EUROGRAPHICS 2010 / M. Chen and O. Deussen
(Guest Editors)

Volume 30 (2011), Number 2

Prior Knowledge for Part Correspondence

Oliver van Kaick[1], Andrea Tagliasacchi[1], Oana Sidi[2], Hao Zhang[1], Daniel Cohen-Or[2], Lior Wolf[2], and Ghassan Hamarneh[1]

[1] School of Computing Science, Simon Fraser University, Canada
[2] School of Computer Science, Tel Aviv University, Israel

Abstract

Classical approaches to shape correspondence base their computation purely on the properties, in particular geometric similarity, of the shapes in question. Their performance still falls far short of that of humans in challenging cases where corresponding shape parts may differ significantly in geometry or even topology. We stipulate that in these cases, shape correspondence by humans involves recognition of the shape parts where prior knowledge on the parts would play a more dominant role than geometric similarity. We introduce an approach to part correspondence which incorporates prior knowledge imparted by a training set of pre-segmented, labeled models and combines the knowledge with content-driven analysis based on geometric similarity between the matched shapes. First, the prior knowledge is learned from the training set in the form of per-label classifiers. Next, given two query shapes to be matched, we apply the classifiers to assign a probabilistic label to each shape face. Finally, by means of a joint labeling scheme, the probabilistic labels are used synergistically with pairwise assignments derived from geometric similarity to provide the resulting part correspondence. We show that the incorporation of knowledge is especially effective in dealing with shapes exhibiting large intra-class variations. We also show that combining knowledge and content analyses outperforms approaches guided by either attribute alone.

1. Introduction

Most efforts on geometry processing in the graphics community have relied on low-level reasoning and operated on low-level features. Recently, there seems to be a research trend towards higher-level geometry processing, particularly studies of shapes at a more semantic level [ARSF09, GSMCO09, SNKS09, XWY*09, KHS10, MYY*10]. Shape correspondence is a fundamental problem that often requires a higher-level understanding of shapes. Applications such as attribute transfer [SP04, BVGP09], morphing [Ale02], shape synthesis [ASK*05, XLZ*10], and object recognition [BBM05] are often meant to employ correspondences between shape parts which possess the same meaning or functionality rather than mere geometric similarity.

Classical approaches to shape correspondence are mainly *content-driven* [FS06, JZvK07, ZSCO*08, LF09, ACOT*10, vKZHCO10], focusing solely on the geometrical and structural similarities between the query shapes (shapes to be matched). However, both criteria can be violated in challenging scenarios where there are large variations in the geometry or topology of the corresponding parts, as the examples in Figures 1 and 2 show. Shape correspondence under these circumstances is simply beyond pure geometry analysis and requires a semantic analysis of the shapes. Such an analysis necessitates the use of prior knowledge: to find a correspondence between parts that may be highly dissimilar geometri-

Figure 1: *Meaningful correspondence between shape parts under significant geometric (right pair: missing mouth) and topological (both pairs: one vs. multiple handles) discrepancies is made possible by incorporating prior knowledge. Corresponding parts are implied by matching colors.*

Figure 2: *A set of man-made shapes used in our work. Note the large intra-class geometric and topological variability.*

cally, we need to invoke our stored memory of similar parts (a *recognition* process) that are known to correspond to each other and use such knowledge to establish a correspondence between the unknown parts. Incorporating recognition into shape correspondence would result in a *knowledge-driven* approach. The power of knowledge is exemplified when pure geometry analysis simply cannot succeed (Figure 1).

In this paper, we introduce an approach to shape correspondence which incorporates prior knowledge. Aiming to mimic the human cognitive process where recognition is known to be primarily part-based [Mar82, HR87], we compute part labels (a recognition process) in conjunction with correspondence. The result is thus a *part correspondence* in contrast to correspondences between low-level feature points, as done in most works to date [vKZHCO10]. The prior knowledge is imparted through a training set of pre-segmented models with semantic labels. The training set serves as a knowledge medium allowing to find the correspondence between geometrically dissimilar parts. With the prior knowledge, we learn a *probabilistic semantic label* for each face of a query shape, where the labeling is derived, with the aid of a classifier, from the similarity of the descriptors of the faces to those of a class in the training set.

Part correspondence may be established based solely on knowledge and individual labeling of the query shapes. However, this may lead to an unsatisfactory outcome when the knowledge set is incomplete or produces indeterminate recognition results. Such cases may have to rely, at least partially, on a direct comparison between the query shapes. We combine the use of knowledge with content-driven analysis, where the latter is based on geometric similarity between the query shapes. The final correspondence is thus obtained through a *joint labeling* of the query shapes. The joint labeling makes use of the knowledge-driven probabilistic labels as well as feature pairings between query shapes to extract the actual parts that we seek to recognize and match across the shapes. The pairing of features incorporates the direct similarity between local regions of the query shapes.

Our contribution is two-fold. First, we show that the incorporation of prior knowledge to shape correspondence is effective. In particular, it leads to significant improvement over classical approaches on query shapes exhibiting large intra-class part variability in geometry or topology (see Figures 1 and 2). Second, we show that the joint labeling approach is effective. Specifically, knowledge-driven and content-driven analyses work synergistically and complement each other in instances where using solely the knowledge or the local geometric similarity of the shapes is insufficient.

2. Related work

Shape correspondence is a fundamental and well-studied problem in geometry processing [vKZHCO10]. The same can also be said about the related problems of shape retrieval and recognition [TV04, IJL*05, ABM*06]. Here we only discuss previous work most closely related to ours. More focus is given to automatic methods which study semantics of shapes or those leaning towards that direction.

