Texture Reconstruction Of Archaeological Textiles Using Hyperspectral Images

Truc Luong Phuong HUYNH[1]

[1]Norwegian University of Science and Technology

Supervisors: Davit Gigilashvili, Sony George, Jon Yngve Hardeberg

Summary—*Reconstructing the Oseberg tapestries is challenging due to their severe deterioration, fragmentation, and the lack of ground truth for identifying relationships between pieces. This project uses hyperspectral imaging to enhance the analysis and reconstruction of these archaeological textiles. By extracting spectral features with techniques like Principal Component Analysis (PCA), Independent Component Analysis (ICA), and Minimum Noise Fraction (MNF), and applying K-means clustering, the study uncovers patterns and hypothesizes connections between fragments. Psychophysical experiments show MNF provides the clearest visualizations. Using the Jaccard Index, the method effectively analyzes fragment relationships, demonstrating the value of computational tools for solving such complex heritage reconstruction problems.*

Keywords—Hyperspectral imaging, archaeological textiles, Oseberg tapestries

1. Introduction

The Oseberg burial mound in Vestfold, Norway, excavated in 1904, revealed a rich collection of Viking-era artifacts, including hundreds of textile fragments. Dendrochronological analysis dated the mound to AD 834, situating these artifacts within the Viking Age's cultural and artistic traditions. Among the discoveries were the Oseberg tapestries—intricately woven textiles featuring processions, mythical creatures, and geometric patterns—providing invaluable insights into Viking storytelling and societal structures. However, severe deterioration of the textiles at the time of excavation rendered it impossible to discern their original length or connections between fragments. Figure 1 illustrates the degraded state of one fragment, highlighting the challenges of interpretation. Despite over a century of documentation and preservation efforts, many fragments remain difficult to analyze [3]. A notable contribution to their documentation is Mary Storm's 1936 hand-drawn illustration, shown in the same figure.

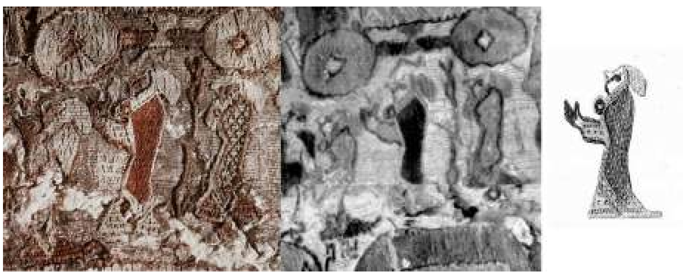


Figure 1. Visualizations of Fragment 1. From left to right: a photograph by E. Johnsen (MCH, UiO) [3], showing a female figure with white headgear, a red robe lined with white, and two wagon wheels above her; an MNF visualization; and a 1936 hand-drawn illustration by Mary Storm (MCH, UiO) [3].

Reconstructing the fragmented textiles is challenging due to their deterioration and the difficulty of identifying hidden patterns and relationships between fragments. The Oseberg tapestries exemplify this complexity, with faded colors, missing sections, and altered structures obscuring their designs. Traditional analysis methods, relying on manual examination and expert intuition, are time-intensive and limited in detecting subtle features [3]. Hyperspectral imaging, a non-invasive technique, overcomes these limitations by capturing data across a wide spectral range, revealing details invisible to the human eye [1], [2]. This approach offers insights into material composition, dye characteristics, and hidden patterns.

The project aims to reconstruct and analyze these textile fragments using advanced hyperspectral imaging and computational techniques. It focuses on enhancing visualization, uncovering hidden details, and aiding archaeologists in identifying and matching fragments, ultimately contributing to reconstructing the original designs and stories embedded in these artifacts (Figure 2).

The methodology integrates spectral visualization and classification techniques. Dimensionality reduction methods such as PCA, ICA, and MNF, along with false-color imaging, are used to highlight key features within hyperspectral data. Classification techniques, including K-means clustering, group fragments based on spectral similarities, aiding in relationship identification. These approaches are evaluated through psychophysical experiments involving archaeologists and naive observers to ensure practical and relevant results.

