International Journal of Computer Vision (2023) 131:1877–1891
https://doi.org/10.1007/s11263-023-01800-2

Geometric and Statistical Models for Analysis of Two-Object Complexes

Zhiyuan Liu[1] · James Damon[2] · J. S. Marron[3] · Stephen Pizer[1]

Received: 28 April 2022 / Accepted: 12 April 2023 / Published online: 26 April 2023

Abstract
Correlated shape features involving nearby objects often contain important anatomic information. However, it is difficult to capture shape information within and between objects for a joint analysis of multi-object complexes. This paper proposes (1) capturing between-object shape based on an explicit mathematical model called a linking structure, (2) capturing shape features that are invariant to rigid transformation using local affine frames and (3) capturing Correlation of Within- and Between-Object (CoWBO) shape features using a statistical method called NEUJIVE. The resulting correlated shape features give comprehensive understanding of multi-object complexes from various perspectives. First, these features explicitly account for the positional and geometric relations between objects that can be anatomically important. Second, the local affine frames give rise to rich interior geometric features that are invariant to global alignment. Third, the joint analysis of within- and between-object shape yields robust and useful features. To demonstrate the proposed methods, we classify individuals with autism and controls using the extracted shape features of two functionally related brain structures, the hippocampus and the caudate. We found that the CoWBO features give the best classification performance among various choices of shape features. Moreover, the group difference is statistically significant in the feature space formed by the proposed methods.

Keywords Joint shape representations · Multi-object shape analysis · Local reference frames · Shape classification · S-reps

1 Introduction

Joint shape analysis of multi-object complexes in the human body can often yield additional insights relative to single object analysis. Especially interesting is to understand positional and geometric relations between multiple brain structures (Bossa & Olmos, 2007)—the positional relations include relative pose and orientation between the two objects, while the geometric relations include local shape captured using geometric properties such as boundary curvature. In the development of Autism Spectrum Disorder (ASD), for example, multiple brain subcortical structures can together have morphological variation due to the disorder. There has been a number of research projects on the association between the development of ASD and single objects (e.g., the hippocampus) morphology. Yet, few of these projects reveals the positional and geometric relations between the brain structures in ASD.

The challenges of extracting the positional and geometric relations between 3D anatomic objects come from many directions. First, instead of correlations between global morphological measures (e.g., volume-to-volume correlations), the brain structures interrelate in a subtle way and in local regions (Hazlett et al., 2017). This requires not only sensitive shape representations that can well capture local shape features but also good correspondences across samples. Second, the anatomically important interrelations between objects are often non-linear. The average of two-object complexes, for example, is by no means the concatenation of two averages of single objects. Third, such interrelations between objects can hardly be extracted via the analysis of the concatenated single objects because the concatenation can ignore the interrelations between objects. Fourth, it is difficult to separate geometric from positional relations in multi-object

Communicated by Laurent Najman, PhD, HdR.

✉ Zhiyuan Liu
zhiy@cs.unc.edu

[1] Department of Computer Science, University of North Carolina (UNC), Chapel Hill 27599, NC, USA

[2] Department of Mathematics, UNC, Chapel Hill 27599, NC, USA

[3] Department of Statistics and Operations Research, UNC, Chapel Hill 27599, NC, USA

complexes due to the difficulties of alignment (see e.g., Gollmer et al., 2012.

To address the above difficulties, this paper proposes a geometric model that explicitly captures within- and between-object shape of two-object complexes. In such a complex there exist two parallel disjoint objects. In this setting within-object features refer to single object information – including position, width, and orientation of an object; while between-object features refer to the above-mentioned relations. The features of the model are examined in a statistical analysis to find significant correlation that forms the basis of classification and hypothesis testing.

The geometric model is based on skeletal representations (s-reps) of single objects (Liu et al., 2021). In addition to within-object shape features provided by each s-rep, we also focus on how two s-reps link to each other. The link features are designed to capture both positional and geometric relations between the two objects, as described below.

To obtain desirable link features, we propose a modified version of the linking flow given by Damon and Gasparovic (2017). The key idea is to create a smooth mapping from skeletons to an external linking surface that is formed to bisect the two boundaries. Our linking flow takes in two s-reps and outputs a link vector field defined on each s-rep. From a link vector field, we sample discrete link vectors that have good correspondences across a population. These link vectors provide between-object link features, including (1) the link distances (i.e., lengths of link vectors) between a skeleton and the external linking surface and (2) the directions of the link vectors. Link lengths can capture the positional relation between the two objects, while the link directions can capture the relations of local geometry. This paper explores and exploits the above link features in statistical analysis. Different from the Damon-Gasparovic linking flow, our method yields a smooth non-branching external linking surface between the objects (the hippocampus and the caudate) in our dataset. Such consistent topology allows us to have link vectors in good correspondences across a population, which is specially advantageous in 3D anatomic shape analysis where the between-object shape often varies dramatically.

Moreover, using *local affine frames* we capture geometric features that are invariant to rigid transformations yet are sensitive to local deformations. Fitting local affine frames allows us to extract differential geometry of an s-rep. We map s-rep features to these local affine frames to obtain shape features that are invariant to rigid transformations of the multi-object complex. Not only for within-object shape features, the affine frames can be also used in capturing between-object shape features that are invariant to rigid transformations.

Finally, we jointly analyze within- and between-object shape features using a statistical method called NEUJIVE (Liu et al., 2022), yielding Correlation of Within- and Between-Object (CoWBO) shape features. While between-object shape can vary much more dramatically than within-object shape, our method can reveal the correlation regardless of the different variability of the two.

The proposed methods (illustrated in Fig. 1) show statistical advantages in classifying autism vs. non-autism with the shape of subcortical structures including the hippocampus and the caudate. We found that the CoWBO features made from (1) s-rep features of the hippocampus and (2) link features between the hippocampus and the caudate yield the best classification performance among various choices of shape features. In addition to the application, our contributions in methodology can be summarized as follows.

1. We proposed explicitly capturing positional and geometric relations between objects with modified linking structures.
2. We proposed using affine frame fields on s-reps. These affine frame fields are fitted via a deformation-based method to better characterize the geometry of objects. We map within- and between-object shape features to the fitted local affine frames such that the shape features are invariant to rigid transformations.
3. We proposed joint analysis of within- and between-object shape features using NEUJIVE. We found that the resulting correlated shape features give good performance in classifying ASD vs. non-ASD.

The remainder of this paper is organized as follows. In Sect. 2 we describe geometric and statistical models that are relevant to this paper. Section 3 discusses the joint analysis of within- and between-object shape features, which yields the CoWBO features. Then in Sect. 4 we detail the between-object shape features based on a modified linking flow. In Sect. 5 we detail within- and between-object shape features using local affine frames. In Sect. 6 we evaluate our proposed methods. Section 7 discusses our contributions and potential future work in this research direction.

2 Background

In this section we first review two geometric models on which this paper is based. The two models are (1) s-reps, which are skeletal representations capturing rich geometric features within single objects (see Sect. 2.1) and (2) linking structures that capture how neighboring objects relate to each other (see Sect. 2.2).

Instead of capturing features in a global coordinate system, we extract shape features with respect to local frames, which allow us among other things to represent the shape in a way insensitive to rigid transformations of multi-object complexes. To this end, Sect. 2.3 discusses the method of

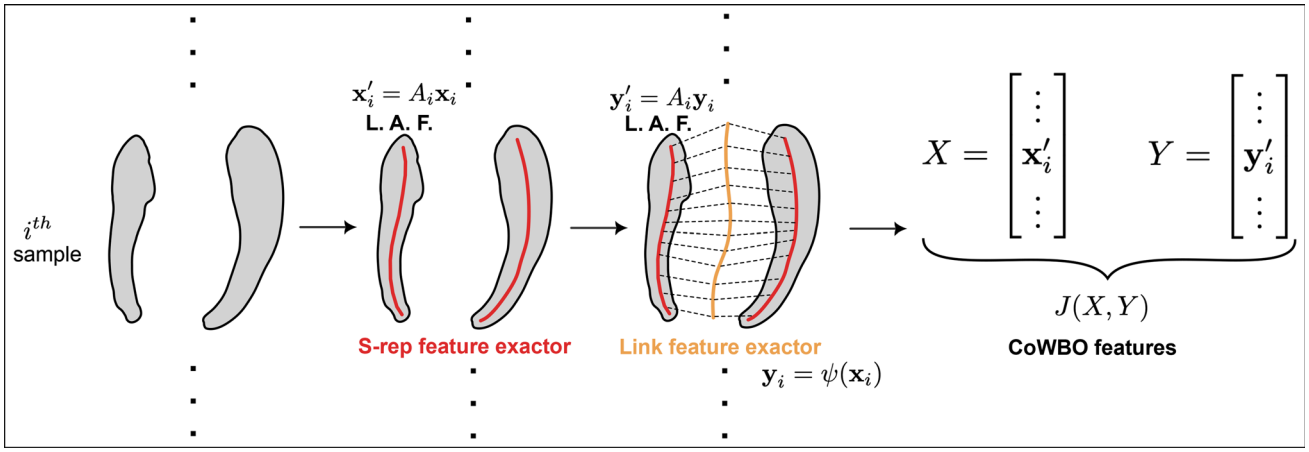

Fig. 1 Illustration of generating CoWBO features in two-object complexes. We extract s-rep features $\mathbf{x}$ from fitted s-reps (Liu et al., 2021). Then we fit a Local Affine Frame (L. A. F.) field on each skeleton (see Sect. 5). The s-rep features are mapped to the L. A. F. via linear mapping A_i, where i indexes the sample. Given the neighboring s-reps, we extract link features from the modified linking flow ψ, as discussed in Sect. 4. Likewise, these link features are mapped to the L. A. F., producing between-object features invariant to rigid transformations. Finally, we extract the joint variation from the transformed within- and between-object features (denoted as X and Y, respectively), forming CoWBO features; see Sect. 3 (Color figure online)

local orthonormal frames, which inspires our method of local affine frames in Sect. 5.1.

