Image and Vision Computing 32 (2014) 414–423

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Morphological analysis for investigating artistic images ☆

Antonio Monroy [a], Peter Bell [a,b], Björn Ommer [a,*]

[a] HCI & Interdisciplinary Center for Scientific Computing, University of Heidelberg, Germany
[b] Institute of European Art History, University of Heidelberg, Germany

ARTICLE INFO

Article history:
Received 23 July 2013
Received in revised form 5 February 2014
Accepted 2 April 2014
Available online 13 April 2014

Keywords:
Shape analysis
Registration
Grouping
Model selection
Stability analysis
Cultural heritage

ABSTRACT

This paper describes an approach for automatically analyzing the alterations of an original artwork during its reproduction. The overall deformation of the artwork is modelled by a piecewise linear model, where regions of the artwork that feature similar alterations are automatically inferred and assigned to the different model components. Model complexity, that is, the required number of affine components required for registration, is automatically estimated using a statistical stability analysis. The main challenge is to simultaneously solve three tasks: (i) inferring the correspondences between both shapes, (ii) identifying the groups in the image that share the same transformation, and (iii) estimating the transformation of these groups. Our approach is tested on controlled scenarios as well as on real historical images.

1. Introduction

Although some stylistic movements in art like impressionism or pointillism define themselves by color, shape has been the predominant way to perceive an artwork. The theory about the primacy of shape can be traced back to Giorgio Vasari (1511–1574) who propagated the line drawing as the predominant technique of all visual arts. His use of the term *disegno* (conceptual design) can be read as the assignment of ideas to shapes. This "shaped idea" is represented through shapes in preparatory drawings, in the artwork itself as well as in drawn reproductions. Based on this observation, changes in shape between artworks and their reproductions or preparatory drawings can be associated with changes in ideas and concepts that reveal artistic choices and stylistic variations. Thus, the analysis of these changes helps art historians to understand the impact of historical influences in the creation and reproduction of art. However, in many cases, these alterations between shapes are very subtle and thus, it becomes extremely difficult, even for experts, to determine the nature and extent of the deformations suffered by different parts within an artwork. The automatic solution of such a shape analysis poses an ambitious computer vision task, and its solution is the focus of the present paper. The nature of the artwork deformations analyzed in this work arises either due to deliberate alterations or due to geometric errors accumulated during the drawing process. For instance, a typical example for a deliberate alteration between a preparatory drawing and the finished work is a subtle conceptual change that induces small alterations in the relative position of extremities in a human pose. These conceptual changes may have personal, cultural, or historical reasons, and thus, it is of interest for art historians to recognize the parts that feature similar transformations and to determine to which extent these parts differ from other regions in the image.

1.1. Piecewise transformation model

The system presented in this contribution addresses the description of an overall non-linear deformation as featured between an original artwork and its reproduction, and at the same time, it gives insights about the local structure of the shape deformation. Shape transformation models within computer vision can be classified into linear and non-linear models. Since *global* linear models cannot be used for describing complex shape changes due to their limited description power, a common choice for describing non-linear changes has been the usage of splines like the TPS [5]. However, besides requiring the estimation of a high number of parameters (proportional to the number of points in the shape) to determine the model, its complexity is regularized by a single manually set parameter for the entire shape. The global nature of this parameter makes it impossible for the model to locally adapt its complexity according to the shape deformation. Therefore, the present paper presents a piecewise linear registration model that adapts the complexity of each component according to the

☆ This paper has been recommended for acceptance by Thomas Pock.

* Corresponding author at: Speyerer Straße 6, 69115 Heidelberg, Germany. Tel.: +49 6221 547853.
E-mail address: ommer@uni-heidelberg.de (B. Ommer).

http://dx.doi.org/10.1016/j.imavis.2014.04.004
0262-8856/

shape deformation in the underlying region. Moreover, the assignment of regions in the shape to different model components induces a clustering which is used in turn to visualize the structure and geometry of the deformation introduced by the artist during the reproduction procedure.

1.2. Automatic complexity estimation

However, a challenge of using piecewise linear models is to automatically determine the number of components required for registration. In the absence of prior knowledge about the shape deformation, the answer to this question represents an important part of the analysis. Nonetheless, an indispensable requirement for selecting the number of components is the robustness of the registration solution. The present paper considers this robustness or stability from a statistical point of view. A stable registration solution for a given number of components is understood as a solution that is reproducible on different subsampled versions of the shape and does not too sensitively dependent on the sample set at hand. Thus, the "correct" number of transformations is defined as the number that yields the most stable solution capable of handling the trade-off between a too rigid transformation and an overparameterization of the transformation model.

