

Retrieval of vase motifs across different representations using a combination of visual features and poses

Yaping Qu

Supervisors: Mathias Zinnen, Dr. Sabine Lang
Prof. Dr.-Ing. Habil. Andreas Maier

Thesis Description

This master thesis focuses on the task of retrieving images in an image dataset of Greek vases. Image retrieval is a fundamental task in the field of computer vision. It aims to identify and return images from large-scale datasets that closely match a given query image or a textual description [1]. Deep learning has become one of the dominant approaches for image retrieval. Human pose estimation is commonly defined as the task of localizing human joints in images or videos, where joints such as elbows and wrists are represented as keypoints [2]. The dataset used for exploring image retrieval and pose estimation consists of paired images of Greek vases represented in two different styles: full-color images and sketch representations. Greek vases provide valuable evidence for understanding ancient Greek mythology, rituals, and daily life. [3] As digital collections expand, image retrieval and pose analysis can help archaeologists identify related motifs across large datasets. The planned approach includes: (1) extracting image fields from vase images, (2) evaluating the performance of different existing retrieval methods, and (3) training and fine-tuning human pose estimation and detection models to enable queries beyond visual feature similarity. The research problem of this thesis is how to achieve robust image retrieval for Greek vase motifs across different visual representations while accounting for variations in human poses. To address this problem, the thesis is structured into three progressive steps:

Firstly, it will be researched how to retrieve the images only using visual appearance features.

The second step will train a deep learning model for human pose estimation in vase images. Although existing deep learning-based human pose estimation methods have achieved strong performance, several challenges remain to be further addressed. These challenges include occlusion in crowded scenes, difficulties in keypoint association, and limited training data for rare or unusual poses [4] [2]. To accomplish this task, the image dataset of Greek vases will be used to select and train a deep learning model in order to improve pose estimation accuracy.

Building upon the previous steps, the third step is to combine the image retrieval model with the pose estimation model, enabling retrieval that takes both visual appearance and pose information into account. Furthermore, if this combination proves successful, the approach will be applied to an additional dataset, such as an urn image dataset, to evaluate its generalization capability. From a digital humanities perspective, this work supports motif comparison across archaeological collections. From a computer science perspective, it investigates image retrieval on ancient vase dataset, contributing to both cultural heritage research and retrieval methods for non-photographic images.

Literatur

- [1] Longye Du, Shuaiyu Deng, Ying Li, Jun Li, and Qi Tian. A survey on composed image retrieval. *ACM Transactions on Multimedia Computing, Communications and Applications*, 2025.
- [2] Shradha Dubey and Manish Dixit. A comprehensive survey on human pose estimation approaches. *Multimedia Systems*, 29(1):167–195, 2023.
- [3] Dimitris Paleothodoros. Ancient greek vase iconography: theories and practice. *Pharos*, 22(1):49–72, 2016.

- [4] Ce Zheng, Wenhan Wu, Chen Chen, Taojiannan Yang, Sijie Zhu, Ju Shen, Nasser Kehtarnavaz, and Mubarak Shah. Deep learning-based human pose estimation: A survey. *ACM computing surveys*, 56(1):1–37, 2023.