ISSN 1213-6972 (print)
ISSN 1213-6964 (online)
Journal of WSCG
http://www.wscg.eu
Vol.31, No-1-2, 2023

Visual Exploration of Repetitive Patterns on Ancient Peruvian Pottery

Stefan Lengauer[1], Lin Shao[1,2], Magdalena Mayerhofer[1], Reinhold Preiner[1], Stephan Karl[3], Elisabeth Trinkl[3], Ivan Sipiran[4], Benjamin Bustos[4], and Tobias Schreck[1]

[1]Graz University of Technology, Institute CGV, Inffeldgasse 16c, 8010 Graz, Austria
[2]Fraunhofer Austria Center for Data Driven Design, Inffeldgasse 16c, 8010 Graz, Austria
[3]University of Graz, Institute of Classics, Universitätsplatz 3/II, 8010 Graz, Austria
[4]University of Chile, Dept. of Computer Science, Beauchef 851, Santiago, Chile

ABSTRACT

The analysis and understanding of artifact properties and their relationships is a key objective in the archaeological analysis of cultural heritage objects. There are many aspects of concern, including both shape properties of the objects as a whole and appearances stemming from paintings and ornamentations on the object surfaces. To date, experts consider those mostly holistically on a per-object basis. We present an approach for the interactive visual exploration and correlation of shape- and ornament-based properties of a large collection of ancient vessels. Our approach allows us to group objects by said properties, and to relate them in side-by-side and bipartite graph displays. To this end, we define an encompassing set of feature descriptors, which are leveraged to cluster the objects by user-selected properties. Case studies show that a comparative overview of all objects effectively supports the discovery of interesting co-occurrences of shapes and ornaments. This way, our tool opens new possibilities for the domain analysis of cultural heritage object collections by data-driven visual exploration.

Keywords

Visual exploration, Pattern analysis, Pattern descriptors

1 INTRODUCTION

The analysis of ancient pottery is an essential task for the understanding of ancient cultures and lifestyles. Of particular interest are the lavish surface decorations exhibited by the majority of pottery artifacts. These decorations – so called *vase paintings* – comprise both repetitive ornaments and motifs depicting mythological scenes. They provide important information for an artifact's attribution to a specific epoch, culture, workshop of even painter [ES16].

The concurrent exploration a large collection allows us to reveal clusters of objects with common traits, if these objects are appropriately arranged in a structured manner, based on relevant properties. Potential properties comprise intrinsic traits like shape, material, capacity and such, such as well as derived traits like culture, dating, etc. However, also more complex traits describing an object's vase painting, e.g., variability, distribution and positioning on the surface or colorization, can be extracted automatically using customized data processing techniques.

Within the scope of this paper, we focus on two of the most important properties: (i) object shape and (ii) vase painting. We present a novel visualization method that groups similar elements along these properties and presents the resulting groups in separate cluster visualizations (Fig. 1). These are connected in a bipartite graph, revealing relationships and co-occurrences between different shapes and paintings. The thickness of graph edges between shape and painting clusters reveal inter-cluster correlations in the collection. This visualization is the core component of an overarching interactive exploration system, supporting various degrees of visual granularity – from a broad overview down to closeup. To this end, we conceptualize and implement tailored views for the individual levels with customized object previews for different properties (e.g., positional glyphs in Fig. 1, bottom, indicating the positioning of patterns on the object surface).

The contribution of our paper is the interactive visualization concept which we evaluate using prototype implementation together with a real-world dataset of ancient pottery objects. Moreover, we present a novel feature descriptor which is able to capture the arrangement of repetitive patterns (e.g., regarding regularity) in a quantifiable manner.

In the following, we report related visualization techniques (Sec. 2), before defining the domain analysis task (Sec. 3) relevant for experts. In Sec. 4 we discuss the proposed concept in detail, before we present the major insights gained with this tool in Sec. 5. We conclude the paper with a discussion (Sec. 6) including feedback from archaeologists.

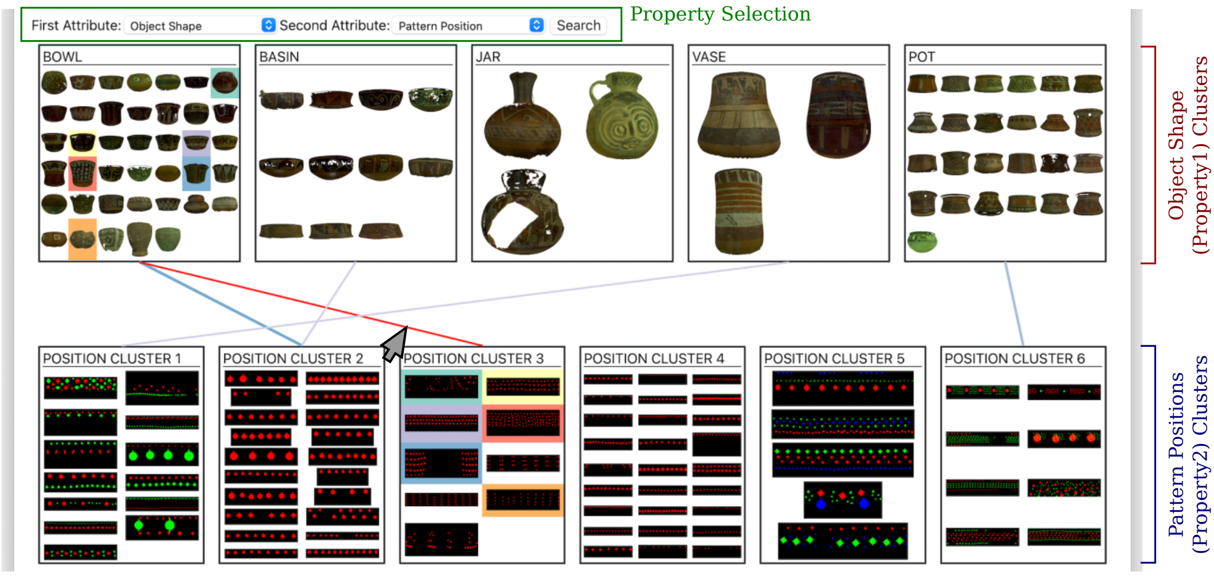


Figure 1: A collection is explored along two selected properties, i.e., object shape and pattern position, which are clustered individually (top and bottom row). Visual links in-between reveal correlations inter-cluster correlations.

