

# Compositional Hierarchical Representation of Shape Manifolds for Classification of Non-manifold Shapes

Mete Ozay

Graduate School of Information Sciences, Tohoku University, Sendai, Miyagi, 980-8579, Japan.

ozay.mete.b4@tohoku.ac.jp

Umit Rusen Aktas, Jeremy L. Wyatt and Aleš Leonardis

School of Computer Science, The University of Birmingham, Birmingham, B15 2TT, United Kingdom.

{u.aktas, j.l.wyatt, a.Leonardis}@cs.bham.ac.uk

## Abstract

We address the problem of statistical learning of shape models which are invariant to translation, rotation and scale in compositional hierarchies when data spaces of measurements and shape spaces are not topological manifolds. In practice, this problem is observed while modeling shapes having multiple disconnected components, e.g. partially occluded shapes in cluttered scenes. We resolve the aforementioned problem by first reformulating the relationship between data and shape spaces considering the interaction between Receptive Fields (RFs) and Shape Manifolds (SMs) in a compositional hierarchical shape vocabulary. Then, we suggest a method to model the topological structure of the SMs for statistical learning of the geometric transformations of the shapes that are defined by group actions on the SMs. For this purpose, we design a disjoint union topology using an indexing mechanism for the formation of shape models on SMs in the vocabulary, recursively. We represent the topological relationship between shape components using graphs, which are aggregated to construct a hierarchical graph structure for the shape vocabulary. To this end, we introduce a framework to implement the indexing mechanisms for the employment of the vocabulary for structural shape classification. The proposed approach is used to construct invariant shape representations. Results on benchmark shape classification outperform state-of-the-art methods.

## 1. Introduction

Representation of shapes using manifolds has been studied for a long decade in Computer Vision [6, 14, 15, 19, 22, 26, 27, 29, 31, 36] and Statistical Shape Theory (SST) [6, 11, 16, 17, 30]. In these works, shapes are assumed

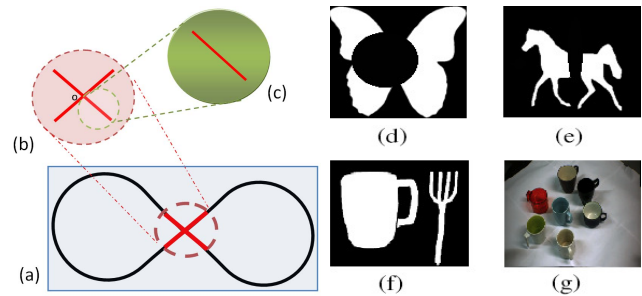


Figure 1: Consider the figure of  $\infty$  in (a), and subparts of it in (b) and (c). Each red line is a set of measurements. Of these neither (a) nor (b) can be modeled on a shape manifold, however the line in (c) can be modeled on a one dimensional manifold. Many such hard to model shapes are observed in the natural world, due to having multiple disconnected components, e.g. due to occlusion in (d, e), or occurrence of multiple objects as in (f, g).

to be represented in shape spaces with manifold structures. However, this assumption fails if the shape spaces are not manifolds because of the *local singularities* that occur in the *shape spaces* [6, 16, 30].

In addition to the aforementioned singularity problem [6, 16, 30], we address a case where the manifold assumption<sup>1</sup> of data<sup>2</sup> and shape spaces fails as depicted in Figure 1. For instance, we consider a scenario where data (measurements) are obtained from a Receptive Field (RF)<sup>3</sup> depicted

<sup>1</sup>In this work, a space is called a topological manifold if it is a i) locally Euclidean, ii) second countable, iii) Hausdorff space.

<sup>2</sup>A set of all possible measurements is called a data space.

<sup>3</sup>We define a Receptive Field (RF) as co-domain of a shape manifold, i.e. a spatial region covering measurements made from an image which is used to represent a *local* component of a shape on a manifold in a shape vocabulary. See Section 2.1.2 for the formal definition.

with the light blue region to model a shape  $\infty$  given in Figure 1.a. However, the shape cannot be modeled on a shape manifold since a measurement at the intersection point  $O$  of two loops has no neighborhood homeomorphic to an open subset in  $\mathbb{R}$ . If we consider another RF with different size and structure depicted by the red region in Figure 1.b, we also observe that the red cross  $\times$  cannot be modeled on a shape manifold. The reason is that still the subspace topology is not locally Euclidean at the intersection point  $O$  due to the occurrence of multiple disconnected components (e.g. edge-like structures) in a neighborhood of  $O$  [33]. If the data and shape spaces each consist of multiple disconnected components due to the structure of the shapes or occlusion, then the spaces may be *disconnected* and non-smooth [21] (see Figure 1.a-b and d-g). In this case, the methods which model shape deformations in connected shape manifolds may fail.

Our motivation for modeling shapes using manifolds is twofold. First, we propose a framework that will be employed for construction of different shape representations with different geometric invariants, such as invariance to Euclidean Similarity and Rigid-body Transformations. Second, we aim to model the topological relationship between shape components, such as connections between parts and compositions deformed under geometric transformations in a shape vocabulary on a *Shape Manifold* (SM). Therefore, we address the aforementioned shape representation problem by providing a compositional hierarchical framework by solving each of its subproblems, recursively. The novelty of the paper can be summarized as follows:

**1. Construction of SMs when data spaces are not manifolds at the first layer  $l = 1$  of a Compositional Hierarchical Shape Vocabulary (CHSV):** We reformulate the relationship between data and shape spaces considering the interaction between RFs and SMs at  $l = 1$  of a CHSV. For this purpose, we first define a RF as a co-domain of an SM that is represented by a shape model in the CHSV. Then, we suggest a method to design structures of RFs considering both the statistical properties of the data, the geometric structure of the SMs, and the topological structure of the models. We analyze the conditions that should be satisfied by the proposed methods to model shapes on SMs in CHSVs.

**2. Construction of invariant SMs and modeling topological properties of models at the higher layers  $l > 1$  of a CHSV:** A disjoint union topology [9] of components of shape representations is designed using an indexing mechanism for the recursive formation of shape models on SMs in a CHSV. More precisely, we introduce a mathematical framework to employ the indexing mechanism for *i*) construction of SMs with various invariance properties for statistical learning of models at  $l > 1$  of CHSVs, *ii*) modeling topological relationships between parts and compositions of

shape models, and *iii*) the employment of the CHSV for structural shape classification.

## 1.1. Related Work and Background

**Statistical learning of shape representations on manifolds:** Manifold embedding methods, such as Kernel methods and Density Representations (DR), represent shapes on the embedded submanifolds [6, 10, 14, 15, 26]. However, the methods suffer from the kernel selection and parameter estimation problems. DR [10, 26] represent a shape using a density function as a point on a manifold whose domain is defined by RFs unlike our proposed model. Therefore, topological relationships between parts and compositions are not modeled by DR, and DR consider a statistical model for shapes with multiple disconnected components.

