Recognizing the Royals - Leveraging Computerized Face Recognition for Identifying Subjects in Ancient Artworks

Ramya Srinivasan, Amit Roy-Chowdhury
Department of Electrical Engineering
University of California, Riverside

Conrad Rudolph, Jeanette Kohl*
Department of Art History
University of California, Riverside

ABSTRACT

We present a work that explores the feasibility of automated face recognition technologies for analyzing identities in works of portraiture, and in the process provide additional evidence to settle some long-standing questions in art history. Works of portrait art bear the mark of visual interpretation of the artist. Moreover, the number of samples available to model these effects is often limited. From a set of portraiture of the Renaissance and Baroque periods, where the identities of subjects are known, we derive appropriate features that are based on domain knowledge of artistic renderings, and learn and validate statistical models for the distribution of the match and non-match scores, which we refer to as portrait feature space (PFS). Thereafter, we use this PFS on a number of cases that have been "open questions" to art historians. They are usually in the form of validating two portraits as belonging to the same person. Using statistical hypothesis tests on the PFS, we provide quantitative measures of similarity for each of these questions. It is, to the best of our knowledge, the first study that applies automated face recognition technologies to the analysis of portraits of multiple subjects in various forms - paintings, death masks, sculptures.

Keywords

Face recognition; Portrait art; Feature selection

Categories and Subject Descriptors

I.4.9 [**Image Processing and Computer Vision**]: Applications

1. INTRODUCTION

Portraits have had a long history encompassing a wide range of art works from ancient sculptures to modern day paintings. Most premodern and early modern portraiture were depictions of people important in their own worlds – from kings and queens to other prominent aristocrats in the society. However, a large number of portraits from before the nineteenth century have lost identities of their subjects through the fortunes of time.

Analysis of faces in portraits can offer significant insights into the personality, social standing, profession and age of the subject they represent. However, this is not a simple task since a portrait can be "subject to social and artistic conventions that construct the sitter as a type of their time" [9], thus resulting in large uncertainty about their identities. Art historians have some lingering ambiguities (based on resemblance of facial features) with respect to certain pairs of portraits whose identities are unknown. Face recognition technologies can be very valuable in providing a quantifiable measure of similarity and an independent source of evidence that can help resolve these long-standing questions.

Challenges: Face recognition in artworks has challenges apart from typical ones such as variations in pose, illumination and facial expression. Some are described below.

1) Choice of features: The chosen features should best justify artists' renditions and possess high discriminating power. Although there has been some preliminary work on this for paintings in general [3], there is little to no work on understanding how to model the style in face portraiture across artists. This leads to interesting questions in machine learning on combinations of various algorithms that are pertinent to the present scenario.

2) Lack of sufficient training data: Many existing face recognition methods assume that extensive training data is available. This is rarely the case here since we are asked to identify similarity between two faces drawn by different artists without having a variety of training examples for these artists. This is due to the difficulty in acquisition of such images from musuems owing to their limited availability, cost and established authenticity. Merely pulling images off the internet would lack scientific integrity. Further, we need to logically choose a set of related images that are relevant to art historians and adhering to specific period styles.

Overview: Figure 1 illustrates the procedure adopted in our work. Artists' renditions are examined to arrive at relevant features for analysis (Sec 2.1) - these being local features (LF) and anthropometric distances (AD). For the pairs of images with *known* identities, we compute LF and AD similarity. Similarly, set of non-match scores are obtained from instances that are known to not match (Sec 2.2). Using Fischer linear discriminant analysis, scores from LF and

*This work was partially supported by an award from the National Endowment for the Humanities, Award No: HD-51735-13.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
MM'13, October 21–25, 2013, Barcelona, Spain.
Copyright 2013 ACM 978-1-4503-2404-5/13/10 ...$15.00.
http://dx.doi.org/10.1145/2502081.2502153 .