2. Literature Review

2.1. Previous Works on Oseberg Tapestries

Gigilashvili et al. [4] focus on reconstructing fragmented archaeological textiles using clustering techniques applied to RGB images. Their pipeline begins with a comprehensive preprocessing stage, which includes steps such as image downsampling, segmentation, patch division, and data augmentation to prepare fragments for analysis. Texture features are extracted using methods like Opponent Color Local Binary Patterns (OCLBP) and a combination of OCLBP with the Center of Mass (CoM), as well as feature extraction from pre-trained deep learning models such as AlexNet and VGG. In the clustering stage, algorithms like K-means, Mean Shift, and Hierarchical Clustering are employed to group fragments based on visual texture characteristics, with the goal of simplifying the identification of related pieces.

One of the notable contributions of their work is the development of a digital canvas. This tool allows archaeologists to virtually assemble fragmented textiles, reducing the need for physical handling and helping preserve delicate artifacts. However, their reliance on RGB data and texture-based clustering introduced significant limitations. Archaeologists noted instances where fragments with distinctly different weaves

Figure 2. One of the largest tapestry fragments depicting a solemn procession of people and animals moving leftward, centered around two covered carriages drawn by horses, with male and female figures carrying spears walking between them. The imagery has been interpreted as either an actual event, such as the Oseberg funeral procession, or a mythological or ritualistic scene. In Viking belief, the afterlife varied depending on the manner of death, with warriors going to Valhalla or Folkvang, others to Helheim, and those lost at sea to Rán's hall, suggesting that this tapestry may symbolize such journeys to the afterlife or an idealized ritual procession.

and motifs were grouped into the same cluster, leading to false positives that reduced the accuracy and interpretability of the clustering results.

Our project takes a different approach by utilizing hyperspectral imaging instead of RGB images. Hyperspectral imaging captures surface texture, material composition, and subtle spectral features invisible to traditional methods. This enables a more detailed analysis of textile fragments, improving pattern and relationship identification.

2.2. Spectral Image Visualization

Spectral image visualization is essential for interpreting hyperspectral data by converting complex datasets into accessible formats. Various methods serve different analytical needs, from basic band mapping to advanced dimensionality reduction techniques [1], [2].

2.2.1. RGB Visualization. RGB visualization maps three spectral bands to the red, green, and blue channels to create pseudocolor images. It is efficient and useful for quick overviews but relies on careful band selection to avoid missing significant features.

2.2.2. Bandwise Analysis. This method examines individual spectral bands as grayscale images, isolating specific wavelengths to identify material properties. Though effective for detailed studies, its labor-intensive nature makes it challenging for large datasets.

2.2.3. False-Color Imaging. False-color imaging assigns arbitrary colors to selected spectral bands, highlighting spectral properties like reflectance or absorption. It reveals details not visible in natural color representations, often uncovering hidden structures or material differences.

2.2.4. Spectral Cube Visualization. This method retains the multidimensional nature of hyperspectral data, allowing for interactive exploration across multiple wavelengths simultaneously. However, a spectral cube contains a large amount of information, making it tedious to analyze and visualize.

2.2.5. Dimensionality Reduction Techniques. Principal Component Analysis (PCA) reduces dimensionality by identifying and ranking components based on variance. Independent Component Analysis (ICA) extends this capability by isolating statistically independent components, making it more effective in separating overlapping signals and emphasizing subtle patterns [1]. PCA and ICA were used in the conservation of Edvard Munch's The Scream to highlight pigment deterioration and hidden details [2]. Meanwhile, Minimum Noise Fraction (MNF) enhances visualization by maximizing the signal-to-noise ratio, removing noise, and preserving essential features. MNF is particularly beneficial for degraded datasets, such as ancient textiles, where noise can obscure valuable details.

