From Scar to Scar: Reconstructing Operational Sequences of Lithic Artifacts using Scar-Ridge-Pattern-based Graph Models

Florian Linsel[1], Jan Philipp Bullenkamp[1] & Hubert Mara[1,2]

[1] Institute of Computer Science, Martin-Luther University of Halle-Wittenberg – Halle, Germany
[2] Institute for Computational Ancient Studies (CompAS), Free University of Berlin – Berlin, Germany
Correspondence: florian.linsel@informatik.uni-halle.de

Abstract

Motivated by the importance of lithic artifacts in the development of our ancestors, this study introduces a novel method combining 3D mesh segmentation and graph modelling to determine distinct features of operational sequences being relevant for creating lithic technology analyses.
For analysing the *operational sequence* of scars, manual scar segmentation was utilized on 3D meshes from both open-access and in-house datasets, to construct directed graph models. These models allow the examination of adjacency and sequential relations among scars, represented as nodes and edges in the graph. Our approach, verified against manual graph models, demonstrates the potential for analysing artifacts digitally and enhancing the understanding of early human technological advancements.
Building on existing practices, we created and approach for determining the relative order of scars. We focus on parameters that are approximations of archaeologically used attributes. For the approximations, we used *Multi-Scale Integral Invariants* (*MSII*) curvature values, integral invariants of polylines, and a *MSII* curvature sampling method along the polylines. In addition, we used parameters derived from either the meshes, like surface area, or the graph models, such as node degree or betweenness centrality.
The method's accuracy was tested on various archaeological samples, including Upper Palaeolithic blades and cores from *Grotta di Fumane*, a Bronze Age blade and an experimental knapping series. Preliminary results for high-resolution 3D models are promising, although the overall accuracy varies. While the results now highlight the need for real ground truth data with low interpretation bias and for improvements of the overall approach, they also demonstrate the potential of graphs modelling *operational sequences* to predict the temporal relations between adjacent scar. Further, this approach is a step to make the analysis of *operational sequences* quantifiable, reproducible and reliable beyond the scope of visual comparisons.

Keywords: Graph Modelling; 3D Meshes; Lithic Technology; Operational Sequences; Computational Archaeology

Introduction

Lithic artifacts are of particular interest due to their durability and their significance for all human species. This study will concentrate on the marks resulting from the procedural steps involved in transforming lithic raw material into the finished artifact. One of the long-term objectives is to identify these marks directly on a 3D surface model and order them according to their procedural position.

These procedures consists of a series of strikes to a block of raw material, resulting in the concave removal of material. All strikes have a temporal linear order, which can be analysed as an *operational sequence* (abbr. *OS*, fr. *chaîne opératoire*). As a consequence of a strike, a negative scar is produced on the raw material, while a positive is formed on the removed flake. All scars are connected to the earlier ones by their border, the ridges (Fig 1). For obtaining the scars as surface features the 3D model must be segmented either (semi-) automatically (Bullenkamp, Kaiser, et al., 2024; Bullenkamp, Linsel, et al., 2022) or manually using either Blender (Blender Online Community, 2018) or MeshLab (Cignoni et al., 2008) as interim solutions.

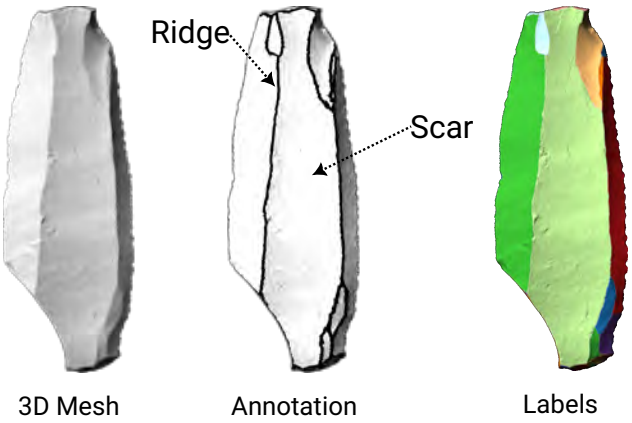


Figure 1. Schematic representation of a solid rendered 3D model, the annotated 3D model (Blender Online Community, 2018) and the vertex labels mapped on the 3D mesh of *ROB* 1241.2.

As demonstrated in Linsel et al. (2024), the segmented 3D model and the manual graph model can be combined, visualized as a 3D graph and directed by the relative properties of adjacent scars to predict their temporal relation. Building on this work, this study includes new approximations of four archaeologically derived properties, including the *Multi-Scale Integral Invariants* (*MSII*) of 3D models (Mara and Krömker, 2013) and the *Integral Invariants of Polyliness* (*IIoPs*) , as well as method sampling parameters along the ridges.

To test the performance of each parameter prediction, a test dataset consisting of different artifacts types belonging to two datasets was created. The dataset consists of Upper Palaeolithic blades and cores from *Grotta di Fumane* (*GdF*) (n = 45), a Bronze Age blade from Troy (*ROB*), and an experimental knapping series (n = 14). The total number of analyzed artifacts is 60. Even though the results should be considered preliminary because the annotations and the *operational sequences* are manually create leading to an interpretation bias, the sample size is low and the variation of records is high, however it demonstrates already the scalability of the approach by integrating new parameters, and the potential of creating graph models of *operational sequences*, which are enriched with 3D data, making the results more comparable, reproducible and transparent.

Related Work

Lithic artifact analysis can be separated in two groups, those with an analogue approach using records for visual analyses and data driven approaches using mostly their digital pendants. Relying on three record types, 3D models, as well as drawings, *operational sequences* (*OS*) and their digital twins, manual segmentations and graph models , this study relates more to the data driven group but relies on data created by the analogue group. While it relates more to digital approaches using 3D models for analysing *operational sequence* of lithic artifacts, a comparison to both categories is needed.

Manual Analyses of Lithic Technologies

One central goal of analysing lithic technologies is to model the knapping sequences needed to create the artifacts (e.g. Soressi and Geneste (2011)), but determining the temporal order of scars and, consequently, their strikes is challenging. Here two different approaches can be applied, which are based on different scales Soressi and Geneste (2011). The "refitting" creates a temporal relation between artifacts based on fitting positive and negative scars of a removal and the core, from which it was removed (Soressi and Geneste, 2011). But if a refitting is not possible, a scar pattern or diacritic analysis is applied on an artifact level analysing the temporal relation of adjacent scars (Pastoors et al., 2015; Richter, 2004; Tafelmaier et al., 2022). Without having explicit knowledge of their order, researchers rely on relative properties between neighbouring scars. A list of 5 attributes are commonly defined, which vary between researchers (Pastoors et al., 2015; Richter, 2004; Tafelmaier et al., 2022) and describe the relation between the younger to the older scars. From a structural standpoint, these attributes often represent a combination of multiple properties. For instance, as stated in Tafelmaier et al. (2022), the first attribute is defined as "The younger negative cuts deeper into the raw material [...]" and "[...] [The younger negative] shows a stronger concavity than the older one [...]", which describes two distinct properties: first the relative location of the scars, and second their difference in concavity. In the archaeological practice, these features can clearly be combined; however, for a computational approach, they must be separated because they are based on different surface features. For a better overview, three lists with each 5 attributes (Pastoors et al., 2015; Richter, 2004; Tafelmaier et al., 2022) were summarized and if the attributes contained more than one property, an attribute is separated in individual properties, resulting in 10 properties (Tab 1).

Property of Younger Scars	Category	Pastoors et al. (2015)	Richter (2004)	Tafelmaier et al. (2022)
is more concave along ridges	*RRP*-1	1	1	1
ridge follows older one	*RRP*-2	4	4	5
ridge cuts across	*RRP*-3	4	-	-
lies deeper	*RSP*-1	1	-	1
is more convex	*RSP*-2	-	-	1
has lances along ridges	*BP*-1	2	2	2
has lance-shaped (multistage) microchips	*BP*-2	3	3	-
has ripple lines in terminal area	*BP*-3	5	-	3
has splinters on the ridge	*BP*-4	-	-	4
is steeply convex in terminal area	*BP*-5	-	5	-

Table 1. Properties to determine the younger and the older scar of two adjacent scars according to Pastoors et al. (2015), Richter (2004) and Tafelmaier et al. (2022). The numbers reflects their list position in the attribute lists. Properties belonging to the same attribute share the position.

