https://doi.org/10.2352/EI.2023.35.13.CVAA-211

Computational tracking of head pose through 500 years of fine-art portraiture

Jean-Peïc Chou[a] and David G. Stork[b]

[a]Program in Symbolic Systems
Stanford University, Stanford, CA

[b]Independent Consultant
Portola Valley, CA

ABSTRACT

We used deep neural network image analysis to automatically extract head pose angles—roll, yaw, pitch (or tilt)—and figure display length (quarter-length, half-length, full-length) from 11,000 digital images of portrait paintings in a wide variety of styles, from the early Renaissance through Modern eras. We tracked trends and exposed anomalies in such formal properties of these portraits, and this sheds light upon the social and aesthetic forces to which portrait artists respond. For example, we find that the so-called Primitive or Naive portraitists favor a highly restricted range of pose angles (primarily frontal) while Expressionist, Mannerist, and Ukiyo-e portraitists employ a far greater range of angles. We also analyzed these formal properties to reveal the different trends throughout the careers of several individual artists, such as Paul Cézanne, Édouard Manet, and Francisco Goya. Our methods can be expanded to incorporate additional computed visual and contextual information—such as genders and ages of figures—and thus form a foundation for addressing a large range of problems in the history of art.

1. INTRODUCTION

Portraitists exploit numerous strategies in the service of their artistic goals, including setting the formal properties of head pose, specified by roll, yaw, and pitch angles, as illustrated in Fig. 1. Traditionally, formal portraits of political and cultural figures as well as matrimonial records from the early Renaissance favored figures in profile, gazing horizontally, as for example Piero della Francesca's *Duke of Urbino*, Antoniazzo Romano's *Profile Portrait of Cardinal Philippe de Levis*, and Leonardo's *La bella principessa*. Such poses often connote stasis, order, and clarity, and may help reveal the appearances and hence identities of the portrait subjects. Alternatively, head pose can be set for expressive purposes, to convey dynamism, motion, and psychological turmoil, as in Lucian Freud's *Reflection with two children (self portrait)*, Élisabeth Vigée Le Brun's *Woman artist in Revolutionary France*, Pablo Picasso's *Gertrude Stein*, and Elizabeth Sirani's *Allegory of painting*. Ukiyo-e woodblock portraits of kabuki actors frequently depict *mie* poses; here the head rotations enhance the drama at highly charged moments in kabuki plays.

How have the formal properties of portrait poses changed, broadly speaking, throughout the development of art? Do different art movements favor different head poses? Have some individual artists altered these formal properties systematically during the course of their careers? How might the answers to these questions help us understand the cultural, aesthetic, and associated forces shaping portraiture? These are the questions we address through deep neural network image analysis of portrait images.

Computational methods have been applied to the modeling of marks and artistic style in painted portraits, specifically from the Renaissance,[1,2] and can also reveal the geometry of arm placement in portraits.[3] More directly relevant to our study is that computational methods can be applied to large corpora of art images to reveal large-scale *trends* in formal properties of artworks, thereby revealing principles of composition, as for instance in landscapes.[4] Analogous trends based on color and form can be revealed in latent spaces learned through deep networks.[5] Similar methods can learn "emotion" (e.g., awe, fear, ...) in service of trend analysis in art.[6]

Send correspondence to `jeanpeic@stanford.edu` and `artanalyst@gmail.com`.

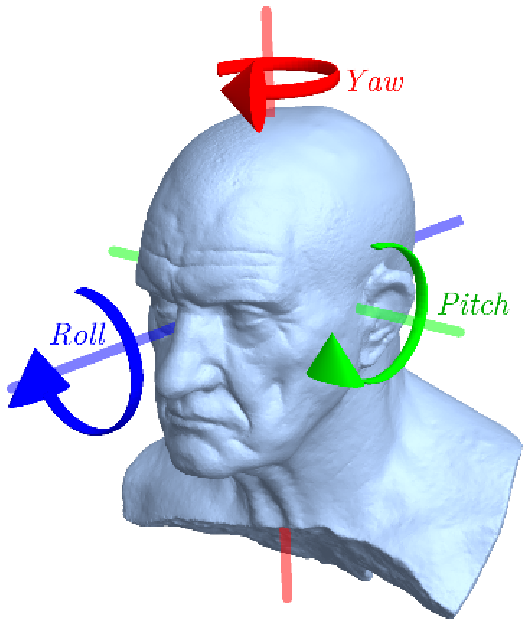


Figure 1. The posture of a portrait head can be expressed by three angles: roll, yaw, and pitch.