2 RELATED WORK

One goal of Information Visualization (InfoVis) techniques is to convert abstract information into visual representations, and thus gain knowledge about internal structures of a dataset. In cases where corresponding data elements have inherent relationships among each other, graph-based visualization techniques are commonly used. A graph visualization is often encoded by a set of nodes and edges and allows the visual analysis of structures and grouping of nodes and/or edges (Compound Graph Visualization) [HSS15]. Further application areas and examples of graph-based visualizations are given in surveys [vLKS*11; TKE12].

Visualization techniques have also become important tools for the research of cultural heritage (CH) objects. Although many scientific visualization techniques in the CH domain focus on the realistic rendering of 3D objects, there is a growing number of interactive visual systems for analyzing CH data [WFS*19]. In recent works, systems and interface designs were introduced that utilize InfoVis designs and Visual Analytics approaches for representing multidimensional and temporal information of CH collections. For instance, the PolyCube framework by Windhager *et al.* [WSL*20] uses space-time cube representations to visualize multidimensional, time-dependent properties of collections. By connecting different visualization techniques, like map, set, and network visualizations, they revealed spatial, categorical, and relational collection aspects. Simon *et al.* [SIBdS16] introduce Peripleo, an open source tool to explore the geographic, temporal and thematic composition of distributed digital collections. Lengauer *et al.* [LKK*20] present an interactive visual exploration system for artifact collections, dubbed Linked Views Visual Exploration System (LVVES), that supports task-oriented analysis and exploration along temporal, spatial, and shape modalities.

To visualize sets of images, the overall layout for arranging the images is often a crucial part. In this regard, Brivio *et al.* [BTC10] propose a Voronoi-based layout to visualize photographic campaigns in CH. In Glinka *et al.* [GPD17] the authors show the potential of details-on-demand techniques for the exploration of large CH collections including images, keywords and textual data. They employ a zoomable timeline visualization to link a “distant-viewing” and “close-viewing” mode for the exploration.

In Mauri *et al.* [MPCC13] a graph of actors and projects is used to explore collaborations between architects, while Tortora *et al.* [DPT*12] use graph views to support archaeologist in finding new correlations from ontology-driven metadata. An extensive survey on visualization techniques for CH collection data is given in [WFS*19].

As opposed to existing approaches, our proposed design should be particularly useful in revealing correlations between different object properties. To this end, we use a bipartite graph layout, connecting different similarity clusterings – a design which has, to our knowledge, not been used before. Moreover, we provide two details-on-demand feature: a side-by-side view for comparing two clusters from different traits, and a closeup view for additional information on the lowest level of visual granularity.

3 DOMAIN ANALYSIS TASKS

The prototype system is designed to be used by domain experts having well-defined research questions. W.r.t.

the analysis of repetitive patterns on ancient pottery those include (but are not limited to): (1) How regular are the ornament patterns within a pattern class? (2) Do similar ornament patterns exhibit similar spatial arrangements? (3) Are ornament patterns correlated with the vessel shape? (4) Are properties of ornament pattern or other shape properties generally correlated with each other?

To date, such questions are mostly answered using pairwise visual comparisons of artifacts. Based on this established workflow, we define the following domain analysis tasks, for which we support a domain expert with customized visualizations: (**T1**) Select two properties and discover inter-trait correlations, (**T2**) Show the detailed correlations between two selected clusters, and (**T3**) Show the properties of a single record.

4 CONCEPT

In the following, our proposed design is discussed in depth. After a broad overview of the idea (Sec. 4.1), we present the dataset used in our experiments (Sec. 4.2). In Sec. 4.3, we give a detailed formal description of our descriptors designed for describing pattern arrangements, while Sec. 4.4 discusses the employed shape features for describing pattern shapes. Sec. 4.5 concludes the section with a description of the applied clustering as well as the similarity computation between clusters.

4.1 Overview

As the starting point for the visual exploration process, we provide the user with a bipartite cluster view, for which the user has to select two object properties via respective drop-down dialogues (Fig. 1, green). The selected properties are used to cluster the objects of the given collection separately, e.g., by 'vessel shape' and 'pattern positioning' (Fig. 1, red and blue). Based on the visual links between the two sets of clusters, a user can derive the presence and strength of possible correlations, e.g., between BOWL shapes and POSITION CLUSTER 2 and 3 w.r.t. Fig. 1, and can further investigate which objects of a pair of clusters is responsible for a correlation by hovering the mouse over the respective inter-cluster link. Clicking on such a link switches the visualization to a side-by-side view of the respective clusters (Fig. 2) and clicking an item in any of the clusters switches to a closeup view (Fig. 3) of the object in question. Return buttons allow one to undo such a transition and allow for a continuous exploration process. Design details on these different views and information on how they support the different domain analysis tasks (**T1–T3**) are given in Sec. 4.6.

