & 0 & \frac{2}{\sqrt{6}} & 0 & 0 & 0 \\ 1610 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1611 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1612 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1613 & 1 & 0 & 0 & 0 & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & 0 & 0 & 0 \end{pmatrix}.$$

1615 Let

$$1617 \quad R = \begin{pmatrix} \mu_r & 0 & 0 \\ 1618 & 0 & 0 \\ 1619 & 0 & \mu_s I_5 \end{pmatrix}.$$

1620 It follows that

$$1622 \quad e(\mathbf{g}^\theta(z), \mathbf{d}^v(z)) = \sum_{i=1}^n \left(\sum_{j \in \mathcal{N}_i} \|((\mathbf{p}_i^\theta(z) - \mathbf{p}_j^\theta(z)) \otimes)^T \bar{J} \mathbf{y}_i \right. \\ 1623 \quad \left. - (\mathbf{d}_i^v(z) - \mathbf{d}_j^v(z))^T\|^2 + \mathbf{y}_i^T R \mathbf{y}_i \right)$$

1627 It follows that K is the Laplacian matrix of the graph whose edges
1628 are $\{(i, j) | j \in \mathcal{N}_i, 1 \leq i \leq n\}$; the diagonal blocks of G^θ are

$$1630 \quad G_{ii}^\theta(z) = \bar{J}^T \left(\sum_{j \in \mathcal{N}_i} (\mathbf{p}_i^\theta(z) - \mathbf{p}_j^\theta(z)) (\mathbf{p}_i^\theta(z) - \mathbf{p}_j^\theta(z))^T \otimes I_3 \right) \bar{J} + R;$$

1633 The ij -th block of $E^\theta(z)$ are given by

$$1634 \quad E_{ij}^\theta(z) = \begin{cases} -\sum_{j' \in \mathcal{N}_i} (\mathbf{p}_i^\theta(z) - \mathbf{p}_{j'}^\theta(z))^T \otimes I_3 & j = i \\ 1635 & (\mathbf{p}_i^\theta(z) - \mathbf{p}_j^\theta(z))^T \otimes I_3 & j \in \mathcal{N}_i \setminus \{i\} \\ 1636 & 0 & \text{otherwise} \end{cases}$$

1639 A.2 Quadratic Norm Expression of Eq.(5)

1641 Given $\mathbf{d}^v(z)$, the optimal solution to

$$1643 \quad \min_{\mathbf{y}} (\mathbf{d}^v(z)^T, \mathbf{y}^T) \begin{pmatrix} K & E^\theta(z) \\ E^\theta(z)^T & G^\theta(z) \end{pmatrix} (\mathbf{d}^v(z); \mathbf{y})$$

1645 is given by

$$1647 \quad \mathbf{y} = -G^\theta(z)^{-1} E^\theta(z)^T.$$

1649 Therefore,

$$1650 \quad L^\theta(z) = K - E^\theta(z)^{-1} E^\theta(z)^T \quad (20)$$

1652 □

1654 B SPECTRAL CONSISTENT SEGMENTATION

1655 This section presents details of the spectral consistent segmentation
1656 procedure used in GenAnalysis. The procedure takes a generalized
1657 adjacency matrix $A \in \mathbb{R}^{nm \times nm}$ as input and outputs K partitions
1658 $\{1, \dots, nm\} = \mathcal{I}_1 \cup \dots \cup \mathcal{I}_K$. The optimal value for M is determined
1659 later.

1660 Denote $\mathbf{d} = A\mathbf{1}$ and $D = \text{diag}(\mathbf{d})$. We first consider the normalized
1661 adjacency matrix

$$1662 \quad \bar{L} = I_{nm} - D^{-\frac{1}{2}} A D^{-\frac{1}{2}}.$$

1663 With \mathbf{u}_l and $\bar{\lambda}_l$ as the l -th leading eigenvector and eigenvalue of \bar{L} ,
1664 where $1 \leq l \leq L$. In our experiments, we choose $L = 10$.

