



**HAL**  
open science

## The 2D shape structure dataset: A user annotated open access database

Axel Carlier, Kathryn Leonard, Stefanie Hahmann, Géraldine Morin, Misha Collins

### ► To cite this version:

Axel Carlier, Kathryn Leonard, Stefanie Hahmann, Géraldine Morin, Misha Collins. The 2D shape structure dataset: A user annotated open access database. *Computers and Graphics*, 2016, (SMI 2016), 58, pp.23-30. 10.1016/j.cag.2016.05.009 . hal-01322964

**HAL Id: hal-01322964**

**<https://inria.hal.science/hal-01322964>**

Submitted on 29 May 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The 2D Shape Structure Dataset: A User Annotated Open Access Database

Axel Carlier<sup>a,\*</sup>, Kathryn Leonard<sup>b</sup>, Stefanie Hahmann<sup>c</sup>, Geraldine Morin<sup>a</sup>, Misha Collins<sup>d</sup>

<sup>a</sup>University of Toulouse, CNRS (Laboratoire IRIT), France

<sup>b</sup>Department of Mathematics, California State University Channel Islands, Camarillo, CA, USA

<sup>c</sup>Université Grenoble-Alpes, CNRS (Laboratoire Jean Kuntzmann), Inria, Grenoble, France

<sup>d</sup>GISHWHES, Inc., Los Angeles, CA, USA

---

## Abstract

In this paper we present the *2D Shape Structure* database, a public, user-generated dataset of 2D shape decompositions into a hierarchy of shape parts with geometric relationships retained. It is the outcome of a large-scale user study obtained by crowdsourcing, involving over 1200 shapes in 70 shape classes, and 2861 participants. A total of 41953 annotations has been collected with at least 24 annotations per shape. For each shape, user decompositions into main shape, one or more levels of parts, and a level of details are available. This database reinforces a philosophy that understanding shape structure as a whole, rather than in the separated categories of parts decomposition, parts hierarchy, and analysis of relationships between parts, is crucial for full shape understanding. We provide initial statistical explorations of the data to determine representative (“mean”) shape annotations and to determine the number of modes in the annotations. The primary goal of the paper is to make this rich and complex database openly available (through the website <http://2dshapesstructure.github.io/index.html>), providing the shape community with a ground truth of human perception of holistic shape structure.

*Keywords:* 2D shape, shape decomposition, shape hierarchy, shape structure, user study, datasets, crowdsourcing, medial axis,

---

## 1. Introduction

This paper presents a large user study of geometric shape structure annotation for the MPEG-7 database of 2D shapes. We define a *geometric shape structure* as a shape representation containing the following three components: (1) a decomposition of the shape into parts based as much as possible on geometry rather than semantics, (2) a hierarchy of parts representing the scale at which each part becomes significant, and (3) an understanding of the relationship between parts at various scales and locations.

Each of the three components is of interest on its own. 2D shape decomposition into meaningful parts is a fundamental problem that has been studied in different fields such as shape modeling, computer vision, psychology and cognition with a broad number of application domains such as shape analysis, shape matching, shape retrieval, shape representations and modeling. Part hierarchies allow for multiscale shape decomposition and shape comparison. Part relationships allow for shape alignment and recognition. Together, the geometric shape structure allows for full shape understanding, which is essential for structure-aware shape processing applications such as shape synthesis, modeling or animation [1].

There has been extensive research done the last two decades on shape decompositions [2-6,8-18]. The different criteria used to guide a partition of a shape into “meaningful parts” may vary between geometric, functional, probabilistic and semantic

or may be expressed in terms of minimizing some cost function. These criteria are generally derived from or related to psychological studies showing that humans use generic perception rules when they are asked to decompose an outline into parts. As stated recently by Lai and Wang [2], however, human decomposition behavior is so complex that despite an impressive number of existing methods there is still a need for improvement in order to comprehensively describe human behavior and to derive computationally faithful algorithms. Validation of segmentation algorithms thus becomes a tedious task.

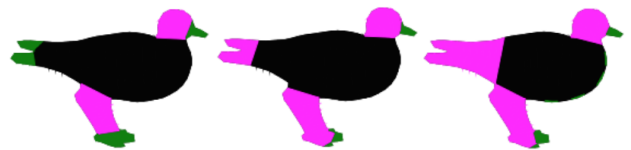


Figure 1: Annotations from our user study: main shape (black), parts (magenta), details (green). These annotations provide three possible decompositions of the shape. While users agree about the hierarchy of the selected parts, they differ in their selection of parts.