Last, we review a statistical method, called NEUJIVE, that can jointly analyze multi-object shape data (see Sect. 2.4). In this paper, we use NEUJIVE to extract the correlation of within-object shape variation and between-object link variation.

2.1 Skeletal Representations (S-reps)

Interiors of objects provide richer and more robust features than boundaries, especially in anatomic shape analysis, where it is difficult to obtain accurate boundaries. Figure 2 (left) gives an example caudate boundary whose tail (highlighted with the red circle) is corrupted. Such a corruption can make this sample misleading or even useless in anatomic shape analysis. In such scenarios we can still obtain robust shape features by making use of the interior geometry of the object (see e.g., Fig. 2 right).

To explore and exploit interiors of objects, Damon (Damon, 2003, 2004) has developed mathematics of general skeletal representations, paving the way for statistical shape analysis with skeletal representations. Pizer et al. (Siddiqi & Pizer, 2008; Pizer et al., 2020) took advantage of those mathematical conditions and developed "discrete s-reps" (*d-s-reps* or *s-reps*). There, an s-rep consists of a skeletal sheet and connection vectors ("spokes") pointing from the skeletal sheet to the boundary of an object.

The skeletal sheet can be understood as a deflated boundary. Strictly, a skeletal sheet is a discrete set that can be stratified into (1) a positive side that connects with one half of the object, (2) a negative side that connects with the other half of the object and (3) an edge curve that bounds the two sides. Because each side is a smooth open manifold (i.e., with no boundaries within each stratum), we can differentiate geometric entities (e.g., spokes) defined on each side. As such, the skeletal points on the two sides are also called smooth skeletal points.

Using the method of Liu et al. (2021), we can map the geometry of the skeletal sheet to that of the boundary with non-crossing spokes. As opposed to methods based on a "boundary-to-skeleton" mapping (Serra, 1986; Saha et al., 2016), Liu's method automatically fits s-reps to non-branching objects (e.g., the hippocampus, the caudate, etc.) so that the fitted s-reps can provide good correspondence across a population. The essence of that method is to fit a skeletal representation with desirable topology (i.e., non-branching topology) to target objects. A fitted s-rep gives rise to a smooth mapping, so-called *radial flow*, from the skeleton to a set of level surfaces of radial flow (see e.g., the right figure in Fig. 2). The radial distances are measured as fractions (i.e., $\tau \in [0, 1]$) of spokes' lengths. That radial flow acts as a shape function describing the shape of an object, including the interior and the boundary of the object. Those non-crossing spokes can also be understood as basis vectors of the shape function, providing rich shape features of an object.

The rich shape features from spokes with good correspondence have shown to be useful in previous research on anatomic shape analysis. Specifically, s-reps have shown promise in many tasks including classification (Hong et al., 2016), segmentation (Vicory, 2016) and hypothesis testing

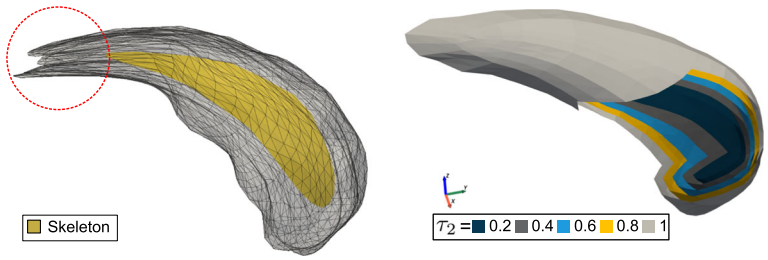

Fig. 2 Left: a caudate with partially corrupted boundary (highlighted by the red circle) from SPHARM-PDM (Styner et al., 2006). The boundary of the caudate is shown as a translucent triangular mesh. The fitted skeleton of this caudate gives more robust geometric features, as shown in yellow. Right: The level surfaces of the radial flow of an s-rep of a caudate. These level surfaces describe the interior geometry of the object (Color figure online)

(Schulz et al., 2016; Taheri & Schulz, 2022). Moreover, compared to the point distribution model (e.g., Styner et al., 2006), shape features from s-reps give more statistically efficient representations of a population of anatomic objects, as demonstrated in Tu et al. (2018) and Pizer et al. (2022).

2.2 Linking Structures

We aim to capture positional and geometric relations that are anatomically important. For example, a pair of abutting objects has close relations in terms of local geometry on their boundaries (Krishna, 2021). Additionally, the variation of each single object should be constrained by the non-penetration condition. Instead of analyzing separate objects, it is useful to explicitly analyze within- and between-object shape.

Many existing research projects (Gollmer et al., 2012; Akhoundi-Asl & Soltanian-Zadeh, 2007; Schwarz et al., 2010; Gori et al., 2017) have obtained implicit between-object relations through statistical analysis of a population of data. However, those relations often ignore important and consistent variation patterns due to sampling bias. Also, there exists research on spatial relations based on fuzzy sets (Bloch, 2005), convolutional activations (Qi et al., 2019) and lingusitic variables (e.g., e.g., Mota & Sridharan, 2018). However, such methods cannot provide good localized correspondence for statistical analysis of a population; see (Liu, 2022) for more details.

To capture between-object shape features, Damon and Gasparovic (see Damon & Gasparovic, 2017) have proposed linking structures that can be built upon general skeletal structures and hence specifically upon s-reps. In essence, the researchers regard between-object relations as shape features of the between-object space (i.e., the bounded complementary space of multiple objects). Accordingly, the linking structures are equivalent to skeletal representations of the between-object shape. The computation of an s-rep linking structure involves multiple neighboring objects, each of which has a fitted s-rep. First, a target object of interest is selected. Then, each spoke in the target object is extended to potentially intersect other extended candidate spokes. Each intersected pair (including the target spoke and a candidate spoke) is associated with an extension. Finally, the linking spoke is selected among all the candidate spokes for which the extension is minimum.

While the Damon-Gasparovic linking structures between s-reps provide a means of relating the geometry of nearby objects, it is challenging to use these linking structures in statistical analysis of 3D multi-object complexes. Due to relaxation of the Blum conditions (Liu et al., 2021), spokes in s-reps may have no candidate links at all, yielding inconsistent empty links across a population of s-reps. In addition, the *self-linking spokes* (see Sect. 4) that are two linked spokes from the same object introduce inconsistent branching topology of between-object shape, which can harm the quality of correspondences across a population. Such self-linking spokes often are associated with dents on a boundary. We, therefore, regard them as inappropriate links in analyzing between-object relations. In Sect. 4 we develop a modified version of the Damon-Gasparovic linking structure for statistical analysis of multi-object complexes.

2.3 Local Orthonormal Reference Frames

In many traditional research projects, the above geometric models are established within a global coordinate system (see e.g., Hong, 2019; Miolane et al., 2021; Vicory et al., 2018). Those geometric features are sensitive to rigid transformations including translation and rotation. Such sensitivity can hardly be removed in multi-object shape analysis because of the difficulty of aligning multi-object complexes. In this regard, local reference frames show promise of extracting shape features that are insensitive to rigid transformations in object recognition (Guo et al., 2013), shape matching (Petrelli & Di Stefano, 2011) and registration (Yang et al., 2016; Lei et al., 2017; Malassiotis & Strintzis, 2007).

Our team's previous research (see Pizer et al., 2022; Taheri & Schulz, 2022) has used local frames in describing within-

object geometry though this paper shows them also to be of use in describing between-object geometry. Those local frames are orthonormal frames defined in the interior of objects. Specifically, Pizer et al. (2022) parametrized the interior of an object using coordinates with respect to the *spine*, which is the skeleton of the skeletal sheet. Like a 2D skeleton that dwells in the quasi-medial place of a closed surface, the spine is a 1D curve that dwells in the quasi-medial place of the 2D skeletal sheet. In Pizer's work, a spine is parametrized by a cyclic variable θ, radial distance on the skeleton from the spine is specified by the parameter τ_1, and radial distance from the skeleton is specified by the parameter τ_2. Accordingly, an orthonormal local frame consists of (1) a unit vector $\widehat{\nabla\tau_1}$ (on the skeleton) that points away from the spine (2) a unit vector $\widehat{\nabla\tau_2}$ that points away from the skeleton and (3) the cross product of the two vectors, i.e., $\widehat{\nabla\theta} = \widehat{\nabla\tau_2} \times \widehat{\nabla\tau_1}$. Geometrically, the direction $\widehat{\nabla\theta}$ (on the skeleton) accounts for circulating the skeletal edge curve in a counter-clock-wise way.