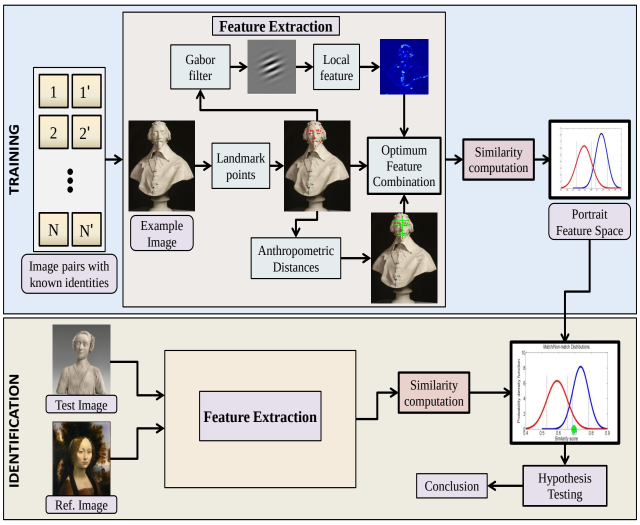

Figure 1: Overview of the Methodology

AD are fused in a way so as to maximize the variance between classes (match/non-match) while maintaining a small variance within each class. Thereafter, using RANSAC algorithm, we fit probability density functions (PDF) to the match and non-match scores and prune outliers to obtain distributions of match and non-match scores (Sec 2.3). We call these learned distributions as Portrait Feature Space (PFS),which is then validated on known instances. For identification purposes, the position of the similarity score between test and reference image in the PFS is used to arrive at conclusions using statistical hypothesis tests (Sec 3).

Contributions: The following are the main contributions of the work :

1. Based on domain knowledge of artists' styles, we identify relevant features that possess high discriminating power.
2. Using statistical pattern recognition tools, we learn and validate PFS from instances with known subjects.
3. We show how unknown subjects can be compared against a reference image using hypothesis testing on the learned distributions.

Related work: Face recognition can be categorized under the broad heading of biometric identification [2]. A survey of still and video based face recognition research is provided in [11]. The vast majority of face recognition techniques have been employed in surveillance and entertainment applications.

Analysis of paintings using sophisticated computer vision tools has gained popularity in recent years [6]. A recent work has explored application of computer based facial image analysis in artworks [7]. The proposed approach used a statistical method for 3D face shape estimation to qualitatively evaluate similarity. While [7] focussed on validating one subject against 4 candidates, our problem is broader. We discriminate between multiple subjects (both genders) across many art-forms/artists taking into account artistic conventions of those times to learn PFS and analyze test cases with respect to PFS. Also, for the present analysis, shape information was found to be less discriminative when compared to other features such as AD and LF. This can be partly attributed to the evidence that artists often focused on LF and took some liberties with shape [9]. This is further substantiated by the use of local features in matching forensic sketches (an art-form) to human faces in works such as [12].

Description of the Dataset: We were provided 34 pairs of images where the identities of the subjects were known. A part of these images was used to learn the PFS and the rest was used to validate it. We were also provided 11 pairs of reference and test images wherein we had to find whether the subject depicted in test image(whose identity is unknown)is same as that depicted in reference image (whose identity is known).The art works consisted of death masks, paintings and sculptures of several aristocrats. A sample representation is provided in Figure 2. The images were carefully vetted for authencity and relevance to a particular artistic style.

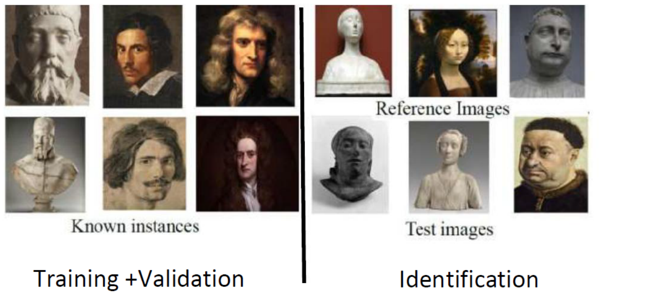

Figure 2: Sample representation of dataset—Columns 1-3 denote images used to learn and validate PFS while columns 4-6 denote images used in testing.