In summary, basic methods like RGB and false-color imaging are accessible but depend on careful selection of bands from the hyperspectral data. Bandwise and spectral cube visualizations offer more detailed analysis across multiple wavelengths separately and simultaneously but are labor-intensive. Techniques such as PCA, ICA, and MNF provide robust solutions for extracting meaningful features, balancing efficiency, interpretability, and analytical precision. In this work, we utilize all these methods to extract the most useful information for effectively visualizing and grouping textile fragments.

2.3. Spectral Classification

Spectral classification involves assigning pixels in hyperspectral images to categories based on their spectral signatures. This process is crucial for identifying materials, patterns, and regions within complex datasets. In this section, we explore commonly used methods, focusing on unsupervised clustering and advanced spectral classification techniques.

2.3.1. K-means Clustering. K-means clustering is a widely used unsupervised algorithm that groups pixels with similar spectral features. It segments hyperspectral data into distinct clusters, aiding in material identification and pattern analysis. Despite its effectiveness in many datasets, K-means relies on the proper selection of the number of clusters and initialization, which can impact its robustness in datasets with high variability. In this project, we utilized K-means for its simplicity and computational efficiency in segmenting hyperspectral data, focusing on generating initial insights into the relationships between fragments.

2.3.2. Advanced Methods. More precise spectral classification techniques include Spectral Angle Mapper (SAM), Spectral Correlation Mapper (SCM), and Spectral Information Divergence (SID). SAM measures the angle between spectral vectors, making it effective for identifying materials with distinct spectral profiles. SCM and SID build on this by incorporating correlation and divergence metrics, enabling better differentiation in datasets with overlapping spectral signatures. These methods are particularly useful for distinguishing subtle spectral variations, such as differences in material composition.

Although SAM, SCM, and SID offer advanced capabilities, their implementation and optimization require significant computational resources and time. Due to these constraints, we prioritized methods that aligned with the project's immediate objectives. Future work may integrate these advanced techniques to enhance spectral classification and achieve more refined analysis of hyperspectral data.

3. Methodology

3.1. Oseberg Tapestries Dataset

The dataset comprises hyperspectral images of textile fragments captured in two spectral ranges: Visible Near-Infrared (VNIR) and Short-Wave Infrared (SWIR). VNIR images cover wavelengths from 400 nm to 1000 nm, offering insights into material properties visible to the human eye and slightly beyond. SWIR images extend from 1000 nm to 2500 nm, uncovering unique spectral features related to material composition and structural properties that are invisible to the human eye.

The dataset includes 26 fragments, each of which may be further divided into sub-fragments. These sub-fragments within a fragment provide a reliable set of samples for evaluating classification algorithms. Additionally, archaeologists have hypothesized the content and relationships among fragments based on texture, motifs, and weaving styles. These hypotheses serve as a reference for evaluating the classification algorithm's performance, as no definitive ground truth exists for fragment groupings.

3.2. Spectral Visualization

The spectral visualization for this project was primarily performed using ENVI software, a specialized tool for processing and analyzing hyperspectral data. ENVI supports various spectral image visualization techniques, including RGB visualization, bandwise analysis, and dimensionality reduction methods such as PCA, ICA, and MNF.

For RGB visualization, ENVI automatically selects the three most informative spectral bands and maps them to the red, green, and blue channels to generate pseudo-color images. These visualizations provide a quick overview of the hyperspectral data and serve as full reference images in our psychophysical experiment (Figure 4). Additionally, PCA, ICA, and MNF transformations in ENVI were applied to the hyperspectral data, with the most significant information captured in the first few bands. These transformed bands were then visualized as grayscale images (Figure 5) or assigned arbitrary colors for false-color imaging (6). Python's spectral library was used to generate spectral cube visualizations as in Figure 3.