4.2 Dataset: Peruvian vessels

The dataset we use for our experiments stems from the 2021 SHREC track on "Retrieval of cultural heritage objects" by Sipiran *et al.* [SLL*21], containing almost 1,000 3D models of ancient pottery artifacts. The real artifacts are kept in the Josefina Ramos de Cox museum in Lima, Perú, where they were digitized as part of a research project[1]. The collection comprises objects from several pre-Columbian cultures, like Chancay, Lurin or Nazca, featuring varied geometry and surface decoration. In a later documentation effort, the boundaries of ornament elements were annotated and all occurring surface patterns were grouped for a subset of the collection, exhibiting well preserved and lavish vase paintings. This annotation by Lengauer *et al.* [LSP*21] is publicly available[2] and contains detailed outlines of all surface patterns, together with pattern-wise properties like orientation, scale, position on the surface, and n-foldness, giving an encompassing data basis for an attribute-driven exploration. In total, the dataset comprises 2,529 *pattern entities* from 82 textured models, which are grouped into 102 distinct similarity classes (referred to by *pattern archetypes*).

The dataset also comprises a varied collection of intrinsic and derived traits, such as vessel shape, pattern variability, colorization, etc., which we need to describe quantitatively in order to cluster the objects by them. The properties are defined at three different levels of detail: (1) *Per-object* properties were provided by experts via a categorization of the vessel shape (i.e., 'bowl', 'basin', 'jar', 'vase' and 'pot'). Surface colorization is also described on a per-object level through the computation of color histograms. (2) *Per-archetype* properties comprise several carefully designed attributes describing the distribution of entities across the surface and the relations among themselves (Sec. 4.3). On a (3) *per-entity* level, we use an abstract description of a pattern entity's shape through established shape features (Sec. 4.4).

4.3 Quantifying Pattern Arrangements

We design a set of custom properties pertaining to the distribution, regularity, overall variability and other important traits of a pattern archetype. All properties are given as scalar, normalized to the range $[0,1)$, so that they can be combined into a feature vector. Specifically, we define and compute the following measures:

Occurrence Frequency. This value describes a pattern archetype's quantity of entities and is given by

$$\tilde{n}_p = \left(\frac{n_p - n_{p_{min}}}{n_p} \right)^{\alpha}, \qquad (1)$$

[1] Project 02-2018-FONDECYT-BM-IADT-AV (Concytec-Perú): Restoration and conservation of archeological pieces using deep learning and convolutional auto-encoder on graphs.

[2] `https://datasets.cgv.tugraz.at/pattern-benchmark/` [Accessed 2023-04-26]

with n_p as the absolute number of pattern entities, $n_{p_{min}} = 2$ as the minimal number of entities for an archetype, and $\alpha = 3$ denoting an empirically determined non-linear scaling factor. Note that $\tilde{n}_p$ asymptotically approaches 1 for $n_p \to \infty$ as there is no theoretical limit for the number of entities.

Fold Symmetry. The patterns' n-fold symmetry (the number of symmetry planes) is already provided by the annotated dataset and is, similar to the **Occurrence Frequency**, normalized to $[0,1)$.

Scale Variability. This property measures the variation of a pattern's size among its entities. To this end, a Normal distribution $\mathcal{N}(\mu, \sigma^2)$ is fitted to the entity scales $S = \{s_i\}_{i \in I}$, with $I = [1\,..\,n_p]$ as the index set of a pattern archetype's entities. S is also provided by the given annotation. Similar to the **Occurrence Frequency**, the variance σ^2 is non-linearly normalized to $[0,1)$ by $\tilde{n}_\sigma = \sigma^2/(\sigma^2+1)$.

Regularity. This property aims to capture the amount of 'structure' exhibited by the patterns' distribution on the surface. More specifically, this metric ranks the distribution from completely random (i.e., no perceivable structure at all, like in Fig. 3) to perfectly aligned in a grid-like manner (e.g., green pattern in Fig. 4, top). To this end, we detect the presence of rows and column arrangements in the set of unordered pattern entities. They are determined based solely on the patterns' center points $C = \{\langle x_i, y_i \rangle\}_{i \in I} \subset \mathbb{R}^2$. Rows and columns are obtained independently by fitting *Gaussian Mixture Models* (GMM) with an optimal number n of components to the x and y components of C, respectively. n is determined via the *Bayesian Information Criterion* (BIC) [BK10], measuring how well a particular GMM models the given data. To this end, we evaluate the BIC for $n \in [1\,..\,\lfloor |C|/n_{min} \rfloor]$, with $n_{min} = 3$ as we require a row or column to feature at least three pattern entities (fewer elements do not reflect a domain expert's notion of rows or columns). If the BIC exceeds an empirically defined threshold, it is assumed no rows or columns are present in the pattern layout.

We assume that the regularity of a distribution increases with its number of rows and columns (normalized quantity $\tilde{m}$) and decreases with the pattern entities offsets from their respective rows and columns (normalized error $\tilde{\varepsilon}$). Based on these assumptions, we define the three tiers of regularity of a distribution: *(i)* 'has rows and columns', *(ii)* 'has either rows or columns', and *(iii)* 'has no identifiable arrangement'. Representatives of these tiers are normalized within the intervals $[0, \frac{1}{3})$, $[\frac{1}{3}, \frac{2}{3})$ and $[\frac{2}{3}, 1)$ respectively. Let $Y = \{y_k\}_{k \in K} \subset \mathbb{R}_+$ be the y-positions of rows K and $X = \{x_l\}_{l \in L} \subset \mathbb{R}_+$ be the x-positions of columns L.