Statistical learning algorithms, which implement extrinsic methods on manifolds, model shapes by *(i)* first mapping manifolds to tangent spaces, *(ii)* then computing the statistical properties on tangent spaces, and finally *(iii)* mapping the computed statistical models into manifolds [1, 6, 14, 15, 16, 26]. However, the mappings may not exist, and the assumption of the algorithms to construct the models using the mappings may fail. Moreover, the mappings may not be invertible or unique [36]. The algorithms proposed for computing the mappings have several limitations such that the suggested representations and metrics may not be invariant to translations, and they lack of physical interpretations and topological models [15, 19, 26].

**Topological properties and assumptions of models:** The manifold assumption of the aforementioned algorithms regarding the structure of a  $D$ -dimensional shape space  $\Sigma_D^N$  defined by  $N$  landmarks may fail in cases where  $\Sigma_D^N$  is not a manifold. For instance, multiple length minimizing geodesics between the *redundant* shape representations are observed on  $\Sigma_3^N$  [6, 16]. Although a Riemannian manifold can be obtained by removing *redundant* representations and constructing shape spaces using the non-redundant representations [6, 16, 30], it is not complete. Therefore, the methods provide approximate solutions [6, 14, 15].

In the next section, we first analyze the aforementioned approaches and methods used for Multi-component shape representation problem. Then, we propose a new approach to resolve the challenges of their assumptions by re-defining the structures used for shape representation and developing mathematical models of receptive fields, indexing mechanisms and model-based compositional hierarchical shape representations. In Section 3, we suggest a method to employ a structural classification algorithm using the proposed shape representation approach. In Section 4, the proposed methods are examined on benchmark datasets for representation and classification of multi-component shapes by constructing shape models equipped with different invariants. Section 5 concludes the paper.

## 2. Multi-component Shape Representation

Multi-component (MC) shape representation problem consists of three sub-problems, namely *i*) modeling SMs with geometric invariants, *ii*) statistical learning of shape representations on invariant SMs, and *iii*) modeling topological properties of invariant SMs.

### 2.1. Multi-component Shape Representation Model

We assume that a subset of measurements are obtained from an open ball  $\mathcal{B}(\mathbf{x}_c, r) = \{\mathbf{x}' : \|\mathbf{x}_c - \mathbf{x}'\|_2 < r, r \in \mathbb{R}, \mathbf{x}' \in A\}$  which is represented by a Receptive Field (RF) centered at  $\mathbf{x}_c \in \mathbb{R}^D$ , where  $A \in \mathbb{R}^D$  is an open subset. In this work, we use two-dimensional images, therefore  $D = 2$  will be considered in the rest of the paper, if not stated otherwise. Additionally, we consider landmarks as Gabor features extracted from images at each image point  $\mathbf{x} \in \mathbb{R}^2$  using a two-dimensional Gabor kernel  $G(\mathbf{x}, \mathcal{R})$ , where  $\mathcal{R}$  is a set of parameters [12]. We construct SMs with different invariants, such as invariance to scale, translation and rotation, at each layer of a Compositional Hierarchical Shape Vocabulary (CHSV). Therefore, we first provide an overview of the geometric properties of SMs that will be employed for representation of shape models in the following subsection.

#### 2.1.1 Geometry of Shape Representation Models

In order to employ group operations on the measurements, we denote a matrix representation of the measurements as  $X \in \mathbb{R}^{N_j \times D}$ , where  $N_j = |A_j|$ ,  $\forall j \in \mathcal{J}$ , and the  $i^{th}$  row of  $X$  is  $\mathbf{x}_i \in A_j$ . In this work, we address the problem of construction of shape manifolds for representation of shapes which are invariant to scale, translation and rotation. Therefore, we first define a set of translated, rotated and isotropically rescaled observations of  $X$ , called the Euclidean Similarity Transformations (EST) of  $X$ , as

$$[X]_{est} = \{\beta X \Gamma + I_{N_j} \gamma^T : \beta \in \mathbb{R}^+, \Gamma \in SO(D), \gamma \in \mathbb{R}^D\}, \quad (1)$$

where  $\beta \in \mathbb{R}^+$  is a scale variable used for isotropic scaling,  $\Gamma \in SO(D)$  is a rotation matrix, and  $\gamma \in \mathbb{R}^D$  is a translation vector. A space consisting of all possible EST of  $X$  is called EST-Shape space and denoted as  $\mathcal{S}_{est}$ . For  $\beta = 1$ , we obtain the Rigid-body Transformations (RT) of  $X$  as

$$[X]_{rt} = \{X \Gamma + I_{N_j} \gamma^T : \Gamma \in SO(D), \gamma \in \mathbb{R}^D\}. \quad (2)$$

The shape space consisting of all possible  $[X]_{rt}$  is denoted as  $\mathcal{S}_{rt}$ . For an identity matrix  $\Gamma$ , we obtain a scaled and translated (ST) shape of  $X$ , called a pre-shape, as

$$[X]_{st} = \{\beta X + I_{N_j} \gamma^T : \beta \in \mathbb{R}, \gamma \in \mathbb{R}^D\}. \quad (3)$$

Note that the ST space  $\mathcal{S}_{st}$ , which is the space of all possible pre-shapes, is scale and translation invariant [11, 16]. In

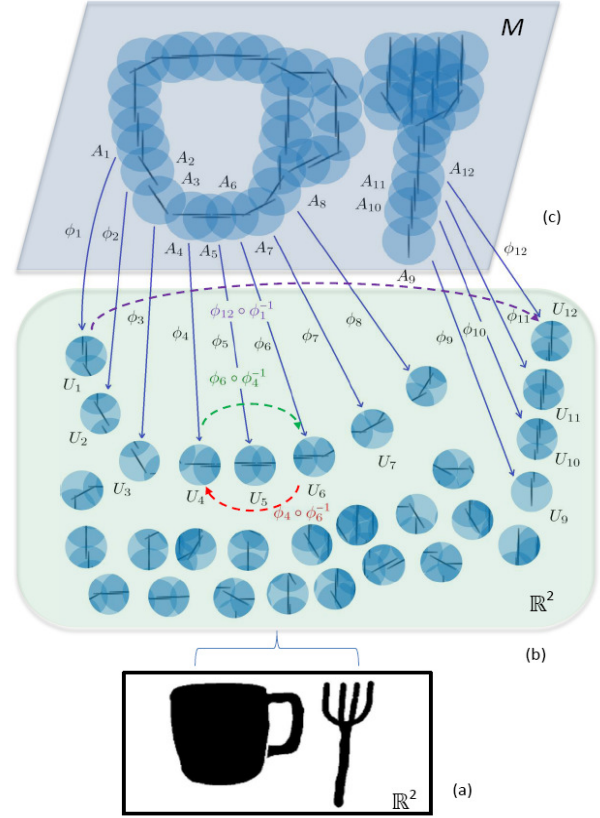


Figure 2: Measurement of shapes of two objects (a) in RFs each of which is defined as co-domain of a mapping from a subset of a manifold  $M$  (c) to a subset of  $\mathbb{R}^2$  (b). See text for details.

order to obtain a rotation invariant representation of pre-shapes, we employ actions from the rotation group as

$$[X]_s = \{X_{st} \Gamma : \Gamma \in SO(D), X_{st} \in \mathcal{S}_{st}\}. \quad (4)$$

The shape space consisting of all possible shapes  $[X]_s$  is denoted by  $\mathcal{S}_s$ . For  $D = 2$ ,  $\mathcal{S}_{st}$  is a  $2k-3$  dimensional sphere  $S^{2k-3}$ , and  $\mathcal{S}_s$  is a quotient space  $S^{2k-3}/SO(2)$  which is a compact manifold. Once we obtain a shape manifold, we employ extrinsic and intrinsic analysis on the manifold for statistical learning of shape representations. The distance functions that will be employed in the learning algorithms are designed according to the geometric and topological structure of the spaces as explained in the following section.

