Head Pose Estimation in Painted Portraits used for Comparison [1)]

Paul Kammerer, Robert Sablatnig, Ernestine Zolda
Institute for Computer Aided Automation,
Pattern Recognition and Image Processing Group
Vienna University of Technology,
Favoritenstr. 9/4/1832, A-1040 Vienna, Austria
paul,sab,zolda@prip.tuwien.ac.at, www.prip.tuwien.ac.at

Abstract:
To compare characteristic features for an artist between faces in painted portraits, the pose of the depicted head has to be considered. This paper describes the estimation of the pose of a depicted head and the mapping of facial regions in faces of different poses to one another. A simplified 3D head model is used on the one hand to estimate the pose, by extracting the model parameters from the 2D image and on the other hand as a reference system to relate regions between different images and to anchor pose independent characteristics.

1 Introduction

The assignment of an unsigned painting to an artist is a difficult problem for art historians and results often in different opinions of experts. A computer aided analysis tool will support the art historian by providing objective and reproducible information to characterize artists. This paper discusses how the pose of a head influences the comparison of painted portraits and presents a method for pose independent analysis of artist specific characteristics. The method presented is part of a hierarchical classification scheme (see [6]) for analyzing painted portraits.

Artists are identified due to their way of painting different facial regions. Since the pose of heads depicted in miniatures varies from frontal view to "three-quarter" and in some cases to profile view, the analysis has to consider the variations of the artist specific characteristics caused by the pose of the head. Characteristics are variations in the arrangement of brush-strokes, shape, and proportion variations of facial features [6]. The frontal part of the head, i.e. th face, will be divided into facial regions like like nose, eyes, mouth and cheeks, shaded parts and highlighted parts. Each region has its own characteristics and will be analyzed

[1)] This work was supported by the Austrian Science Foundation (FWF) under grant P12028-MAT.

individually. Depending on the pose of the head the shape and stroke characteristics within the individual regions will vary. To relate artist specific characteristics to a specific artist, *artist models* are developed, which allow pose independent comparisons within facial regions. The artist models are based on a simple 3D-geometrical model of a human head. A set of transformation parameters allows to establish a relation between the facial parts in the miniature and its corresponding regions in the head model.

Section 2 gives an overview of the classification model and Section 3 briefly describes the pose estimation algorithm and reconstruction of the head model respectively. An empirical estimation of the accuracy of the method and examples with an artificial head and portrait miniatures are given in Section 4. An outlook in Section 5 concludes the paper.

2 The Classification Model

The classification model consists of two parts, the geometrical model of the head and the face on the one hand and the artist model – a set of characteristic features of a certain artist – that are modeled in terms of the head model parameters on the other hand. In the following the three components of the model are described more detailed. The illustration in Figure 1 shows the connection between *image, 3D head model, face model* and *artist model.*

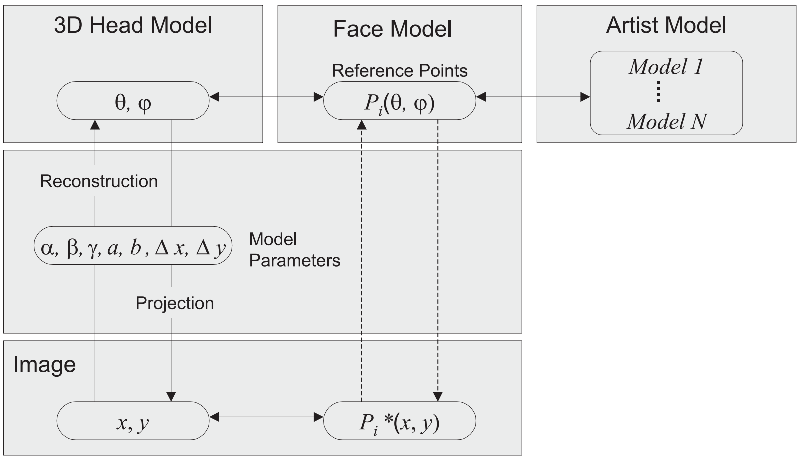
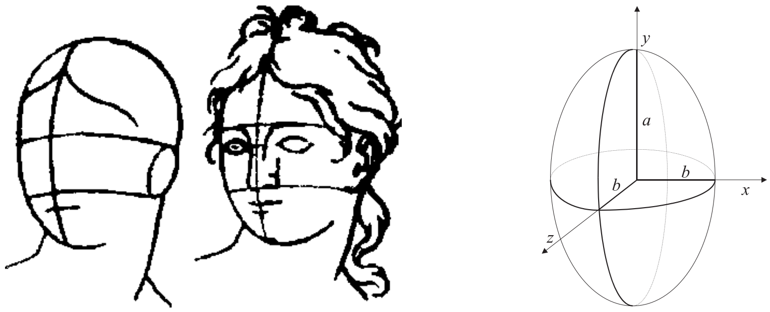