Content-driven approaches to shape correspondence compare a pair of shapes based on geometric similarities between matched features, an approximate isometry criterion, or a combination of both [FS06, BBBK08, ZSCO*08, LF09, vKZHCO10]. The computational paradigm is an optimization or discrete search guided by these criteria. The recent methods of Zhang et al. [ZSCO*08] and Au et al. [ACOT*10] both allow large shape variations, but only to a certain extent, as they are still confined by the premise of geometric similarity and do not model shape semantics.

Part correspondence brings relevance to the segmentation problem. Many approaches to meaningful shape segmentation [Sha08] follow the minima rule [HR87], where cut boundaries are defined near concave regions. Other methods identify shape parts based on their geometric characteristics such as convexity and compactness [KJS07]. Clustering using an intrinsic surface metric or curvature is also common [KT03, LZ07]. Structural approaches mainly focus on skeleton topology [ACOT*10, SSCO08]. While satisfactory individual segmentation results can be obtained, these approaches are not designed to find a consistent segmentation between the shapes, i.e., a segmentation that divides the models into similar parts that correspond across the shapes.

Golovinskiy and Funkhouser address the consistent segmentation problem [GF09] where the connection between matching parts is initially built by a global rigid alignment using iterative closest point. No shape semantics are incorporated and their approach is not designed to handle large intra-class geometric variations such as stretching. Recent work of Xu et al. [XLZ*10] handles non-homogeneous part stretching by grouping the shapes based on their style and then performing part correspondence. Perhaps the first work on explicitly incorporating prior knowledge into shape segmentation is that of Simari et al. [SNKS09], where a multi-objective optimization is performed to segment and label a given shape. However, the user is required to provide semantic knowledge specific to the shape or shape parts and formulate such knowledge to fit the optimization framework; no training set or learning is used.

The ability to tag, annotate, or label a shape lies at the heart of semantic shape analysis. Manual annotation allows the user to either semantically label parts or to apply an ontology to the structure of the shapes [ARSF09]. Another group of approaches is based on correspondence analysis guided by shape geometry. As an application of their skeleton-driven correspondence algorithm, Au et al. [ACOT*10] assign semantic tags to the skeleton features of a subject shape by fully matching it to each labeled shape in a training set and then applying a simple majority voting to determine the tags. Shapira et al. [SSS*09] tag parts in a similar manner, but rely on a part-in-whole contextual signature to retrieve the most relevant parts.

Recently, Kalogerakis et al. [KHS10] present a method to learn the semantic labels of a shape based on training data. This work is significant as it is the first generic learning-based method for semantic shape segmentation. Our work shares similarities with this method, such as the use of a training set of pre-segmented and pre-labeled shapes and classifiers to recognize the shape parts. However, we apply prior knowledge to solve the correspondence problem and augment knowledge-driven analysis with content-driven analysis via the joint labeling approach.

Recognition by correspondence is a classical paradigm in computer vision [BJ97], as are semantic segmentation, labeling, and classification in images based on learning [LW06, SCCOL06]. Such approaches are rare in shape analysis however. Images are typically feature-rich, with color and texture cues as well as foreground and background contrasts, which exemplify the usefulness of local feature patches. However, for typical 3D models of the kind we consider (Figure 2), the distinctiveness of local surface features is significantly reduced. Also, one may be confined to a limited training set which still contains diverse intra-class shape variations. While image-based methods can typically benefit from the availability of large data collections to form adequate training sets or the knowledge base, the same cannot be said about 3D model collections. These are some of the challenges we wish to address in our work.

3. Overview

In this section, we present an overview of our approach to incorporate prior knowledge for part correspondence; see Figures 3 and 4 for an illustration. Full details of the algorithmic components are covered in subsequent sections.

Problem setting. The central problem that we are addressing here is the computation of a meaningful correspondence between a source shape S and a target shape T (the query shapes). In our case, we require a correspondence that is defined at the part level, i.e., it maps groups of faces on S to groups of faces of equivalent parts on T, as opposed to a mapping of feature points from one shape to the other.

Content-driven analysis. A straightforward way to establish the correspondence between S and T is by computing a set of shape descriptors and matching the faces based on these descriptors, which we refer to as the content-driven approach. In this context, one extracts for each face a set of descriptors that capture information about its geometric properties and context. This analysis associates a descriptor vector D_{s_i} with a face $s_i \in S$ (and D_{t_i} similarly for $t_i \in T$). Ideally, by measuring the distance between two such vectors we obtain an indication of the similarity of two faces (or their surrounding regions). Now, a correspondence between S and T can be derived from their descriptor vectors $D_S = \{D_{s_1}, D_{s_2}, \ldots, D_{s_n}\}$ and $D_T = \{D_{t_1}, D_{t_2}, \ldots, D_{t_m}\}$, with $n = |S|$ and $m = |T|$. The correspondence can be computed by any algorithm $\mathcal{A}$ that accepts as input two sets of vectors and computes a correspondence $\mathcal{C} = \mathcal{A}(D_S, D_T)$ based on the descriptor similarities (examples for $\mathcal{A}$ include bipartite matching or quadratic optimization [vKZHCO10]).