The normalized regularity $\tilde{r}$ is given by

$$\tilde{r} = \frac{1}{3} \begin{cases} \frac{1}{1+\frac{\tilde{\varepsilon}}{\tilde{m}}} + 2 & |X| > 0 \text{ and } |Y| > 0 \\ \frac{1}{1+\frac{\tilde{\varepsilon}}{\tilde{m}}} + 1 & |X| > 0 \text{ or } |Y| > 0 \\ \frac{n_p}{n_p+10} & \text{otherwise.} \end{cases} \tag{2}$$

The normalized quantity $\tilde{m}$ we define as

$$\tilde{m} = \begin{cases} \frac{|X|+|Y|}{|X|+|Y|+10} & |X| > 0 \text{ and } |Y| > 0 \\ \frac{|X|}{|X|+1} & |X| > 0 \\ \frac{|Y|}{|Y|+1} & \text{otherwise,} \end{cases} \tag{3}$$

and the normalized error $\tilde{\varepsilon}$ as

$$\tilde{\varepsilon} = \frac{1}{|I|} \sum_{i \in I} \begin{cases} \left\| \langle x_i, y_i \rangle - \langle \hat{x}_i, \hat{y} \rangle \right\| & |X| > 0 \text{ and } |Y| > 0 \\ \|x_i - \hat{x}_i\| & |X| > 0 \\ \|x_i - \hat{y}_i\| & \text{otherwise,} \end{cases} \tag{4}$$

with $\hat{x}_i = \arg\min_{x_l \in X} \|x_i - x_l\|$ and $\hat{y}_i = \arg\min_{y_k \in Y} \|y_i - y_k\|$ as a pattern entities offset from its designated column and row respectively.

Alternatingness. Pattern entities exhibit an orientation – or for an n-fold symmetry larger than one, even multiple equivalent orientations – provided by the annotation. With the 'alternatingness' property, we aim to quantify how a pattern's orientation deviates on average from its predecessor if an archetype's pattern entities exhibit any kind of sequence. The rationale behind this is that patterns which appear in an alternating fashion (e.g., always rotated by 180 degrees from one entity to the next) stand out from those which have a uniform orientation or those which have a completely random orientation.

Let $I_k, I_l \subset I$ denote the index sets of patterns belonging to the k-th row and l-th column, respectively. For a pattern archetype with n-fold symmetry n_f and $O = \{o_i\}_{i \in I} \subset \mathbb{R}_+$ as the orientations of its pattern entities, we define the normalized alternatingness as

$$\tilde{o} = \begin{cases} 0 & \substack{n_f = \infty \\ \text{or } |X|+|Y|=0} \\ \frac{1}{\Delta o_{max}} \left(\sum_{k \in K} \frac{\Delta_k}{|I_k|} + \sum_{l \in L} \frac{\Delta_l}{|I_l|} \right) & \text{otherwise.} \end{cases} \tag{5}$$

Here, $\Delta o = 2\pi/n_f$ denotes the maximum orientation difference between consecutive patterns with $\Delta o_{max} = \Delta o/2$. The row-wise and column-wise differences are given by $\Delta_k = \sum_{j \in [1..|I_k|]} \angle_{min}(o_{I_k[j]}, o_{I_k[j+1 \mod |I_k|]})$ and $\Delta_l = \sum_{j \in [1..|I_l|]} \angle_{min}(o_{I_l[j]}, o_{I_l[j+1 \mod |I_l|]})$, respectively, where $\angle_{min}(a,b) = \min_{i,j \in [1..n_f]} \{|a + i\Delta o - b + j\Delta o|\}$ defines the minimum angle between two ambiguous pattern orientations

The attributes for the given dataset, as well as the source code used to extract them from the annotated dataset, is available on the website hosting the pattern annotations.

4.4 Quantifying Pattern Entities' Shapes

A description of a pattern's shape is computed from the polygon describing its silhouette. To obtain a fixed-length numerical representation of this input, we employ the *Shape Context* feature descriptor by Belongie *et al.* [BMP02], as it is invariant w.r.t. all affine transformations. The underlying idea of this approach is to, first of all, extract a small sample of a contour with a roughly uniform sampling. For each of these sample points, a histogram describing the directivity and distance to all other points is computed. This description of a point by means of a histogram allows to determine the similarity between two points using the χ^2 metric. The similarity of two input shapes with the same number of sample points can then be inferred from the assignment costs of assigning all point pairs in an optimal fashion. This well-established minimization problem is referred to as square assignment problem and can be solved using, e.g., the Hungarian method [PS98]. As this kind of assignment is computationally expensive, we take a very small sample size of 20 sample points, which however is sufficient for the task at hand.

4.5 Clustering

Different clustering algorithms are applied, depending on the type of property (categorical, abstract) and level of detail (per-object, per-archetype, per-entity). For the object's shape class and the fold symmetry, no clustering algorithm is necessary as these properties are already grouped into five and four classes, respectively. For the pattern shape (Sec. 4.4) we employ a hierarchical clustering, since it has the advantage that we can provide our own distance function, which is necessary for our used feature descriptor. For all other properties (Sec. 4.3) Lloyd's K-means clustering [Llo82] is used to obtain six similarity clusters.