Without a ground truth data set, existing methods must settle for comparison to previous methods, but the goal should be to provide a segmentation coherent with human perception. Given that human behaviors in shape decomposition cannot be uniquely defined in terms of geometry [2], a ground truth is best obtained by a user study where the number and variety of participating subjects is large enough to be as meaningful as possible. On one hand, there may exist two or more valid decompositions for the same shape (Figure 1). On the other hand,

---

\*Corresponding author

Email address: [Axel.Carlier@enseeiht.fr](mailto:Axel.Carlier@enseeiht.fr) (Axel Carlier)

the same partition of the shape into parts can lead to different hierarchies (Figure 2). In the literature, existing studies lack significant participant size or shape variety to capture this nuance, as the experiments are usually filtered to remove singular users (outliers), and one ground truth segmentation is retained (typically the majority vote) [3, 4, 5, 6].



Figure 2: Two meaningful annotations for a butterfly from our user study: main shape (black), parts (magenta), details (green). These annotations give similar decompositions but different shape structures due to the difference in parts hierarchy. Whereas the users agree about the shape details (green), they diverge in determining the main shape.

In this paper we present the *2D Shape Structure Dataset* collected by an extended user study via crowdsourcing to obtain 2D shape partitioning into a hierarchy of shape parts. Users in the study provided a shape decomposition based on the Delaunay triangulation of the shape, which is known to capture human shape perception [7], and labeled each triangle to indicate its level in the part hierarchy. As shown in Figure 3, the edges of the Delaunay triangulation accommodate part transitions well. These edges correspond to the local axes of symmetry, and will occur for any salient shape protrusion.

Because the Delaunay triangulation is dual to the discrete medial axis of the shape, the annotation provides all three components of a geometric shape structure: parts decomposition, parts hierarchy, and relationships between parts via the medial axis graph structure. Indeed, the link to the medial axis provides (a) local feature size via the medial axis radius function, (b) feature depth via any one of a number of measures such as SAT [8], (c) sub-tree structures capturing interaction between parts at different levels, and (d) geometric relationships between parts via the angle of connection of the corresponding medial branches.

The goals of this dataset are to collect and share user-defined shape parts and hierarchies in a format that encodes part relationships, to provide ground truth for shape decomposition for a popular 2D shape database, and to focus on geometric interpretation rather than semantics, function, or other categorization of parts. By sharing the data with the research community, we hope to prompt important discussions about geometric shape structures themselves, as well as their relationship to semantics, functionality, and other modalities.

All collected and processed data are publicly available at <http://2dshapesstructure.github.io/index.html>. The proposed benchmark Shape Structure Database has the following relevant features.

- A collection of over 1200 shapes has been annotated by 2861 users of various demographics resulting in 41953

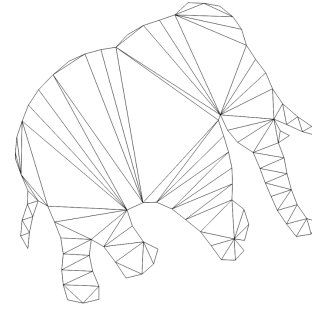


Figure 3: Delaunay triangulation for the *elephant-1* shape. Note that the Delaunay edges provide clean cuts for each salient part.

annotated shapes. For each shape, a set of at least 24 annotations are available giving a decomposition into a main shape, one or more levels of parts, and a level of shape details. The complete raw data is available.

- The input curves and shape discretizations, along with associated triangulations, are available in two formats, \*.obj and \*.json.
- All annotation results are available in a .csv file.
- Extracted shape statistics, such as the majority vote shape and a recommended number of modes for each shape class, are available.

The remainder of the paper is organized as follows: Section 2 details previous work on 2D shape decomposition, including articles, available datasets, and user-studies results; Section 3 describes the user study setup and results; and Section 4 proposes an analysis of the dataset.

## 2. Related Work

### 2.1. Shape decomposition

Decomposing 2D shapes into meaningful parts has been an active research area for decades. Existing methods include convex and near-convex parts decomposition [9, 10, 11, 12, 13, 5], methods based on outline-related factors such as curvature and cut length [4, 14], or methods based on skeletons and graphs [3, 15, 16], convolutions combined with a persistence measure [17], or optimization and probabilistic methods [13, 2]. Combining these different tools has also led to efficient shape decompositions, such as in [2].