#### 2.1.2 Topological Properties of Shape Manifolds

In order to assure the representation of shapes on SMs, we consider a shape space as a manifold, and the measurements obtained from a RF as a co-domain of the manifold at the first layer of a CHSV.

The co-domain  $U$  determined by a RF is defined together with a manifold

$$\mathcal{P} = \{(A_j, \phi_j) | \phi_j : A_j \subset M \rightarrow U_j \subset \mathbb{R}^D, \forall j \in \mathcal{J}\}, \quad (5)$$

where  $\mathcal{J}$  is a nonempty countable set (see Figure 2). We call the manifold  $\mathcal{P}$  a *part*, and  $(A_j, \phi_j)$  is a chart of the  $\mathcal{P}$ ,  $\forall j \in \mathcal{J}$ . A part  $\mathcal{P}$  can be considered as an element of a shape space, which is also designed to be a manifold. Therefore, we should first show that  $\mathcal{P}$  is a manifold which is a non-degenerate representation model. For this purpose, we will first analyze the conditions which enable us to model the manifold structure of a part  $\mathcal{P}$  when there is an overlap among the charts of the  $\mathcal{P}$ , and among the corresponding RFs in the data spaces.

Given any two charts  $(A_i, \phi_i)$  and  $(A_j, \phi_j)$  on  $\mathcal{P}$ , where  $A_i \cap A_j \neq \emptyset$ , the transition maps are defined as [13]

$$\phi_{ji} : \phi_i(A_i \cap A_j) \rightarrow \phi_j(A_i \cap A_j), \quad (6)$$

$$\phi_{ij} : \phi_j(A_i \cap A_j) \rightarrow \phi_i(A_i \cap A_j), \quad (7)$$

$$\phi_{ji} = \phi_j \circ \phi_i^{-1} \text{ and } \phi_{ij} = \phi_i \circ \phi_j^{-1}, \forall i, j \in \mathcal{J}. \quad (8)$$

Defining  $U_{ij} = \phi_i(A_i \cap A_j)$  and  $U_{ji} = \phi_j(A_i \cap A_j)$ , we compute a map  $\phi_{ji} : U_{ij} \rightarrow U_{ji}$  which is called a *gluing map* in [13]. Then, a set of gluing measurements, which is a superset of RFs used on the images, is defined as [13]

$$\mathfrak{R}\mathfrak{F} = \{\{U_i\}_{i \in \mathcal{J}}, \{U_{ij}\}_{(i,j) \in \mathcal{L}}, \{\phi_{ji}\}_{(i,j) \in \mathcal{L}}\}, \quad (9)$$

where  $\mathcal{L} = \{(i, j) \in \mathcal{J} \times \mathcal{J} : U_{ij} \neq \emptyset\}$  is an *index set*.

In order to assure that there is a manifold with co-domain and maps defined by  $\mathfrak{R}\mathfrak{F}$ , the maps should satisfy the following conditions [13]:

1.  $\phi_{ii} = id_{U_i}$ , where  $id_{U_i}$  is an identity map of  $U_i$ ,  $\forall i \in \mathcal{J}$ .
2.  $\phi_{ij} = \phi_{ji}^{-1}$ ,  $\forall (i, j) \in \mathcal{L}$ .
3. If  $U_{ji} \cap U_{jk} \neq \emptyset$ , then  $\phi_{ij}(U_{ji} \cap U_{jk}) = U_{ij} \cap U_{ik}$ ,  $\forall i, j, k \in \mathcal{J}$ , and  $\phi_{ki}(\mathbf{x}) = \phi_{kj} \circ \phi_{ji}(\mathbf{x})$ ,  $\forall \mathbf{x} \in (U_{ji} \cap U_{ik})$ .
4. There are open balls  $V_{\mathbf{x}}$  and  $V_{\mathbf{x}'}$  centered at  $\mathbf{x}$  and  $\mathbf{x}'$  such that no point of  $V_{\mathbf{x}'} \cap U_{ji}$  is the image of any point of  $V_{\mathbf{x}} \cap U_{ij}$  by  $\phi_{ji}$ ,  $\forall (i, j) \in \mathcal{L}$  with  $i \neq j$ , and  $\forall \mathbf{x} \in \partial(U_{ij}) \cap U_i$ ,  $\forall \mathbf{x}' \in \partial(U_{ji}) \cap U_j$ , where  $\partial(U_{ij})$  is the boundary of  $U_{ij}$ .

By the employment of the conditions (1-4), we define a set of algebraic operations in order to provide an algebraic framework for operations on parts and compositions on a shape manifold  $M$  at the  $l^{th}$  ( $l > 1$ ) layer of compositional hierarchical representation. In the first condition (1), we assure the existence of an idempotent map of composition of maps, and, thereby, the existence of inverse maps.

The relationship between a composition of maps and its inverse induced by an index set is defined in the second condition (2). Together with the third condition (3), we obtain a property of *composition of maps*  $\phi_{ij} = \phi_{ik} \circ \phi_{kj}^{-1}$ . For example, we model  $\phi_{64} = \phi_6 \circ \phi_4^{-1} : U_{46} \rightarrow U_{64}$  and  $\phi_{46} = \phi_4 \circ \phi_6^{-1} : U_{64} \rightarrow U_{46}$ , in Figure 2<sup>4</sup>. Following the composition property, we obtain  $\phi_{64} = \phi_{65} \circ \phi_{54}^{-1}$ . In addition, we model the relationship  $\phi_{12} \circ \phi_1^{-1}$  between  $U_1$  and  $U_{12}$  by employing the composition property for  $i = 1, j = 12$  and for each  $k = 2, 3, \dots, 11$ , in our proposed approach<sup>4</sup>.

The conditions (1-3) show that a shape space  $M$  is second-countable. Considering the fourth condition (4), we assure that  $M$  is Hausdorff, thereby,  $M$  is a topological manifold. Therefore, the conditions (1-4) and the proposed approach provide a framework to associate the measurement spaces and pre-shape spaces defined by RFs to  $D$  dimensional shape manifolds. Next, we provide a theorem by formalizing the observations depicted in Figure 2.