Figure 1: Mapping between image and 3D reference model. Embedding of artist and face model.

The Head Model

A head depicted in a painting can be seen similarly like a photograph as the projection of a real face onto a two dimensional plane. While the acquisition geometry of an image taken by a camera can be calculated very exactly, the acquisition geometry of a painting can only be estimated. From the literature (e.g. [4]) it is known that the use of formulas or schemata depicted in drawing books was common. An example (see Figure 2a) shows that schemata of heads have an oval shape, which can be approximated by an ellipse. In 3D space, the head of a person can be modeled by a rotational ellipsoid. Assuming an orthographic projection, the projection of the ellipsoid to a plane is an ellipse. This simple geometry will allow an easy reconstruction (see Section 3) of the 3D model from the 2D image on the one hand and is complex enough to model the the pose of a 3D head on the other hand.



(a) (b)

Figure 2: (a) Schematic heads from 18th century drawing books, (b) head model embedded in 3D coordinate system.

To overcome the problem of characterizing a human face under general viewing conditions a number of different approaches have been proposed. Most of these methods are based on a 2D representation of the face and deal with the multiple view-points in an example-based manner. Detailed 3D-facial model have been used e.g. for analyzing expression [2]. Simpler 3D-geometrical face models have been mainly used in model-based face tracking [1].

Our model consists of a rotational ellipsoid with the major axes $a = r_y, b = r_x = r_z$. The geometrical setup is shown in Figure 2b. Any point x_i on the surface of the ellipsoid is defined by two angles θ, φ: $x_i = (b * sin(\theta) * sin(\varphi), a * cos(\theta), b * sin(\theta) * cos(\varphi))^T$, while $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi]$.

The model parameters, which map a coordinate x, y from the two dimensional image space to its corresponding coordinate in the model space θ, φ are: the translation $\delta x, \delta y$, the semi-axes of the ellipsoid a, b and the rotation with respect to the $x-, y-, z-axis$: α, β, γ.

Face Model

The face is the frontal part of the head. In our simplified head model, we define the face as the frontal surface of the ellipsoid, $F \in x_i$ where $\theta \in [0, \pi]$ and $\varphi \in [\pi, 2\pi]$. This model ignores the three-dimensional signature of a real face, but it is an approximation which allows to segment the face into facial regions and to assign artist specific features to a certain region. The outlines of the facial regions and characteristic points are defined by a set of reference points $P_i(\theta, \varphi)$ (see also examples in Section 4). Using the ellipsoid head model these reference points can be mapped to the corresponding image points $P_i^*(x, y)$.

Artist Model

To classify the portrait miniatures artist specific models are developed which identify the artist specific characteristics like proportions of eyes, nose, mouth; stroke parameters (length, orientation, color...); illumination (shadowed vs. not-shadowed half of the face); Artist specific characteristics are assigned to facial regions which are defined in the 3D face model. The coordinate space spans the surface of a rotational ellipsoid, which is the 3D model of a human head. Artist specific modulation of these parameters (e.g. a typical characteristic of a painter are oversized eyes) and the specific painting technique applied to specific areas of the face (e.g. stroke type of eye brows) are mapped from the image to the reference-model and compared independently of the projection on this reference model.