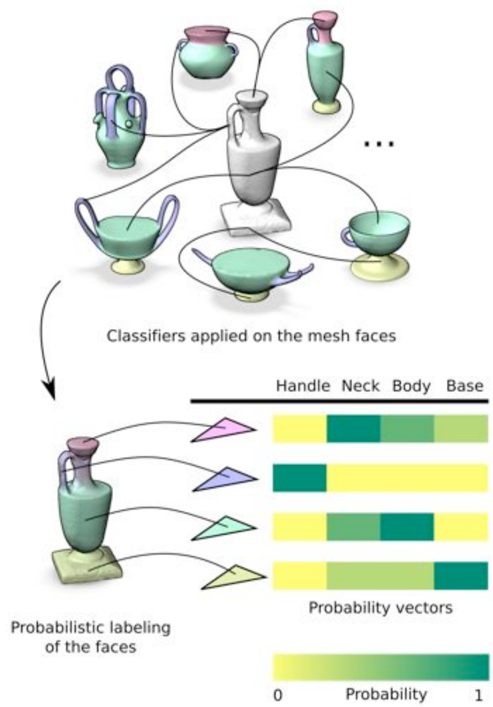

Figure 3: *Probabilistic semantic labeling: given a query shape (the gray vase), we first label each of its faces with per-label classifiers. The result is a label probability vector per face (shown as a color-blend on the surface of the bottom shape). The classifiers are learned from the training set (shown around the query with their ground-truth labeling).*

Incorporation of semantic knowledge. One of our goals is to show that a more meaningful correspondence is obtained when semantic information is added to the solution. This is especially effective when the shapes vary greatly in geometry and part constitution. In the knowledge-driven component of our approach, semantic information is derived from a set of training shapes and characterized by classifiers for part labels. The shapes in the training set are pre-segmented and tagged with semantic labels. The labeling is defined at the face level. Next, we compute shape descriptors for the faces of each training shape and, based on the descriptors, we train a classifier $\mathcal{K}_l$ for each label l, using the faces labeled l as positive examples of the label, and the remaining faces as negative examples. The learning phase results in one model per part type, each capturing the discriminative properties of the faces that belong to the label.

Probabilistic semantic labeling. Now, given a face s_i on a query shape, we compute the same set of descriptors and estimate $L_{s_i,l}$, the probability of s_i having the label l. $L_{s_i,l}$ is obtained by applying the classifier $\mathcal{K}_l$ to the descriptors of face s_i. By applying all the classifiers, a face receives a label probability vector L_{s_i} composed of the probabilities for all possible labels l (Figure 3). We denote the semantic labels for the whole shapes as L_S and L_T, with $L_S = \{L_{s_1}, L_{s_2}, \ldots, L_{s_n}\}$, and L_T similarly defined. In Section 4, we elaborate on the training and probabilistic labeling procedures.

Joint labeling and part correspondence. Each face belonging to each of the query shapes has now an associated probabilistic label. It would be possible to use the labels in place of our original descriptors and obtain the correspondence via a selected algorithm $\mathcal{C} = \mathcal{A}(L_S, L_T)$. However, we are interested in combining knowledge and geometric shape content, and we also require a correspondence that is defined at the part level. Therefore, we find the solution with a method that extracts parts from the shapes while simultaneously considering their correspondence. This step of the algorithm is achieved with our *joint labeling* (Figure 4).

The joint labeling takes into simultaneous consideration the probabilistic semantic labels of each face (based on knowledge/training), the mesh connectivity (captured by intra-mesh arcs between pairs of faces on the same shape), and connections between the faces of the two query shapes (feature pairing by inter-mesh arcs). The output of the joint labeling is a set of parts for each shape and their correspondence. The inter-mesh arcs come from candidate assignments extracted from the similarity of shape descriptors. Details on the joint labeling are described in Section 5.

The joint labeling has two advantages over the labeling based only on knowledge. First, there are practical limitations for the knowledge representation, such as the size and variability of the training set and accuracy of the shape descriptors. In this case, the inter-mesh edges complement the knowledge with extra information on the direct similarity between portions of the shapes. Moreover, even if the training set were large enough, covering an infinitude of variations, there would still be cases where a purely knowledge-driven approach could fail, such as when identical parts appear on different locations of a shape, or when there exist ambiguities in the functional role of the parts. The addition of content analysis contributes to the disambiguation of such cases.

4. Prior knowledge and probabilistic semantic labeling

Training set. The first step towards the development of our correspondence approach is to create a knowledge base from a dataset of training shapes. This dataset contains manually segmented shapes and the semantic labeling of each part (e.g., labels such as "leg" and "seat" for chairs). Thus, each mesh face is associated to a semantic label. An example dataset, used in part of our experiments, is shown in Figure 2.

Shape descriptors. Next, we compute a collection of descriptors for each face of all the training shapes. These descriptors should capture different properties of the faces, such as the local geometry of the shape around the face and

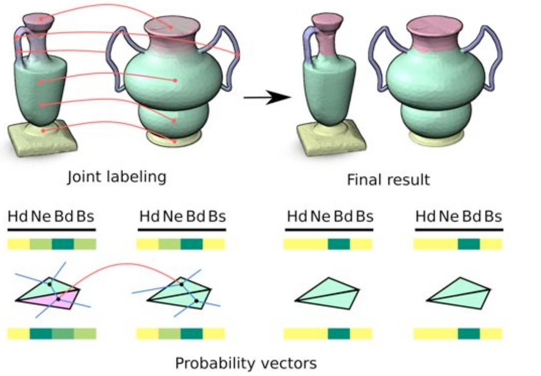

Figure 4: *Joint labeling: given a pair of shapes (to the left), we obtain their optimized labeling by making use of their probabilistic labels (Figure 3), intra-mesh arcs (blue lines) coming from the shapes' connectivity, and feature pairings given by the inter-mesh arcs (red lines) added between faces with similar descriptors. The result (to the right) is a segmentation of the shapes and correspondence at the segment level (indicated by matching colors). The probability vectors and the final deterministic labels are shown for two pairs of faces (before and after the joint labeling). The classes are: Handle (Hd), Neck (Ne), Body (Bd), and Base (Bs).*

its context in relation to the whole shape. The purpose of using a collection of descriptors is that their union should be rich enough to distinguish the faces of different classes. We extract descriptors similar to those that appear in the learning approach of Kalogerakis et al. [KHS10], more specifically, the ones based on principal component and curvature analysis in the neighborhood of a face, the shape diameter function [SSCO08], average of geodesic distances, and the binning of face areas into geodesic shape contexts.