**Theorem 2.1** ([13]). *For every set  $\mathfrak{R}\mathfrak{F}$  defined in (9), there is a  $D$  dimensional manifold whose transition maps are  $\{\phi_{ji}\}_{(i,j) \in \mathcal{L}}$ , where  $\mathcal{L}$  is an index set. More precisely, there exists a manifold  $\mathfrak{M} = (\mathfrak{R}\mathfrak{F}, \{\psi_i\}_{i \in \mathcal{J}})$ , where  $\psi_i : U_i \rightarrow \mathbb{R}^D$  is defined as*

$$\psi_i \triangleq \psi_j \circ \phi_{ji}, \forall (i, j) \in \mathcal{L}. \quad (10)$$

In the following **Corollary**, we show that a part  $\mathcal{P}$  is a manifold in our proposed framework.

**Corollary 1.**  *$\mathcal{P}$  is a manifold whose co-domain and transition maps are defined by  $\mathfrak{R}\mathfrak{F}$  with measurements on two-dimensional images.*

Following the shape space construction approach discussed in the previous section, parts will be constructed using equivalence relations between shape representation models through quotient maps of data and shape spaces. Additionally, we have shown that a part  $\mathcal{P}$  can be constructed as a manifold which is invariant to scale, translation and/or rotation using quotient maps at the first layer  $l = 1$  of CHSV. Next, we show that a composition of parts constructed at  $l > 1$  of the CHSV is a manifold.

**Theorem 2.2.** *Given a collection of parts  $\{\mathcal{P}_n\}_{n \in \mathfrak{N}}$ , where  $\mathfrak{N}$  is a set of indices, a composition  $\mathcal{C}$  of parts defined as*

$$\mathcal{C} = \coprod_{n \in \mathfrak{N}} \mathcal{P}_n \quad (11)$$

*is a manifold.*

**Theorem 2.2** assures that a composition constructed in the higher layers of the compositional hierarchical vocabulary is a manifold. The formal proofs of the **Theorems** and **Corollary** are given in the Supplemental Material.

<sup>4</sup> The colored text given in the example refers to the sample maps induced by the proposed indexing mechanism as depicted in Figure 2.



## 2.2. Statistical Learning of Representations on SMs

In our statistical learning model, we aim to compute the joint observation probability of two parts with different part indices or ids in locally defined open sets on Shape Manifolds (SMs), which are clusters. In other words, we compute a set of disjoint clusters  $\Lambda_k \subset \mathcal{C}$ ,  $\forall k = 1, 2, \dots, K$ , such that  $\bigcup_{k=1}^K \Lambda_k = \mathcal{C}$ , assuming that  $Q$  is a probability distribution on  $\mathcal{C}$ ,  $\{\Omega_n\}_{n \in \mathfrak{N}}$  are random variables associated to  $\{\mathcal{P}_n\}_{n \in \mathfrak{N}}$  and drawn independent and identically distributed (i.i.d.) from  $Q$ . Moreover, we require that if  $\Lambda_k \subseteq \Lambda_l$ , then  $\sigma(\Lambda_k) \leq \sigma(\Lambda_l)$ , where  $\sigma(\cdot)$  is a cost function used to avoid obtaining redundant subsets of clusters. This condition has been defined and used as the Monotone Clustering Property (MCP) for clustering [8]. In our proposed model, we define the cost function of an instance of a cluster as

$$\sigma(\lambda_k, \omega_n, \omega_m) = \Pr(\lambda_k \in \Lambda_k | \omega_n \in \Omega_n, \omega_m \in \Omega_m) \quad (12)$$

which is the probability of observing two given part instances  $\omega_n \in \Omega_n$  and  $\omega_m \in \Omega_m$  in a cluster instance  $\lambda_k \in \Lambda_k$ ,  $\forall k$ . Computation of a collection of clusters which maximize (12) is considered as a minimum entropy clustering problem [8, 23] and defined as

$$\mathcal{K} = \arg \min_{\mathfrak{K}} \sum_{\omega_n \in \Omega_n} \sum_{\omega_m \in \Omega_m} \sum_{\lambda_k \in \Lambda_k} \Pr^\alpha(\lambda_k | \omega_n, \omega_m) - 1, \quad (13)$$

where  $\mathfrak{K}$  is the set of all possible partitions of the dataset among  $K$  clusters. We choose  $0 < \alpha < 1$  unlike the clustering problem proposed in [2] to satisfy the MCP.

In a clustering algorithm [23] which computes (13), a part  $\mathcal{P}_m$  that resides in the neighborhood of a given part  $\mathcal{P}_n$  is selected to minimize the conditional entropy associated with the conditional probability defined in (12). We compute four distance functions [11] for the computation of the neighborhood of  $\mathcal{P}_n$  in each of the shape spaces:

**Distance on scaled shape space  $\mathcal{S}_s$ :** Given two measurements  $X_1, X_2 \in \mathbb{R}^{N_j \times D}$ , the full Procrustes distance  $0 \leq d_f(X_1, X_2) \leq 1$  is computed as

$$\begin{aligned} d_f(X_1, X_2) &= \inf_{\Gamma \in SO(D), \beta \in \mathbb{R}} \|\hat{X}_2 - \beta \hat{X}_1 \Gamma\| \\ &= \left(1 - \left(\sum_{i=1}^D \nu_i\right)\right), \end{aligned} \quad (14)$$

where  $\nu_1 \geq \nu_2 \geq \dots \geq |\nu_D|$  are square roots of the eigenvalues of  $\hat{X}_1^T \hat{X}_2 \hat{X}_2^T \hat{X}_1$ , and  $\hat{X}_1$  and  $\hat{X}_2$  are pre-shapes of  $X_1$  and  $X_2$ , respectively. For  $\beta = 1$ , the partial Procrustes distance  $0 \leq d_{pp}(X_1, X_2) \leq \sqrt{2}$  is computed on  $\mathcal{S}_s$  as

$$\begin{aligned} d_{pp}(X_1, X_2) &= \inf_{\Gamma \in SO(D)} \|\hat{X}_2 - \hat{X}_1 \Gamma\| \\ &= \sqrt{2} \left(1 - \left(\sum_{i=1}^D \nu_i\right)\right)^{\frac{1}{2}}. \end{aligned} \quad (15)$$

**Distance on pre-shape space  $\mathcal{S}_{st}$ :** On the unit sphere, the Procrustes distance between two pre-shapes  $0 \leq d_p(X_1, X_2) \leq \frac{\pi}{2}$  is computed as

$$d_p(X_1, X_2) = \arccos\left(\sum_{i=1}^D \nu_i\right). \quad (16)$$

**Distance on  $\mathcal{S}_{est}$ :** The ordinary Procrustes distance between two shapes is computed as

$$\begin{aligned} d_{op}(X_1, X_2) &= \inf_{\Gamma, \beta, \gamma} \|X_2 - \beta X_1 \Gamma - I_N \gamma^T\| \\ &= \|X_2\|^2 \sin^2(d_p(X_1, X_2)), \end{aligned} \quad (17)$$

where  $\|X_2\| = (Tr(X_2^T X_2))^{\frac{1}{2}}$  is the Euclidean norm,  $Tr(\cdot)$  is the matrix trace, and  $X_1$  and  $X_2$  are centered but may not be unit sized matrices of the measurements. Note that,  $d_{op}(X_1, X_2) = d_f^2(X_1, X_2)$  for unit size measurements. However, the parts and compositions may not be unit size especially when Gabor features with different geometric structures are used as landmarks (see Figure 2).