2. LEARNING PORTRAIT FEATURE SPACE

2.1 Feature Selection

Understanding artists' styles: Prominent facial features of a subject are retained in art works of the same subject by various artists [9]. Evidence regarding preserving certain salient body proportions (also known as anthropometric distances) in art works can be obtained from [8]. The importance of these distances is evident in various cultures starting from the ancient Egyptian era to the more recent Renaissance era.

Choice of Features: We find the following aspects emphasized by various artists particularly useful for analysis.

1. Local Features: These include features such as eye corners, nose tip, etc. which are specific to a person. One of the most well-known approaches to analyze local facial features is by Elastic Bunch Graph Matching [10] where faces are represented as image graphs based on fiducial points on the face (e.g., eyes, nose) extracted using Gabor filters. Gabor wavelets are very robust and biologically relevant since they mimic the behavior of visual cortex. This method is useful in extracting the features which artists emphasize.

2. Anthropometric Distances: These include salient distances such as width of forehead, upper face, etc. We leverage upon methods that have provided normative models of facial measurements with the degree of deviation in a population [1].

2.2 Feature Extraction

Each of the features mentioned above is extracted from the portrait images, the details of which, we explain below.

1. Local features: A set of 22 fiducial points is used to represent each face. These include forehead tips (left, right), forehead center, chin bottom, eye corners (right, left of each eye), iris (left, right), cheekbones (left, right), nose top, nose bottom, mouth corners (left, right), chin ear corners (left, right), points on temple (left, right), and points on chin (left, right). The precise location of these points is determined by registering a generic mesh on the face and finding the corresponding points between them.

Gabor jets are evaluated at each of these fiducial points. A jet describes a small patch of grey values in an image around the fiducial points described above. It is based on convolution of the image with Gabor wavelets corresponding to 5 frequencies and 8 orientations.

We evaluate similarities between jets across corresponding fiducial points n in two faces as described in [10]. LF similarity score between two portraits is evaluated as the average of jet similarities over all fiducial points N.

Anthropometric Distances: All images are normalized with respect to scale and orientation. A set of 11 salient anthropometric distances, represented as a vector, characterizes each face. These distances include distance between iris, between forehead center and chin bottom, between forehead tips, between nose top and bottom, between chin ear cornors, between mouth corners, between cheekbones, between points on chin, between forehead center and nose bottom, between points on temples and width of nose. The similarity between two AD vectors is evaluated by converting the distance into a similarity measure as $s_{AD}(m, n) = e^{-\beta * d}$, where d is the 2D Euclidean distance between the AD vectors $\vec{m}$, $\vec{n}$, and β is a co-efficient that is chosen suitably to obtain a discriminative dynamic range of values. In our experiements, we set β to be 5.

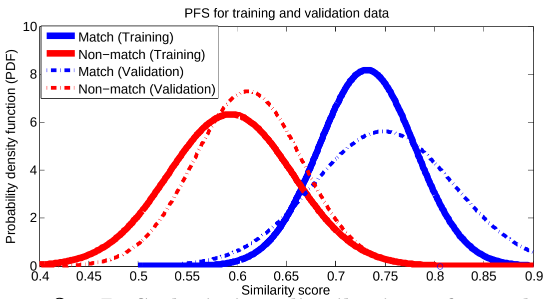

Figure 3: PFS depicting distribution of match and non-match scores. For training, match distribution was $G(.73, .048)$ and for non-match was $G(.59, .063)$. The validation curves for match was $G(.75, .055)$ and $G(.61, .071)$ for non-match. $G(\mu, \sigma)$ denotes Gaussian distribution with mean μ and standard deviation σ.

2.3 PFS Learning Framework

A set of portrait pairs authenticated to be of the same subject by our collaborators in art history are used as *training examples* to learn PFS and the remaining is used to validate it. We fuse scores obtained from LF and AD features of these images in a way such that the resulting distribution of match and non match scores are as peaked and disjoint as possible so as to enable efficient decision making in identification cases. Towards this, we employ the following methodology.