Moreover, the rotation between those frames captures differential geometry within and across level surfaces of radial flow. Pizer's local frame field of a target object can be obtained by deforming a template object (e.g., an ellipsoid) along with its local orthonormal frames. Because this deformation can produce non-orthonormal frames in the target object, an orthonormalization step is needed. As a result, the orientations of these orthonormalized frames are determined by local geometry at the frames' origins.

In our work, we realized that the above local orthonormal frames can not fully capture geometric differences within and between objects. As a comparison, non-orthonormal frames resulting from the aforementioned deformation can give more expressive geometric features, as discussed in Sect. 5.

2.4 Non-Euclidean Joint and Individual Variation Explained

Because ASD simultaneously affects multiple brain structures, the correlated shape features between multiple brain structures should yield important information of the development of ASD. Also, because these multiple brain structures can be differently affected in terms of spatial scales, the correlated shape features should be effectively obtained via multi-block data analysis (see e.g., Feng et al., 2013; Hardoon et al., 2004; Lock et al., 2018), as opposed to the analysis of the concatenation of multiple objects. However, many of the existing methods (e.g., Angle-based Joint and Individual Variation Explained or AJIVE) assume data in various blocks live in Euclidean space. This assumption can fail in the analysis of shape data.

Considering the non-Euclidean properties of shape data, Liu et al. (2022) have proposed a statistical method, called Non-EUclidean Joint and Individual Variation Explained or NEUJIVE, for extracting shape correlation. There, shape descriptors of an object form a block of data (i.e., a data matrix). NEUJIVE takes in multiple blocks and outputs (1) joint components that capture the correlated shape features between objects, (2) individual components that capture the shape variation specific to each object and (3) residual components. These components together form comprehensive understanding of multi-object shape variation.

There are two critical steps in NEUJIVE. First, NEUJIVE converts each (pre-)shape space to a Euclidean space. As a result, each (pre-)shape can be represented as a vector in Euclidean space. This step is called *Euclideanization*. Second, NEUJIVE uses an existing Euclidean method AJIVE (Feng et al., 2018) to extract the joint components and individual components within the multiple Euclidean spaces. The joint components represent correlated features between objects, living in a low dimensional Euclidean subspace.

Though NEUJIVE yields statistical correlation from a population of data, the resulting between-object geometry is ignored in the analysis. The statistical analysis does not guarantee the non-penetration condition between the nearby objects. An explicit modeling of the geometric relation between objects is needed because it can not only give a direct understanding of between-object geometry, but also can respect the non-penetration condition.

In this paper, we propose forming the within- and between-object shape features as two blocks of data. The joint analysis of these two blocks via NEUJIVE, yielding CoWBO features, leads to comprehensive analysis of multi-object complexes, as discussed below.

3 CoWBO Features: Correlation of Within- and Between-Object Features

As noted above, the joint analysis of multi-object complexes gives comprehensive understanding of the population. This paper focuses on the joint analysis of within- and between-object geometry via NEUJIVE (see Sect. 2.4). In using NEUJIVE, we formulate within- and between-object shape features as two input blocks of data.

Specifically, in a population of n samples, assume each sample has two non-penetrating objects. We use $X \in \mathbb{R}^{d_1 \times n}$ to denote the feature matrix from n s-reps of an object; this feature space can be embedded in a d_1-dimensional Euclidean space. Likewise, we use $Y \in \mathbb{R}^{d_2 \times n}$ to denote the between-object feature matrix from the n samples; this feature space can be embedded in a d_2-dimensional Euclidean space. Also, we use $J(\cdot, \cdot)$ to denote the operation of extracting the correlated features by NEUJIVE. Altogether, the

CoWBO features are written as

$$J(X, Y) \in \mathbb{R}^{r \times n}, \tag{1}$$

where r is the dimension of the joint variation subspace, in which the correlated features are extracted. The dimension r is determined by NEUJIVE depending on the degree to which X and Y are correlated. We only choose the within-object shape features of a target object because in our target problem the shape features of the other object have distracting factors and thus would harm the performance of statistical tasks.

CoWBO features represent how the within- and between-object shape jointly vary. These correlated shape features are often in a rather low dimensional feature space, as compared to the original feature space of either X or Y. Moreover, CoWBO features can give robust estimation of the correlation despite that between-object shape can vary more dramatically than within-object shape.

As we will show in Sect. 6, we classify two groups of samples using CoWBO features. Sections 4 and 5 detail the methods for generating block Y and X in eq. (1).

4 Modified Linking Flow Captures Between-Object Shape Features

At first sight of multi-object complexes humans' eyes tend to pay more attention to relations between objects rather than to the details on each boundary (see Siddiqi & Pizer, 2008, Chapter 1). Inspired by this psychological study, we propose in this section a modified version of the Damon-Gasparovic linking flow, which focuses more on between-object relations than on boundary details of individual objects.

We will use a 2D example to illustrate the modification we propose. As shown in Fig. 3 left, the Damon-Gasparovic method (Damon, 2019) can develop self-links that are formed by two spokes from the same object, as highlighted in the gray box. Such self-links are found in notable dent regions. Because these self-links are barely relevant to between-object relations and because they are inconsistent across a population, we modify such links to be less sensitive to boundary details (e.g., dents), as highlighted on the right of Fig. 3.

In the following, Sect. 4.1 first gives a formal definition of our modified linking flow. Then we describe the method and algorithm of the modified linking flow. Finally, we discuss the properties of our linking flow. Section 4.2 discusses between-object shape features that are extracted from the modified linking flow.

4.1 Modified Linking Flow

Modified linking flow aims to relate the geometry from multiple objects, focusing on how the objects link with each other. The flow leads to an external linking surface that is analogous to a skeletal representation of between-object space.

Definition. Our modified linking flow ψ is defined as a smooth bijective mapping from a skeleton $\mathcal{M}$ to the external linking surface $\mathcal{L}$ between two objects, i.e.,

$$\psi : \mathcal{M} \mapsto \mathcal{L}. \tag{2}$$

In our modified linking flow, the resulting external linking surface is a smooth non-branching surface. Given a skeletal point $p \in \mathcal{M}$, the image $\psi(p) \in \mathcal{L}$ of p is called a link point. Accordingly, the vector pointing from p to $\psi(p)$ is called a link vector, whose magnitude and direction are called a link length and a link direction, respectively. The following approach can construct a mapping ψ such that (1) the mapping ψ is well-defined everywhere on an s-rep; i.e., $\psi(p) \in \mathcal{L}, \forall p \in \mathcal{M}$ and (2) the external linking surface contains no self-linking and is one-to-one.

Approach. Our method for constructing a desirable mapping ψ for every two-object complex has 3 steps (illustrated in Fig. 4):

1. Find *regularly linked spokes* using the method given by Damon (2019), as described below;
2. Infer the mapping ψ from the regularly linked spokes;
3. Apply the mapping ψ to the skeleton.

A pair of regularly linked spokes is formed by two spokes from the s-reps of two objects; the two spokes link at a place between the two objects. Specifically, in a complex of two objects Ω_1 and Ω_2, let $\mathbf{s}_1$ and $\mathbf{s}_2$ denote[1] two spokes that are in the objects Ω_1 and Ω_2, respectively. The directions of the two spokes are $\mathbf{u}_1$ and $\mathbf{u}_2$; the end points of the two spokes are $\mathbf{y}_1$ and $\mathbf{y}_2$. The two spokes are candidate linked spokes *only if* there exists a non-negative scalar t such that the extensions of the two spokes lie outside the objects and intersect at their endpoints, i.e.,

$$\mathbf{y}_1 + t\mathbf{u}_1 = \mathbf{y}_2 + t\mathbf{u}_2, \quad s.t. \quad t \geq 0, \tag{3}$$

where t denotes the equal lengths of the extension of the two spokes.

That intersection point, also called a link point, $\mathbf{y}_1 + t\mathbf{u}_1$ (or equivalently, $\mathbf{y}_2 + t\mathbf{u}_2$) is on the external linking surface. Importantly, there can exist zero or many candidate spokes in Ω_1 and Ω_2 that link to $\mathbf{s}_1$, satisfying eq. (3) with different extensions t. A linking point is the one where the length "t"

[1] We use bold letters to denote vectors in this paper.