**Distance on  $\mathcal{S}_{rbt}$ :** The ordinary Procrustes distance between two shapes is computed as

$$\begin{aligned} d_{opp}(X_1, X_2) &= \inf_{\Gamma, \gamma} \|X_2 - X_1 \Gamma - I_N \gamma^T\| \\ &= Tr(X_1^T X_1) + Tr(X_2^T X_2) \\ &\quad - 2\|X_1\| \|X_2\| \cos(d_p(X_1, X_2)). \end{aligned} \quad (18)$$

## 2.3. Compositional Formation of Hierarchy of SMs

In the proposed approach, we construct parts  $\{\mathcal{P}_n\}_{n \in \mathfrak{N}}$ , and a shape manifold  $\mathcal{C}$  at each layer  $l$  of a compositional hierarchical shape vocabulary by employing a disjoint union operation and the indexing method proposed in Section 2.1.2. In addition, we integrate a set of shape manifolds  $\{\mathcal{C}_m\}_{m \in \mathfrak{M}}$  to construct a shape vocabulary  $\Upsilon = \bigsqcup_{m \in \mathfrak{M}} \mathcal{C}_m$ , where  $\mathfrak{M}$  is a nonempty countable set of indices. Therefore, we consider a compositional approach for the formation of the vocabulary  $\Upsilon$ . We represent a shape manifold  $\mathcal{C}_m^l$  constructed at the  $l^{th}$  layer by a graph  $\mathcal{G}_m^l = (\mathcal{V}_m^l, \mathcal{E}_m^l)$ ,  $\forall m \in \mathfrak{M}$ .  $\mathcal{V}_m^l$  is the set of nodes where each node  $v_i^{l+1} = n$  represents an index label or id  $n \in \mathfrak{N}$  of a subpart  $\mathcal{P}_n^l$ ,  $\forall m$ .<sup>5</sup>

At  $l = 1$ , we have disconnected parts, therefore  $e_{ij} = \emptyset$ ,  $\forall e_{ij} \in \mathcal{E}_m^1$ ,  $\forall m$ . For  $l > 1$ , if two parts  $\mathcal{P}_n^l$  and  $\mathcal{P}_{n'}^l$  co-occur in a cluster  $\lambda_k$  with probability  $\sigma(\lambda_k, \omega_n, \omega_{n'}) > 0$ ,  $v_i^l = n$  and  $v_j^l = n'$ , then  $e_{ij} = k$  (Figure 3). Following the recursive definition of the vocabulary, a shape manifold  $\mathcal{C}_m^l$  is considered as a part of another shape manifold  $\mathcal{C}^{l+1}$  which is referred as a *composition* constructed at the  $(l+1)^{st}$  layer (Figure 3.c). At each layer  $l$ , we can approximate geodesic between two parts by comparing their

<sup>5</sup>Note that, we can employ  $\mathfrak{N} = \mathfrak{I}^l$  and  $\mathfrak{M} = \mathfrak{I}^{l+1}$ , at each  $l$  of a CHSV, where  $\mathfrak{I}$  is used to define parts in (5). For the sake of simplicity, we updated the notation in this section.

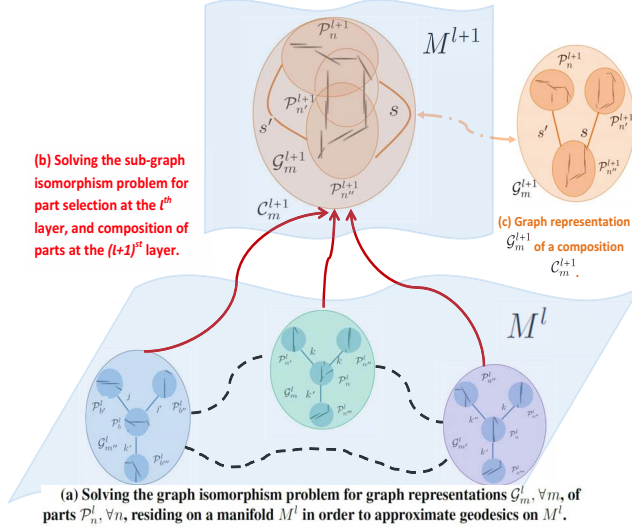


Figure 3: At the  $l^{th}$  layer, (a) the graph isomorphism problem is solved to approximate geodesics [5] between parts and compositions on a manifold  $M^l$ , (b) isomorphism between parts and compositions is achieved between consecutive layers  $l$  and  $l+1$  by solving the sub-graph isomorphism problem. (c) A graph representation  $G_m^{l+1}$  of a composition  $C_m^{l+1}$  constructed at the  $(l+1)^{st}$  layer.

graph representations considering the manifold structure of shape spaces (Figure 3.a). Given two graphs  $G_m^l$  and  $G_{m'}^l$ , deciding whether they are isomorphic is defined as a graph isomorphism problem [28, 34].

At the first layer  $l = 1$  in the vocabulary, we assure the existence of a bijection  $b: \mathcal{V}_n^l \rightarrow \mathcal{V}_{n'}^l$  between two graph representations  $G_n^l$  and  $G_{n'}^l$  of two parts  $P_n^l$  and  $P_{n'}^l$  in a SM  $C_m^{l+1}$ , respectively. In other words,  $G_n^l \sim G_{n'}^l$  iff  $G_{n'}^l \in b(G_n^l)$ ,  $\forall n, n' \in \mathfrak{N}^6$  following the compositional formation of parts as described in the previous section. For  $l > 1$ , we search and select the graphs  $G_n^l$  of the parts  $P_n^l$  that are subgraphs and isomorphic to the graphs  $G_m^{l+1}$  of compositions  $C_m^{l+1}$ ,  $\forall m \in \mathfrak{M}$ . This problem has been defined and studied as a sub-graph isomorphism problem [2, 28, 34]. In this work, we select the subgraphs of compositions  $C_m^{l+1}$ ,  $\forall m \in \mathfrak{M}$ , solving the sub-graph isomorphism problem [2]. We search the SMs and increment the hierarchy until we obtain a compact connected SM which contains a single component at the highest layer of the hierarchy (Figure 3.b).

## 2.4. Inference of Shape Representations on Test Data

We first employ the proposed indexing mechanism to induce a disjoint union topology in an inference tree  $\mathcal{T}$ .

<sup>6</sup>This condition is equivalent to the membership of node ids to  $\mathcal{L}$  which is employed to define RFs in (9), i.e.  $\forall (n, n') \in \mathcal{L}$ .

Thereby, we can represent measurements obtained from test images and parts obtained from the learned vocabulary on a manifold  $M^l$  at each layer  $l$  of the hierarchy (Figure 3).