Such trend analyses can expand art scholarship in several ways. Changes over time in formal properties of artworks, such as portraits, may indicate changes in patronage, medium (e.g., from easel paintings to murals), social, political, and aesthetic forces, the stylistic influence of individual artists, and more.[7] Computational trend analyses often reveal information in new ways so that art scholars can better incorporate their knowledge of context in order to craft deeper interpretations supported by expanded classes of evidence.

We begin in Sect. 2 by describing our computational methods for extracting head pose of portrait figures and then present in Sect. 3 some of the trends and comparisons revealed by these methods. We mention briefly some guidelines to using these analytic methods in Sect. 4. We conclude in Sect. 5 with some broad lessons learned and suggestions for extensions and applications of our methods to problems elsewhere in the history of art.

2. DEEP NET HEAD POSE ESTIMATION

The computational pose estimation task is moderately challenging in the domain of high-quality natural portrait photographs but more difficult in the domain of portrait paintings in different media (oil, acrylic, watercolor, egg tempera, ...), natural and unnatural colors, and varieties of brush strokes. Portrait paintings are particularly problematic because artists often distort or caricature faces for expressive ends, as for instance in Francis Bacon's mangled and distorted *Self portrait* of 1969 and Juan Gris' *Portrait of Pablo Picasso*, to take just two of innumerable examples.

The rather obvious and well-established computational approach to similar problems—using very large training sets—simply will not work in most art analysis tasks such as portrait analysis.[8] The variety of existing styles would be difficult to account for in a single model and the total number of portrait artworks (perhaps a few tens of thousands of images) is far smaller than that for natural photographic portraits (many tens of billions of images). Moreover, there are currently no datasets of art portraits labeled with pose information equivalent to those for photographs.

In light of the lack of adequately large database of labeled portrait paintings, our approach is to test *two* existing computational pose estimation systems, each trained with natural portrait photographs, and rely on the system that proves most accurate for paintings.

The two computational methods for estimating head poses of portrait figures we tested are *Perspective-n-point* and *FSA-Net*.

2.1 Perspective-n-point

Here and throughout our work, we model the artist's rendering process as a projection of the three-dimensional subject's head onto the picture plane, as illustrated schematically in Fig. 2. The Perspective-n-point method relies on features or keypoints extracted from images of faces and finds the pose angles consistent with this evidence based on a generic 3D model of a head.[9] This first method has the advantage of offering some flexibility since a head rotation can always be averaged based on keypoints. In some cases where the painted face is unrealistic other end-to-end head pose estimation models would likely fail whereas this method would not.