The goal of these methods is to automatically partition a shape into independent parts, mimicking as faithfully as possible what would be done by a human, to recover partitions coherent with perception. Such partitions of shapes, however, are not structure-aware: they lack both a hierarchy of parts and an understanding of the relationship between parts. As such, these methods fail to capture the relevant importance of the identified parts. They also usually require additional shape processing such as noise removal. Figure 4 shows the difference between classical decomposition methods and the results of our user study. Whereas classical decomposition methods give a



Figure 4: Results from [10] (1st row), [9] (2nd row), [14] (3rd row) compared to majority vote results from our user study for four shapes.

partition of the shape and part neighborhoods through cell adjacencies, our user annotations partition the shape into a hierarchy that inherits rich topological and geometric information from the medial structure, as discussed in Section 1. Note also that our user annotations occasionally violate some of the guidelines generally followed in algorithms, such as decomposition into convex regions as seen in the main shape for the elephant.

Similarly, most of the work on parts hierarchy draws on the medial axis [3, 18] and lacks an accompanying shape decomposition. Consequently, we believe the shape community would benefit from results that establish a connection between parts decompositions and their hierarchies for standard shape databases such as MPEG-7. It is such a connection that our database provides.

## 2.2. Existing shape databases

Most of the shape decomposition literature from the last two decades refers to one of two 2D shape databases for testing and validation purposes: the *Kimia*<sup>1</sup> dataset (1995) containing binary images of shape silhouettes for birds, bones, brick, camels, car, children, classic cards, elephants, etc., and the *MPEG-7*<sup>2</sup> database (2000), developed as part of the MPEG-7 standards project [19], containing 70 classes of shapes given as binary images with 20 examples of each class, for a total of 1,400 shapes. Another database of hand drawn black-and-white pictures, called S&V database, is available by Snodgrass and Vanderward [20]. The ETHZ shape database gives original natural scene images and their associated shape contours [21].

## 2.3. User studies

The literature shows several user studies related to shape decomposition [3, 4, 22, 13]. For example, [13, 23] evaluate 12 user segmented shapes on 10 shapes for 20 categories of

the MPEG-7 database for use as ground truth for the proposed shape decomposition methods. None of the data from these user studies is publicly available, and none presents a full shape structure. The only user study with collected data available<sup>3</sup> to the community was conducted by De Winter and Wagemans in 2006 [6]. They provide a benchmark that contains the cuts proposed by 201 users on 88 shapes. A set of cuts correspond to a decomposition, but no information on relative importance of the parts appear here.

Some work exists for 3D as well, such as the Princeton 3D shape dataset [24] that provides a benchmark on 380 meshes for 19 categories (humans, chair, bird, hand, etc.) annotated on average by 11 humans who indicated cuts using an interactive segmentation tool developed by the authors. A survey on 3D segmentation can be found in [25].

## 3. Data collection

Our goal is to establish a ground truth for human-derived geometric shape structure, a perceptual decomposition into a hierarchy of parts in such a way that relationships between parts are retained. We test on a large number of shapes, including regular and irregular shapes, simple and complex shapes, as well as shapes with and without semantic meaning. The MPEG-7 dataset offers 1440 shapes and 70 shape categories, and we add to that a selection of artificial shapes without semantic meaning.

With the augmented MPEG-7 dataset, we conducted a crowd-sourced user-study to guarantee a large number of participants. In contrast to previous work, where users have been asked to decompose a shape into parts, we collected information about full shape structure: parts decomposition, parts hierarchy, and relationship of parts. We asked users to label the interior regions of shapes by coloring triangles in the associated Delaunay triangulation to indicate membership in one of three or more levels: main shape, parts (with possibly multiple part levels), and details. In the two following subsections, we describe the process used to collect the dataset.

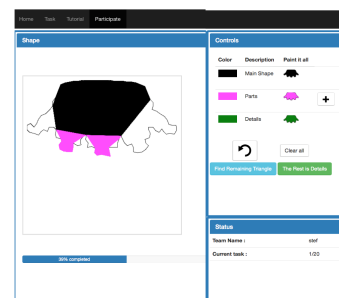


Figure 5: Screenshot of our web-based shapes annotation interface, used during the scavenger hunt. The crown shape is displayed. Similar to a digital coloring system, the users drag the mouse across the shape region to color it with the selected color. The “+” icon on the far right allows users to add additional part levels (though very few did).