4.5.1 Inter-cluster Similarities

For the bipartite cluster view (Fig. 1) we also require a measure of the similarity between clusters obtained from different properties, since we want to display how strong the selection of objects between cluster-pairs varies. For the similarity between clusters, which are given as a set of objects we adopt the well known Jaccard index [Jac01], such that a small cluster with a high overlap with a large cluster still has a high similarity, to account for unevenly sized clusters. Let c_a, c_b be two clusters (sets of objects) obtained by clustering along property a and b, respectively. We define the similarity between c_a and c_b by

$$sim(c_a, c_b) = \frac{1}{2}\left(\frac{c_a \cap c_b}{c_a} + \frac{c_a \cap c_b}{c_b}\right). \qquad (6)$$

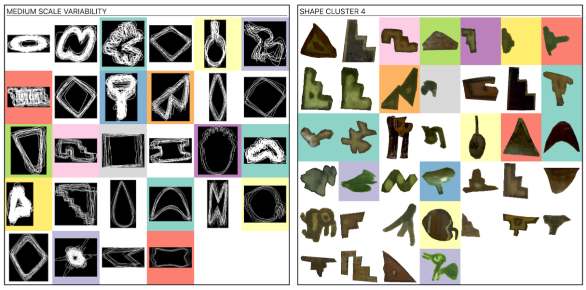


Figure 2: The side-by-side cluster view shows two clusters obtained by clustering along different object traits. The highlighting in different colors marks objects which appear in both clusters, with the same color indicating common objects.

4.6 Visual Design

Our prototype visualization system comprises three appropriate views, supporting different levels of visual granularity as well as dedicated visualizations for some of the pattern attributes. That is, the views, in descending order of visual granularity, are the following.

Bipartite Clusters Graph. This view (Fig. 1), which is the centerpiece of our exploration system, supports the discovery of correlations between traits. Hence, it is tailored to fulfill the requirements of task **T1**. From two drop-down menus at the top, users are able to select the two properties they want to compare. Based on this selection, two groups of clusters are presented in a row-wise manner. The clusters are displayed as containers, framing their belonging objects which are visualized (depending on the selected attribute) as glyphs, which are arranged in a grid-like layout. Depending on the number of objects within the cluster, the glyphs are scaled such that the space within the container is optimally used. Note, however, that the sequence in which the objects appear within a cluster is not deterministic in our current implementation. Future work might include a sorting according to inner-class similarity or other attributes. Containers also bear a label, which is a class name, e.g., 'bowl', in the case of categorical attributes like vessel shape, or an abstract term, e.g., 'position cluster n', for derived properties. In-between the two rows of clusters, we display visual links whose color saturation and thickness are relative to the cluster similarities (Sec. 4.5.1).

Additional information regarding cluster similarity is revealed through interaction. More specifically, upon hovering over a link, all objects (common across the clusters connected by the link) are highlighted. Since the appearance of the object can differ, pair-wise affiliations are established through a qualitative color coding [Bre]. From this view, a user can also transit to the side-by-side clusters view

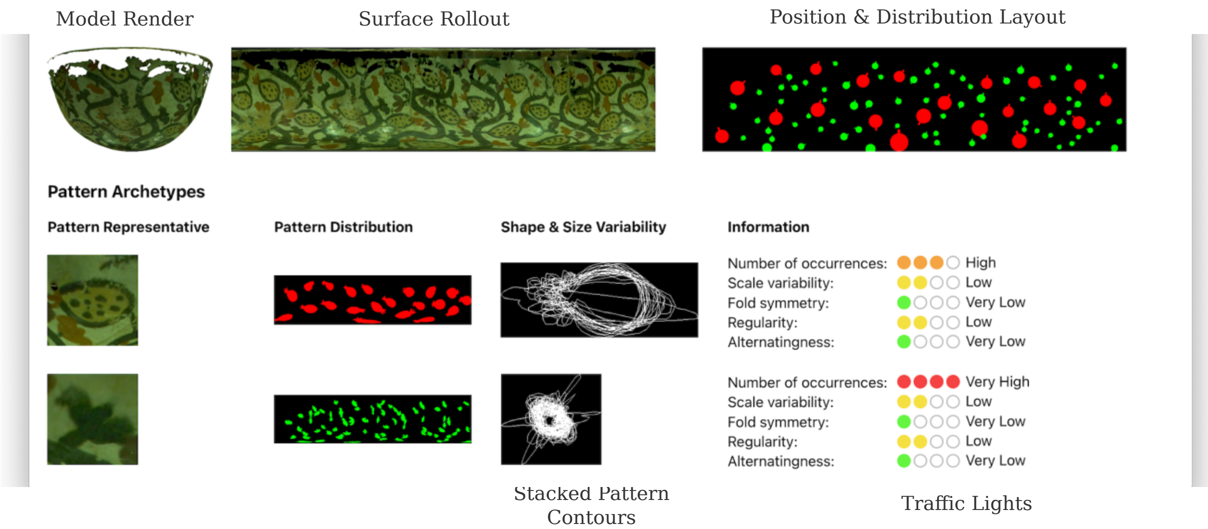


Figure 3: Closeup view of a bowl object with the various object and pattern attributes being visualized with our custom designs.

by clicking a link, or to the closeup view by clicking any of the objects within a container.

Side-by-side Clusters. The side-by-side cluster view (Fig. 2), addressing task **T2**, shows two clusters from different traits side-by-side. The container representation from the bipartite cluster graph is reused and the connection between the items is similarly established through the same color coding. From this view, a user can either go back to the bipartite clusters graph or move forward to the closeup view by clicking one of the items.

Closeup. The lowest level of visual granularity, representing just a single object, is given by the closeup view (Fig. 3). This view illustrates all the information and properties, with their respective visualizations, available for a specific object as required by task **T3**. From here, a user can return to the side-by-side cluster view or the bipartite clusters graph.