1.3. Historical analysis of image reproductions

Finally, we utilize the proposed approach to analyze prominent reproductions from different periods of art history. At first, images coming from the Codex Manesse illustrated between ca. 1305 and ca. 1340 in Zürich and their reproductions commissioned by Bodmer/Breitinger in 1746/1747 are considered. This image collection is important for art history since the Codex Manesse is the single most comprehensive source of Middle High German Minnesang poetry [3] and represents an outstanding source for understanding the visual interpretation of the Middle Ages in early modern and modern times. Whereas the tracings from book illustrations like the reproductions of the Codex Manesse exhibit only slight changes, the differences between a drawing and a mural painting are obviously greater. Therefore, we also analyze parts of Michelangelo's ceiling fresco in the Sistine Chapel (1508–1512) with sketches, which were made in the artists surroundings, probably after Michelangelo's own preparatory drawings or by Dutch artists after the original artwork had been completed.

2. Related work

In the study of Monroy et al. [18], the temporal drawing process of how an image is reproduced was analyzed. It was assumed that parts drawn in closed succession in the reproduction exhibit similar deformations between the images. A limitation is the manual location and matching of landmark points. Furthermore, the approach lacks a unified model since two different clustering algorithms were applied for estimating the parameters of local affine transformations assuming perfect point correspondences, thus making this procedure very susceptible to noise. The present paper formulates a single optimization problem where affine transformations are estimated and points are grouped within the same procedure.

In the work of Monroy et al. [17], we proposed to solve for the groups and affine transformations by formulating a single optimization problem that was solved using deterministic annealing (DA). However, at the beginning of the optimization procedure, shape points were assigned with almost the same probability to the initial affine transformations. Thus, after updating the transformations, all affine parameters became equal and the algorithm got trapped in a local minimum. A further limitation, which is also shared by Monroy et al. [18], was the inclusion of a Euclidean distance term in the energy function to force the compactness of the groups. Thus, a bias in the solution was introduced since groups were clustered due to proximity and not depending on the registration quality. In the study of Monroy et al. [17], we also assumed for simplicity to have fix point correspondences between shapes, and their calculation was not related to the main optimization procedure. The current approach substitutes the DA technique by a linear program (LP) formulation for assigning points to groups. Moreover, we eliminate the Euclidean distance term in the energy function, and groups are found only by the accuracy of registration. In addition, our method also optimizes point correspondences between shapes along with the groups and the transformation within the same procedure.

In the field of sparse motion segmentation for instance, Wang and Adelson [25] presented a method for decomposing videos into similarly moving layers. This method estimates affine motion models for segments on a regular grid. Due to clutter and missing contours, the accurate estimation of small and continuous deviations in transformations cannot be estimated with this approach. In the study of Delong et al. [8], a regularized energy function was minimized with Graph-Cuts ([2]), which also included a pairwise regularization and thus a bias in the result. This regularization led in practice to poorer registration quality since parts in the shape belonging to different model components were mixed. Furthermore, Komodakis et al. [12] presented an LP formulation of a central clustering in which the number of clusters is determined indirectly by a hard to determine penalty term for each data point. Lazic et al. [14] also indirectly determined the number of clusters through the weighting of the different randomly subsampled linear subspaces. Normally, (rigid) motion segmentation can be seen as an application of the more general task of subspace segmentation [14,26]. This latter task commonly assumes that the data points lie on several distinct *linear subspaces* [9,26,7,24,11]. However, the linearity assumption does not hold in our setting: Whereas shape points lie in a 2D vector space, each of the shape parts that were similarly altered by the artist are represented through elements of the affine group. Therefore, the task consists not only of clustering points that define a linear subspace, but three tasks need to be solved jointly: the correspondence between both shapes, the groups in the image that share the same transformation, and the estimation of the transformations of those groups.

In the field of computer graphics, Sýkora et al. [21] embedded each shape in a lattice consisting of several connected squares and registered them by estimating a rigid transformation for every square. Since the registration is only on the level of rigid squares, a grouping into flexibly shaped regions with related modifications is not part of this contribution. Furthermore, Sýkora et al. [21] are not able to handle deformations that do not preserve local rigidity (e.g., scaling or shear), and it requires a significant overlap between shapes for registration. Additionally, in our setting, background clutter needs to be handled, whereas the method of Sýkora et al. [21] is only applied to cartoons without any clutter. Another interesting related work is by Commowick et al. [6], which presented a piecewise affine regularization method for medical image registration. The drawback of this method is that the affine-registered areas need to be estimated manually by the user. Related to piecewise affine registration, Hongsheng et al. [10] recently introduced a matching algorithm based on affine transformations calculated on a triangulation of the shape. In this case, to match articulated objects, it is required to manually select the groups and their articulation in order to match the scene images. Two different works that are related to estimating transformations between artworks are by Chang and Stork [4] and Usami et al. [22]. While Chang and Stork [4] tried to ensure consistent perspective in art images, Usami et al. [22] aimed to dewarp image reflections shown in convex mirrors within very specific paintings. Common non-linear registration algorithms like Chui and Rangarajan [5] or Myronenko and Song [19] are also not suited to the purpose of the present task. Whereas Chui and Rangarajan [5] used a thin plate spline (TPS) to model the transformation, Myronenko and Song [19] estimated a displacement vector for each point in the shape. In both cases, these models introduce artifacts in the registration as observed by Monroy et al. [17], which is undesirable for art comparison.