For this purpose, we first extract a set of landmarks or features from a test image. Then, we represent the test features as parts  $\{\tilde{P}_n^l\}_{n \in \mathfrak{N}}$  and compositions  $\{\tilde{C}_m^l\}_{m \in \mathfrak{M}}$  at the first layer  $l = 1$  of an inference tree  $\mathcal{T}$  such that  $\tilde{C}_m^1 = \coprod_{n \in \mathfrak{N}} \tilde{P}_n^1$ ,  $\forall m \in \mathfrak{M}$ . At each consecutive layer  $l > 1$ ,

we search for a set of graphs  $\{\tilde{G}_u^l : u \in \mathfrak{U}\}$  whose members are graph isomorphic to the shape manifolds  $G_n^l$ ,  $\forall n \in \mathfrak{N}$  in the vocabulary  $\Upsilon$ . More precisely, we first compute a set of all graphs  $iso(G_n^l)$  which are isomorphic to  $G_n^l$ ,  $\forall n$ , and then search the set of parts  $\tilde{G}_u^l \in iso(G_n^l)$ ,  $\forall u \in \mathfrak{U}$ , where  $\mathfrak{U}$  is a nonempty countable set of indices (Figure 3.a).

## 3. Shape Classification using Compositional Hierarchical Shape Vocabularies

In the shape classification problem, we consider *contribution* of each part and composition in the vocabulary to the representation of classes using a top-down approach. For this purpose, we employ the topological structure of the vocabulary  $\Upsilon$  which is induced by the indexing mechanism, such as the recursive representation of parts and compositions at each layer  $l > 1$  of  $\Upsilon$ .

After shape manifolds are formed in a hierarchical vocabulary, we first compute training and test inference trees  $\mathcal{T} = \{\mathcal{T}_n : n = 1, 2, \dots, N\}$  and  $\tilde{\mathcal{T}}$  using a set of  $N$  training images and a given test image, respectively. Let  $F \in \mathbb{R}^{N \times H}$  denote an input matrix whose row vector  $\mathbf{f}_n$  represents the detection of a feature by a part or a composition in  $\Upsilon$  on the  $n^{th}$  image provided by an inference tree  $\mathcal{T}_n$ . We encode the class labels using 1-of- $W$  coding of classes using a matrix  $Y \in \mathbb{R}^{N \times W}$ , where  $y_{nw} = 1$  if the  $n^{th}$  image belongs to the  $w^{th}$  class, and  $y_{nw} = 0$ , otherwise. In addition, we assume a linear class model for each class variable as

$$\mathbf{y}_w = F\mathbf{z}_w + \epsilon_w, \forall w = 1, 2, \dots, W, \quad (19)$$

where  $\mathbf{z}_w \in \mathbb{R}^H$  is a vector of model coefficients, and  $\epsilon_w$  is a vector that represents model error. Then the class model estimation problem is defined as [18]

$$\hat{Z} = \arg \min \left( \sum_{w=1}^W (\mathbf{y}_w - F\mathbf{z}_w)^T (\mathbf{y}_w - F\mathbf{z}_w) + \mathcal{L} \right), \quad (20)$$

where  $\hat{Z}$  is the matrix of estimated model variables, and  $\mathcal{L}$  is a structural loss function. In order to employ the structure induced by the vocabulary in the classification models, we define the loss function  $\mathcal{L}$  as [18]

$$\mathcal{L} \triangleq \xi \sum_{h=1}^H \sum_{v_i \in \Lambda_{int}} \mu_{v_i} \|\mathbf{z}_{\Phi_{v_i}}^h\|_2 + \xi \sum_{h=1}^H \sum_{v_j \in \Lambda_{leaf}} \mu_{v_j} \|\mathbf{z}_{\Phi_{v_j}}^h\|_2, \quad (21)$$

where  $\xi$  is a regularization parameter.  $\Lambda_{leaf}$  is a set of leaf nodes which represent the detections by parts in  $\Upsilon$ .  $\Lambda_{int}$  is a set of internal nodes which represent the detections by either parts or compositions in  $\Upsilon$  according to the topology induced by the indexing mechanism. In other words,  $\Lambda_{int}$  consists of shared parts and compositions among different objects and different layers, and defines the overlapping groups.  $\mu_{v_i}$  and  $\mu_{v_j}$  are the regularization weights associated to each node  $v_i \in \Lambda_{int}$  and  $v_j \in \Lambda_{leaf}$ .

$\Phi_{v_i}$  and  $\Phi_{v_j}$  are group structures whose members are the indices of leaf nodes of the subtrees rooted at the nodes  $v_i \in \Lambda_{int}$  and  $v_j \in \Lambda_{leaf}$ . Using the group structures encoded by the indexing mechanism, feature values detected at the corresponding parts and compositions are encoded in  $\mathbf{z}_{\Phi_{v_i}}^h$  and  $\mathbf{z}_{\Phi_{v_j}}^h$ , which are the vectors of variables  $z_{\varrho_i}^h, \forall \varrho_i \in \Phi_{v_i}, \forall v_i \in \Lambda_{int}$  and  $z_{\varrho_j}^h, \forall \varrho_j \in \Phi_{v_j}, \forall v_j \in \Lambda_{int}$  [18]. In the proposed framework, we solve the problem (20) by minimizing (21) for an ensemble of inference trees induced by the learned hierarchical compositional shape vocabulary.

## 4. Experiments

We first examine the robustness and invariants of the shape representations modeled by the proposed algorithms using a benchmark dataset of shapes that are corrupted by noise and geometric transformations for shape classification. Then, we examine classification performance of the proposed statistical learning algorithms on SMs represented by CHSVs that are modeled by the proposed methods using two benchmark datasets. Implementation details of the algorithms are provided in the Supplemental Material. The results are given for three different cases each of which is denoted by **MC-Z**, where  $\mathbf{Z} \in \{1, 2, 3\}$  represents the employment of statistical learning algorithms using distance functions  $d_f$  (14),  $d_{op}$  (17) and  $d_{opp}$  (18) in the shape spaces  $\mathcal{S}_s$ ,  $\mathcal{S}_{est}$  and  $\mathcal{S}_{rbt}$ , respectively.

### 4.1. Experimental Analysis on Artificial Datasets and Comparison with State-of-the-art Hierarchical Deep Representations

In this section, we examine the suggested approach and the algorithms on three different benchmark artificial datasets. The datasets are generated such that the shapes consisting of single and/or multiple components are corrupted by noise and geometric transformations [20].

Convex Shapes (**CS**) dataset [20] consists of 8000 training and 50000 test images each of which contains convex and non-convex regions, and belong to two classes. In this dataset, the classification problem is defined as a binary classification of convex and non-convex regions. Mnist-Rotation (**M-R**) dataset [20] consists of 12000 training and 50000 test images of the Mnist digits which belong to ten classes and are rotated by an angle generated uniformly be-

Table 1: Classification performance (%) of the algorithms on the benchmark shape datasets.