For technical purposes, these properties can be categorized into three groups: *Relative Scar Property* (*RSP*), *Rela-*

tive Ridge Property (*RRP*), and *Binary Property* (*BP*). Although only one stated that the ranking reflects the relevance of the properties in archaeological practice (Tafelmaier et al., 2022), the significance of the attribute ranking was implicit in the cited sources. The scars are set in relation to each other, creating a graph model, in which scars are the nodes, ridges are the edges and the edge direction represents the temporal relation. The temporal relation is an interpretation, hence it can vary between researchers. After applying these attributes for all scar relations, these result cumulatively in a complete manual *operational sequence*. One implication of these lists is that a graph model can be completely directed by one or a combination of multiple attributes.

From today's archaeological perspective, graph models of complex scar patterns with 100 - 200 scars, e.g. of bifacial tools, are "hardly informative [for the visual comparison of human experts]" (Kot, 2016). The contraction of scars to working stages is here done either by researchers or in an experiment by the knapper (Kot, 2016; Pastoors et al., 2015; Richter, 2004).

Digital Analyses of Lithic Technologies based on 3D Models or Graph Models

Refitting or 3D puzzling is becoming a topic of interest in the 3D community. For lithic artifacts, manual 3D refitting is becoming popular (e.g. Delpiano et al. (2017) and Sánchez-Martínez et al. (2024) in some instances also automatic refitting (Yang, Matsuyama, et al., 2016). These refittings can then be used to animate the reduction sequence (Yang, Konno, et al., 2019). In related fields, 3D puzzling is also an emerging field, which can be underlined by the increased size of published datasets like the RePAIR benchmark dataset (Tsesmelis et al., 2024) or (semi-) automatic approaches like the refitting of 3D point clouds of marble slabs using varying degree of user input (Houska et al., 2024). As shown in Huang et al. (2006), graph models can be useful to simplifying the registration of 3D models. But most approaches stop after performing the refitting without deriving additional data of the reconnected artifacts, but they can result in graph models like shown by Yang, Konno, et al. (2019).

Contrary to 3D refitting, scar pattern analyses using 3D models has only been demonstrated in Richardson et al. (2014), in which 3D mesh data and an automatic segmentation were used to create an undirected adjacency graph. But Richardson et al. (2014) proposed that additional data could be integrated in the graph model e.g. the "area of the scar" as node parameter and "the mean curvature along the borders" as edge parameter. Following that reasoning, Grosman (2016) highlighted the potential of with "inaccessible data" enriched graph models. Grosman (2016) emphasized that scar and ridge parameters like the distribution of scar counts, their areas, shapes, and mean concavity can lead to a "more precise description and analysis of the lithic artifact surface and the underlying production technology". An analysis of an enriched graphs was referenced but never has been published (Grosman, 2016). To our knowledge, our previous study (Linsel et al., 2024) is the first attempt to analyse *operational sequence*s of lithic artifacts by combining graph models with annotated 3D models.

Graph models are commonly seen as the result of an archaeological study but not the data source for further investigations. One exception is the study of Kot et al. (2024), where they created temporal graph models of scars and working stages. They used experimental knapping sequences and created a strict linear graph model between the documented scars. This study shows the potential of graph based approaches to analyse parameters like the removal rate of scars. It displays how many scars and their temporal relations are still visible in bifacial experimentally reduced artifacts and found out that only 14 % of all removals and 28 % of knapping stages are still visible at the end. In a second study, Mahaney (2014) created and compared action based graph models of several *operational sequences*. They compared graph models of an experimental knapping (hand axe), a hypothetical Oldowan tool and five primate action to evaluate the complexity needed to generate them. Both studies illustrate

an important aspect of graph modelling: graph models can be adapted to the purpose of the study, even if the object of interest is similar, allowing for an argument driven and diverse field of research. But both studies are limited to the graph models without using additional data derived from 2D or 3D models.

Material and Methods

For conducting this research, three different data types are needed, 3D meshes, manual segmentations and the temporal order of scars as colour-coded drawing or as graph model. Due to the time-consuming nature of 3D scanning, annotating and creating graph models of the *operational sequence* as well as prohibitive data policies, no datasets fulfil all three criteria.

Dataset

A dataset was collected consisting of self-created models of an experimental knapping series and one artifact from Troy, as well as artifacts from the northern Italian site *Grotta di Fumane* (*GdF*), which has published 3D models (Falcucci and Peresani, 2023) and colour-coded drawings (Falcucci, Conard, et al., 2017; Falcucci and Peresani, 2018). Even though these drawings and manual *operational sequence* reflect a subjective interpretation, they are at the moment the closest model of a real ground truth dataset and will be used to create and ultimately test the results.

Self-Created 3D Models: All self-created 3D models were scanned using a high resolution structured light scanner (GOM 1 Scan 100). The experimental knapping series was created as an exemplary educational material in Cologne. The series was due to it's use incomplete. A subset of 14 artifacts was scanned in 2023, manually segmented and analysed according to the temporal relations between scars. The series includes multiple flakes, as well as the exploited core. The artifact from Troy is a laterally retouched blade (*ROB* 1241.2), found in the layer Troy II, dated between 4,450 and 4,300 cal BP (Weninger and Easton, 2014). It belongs to the collection of the Archaeological Museum at the Martin-Luther University Halle-Wittenberg (*ROB*).

Externally-Sourced 3D Models: The used artifacts of the open-dataset are from the northern Italian site *Grotta di Fumane* (*GdF*)(Falcucci and Peresani, 2023). The dataset consists of blades and bladelets, cores and flakes from the complete Upper Palaeolithic sequence (Aurignacian, Gravettian) and is dated between 41,000 and 33,000 cal BP (Higham et al., 2009). The 948 3D scans were created using either a structured light scanner (Artec Spider) or a micro CT with a noticeably low resolution of around 40,000 to 400,000 vertices per scan.

62 3D scans were previously annotated and published (Linsel et al., 2023) but unfortunately, for all of these artifacts there are no published *operational sequence* available. For 44 other artifacts an *operational sequence* has been published (Falcucci, Conard, et al., 2020; Falcucci and Peresani, 2018) following the diacritic analysis by Dauvois (1976). Additionally, an *operational sequence* was created for one artifact (*GdF* b-207), bringing the total to 45 artifacts.

Methods

This section consists of the steps needed to convert 3D mesh data to an enriched graph model requiring preprocessed and annotated 3D meshes. These graph models are used to model the spatial and to predict the temporal relation of adjacent scars. For doing so, 10 parameters, 5 approximating archaeologically determined attributes,

are created using either mesh, polyline, or graph-based methods (Tab. 2). All methods are implemented using Python or executed using software like GigaMesh or Blender.

Data Source	Parameter	Archaeological Attribute
Mesh	Surface Area	-
	MAX *MSII* Curvature	*Younger scar is more convex than older scar (*RSP-*2)*.
	MSII Curvature	*Younger scar is more convex than older scar (*RSP-*2)*.
Polyline	*Length of Polyline (*IIoP *1)*	*Younger scar ridge follows older one (*RRP-*2)*.
		*Younger scar ridge cuts across older scar (*RRP-*3)*.
	Angle of Polyline (*IIoP* 2)	*Younger scar ridge follows older one (*RRP-*2)*.
		*Younger scar ridge cuts across older scar (*RRP-*3)*.
	Curvature along Polylines (samp$_{in}$ - *MSII*)	*Younger scar is more concave along ridges than older scar (*RRP-*1)*.
Graph	Degree	-
	Betweenness Centrality	-
	Degree Centrality	-

Table 2. List of parameters used to approximate the archaeologically determined scar properties.