<sup>1</sup>Kimia datasets: <http://yacvid.hayko.at>

<sup>2</sup>MPEG-7 dataset: <http://www.dabi.temple.edu/~shape/MPEG7/dataset.html>

<sup>3</sup>De Winter/Wagemans dataset: <http://www.gestaltrevision.be/en/resources/supplementary-material/129-de-winter-a-wagemans-brmic-2004-stimuli-sets-a-datasets>

### 3.1. Preprocessing of shapes

The experiment was conducted online. All shapes underwent the same four preparation steps. Shape outlines (closed polygonal curves) were computed by applying contour detection to the binary images. The interiors were discretized into triangles via Delaunay triangulation. Note that we limited the number of triangles to about a hundred per shape so that triangles were sufficiently large to be distinguished by the user on a screen. We accomplished this by sub-sampling the boundary curves accordingly before computing a triangulation [26]. Finally, we removed all shapes (~200 shapes) exhibiting features with triangles that became degenerate after subsampling (for example, some beetles with extremely thin antennae). 1250 shapes remained, belonging to 69 categories. We also added a few shapes with no semantic meaning to capture purely geometric decomposition.

### 3.2. The user study - a crowdsourcing experiment

**Setup.** We implemented a web-based user interface to collect user annotations describing perceptual shape structure. The interface allows users to label each triangle of the Delaunay triangulation of the shape as belonging to one of three or more categories by color coding. By default, three part levels are defined as Main Shape, Parts, Details, and associated with black, magenta, and green. The user was allowed to define additional levels of parts. Figure 5 shows our interface, which looks like a digital coloring system. Once a color has been selected, the user colors the triangles by dragging the mouse across a shape region or by clicking individual triangles which appear progressively during coloring. Additional actions provided to ease the interactions include:

- undo previous action (circle with arrow button);
- remove a triangle’s annotation (right click);
- replace a triangle color with a new one;
- initialize the hierarchy with a single color;
- finish coloring all remaining triangles with green
- find remaining triangles and zoom into them;
- zoom and pan (mouse wheel);
- follow progress indicating the ratio of colored triangles.

Each user is given 20 shapes sequentially to color, 16 randomly selected shapes and 4 (so-called) “Gold Standard” shapes used to evaluate user reliability appearing at position 1, 6, 11 and 16.

**Tutorial.** An interactive tutorial guides the user through all possible interactions with the system. The purpose of the tutorial is to teach the use of the interface by forcing the user to use every action at least once. Each step consists of either reading a short text or performing an action. Explanatory text pops up around specific widgets. The user is not told what a “good” shape annotation is nor how to decompose a shape.

**Scavenger hunt.** We launched our user study as part of the one-week, international online scavenger hunt GISHWES 2015 with tens of thousands of participants of all ages from across the world. Our study was one task among more than 200 tasks for the scavenger hunt, and consisted of annotating 20 shapes from

our shape dataset. Users were strongly encouraged, although not forced, to view the tutorial before starting the task. They were told the task has scientific purposes: “*providing shape decompositions according to your human brain will allow the researchers to know what humans perceive as main body, parts, and details.*” The task was described as: “*You will be shown a sequence of shapes and asked to color-code regions according to whether they are main body, parts, or details [...] Try to determine which regions are in which level based on visual shape alone instead of function.*” In order to help users abstracting semantics, we did four things: 1. We provided online training to the users. 2. We did not display the names of the shapes. 3. We rotated the shapes upside down. 4. We added some artificial shapes. The task description also emphasized the necessity of providing good shape decompositions in order to receive points on the task for the scavenger hunt.

**Statistics.** During four days, we collected more than 40,000 annotations from 2,861 teams participating to the hunt. Every shape has been annotated at least 24 times.

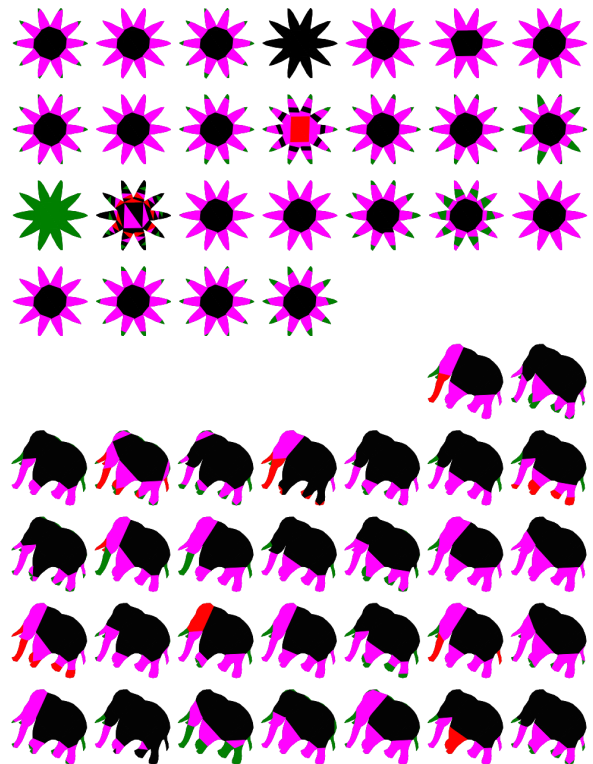


Figure 6: Two shapes with user annotations. Top: *device7-17* shape. Bottom: *elephant-1* shape. This is the complete raw data collected for these two shapes. No post-processing, such as cleaning, filtering or transformation, has been applied.