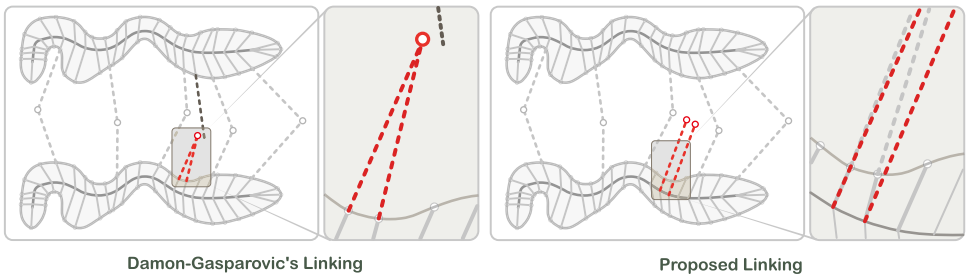

Fig. 3 Simulated links between two 2D objects. Compared to the Damon-Gasparovic linking structure (left), our modified linking structure (right) changes the link properties to avoid self-linking, as shown in the gray boxes (Color figure online)

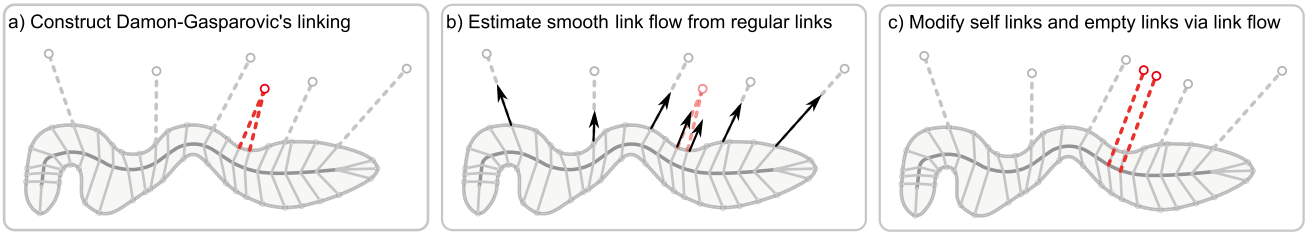

Fig. 4 Illustration of the algorithm for modified linking structures; see Algorithm 1 for more details (Color figure online)

for the pair of specific candidate linked spokes is minimal among all candidate linking points involving either of the given pair of spokes.

Next, we infer the mapping ψ using the regularly linked spokes found by the above method. Let S_1 denote a set of regularly linked spokes in Ω_1. The skeletal points of S_1 give a subset $\widehat{\mathcal{M}}$ of the skeleton $\mathcal{M}$. In addition, the link points of S_1 give a subset $\widehat{\mathcal{L}}$ of the external linking surface $\mathcal{L}$. We use the Thin Plate Spline (TPS) algorithm with landmarks on $\widehat{\mathcal{M}}$ and $\widehat{\mathcal{L}}$ to derive a diffeomorphism $\psi : \widehat{\mathcal{M}} \mapsto \widehat{\mathcal{L}}$, thereby avoiding the many-to-one links in the Damon-Gasparovic method produced by self-links; see the middle panel in Fig. 4.

Last, we apply ψ to the complement of $\widehat{\mathcal{M}}$ to form a smooth external linking surface $\mathcal{L}$. By doing so, we have a link point $\psi(p)$ for every skeletal point $p \in \mathcal{M}$. Algorithm 1 gives the detailed algorithm.

Algorithm 1 Modified linking flow

Require: A pair of s-reps fitted to two nearby objects.
1: Select one object as the target object.
2: Compute the Damon-Gasparovic linking structure P (see Damon, 2019, Section 4) for the target object.
3: Create an empty set $S_1 = \emptyset$ of pairs of points.
4: **for** each link l in P **do**
5: **if** l regularly links to another object **then**
Add the two ends of l to S_1.
6: **end if**
7: **end for**
8: Apply TPS to infer deformation ψ using pairs of points in S_1.
9: Apply ψ to all skeletal points on $\mathcal{M}$ of the target object.
10: **return** the result of $\psi(\mathcal{M})$.

Our modification makes use of regularly linked spokes to modify both the self-linking and unlinked spokes such that two-object complexes consistently yield non-branching smooth external linking surfaces. This consistency allows us to extract between-object shape features with good correspondences. Also, because the modified linking flow is derived from regularly linked spokes of two objects, the mapping ψ can capture positional and geometric relations between objects. In order to represent these relations for statistical analysis, we discretize the link vector field and extract corresponding link features from the discrete link vectors, as described below.

4.2 Link Vectors and Between-Object Shape Features

Like spokes in an s-rep (see Sect. 2.1), link vectors can be understood as a discretized vector field defined on a smooth manifold (i.e., the skeleton $\mathcal{M}$). To facilitate statistical analysis, these link vectors should be sampled to have good correspondences across a population.

To this end, we sample link vectors at each skeletal point in a discrete s-rep. These link vectors have good correspondence because the base manifold is sampled with good correspondence. Each s-rep is deformed from an s-rep of a similar ellipsoid. The skeletal points are sampled with respect to the relative positions within the ellipsoid. Such consistent discretization of skeletons gives good correspondences across samples.

As we apply TPS to produce link vectors (see Algorithm 1 steps 8 and 9), the interpolated link vectors can have directions different from the spokes associated at the skeletal points. Nevertheless, the link directions we used include the interpolated ones. In addition to the link directions, we also extract the magnitudes of the link vectors that represent distances between a skeleton and the external linking surface.

In order to obtain statistical correlation with within-object shape, it is useful to Euclideanize between-object shape features in NEUJIVE. The Euclideanization of between-object shape features should consider both Euclidean variables (i.e., log link lengths) and non-Euclidean directional data (i.e., link directions). For the link lengths, we convert these non-negative scalar values to Euclidean variables distributed in $\mathbb{R}$ with the log map (Tu et al., 2018). For the directional data that live on a unit 2-sphere, we use Principal Nested Spheres (Jung et al., 2012) to convert the data into Euclidean coordinates.

A subtlety of these features is their sensitivity to positions and orientations of objects. To overcome this problem, we capture between-object shape features in local affine frames to achieve invariance properties, as described below.

5 Local Affine Frames on Skeletal Representations

It is often challenging to align objects for analysis of geometric features, especially when there are multiple objects in every sample. For a population of multi-object complexes, aligning individual objects can sacrifice between-object relations, while aligning the combination of multiple objects can bias the analysis of individual objects.

To address the above dilemma, this section describes geometric features that are with respect to local affine frames, as described in Sect. 5.1. These features, including within- and between-object shape features (see Sect. 5.2), are insensitive to rigid transformations. Importantly, because the local affine frames for each instance of an object are fitted using a consistent deformation from an ellipsoid, they also capture local geometry which is useful in distinguishing the shape.

5.1 Local Affine Frames

In this paper, a local affine frame refers to a general frame centered at a point on a 2D surface. Such a frame consists of 3 basis vectors that are not necessarily orthogonal to each other. Moreover, the 3 frame vectors are not necessarily of unit lengths.

Specifically, at each skeletal point, we have an affine frame that allows us to map features at the skeletal point to this local frame. Let $\mathbf{x} \in \mathbb{R}^3$ be a unit directional vector (e.g., a spoke's direction) associated with a skeletal point p; assume $\mathbf{x}$ is represented in global coordinates. An affine frame at p can be represented as a matrix $A \in \mathbb{R}^{3\times3}$. In a special case where this frame is an orthonormal frame, the matrix A is an orthogonal matrix of determinant 1. Generally, we can map the directional feature $\mathbf{x}$ into the local affine frame via

$$\mathbf{x}' = A\mathbf{x}, \tag{4}$$

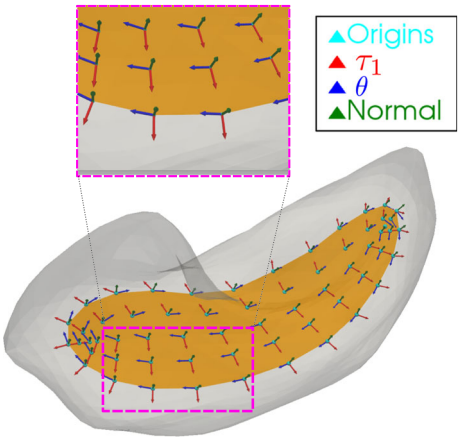

Fig. 5 Fitted local affine frames on the skeleton of a hippocampus. The skeleton is shown as the orange surface. At each skeletal point, there is an affine frame consisting of a τ_1, θ and normal vector (see the zoomed picture above). The vectors of τ_1 and θ are not necessarily orthonormal vectors (Color figure online)

where $\mathbf{x}' \in \mathbb{R}^3$ is a feature vector that is insensitive to rigid transformations.

To better capture geometry of objects, we take into account the following objectives in constructing local affine frames on s-reps:

1. the origins of these local frames should have good correspondences across samples;
2. the basis vectors of local frames within a sample should reflect within-object geometry;
3. the basis vectors of local frames across samples should capture geometric difference.

As skeletal points in s-reps provide good correspondence across a population, the first objective can be satisfied by fitting a local frame at each skeletal point in an s-rep. The second and third objectives are achieved by fitting local frames via a deformation-based method as follows. We choose an ellipsoid as a template shape which is deformed into an object. Then we construct orthonormal local frames (see Sect. 2.3) on the ellipsoidal s-rep. We construct a diffeomorphism via TPS to deform this template s-rep to the s-rep of each target object, mapping skeletal points on the ellipsoidal s-rep to corresponding skeletal points on the target s-rep. Such a deformation stretches and rotates the orthonormal frame vectors of every local frame, resulting in basis vectors of a local affine frame.