Preprocessing

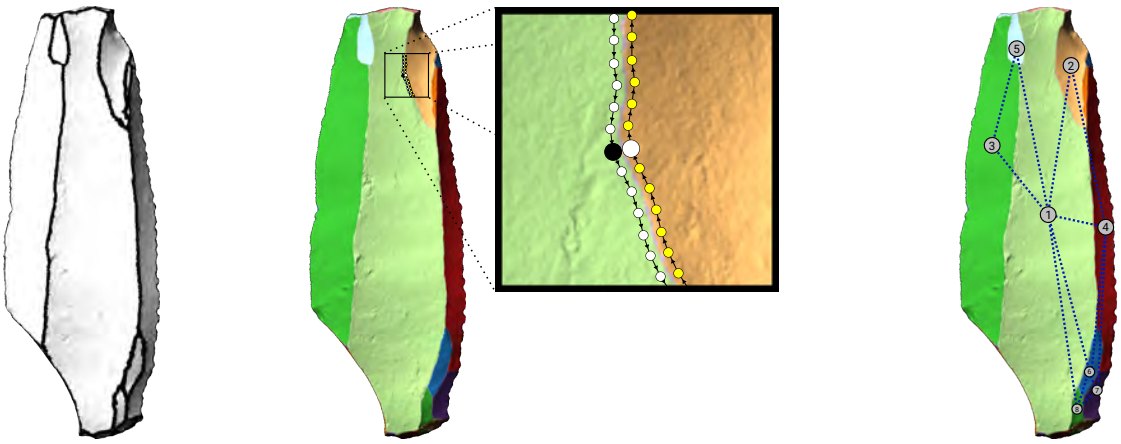
Preprocessing using GigaMesh: All meshes are orientated, cleaned and filled ensuring that the surface is a differentiable 2-manifold (Mara, Krömker, et al., 2010), which can be defined as a pair $M = (V, F)$, where V represents the vertices and F the faces of a mesh M. After this routine, the *MSII* as surface curvature approximation is calculated using 16 equidistant radii scales in the interval $(0, 1.0 \text{ mm}]$ (Mara and Krömker, 2013). The MSII values are used to calculate the convexity of scars (RSP-2).

Manual Segmentation: An annotation workflow using Blender was developed by binary colour-coding the vertices (Blender Online Community, 2018). These colours are then transformed to labels (Fig 1, Fig 2). Previous annotations were done using the MeshLab Z-paintbrush tool (Cignoni et al., 2008), which was discontinued due to its time-consuming and error-prone nature, as all vertices had to be coloured correctly, because vertex colours are referencing to each scar individually and can be directly converted into labels. Only the test knapping artifacts were annotated using MeshLab, all other annotations were created using Blender.

For each label l a connected component S_l is defined as a subset of the original mesh $S_l \subseteq M$, which consists of all connected and uniformly labelled vertices and their faces. A scar is defined as a connected component similar to a polygon directly on the surface of the mesh, where its outline is a closed polyline . In a later step, when two polylines are adjacent, the connected polylines are segmented in two open polyline segments, representing a ridge (Fig 2).

Extract Polyline-based Parameters

The labels allow to compare properties with parameters which are derived from attributes either of the labels or the polylines between two adjacent labels (Tab 2). An important feature of a label is its outline that can be used to analyse the temporal relation between scars. The outline of a label S_l is a polyline of the length n, which is defined as a polygonal chain of vertices $P_l = (v_1, ..., v_n)$, where each v_i belongs to the label l and has at least one

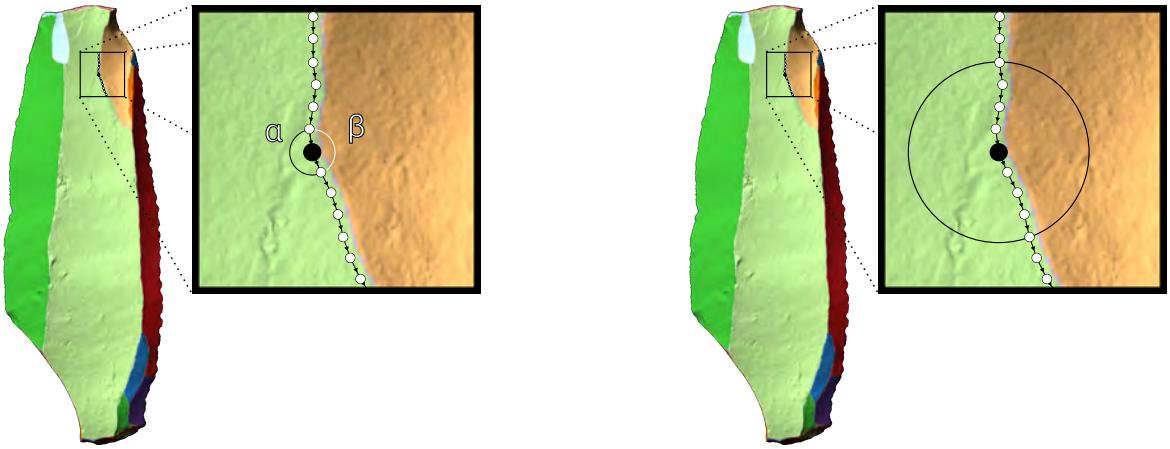


(a) **(b)** **(c)**

Figure 2. Renderings of mesh highlighting key features (*ROB* 1241.2) (a) colour-coded annotation of the ridges using Blender (black:ridge; white=scar); (b) schematic representation of polylines of two adjacent labels; (c) schematic representation of a scar-ridge pattern graph model.

neighbouring vertex with a different label than l.

Integral Invariants of Polylines: The *Integral Invariants of Polylines* (*IIoPs*) are a curvature approximation of a 1-manifold polyline and also the related ridge properties (RRP-2, RRP-3). For calculating them. For each vertex v two points need to be identified, one upstream v_u and one downstream v_d of the polyline, where the polyline first intersects with a sphere of radius r. These points are used to calculate both *IIoPs*, *IIoP 1* and *IIoP 2*.



(a) **(b)**

Figure 3. (a) Estimating the curvature of the polyline by using calculating the angle of polyline (*IIoP 1*) , which is applied on multiple scales (b).

IIoP 1 relies on the angle α between the intersection points v_u and v_d of v (Fig 3), so that α is defined as

$$\alpha(v) = \arccos\left(\frac{d_u \cdot d_d}{\|d_u\|\|d_d\|}\right), \tag{1}$$

where d_u is define as $d_u = v_u - v$ and d_d is defined as $d_d = v_d - v$.

The *IIoP 1* is the signed angle or curvature $\bar{\kappa}$ of the polyline and is defined as

$$\bar{\kappa}(v) = \text{degrees}(\alpha) \times \text{sgn}(v \cdot R), \tag{2}$$

where the rotation axis R is defined as $R = d_u \times d_d$. The range of the curvature $\bar{\kappa}$ is (-180°, 180°). Negative values represent convex, 0° straight and positive values concave polylines.

The *IIoP* 2 is the length-to-radius ratio of the length of the polyline segment within a sphere of radius r. For the length-to-radius ratio, a function len is defined for each $v \in P_l$ as

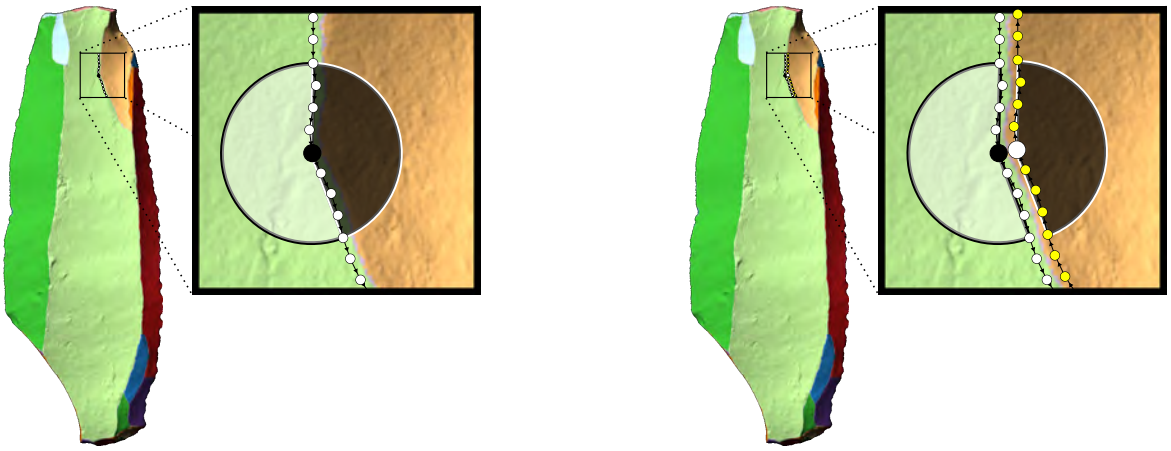
$$len(v) = \frac{1}{2r} \sum_{i \in I} d(v_i, v_{i+1}), \tag{3}$$

where d is the Euclidean distance between the points v_i and v_{i+1}, r is the radius of the sphere, and I are the indices of the vertices along the polyline from v_u to v_d. This parameter can be used to estimate the roughness of a polyline.