Classifier training. Finally, we group the faces from the training set according to their labels. Suppose that for each label l, we are given n_l training faces coming from different shapes and grouped into the set $F_l = \{f_{1,l}, f_{2,l}, \ldots, f_{n_l,l}\}$. We compute the descriptors for each face $f_{i,l}$, denoted as $D_{f_{i,l}}$, and the full collection of descriptors for label l is given by $D_l = \{D_{f_{1,l}}, D_{f_{2,l}}, \ldots, D_{f_{n_l,l}}\}$. Then, for each label l, we train a classifier $\mathcal{K}_l$ with the descriptors D_l as positive examples of the label, and with the descriptors $D_{\bar{l}} = \cup_{D_{l'}}$, with $l' \neq l$, as negative examples. Notice that instead of training a set of per-label classifiers, a single multi-class classifier could also be used to take advantage of shared decision rules.

To train a classifier, we use the "gentleboost" algorithm, which has several advantages in relation to other choices, such as performing automatic feature selection, being a time-efficient training algorithm, and attaching confidence values to each classification decision [KHS10]. By adjusting the importance weight given to each training sample, it is also possible to account for unbalanced datasets. Details on this algorithm can be found in [FHT00]. The unnormalized confidence values returned by each classifier can be transformed into probabilities with the softmax activation function (a generalization of the logistic function to multiple variables).

Probabilistic semantic labeling. Now, given an unknown face s_i on a query shape S, we compute its associated set of descriptors D_{s_i} and provide them as input to the classifier $\mathcal{K}_l$. The classifier estimates the probability $L_{s_i,l}$ that face s_i should belong to class l. The probabilities for all possible labels are grouped into the vector L_{s_i}. The probabilistic labeling is performed for all faces of the two query shapes S and T. One of our key contributions here is to show that computing a correspondence based on the semantic labeling L_S provides superior results when compared to those computed directly from the descriptors D_S (see Section 6).

5. Part correspondence via joint labeling

After obtaining the probabilistic labeling of the faces on the query shapes, we utilize this information in a joint labeling scheme (i.e. labeling both query shapes simultaneously) to obtain the final result. By posing the correspondence problem as that of label optimization, we are able to simultaneously incorporate both the semantic information and local similarity of the two query shapes into the computation, while also extracting as a result semantic parts from the query shapes and their correspondence. This process is inspired by methods for consistent segmentation and labeling [GF09,XLZ*10,NAH10], although we deviate from their scheme and do not make use of rigid alignment.

Joint labeling. The label optimization problem is defined as the assignment of deterministic labels to the nodes of a graph such that a given energy is minimized [NAH10, SSS*09, KHS10]. The set of assigned labels is the same one used in the definition of the semantic information. Given a query shape S, we define a graph $G_S = \{V_S, E_S\}$, where the nodes V_S are the faces of the mesh and the arcs in E_S connect two faces if they are adjacent on the mesh. A similar graph can be defined for a shape T. We perform the joint labeling on a graph $G = \{V, E\}$, where $V = V_S \cup V_T$ and the connectivity of the graph is given by two types of arcs, $E = E_{\text{intra}} \cup E_{\text{inter}}$. The intra-mesh arcs are simply given by $E_{\text{intra}} = E_S \cup E_T$. The inter-mesh arcs E_{inter} connect faces in S to faces in T.

Feature pairing. We select a set of pairwise assignments from the faces on S to the faces on T, based on the similarity of shape descriptors, to constitute the inter-mesh arcs. However, to increase the quality of the assignments, we also incorporate a learning procedure into this step. We learn which shape descriptors are most discriminative in distinguishing correct from incorrect assignments. The learning is performed on assignments derived from the training set. Since a large collection of descriptors is available, a learning procedure also has the advantage of properly weighting the influence of each descriptor in the similarity computation.

First, we derive a set of training assignments $\mathcal{T}$ from the dataset of training shapes. For each individual shape descriptor, we compute its similarity for all possible pairs of faces in $S \times T$. Next, for each face in $\mathcal{S}$, we select the k first pairwise assignments with the highest similarity and add them to the set $\mathcal{T}$. The similarity is given by the inverse of the Euclidean distance between the descriptors of the two faces, while k is set to 10 throughout our experiments By choosing a small k, we select the assignments that are most likely to capture the correct relation between the different shape descriptors and the correspondence. Each assignment in $\mathcal{T}$ is either labeled as *true*, if it maps two faces with the same label, or *false*, if the labels are different. Next, we train a classifier based on a set of vectors derived from $\mathcal{T}$, also using the "gentleboost' algorithm. Each vector is associated to an assignment in $\mathcal{T}$ and has an entry for the similarity of each descriptor.

The outcome of the training is a classifier that allows us to label a candidate assignment with *true* or *false*, based on the vector of similarities obtained by considering multiple shape descriptors, while also assigning a confidence to this decision. Finally, given a pair of query shapes, we select candidate assignments with the same procedure used for the training data. We apply the classifier and select the top 20% assignments with the most confidence of having the label *true*, to obtain a small yet reliable set of inter-mesh arcs.