In a single 3D object like a hippocampus (see Fig. 5), the transformation between our affine frames is achieved by a linear operator that combines the rotation and the stretching of basis vectors. In Sect. 6.4 we demonstrate that the local geometry captured by local affine frames is useful in classifying the ASD and non-ASD groups.

An affine frame field is promising in multi-object shape analysis not only because it captures geometric features within objects invariant to rigid transformations, but also because it enriches between-object shape features, as detailed in the following.

5.2 Within- and Between-Object Shape Features in Local Affine Frames

Due to the above methods, there are 3 geometric entities associated with each skeletal point in an s-rep: a spoke, a link vector and an affine frame. To obtain geometric features that are invariant to rigid transformations, we map vectors at each skeletal point into the respective local affine frame using the linear transformation eq. (4).

Within-object shape features from spokes include the skeletal points, the spokes' lengths and directions. To make the coordinates of skeletal points independent of the coordinate system, we convert the global coordinates of skeletal points into the coordinates relative to the skeletal center's local frame. Moreover, we map the spoke direction unit vectors to the local frames at the tails of the associated spokes. In order to retain the unit length of the projected directional vector, we normalize the three coordinates of $\mathbf{x}'$. Likewise, we map the link features from the link vectors to the local affine frames at the tails of the associated link vectors.

The resulting within- and between-object shape features are sent to X and Y, respectively, in eq. (1) for the joint analysis. In NEUJIVE, the Euclideanization of within-object shape features is the same with the Euclideanization of between-object shape features, as discussed in Sect. 4.2.

6 Experiments

This section demonstrates our proposed methods using a database of infants' MRI brain images, consisting of two groups of infants. Section 6.1 describes the images from these two groups. In Sect. 6.2, we show the between-object link features in two-object complexes, including the hippocampus and the caudate. We collect such link features and within-object s-rep features, pooling the two groups. These pooled features form two data blocks, as discussed in Sect. 3. Then we apply NEUJIVE to the two blocks. Section 6.3 shows (1) the generalizability of a linear classifier learned on our proposed features and (2) the statistical significance of the group difference found in our feature space. To demonstrate the benefit from local affine frames, Sect. 6.4 describes the features from local affine frames and shows the classification results using these features. Finally, Sect. 6.5 lays out an extensive comparison of classification performance using various features of multi-object complexes.

6.1 Data

We test our proposed methods using MR images from the Infant Brain Imaging Study (IBIS) network (St John et al., 2016). The data we are using involve 176 6-months-old infants. Among these infants, there are 34 of the children who were diagnosed as autistic later and 142 of these who were shown not to have developed autism.

Among various subcortical structures, the left hippocampus and the left caudate nucleus were segmented. Then these segmented surfaces were fitted with triangular meshes using SPHARM-PDM (Styner et al., 2006). As described in Liu et al. (2021), we fit an s-rep to every triangular mesh. In the resulting s-rep, we have 72 smooth skeletal points and 24 skeletal edge points. A smooth skeletal point is associated with two spokes pointing toward two sides of an object, while an edge point is associated with a fold spoke pointing away along the tangent direction of the skeleton.

6.2 Visualization of Between-Object Shape Features from Modified Linking Structures

Between-object shape features from modified linking structures are link features defined on objects' skeletons. Each skeletal point is associated with a spoke length, a spoke direction, a link length and a link direction. To discount the object widths in the links, we show the extensions of spokes (i.e., the values of t's in eq. (3)), subtracting the spoke length from the link length at a skeletal point. Though the extensions are distances from the boundary to the external linking surface, we can map such scalar features onto the corresponding skeletal points, as shown in Fig. 6.

In Fig. 6 the top row shows the extensions as heat maps on skeletons of the hippocampus (bottom) and of the caudate (top) in 5 autistic cases. These heat maps share similar patterns: (1) because the caudate bends away from the hippocampus, the links in the head and the tail have smaller extensions than the links in the body, and (2) because the caudate and the hippocampus are not perfectly parallel to each other – one side of the caudate is farther from the hippocampus than the other side, the extensions gradually vary from the near side to the far side of the caudate skeletons. The bottom row shows the link directions in the same 5 autistic cases. While these directions can be slightly different from spokes' directions, the results suggest that the link vectors smoothly swing along the skeletons.

6.3 Classify Autism Versus Non-autism with Two-Object Complexes

In this section, we show the results from classifying ASD vs. non-ASD using the shape of the hippocampus and of the caudate. Specifically, we intend to investigate the

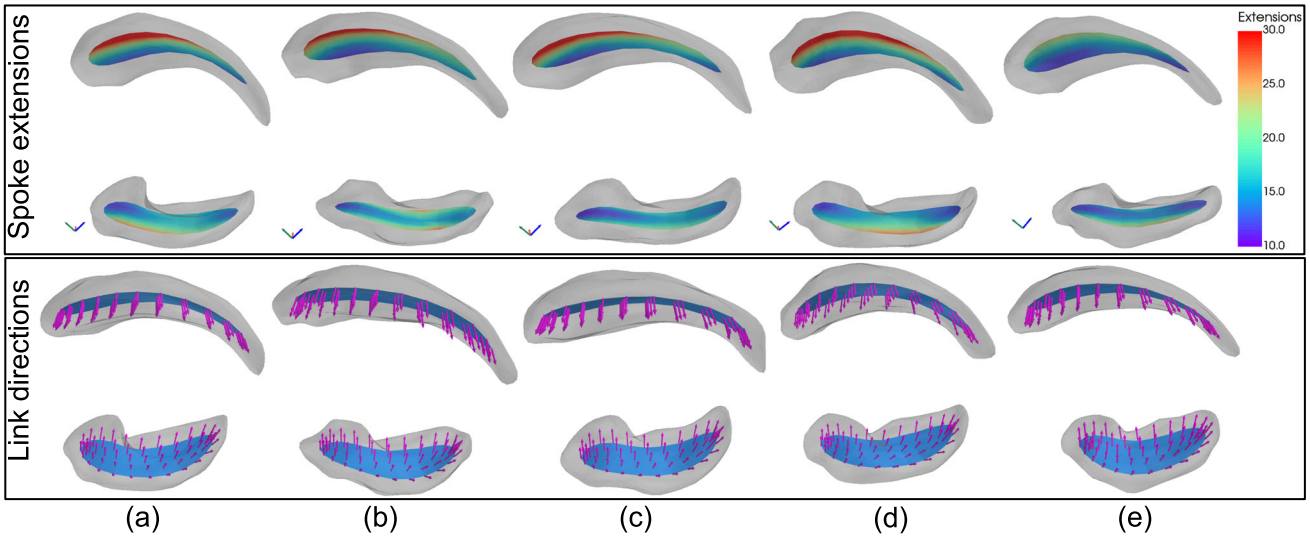

Fig. 6 Linking features of 5 caudate-hippocampus pairs: **a**–**e**. The transparent surfaces are the caudate (above) and the hippocampus (below). Top row: spoke extensions (in mm) from a boundary to the external linking surface. The heat maps on the skeletons show the values of spoke extensions. Bottom row: link directions of link vectors at discrete skeletal points, shown as the arrows in magenta. The skeletons are shown as blue surfaces (Color figure online)

generalizability of a linear classifier trained in our proposed feature space. To this end, we train the classifier with varying training data size and evaluate the corresponding test performance. Moreover, we investigate the statistical significance of the group difference using our proposed features. We perform hypothesis testing with a large number (1000) of permutations to avoid bias of sampling.

We start our methods from SPHARM-PDM models of objects. Each sample in the dataset contains a triangular mesh of a left hippocampus and that of a left caudate. We fit s-reps to these surface meshes, extracting the within-object shape features $\mathbf{x}_i \in \mathbb{R}^{d_1 \times 1}$ for the i^{th} sample (see Sect. 5.2). Then we develop the modified linking structure of each sample, extracting the between-object shape features $\mathbf{y}_i \in \mathbb{R}^{d_2 \times 1}$ (see Sects. 4.2 and 5.2). In developing the modified linking structure, we interpolate the s-reps using the method in Liu et al. (2021) to have a moderate number (about 51) of regular links in every sample. This operation is similar to smoothing the external linking surface because an even finer interpolation tends to "overfit" the shape of individual objects. Finally, we apply NEUJIVE to extract the correlated shape features of $X = \{\mathbf{x}_i\}$ and $Y = \{\mathbf{y}_i\}$ (see Sect. 3), where $i = 1, \cdots, 176$. We set the initial ranks for both X and Y in NEUJIVE as 10, resulting in the joint shape feature matrix. Following the post-feature-selection idea (Liu et al., 2022), we perform classification and hypothesis testing using that feature matrix as follows.

Classification. We train and test a linear classifier called Distance Weighted Discrimination (DWD) (Marron et al., 2007) within our proposed feature space. We randomly select k samples from 34 autistic samples and k samples from 142 non-autistic samples to form the training dataset. The remaining are used as the test data. To avoid the bias introduced by the partitioning of training and test data, we perform 1000 random partitions and classifications for each k. From these 1000 experiments, we obtain an average test Area Under Curve[2] (AUC).