Sampling along Polyline: To sample attributes along the polyline, a function $samp_{in}$ is defined to sample an attribute around a vertex v of P_l in a radius r. Only vertices with the same label l as the polyline are considered, so the set of indices to be sampled is given by $I_{l,r}(v) = \{i \mid v_i \in S_l \text{ AND } d(v_i, v) < r\}$. Let

$$samp_{in}(v) = \frac{1}{\mid I_{l,r}(v) \mid} \sum_{i \in I_{l,r}(v)} attr(v_i), \tag{4}$$

where the function $attr$ retrieves the sampled parameter. Generally, Eq 4 can be used for sampling specific attributes of a mesh. In this study, it was used to sample *MSII* surface curvatures of all surrounding vertices along the polyline (RRP-1).



(a) **(b)**

Figure 4. Schematic overview of the sampling method. (a) is showing the sampling along a polyline an attribute inside of label S_1 (white area = $samp_{in}(v_i)$ $v_i \in P_1$) and outside (black area) of a radius; (b) displays the sampling along two adjacent polylines (white area = $samp_{in}(v_i)$, $v_i \in P_1$; black area = $samp_{in}(v_j)$, $v_j \in P_2$).

Operational Sequence and its Graph Model Representations

The following section introduces the concept of graph models based on scars and their adjacencies, which are used to model the relation between scars. Additionally, the potential of graph simplification and a parameter-based prediction of the temporal relation between scars will be introduced.

Scar Graph Model: Creating an undirected graph model $G = (\mathcal{S}, \mathcal{R})$, where scars $\mathcal{S}$ represent the nodes and the ridges $\mathcal{R}$ are the edges that are derived from their adjacency (Fig 5), so that

$$\mathcal{R} = \{\{\mathcal{S}_l, \mathcal{S}_m\} \mid \mathcal{S}_l, \mathcal{S}_m \in \mathcal{S} \text{ AND } \mathcal{S}_m \in N(\mathcal{S}_l)\}, \tag{5}$$

where N returns adjacent scars of a scar on the mesh.

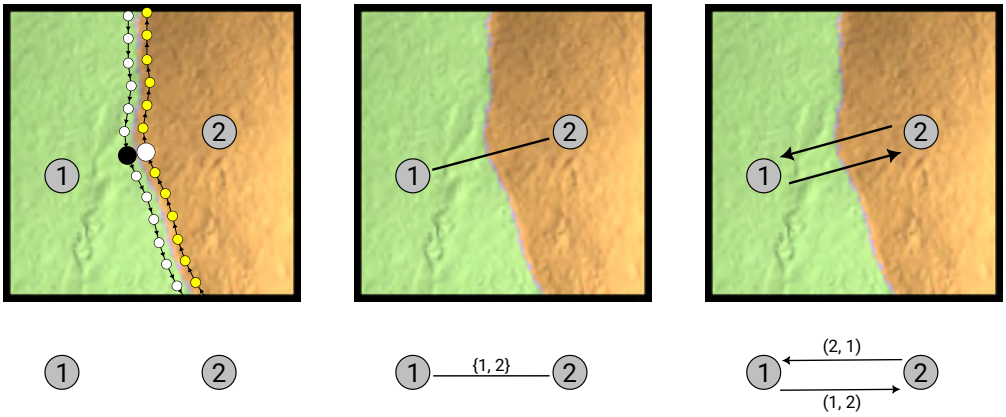


Figure 5. Workflow from the adjacent labels S_1 and S_2, which starts with detecting adjacent polylines (left), over create an undirected edge from the polylines (centre), to the final directed edge (right).

For each edge $\{\mathcal{S}_l, \mathcal{S}_m\} \in \mathcal{R}$ exists an open polyline $C_{l,m}$ (Fig 5 (left)) which is a segmented subset of the closed scar polyline P_l. Thise polyline is defined as $C_{l,m} = (v_1, ..., v_n)$ with $P_m \in N_l(v_i)$ for $v_i \in P_l$. The function N_l returns the neighbouring labels of the vertex v_i.

Parameter-based Directions: By creating the graph G (Fig 5 (centre)), the data of the connected components with the label S and their polylines P are directly assigned to the nodes $\mathcal{S}$ (Tab 3) and those of the ridge segments C are assigned as edge attributes (Tab 4). These mesh features are assigned to the node. This enables to calculate parameters like the mean curvature using the MSII values stored in the vertices or the surface area of each label. Additionally, three network properties (degree, betweenness centrality, degree centrality) are calculated for each node.

- Node $\mathcal{S}_1$
  - Connected Component S_1
    - Vertices
      - *MSII* curvature values
    - Scar Area
  - Polyline P_1
    - *IIoPs*
    - Curvature along the polyline
  - Network Parameters
    - Degree
    - Degree Centrality
    - Betweenness Centrality

Table 3. Mesh and network attributes assigned to the node $\mathcal{S}_1$.

These scar attributes can be used for comparing connected scars relatively and use their difference for predicting their temporal relation, which is based on the direction of the edges. For directing the edges $\mathcal{R}$ of the graph G, a function dir gets defined as $dir : \mathcal{R} \rightarrow \mathcal{R}_{dir}$ such that an undirected edge $\{\mathcal{S}_l, \mathcal{S}_m\}$ in $\mathcal{R}$ is mapped to a directed edge according to the comparison of attributes

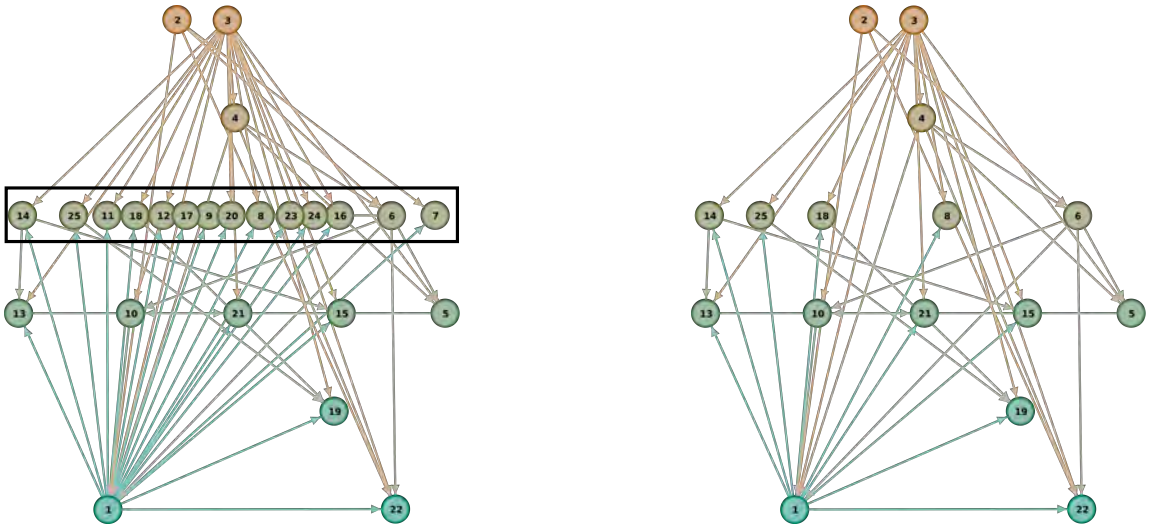
$$dir(\{\mathcal{S}_l, \mathcal{S}_m\}) = \begin{cases} (\mathcal{S}_l, \mathcal{S}_m) \mid \text{attr}(\mathcal{S}_l) < \text{attr}(\mathcal{S}_m) \\ (\mathcal{S}_m, \mathcal{S}_l) \mid \text{attr}(\mathcal{S}_l) > \text{attr}(\mathcal{S}_m) \end{cases} . \tag{6}$$

This function essentially returns an ordered pair of nodes representing the direction of the edge pointing from the node with the lower attribute value to the node with the higher attribute value. If $\mathcal{S}_l$ and $\mathcal{S}_m$ have the same attribute values, the direction is undefined, yet. By applying the dir function to the edges $\mathcal{R}$ (Eq 6) a directed graph DiG can be defined as $DiG = (\mathcal{S}, \mathcal{R}_{\text{dir}})$ (Fig 5 (right)).