1666 We form an embedding space

$$1667 \quad U = (\mathbf{u}_2, \dots, \sqrt{\frac{\bar{\lambda}_2}{\bar{\lambda}_L}} \mathbf{u}_L) \in \mathbb{R}^{nm \times (L-1)}.$$

1668 where each row of U provides the coordinate of the corresponding
1669 over-segment. We then perform probabilistic K-means with an
1670 isotropic covariance matrix to compute the resulting clusters. The
1671 initial centers are determined by farthest point sampling (FPS).

1672 We determined the optimal value K with the maximum value in
1673 $\frac{r_{K-1}}{r_K}$ where r_K is the maximum distance between each point of the
1674 cluster center in FPS with K clusters.

1677 C MORE QUALITATIVE RESULTS

1678 We show consistent shape segmentation results on chair, table,
1679 airplane, guitar, knife, lamp, pistol category in ShapeNet in Figure
1680 15, 16, 17, 18, 19, 20 and 21 respectively. We show more results
1681 on part label transfer and key point transfer in Figure 22 and 23.

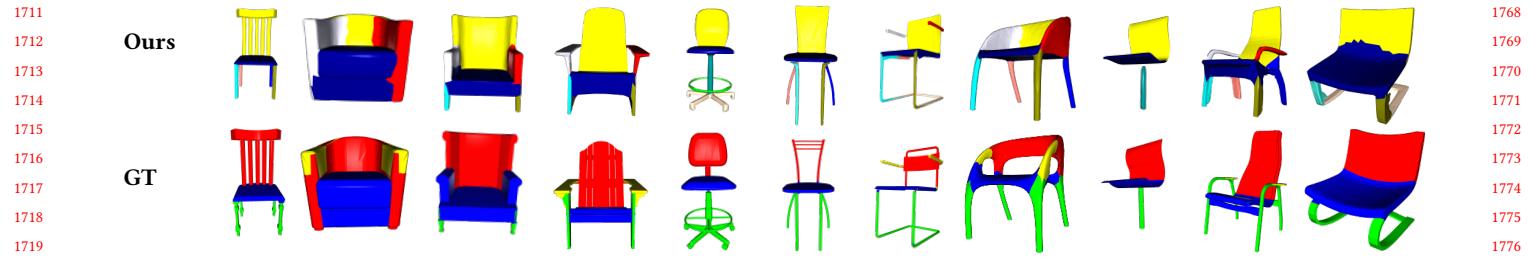


Fig. 15. More qualitative results of shape co-segmentation on chair category in ShapeNet.



Fig. 16. More qualitative results of shape co-segmentation on table category in ShapeNet.

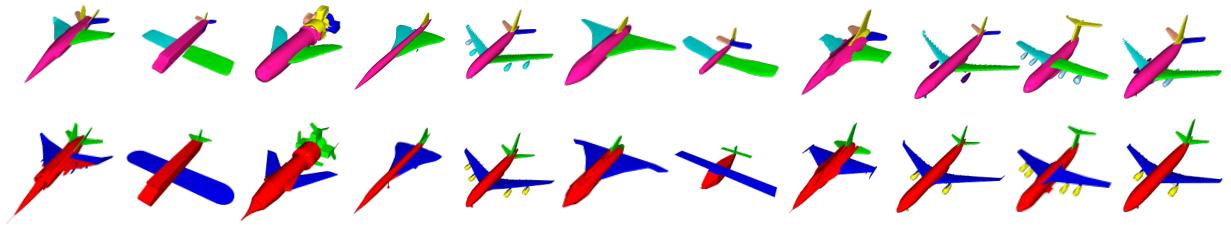


Fig. 17. More qualitative results of shape co-segmentation on airplane category in ShapeNet.