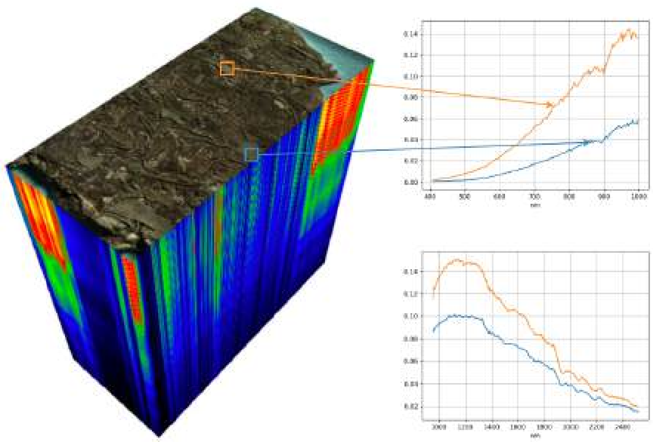


Figure 3. Illustration of the hyperspectral data cube of a subarea of the dataset. Spectral refectance curves of VNIR (aboved graph) and SWIR (belowed graph) at two locations are visualised as examples.

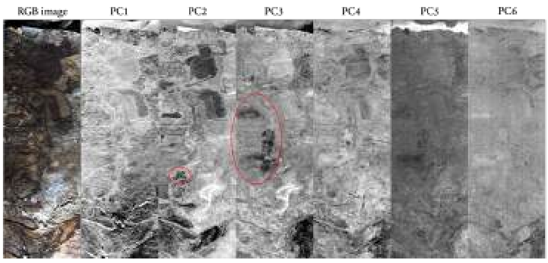
3.3. Spectral Classification

K-means clustering was employed to segment hyperspectral data by grouping pixels into K clusters based on spectral similarity. To ensure accurate analysis, a manual mask was created to exclude background areas (Figure 7). The optimal number of clusters (K) was determined using the following metrics:

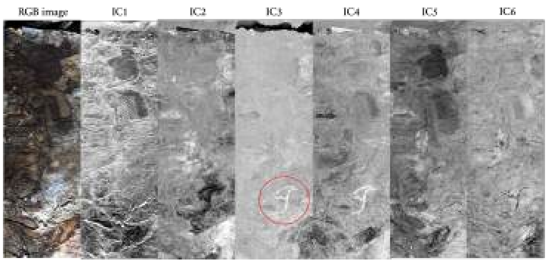
- **Elbow Method**: Visualizes the trade-off between the number of clusters K and the inertia (sum of squared distances within clusters). The optimal K is identified at the "elbow" point, where the decrease in inertia begins to diminish significantly, balancing compactness and simplicity.

- **Silhouette Score**: Assesses clustering quality by comparing the average distance between points in the same cluster (cohesion) to the distance from points in other

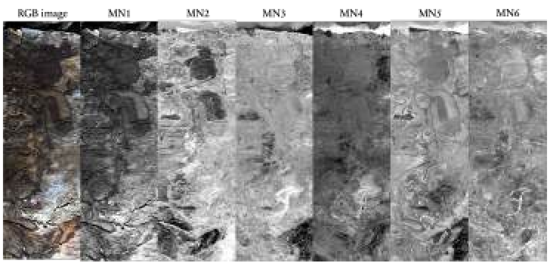
Figure 4. Some RGB Visualziation served as reference images in the psychophysical experiment.



(a) PCA on Texrec1Atop_VNIR



(b) ICA on Texrec1Atop_VNIR



(c) MNF on Texrec1Atop_VNIR

Figure 5. Grayscale visualizations of hyperspectral fragment using PCA, ICA, and MNF components. Each visualization highlights distinct features of the fragment. Red circles show distinctive details of specific components.

clusters (separation). A high average silhouette score indicates well-separated and cohesive clusters, with the optimal K corresponding to the highest score.

- **Davies-Bouldin Index**: Evaluates clustering quality based on the average similarity between each cluster and its most similar cluster, combining compactness and separation. The optimal K minimizes the index, indicating clusters that are both dense and distinct.
- **Calinski-Harabasz Index**: Measures the ratio of between-cluster dispersion to within-cluster dispersion, highlighting how well-defined the clusters are. The optimal K maximizes the index, reflecting clear separation and minimal overlap between clusters.