- Edge $\{1, 2\}$
  - Polyline $C_{1,2}$ and $C_{2,1}$
    - Curvature of polyline (*IIoPs*)
    - Curvature along polyline (samp - *MSII*)

Table 4. Attributes assigned to the edge $\{1, 2\}$ of the polylines segments $C_{1,2}$ and $C_{2,1}$.

Graph Simplification: In the archaeological practice scars are assigned to working stages, phases or procedural units (e.g. Clarkson, 2002; Richter, 2016; Richter, 2004) of the tool manufacturing sequence. This information is used to contract nodes belonging to the same working stage and hence a simplification of the original graph. In the last step before using the tool, their functional edge can be retouched (Clarkson, 2002; Richter, 2004). These retouches follow the functional edges and are in the case of blades and bladelets isolated between two scars (degree < 3) or additionally connected to each other (Fig 6a), making them detectable. One key assumption is that the temporal relations between the retouched scars as the younger scar and their neighbours as older scars are clear, hence the edge can be contracted.



(a) Graph created from manual *operational sequence* **(b)** Simplified Graph

Figure 6. Graph models of *GdF* b-207 manually created *operational sequence* vertically ranked by Indegree (0 = Top; 6 = Bottom). The black box marks mostly edge retouches (left) (Visualized using Gephi (Bastian et al., 2009)).

By deleting the isolated nodes, the original directed graph DiG can be simplified to reflect the stages of the blank production $DiG_{simp} = (\mathcal{S}_{simp}, \mathcal{R}_{simp})$. To create the simplified directed Graph DiG_{simp} (Fig 6b) and its edges $\mathcal{R}_{simp}$, the nodes $\mathcal{S}_{simp}$ are defined as

$$\mathcal{S}_{simp} = \{\mathcal{S}_l \in \mathcal{S} \mid deg(\mathcal{S}_l) > 2\}, \tag{7}$$

where deg calculates the node degree of a node $\mathcal{S}$.

Graph Simplification - Detecting Retouched Edges: An additional application of the graph simplification is that edges, which are formally interrupted by edge retouches, can be identified and weighted by the amount of cancelled edge retouches between them. For doing so, let $(ret\_edge : \mathcal{R}_{dir} \rightarrow \mathbb{N})$ be a weight function on the edges R_{dir}, which is defined as

$$ret\_edge((\mathcal{S}_l, \mathcal{S}_m)) = \sum_{\mathcal{S}_n \in N(\mathcal{S}_l) \cap N(\mathcal{S}_m)} 1 \text{ if } deg(\mathcal{S}_n) < 3, \tag{8}$$

where the sum of nodes $\mathcal{S}_n$ neighbouring exclusively $\mathcal{S}_l$ and $\mathcal{S}_m$ gets calculated.

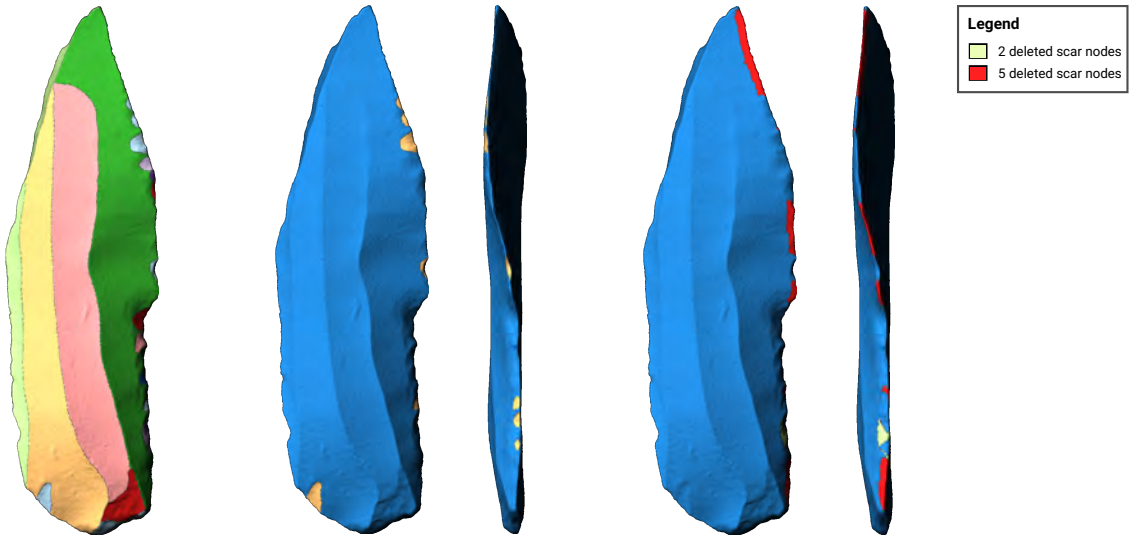


Figure 7. From left to the right, mesh with coloured labels (left), deleted nodes (blue: remaining labels; gold: deleted labels) (centre) and edges previously interrupted by deleted scars and highlighted by amount of deleted scars (right) (*GdF* b-207).

Deducing from the aforementioned reasoning that the edge with the most cancelled nodes is a retouched edge, the marked edge is a functional edge (Fig 7). An additional benefit is that if the artifact has a museum label, or if the surface is altered by adding an object identifier, this can be extracted as long as the alteration does not cross too many ridges.

Results and Evaluation

For each artifact, a directed graph (Fig 6a) and a simplified graph (Fig 6b) was generated based on the manually segmented mesh data and the manually created graph models. Similar to the *MSII* surface curvature, all polyline based methods are tested on 16 equidistant radii scales in the interval $(0, 1.0\ \text{mm}]$. For all multi-scale parameters including the *MSII* curvature, each individual scale got assigned separately. All parameters belonging to labels and ridge polylines are based on vertex data and therefore vary in scale. For all parameters relying on vertex data, like *MSII*, the *IIoPs* or the sampling along polylines, their mean was calculated and assigned either as node or as edge property. By applying Eq 6 for each node and edge property, the edges of a graph model get directed.

The parameter-based direction predictions are tested using an evaluation metric, which compares the direction predictions of the edges with the manually directed edges. Then, the prediction performance for the self-created and the *GdF* datasets get showcased separately. For all multi-scale parameters each individual scale is evaluated and the scale with the highest accuracy is selected for the final results.

Evaluation Metric

In order to assess the efficacy of the temporal prediction, the parameter directed edges $\mathcal{R}_{dir}$ (Eq 6) are compared with the archaeologically directed edges $\mathcal{R}_{gt}$, here defined as ground truth dataset. To evaluate the parameter directed edges $\mathcal{R}_{dir}$ against the ground truth dataset $\mathcal{R}_{gt}$, a function $eval$ is defined as

$$eval(\mathcal{R}_{dir}, \mathcal{R}_{gt}) = \frac{|\mathcal{R}_{dir} \cap \mathcal{R}_{gt}|}{|\mathcal{R}_{gt}|}, \tag{9}$$

where $\mathcal{R}_{gt}$ are the manually created directed edges.

Simplification Rate

The original *GdF* graph models had 7 to 70 nodes and 14 to 193 edges, which were reduced to 6 to 61 nodes and 12 to 175 edges after simplification. Similarly, the self-created graphs originally had 15 to 85 nodes and 37 to 235 edges, reduced to 13 to 76 nodes and 33 to 217 edges after simplification (Fig 8). Simplifying the graphs reduced the differences between the datasets and hence results in a lower cancellation rate for the *GdF* graph models (Fig 9), but the differences still remained prevalent.

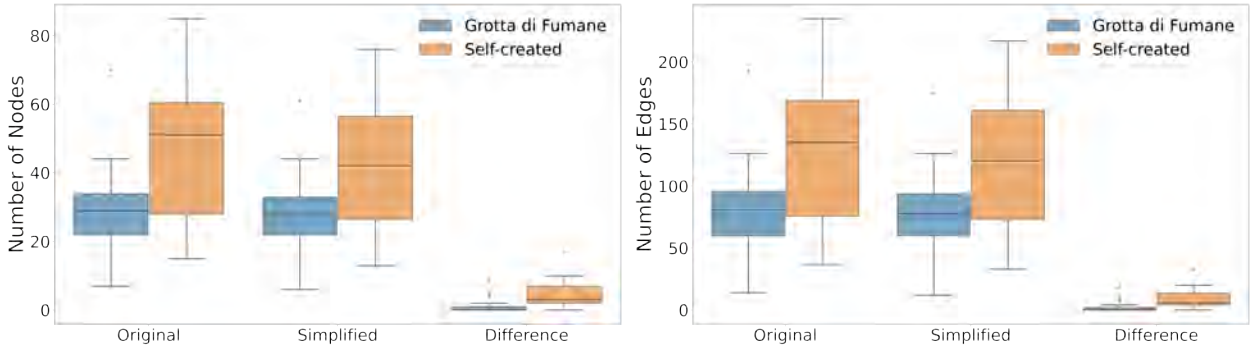


Figure 8. Number of Nodes and Edges in the Original, the Simplified Graph Model, and the Difference between them.