### 3.3. Results and data format

The dataset is available through a web interface. Each of the 70 categories defines a different webpage showing all shapes in the category as images, partitioned according to a majority vote after filtering, with annotations for each shape associated

to a Delaunay triangulation. For each shape, the results of all segmentations are shown. We provide also a csv file containing all annotations in the form:

#*annotation, user\_name, shape\_name, {0, 1, 2, 3}*\*

where the sequence of digits gives the color of the Delaunay triangles ordered according to the shape file. The Gold Standard shapes have been annotated by all 41,953 users. These annotations are too numerous to display, however a separate csv file is provided for each of them.

## 4. Data analysis

### 4.1. Distance between annotations

Once the annotations are collected, we wish to combine and analyze them. To measure the similarity of two annotations, we propose a simple, intuitive distance measuring whether there is a color difference between annotations of a triangle, and weighting the contribution by the area of each triangle.

Let  $s$  be a shape of area  $\mathcal{A}_s$ , partitioned in  $N$  Delaunay triangles  $t_i$ ,  $i \in 1 \dots N$ . Each triangle  $t_i$  has an area of  $\mathcal{A}_i$ . We associate each classification to an integer value (no user included more than four classes):

$$\begin{cases} 0 & \text{Main Shape} \\ 1 & \text{Parts} \\ 2 & \text{Details} \\ 3 & \text{Parts Level 2} \end{cases} \quad (1)$$

We define the annotation of a user  $f$  as the following function:

$$f : s \rightarrow \{0, 1, 2, 3\} \\ t_i \mapsto f(t_i).$$

Let  $f_1(s)$  and  $f_2(s)$  be two annotations of shape  $s$ , we define the distance  $d_1$  between these two annotations as:

$$d(f_1(s), f_2(s)) = 1 - \sum_{i=1}^n \delta(f_1(t_i), f_2(t_i)) \frac{\mathcal{A}_i}{\mathcal{A}_s}, \quad (2)$$

$$\text{where } \delta(a, b) = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b \end{cases}$$

The distance varies between 0 and 1, and by construction is 0 only if the annotations are identical, and is 1 if all triangles are of different colors. For example, the distance between the annotations of the butterflies of Figure 2 is almost 1 (1 minus the area of the antennas relative to the total area).

### 4.2. Filtering using Gold Standard shapes

In any crowdsourcing experiment, there are typically three sources of errors [27] (see Figure 7 for illustration):

- (a) occasional mistakes, that can be due to lack of attention or user exhaustion;
- (b) errors resulting from a poor understanding of the task, or from difficulties mastering the interface;

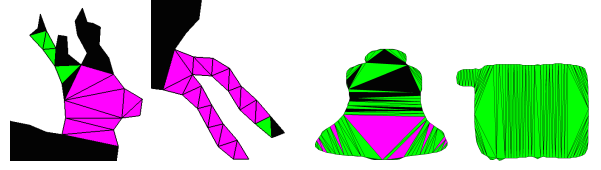


Figure 7: Left: Two user annotations on a deer shape that include small “random” mistakes of type (a). Right: unsatisfying user annotations on a bell and a cellphone that could be due to either a poor understanding of the task, type (b), or a spammer, type (c).

- (c) voluntary errors that come from users (spammers) who try to cheat the system to get the reward as quickly as possible.

When crowdsourcing a task, filtering users who do not understand the task (type b) or spam the system (type c) is required. A widely used approach is to introduce Gold Standard data, i.e. data for which we already have an expert annotation (e.g. [27]). All users have to annotate the Gold Standard (GS) shapes. These annotations are then used to measure a user’s quality by comparing her annotations to the expert annotations. We call the set of experts  $E$ . The authors and their colleagues provided 7 expert annotations for the four GS shapes. See Figure 8 for the GS shapes and some of their expert annotations.