In Figure 8, the Elbow Method shows a clear "elbow" at $K = 4$ or $K = 5$, where the rate of inertia reduction slows, indicating that additional clusters provide limited improvement. The Silhouette Score is highest at $K = 2$, suggesting well-defined clusters; however, the score declines as K increases, with clusters becoming less distinct at higher values. The Davies-Bouldin Index reaches its lowest point at $K = 3$, indicating compact and well-separated clusters, but beyond $K = 3$, cluster quality declines as the score increases. The Calinski-Harabasz Index continuously increases with higher K, reflecting improved separation between clusters, although the rate of improvement slows around $K = 7$. Balancing all metrics, $K = 4$ or $K = 5$ appears to be the optimal choice.

For our experiment, we chose $K = 4$. The clustering results for the selected K were visualized as segmentation maps, and histograms were generated for each map. The Jaccard Index was then applied to evaluate the similarity between histograms across fragments. A high Jaccard Index between two histograms suggests potential relationships between the corresponding fragments.

4. Results

4.1. Psychophysical Experiment

A psychophysical experiment was conducted with archaeologists and naive observers to evaluate the visualizations. Participants were shown 30 pairs of images where each pair is a reference image using RGB visualization and grayscale visualizations of the chosen components from PCA, ICA, and MNF. They rated the visualizations on:

- **Clarity:** How clear is the visualization in terms of readability, labeling, and general understanding? Rated on a 5-point scale (1 = very unclear, 5 = very clear).
- **Ease of Recognition:** How easily can specific hypothesized elements (e.g., people holding spears, horses, carts) be identified? Rated on a 5-point scale (1 = very difficult, 5 = very easy).

Due to time constraints, we were unable to obtain evaluations from the archaeologists. Although the experiment was sent to them two weeks in advance, their schedules did not allow for participation. Therefore, Table 1 presents the evaluations from naive observers. The results indicate that MNF outperformed other methods. Furthermore, the minimum, maximum, and standard deviation values demonstrate that the evaluations were consistent among different observers.

4.2. Histogram-Based Cluster Evaluation

4.2.1. Dataset and Splitting. Below are chosen fragments for this task, subfragments (e.g. Abot, Atop) of a fragment are

	Clarity			Ease of Recognition			Overall		
	PC	IC	MN	PC	IC	MN	PC	IC	MN
Average	2.588	3.023	**3.274**	3.012	2.881	**3.302**	2.800	2.953	**3.288**
Min	1	1	1	1	1	1	1	1	1
Max	5	5	5	5	5	5	5	5	5
Stdev	1.089	1.187	1.125	1.119	1.257	1.084	1.104	1.226	1.108

Table 1. Statistical Measures for Clarity, Ease of Recognition, and Overall

definately belong to each others. Apart from that: Archaeologists hypotheize relationships between Fragments 12 and 90, and Fragments 86 and 87.

- **Fragment 1**: Abot, Atop, Bbot, Btop
- **Fragment 4**: bot, top
- **Fragment 12**: bot, top
- **Fragment 90**: bot, top
- **Fragment 86, 87**: Single pieces each

Using 66 posible pairs among these fragments and subfragments, we split them into 53 pairs for train set and 13 pairs for validation set.

4.2.2. Jaccard Index and Threshold. The Jaccard Index (J) was calculated for all pairs in the training set to analyze the similarity between subfragments. The results show a clear distinction in the Jaccard Index ranges for related and unrelated subfragments, as outlined in Table 2.

Pair Type	Jaccard Index Range	
	Min	Max
Related Subfragments	0.65	0.85
Unrelated Subfragments	0.15	0.55

Table 2. Jaccard Index ranges for related and unrelated subfragments in the training set.

To effectively classify the pairs, a threshold J_t of 0.6 is selected. This value ensures that pairs with $J \geq 0.6$ are classified as related, while those with $J < 0.6$ are considered unrelated.