Labeling energy. The energy to be minimized by the labeling is composed of two types of terms: unary and binary. The unary term takes into consideration how likely it is that a given node has a specific label. This is encoded by a labeling cost assigned to each face. The binary terms consider the connectivity between faces (intra- and inter-mesh) and quantify how likely it is that two neighboring nodes have a specific pair of labels, according to a pairwise cost. We define the energy of the labeling $\mathbf{l}$ as

$$\mathcal{E}(\mathbf{l}) = \sum_{i \in V} \mathcal{U}(i,l_i) + \sum_{ij \in E_{\text{intra}}} \mathcal{B}_{\text{intra}}(i,j,l_i,l_j) + \sum_{ij \in E_{\text{inter}}} \mathcal{B}_{\text{inter}}(i,j,l_i,l_j), \quad (1)$$

where l_i and l_j are the labels of nodes i and j, respectively, and $\mathcal{U}$ and $\mathcal{B}_{\text{intra}}$, $\mathcal{B}_{\text{inter}}$ are the unary and binary terms.

The unary term is given by

$$\mathcal{U}(i,l_i) = -a_i \log P(l_i|\mathbf{x}_i), \quad (2)$$

where $P(l_i|\mathbf{x}_i)$ is the probability of node i having label l_i, based on the classifiers applied to the face descriptors $\mathbf{x}_i$, and a_i is the area of the face corresponding to node i. The weight a_i ensures that the cost is given in terms of labeling the total shape area.

The intra-mesh binary term is defined as

$$\mathcal{B}_{\text{intra}}(i,j,l_i,l_j) = L(l_i,l_j)[\lambda\alpha_{ij} + \mu\ell_{ij}], \quad (3)$$

where we take into account the compatibility $L(l_i,l_j)$ between two labels, as well as the edge length ℓ_{ij} and dihedral angle α_{ij} between faces i and j, similarly as done in [SSS*09] and [KHS10]. The label compatibility term L is derived from the training data in the form of statistics that quantify how likely it is that two labels appear neighboring each other. $L(l,l) = 0$ if two faces share the same label l. The parameters λ and μ regulate how much the angle and edge length contribute to the total energy.

Finally, the inter-mesh term is given by

$$\mathcal{B}_{\text{inter}}(i,j,l_i,l_j) = L(l_i,l_j)[\nu\sigma_{ij}], \quad (4)$$

where σ_{ij} is the confidence that the assignment between faces i and j is correct, and ν regulates the influence of the inter-mesh term to the total energy. Thus, the higher the confidence value attached to the assignment, the more the cost is increased if the labels are different.

Graph-cut optimization. We use multi-label graph-cuts to assign labels to the nodes in an optimal manner. More specifically, the α-β swap algorithm is utilized [BVZ01], since the pairwise costs L do not define a metric. By minimizing the given energy, we obtain the most likely label for each face while also avoiding the creation of small disconnected segments. The optimal parameters λ, μ, and ν are obtained by performing a grid search on training data separated for this purpose. This procedure is explained in Section 6.

6. Experimental results

In this section, we present a set of experiments aimed at evaluating our approach for shape correspondence. We also contrast our results to other state-of-the-art methods.

Datasets. We utilize two datasets in our experiments. All the shapes are pre-segmented and labeled, implying that the ground-truth label for each mesh face is known. We designed the first dataset, shown in Figure 2, composed of four classes of man-made shapes with large geometric and topological variability. Notice that the presence of some of the semantic parts is optional, and certain parts can appear more than once on each shape. The second dataset consists of a selected subset of classes from the mesh segmentation benchmark [CGF09]. The selected classes are listed in Table 1 and examples appear in Figure 5 (a)-(e). We utilize the segmentations and labelings created for these shapes by Kalogerakis et al. [KHS10]. We selected classes that possess shapes in different poses (e.g., Human or Hand) and shapes with considerable structural and geometric variability (e.g., FourLeg), as opposed to models with a predictable structure. Also note that the segmentations for the first dataset were created by a single user with a pre-defined goal, while the second dataset is given by the average segmentation for each class. With this setting, we demonstrate that our method is robust against variations in the dataset design process.

Figure 5: *Part correspondence results via joint labeling. Aside from the failure cases in (j), (t), and (y), our method succeeded even under significant geometric and topological variations between the matched parts, while purely content-driven methods would fail (see Figure 7 for a comparison). Corresponding segments between a pair of shapes are shown with the same color.*

Correspondence results. Figure 5 shows a set of visual results. To generate these correspondences, we selected for each class a random subset of 60% of the shapes as the training set, and delegated the remaining shapes to be test cases. A random subset of 3/4 of the training shapes was used to learn the classifiers, while the remaining 1/4 of shapes was used to select the best parameters for the joint labeling. We perform a two-level grid search on the three parameters λ, μ, and ν and select those that result in the best labeling of the selected 1/4 of the training data, according to the ground truth. Next, we compute the joint labeling for all the pairs of test shapes. Some of these test pairs are shown in Figure 5.

Firstly, we see in Figure 5 (a)-(e) that our method is able to establish meaningful correspondences for queries that are also handled by content-driven methods. In addition to that, and more importantly, we also observe that the prior knowledge is effective in matching shapes whose parts differ by significant geometric changes. Examples include matching different types of candle holders in (f)-(j), different torches and bases in (k)-(o), variations in the recipients and bases in (p)-(t), and seat rests with different railings in (u)-(x). Our method is also able to handle topological variations, as the different lamp supports in (k)-(m), and variations in chair legs as in (u) and (y). Finally, our method is also successful in matching shapes with different numbers of parts, such as the multiple wax candles in (h)-(j), and the different numbers of handles in (r) and (s). Notice that the presence of additional parts, such as the handles in (p) and (q), do not affect the accuracy of the results. Figure 5 also shows cases where our method did not provide the most accurate correspondence, mainly due to the lack of sufficient prior knowledge and limitations in the shape descriptors. Examples are the missed candle flame in (j), the extra handle added to the top of the vase in (t), and the rest and arms of the chair in (y) that were not properly separated from the seat.