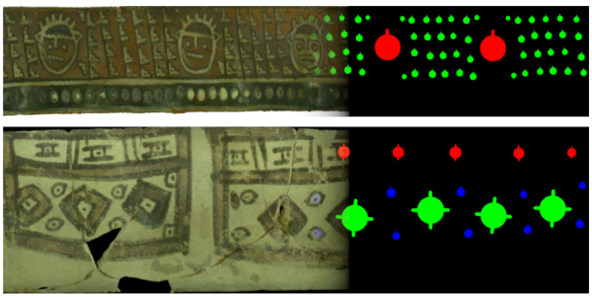


Figure 4: The projected model surfaces of two objects (left) blending over into the pattern position & distribution layout (right).

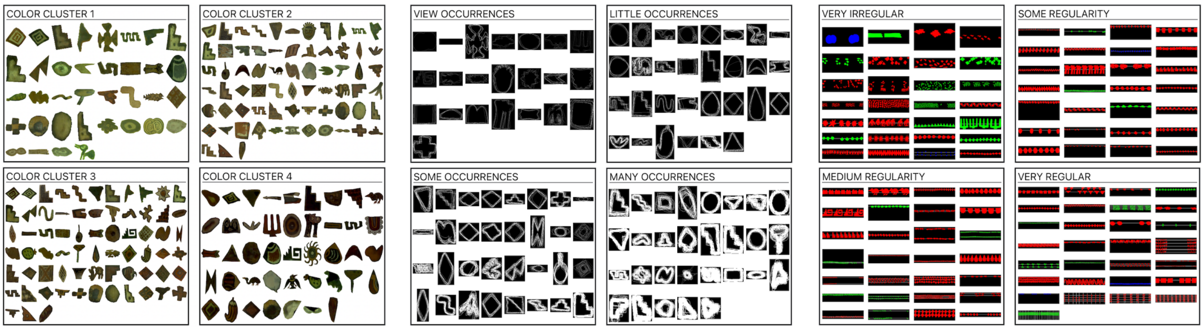
For some of the derived pattern attributes, dedicated previews are devised for displaying them in the clusters. Those comprise the following.

Model Rendering. For each 3D model, one static rendering is generated, which is used as a thumbnail image in the close-up view.

Surface Rollout. One visualization that is already given by the used dataset is a surface rollout (Fig. 4, left). Such a representation is able to visualize a model's surface as a whole and can be obtained by fitting a proxy geometry to the 3D model, which is subsequently projected, cut open, and flattened. The rollouts used in our experiments are based on a variant of the cylindrical unwrapping by Karras *et al.* [KPP96].

Position & Distribution Layout. For the regularity property we implement a glyph showing the positions, scales, and orientations of pattern entities but omitting any distracting textual information (Fig. 4, right). Starting from an empty image, with equal size to the rollout, all pattern entities are drawn as circular markers with the radii indicating their respective sizes. Color coding is used to group the patterns by their archetypes and the individual orientations are visualized by an outwards pointing straight line. Note that, in the case of an n-fold symmetry larger than one, multiple such lines are drawn for each of the equivalent orientations.

Stacked Pattern Contours. To visualize the variety of shapes belonging to one and the same pattern archetype, we conceptualize a stacked outline image comprising the silhouettes of all its pattern entities (Fig. 2). To this end, we leverage the polylines marking the outline of a pattern entity, given within the annotation. For a meaningful overlay, we rotate them in the inverse direction of their given orientation property and scale them relative to their inverse scale property. All these registered polylines



(a) Color (b) Occurrence Frequency (c) Regularity

Figure 5: Clusters resulting from clustering by the properties color (a), occurrence frequency (b) and regularity (c).

are then combined in an additive manner on an empty image.

Traffic Lights. Other (abstract) pattern properties are visualized with a traffic light analogy, after sorting them globally into four bins ranging from ‘very low’ to ‘very high’ (Fig. 3, lower right half).

5 RESULTS

Our proposed visualization concepts are implemented in an interactive prototype, allowing us to evaluate usability and effectiveness aspects. In the following we briefly describe the implementation (Sec. 5.1), before we discuss some of the findings obtained with the system (Sec. 5.2).

5.1 Implementation

For the prototype, we use a web-based implementation relying on React[3] for the visualization frontend and a backend written in Kotlin[4]. Data processing as clustering and image processing, is conducted using Python scripts, relying on the SciPy[5] and OpenCV[6] libraries. Extracted feature descriptors, pattern attributes, and cluster similarities are cached in a MySQL database.

5.2 Findings in the Dataset

In the following, we present some of the datasets’ intrinsic property structures and correlations, revealed by our visualization. Firstly, the kind of clusters obtained by clustering along a single object property. In this regard, we have selected the three varied properties: color, occurrence frequency, and regularity (Sec. 4.3), which are illustrated in single cluster views in Fig. 5a, 5b, and 5c, respectively. All of them comprise – depending on the property in question – different object representatives (Sec. 4.6).