3 Reconstruction of the Head model

The rotational ellipsoid – the 3D head model – is reconstructed in two steps. In the first step the position $\delta x, \delta y$ the size a, b and the rotation γ with respect to the $z - axis$ is calculated by fitting an ellipse to the contour of the depicted head. The rotation angle α with respect to the $x - axis$ is very small and thus set to zero. Since it was the artists intention to show persons looking at the viewing person. To obtain the ellipse parameters, we use a least squares minimization algorithm ([3]), which fits the ellipse to a set of points located on the face contour, which are currently marked manually.

In the second step the rotation angle β with respect to the $y - axis$ is calculated by using the symmetry property of a human head. Faces are "almost" symmetric with respect to a symmetry plane. The same holds for "projected" faces in paintings and for certain facial features like eyes, eyebrows, nose, mouth. The rotation angle β is estimated by finding the most symmetric constellation of points from symmetric facial features (see [5]).

4 Experimental Results

This section presents some preliminary results achieved with the method presented. Section 4.1 gives experimental results of an accuracy estimation for the pose estimation algorithm and the mapping algorithm. Section 4.2 shows some examples of region mappings in portrait miniatures.

4.1 Accuracy Estimation Using an Artificial Head

As reported in [5] the algorithm to estimate the rotation angle depends on the selection of symmetric features (i.e. a set of corresponding points). Experiments with an artificial head (see Figure 3) have shown an average angular deviation of 2.6 degree between the actual and the estimated angle within a rotation range of 0 to 40 degree (see Table 1 for details).

actual angle	0	5	10	15	20	25	30	35	40
estimated angle	1	3	16	19	24	25	29	31	39
abs dif	1	2	6	4	4	0	1	4	1

Table 1: Absolute differences between estimated and actual angle.

To estimate the accuracy of the mapping of individual points in an image to the head model, the results of the pose estimation are combined with the mapping procedure. Significant positions in the eyes and the mouth in the artificial head are marked by points and mapped to the normalized head model. Figure 3a-d and 4a show the head in different position, the ellipses fitted to the head and the marked points. The normalized head model and the mapped points are illustrated in Figure 4b. A numerical analysis of the accuracy of the mapping is based on the ϕ-angles of the points mapped to the model. We measured the absolute difference between the points of the frontal head $P^0_{1...6}$ with the points of the rotated heads $P^{10,20,30,40}_{1...6}$ (the superscripts the angle corresponding to the rotation angle of the head,the subscripts denote Points 1 to 6) the absolute difference between. Table 2 shows the differences measured.

$\delta\phi$	left eye left point (1)	left eye right point (2)	right eye left point (3)	right eye right point (4)	mouth left point (5)	mouth right point (6)
$\|a-b\|$	3.3	1.4	2.8	3.1	0.0	0.8
$\|a-c\|$	5.8	5.0	7.3	3.8	0.2	1.3
$\|a-d\|$	6.8	9.7	11.6	10.3	0.9	0.1
$\|a-e\|$	9.1	12.9	15.2	16.9	1.3	5.3

Table 2: Measured differences of reference points.

While point 5 and 6 in the mouth is mapped within an angular difference of 5.3 degrees, the error of the eye points grows to 16.9 degrees in case of point 4. This is an effect which results from the imprecise modeling of the 3D surface of a realistic face on the one hand and instabilities of the projection near the edge of the ellipse.

4.2 Mapping of Regions in Portrait Miniatures

To demonstrate the mapping method six portrait miniatures assigned to one artist (see Figure 3) have been selected. In all six images regions are defined by a set of reference points defining a polygon. After having extracted the head model parameters from each image, the reference points are mapped to the 3D reference model as illustrated in Figure 5a. Figure 5b shows in an example the mapping of the region from image 326 to the image 310.