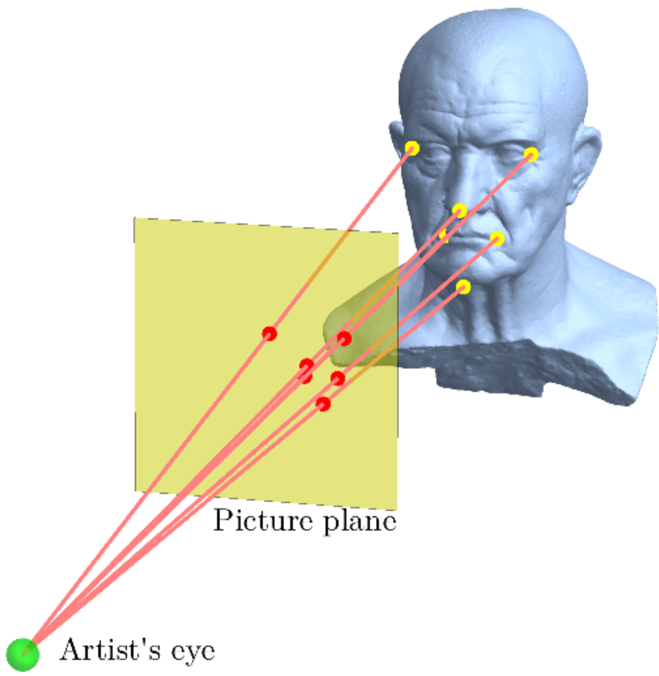


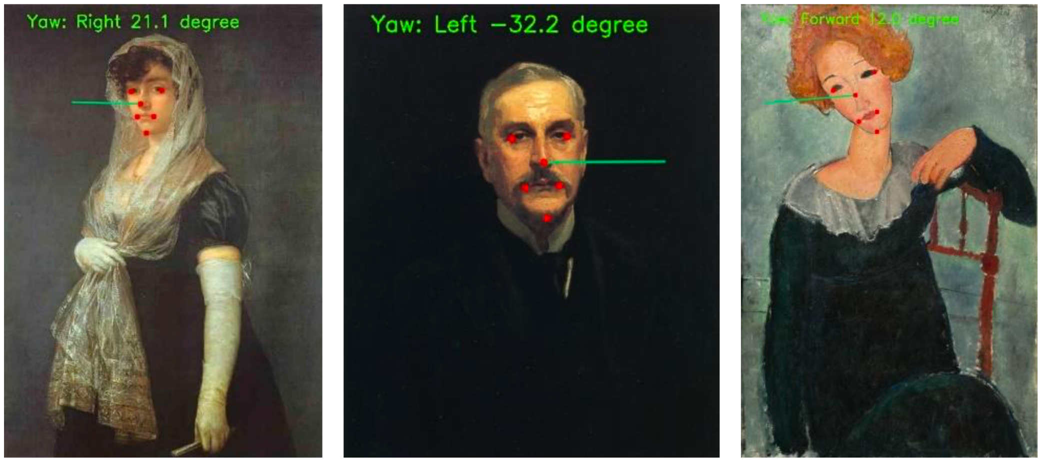
Figure 2. Throughout we modeled an artist's execution as a projection from the three-dimensional figure to the two-dimensional picture plane of the painting, where the center of projection is the artist's eye. Given the location of visual keypoints in the picture plane (red points) corresponding to the subject's eyes, nostrils, chin, and so forth, the Perspective-n-point method estimates the roll, yaw, and pitch of the subject.

More specifically, this method makes use of an algorithm from the *dlib* library that extracts face keypoints from each portrait based on a model pretrained with a large number of faces in natural photographs.[10] The alignment of the face's features from the painting and the 3D model is then computed using the *solvePnP* function in *OpenCV*.

Figure 3 shows example portrait paintings and yaw angles estimated by the Perspective-n-point method. Note that this method works quite well, even on the Modigliani portrait in the right panel despite the fact that the figure is not particularly realistic or three-dimensional. Note that this method often fails when a face is rotated (about any axis) so far that too few feature points are extracted to solve the angle estimation problem. We found, through small hand-tested cases, that pitch and roll were rather accurately estimated by this method, though the estimation of yaw was occasionally unreliable, primarily when the yaw angle was so large that a keypoint such as the outer corner of an eye became occluded.

2.2 Fine-Grained Structured Aggregation (FSA-Net)

The second estimation method we employed was FSA-Net, short for Fine-Grained Structured Aggregation network.[11] This method employs regression applied to aggregated visual features in the following way: feature maps are extracted from the images, much as in Perspective-n-point, but the keypoints are then aggregated into ever finer sets. The resulting feature vectors then serve as inputs to a regression model trained to yield the yaw, roll, and pitch of the portrait face.



21.1° −32.2° 12.0°

Figure 3. The Perspective-n-point method accurately estimates the pose (yaw is listed) for portraits in a wide range of styles, and is generally more accurate the larger the number of keypoints are visible.

The Fine-Grained Structural Aggregation method also provides reliable bounding boxes of the portrait faces, and such information permits the calculation of the percentage of the picture plane that is devoted to the portrait figure. We thus can compute automatically whether any given portrait contains just the head, or is half-length, or is full-length, or any intermediate proportion. Trends and anomalies in the proportion of figures portrayed in portraits can be analyzed much as are head poses, as we shall see below.

2.3 Data selection and preprocessing

Our portrait images come from the Kaggle challenge, *Painter by numbers*,[12] where the original images were mined from WikiArt.org. Each painting is annotated with title, artist, date of execution, genre (portrait, self portrait, landscape, ...), and style (art movement such as High Renaissance or Art Deco), as well as the image height and width, in pixels. There were $13,045$ portraits and $1,272$ self portraits in that dataset.

The dataset consists primarily of Western European easel paintings but also includes drawings, miniatures, prints, or other forms of pictorial representations. In this study, we chose to keep this diversity since we decided to focus on trends in general portraiture. We also noted that some pictures were cropped or extracted from larger images. We observed that these specific cases represented a very small proportion of the full dataset and did not present considerable modifications impacting the trends of our study. Moreover, some of the provided annotations were manually checked. Even though some annotated information, such as the art movements, was debatable, we assessed most of them as sufficiently reliable for a large-scale analysis.