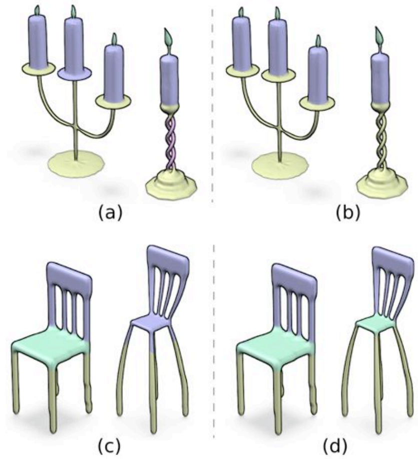

Figure 6: *Joint labeling improves upon the use of prior knowledge alone. (a) and (c): correspondences obtained using probabilistic labels only, without content analysis. (b) and (d): results from joint labeling. The latter is effective in implying the correct correspondence for mislabeled parts when there is sufficient part-to-part similarity.*

In cases of insufficient knowledge, an accurate correspondence can still be obtained via joint labeling when there is enough similarity between parts of the shapes. Such examples can be seen in Figure 6. In (a), the correspondence is computed based purely on the knowledge, i.e., only the unary and intra-mesh terms enter the labeling. Part of the support of the candle on the right is mistakenly labeled as a handle and, therefore, does not have a corresponding part on the shape on the left. However, when the content of the shapes is also taken into consideration (i.e., the inter-mesh term is added), we obtain the more accurate correspondence in (b), since the descriptors of the thin structures on both shapes are similar. We observe an analogous situation in (c), where the seat of the chair is mistakenly labeled as a backrest, due to the geometric distortion present in the model. In (d), the joint labeling is also able to obtain the correct part correspondence, guided by the similarity of the seats. We point out that both knowledge and content are essential for finding the correct correspondence in these cases.

In Table 1, we show a statistical evaluation of the correspondence results we obtained. For each shape class, we performed 5 experiments in the same manner as described before (60% training shapes and 40% test shapes). We average the accuracy over all the pairs of test shapes in each experiment. The accuracy for a single pair of shapes S and T is calculated with a function that measures the quality of the correspondence between the two shapes. It is given by

$$\text{Accuracy}(l) = \frac{\sum_{i \in S} \sum_{j \in T} a_i a_j \, |\delta(l_i = l_j) + \delta(t_i = t_j) - 1|}{\sum_{i \in S} \sum_{j \in T} a_i a_j}, \tag{5}$$

where a_i is the area of face i, l is the labeling of the shapes returned by our method, t is the ground-truth labeling, and $\delta(x = y)$ is 1 only if $x = y$. This quantity measures how many of the faces that have the same label in the ground-truth are also found in corresponding segments, independently of which label was assigned to the faces. It also captures the notion that two faces that do not possess the same label in the ground-truth should not be assigned to corresponding segments. The weighting is chosen so that the correspondence accuracy is given in terms of the total shape area.

We can see that the average accuracy for the first dataset (classes of man-made shapes with significant variability) is 90%, while for the second dataset (organic shapes), it is 89%. We attribute the failure cases to a few factors. Firstly, shapes that have parts with no counterparts in the prior knowledge can occur in the test set. An example is the chair shown before in Figure 5 (y), contrasted to the other shapes in the dataset (Figure 2). Secondly, the descriptors may not be sufficient to distinguish certain parts of the shapes, such as the different fingers in the hands. Finally, a small fraction of the errors is also due to imperfections in the labeling, such as dislocated borders which happen when the inter-mesh and intra-mesh terms compete in the optimization. Such cases can be adjusted, for example, with post-processing that displaces the borders to concave regions of the shapes [Sha08].

When comparing the statistical results of the approach using only prior knowledge (unary and intra-mesh terms) to that of the joint labeling (incorporating the inter-mesh term), we find that both approaches are comparable, with a deviation of $\pm 5\%$ in the accuracies. The joint labeling does not lead to a significant increase in the accuracies since the inter-mesh term is primarily designed to handle cases such as those in Figure 6, where the content aids in improving the correspondence for certain portions of the shapes that do not appear in the knowledge. Such cases are not very frequent in the datasets, but could be prevalent in certain situations (e.g., a collection of shapes modeled by reusing existing parts).

Comparison to classical approach. We compare our approach with two state-of-the-art content-driven methods that are considered to be comparatively competitive at handling general deformation between models. In Figure 7, we show a comparison with the deformation-driven approach of Zhang et al. [ZSCO*08], which finds a matching between sparse sets of feature points. As we can see, when the corresponding parts differ sufficiently in their scales or geometric properties, this method fails as it is still formulated on matching the geometry of the shapes. Our method on the other hand finds a correspondence between two shapes by using knowledge as the medium; it succeeds since the resulting correspondences have sufficient support from the training set. We also compared our method to that of Shapira et al. [SSS*09]. However, when presented with the query examples shown in Figure 7, the partitioning of the shapes obtained with this method differed significantly from one shape to the other,

Figure 7: *Comparison to content-driven correspondence on geometrically dissimilar models. (a) and (c): results from the deformation-driven method of Zhang et al. [ZSCO*08], including the induced deformation and the matching feature points. (b) and (d): results from our joint labeling method.*

preventing the method from establishing a meaningful correspondence. Other recent works on content-driven shape correspondence, e.g., [LF09, ACOT*10], are not expected to succeed on these examples either as they are also based on geometric similarities between matched features, an approximate isometry criterion, or a combination of both.