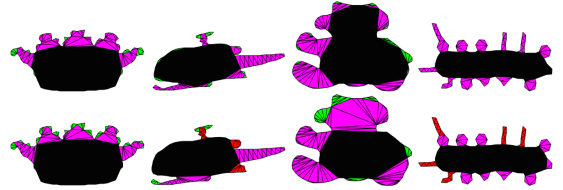


Figure 8: Annotations from two (of our 7) experts of the 4 Gold Standard shapes (crown-10, chopper-12, teddy-12 and an additional image added with no semantic information). Note again human variability in annotation, even among the experts.

For each user, and for each GS shape, we compute the distance to each expert annotation of the same shape. We keep the minimum value of these distances to experts: since there are multiple possible interpretations of a single shape, this step allows us to detect which expert has the closest interpretation of the shape to the user. Next, we average the minimum expert distances for all four GS shapes. Finally, a quality score  $Q(u)$  for user  $u$  is defined as 1 minus this average distance:

$$Q(u) = 1 - \frac{1}{|\text{GS}|} \sum_{s \in \text{GS}} \min_{e \in E} d(f_u(s), f_e(s)) \quad (3)$$

Figure 9 shows the annotations of users who obtain the best and worst scores with respect to GS shapes. Typically, around 20% of the annotations for each shape are filtered out.

### 4.3. Majority vote

To obtain a single representative annotation for each class, we apply a simple majority vote to eliminate occasional mistakes (type a). Majority vote consists in choosing the label for

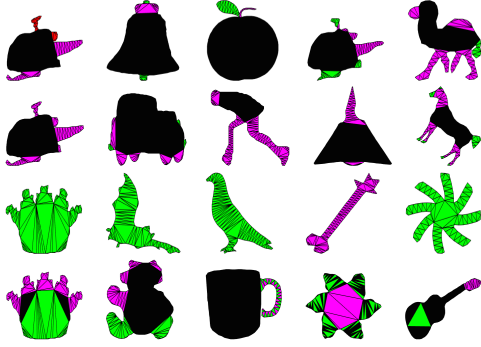


Figure 9: Each row shows annotations for a single user: the top two rows correspond to the two users with the best quality scores, the bottom rows to the two users with the worst scores. They appear clearly as a spammer (third row) and as someone that did not understand the task (last row).

a triangle to be the color appearing the most in the collected annotations after filtering. See, for example, Figure 11-left.

This method works well for shapes which exhibit a consensus for a valid hierarchy. However, for shapes with multiple contradictory annotations, like the butterfly in Figure 2 (see also the entire dataset), a majority vote may give results far from any annotations. In such a case the majority is neither representative nor robust, see Figure 11-left. There has been work focusing on how to determine one resulting annotation out of multiple potentially noisy users annotations. Dawid and Skene [28], in their pioneering work, address this issue using a version of the Expectation-Maximization algorithm that simultaneously evaluates the annotators and builds the most likely single representative annotation from all users.

For our purposes, however, obtaining a single representative annotation, via majority vote or Expectation - Maximization, is often not sufficient. Indeed, the user study reveals the existence of several valid, meaningful annotations, as seen Figures 2 and 11-left. We therefore propose in the next section to use a method able to determine multi-modal behavior of the data, thereby identifying all perceptually meaningful annotations.

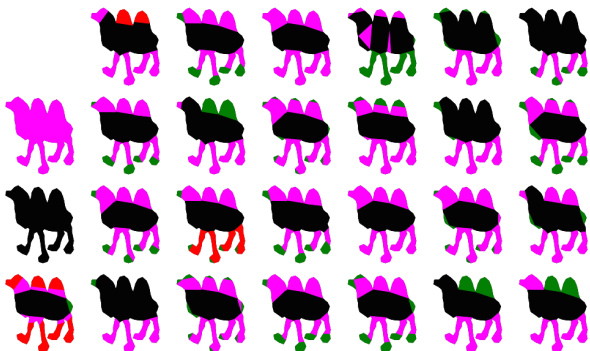


Figure 10: Annotations of the shape *camel18*.

#### 4.4. Identifying shapes with multiple structures

Unlike previous work, and other comparisons of decomposition to human input [13, 22, 3], we provide further analysis

| percentage of shapes and their # of modes |         |         |         |           |
|---|---------|---------|---------|-----------|
| 1 (unimodal)                              | 2 modes | 3 modes | 4 modes | 5-8 modes |
| 18%                                       | 70%     | 8%      | 2%      | 2%        |

Table 1: The percentage of shapes modeled well with the given number of modes: a large majority of shape have two different, meaningful representative annotations given by the two modes. Only Gold Standard shapes that have more than 2000 annotations generated four or more modes.

on the user segmentations to take into account that a shape may have more than one good structural interpretation such as those shown in Figure 2 for the butterfly.