Periods	1400-1450	1450-1500	1500-1550	1550-1600	1600-1650	1650-1700
Portraits	32	131	389	143	431	178

Periods	1700-1750	1750-1800	1800-1850	1850-1900	1900-1950	1950-2000
Portraits	119	1079	1302	2696	3225	431

Table 1. Number of portraits in our final culled database versus time, binned by half-century periods.

We found that the FSA-Net method failed to determine the head pose in roughly 18% of works in the above Kaggle images—specifically in $2,403$ portraits and in 208 self portraits. Furthermore, that method detected more than one figure in 820 portraits and in 147 self portraits. We eliminated such problematic and multi-figure works to better focus on statistical analysis of trends in poses of individual portrait figures. In this way, we were

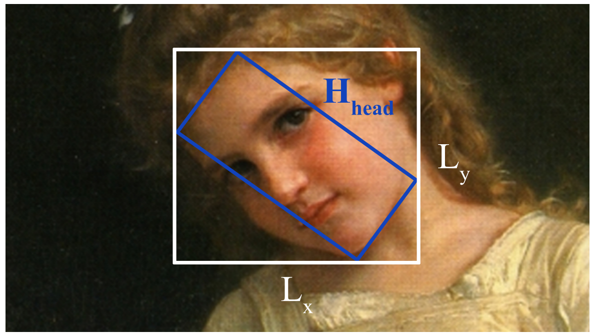


Figure 4. The FSA-Net method computes L_x and L_y, the horizontal and vertical dimensions of the extracted bounding box for the subject's head. Given the computed roll angle, the intrinsic size of the head, H_{head}, can be computed according to Eq. 1.

left with 9,822 portraits and 917 self portraits. Table 1 shows the number of portraits (including self portraits) in our final database, binned into half-century periods. The uneven distribution over time should be noted and will affect our analysis.

As mentioned, in addition to the yaw, roll, and pitch, the FSA-Net method provides the bounding box of the detected face. We use such a bounding box to determine the size of the head, the separation between the top of the head and the top of the painting, and the proportion of the body that is represented, relative to the height of the painting.

Figure 4 shows how the size of the head, H_{head}, is computed based on the size of the bounding box and the angle of the head roll, ϕ. A geometric analysis based on the figure reveals that the intrinsic height of the portrait head is governed by the extracted properties according to

$$H_{head} = \frac{L_y \cos(\phi) - L_x \sin(\phi)}{\cos(2\ \phi)}. \tag{1}$$

The proportion of the depicted body is computed by assuming that the size of the head is 1/8 of the whole body. We normalized this proportion to the range 0 to 1. The computation of the relative height of the head in the picture plane, given the location of the bounding box, is straightforward.

3. RESULTS AND TRENDS

The following are some of the principal trends and results displayed as box-whisker plots. Recall that at each entry in such a plot the box displays the middle quartiles of the empirical distributions, the horizontal line segment displays its average, and the whiskers show the full range of the distribution.[13]

3.1 Trends and properties in large, diverse corpora of portraits

We begin with two plots that illustrate our results and inferences generally, then move to more specific results, including trends in portraits executed by particular artists. The left panel in Fig. 5 shows a box-whisker plot of the pitch of our entire portrait corpus versus time, partitioned into half-century eras or periods. Notice that throughout the half millennium considered, the average pitch is slightly negative, that is, portrait heads tend to tip downward slightly.

This systematic negative average in the pitch angle may arise from a slight bias in estimation algorithm, though the developers of the FSA-Net method provide evidence countering that hypothesis.[11] It seems more likely that overall portrait subjects tend to tip their heads slightly downward or the heads are slightly below the height of the artist's eye. While this result deserves further analysis, most of our conclusions derive from the *variations* in angles, so some small putative bias is generally of minor importance to our conclusions, as we discuss in Sect. 4.