1. We consider a convex combination of the scores from the two measures LF and AD as $\lambda * s_{LF} + (1 - \lambda) * s_{AD}$, λ being varied from 0 to 1 in steps of 0.1.
2. For every λ, we evaluate the mean and standard deviation of match and non-match scores using the RANSAC algorithm [4] to prune outliers.
3. At each λ, we evaluate the Fisher linear discriminant function, $J = \frac{S_b}{S_w}$, where S_b is between class variance and S_w is within class variance. We choose that value of $\lambda = \lambda_{opt}$ that gives the maximum value of J.
4. The distributions of match and non-match scores, with λ_{opt} obtained in Step 3, for the combined (LF and AD) feature set are modelled as Gaussians distributions with means and standard deviations estimated from Step 2.

Validation of learned PFS: We perform two-fold cross validation on the set of images, i.e., we divide the set of instances into two groups F_1 and F_2. In fold 1, we learn the PFS from F_1 and validate on F_2. In fold 2, we learn PFS from F_2 and validate on F_1. μ and σ of match and non-match scores from two folds are averaged to obtain the resulting curves shown in Fig 3. It is to be noted that these curves are dependent on the data under consideration.

3. IDENTIFICATION EXPERIMENTS

Our collaborators in art history provided us 11 paradigms wherein we were required to verify whether subjects in two potraits were the same or not. These were (1) sculptural portrait of Sforza against a death mask by Laurana, (2) Sforza painitng against a death mask by Laurana, (3) portrait of a lady at the window against Lady with Primroses attributed to Verrocchio, (4) bust of a young woman attributed to Verrocchio against a de' Benci painting by Da Vinci, (5) bust of Strozzi by Fiesole against a painting by Campin, (6a,6b,6c) uncertain painted portrait compared to a drawing, photograph and painting of Mary Queen of Scotts, (7a,7b) uncertain painted portrait compared to painted portraits of James Scott by Kneller and Wissing, and (8a) uncertain photographic portrait compared to painting of Constanze Mozart by Lange.

3.1 Identification Framework

Given the learned PFS, the question now is to verify an unknown test image against a reference image. Towards this, we employ hypothesis testing.

Hypothesis Testing: This is a method for testing a claim or hypothesis about a parameter in a population [5]. Below, we summarize it with respect to the learned PFS in arriving at the conclusion for a match.

1. Null hypothesis claims that the match distribution accounts for the test's similarity score (with reference) better than non-match distribution. The alternate hypothesis is that non-match distribution models the score better.
2. We set level of significance α (test's probability of incorrectly rejecting the null hypothesis) as 0.05, as per common pratice in such problems.
3. We compute the test statistic using one independent non-directional z test [5], which determines the number of standard deviations the similarity score deviates from the mean similarity score of the learned distributions.
4. We compute p values which are the probabilities of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true. If $p \leq \alpha$, we reject the null hypothesis.

Identity Verification: In order to examine the validity of the chosen approach, we consider similarity scores of the

Reference		Distractors		Conclusion
Match	Non-match	Match	Non-match	
$p > \alpha$	$p \leq \alpha$	$p \leq \alpha$	$p > \alpha$	Match
$p \leq \alpha$	$p > \alpha$	$p < \alpha$	$p \geq \alpha$	Non Match
$p > \alpha$	$p > \alpha$	NA	NA	No decision
$p < \alpha$	$p < \alpha$	NA	NA	No decision
$p > \alpha$	$p \leq \alpha$	$p > \alpha$	$p \leq \alpha$	No decision
$p \leq \alpha$	$p > \alpha$	$p \leq \alpha$	$p > \alpha$	No decision

Table 1: Various possibilities for p values of test with reference and distractors. Different colors indicate different reasons (as explained in text) for reaching the corresponding conclusion. NA: Not applicable

test image with artworks known to depict different persons other than the one depicted in reference image. We call these images as distractors. Depending on availability, we choose similar works by the same artist (artist of reference image) as distractors. If a test image indeed represents the same subject as in the reference image, not only should its score with the reference image be modeled through match distribution, but also its scores with distractor faces should be modeled by non-match distribution.