To evaluate the generalizability of the classifier. We vary k from $k = 6$ (20% of the total autism samples) to $k = 27$ (80% of the total autism samples). The resulting average test AUCs of the various k's are shown in the rightmost panel of Fig. 7. In this figure the outperforming (top) curve results from the CoWBO features. The other two features are obtained similarly to the above process but using different blocks in NEUJIVE. Specifically, for NEUJIVE-Sreps we use the correlated features of (1) the s-reps features of the hippocampus and (2) the s-reps features of the caudate. For NEUJIVE-PDMs we use the correlated features of (1) the SPHARM-PDMs of the hippocampus and (2) the SPHARM-PDMs of the caudate.

As we vary the training size, we note (1) that our features can learn useful patterns of the ASD versus the non-ASD even in a small data size, e.g., when the training sample size is 20% and (2) that the increase of the test performance is monotonic with the increase of training data size, which is superior to the classification using NEUJIVE-Sreps and NEUJIVE-PDMs features that are solely from within-object shape features.

[2] Here, the *Curve* refers to the Receiver Operating Characteristic curve.

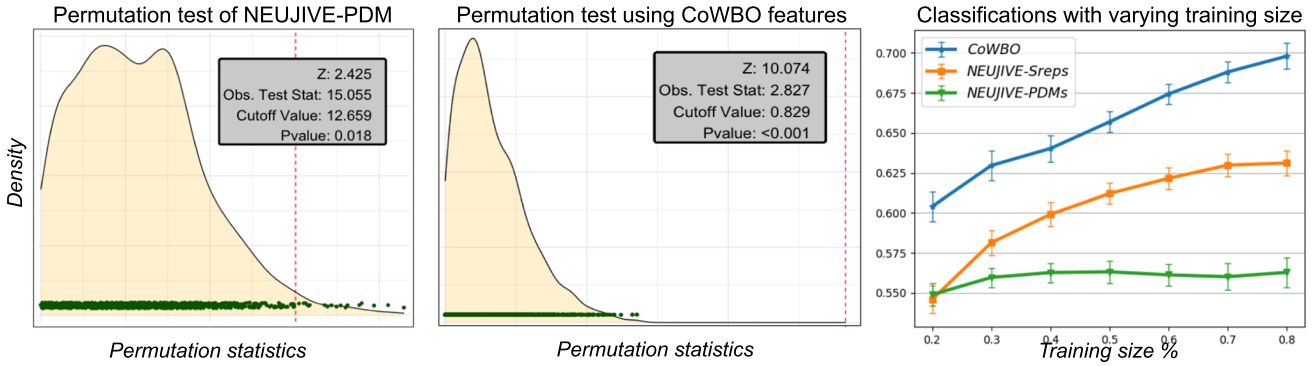

Fig. 7 Left: Hypothesis testing using the correlated shape features of SPHARM-PDMs of the hippocampus and caudate. Middle: Hypothesis testing using the correlated shape features of the hippocampal s-rep features and the link features. Right: Average test AUCs over varying training data size. The error bar at each point shows the standard error of the mean (Color figure online)

Hypothesis testing. To verify if the above group difference is statistically significant, we test the (null) hypothesis that the two groups have the same mean value in the proposed feature space. We use a method called Direction-Projection-Permutation (DiProPerm) (Wei et al., 2016) to perform the hypothesis test for a large number of permutations.

We let DiProPerm generate 1000 permutations of the class labels. The resulting test statistics (mean differences) are shown in the left two panels of Fig. 7 as the green dots. Specifically, the middle panel results from the hypothesis test using the CoWBO features. The observed statistic is shown as the vertical dashed line, which has p-value < 0.001 and z-score 10.074. The p-value and z-score together suggest that the group difference is statistically significant in the feature space formed by our geometric and statistical methods.

As a comparison, the left panel of Fig. 7 results from another feature space that contains the correlated SPHARM-PDM features of the hippocampus and the caudate. Both the p-value (0.018) and z-score (2.425) suggest the group difference there is less statistically significant.

6.4 Local Affine Frames Capture Differentiating Local Geometry

This section demonstrates the benefit from local affine frames in the classification problem. Because ASD affects multiple local regions of a brain structure, geometric features across various local regions should have different importance in classifying ASD and non-ASD. Typically, a classifier assigns different weights to local features to best classify the two groups in a dataset. In our research, the projection to local affine frames (see eq. (4)) is analogous to weighting local geometric features for a better classification because the fitted local frames can capture important local geometry between groups.

Table 1 Test AUCs from using affine frames

Object	$\|\mathbf{n}\|$	$\|\mathbf{n}\| \oplus$ skeletal point
Hippocampus	0.589 (0.004)	0.634 (0.003)
Caudate	0.518 (0.003)	0.527 (0.003)

The numbers outside the parentheses are mean AUCs, while the numbers inside the parentheses are standard errors of the mean.

Specifically, we fit every local affine frame in an object by deforming a local orthonormal frame in the template ellipsoid.[3] To demonstrate the local geometry captured by the frames in the object, we focus on the deformation captured along $\widehat{\nabla\tau_1}$ and $\widehat{\nabla\theta}$ directions on the skeleton (see Sect. 2.3). As a result of the deformation, the unit vectors $\widehat{\nabla\tau_1}$ and $\widehat{\nabla\theta}$ in the template ellipsoid become general vectors $\nabla\tau_1$ and $\nabla\theta$ in the object. The vectors $\nabla\tau_1$ and $\nabla\theta$ together capture local deformation of the skeleton. We compute the cross product $\mathbf{n} = \nabla\tau_1 \times \nabla\theta$ at every skeletal point. The length of $\mathbf{n}$ gives the area of the parallelogram defined by $\nabla\tau_1$ and $\nabla\theta$.

We have extracted the feature magnitude $\|\mathbf{n}\|$ at each skeletal point, and we have used them both alone and in combination with skeletal points' coordinates to classify ASD and non-ASD. These coordinates are centered by subtracting the coordinates of the skeletal center. The results are shown in Table 1.

Table 1 reports test AUCs over 1000 random holdouts, each of which randomly takes 80% of data as training data and the remaining as test data. In the table header, $\|\mathbf{n}\|$ denotes using the lengths as classifying features. The symbol "$\oplus$" denotes the concatenation of features. The last column shows test AUCs from the classification using the concatenation of

[3] We choose the Procrustes Mean ellipsoid as a template shape for each object. We use the General Procrustes Analysis method to obtain this Procrustes Mean of the ellipsoids resulting from Mean Curvature Flow on each object (Dryden & Mardia, 2016).

the lengths of **n**'s and the coordinates of the skeletal points as features.

These results suggest that the cross-products of basis vectors $\nabla\tau_1$ and $\nabla\theta$ already have some discriminatory power (average AUC > 0.5). Moreover, this discriminatory power is slightly boosted by adding the skeletal points at which those local affine frames are centered.

6.5 Comparing with Various Feature Spaces

In order to investigate the difference between using within-object and between-object shape features in classification, we conducted extensive experiments using various sets of features, as listed in Table 2.

In the column of Global Coordinate System (Global C.S.) in Table 2, the points and vectors are expressed with global coordinates after Procrustes alignment, whereas the Local Affine Frames (L.A.F.) columns use the invariant features described in Sect. 5.

The rows in the table compare test AUCs over 1000 random holdouts involving various sets of features. Figure 8 illustrates some of these features. In particular, we categorize these features into 5 groups shown in the leftmost column. The first two groups ("Hippocampus" and "Caudate", respectively) use various choices of within-object shape features of the hippocampus and the caudate, respectively. Each choice is indexed with a circled number. Here, the PDMs of skeletons consist of 3 points sampled at every local affine frame position; they are the origin and two other points that are ϵ-distance away from the origin along the $\nabla\theta$ and $\nabla\tau_1$ axes.

The third group uses shape features from two-object complexes. Specifically, the row "① & ④ PDMs" concatenates coordinates of SPHARM-PDMs of the hippocampus and the caudate; the row "② & ⑤ Spokes" concatenates spokes' features of the two objects; the row "③ & ⑥ PDMs of skeletons" concatenates the skeletal PDMs.

The fourth group uses the resulting two-object joint components from NEUJIVE. We tuned the initial ranks of NEUJIVE (see Sect. 3) for each row in this group, as described in Liu et al. (2022). The PDMs' joint features ("① &④ PDMs' joint features") refer to the NEUJIVE joint components taking the SPHARM-PDMs of the hippocampus and the caudate as two input blocks.

Our major proposal is shown in the last group. The row "⑦ Between-object linking features" contains the hippocampal skeletal points and the link features sampled at these skeletal points. The last row results from the CoWBO features, as discussed in Sect. 3. From this table, we have some key observations as follows.