Given the different annotations of a shape  $s$ , we construct the affinity matrix  $A = [a_{kl}]_{k,l=1\dots n_i}$  of the set of  $n_i$  annotations  $\{f_i\}_{i=1\dots n_i}$ . Using the distance proposed in Equation 3, we set:

$$a_{kl} = d(f_k(s), f_l(s)).$$

The affinity matrix provides the input for a spectral clustering algorithm [29]. By reordering annotations and applying  $k$ -means clustering on the data, the matrix can be given a relative block structure corresponding to different clusters. The number of clusters  $k$  is determined by minimizing the ratio between the intra-cluster and the inter-cluster distance based on the criterion proposed by Calinski et al. [30]. We then discard clusters with fewer than three annotations.

Table 1 shows that for most shapes (70%) the annotations are best clustered into two separate valid clusters. As an example, the annotations of Figure 10 are found to have two representative clusters (Figure 11-right), that highlight different, meaningful ways to partition the shape and to define a structure. The majority vote is also coherent but less informative (Figure 11-left).



Figure 11: "Mean" annotations for the *camel18* annotations shown in Figure 10. Left: the result of majority vote after filtering. Right: the *camel18* has two modes corresponding to these two annotations.

## 5. Contributions and Discussion

We have defined a geometric shape structure as a shape representation that includes a parts decomposition, parts hierarchy, and relationship between parts. Such a structure allows for an integrated approach to shape comparison, articulation, animation, and creation. We have demonstrated the absence of geometric shape structure analysis in the literature. As a first step toward filling that gap, we contribute a ground truth shape structure database generated by a massive user study to collect shape

annotations on the MPEG-7 database. These annotations include a parts decomposition, a parts hierarchy, and a retrievable relationship between parts.

Most existing methods, particularly for segmentation, try to mimic human behavior, but then compare to each other instead of human-generated data. Figures 4 and 6 clearly demonstrate the need for user data in evaluation, immediately offering two insights into shape structure. First, many shapes admit more than one valid underlying structure. The users who segmented the shape *elephant-1* in Figure 6 can be divided into two classes: those who think the head is part of the main shape, and those who do not. Automatic algorithms would benefit from taking this into account. Second, we observe that the cuts chosen by the users do not always define convex regions. The segmented elephants of Figure 6 once again provide a good illustration, showing that convexity is not a universal principle for part decomposition.

The dataset also allows for a range of potential applications. As thousands of users interpreted and decomposed shapes, the dataset could be used to train machine learning algorithms for shape segmentation and hierarchy, as well as to evaluate shape structure algorithms. We further believe the results of the user study offer some tantalizing questions about human shape understanding, and we are eager to see what the research community can extract from our data.

**Acknowledgements:** The authors gratefully acknowledge the support of Marie-Paule Cani and advanced grant no. 291184 EXPRESSIVE from the ERC (ERC-2011- ADG\_20110209), support from NSF award IIS-0954256, and from CIMI, program ANR-11-LABX-0040-CIMI. We also thank the 2015 Gish superfans for beta-testing and the Gishers themselves for sharing their shape perception abilities.