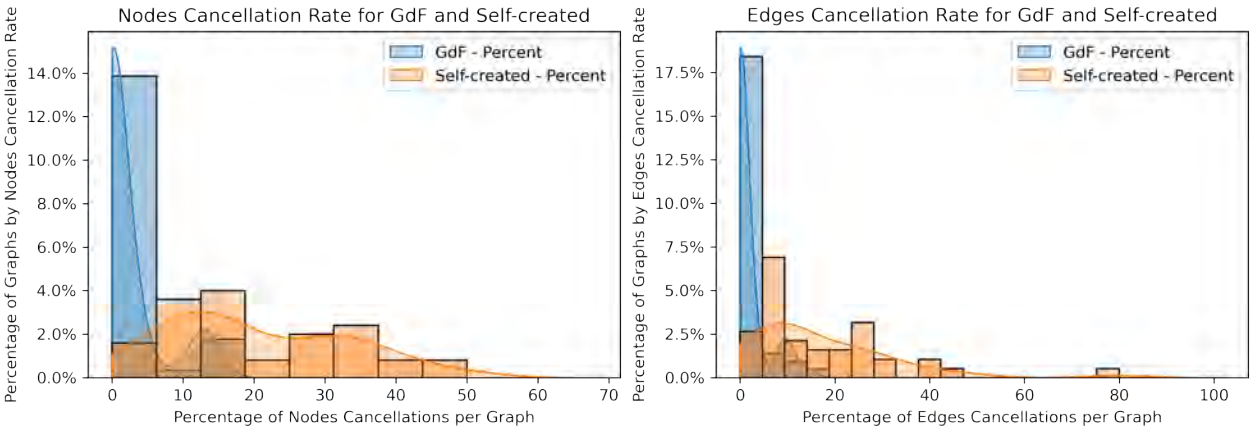


Figure 9. Cancellation rate of all nodes and edges in the simplification.

Parameter Performance

Due to the differences in performance between the self-created dataset and the *GdF* dataset, the results will be introduced separately as shown in Tab 5. However, the sample size of both datasets individually make an in detail evaluation and interpretation of all results impossible and will not yield any final scrutiny. This is particularly true as both datasets are based on manual input.

			Self-created		Grotta di Fumane	
Data Source	**Parameter**	**Property**	**Original**	**Simplified**	**Original**	**Simplified**
Mesh	Surface Area	-	64.41 %	**86.44 %**	**62.06 %**	61.80 %
	MAX *MSII* Curvature	*RSP*-2	58.85 %	**63.65 %**	55.94 %	**56.13 %**
	MSII Curvature	*RSP*-2	58.66 %	**59.36 %**	59.61 %	**59.75 %**
Polyline	Curvature along Polylines	*RRP*-1	61.80 %	**62.18 %**	**62.30 %**	62.16 %
	Angle of Polylines (*IIoP* 1)	*RRP*-2, *RRP*-3	**59.32 %**	58.43 %	62.42 %	**62.73 %**
	Length of Polylines (*IIoP* 2)	*RRP*-2, *RRP*-3	64.95 %	**71.15 %**	62.06 %	**62.28 %**
Graph	Degree	-	58.11 %	**89.87 %**	**55.37 %**	55.14 %
	Betweenness Centrality	-	61.22 %	**90.84 %**	**60.29 %**	60.12 %
	Degree Centrality	-	58.11 %	**89.87 %**	**55.37 %**	55.14 %

Table 5. Mean accuracy of parameters used to approximate the temporal relation of adjacent scars.

Self-created dataset: The mean accuracy of the direction prediction displays a vast spectrum ranging for the original graph model between 58.11 % and 64.95 %, and for the simplified graphs between 58.43 % and 90.84 % (Tab 5). The best performance was achieved using the simplified graph model and the betweenness centrality which ranged between 84.30 % and 96.97 % (Fig 10). For this dataset the graph simplification had a huge impact on the performance yielding an increase of up to 31.76 % in median accuracy (Tab 6).

Data Source	**Parameter**	**Self-created**	**Grotta di Fumane**
Mesh	Surface Area	22.03 %	-0.26 %
	MAX *MSII* Curvature	4.80 %	0.19 %
	MSII Curvature	0.70 %	0.14 %
Polyline	Curvature along Polylines	6.20 %	0.22 %
	Mean Angle of Polylines (*IIoP* 1)	-0.89 %	0.32 %
	Length of Polylines (*IIoP* 2)	-0.98 %	-0.14 %
Graph	Degree	31.76 %	-0.23 %
	Betweenness Centrality	29.63 %	-0.18 %
	Degree Centrality	31.76 %	-0.23 %

Table 6. Percentage differences in accuracy between simplified and original graph models.

***GdF* dataset**: For the *GdF* dataset the best performing parameter was the first *IIoP* $\bar{\kappa}$ with 62.73 %. For the original the overall mean performance ranged only between 55.37 % and 62.42 % and for the simplified graph models, it is between 55.14 % and 62.73 % (Tab 5). The difference between the original and the simplified graph ranged between -0.26 % and 0.32 % (Tab 6).

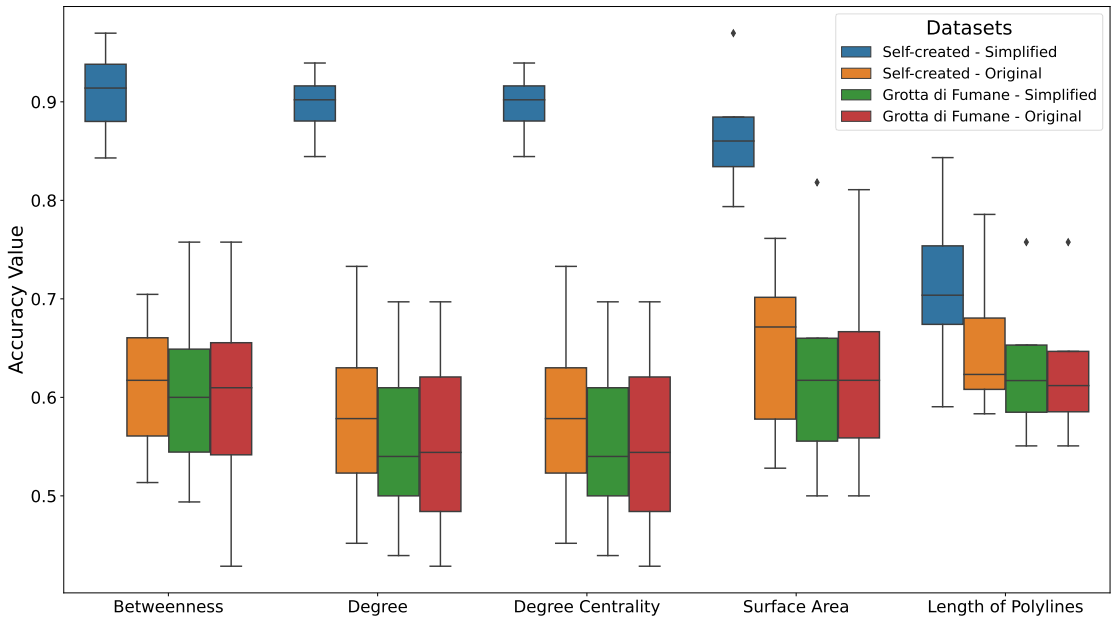


Figure 10. Top 5 performing parameters based on median accuracy.

Even if considering that the *GdF* dataset consists of blades and cores, which could have impacted the result, the performance across both categories stayed low. The best performance for blades are 61.88 % (*MSII* Curvature) for the original and 61.65 % (Curvature along Polylines) for the simplified graph model, and the best performance for cores is 62.69 % (Angle of Polylines) for the original and 63.01 % (Angle of Polylines) for the simplified graph model.