4.2.3. Evaluation on Test Set. The evaluation was conducted on 13 pairs in the test set, with results reported in terms of accuracy, precision, recall, and F1-score. These metrics reflect the model's ability to correctly classify related and unrelated fragment pairs based on the selected Jaccard Index threshold ($J_t = 0.6$).

- **Accuracy**: 0.85
- **Precision**: 0.89
- **Recall**: 0.80
- **F1-Score**: 0.84

Actual/Predicted	Predicted Related	Predicted Unrelated
Actual Related	8	2
Actual Unrelated	1	2

Table 3. Confusion matrix for the test set evaluation. Each cell represents the count of pairs classified under specific categories.

5. Discussion

The results demonstrate the effectiveness of spectral visualization techniques and histogram-based clustering for analyzing fragmented textiles. False-color imaging (Figure 9) reveals subtle details, such as weave patterns and material degradation, that are not visible in RGB visualizations. These insights help match fragments, reducing the need for manual analysis and minimizing the risk of missing critical details.

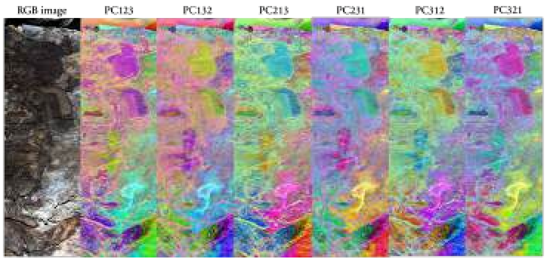
The psychophysical experiment showed that MNF visualizations consistently outperformed PCA and ICA in clarity and ease of recognition. MNF's enhanced signal-to-noise ratio proved especially useful for analyzing degraded textiles. Histogram-based clustering, with a Jaccard Index threshold ($J_t = 0.6$), effectively distinguished related from unrelated subfragments. The high precision (0.89) highlights the model's ability to minimize false positives, while the F1-score (0.84) indicates overall robustness. However, the recall (0.80) suggests room for improvement in identifying all true relationships, potentially through more advanced techniques like SAM, SCM, or SID.

Several limitations affected the workflow and results. The manual selection of components for grayscale visualizations introduced subjectivity, and experimenting with RGB combinations for false-color images was inefficient and time-consuming. The lack of automation in determining the Jaccard Index threshold added complexity and potential inconsistencies. Additionally, the massive size of the Oseberg dataset (40TB) limited the application of these methods, and heavily degraded fragments often yielded little useful information. Limited expertise in interpreting historical patterns and insufficient discussions with archaeologists further restricted the evaluation of results.

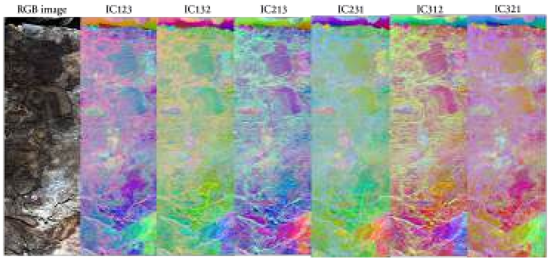
Despite these limitations, the study demonstrates the potential of these methods for addressing the challenges of fragmented textile analysis. Addressing the noted issues will be crucial for refining and scaling the methodology.

6. Conclusion

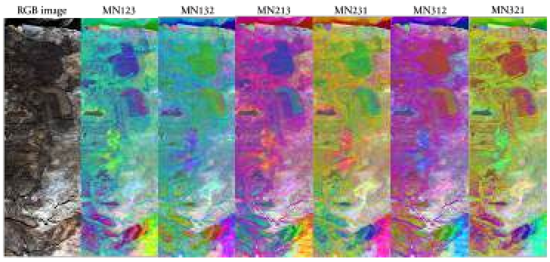
This study demonstrates the effectiveness of combining hyperspectral imaging with computational techniques for analyzing and reconstructing fragmented archaeological textiles, particularly the Oseberg tapestries. Using spectral visualization