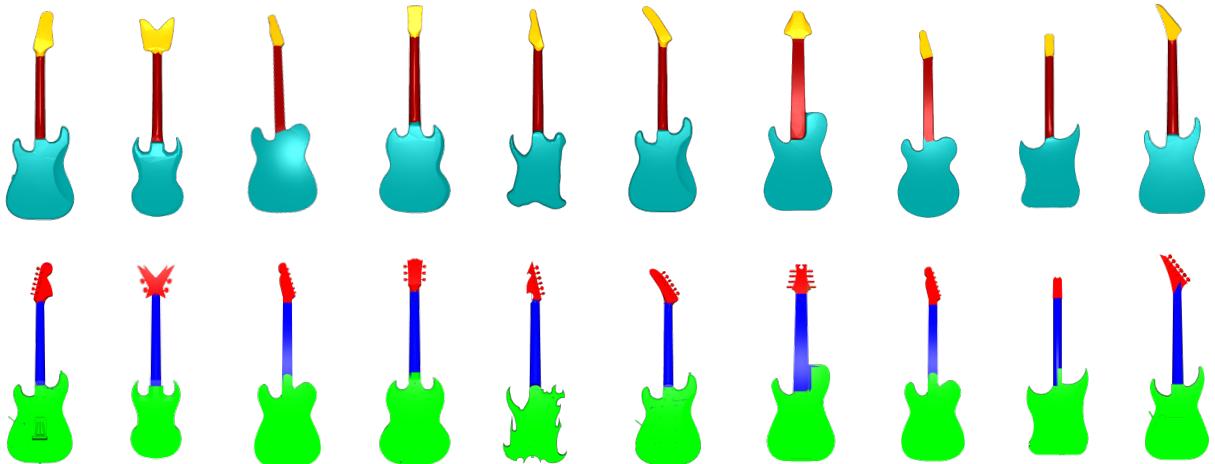


Fig. 18. More qualitative results of shape co-segmentation on guitar category in ShapeNet.

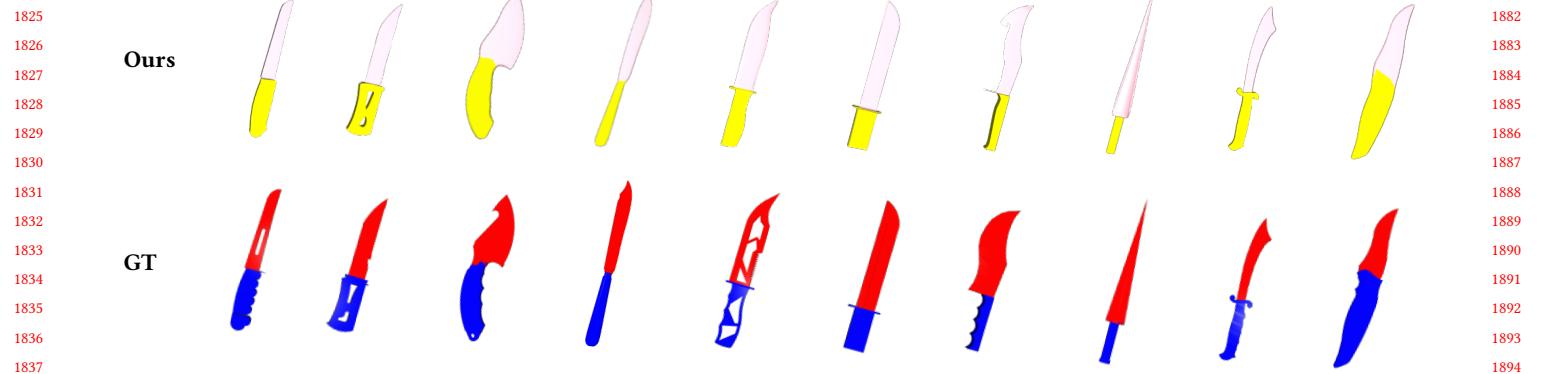


Fig. 19. More qualitative results of shape co-segmentation on knife category in ShapeNet.