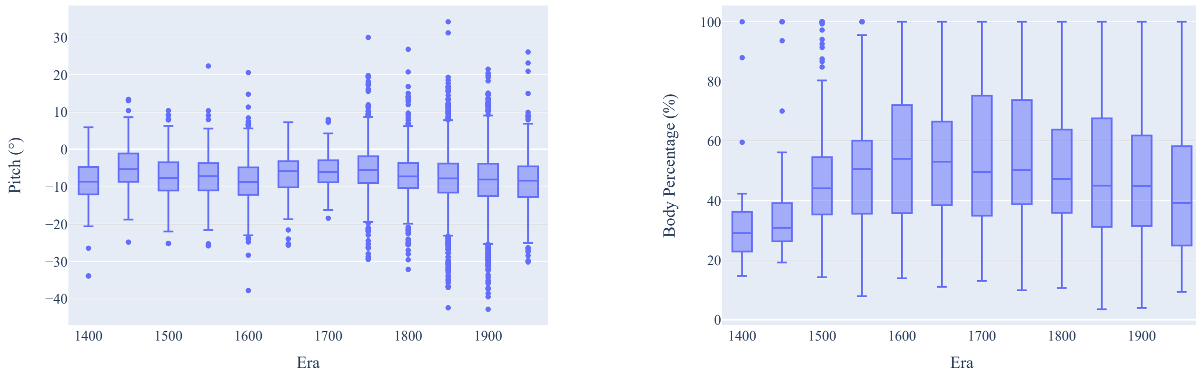


Figure 5. (L) The pitch angle, measured in degrees, of all portraits in our database versus time, partitioned into half-century periods. Notice that the average pitch throughout the half millennium is consistently negative, most likely due to near-universal stylistic strategies. (R) The distribution of the proportion of subjects' bodies depicted in portraits versus time. The late Medieval and early Renaissance portraitists seem to favor low proportions of bodies (that is, showing primarily heads) while the proportion increases and peaks near the beginning of the 18th century.

The right panel in Fig. 5 reveals the unimodal trend in the proportion of the body represented in portraits over time, where the broad peak occurs near the beginning of the 18th century. There are of course many personal, aesthetic, cultural, political, and financial forces that affect the ultimate appearance of any portrait. Nevertheless plots such as in Fig. 5 and those below prompt hypotheses that can be explored and tested through further analysis. For example, the data suggests that early portraits focused primarily on just faces because early artists, less versed in the techniques of accurate realism that would arise later, found that the identities were hard to convey and so it was important to devote as much of the picture plane as possible to the face. One relevant trend revealed by the right panel in Fig. 6 is that the range and variance of the distributions increase somewhat over the half-millennium period. It seems likely that as artists experiment with new styles, media, and compositional strategies that the range of body proportions will increase.

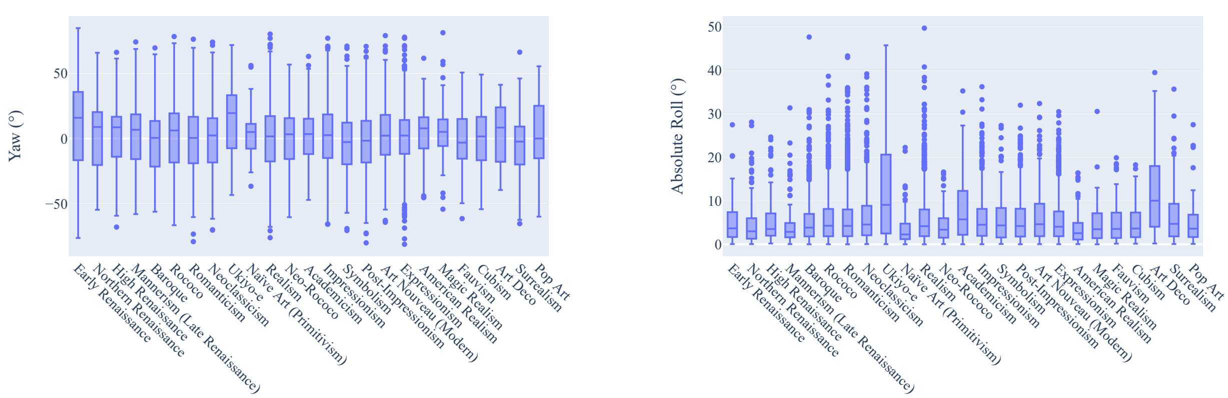


Figure 6. (L) Yaw angle partitioned by style or art movement. Notice here and below that the large and statistically significant difference between Naive Art and Ukiyo-e, for example. (R) Roll angle, partitioned by style or art movement. Notice that the average roll angle, broadly speaking, is rather small, that is, portrait subjects rarely tip their heads side-to-side by large angles. Many of the portraits with large roll angles depict subjects reclining or lying supine.