3.2 Analysis Scenarios

We computed similarity scores of test cases with corresponding reference image and with 10 distractors. Table 1 lists various hypothesis test scenarios that can arise [5] and the corresponsing conclusions that one can infer. Match and non-match cases are straight forward to infer from Table 1. In cases where both match and non-match distributions are likely to account for the test data in the same way, it can be said that the learned PFS cannot accurately describe the test data (black rows in Table 1). If either match or non-match distribution is more likely to account for both test as well as distractors (magenta rows in Table 1), it can be inferred that the chosen features do not possess sufficient discriminating power to prune outliers. Thus in these scenarios, it is not possible to reach any conclusion.

3.3 Case Studies of Test Images

Table 2 lists the various p values for the identification attempts. Scores from 10 distractors are averaged to evaluate p values with match and non-match distributions. The following are the main categories of conclusions that can be inferred.

Match: Since p values satisfy conditions in blue row of Table 2, cases 1, 2, 3 and 8a are more likely modeled by match distribution.

Non match: Since case 4 satisfies conditions in red row of Table 2, it is more likely modeled by non-match distribution.

No conclusion: For cases 5, 6a, 6b and 7a both match/non-match distributions were accountable for the test scores obtained, and for cases 7b and 6c, distractors caused confusion.

4. CONCLUSIONS

We presented a work that explores the feasibility of computer based face recognition in identifying works of portriature. We provided experimental results validating the proposed approach. We also analysed the identification attempts against the backdrop of known hypotheses. We believe that these results can serve as a source of complementary evidence to the art historians. Future work will analyze larger datasets and model artistic styles.

Test	R_m	R_{nm}	D_m	D_{nm}	Conclusion
1	.485	.05	.05	.34	Match
2	.298	.002	.04	.15	Match
3	.764	.015	.04	.56	Match
4	.038	.562	.03	.5	Non-match
5	.322	.1552	.03	.55	No Decision
6a	.267	.187	.04	.4	No Decision
6b	.275	.18	.41	.3	No Decision
6c	.45	.05	.21	.02	No Decision
7a	.144	.28	.03	.5	No Decision
7b	.352	.05	.44	.03	No Decision
8a	.187	.3	.05	.56	Match

Table 2: p values for test cases. R_m and D_m denote p values of reference and distractor images with respect to match while R_{nm} and D_{nm} denote p values of reference and distractor images with respect to non-match distribution.

5. REFERENCES

[1] L. Farkas. International anthropometric study of facial morphology in various ethnic groups/races. *The Journal of Craniofacial Surgery*, 16(4):615–646, 2005.

[2] A. K. Jain, A. Ross, and K. Nandakumar. Introduction to biometrics: A textbook. *Springer Publishers ISBN: 978-0-387-77325-4*, 2011.

[3] J. Li, L. Yao, and J. Wang. Rhythmic brushstrokes distinguish Van Gogh from his contemporaries: Findings via automated brushstrokes extraction. *IEEE Trans. Patt. Anal. Mach. Intell.*, 34(6):159–176, 2012.

[4] M.Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. of the ACM*, 24(6):381–395, 1981.

[5] J. Shao. Mathematical statistics. *Springer New York,ISBN-13: 9781441929785, 2nd Edition*, 2012.

[6] D. Stork. Computer vision and computer graphic analysis of paintings and drawings: An introduction to the literature. *Proc. Comp Anal. Images. Patt.*, 2009.

[7] C. Tyler, W. Smith, and D. Stork. In search of Leonardo : Computer based facial image analysis of renaissance artworks for identifying Leornado as subject. *Human Vision and Electronic Imaging*, 2012.

[8] F. Vegter and J. Hage. Clinical anthropometry and canons of the face in historical perspective. *History of Facial Anthropometry*, 106(5):1090–1096, 2005.

[9] S. West. Portriature. *Oxford University Press*, 2004.

[10] L. Wiskott, J. Fellous, N. Kruger, and C. Malsburg. Face recognition by elastic graph bunch graph matching. *IEEE Trans. Patt.Anal. Mach. Intell.*, 7:775–779, 1997.

[11] W.Zhao and R. Chellappa. Face processing: Advanced modeling and methods. *Academic Press*, 2006.

[12] L. Zhifeng and A.K.Jain. Matching forensic sketches to mug shot photos. *Patt. Anal. Mach. Intell.*, 33(3):639–646, 2011.