## References

- [1] N. J. Mitra, M. Wand, H. Zhang, D. Cohen-Or, V. Kim, Q.-X. Huang, Structure-aware shape processing, in: ACM SIGGRAPH Courses, 2014.
- [2] C. Wang, Z. Lai, Shape decomposition and classification by searching optimal part pruning sequence, *Pattern Recognition* 54 (2016) 206–217.
- [3] K. Siddiqi, B. B. Kimia, Parts of visual form: Computational aspects, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (3) (1995) 239–251.
- [4] M. Singh, G. D. Seyranian, D. D. Hoffman, Parsing silhouettes: The short-cut rule, *Perception & Psychophysics* 61 (4) (1999) 636–660.
- [5] G. Liu, Z. Xi, J.-M. Lien, Dual-space decomposition of 2D complex shapes, in: *IEEE Conf. on Computer Vision and Pattern Recognition*, 2014, pp. 4154–4161.
- [6] J. De Winter, J. Wagemans, Segmentation of object outlines into parts: A large-scale integrative study, *Cognition* 99 (3) (2006) 275–325.
- [7] B. S. C. Firestone, "please tap the shape, anywhere you like": Shape skeletons in human vision revealed by an exceedingly simple measure, *Psychology Science* 25 (2014) 377–386.
- [8] J. Giesen, B. Miklos, M. Pauly, C. Wormser, The scale axis transform, in: *Proc. of the 25th Symposium on Computational Geometry, SCG '09*, ACM, 2009, pp. 106–115.
- [9] J.-M. Lien, N. M. Amato, Approximate convex decomposition of polygons, *Computational Geometry* 35 (12) (2006) 100–123.
- [10] H. Liu, W. Liu, L. J. Latecki, Convex shape decomposition, in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
- [11] L. L. Walker, J. Malik, Can convexity explain how humans segment objects into parts?, *Journal of Vision* 3 (9) (2003) 503–503.
- [12] Z. Ren, J. Yuan, C. Li, W. Liu, Minimum near-convex decomposition for robust shape representation, in: *IEEE International Conference on Computer Vision (ICCV)*, IEEE, 2011, pp. 303–310.
- [13] T. Jiang, Z. Dong, C. Ma, Y. Wang, Toward perception-based shape decomposition, in: *Computer Vision, ACCV 2012*, Springer, pp. 188–201.
- [14] L. Luo, C. Shen, X. Liu, C. Zhang, A computational model of the short-cut rule for 2D shape decomposition, *IEEE Transactions on Image Processing* 24 (1) (2015) 273–283.
- [15] X. Mi, D. DeCarlo, Separating parts from 2D shapes using reliability, in: *IEEE International Conference on Computer Vision*, 2007, pp. 1–8.
- [16] D. Macrini, K. Siddiqi, S. Dickinson, From skeletons to bone graphs: Medial abstraction for object recognition, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1–8.
- [17] Y. Lu, J.-M. Lien, M. Ghosh, N. Amato,  $\alpha$ -decomposition of polygons, *Computers & Graphics (SMI'12)* 36 (5) (2012) 466–476.
- [18] B. B. Kimia, A. R. Tannenbaum, S. W. Zucker, Shapes, shocks, and deformations i: The components of two-dimensional shape and the reaction-diffusion space, *Int. Journal of Computer Vision* 15 (1994) 189–224.
- [19] L. J. Latecki, R. Lakamper, T. Eckhardt, Shape descriptors for non-rigid shapes with a single closed contour, in: *IEEE Conference on Computer Vision and Pattern Recognition'00*, 2000, pp. 424–429.
- [20] J. G. Snodgrass, M. Vanderwart, A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity, *Journal of Experimental Psychology: Human Learning and Memory* 6 (2) (1980) 174–215.
- [21] T. Quack, V. Ferrari, B. Leibe, L. J. V. Gool, Efficient mining of frequent and distinctive feature configurations, in: *IEEE International Conference on Computer Vision, ICCV*, 2007, pp. 1–8.
- [22] E. Cohen, M. Singh, Geometric determinants of shape segmentation: Tests using segment identification, *Vision Research* 47 (22).
- [23] C. Ma, Z. Dong, T. Jiang, Y. Wang, W. Gao, A method of perceptual-based shape decomposition, in: *IEEE International Conference on Computer Vision (ICCV) 2013*, 2013, pp. 873–880.
- [24] X. Chen, A. Golovinskiy, T. Funkhouser, A benchmark for 3D mesh segmentation, *ACM Transactions on Graphics* 28 (3) (2009) 73:1–73:12.
- [25] A. Shamir, A survey on mesh segmentation techniques, *Computer Graphics Forum* 27 (6) (2008) 1539–1556.
- [26] N. Amenta, M. Bern, M. Kamvyselis, A new Voronoi-based surface reconstruction algorithm, in: *Proceedings of the 25th annual conference on Computer graphics and interactive techniques, SIGGRAPH '98*, ACM, New York, NY, USA, 1998, pp. 415–421.
- [27] D. Oleson, A. Sorokin, G. P. Laughlin, V. Hester, J. Le, L. Biewald, Programmatic gold: Targeted and scalable quality assurance in crowdsourcing, *Human computation* 11 (11).
- [28] A. P. Dawid, A. M. Skene, Maximum likelihood estimation of observer error-rates using the em algorithm, *Applied statistics* 28 (1) (1979) 20–28.
- [29] A. Ng, M. Jordan, Y. Weiss, On spectral clustering: Analysis and an algorithm, *Advances in neural information processing systems* 2 (2002) 849–856.
- [30] T. Caliński, J. Harabasz, A dendrite method for cluster analysis, *Communications in Statistics-theory and Methods* 3 (1) (1974) 1–27.