Figures 6 shows yaw and roll angles, partitioned by style (as labeled in the WikiArt.com dataset). Particularly salient are the statistical differences between the distributions for portraits in Ukiyo-e and so-called Primitive or Naive Art, as practiced by Henri Rousseau, Edward Hicks, Paul Gaugin, Marc Chagall, Frida Kahlo, and others.

Ukiyo-e portraits frequently depict geishas and actors, the latter often in dramatic *mie* poses at key moments in kabuki plays. Here the large rotation angles of the head enhance the emotional power of these moments. By contrast, portraits by Naive artists are often fairly direct and easy to read. After all, nearly all artists find depicting portrait heads in dynamic, tilted, or rotated poses to be a technical challenge—a challenge not widely attempted by Naive artists, many of whom were self taught.

The left panel in Fig. 7 shows the absolute value of the yaw computed in 11,000 portraits, grouped by art movement. While most movements lead to similar distributions of yaw angles, a few movements stand out. Most conspicuous is that of all movements Naive or Primitive art has both the lowest average yaw angle and the smallest variance. In short, Primitivist artists such as Henri Rousseau strongly favor frontal portraits. Ukiyo-e portraits, such as by Katsushika Hokusai and Utamaro Kitagawa, reveal large average and large variance in yaw angle (as well as roll and pitch, not shown).

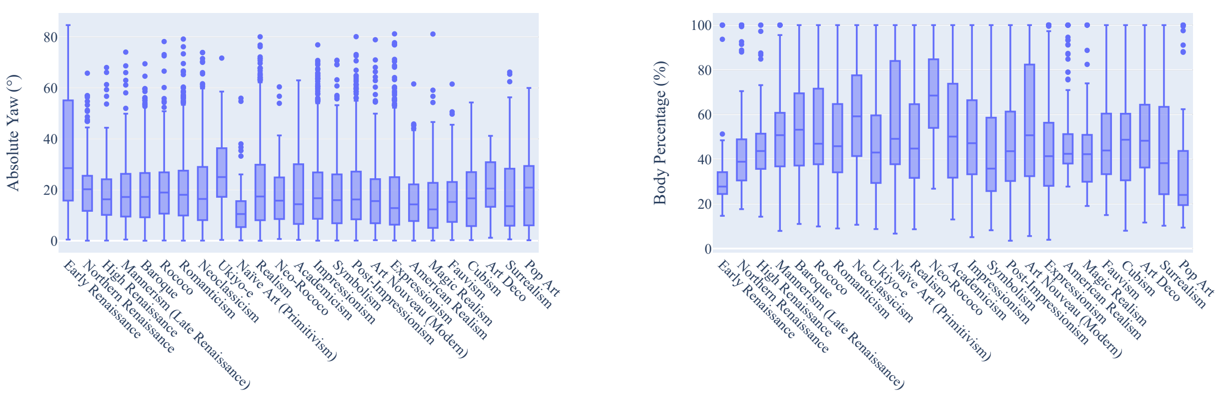


Figure 7. (L) The absolute value of yaw, grouped by art movement. Notice that the distribution of yaw rotations in Naive or Primitivist art has a small average and small variance, while Ukiyo-e portraits span a large average and large variation. (R) The proportion of figures' bodies depicted within portrait paintings, partitioned by style or art movement. Note that this proportion is rather low during the High Renaissance (which focused on portrait heads), while rather large in Neo-Rococo.

While the differences are more subtle between other movements, other features revealed by the plot of yaw angle by style may prompt further hypotheses, inquiry, and analyses. For example the average yaw angle and especially the variance for Art Deco portraits is lower than those for other styles, generally speaking, and this may be due to the stylistic conventions of that movement which stressed relatively flat graphic presentations for which large yaw angles might lead to portraits that are difficult to read.

The plot in the left panel in Fig. 7 is representative of trends in other formal properties of portraits in our dataset. The small variances found in distributions such as yaw angle in late Medieval portraiture and Primitivist art are consistent with the more constrained nature of these styles, while the large variances found in Ukiyo-e portraits are consistent with their greater concern with expressing emotionally charged moments from kabuki plays.

Figure 7 shows the distribution of the proportion of figures' bodies depicted within artworks for over a half millennium, divided into half-century periods. Here a small portion would be just the head, a larger portion a half-length portrait, and the largest portion a full-length portrait. Note the rather small portion in the late-Medieval era, and the gradual rise and fall up to the Modern era.