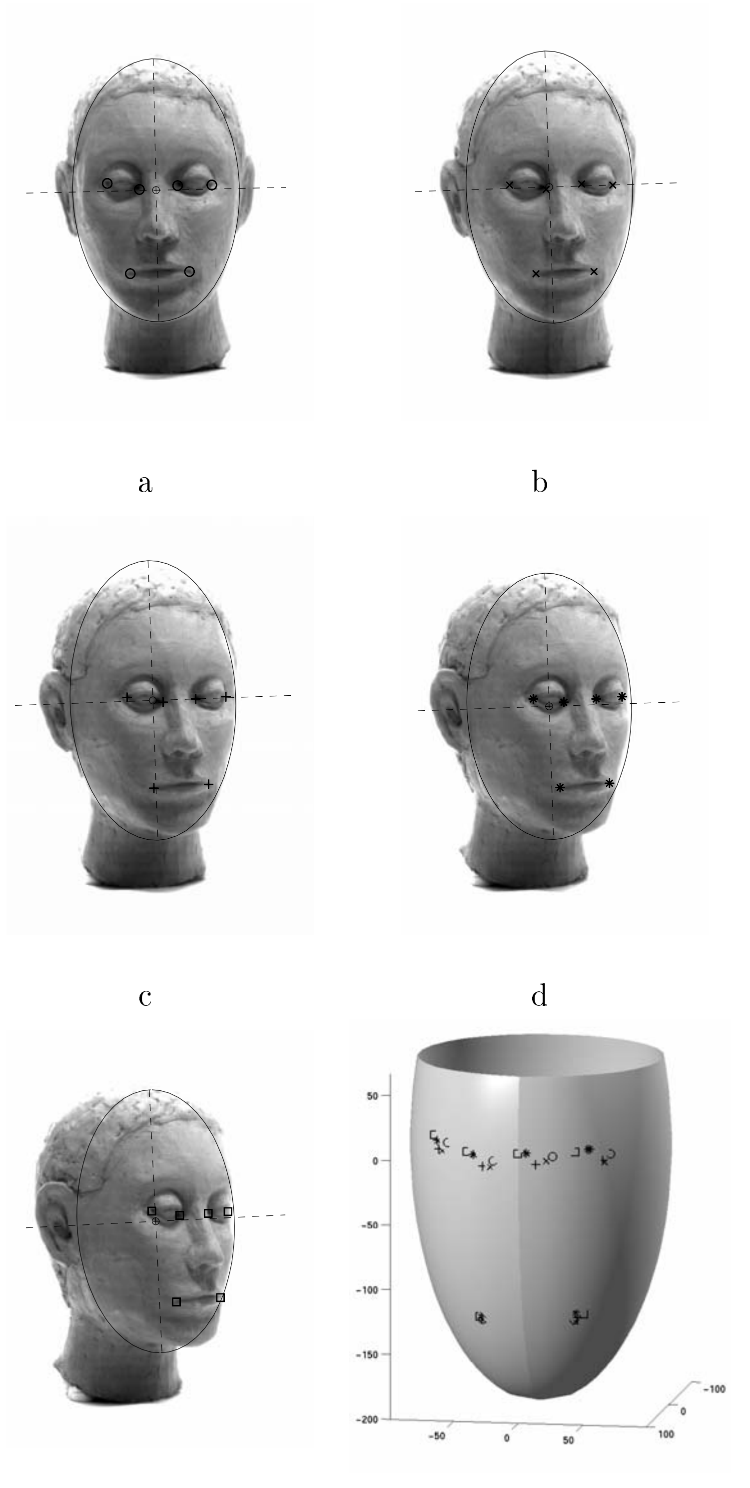