Algorithm	CS	M-R	M-RBI
<b>DBN-3</b> [20]	81.37	87.70	71.49
<b>DAE2</b> [7]	81.56	88.06	55.08
<b>CAE2</b> [7]	80.70	86.38	51.75
<b>mDAE2</b> [7]	81.90	89.64	53.88
<b>MC-1</b> ( $\mathcal{S}_s$ )	83.20	89.29	78.26
<b>MC-2</b> ( $\mathcal{S}_{est}$ )	84.15	90.16	77.15
<b>MC-3</b> ( $\mathcal{S}_{rbt}$ )	81.07	91.71	74.08

tween 0 and  $2\pi$  radians. In the Mnist-RBI (**M-RBI**) dataset [20], a random patch from a gray scale image is used as the background for the rotated images obtained from the Mnist-Rotation dataset. The size of each image is  $28 \times 28$ . The proposed algorithms are compared with Deep Belief Networks with three layers (**DBN-3**), Denoising (**DAE-2**), Contractive (**CAE-2**) and marginalized Denoising auto-encoder (**mDAE-2**) algorithms with two layers.

The results given in Table 1 show that the proposed algorithms outperform the state-of-the-art algorithms. In addition, we observe that the proposed algorithms can employ invariant properties of shape manifolds on shape representations. For instance, EST invariant shape spaces are used in **MC-2** for encoding information about convexity of shapes. Therefore, the best classification performance is observed in **MC-2** on the **CS** dataset. In addition, we observe that Rigid-body Transformation invariant shape manifolds used in **MC-3** provide better representations of rotated shapes obtained from the **M-R** dataset than the other shape spaces.

The proposed methods which model representations in  $\mathcal{S}_{est}$  and  $\mathcal{S}_s$  outperform the algorithms in noisy images which contain corrupted shapes (**M-RBI** dataset). The reason of performance boost is that the patterns observed at different scales in the background images are recognized and removed from the shape representations by the proposed graph compression algorithms. Moreover, **MC-1** outperforms **MC-2** on **M-RBI**. The reason follows the hierarchical employment of invariants in **MC-1** by first removing scale and translation in  $\mathcal{S}_{st}$ , and then removing rotation in  $\mathcal{S}_s$ . Therefore, statistical relationship between parts is modeled on  $\mathcal{S}_s$  by a *spectral analysis of geometric information* obtained from *pre-shapes* for computation of  $d_f$  (14). However, **MC-2** aims to construct a scale, translation and rotation invariant representation in  $\mathcal{S}_{est}$ , and model the statistical relationship by a *geometric analysis of spectral information* obtained from *measurements* for computation of  $d_{op}$  (17). Therefore, we observe different and complementary performances using the proposed methods according to the geometric and statistical properties of the data. This observation is analyzed in the following sections.

Table 2: Classification performance (% Mean  $\pm$  Variance) of algorithms using the Butterfly Dataset.

<b>S-kFP</b> [14]	<b>S-kFPG</b> [15]	<b>MKL</b> [14]	<b>SM</b> [14]	<b>MC-1</b> ( $\mathcal{S}_s$ )	<b>MC-2</b> ( $\mathcal{S}_{est}$ )	<b>MC-3</b> ( $\mathcal{S}_{rbt}$ )
57.75 $\pm$ 2.0	60.37 $\pm$ 1.6	60.84 $\pm$ 2.0	63.98 $\pm$ 1.6	<b>67.83 <math>\pm</math> 2.1</b>	66.17 $\pm$ 2.3	64.09 $\pm$ 1.1

## 4.2. Classification Results on Real-world Datasets

### 4.2.1 Comparison with State-of-the-art Shape Classification Methods Employed on Shape Manifolds

In the first set of the experiments, we examined the classification performance of the proposed algorithms with the state-of-the-art shape classification algorithms that are implemented on shape manifolds using the Butterfly Dataset [35]. The dataset contains 823 images belonging to 10 classes. Following the data setup suggested in [14], we randomly selected 40 shapes from each class for training and used the rest for testing. The procedure was repeated 10 times, and mean and variance of the classification performance values are given in Table 2.

In the real world datasets which contain shapes that are deformed by articulation and non-rigid body transformations, the algorithms need to model higher variance of the geometric patterns distributed among different measurements, objects and classes. Therefore, the algorithms which use shape spaces with *richer* geometric properties perform better than the other algorithms on these datasets. For instance, **S-kFP** and **MKL** employ full Procrustes (FP) on the Kendall’s shape space [14]. Similarly, we have used the FP distance for statistical learning on the scaled  $\mathcal{S}_s$  which corresponds to the Kendall’s shape space in **MC-1**. Meanwhile, Grassmann manifolds are used in **SM**. Since our proposed method considers the contribution of each part and composition to the classification model, we obtain better classification performances using  $\mathcal{S}_s$  in **MC-1**.

### 4.2.2 Comparison with State-of-the-art Shape Descriptors and Vocabularies

In the second set of the experiments, we examined the classification performance of the algorithms using the Animals Dataset [3] consisting of 2000 shapes belonging to 20 categories. We have used this dataset instead of the MPEG-7 dataset since the Animals dataset contains more images that are deformed with more articulated and non-rigid transformations, and larger class-variation than the MPEG-7 dataset [4]. 50 images per class were randomly chosen for training, and the rest of the samples were used for testing as suggested in [4]. The procedure is repeated 10 times for the proposed algorithms and the results are compared with the performances of the shape descriptors.

The results given in Table 3 show that the proposed algorithms perform with higher average performance and less

Table 3: Classification performance (% Mean  $\pm$  Variance) of the proposed algorithms and the shape descriptors on the Animals Dataset.

<b>CSS</b> [32]	<b>IDSC</b> [25]	<b>CSSP</b> [3]	<b>NN-SVM</b> [4]
69.7	73.6	78.4	84.30 $\pm$ 1.0
<b>Lim</b> [24]	<b>MC-1</b> ( $\mathcal{S}_s$ )	<b>MC-2</b> ( $\mathcal{S}_{est}$ )	<b>MC-3</b> ( $\mathcal{S}_{rbt}$ )
80.4	84.2 $\pm$ 2.1	<b>85.9 <math>\pm</math> 2.4</b>	84.9 $\pm$ 2.0

variance compared to the algorithms. Note that similar variance values are obtained for the proposed algorithms. In addition, we report the performances of the proposed algorithms over 10 sets of experiments unlike 100 sets of experiments reported for the **NN-SVM** [4]. Therefore, the higher variance of the performance values can be attributed to the less variance of the data generated in the experimental setup. Additional experimental results are given in the Supplemental Material.

## 5. Conclusion

Multi-component Shape Representation problem has been proposed for design and construction of shape representation models on manifolds in a Compositional Hierarchical Shape Vocabulary (CHSV). Shape representations are modeled with topologies induced by the statistical properties of the data and geometric properties of shape spaces that we want to achieve. The relationship between data and shape spaces has been reformulated using RFs in order construct shape spaces with manifold structure at the first layer of the CHSV.