(a) False color images using PCA on Texrec1Atop_VNIR (PC123 means R = PC1, G = PC2, B = PC3)



(b) False color images using ICA on Texrec1Atop_VNIR (IC123 means R = IC1, G = IC2, B = IC3)



(c) False color images using MNP on Texrec1Atop_VNIR (IC123 means R = IC1, G = IC2, B = IC3)

Figure 6. False color visualizations of hyperspectral fragment using combination of PCA, ICA, and MNF components.

methods such as PCA, ICA, MNF, and false-color imaging, the project reveals intricate details like weave patterns and material compositions that are often invisible to traditional methods. These insights provide critical clues for fragment matching, enhancing both accuracy and efficiency in reconstruction efforts.

The histogram-based clustering approach, supported by the Jaccard Index, successfully identifies relationships among fragments, achieving high precision and balanced performance metrics on the test set. Psychophysical experiments further validate the clarity and usability of the visualizations, particularly highlighting MNF's capability to enhance degraded data.

While the results underscore the potential of these methods, limitations such as manual component selection, large dataset size, and insufficient discussions with archaeologists highlight areas for improvement. These insights pave the way for future advancements in automating workflows, expanding datasets, and incorporating more sophisticated spectral classification techniques.

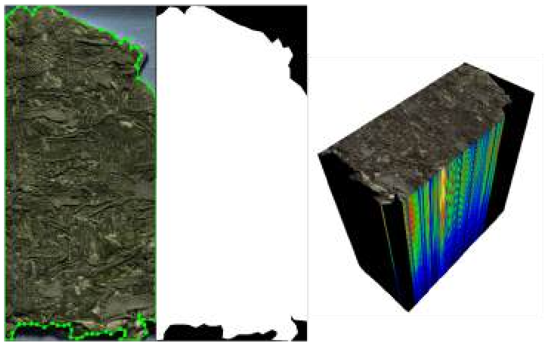


Figure 7. From left to right, RGB visualization, the mask, and hyperspectral cube after applying the mask.

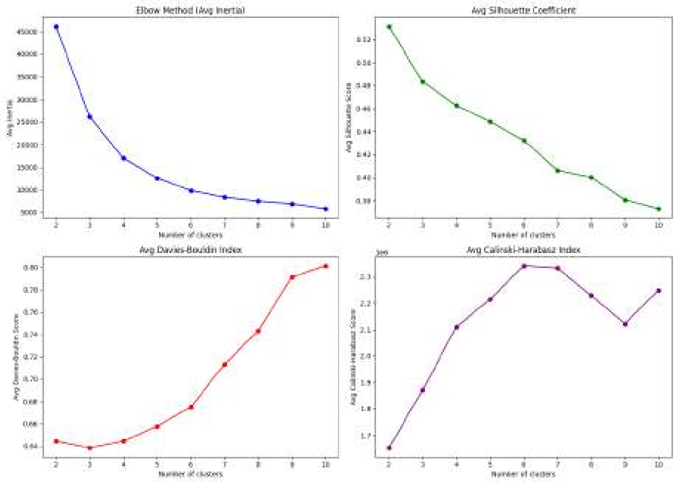


Figure 8. Identifying the optimal number of clusters.

7. Future Work

Future directions focus on refining both the methodology and its applications:

1. Incorporate Spectral Angle Mapper (SAM), Spectral Correlation Mapper (SCM), and Spectral Information Divergence (SID) to improve classification accuracy and detect subtle spectral variations.

2. Develop automated tools for preprocessing, component selection, and threshold determination to streamline workflows and reduce subjectivity.

3. Apply the methodology to larger datasets, including the entire Oseberg collection and other archaeological textiles, to evaluate scalability and robustness.

4. Introduce interactive tools for spectral cube exploration and enhanced false-color imaging to aid archaeologists in interpreting complex data.

5. Strengthen partnerships with domain experts to refine hypotheses and ensure historical accuracy in pattern interpretation.