Figure 8 shows the proportion of figures' bodies shown within the portrait frame (viz., half-length, full-length, ...), and the height of the center of figures' heads within the painting frame, partitioned by style or art movement. While this proportion is rather small for portraits from the early Renaissance (1400–1450), it remains rather large throughout the grand sweep of art in other art movements.

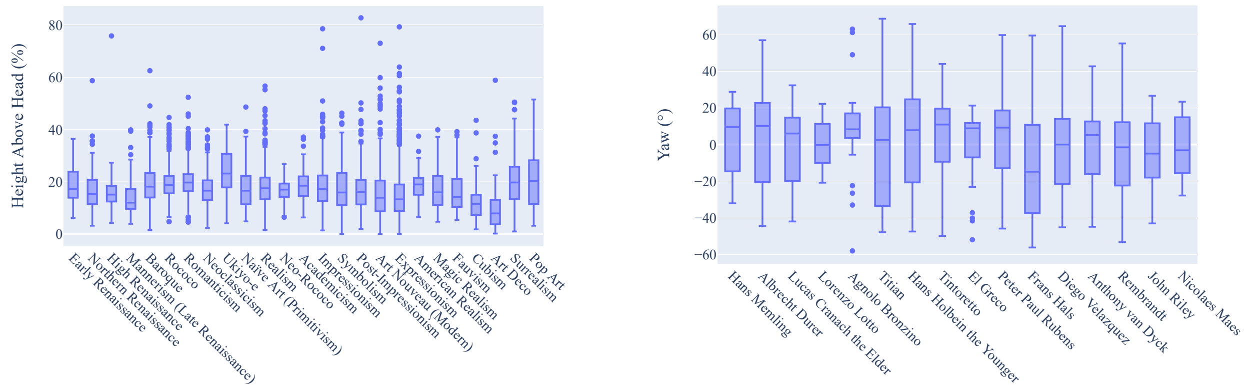


Figure 8. (L) The proportion of the picture plane above portrait subjects' heads is fairly constant across all art movements, suggesting there are broad, nearly universal design principles guiding the this aspect of the composition of portrait paintings. (R) The distributions of yaw angle partitioned by artist. Notice, for example, that El Greco and Lorenzo Lotto have small mean angle as well as a small variance. In other words, their portrait subjects generally face directly toward the artist. By contrast, Rembrandt and Titian depict their portrait subjects in a very wide range of such angles.

3.2 Pose distributions by artist

We now turn to some trends in the careers of individual artists. Table 2 shows the number of portraits by artist, those with sufficiently many portraits that trend analysis is statistically valid.

Artist	Franz Hals	Rembrandt	Diego Velázquez	Titian
Portraits	115	88	80	79

Artist	Lucas Cranach	Anthony van Dyck	Hans Holbein	Peter Paul Rubens
Portraits	76	70	62	60

Artist	Albrecht Dürer	Agnolo Bronzino	John Riley	Tintoretto
Portraits	38	36	35	34

Artist	Lorenzo Lotto	El Greco	Hans Memling	Nicholaes Maes
Portraits	28	26	23	21

Table 2. The number of portraits in our database for the most-represented artists before 1700, listed in descending order of the number of portraits.

The right panel in Fig. 9 shows the distribution of yaw angles of different artists of the 17th century. Note that variance for the yaw angle for El Greco is relatively small and the quartiles are rather close to zero, indicating that his figures tend to face directly toward the artist. The left panel in Fig. 9 shows the proportion of portrait subjects' bodies depicted by different artists. Notice that some artists, such as Hans Memling and Albrecht Dürer, tend to favor heads and shoulders while others, such as Anthony van Dyck, favor full-length portraits including of aristocracy and royalty standing confidently in imposing body postures.