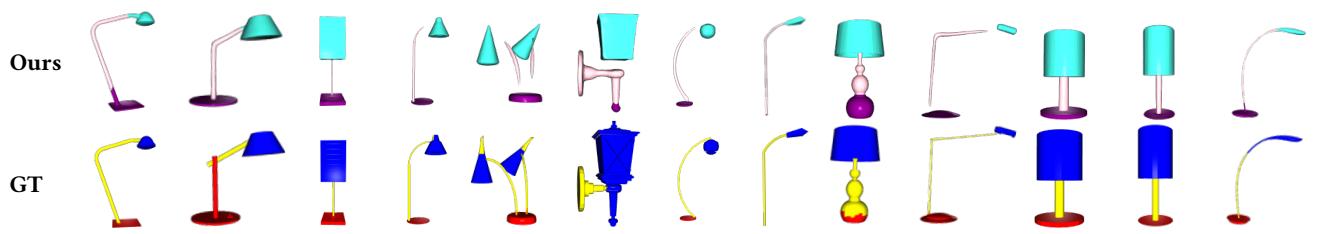
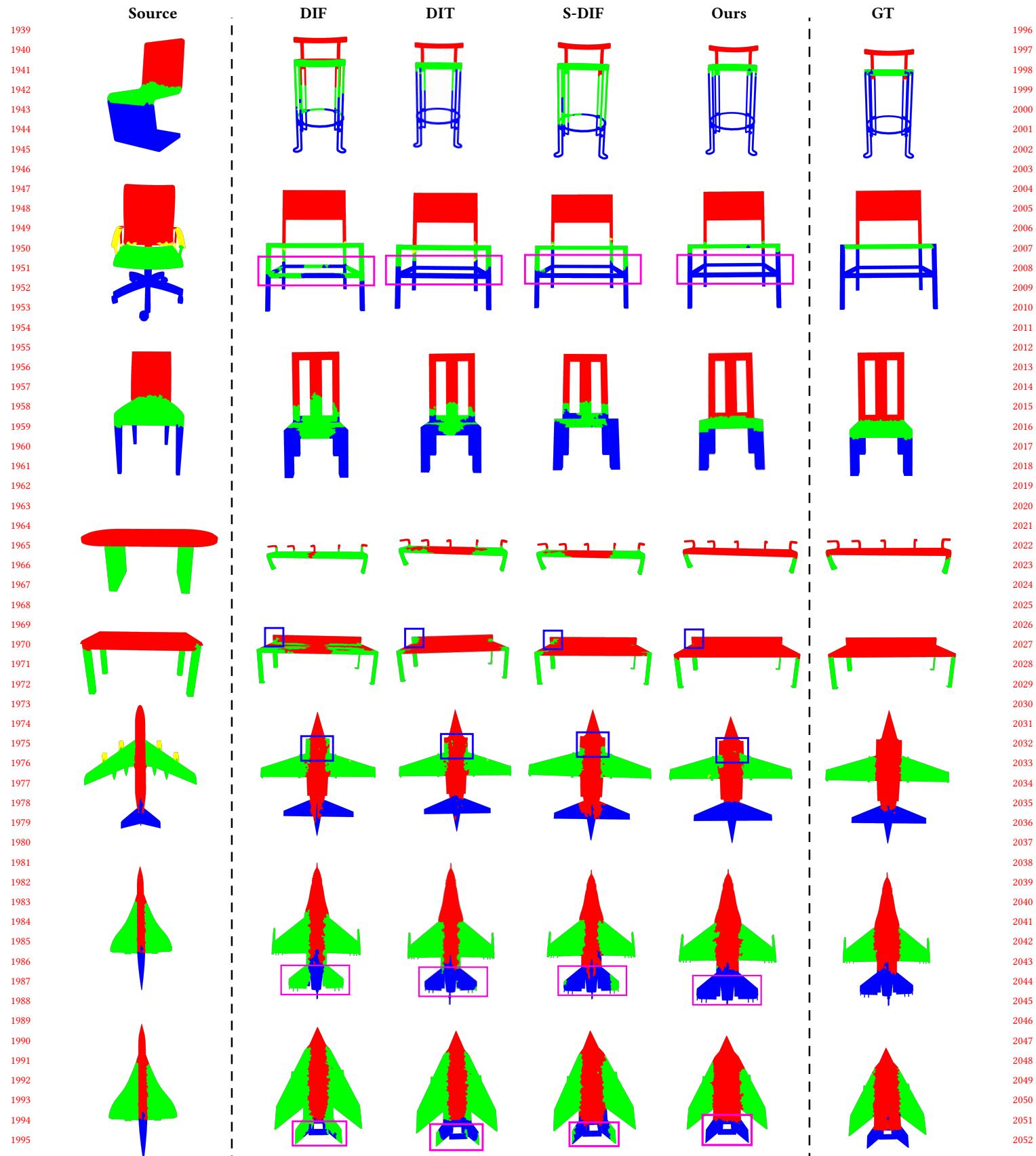


Fig. 20. More qualitative results of shape co-segmentation on lamp category in ShapeNet.