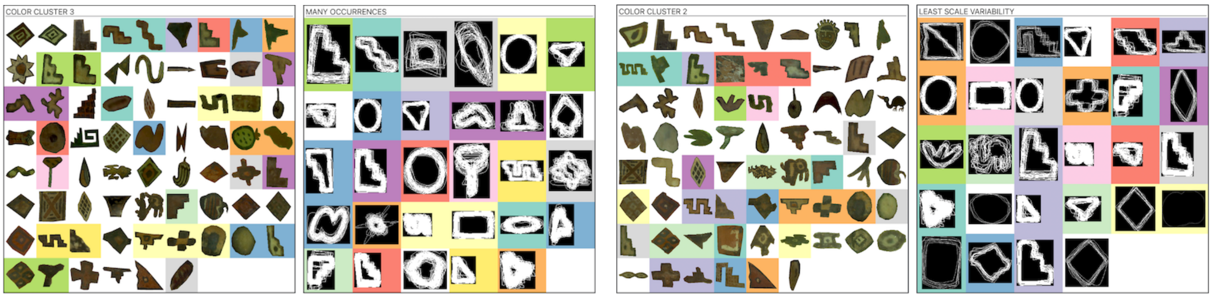
[3] `https://reactjs.org` [Accessed 2023-04-26]
[4] `https://kotlinlang.org` [Accessed 2023-04-26]
[5] `https://scipy.org` [Accessed 2023-04-26]
[6] `https://opencv.org` [Accessed 2023-04-26]

I.e., in the first example – the color clusters (Fig. 5a) – the patterns are represented in their original state, segmented from the unrolled model surface. Four different color clusters are visible with the first one, comprising mostly greenish and yellowish patterns, while the fourth features all the darker patterns. Cluster two and three have mostly light-brown samples. The second cluster view (Fig. 5b) shows the patterns clustered by their occurrence frequency. Here, the stacked pattern contours are employed to visualize the variety and frequency of pattern shapes belonging to a common pattern archetype. From this image, it can be seen that some pattern classes are strongly correlated with the occurrence frequency. More specifically, it appears that cross-like shapes generally have very few occurrences, while staircase-shaped patterns have very many entities. Other shapes like rectangles, tears, circles, etc. are somewhat in-between. For the third cluster view (Fig. 5c), showing the regularity property, we use the position & distribution layout. This helps us to easily spot that the first cluster comprises mostly completely random pattern distributions, while the fourth cluster features clearly features several checkerboard arrangements and other very regular layouts. The examples from cluster two and three exhibit at least either row or column structures.

In another case example, we look into inter-property correlations which are revealed by the Bipartite Clusters Graph. Six of the correlations found in the Peruvian pottery dataset have been selected and are given in Fig. 6 with the side-by-side cluster view. It can be seen (Fig. 6a) that the patterns on the bowl are strongly correlated with the fourth color cluster. I.e., the patterns on bowls generally exhibit darker hues. The bowl shape seemingly also entails a high scale variability (Fig. 6b). Two significant correlations are also established for the pot objects. Firstly, the patterns on this shape are mostly of the light-brownish hue found in color cluster three (Fig. 6c). Secondly, the pot shape also strongly correlates with shape cluster four, which is comprised to a large extent of patterns with a staircase or pyramid-

(a) Vessel Shape ∼ Color (b) Vessel Shape ∼ Scale Variability

(c) Vessel Shape ∼ Color (d) Vessel Shape ∼ Pattern Shape



(e) Color ∼ Occurrence Frequency (f) Color ∼ Scale Variability

Figure 6: Side-by-side cluster views showing the strongly correlated clusters between the properties (a) vessel shape and color, (b) vessel shape and scale variability, (c) vessel shape and color, (d) vessel shape and pattern shape, (e) color and occurrence frequency, as well as (f) color and scale variability.

like outline (Fig. 6d). Those pattern types, in particular, seem to be characteristic for the pot shape.

Comparing the properties color and occurrence frequency also yields a strong interdependency between the color cluster three and the patterns with a very high occurrence frequency (Fig. 6e). The last correlation is between the color and the scale variability properties (Fig. 6f), as it appears that patterns with the least scale variability belong to the second color cluster.

6 DISCUSSION

The examples presented in Sec. 5 show that we are – even without being familiar with the explored domain – able to easily spot several correlations and clusters in the data, which are not revealed by simply looking at the 3D models or images. Besides exploring the data ourselves, we conduct a collaborative walkthrough and subsequent discussion and feedback round with two actual domain experts from the field of archaeology. The exploration workflow was very well received, and it was established that the presented visualization and exploration techniques would also be a valuable tool for other branches of archaeology. One thing that was particularly surprising was the amount of intrinsic object information that could be revealed with basic feature extraction and data processing techniques. Examples for that are the occurrence frequency or scale variability, which makes it easy to identify clusters within the data.

7 CONCLUSION

We introduce an approach for the visual comparison of different properties for sets of ancient pottery objects. Links and color-highlighting allow for identifying relationships between groups in terms of co-occurrence of objects. Our approach supports domain experts to understand vessel shape and ornament elements, and eventually to draw conclusions and help with the interpretation of digital vessel objects. We showed the principal applicability of our concept on a small-sized annotated dataset. The proposed design, however, is also suitable for larger object collections. Future work includes the extension of the shape and ornament features to use, and the inclusion of metadata and textual descriptions of the objects. Data analysis methods as frequent pattern mining, could be a valuable addition to help domain experts search for interesting patterns in large amounts of vessel objects.

ACKNOWLEDGEMENTS

This work was partially co-funded by the Austrian Science Fund FWF and the State of Styria, Austria within the project *Crossmodal Search and Visual Exploration of 3D Cultural Heritage Objects* (CrossSAVE-CH, P31317-NBL).

REFERENCES

[BK10] BHAT, HARISH S and KUMAR, NITESH. "On the derivation of the bayesian information criterion". *School of Natural Sciences, University of California* 99 (2010) 4.

[BMP02] BELONGIE, SERGE, MALIK, JITENDRA, and PUZICHA, JAN. "Shape matching and object recognition using shape contexts". *IEEE transactions on pattern analysis and machine intelligence* 24.4 (2002), 509–522 5.

[Bre] BREWER, CYNTHIA A. *ColorBrewer*. URL: http://www.ColorBrewer.org (visited on 06/16/2022) 5.