Timing. The most expensive procedure in the knowledge-driven framework is classifier learning. For a training set ranging from 20 to 30 models with average size of 30K triangles, this step can take 10 hours in an AMD Opteron 1GHz processor. Then, applying the classifiers on two queries and performing the joint labeling runs in the order of minutes.

7. Conclusions and future work

The main message of our work presented in this paper is that challenging cases of 3D shape correspondence can be solved effectively by incorporating prior knowledge. At the same time, direct geometric comparison between the query shapes is still merited particularly when the knowledge base is incomplete or leads to indeterminate recognition results. An effective approach via joint labeling, combining knowledge-driven probabilistic semantic labeling and content-driven analysis via local geometric similarity between query shapes, is thus introduced. We demonstrate significant improvement on shape correspondence results over classical approaches, particularly when the query shapes exhibit large geometric or even topological variations.

Limitations. While the idea of joint labeling is quite general, the content analysis component of our approach is still fairly primitive in terms of the feature similarity employed. There still remain failure cases as shown in Figure 5. We believe that this can be attributed to our current reliance of low-level shape descriptors in the content-driven analysis as well as in the recognition step, where a query shape is compared to shapes in the training set. More advanced geometric analysis tools incorporating criteria such as shape symmetry [GPF07] or style-content separation [XLZ*10] may lead to improvement in this regard.

Table 1: *Statistical evaluation of the correspondence results.*

Class	Corr.	Class	Corr.	Class	Corr.
Candle	88%	Airplane	92%	FourLeg	82%
Chair	87%	Ant	96%	Hand	80%
Lamp	97%	Bird	86%	Human	90%
Vase	86%	Fish	92%	Octopus	96%

Future work. In addition to improving the quality of the shape descriptors, we also plan to explore additional measures such as structural similarities for joint labeling. We would also like to explore the possibility of avoiding the need to pre-classify the training shapes. Either we integrate the recognition of the training parts with their correspondences, or avoid it altogether by developing generic labels and classifiers that capture the general notion of handle and base, rather than a cup handle or a lamp base, for example. In hindsight, what ultimately makes the correspondence approach effective under challenging circumstances, particularly for many classes of man-made shapes, is the ability to learn the *functionality* of the parts. Recognizing functionality is clearly a difficult problem. It calls for intermediate-level descriptors that are able to capture properties such as flatness, concavity, symmetry, and fuzziness that are required in order to achieve a certain functionality. Learning functionalities of parts for shape analysis is certainly an interesting direction for future work.

Acknowledgements. We thank Evangelos Kalogerakis for providing the dataset of labeled shapes and information on descriptor computation. We thank Xiaobai Chen for the segmentation benchmark, Daniela Giorgi and AIM@SHAPE for the meshes from the SHREC 2007 Watertight Track, and Ran Gal for additional man-made models. We also thank Lior Shapira and Shy Shalom for their help with the SDF. This work was supported in part by grants from the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities, and the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[ABM*06] ATTENE M., BIASOTTI S., MORTARA M., PATANÉ G., SPAGNUOLO M., FALCIDIENO B.: Computational methods for understanding 3D shapes. *Computers & Graphics 30*, 3 (2006), 323–333.

[ACOT*10] AU O. K.-C., COHEN-OR D., TAI C.-L., FU H., ZHENG Y.: Electors voting for fast automatic shape correspondence. *Computer Graphics Forum (Proc. EUROGRAPHICS) 29*, 2 (2010), 645–654.

[Ale02] ALEXA M.: Recent advances in mesh morphing. *Computer Graphics Forum 21*, 2 (2002), 173–198.

[ARSF09] ATTENE M., ROBBIANO F., SPAGNUOLO M., FALCIDIENO B.: Characterization of 3D shape parts for semantic annotation. *Computer-Aided Design 41*, 10 (2009), 756–763.

[ASK*05] ANGUELOV D., SRINIVASAN P., KOLLER D., THRUN S., RODGERS J., DAVIS J.: SCAPE: shape completion and animation of people. *ACM Trans. on Graphics (Proc. SIGGRAPH) 24*, 3 (2005), 408–416.

[BBBK08] BRONSTEIN A. M., BRONSTEIN M. M., BRUCKSTEIN A. M., KIMMEL R.: Partial similarity of objects, or how to compare a centaur to a horse. *Int. J. Comput. Vision 84*, 2 (2008), 163–183.

[BBM05] BERG A. C., BERG T. L., MALIK J.: Shape matching and object recognition using low distortion correspondence. In *Proc. IEEE Conf. on CVPR* (2005), vol. 1, pp. 26–33.

[BJ97] BASRI R., JACOBS D. W.: Recognition using region correspondences. *Int. J. Comput. Vision 25*, 2 (1997), 145–166.

[BVGP09] BARAN I., VLASIC D., GRINSPUN E., POPOVIĆ J.: Semantic deformation transfer. *ACM Trans. on Graphics (Proc. SIGGRAPH) 28*, 3 (2009).

[BVZ01] BOYKOV Y., VEKSLER O., ZABIH R.: Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell. 23*, 11 (2001), 1222–1239.

[CGF09] CHEN X., GOLOVINSKIY A., , FUNKHOUSER T.: A benchmark for 3D mesh segmentation. *ACM Trans. on Graphics (Proc. SIGGRAPH) 28*, 3 (2009).

[FHT00] FRIEDMAN J., HASTIE T., TIBSHIRANI R.: Additive logistic regression: a statistical view of boosting. *Annals of statistics 28*, 2 (2000), 337–374.