Fig. 21. More qualitative results of shape co-segmentation on pistol category in ShapeNet.



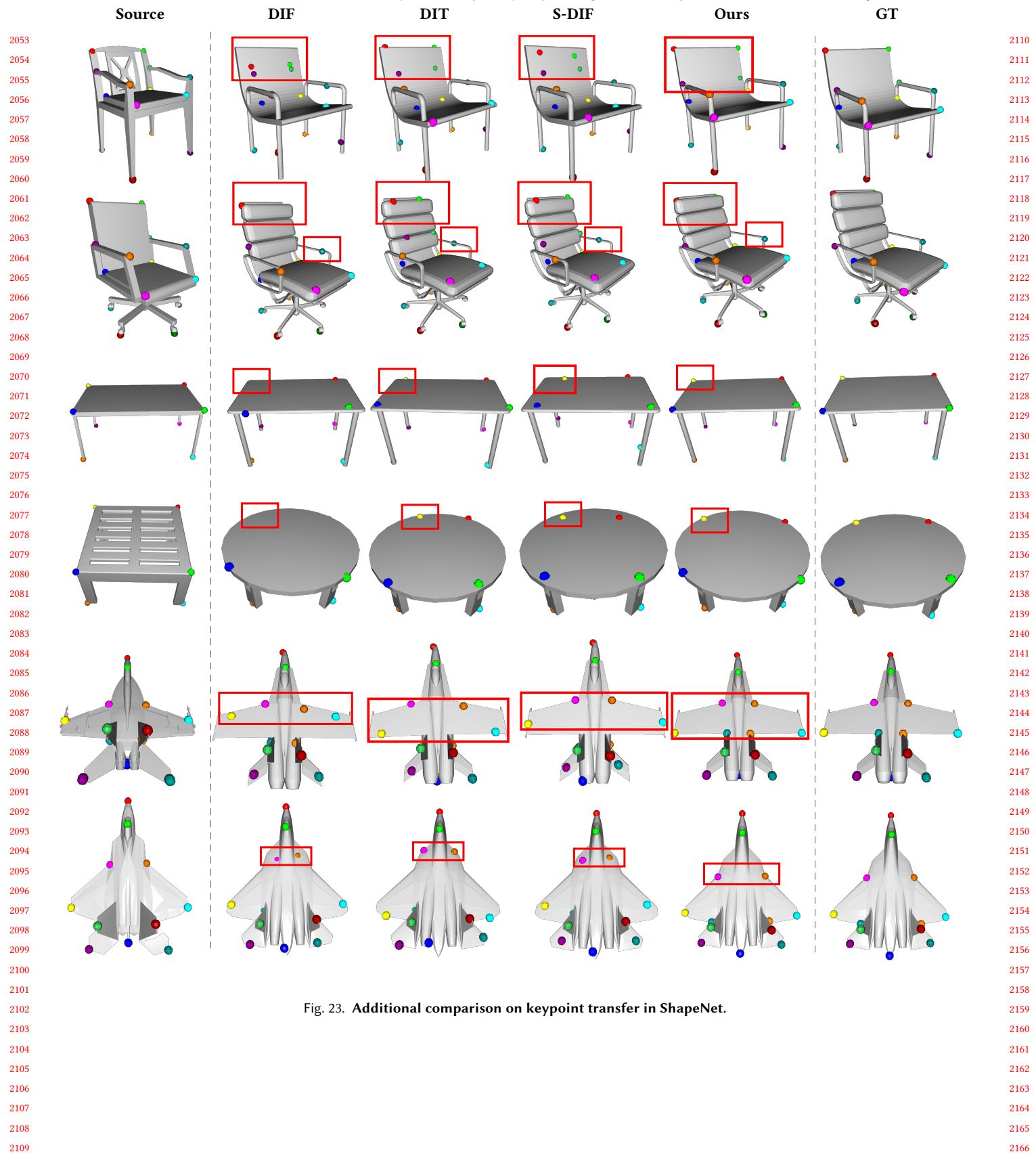


Fig. 23. Additional comparison on keypoint transfer in ShapeNet.