We have suggested a mathematical framework of an indexing mechanism to employ and encode the geometric information obtained from measurements in the shape models. We have modeled formation of shape models at the higher layers of the CHSV using a disjoint union topology of shape manifolds induced by the indexing structure of the CHSV. The proposed methods have been used for structural shape classification. In the experiments, we have constructed different shape models by designing various shape manifolds with different invariance properties for shape classification.

## 6. Acknowledgement

This work was supported in part by the European Commission project PaCMan EU FP7-ICT, 600918.



## References

- [1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, 2008. 2
- [2] U. Aktas, M. Ozay, A. Leonardis, and J. L. Wyatt. A graph theoretic approach for object shape representation in compositional hierarchies using a hybrid generative-descriptive model. In *ECCV*, pages 566–581, 2014. 5, 6
- [3] X. Bai, W. Liu, and Z. Tu. Integrating contour and skeleton for shape classification. In *ICCV*, pages 360–367, Sept 2009. 8
- [4] X. Bai, C. Rao, and X. Wang. Shape vocabulary: A robust and efficient shape representation for shape matching. *IEEE TIP*, 23(9):3935–3949, Sept 2014. 8
- [5] M. Bernstein, V. De Silva, J. C. Langford, and J. B. Tenenbaum. Graph approximations to geodesics on embedded manifolds. Technical report, Technical report, Department of Psychology, Stanford University, 2000. 6
- [6] A. Bhattacharya and R. Bhattacharya. *Nonparametric Inference on Manifolds: With Applications to Shape Spaces*. Cambridge University Press, New York, NY, USA, 2012. 1, 2
- [7] M. Chen, K. Q. Weinberger, F. Sha, and Y. Bengio. Marginalized denoising auto-encoders for nonlinear representations. In T. Jebara and E. P. Xing, editors, *ICML*, pages 1476–1484. JMLR Workshop and Conference Proceedings, 2014. 7
- [8] I. Christou. Coordination of cluster ensembles via exact methods. *IEEE TPAMI*, 33(2):279–293, Feb 2011. 5
- [9] M. Crossley. *Essential Topology*. Springer Undergraduate Mathematics Series. Springer, 2006. 2
- [10] Y. Deng, A. Rangarajan, S. Eisenschenk, and B. Vemuri. A riemannian framework for matching point clouds represented by the schrodinger distance transform. In *CVPR*, pages 3756–3761, June 2014. 2
- [11] I. Dryden and K. Mardia. *Statistical shape analysis*. Wiley series in probability and statistics. Wiley, 1998. 1, 3, 5
- [12] H. G. Feichtinger and T. Strohmer, editors. *Gabor Analysis and Algorithms: Theory and Applications*. Birkhauser Boston, 1997. 3
- [13] J. Gallier, D. Xu, and M. Siqueira. Parametric pseudo-manifolds. *Differential Geometry and its Applications*, 30(6):702–736, 2012. 4
- [14] S. Jayasumana, R. Hartley, M. Salzmann, H. Li, and M. Harandi. Optimizing over radial kernels on compact manifolds. In *CVPR*, pages 3802–3809, June 2014. 1, 2, 8
- [15] S. Jayasumana, M. Salzmann, H. Li, and M. Harandi. A framework for shape analysis via hilbert space embedding. In *ICCV*, pages 1249–1256, 2013. 1, 2, 8
- [16] D. Kendall, D. Barden, T. Carne, and H. Le. *Shape and Shape Theory*. Wiley Series in Probability and Statistics. Wiley, 2009. 1, 2, 3
- [17] D. G. Kendall. A survey of the statistical theory of shape. *Statist. Sci.*, 4(2):87–99, 05 1989. 1
- [18] S. Kim and E. P. Xing. Tree-guided group lasso for multi-task regression with structured sparsity. In *ICML*, pages 543–550, 2010. 6, 7
- [19] S. Kurttek, E. Klassen, J. Gore, Z. Ding, and A. Srivastava. Elastic geodesic paths in shape space of parameterized surfaces. *IEEE TPAMI*, 34(9):1717–1730, Sept 2012. 1, 2
- [20] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, and Y. Bengio. An empirical evaluation of deep architectures on problems with many factors of variation. In *ICML*, pages 473–480, 2007. 7
- [21] J. M. Lee. *Introduction to smooth manifolds*. Springer, New York, Berlin, Heidelberg, 2013. 2
- [22] J. Lewi, R. Butera, D. M. Schneider, S. Woolley, and L. Paninski. Designing neurophysiology experiments to optimally constrain receptive field models along parametric sub-manifolds. In *NIPS*, pages 945–952. 2009. 1
- [23] H. Li, K. Zhang, and T. Jiang. Minimum entropy clustering and applications to gene expression analysis. In *CSB*, pages 142–151, Washington, DC, USA, 2004. 5
- [24] K.-L. Lim and H. K. Galoogahi. Shape classification using local and global features. *PSIVT*, pages 115–120, 2010. 8
- [25] H. Ling and D. W. Jacobs. Shape classification using the inner-distance. *IEEE TPAMI*, 29(2):286–299, Feb 2007. 8
- [26] A. M. Peter and A. Rangarajan. Information geometry for landmark shape analysis: Unifying shape representation and deformation. *IEEE TPAMI*, 31(2):337–350, 2009. 1, 2
- [27] G. Peyré. Manifold models for signals and images. *Comput. Vis. Image Underst.*, 113(2):249–260, Feb 2009. 1
- [28] D. Raviv, R. Kimmel, and A. M. Bruckstein. Graph isomorphisms and automorphisms via spectral signatures. *IEEE TPAMI*, 35(8):1985–1993, Aug 2013. 6
- [29] C. Samir, A. Srivastava, M. Daoudi, and E. Klassen. An intrinsic framework for analysis of facial surfaces. *Int. J. Comput. Vision*, 82(1):80–95, Apr 2009. 1
- [30] C. Small. *The Statistical Theory of Shape*. Springer Series in Statistics. Springer New York, 1996. 1, 2
- [31] A. Srivastava, E. Klassen, S. Joshi, and I. Jermyn. Shape analysis of elastic curves in euclidean spaces. *IEEE TPAMI*, 33(7):1415–1428, July 2011. 1
- [32] K. B. Sun and B. J. Super. Classification of contour shapes using class segment sets. In *CVPR*, CVPR '05, pages 727–733, 2005. 8
- [33] L. Tu. *An Introduction to Manifolds*. Springer, 2007. 2
- [34] J. R. Ullmann. An algorithm for subgraph isomorphism. *J. ACM*, 23(1):31–42, Jan 1976. 6
- [35] J. Wang, K. Markert, and M. Everingham. Learning models for object recognition from natural language descriptions. In *BMVC*, pages 1–11, 2009. 8
- [36] Q. Xie, I. Jermyn, S. Kurttek, and A. Srivastava. Numerical inversion of srnfs for efficient elastic shape analysis of star-shaped objects. In *ECCV*, pages 485–499. 2014. 1, 2