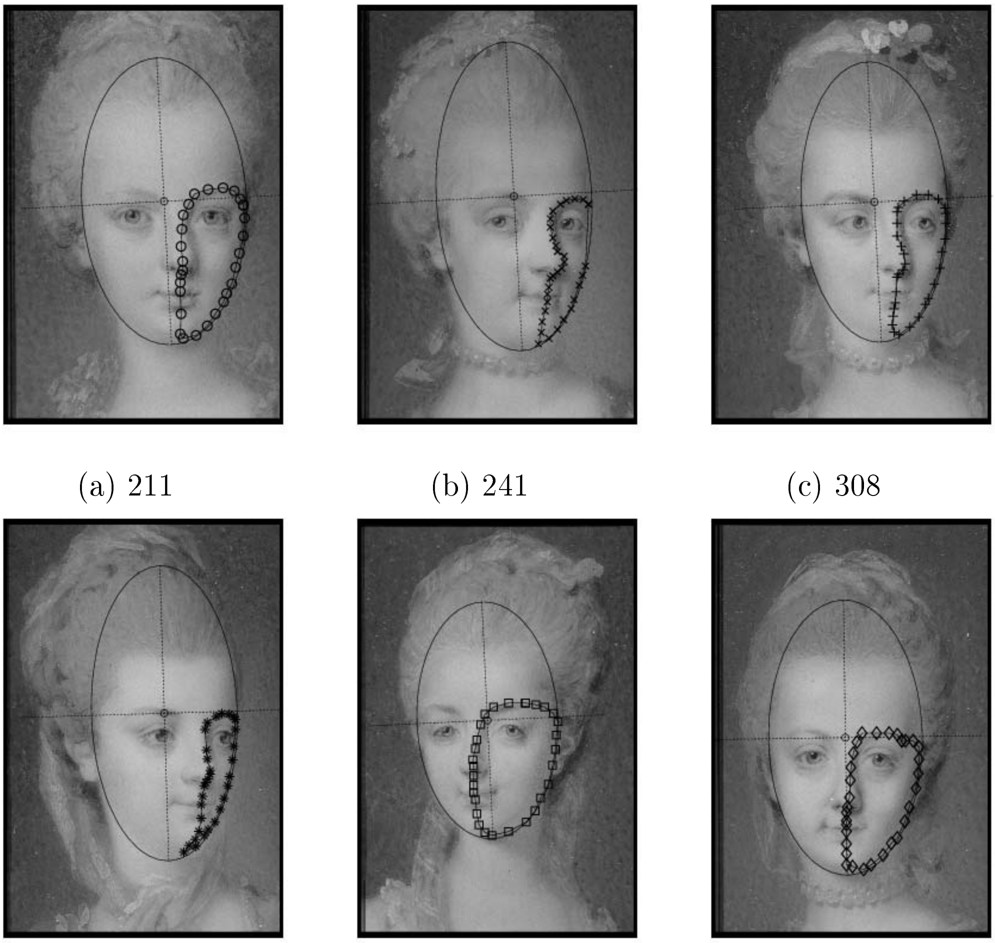