1. **The CoWBO features give notably the best classification performance** Comparing to the result from the spokes features of the hippocampus (0.624) and the result from the between-object link features (0.68), the average test AUC from the CoWBO features (0.698) increases by about 12% and 3%, respectively. Moreover, because of the removal of the individual and residual components (see Sect. 3), the AUCs of using CoWBO features have smaller Standard Errors of the Mean (SEM) compared to those in the top 3 groups.
2. The hippocampal shape features give notably stronger discriminatory power than caudate features, as seen by comparing the corresponding rows in the "Hippocampus" group with the "Caudate" group. Due to this observation, we neglected within-object shape features of the caudate in forming CoWBO features.
3. Using the hippocampal skeletal features is slightly better than using the hippocampal boundary features. However, this is not the case for classification using the caudate shape.
4. Using concatenated shape features from two objects is not very much different from using the hippocampal shape features alone (comparing the corresponding rows in the "Hippocampus" group with the "Two objects" group). This is because the increased classification power that might come from adding the weakly classifying the caudate features can not compensate the loss due to increasing feature dimensions.
5. The between-object shape features (i.e., ⑦) contain critical discriminatory information. Particularly, we used the modified linking structures of the hippocampus that capture the geometric relations with its neighboring caudate, yielding notably better classification performance than the concatenated shape features in the "Two objects" group.
6. The global coordinates give inferior classification results compared to coordinates with respect to local affine frames, at least when within-object shape features are included. Because the template ellipsoids are deformed to each target object separately, the between-object relations are ignored in fitting local affine frames. The lack of those relations can explain the very limited advantages of using local affine frames (L.A.F. column) over using global coordinates (Global C.S. column) in the rows labeled "Two objects".

7 Conclusions and Discussion

In this paper, we proposed a combination of geometric and statistical models to analyze multi-object shape. The proposed methods have been implemented in https://github.com/ZhiyLiu/shanapy.

Table 2 Test AUCs with features from two-object complexes

	Features	Global C.S.*		L.A.F.*	
		Mean	SEM*	Mean	SEM*
Hippocampus	① SPHARM-PDM	0.591	0.097	-	-
	② Spokes	0.596	0.103	0.624	0.01
	③ PDMs of skeletons	0.598	0.102	0.627	0.096
Caudate	④ SPHARM-PDM	0.54	0.088	-	-
	⑤ Spokes	0.514	0.088	0.552	0.01
	⑥ PDMs of skeletons	0.502	0.089	0.521	0.102
Two objects	① & ④ PDMs	0.588	0.089	-	-
	② & ⑤ Spokes	0.594	0.094	0.591	0.008
	③ & ⑥ PDMs of skeletons	0.593	0.095	0.586	0.01
NEUJIVE	① & ④ PDMs' joint features	0.563	0.009	-	-
	② & ⑤ Spokes joint features	0.552	0.01	0.631	0.007
Ours	⑦ Between-object linking features	0.679	0.009	0.68	0.009
	② & ⑦ CoWBO features	0.668	0.009	**0.698**	0.008

C.S. stands for Coordinate System
L.A.F. stands for Local Affine Frames
SEM stands for Standard Error of the Mean

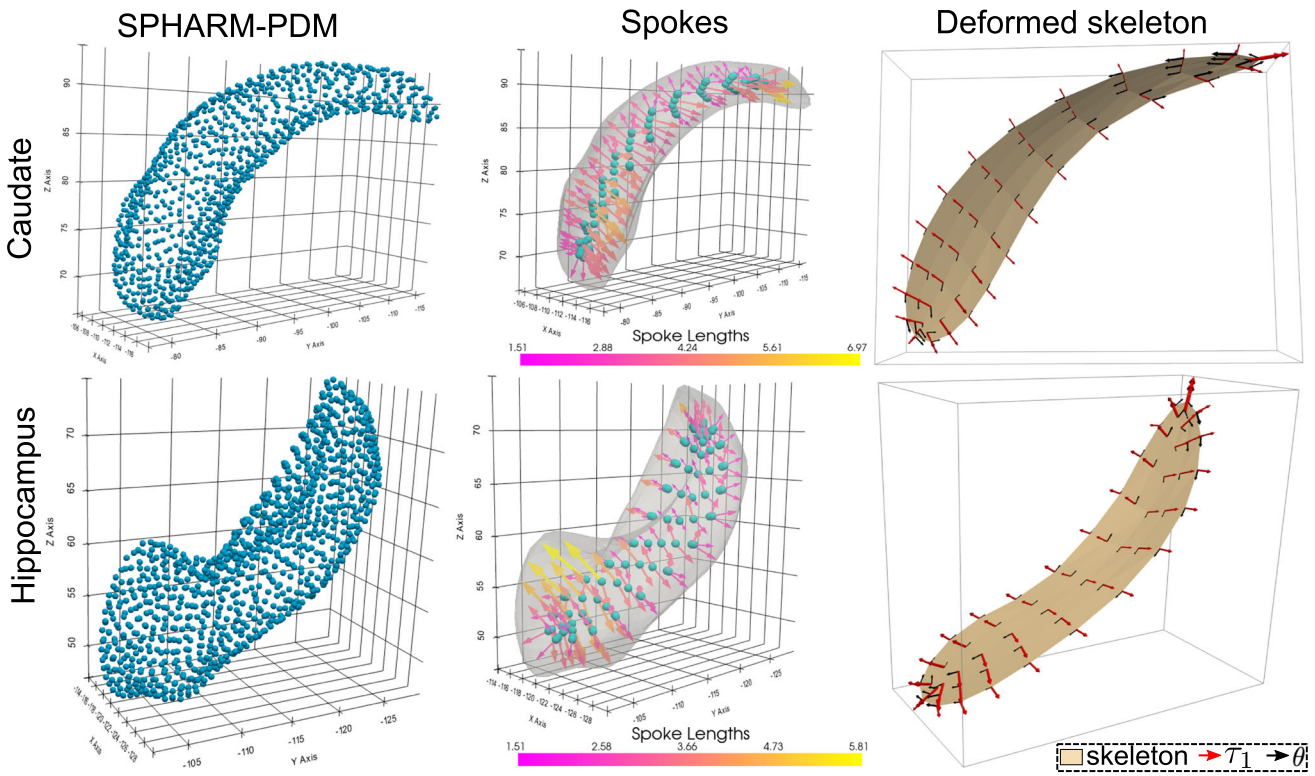

Fig. 8 Various shape features of a two-object complex consisting of a caudate (top row) and a hippocampus (bottom row). The left column shows boundary landmarks from SPHARM-PDM (Styner et al., 2006). The middle column shows spoke features from (Liu et al., 2021)—including coordinates of skeletal points, unit directions of spokes and lengths of spokes. The right column shows deformation field from an ellipsoidal skeleton to the skeleton of a target object (the caudate or the hippocampus). The arrows τ_1 and θ indicate the deformation in two directions (Color figure online)

7.1 Geometric Models

The geometric models including modified linking structures and local affine frames are beneficial in classifying ASD and non-ASD. Especially, the classification using link features between the hippocampus and caudate outperforms the classification using single-object shape features by a large margin. Our modified linking structures capture between-object relations in multi-object complexes. The improved classification performance suggests that the geometric relations between the hippocampus and caudate are important in the development of ASD. Additionally, the good correspondence across our linking structures allows for effective analysis of multi-object complexes.

Moreover, in the classification using single-object shape features, we found that the fitted local affine frames notably boost the performance. Such a boost is due to two reasons: first, the angles and lengths of basis vectors characterize local geometry; second, the projection of both within-object shape features and link features into local affine frames yields shape features that are irrespective of poses and positions of objects.

7.2 Statistical Analysis

The joint analysis of within-object s-rep features and between-object link features further improved the classification performance. Considering the different variability of within- and between-object shape features, we use NEUJIVE to obtain their correlated shape features. To produce robust and generalizable correlations, NEUJIVE makes use of joint variation of the geometry of the complex data spaces in both within- and between-object features. The resulting shape features showed superior statistical power in hypothesis testing and classification.

A limitation of NEUJIVE is that the test data and training data have to be pooled to construct the joint variation subspace. This requirement slows the inference process. However, this problem is negligible in the field of medical image analysis where data sizes tend to be small.

7.3 Future Work

There exist many possible directions for future work. First, we intend to apply our methods to analyze multi-object shape variation in other diseases (e.g., Alzheimer's Disease).

Second, while we tested our methods in two-object complexes in which the two objects are almost parallel to each other, we intend to apply our methods in analyzing other types of between-object relations.

Third, though this paper focuses on two-object complexes, the methods can be applied to multi-object complexes, each of which contains more than two objects. In multi-object complexes, it is desired to capture within-object shape of an anchor object that is of importance in the domain (e.g., the hippocampus in classifying ASD). Also, the relations between the anchor object and its neighboring objects can be captured by the modified linking structure. The CoWBO features of the anchor object provide a comprehensive understanding about the anchor object. We leave this research to our future work.

Fourth, the modified linking structure can also be generalized to analyzing geometric relations between different parts of an object. Take a U-shaped object as an example, the object can be separated into two bar-shaped parts. The within-object shape features can only extract the geometric features within each part. However, the links between the two parts of the object can reveal additional information than the analysis of within-object shape. For example, it can be useful to obtain the variation of relative positions between the two parts.