[BTC10] BRIVIO, PAOLO, TARINI, MARCO, and CIGNONI, PAOLO. "Browsing Large Image Datasets through Voronoi Diagrams". *IEEE Transactions on Visualization and Computer Graphics* 16.6 (2010), 1261–1270. DOI: 10.1109/TVCG.2010.136 2.

[DPT*12] DEUFEMIA, V., PAOLINO, L., TORTORA, G., et al. "Investigative Analysis across Documents and Drawings: Visual Analytics for Archaeologists". *Proceedings of the International Working Conference on Advanced Visual Interfaces*. AVI '12. Capri Island, Italy: Association for Computing Machinery, 2012, 539–546. DOI: 10.1145/2254556.2254658 2.

[ES16] ESCHBACH, NORBERT and SCHMIDT, STEFAN. *Töpfer Maler Werkstatt: Zuschreibungen in der griechischen Vasenmalerei und die Organisation antiker Keramikproduktion*. Beihefte zum Corpus vasorum antiquorum ; Bd. 7. CH Beck, 2016. ISBN: 3406669409 1.

[GPD17] GLINKA, KATRIN, PIETSCH, CHRISTOPHER, and DÖRK, MARIAN. "Past Visions and Reconciling Views: Visualizing Time, Texture and Themes in Cultural Collections". *Digit. Humanit. Q.* 11 (2017) 2.

[HSS15] HADLAK, STEFFEN, SCHUMANN, HEIDRUN, and SCHULZ, HANS-JÖRG. "A Survey of Multi-faceted Graph Visualization". *Eurographics Conference on Visualization (EuroVis) - STARs*. The Eurographics Association, 2015. DOI: 10.2312/eurovisstar.20151109 2.

[Jac01] JACCARD, PAUL. "Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines". *Bull Soc Vaudoise Sci Nat* 37 (1901), 241–272 5.

[KPP96] KARRAS, GE, PATIAS, P, and PETSA, E. "Digital monoplotting and photo-unwrapping of developable surfaces in architectural photogrammetry". *International Archives of photogrammetry and Remote Sensing* 31 (1996), 290–294 6.

[LKK*20] LENGAUER, STEFAN, KOMAR, ALEXANDER, KARL, STEPHAN, et al. "Visual Exploration of Cultural Heritage Collections with Linked Spatiotemporal, Shape and Metadata Views". *Vision, Modeling, and Visualization*. The Eurographics Association, 2020. DOI: 10.2312/vmv.20201196 2.

[Llo82] LLOYD, STUART. "Least squares quantization in PCM". *IEEE transactions on information theory* 28.2 (1982), 129–137 5.

[LSP*21] LENGAUER, STEFAN, SIPIRAN, IVAN, PREINER, REINHOLD, et al. "A Benchmark Dataset for Repetitive Pattern Recognition on Textured 3D Surfaces". *Computer Graphics Forum* (2021). ISSN: 1467-8659. DOI: 10.1111/cgf.14352 3.

[MPCC13] MAURI, MICHELE, PINI, AZZURRA, CIMINIERI, DANIELE, and CIUCCARELLI, PAOLO. "Weaving Data, Slicing Views: A Design Approach to Creating Visual Access for Digital Archival Collections". *Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI*. Trento, Italy: Association for Computing Machinery, 2013. DOI: 10.1145/2499149.2499159 2.

[PS98] PAPADIMITRIOU, CHRISTOS H and STEIGLITZ, KENNETH. *Combinatorial optimization: algorithms and complexity*. Courier Corporation, 1998 5.

[SIBdS16] SIMON, RAINER, ISAKSEN, LEIF, BARKER, ELTON T. E., and de SOTO CAÑAMARES, PAU. "Peripleo: a tool for exploring heterogenous data through the dimensions of space and time". *Code4Lib Journal* (2016) 2.

[SLL*21] SIPIRAN, IVAN, LAZO, PATRICK, LOPEZ, CRISTIAN, et al. "SHREC 2021: Retrieval of cultural heritage objects". *Computers & Graphics* 100 (2021), 1–20. DOI: https://doi.org/10.1016/j.cag.2021.07.010 3.

[TKE12] TARAWANEH, RAGA'AD M., KELLER, PATRIC, and EBERT, ACHIM. "A General Introduction To Graph Visualization Techniques". *Visualization of Large and Unstructured Data Sets: Applications in Geospatial Planning, Modeling and Engineering - Proceedings of IRTG 1131 Workshop 2011*. Vol. 27. OpenAccess Series in Informatics (OASIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2012, 151–164. DOI: 10.4230/OASIcs.VLUDS.2011.151 2.

[vLKS*11] Von LANDESBERGER, TATIANA, KUIJPER, ARJAN, SCHRECK, TOBIAS, et al. "Visual Analysis of Large Graphs: State-of-the-Art and Future Research Challenges". *Comput. Graph. Forum* 30.6 (2011), 1719–1749. DOI: 10.1111/j.1467-8659.2011.01898.x 2.

[WFS*19] WINDHAGER, FLORIAN, FEDERICO, PAOLO, SCHREDER, GÜNTHER, et al. "Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges". *IEEE Transactions on Visualization and Computer Graphics* 25.6 (2019), 2311–2330. DOI: 10.1109/TVCG.2018.2830759 2.

[WSL*20] WINDHAGER, FLORIAN, SALISU, SAMINU, LEITE, ROGER A., et al. "Many Views Are Not Enough: Designing for Synoptic Insights in Cultural Collections". *IEEE Computer Graphics and Applications* 40.3 (2020), 58–71. DOI: 10.1109/MCG.2020.2985368 2.