Figure 3: Artificial head in orientations (a) 0 (b) 10 (c) 20 (d) 30 and (e) 40 degrees, (f) normalized head model with mapped points.

These examples shows how our mapping method relates corresponding regions from different miniatures. This mapping is essential for analyzing the style of an artist since a region to region comparison allows to compare the underlying stroke structure correctly. Otherwise only portraits with similar pose could be analyzed. Using the pose estimation and the rotational ellipsoid as the 3D model allows a pose independent comparison of face regions essential for art-historic analysis.



(a) 211 (b) 241 (c) 308

(d) 310 (e) 326 (f) 328

Figure 4: (a)-(f) Six images assigned to one artist showing the fitted ellipse and detected facial regions.

5 Conclusion

The evaluation of the results is rather complicated since there is no ground truth available and the assumption of the 2D projection of a 3D head does not hold since the heads are drawn manually and have inherent artist specific geometry errors. By using an artificial head errors of the pose estimation and the mapping algorithm have been shown. The deviations from the ground truth information are mainly due to inaccuracies of our 3D head model. A more detailed analyses with a synthetic 3D head (e.g. modeled by a 3D graphics software) will allow to estimate the estimation error between an ellipsoid and a synthetic head. For art historic applications, however, the results are quite reasonable.

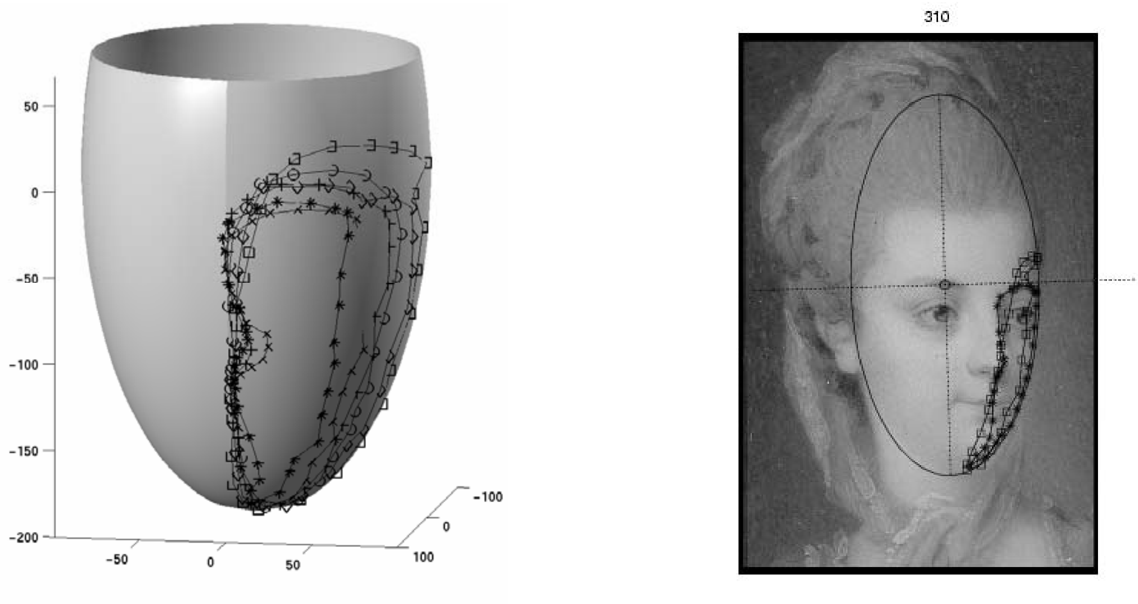


Figure 5: (a) Mapping of facial regions to reference model (b) mapping of region from image 326 to image 310.

References

[1] S. Basu, I. Essa, and A. Pentland. Motion Regularization for Model-based Head Tracking. In *Proceedings of the 13th ICPR, Vienna, Austria*, volume III, pages 611–616, August 25–29 1996.

[2] I. Essa and A. Pentland. A Vision System for Observing and Extracting Facial Action Parameters. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition, Seattle, WA*, pages 76–83, 1994.

[3] A. Fitzgibbon, M. Pilu, and R.B. Fisher. Direct least square fitting of ellipses. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(5):476–480, 1999.

[4] E. Gombrich. *Art and Illusion*, chapter 5, Formula and Experience, pages 146–178. Phaidon Press Limited, Oxford, 1977.

[5] P. Kammerer. Pose estimation and comparison of painted portraits using a 3d head model. In Tomáš Svoboda, editor, *Proceedings of the Czech Pattern Recognition Workshop 2000*, pages 173–178, 2000.

[6] R. Sablatnig, P. Kammerer, and E. Zolda. Hierarchical classification of paintings using face- and brush stroke models. In *14th Int'l Conference on Pattern Recognition, Brisbane, Australia, August 17-20*, pages 474–476, 1998.