Fifth, in addition to the above statistical analysis, our modified linking structures can also be used in other applications. In robotics, for example, the paths of robots should be deliberately planned so that the movement of robots will not collide with critical objects in the environment (see e.g., He et al., 2020). To this end, it is promising to model the objects with s-reps and the between-object space with our linking structures. By doing so, the path planning for robots can be efficient and robust.

Acknowledgements This research is funded by NIH grants R01HD 055741, R01HD059854 and R01HD088125. The author J. S. Marron was partially supported by NSF Grant DMS-2113404. The data for the applications was kindly provided by the IBIS network. We thank G. Gerig (NYU), SunHyung Kim (UNC), D. Louis Collins (McGill University), Vladimir Fonov (McGill University) and Heather Hazlett (UNC) for their help in processing the data. Also, we thank the reviewers for their comments. The authors especially appreciate the collaboration with the following colleagues. We thank Martin Styner (UNC) not only for his help in this project but also for his wise advice during the research. We also thank Mohsen Taheri and Jörn Schulz for their insightful comments. The first author also benefits from collaboration with Na Gao and Yuesheng Deng in Harbin Institute of Technology (Shenzhen), China. Last but not least, we thank Akash Krishna, Ankur Sharma, Nicholas Tapp-Hughes and Ran Li for their great support and suggestions.

References

Akhoundi-Asl, A. & Soltanian-Zadeh, H. (2007). Nonparametric entropy-based coupled multi-shape medical image segmentation. In *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (pp. 1200–1203). IEEE.

Bloch, I. (2005). Fuzzy spatial relationships for image processing and interpretation: a review. *Image and Vision Computing, 23*(2), 89–110.

Bossa, M. N. & Olmos, S. (2007). Multi-object statistical pose+ shape models. In *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (pp. 1204–1207). IEEE.

Damon, J. (2003). Smoothness and geometry of boundaries associated to skeletal structures I: Sufficient conditions for smoothness. *Annales de l'institut Fourier, 53*(6), 1941–1985.
Damon, J. (2004). Smoothness and geometry of boundaries associated to skeletal structures, II: Geometry in the blum case. *Compositio Mathematica, 140*, 1657–1674.
Damon, J. (2019). Extending smooth and discrete medial/ skeletal structures to linking structures. *Preliminary preprint*.
Damon, J., & Gasparovic, E. (2017). Modeling multi-object configurations via medial/skeletal linking structures. *International Journal of Computer Vision, 124*(3), 255–272.
Dryden, I. L., & Mardia, K. V. (2016). *Statistical shape analysis: With applications in r*. Wiley.
Feng, Q., et al. (2018). Angle-based joint and individual variation explained. *Journal of Multivariate Analysis, 166*, 241–265.
Gollmer, S. T. et al. (2012). Multi-object active shape model construction for abdomen segmentation: preliminary results. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 3990–3993). IEEE.
Gori, P., Colliot, O., Marrakchi-Kacem, L., Worbe, Y., Poupon, C., Hartmann, A., Ayache, N., & Durrleman, S. (2017). A Bayesian framework for joint morphometry of surface and curve meshes in multi-object complexes. *Medical Image Analysis, 35*, 458–474. https://doi.org/10.1016/j.media.2016.08.011
Guo, Y., et al. (2013). Rotational projection statistics for 3d local surface description and object recognition. *International journal of computer vision, 105*(1), 63–86.
Hardoon, D. R., et al. (2004). Canonical correlation analysis: An overview with application to learning methods. *Neural Computation, 16*(12), 2639–2664.
Hazlett, H. C., Gu, H., Munsell, B. C., Kim, S. H., Styner, M., Wolff, J. J., Elison, J. T., Swanson, M. R., Zhu, H., Botteron, K. N., et al. (2017). Early brain development in infants at high risk for autism spectrum disorder. *Nature, 542*(7641), 348–351.
He, L. et al. (2020). Efficient multi-agent motion planning in continuous workspaces using medial-axis-based swap graphs. arXiv preprint arXiv:2002.11892.
Hong, J. (2019). *Classification of Neuroanatomical Structures Based on Non-Euclidean Geometric Object Properties*. PhD thesis. Computer Science dissertation, Univ. of North Carolina at Chapel Hill.
Hong, J. et al. (2016). Non-Euclidean classification of medically imaged objects via s-reps. *Medical Image Analysis*, 31.
Jung, S. et al. (2012). Analysis of principal nested spheres. *Biometrika*.
Krishna, A. (2021). Incorporating the geometric relationship of adjacent objects in multi-object shape analysis. https://doi.org/10.17615/3dhj-t044
Lei, H., et al. (2017). Fast descriptors and correspondence propagation for robust global point cloud registration. *IEEE Transactions on Image Processing, 26*(8), 3614–3623.
Liu, Z. (2022). *Geometric and Statistical Models for Multi-Object Shape Analysis*. PhD thesis, The University of North Carolina at Chapel Hill.
Liu, Z. et al. (2021). Fitting unbranching skeletal structures to objects. *Medical Image Analysis*.
Liu, Z., et al. (2022). Analysis of joint shape variation from multi-object complexes. *Journal of Mathematical Imaging and Vision*. https://doi.org/10.1007/s10851-022-01136-5
Lock, E. F., et al. (2013). Joint and individual variation explained (jive) for integrated analysis of multiple data types. *The Annals of Applied Statistics, 7*(1), 523.
Malassiotis, S., & Strintzis, M. G. (2007). Snapshots: A novel local surface descriptor and matching algorithm for robust 3d surface alignment. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 29*(7), 1285–1290.
Marron, J. S., et al. (2007). Distance weighted discrimination. *Journal of the American Statistical Association, 102*(480), 1267–1271.
Miolane, N., Caorsi, M., Lupo, U., Guerard, M., Guigui, N., Mathe, J., Cabanes, Y., Reise, W., Davies, T., & Leitão, A., et al. (2021). Iclr 2021 challenge for computational geometry & topology: Design and results. arXiv preprint arXiv:2108.09810.
Mota, T. & Sridharan, M. (2018). Incrementally grounding expressions for spatial relations between objects. In *IJCAI* (pp. 1928–1934).
Petrelli, A. & Di Stefano, L. (2011). On the repeatability of the local reference frame for partial shape matching. In *2011 International Conference on Computer Vision* (pp. 2244–2251). IEEE.
Pizer, S. et al. (2020). Object shape representation via skeletal models (s-reps) and statistical analysis. *Riemannian Geometric Statistics in Medical Image Analysis* (pp. 233–271).
Pizer, S. et al. (2022). Skeletons, object shape, statistics. *Frontiers in Computer Science*.
Qi, L., et al. (2019). Exploiting spatial relation for fine-grained image classification. *Pattern Recognition, 91*, 47–55.
Saha, P. K., et al. (2016). A survey on skeletonization algorithms and their applications. *Pattern Recognition Letters, 76*, 3–12.
Schulz, J., et al. (2016). Non-linear hypothesis testing of geometric object properties of shapes applied to hippocampi. *Journal of Mathematical Imaging and Vision, 54*(1), 15–34.
Schwarz, T., Heimann, T., Lossnitzer, D., Mohrhardt, C., Steen, H., Rietdorf, U., Wolf, I., & Meinzer, H.-P. (2010). Multiobject segmentation using coupled shape space models. In *Medical Imaging 2010: Image Processing* (Vol. 7623, pp. 76233V). International Society for Optics and Photonics.
Serra, J. (1986). Introduction to mathematical morphology. *Computer Vision, Graphics, and Image Processing, 35*(3), 283–305.
Siddiqi, K. & Pizer, S. (2008). Medial representations: Mathematics, algorithms and applications.
St John, T., Estes, A. M., Dager, S. R., Kostopoulos, P., Wolff, J. J., Pandey, J., Elison, J. T., Paterson, S. J., Schultz, R. T., Botteron, K., et al. (2016). Emerging executive functioning and motor development in infants at high and low risk for autism spectrum disorder. *Frontiers in Psychology, 7*, 1016.
Styner, M., et al. (2006). Statistical shape analysis of brain structures using SPHARM-PDM. *The Insight Journal, 1071*, 242–250.
Taheri, M. & Schulz, J. (2022). Statistical analysis of locally parameterized shapes. *Journal of Computational and Graphical Statistics*.
Tu, L. et al. (2018). Skeletal shape correspondence through entropy. *IEEE Transactions on Medical Imaging*.
Vicory, J. (2016). *Shape deformation statistics and regional texture-based appearance models for segmentation*. PhD thesis, The University of North Carolina at Chapel Hill.
Vicory, J., Pascal, L., Hernandez, P., Fishbaugh, J., Prieto, J., Mostapha, M., Huang, C., Shah, H., Hong, J., & Liu, Z., et al. (2018). SlicerSALT: Shape analysis toolbox. In *International Workshop on Shape in Medical Imaging* (pp. 65–72). Springer.
Wei, S., et al. (2016). Direction-projection-permutation for high-dimensional hypothesis tests. *Journal of Computational and Graphical Statistics, 25*(2), 549–569.
Yang, J., et al. (2016). A fast and robust local descriptor for 3d point cloud registration. *Information Sciences, 346*, 163–179.