Comparison of Datasets and Discussion

The differences in median accuracy between both datasets are noticeable and the observations of the self-created dataset cannot be seen for the *GdF* dataset. Further, the effect of the graph simplification is completely different, marginal or very positive. That can be partly explained by the resolution of the meshes and thus the visibility of their surface features. But the resolution should have no effect on the surface area, the length of polylines as well as all network and mesh-based properties. It is also possible that the mismatch between the original drawings and the manual segmentation was too high, resulting in a change of the temporal order of the scars. One factor could also be errors in the original *operational sequence*, which could have resulted in an incorrect ground truth dataset.

However, a graph model of scars ridge pattern is always incomplete and partly ordered. That means that the artifacts shows only a specific knapping sequence leading to the final form. Due to the overprinting of younger scars, altering the previous scar pattern, the shape, size and curvature distribution will change for some those metrically and for the whole artifact relatively. Older scars can get separated, creating two distinct marks, belonging to the same original scar. Single scars or complete sequences can get completely removed and remain on the removed artifact. These scars are therefore completely unobservable on the remaining artifact (Kot et al., 2024). However, the attributes to determine the temporal relation between scars should be, if the right attribute or combination of attributes are found, be result in a valid graph model. The same argument would also apply to refitting, which can also only be modelled by partly ordered graph models because of e.g. missing artifacts or

parallelly manufactured artifacts belonging to the same source like different blades from the same core which can be altered simultaneously.

Interesting is that the performance of the *RSP* and *RRP* parameters for both datasets are low in comparison to the additional parameters. The best performing archaeologically derived property is the length of polylines with 71.15 %. The importance of the scar area for determining the direction rings especially true while considering it as counter-intuitive. A general rule of size-frequency distributions of lithic debitage is that removals belonging to a later phase shrink in size (Brown, 2001; Gunn and Mahula, 1976). Contrary to this trend, the result of the scar area suggests that younger scars are larger than older ones. This trend can be explained through the observation that younger scars partly or completely remove older scars, which can also explain the importance of the network parameters, indicating a relation between the temporal position of a scar, its connectivity inside of a graph and its size. The older the scar the more likely it is to be reduced by, but also, to be connected to more younger scars. This leads to the assumption that scars can be grouped either into older, small and well connected and younger, larger, and less connected scars.

Under the premises of all factors playing a role in the process, from the interpretational biases in the creation of the drawings and the *operational sequence*, which than be sourced to manually reproduce them as 3D annotations to the variance in 3D mesh resolution and artifact categories, including cores and blades as well as an experimental knapping sequence, the results could have been influenced by multiple factors.

Whether the higher resolution of the self-created dataset, the subjective decisions during the annotation process or a technological difference is determinable remains questionable. Other factors like raw material, post-processing routines or annotation workflows could have also be influential for the overall performance. None of these questions can be answered yet which is mainly due to the lack of comparable studies using similar techniques and openly published datasets.

In the future, due to the shown scalability of this approach, additional properties could be approximated or existing approximations can be updated or exchanged. The current selection of properties was based on finding accurate representations of surface features like ridges and scars but also on the availability of annotated data. For example, the missing *RSP* property, *RSP*-1, which is defined as "younger scar lies deeper than older scar", is not easily definable parameter wise. All other missing properties are from the *BP* category, which are based on specific features (ripple lines, splinters, lances, microchips), the direction of impact (terminal area) or both. For incorporating these properties, these features need to be annotated and, to make the approach scalable, specific detection algorithms need to be trained to find them.

Outlook and further Applications

Graph models as an analytical tool to simplify technological procedures in archaeology are still in it's infancy. However, using graphs as approximation for roads (Soto, 2019) or social network studies (Brughmans, 2013) yield already a great result for understanding complex systems. In lithic studies, graph models are often used to synthesize knapping sequences for a better visual understanding. However, it is rare for graph models to be the focus of an investigation rather than the result (Kot et al., 2024).

However, beside Linsel et al. (2024), this study is the first in which multiple data sources are combined to tackle this problem at hand. In this study, graphs modelling the knapping sequences were combined with manually

segmented mesh data, enabling an analysis of the temporal relation of scars. The results show a trend that scar area and network parameters yield a high potential for reconstructing the temporal relation of each scar, while commonly used attributes are not yet showing the highest accuracy. The mesh resolution and scar annotation could potentially influence the accuracy of the overall performance. However, these trends should be interpreted as preliminary due to to the scope of the study.

Due to the adjustments needed to be done on the labels and assigning the extra labels to fit the original drawings of the *GdF* dataset, errors are inevitable. Hence, relying on second source drawings, creating labels, which are similar to these and adjusting the scar position according to added and/or reevaluated connections, made this study an challenging endeavour. For testing the approach more reliably, openly published annotated 3D models and ideally experimental knapping series with a clear temporal relationship between scars are crucial. It was not yet demonstrated in this study but the annotating 3D models and graph models can be used to compare differing interpretations of researchers, allowing a higher transparency and comparability of research opinions.

At this stage of the study, the direction prediction of edges is based on singular parameter but in the archaeological practice, a multivariate analysis of multiple parameters are used simultaneously. Hence, a natural next step, is to use an multi-variate approach for predicting the edge directions.

The graph simplification shows potential, especially because it is already similarly used as edge contraction of scars into working stages. Other forms of graph modifications like separating front and backside could have a positive impact on the result. Also, the node parameter derived from the archaeological practice like the scar's phase and working stages, is an interesting use case. Even though the graph simplification is prototypical, but it is a fair assumption that improved methods can help to make graph models more comparable and reduces manual oversegmentation reliably.

Creating graph models of lithic artifacts on a larger scale will also allow to compare these simplified representations to find out which artifacts are similar to each other. Doing this for the graphs of hundreds or thousands of artifacts will potentially show temporal and spatial development of knapping techniques that are otherwise difficult to grasp by looking only on the artifacts as a human.

Acknowledgements

We thank Jürgen Richter for his help in creating the *operational sequences* as well as Armando Falcucci for his help with the 3D models used from *Grotta di Fumane* (*GdF*). We want to thank Rebecca Gnau and Naira Khachatryan for their help in creating the annotations and graph models of the artifacts. We would also like to thank the three anonymous reviewers for their detailed comments.

Fundings

This research was funded by the Martin-Luther University Halle-Wittenberg and conducted at the Institute of Computer Science with contributions from the Department of History and Cultural Studies at the Free University of Berlin.

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Data, Script, Code and Supplementary Information Availability

An open access data publication of the annotated 3D models as well as the graph models is in preparation.

References

Bastian M, S Heymann, and M Jacomy (2009). Gephi: an open source software for exploring and manipulating networks. In: *Third international AAAI conference on weblogs and social media*.

Blender Online Community (2018). *Blender - a 3D modelling and rendering package*. Blender Foundation. Amsterdam.

Brown CT (2001). The Fractal Dimensions of Lithic Reduction. *Journal of Archaeological Science* 28, 619–631. DOI: 10.1006/jasc.2000.0602.

Brughmans T (2013). Thinking Through Networks: A Review of Formal Network Methods in Archaeology. *Journal of Archaeological Method and Theory* 20, 623–662. DOI: 10.1007/s10816-012-9133-8.

Bullenkamp JP, T Kaiser, F Linsel, S Krömker, and H Mara (2024). *it - Information Technology*. DOI: 10.1515/itit-2023-0027.

Bullenkamp JP, F Linsel, and H Mara (2022). Lithic Feature Identification in 3D based on Discrete Morse Theory. In: *Eurographics Workshop on Graphics and Cultural Heritage*. Ed. by Ponchio F and Pintus R. The Eurographics Association. DOI: 10.2312/gch.20221224.

Cignoni P, M Callieri, M Corsini, M Dellepiane, F Ganovelli, and G Ranzuglia (2008). MeshLab: an Open-Source Mesh Processing Tool. In: *Eurographics Italian Chapter Conference*, pp. 129–136. DOI: 10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136.

Clarkson C (2002). An Index of Invasiveness for the Measurement of Unifacial and Bifacial Retouch: A Theoretical, Experimental and Archaeological Verification. *Journal of Archaeological Science* 29, 65–75. DOI: 10.1006/jasc.2001.0702.

Dauvois M (1976). *Précis de dessin dynamique et structural des industries lithiques préhistoriques*. Périgueux: Pierre Fanlac.