[FS06] FUNKHOUSER T., SHILANE P.: Partial matching of 3D shapes with priority-driven search. In *Proc. Symp. on Geom. Processing (SGP)* (2006), pp. 131–142.

[GF09] GOLOVINSKIY A., FUNKHOUSER T.: Consistent segmentation of 3D models. *Computers & Graphics (Proc. of SMI) 33*, 3 (2009), 262–269.

[GPF07] GOLOVINSKIY A., PODOLAK J., FUNKHOUSER T.: Symmetry-aware mesh processing. *Princeton University TR-782-07* (2007).

[GSMCO09] GAL R., SORKINE O., MITRA N. J., COHEN-OR D.: iWIRES: an analyze-and-edit approach to shape manipulation. *ACM Trans. on Graphics (Proc. SIGGRAPH) 28*, 3 (2009).

[HR87] HOFFMAN D. D., RICHARDS W. A.: Parts of recognition. *Readings in computer vision: issues, problems, principles, and paradigms* (1987), 227–242.

[IJL*05] IYER N., JAYANTI S., LOU K., KALYANARAMAN Y., RAMANI K.: Three-dimensional shape searching: state-of-the-art review and future trends. *Computer-Aided Design 37*, 5 (2005), 509–530.

[JZvK07] JAIN V., ZHANG H., VAN KAICK O.: Non-rigid spectral correspondence of triangle meshes. *International Journal on Shape Modeling 13*, 1 (2007), 101–124.

[KHS10] KALOGERAKIS E., HERTZMANN A., SINGH K.: Learning 3D mesh segmentation and labeling. *ACM Trans. on Graphics (Proc. SIGGRAPH) 29*, 3 (2010).

[KJS07] KRAEVOY V., JULIUS D., SHEFFER A.: Model composition from interchangeable components. In *Proc. Pacific Graphics* (2007), pp. 129–138.

[KT03] KATZ S., TAL A.: Hierarchical mesh decomposition using fuzzy clustering and cuts. *ACM Trans. on Graphics (Proc. SIGGRAPH) 22*, 3 (2003), 954–961.

[LF09] LIPMAN Y., FUNKHOUSER T.: Möbius voting for surface correspondence. *ACM Trans. on Graphics (Proc. SIGGRAPH) 28*, 3 (2009).

[LW06] LEVIN A., WEISS Y.: Learning to combine bottom-up and top-down segmentation. *Int. J. Comput. Vision (Proc. ECCV) 81*, 1 (2006), 105–118.

[LZ07] LIU R., ZHANG H.: Mesh segmentation via spectral embedding and contour analysis. *Computer Graphics Forum (Proc. EUROGRAPHICS) 26*, 3 (2007), 385–394.

[Mar82] MARR D.: *Vision - A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman, 1982.

[MYY*10] MITRA N. J., YANG Y.-L., YAN D.-M., LI W., AGRAWALA M.: Illustrating how mechanical assemblies work. *ACM Trans. on Graphics (Proc. SIGGRAPH) 29*, 4 (2010), 1–12.

[NAH10] NG B., ABUGHARBIEH R., HAMARNEH G.: Group MRF for fMRI activation detection. In *Proc. IEEE Conf. on CVPR* (2010), pp. 2887–2894.

[SCCOL06] SCHNITMAN Y., CASPI Y., COHEN-OR D., LISCHINSKI D.: Inducing semantic segmentation from an example. In *Proc. of ACCV* (2006), vol. 2, pp. 373–384.

[Sha08] SHAMIR A.: A survey on mesh segmentation techniques. *Computer Graphics Forum 27*, 6 (2008), 1539–1556.

[SNKS09] SIMARI P., NOWROUZEZAHRAI D., KALOGERAKIS E., SINGH K.: Multi-objective shape segmentation and labeling. *Computer Graphics Forum (SGP) 28*, 5 (2009), 1415–1425.

[SP04] SUMNER R. W., POPOVIĆ J.: Deformation transfer for triangle meshes. *ACM Trans. on Graphics (Proc. SIGGRAPH) 23*, 3 (2004), 399–405.

[SSCO08] SHAPIRA L., SHAMIR A., COHEN-OR D.: Consistent mesh partitioning and skeletonization using the shape diameter function. *The Visual Computer 24*, 4 (2008), 249–259.

[SSS*09] SHAPIRA L., SHALOM S., SHAMIR A., COHEN-OR D., ZHANG H.: Contextual part analogies in 3D objects. *Int. J. Comput. Vision 89*, 2–3 (2009), 309–326.

[TV04] TANGELDER J. W. H., VELTKAMP R. C.: A survey of content based 3D shape retrieval methods. In *Proc. IEEE Conf. on Shape Modeling and Applications* (2004), pp. 145–156.

[vKZHCO10] VAN KAICK O., ZHANG H., HAMARNEH G., COHEN-OR D.: A survey on shape correspondence. In *Proc. Eurographics State-of-the-Art Report* (2010).

[XLZ*10] XU K., LI H., ZHANG H., COHEN-OR D., XIONG Y., CHENG Z.: Style-content separation by anisotropic part scales. *ACM Trans. on Graphics (SIGGRAPH Asia) 29*, 5 (2010).

[XWY*09] XU W., WANG J., YIN K., ZHOU K., VAN DE PANNE M., CHEN F., GUO B.: Joint-aware manipulation of deformable models. *ACM Trans. on Graphics (Proc. SIGGRAPH) 28*, 3 (2009), 1–9.

[ZSCO*08] ZHANG H., SHEFFER A., COHEN-OR D., ZHOU Q., VAN KAICK O., TAGLIASACCHI A.: Deformation-driven shape correspondence. *Computer Graphics Forum (Proc. SGP) 27*, 5 (2008), 1431–1439.