By addressing these areas, the project aims to further integrate hyperspectral imaging and computational tools into the preservation and analysis of cultural heritage artifacts.

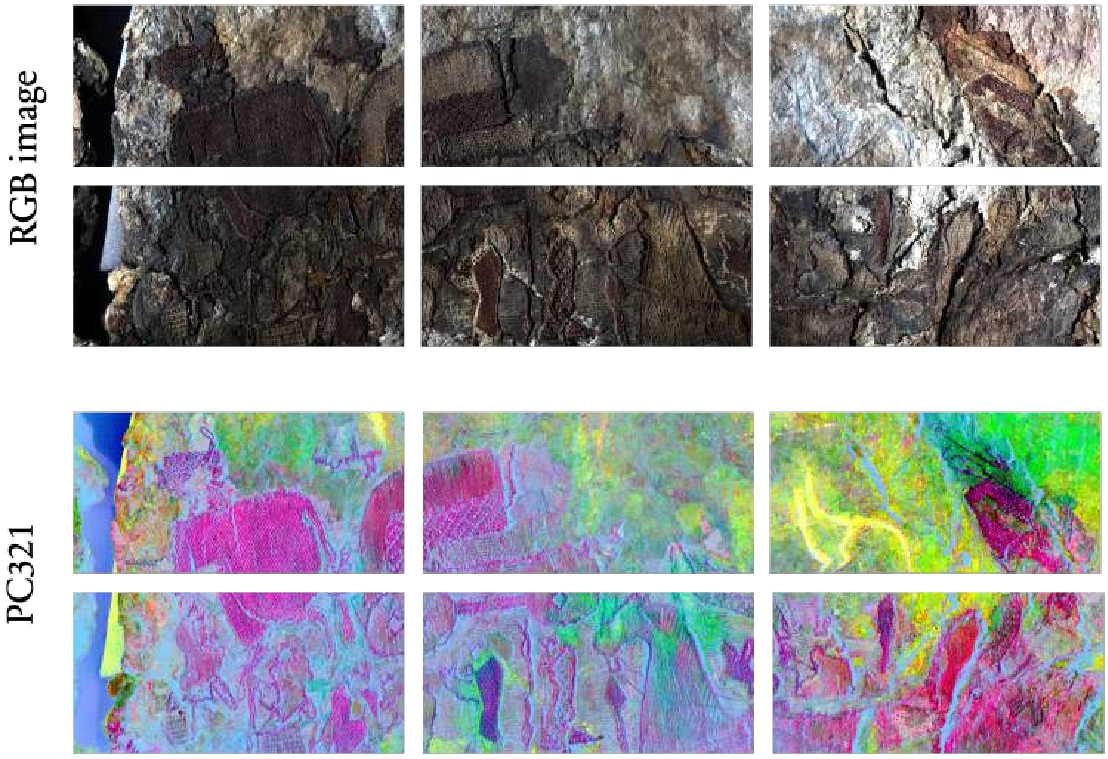


Figure 9. Textile reconstruction as a puzzle problem.

References

[1] Y. Zhu, P. K. Varshney, and H. Chen, "Ica-based fusion for colour display of hyperspectral images", *International journal of remote sensing*, vol. 32, no. 9, pp. 2427–2450, 2011.

[2] J. Y. Hardeberg, S. George, F. Deger, I. Baarstad, and J. E. H. Palacios, "Spectral scream: Hyperspectral image acquisition and analysis of a masterpiece", *Public Paintings by Edvard Munch and His Contemporaries: Change and Conservation Challenges*, pp. 72–83, 2015.

[3] M. Vedeler, *The oseberg tapestries*. Scandinavian Academic Press Oslo, 2019.

[4] D. Gigilashvili, C. F. Gulbrandsen, H. T. Nguyen, M. Havgar, M. Vedeler, and J. Y. Hardeberg, "Toward solving a puzzle of fragmented archeological textiles", 2024.