Delpiano D, M Peresani, and A Pastoors (2017). The contribution of 3D visual technology to the study of Palaeolithic knapped stones based on refitting. *Digital Applications in Archaeology and Cultural Heritage* 4, 28–38. ISSN: 2212-0548. DOI: 10.1016/j.daach.2017.02.002.

Falcucci A, NJ Conard, and M Peresani (2020). Breaking through the Aquitaine frame: A re-evaluation on the significance of regional variants during the Aurignacian as seen from a key record in southern Europe. en. *J Anthropol Sci* 98, 99–140. DOI: 10.4436/JASS.98021.

— (2017). A critical assessment of the Protoaurignacian lithic technology at Fumane Cave and its implications for the definition of the earliest Aurignacian. *PLOS ONE* 12, 1–43. DOI: 10.1371/journal.pone.0189241.

Falcucci A and M Peresani (2018). Protoaurignacian Core Reduction Procedures: Blade and Bladelet Technologies at Fumane Cave. *Lithic Technology* 43, 125–140. DOI: 10.1080/01977261.2018.1439681.

— (2023). *The Open Aurignacian Project. Volume 1: Fumane Cave in northeastern Italy*. Version 2.1.1. Zenodo. DOI: 10.5281/zenodo.7664308.

Grosman L (2016). Reaching the Point of No Return: The Computational Revolution in Archaeology. *Annual Review of Anthropology* 45, 129–145. ISSN: 1545-4290. DOI: https://doi.org/10.1146/annurev-anthro-102215-095946.

Gunn J and R Mahula (1976). The Sollberger Distribution: analysis and application of a tool reduction sequence. *La Tierra* 3, 2–8.

Higham T, F Brock, M Peresani, A Broglio, R Wood, and K Douka (2009). Problems with radiocarbon dating the Middle to Upper Palaeolithic transition in Italy. *Quaternary Science Reviews* 28, 1257–1267. DOI: 10.1016/j.quascirev.2008.12.018.

Houska P, A Masur, S Kloiber, S Lengauer, S Karl, and R Preiner (2024). Computer-Assisted Collaborative Fragment Matching of Incomplete Stone Artifacts. In: *Eurographics Workshop on Graphics and Cultural Heritage*. Ed. by

Corsini M, Ferdani D, Kuijper A, and Kutlu H. The Eurographics Association. ISBN: 978-3-03868-248-6. DOI: 10.2312/gch.20241244.

Huang QX, S Flöry, N Gelfand, M Hofer, and H Pottmann (2006). Reassembling fractured objects by geometric matching. *ACM Trans. Graph.* 25, 569–578. ISSN: 0730-0301. DOI: 10.1145/1141911.1141925.

Kot M (2016). Technological analysis of bifacial leafpoints from Middle/Upper Palaeolithic transitional industries: Technologische Untersuchungen an bifaziellen Blattspitzen aus Technokomplexen am Übergang vom Mittel- zum Jungpaläolithikum. *Quartär* 63, 61–88. DOI: 10.7485/QU63_4.

Kot M, J Tyszkiewicz, and N Gryczewska (2024). Can we read stones? Quantifying the information loss in flintknapping. *Journal of Archaeological Science* 161, 105905. DOI: 10.1016/j.jas.2023.105905.

Linsel F, JP Bullenkamp, and H Mara (2023). *3D Data Derivatives of Grotta di Fumane: GigaMesh-processed, Annotations and Segmentations*. Version 1.0.0. Zenodo. DOI: 10.5281/zenodo.8288967.

— (2024). Linking Scars: Topology-based Scar Detection and Graph Modeling of Paleolithic Artifacts in 3D. DOI: 10.5281/zenodo.10477448.

Mahaney RA (2014). Exploring the complexity and structure of Acheulean stoneknapping in relation to natural language. *PaleoAnthropology* 2014, 586–606.

Mara H and S Krömker (2013). Vectorization of 3D-Characters by Integral Invariant Filtering of High-Resolution Triangular Meshes. In: *2013 12th International Conference on Document Analysis and Recognition*, pp. 62–66. DOI: 10.1109/ICDAR.2013.21.

Mara H, S Krömker, S Jakob, and B Breuckmann (2010). GigaMesh and Gilgamesh 3D Multiscale Integral Invariant Cuneiform Character Extraction. In: *VAST: International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*. Ed. by Artusi A, Joly M, Lucet G, Pitzalis D, and Ribes A. The Eurographics Association. DOI: 10.2312/VAST/VAST10/131-138.

Pastoors A, Y Tafelmaier, and GC Weniger (2015). Quantification of late Pleistocene core configurations: Application of the Working Stage Analysis as estimation method for technological behavioural efficiency. *Quartär* 62, 62–84. DOI: 10.7485/QU62_3.

Richardson E, L Grosman, U Smilansky, and M Werman (2014). Extracting Scar and Ridge Features from 3D-scanned Lithic Artifacts. *Archaeology in the Digital Era: 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA)*, 83–92. DOI: 10.1515/9789048519590-010.

Richter J (2016). Leave at the height of the party: A critical review of the Middle Paleolithic in Western Central Europe from its beginnings to its rapid decline. *Quaternary International* 411. Middle Palaeolithic in North-West Europe: Multidisciplinary Approaches, 107–128. DOI: 10.1016/j.quaint.2016.01.018.

Richter J (2004). Copies of flakes: operational sequences of foliate pieces from Buran-Kaya III level B1. *The Middle Paleolithic and Early Upper Paleolithic of Eastern Crimea* 3, 233–247.

Sánchez-Martínez J, K Calmet, J Martínez Moreno, and XR Gilabert (Aug. 2024). Virtual reconstruction of stone tool refittings by using 3D modelling and the Blender Engine: The application of the "ReViBE" protocol to the archaeological record. en. *PLoS One* 19, e0309611. DOI: 10.1371/journal.pone.0309611.

Soressi M and JM Geneste (2011). The history and the efficacy of the chaîne opératoire approach to lithic analysis. *Paleo Anthropol.*, 334–350. DOI: 10.4207/PA.2011.ART63.

Soto P de (2019). Network Analysis to Model and Analyse Roman Transport and Mobility. In: *Finding the Limits of the Limes: Modelling Demography, Economy and Transport on the Edge of the Roman Empire*. Ed. by Verhagen P, Joyce J, and Groenhuijzen MR. Cham: Springer International Publishing, pp. 271–289. DOI: 10.1007/978-3-030-04576-0_13.

Tafelmaier Y, G Bataille, V Schmid, A Taller, and M Will (2022). Working Stage Analysis. In: *Methods for the Analysis of Stone Artefacts : An Overview*. Wiesbaden: Springer Fachmedien Wiesbaden, pp. 37–48. DOI: 10.1007/978-3-658-39091-4_5.

Tsesmelis T, L Palmieri, M Khoroshiltseva, A Islam, G Elkin, OI Shahar, G Scarpellini, S Fiorini, Y Ohayon, N Alali, S Aslan, P Morerio, S Vascon, E Gravina, MC Napolitano, G Scarpati, G Zuchtriegel, A Spühler, ME Fuchs, S James, O Ben-Shahar, M Pelillo, and AD Bue (2024). Re-assembling the past: The RePAIR dataset and benchmark for real world 2D and 3D puzzle solving. DOI: 10.48550/arXiv.2410.24010. arXiv: 2410.24010 [cs.CV].

Weninger B and D Easton (2014). The Early Bronze Age Chronology of Troy (Periods I-III): Pottery Seriation, Radiocarbon Dating and the Gap. In: *Western Anatolia before Troy: Proto-Urbanisation in the 4th Millenium BC? Proceedings of the International Symposium held at the Kunsthistorisces Museum Wien Vienna, Austria, 21-24 November, 2012*. Ed. by Horejs B and Mehofer M. Verlag der Österreichischen Akademie der Wissenschaften.

Yang X, K Konno, F Chiba, and S Yokoyama (2019). Visualization of Flake Knapping Sequence with Analyzing Assembled Chipped Stone Tools. *The Journal of the Society for Art and Science*. DOI: 10.3756/artsci.18.40.

Yang X, K Matsuyama, and K Konno (2016). A New Method of Refitting Mixture Lithic Materials by Geometric Matching of Flake Surfaces. *The Journal of the Society for Art and Science* 15, 167–176. DOI: 10.3756/artsci.15.167.