Our methods can be applied to individual artists to reveal *trends* in their compositional strategies throughout their careers. The right panel in Fig. 10 shows, for instance, the absolute yaw angle of portrait subjects during five stages in the career of Diego Velázquez. There is a very clear trend to smaller yaw angles at later periods. It

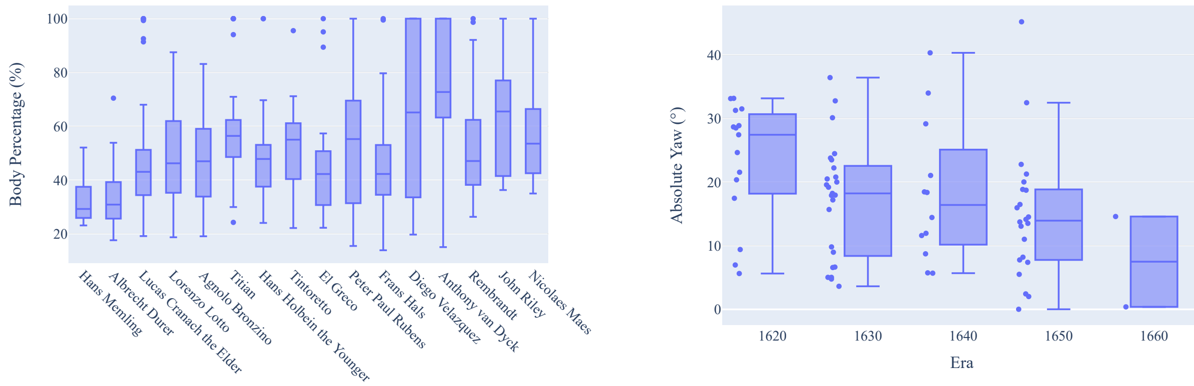


Figure 9. (L) The proportion of a portrait subject's body depicted by different artists. This plot reveals that painters such as Hans Memling and Albrecht Dürer focused on just heads and shoulders whereas Peter Paul Rubens, Diego Velázquez, and Anthony van Dyck focused more on full-length portraits. (R) The absolute yaw in the portraits of Diego Velázquez during his career, from 1620 through 1660. Notice how figures increasingly face the artist, that is, the absolute yaw angle decreases during later stages of this artist's career. This plot then reveals the fact that Velázquez's late portraits of Spanish royalty were more formal than were his earlier works.

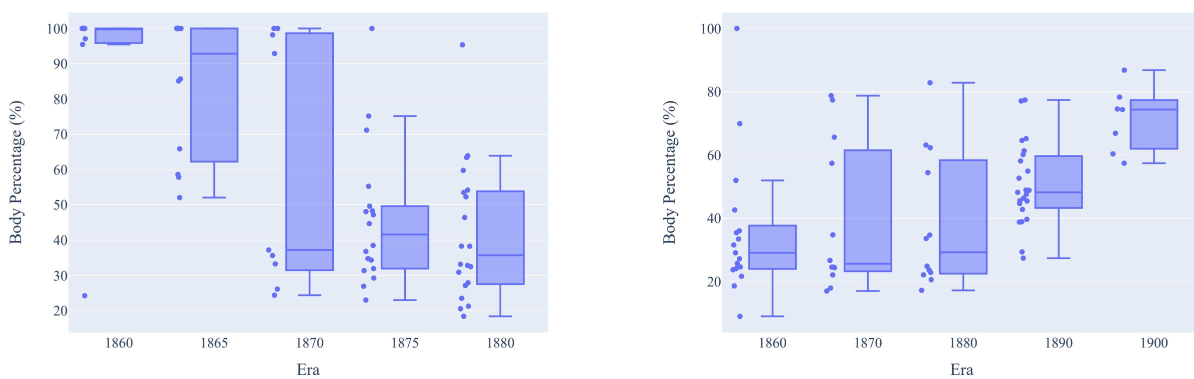


Figure 10. (L) The proportion of figure bodies appearing in portraits by Édouard Manet. In addition to the broad trend of decreasing body proportion throughout Manet's career, we find a clear bimodal distribution of these proportions at the middle of his career (1870). The individual points in the plot shows that during this period his portraits were most likely either full length or of just heads. (R) The trend in the body proportion in the portraits by Paul Cézanne show the opposite trend of those by Manet. Cézanne increasingly favored full-length portraits toward the end of his career.

is unclear the reason for this trend, but one hypothesis deserves consideration. Early in Velázquez's career, when he was in Sevilla, he depicted common folk as in *An old woman cooking eggs*. Later, after 1631, he joined the Royal Court as Royal Usher in Madrid's Alcazar Palace, where he painted many formal portraits of King Philip IV, Queen Mariana, Doña Margarita Theresa of Austria, and others, including in his towering masterpiece *Las meninas*. His formal and official portraits nearly universally depicted subjects with heads upright, as befits the context and as reflected in the data.

The left panel in Fig. 10 shows the proportion of depicted portrait bodies throughout the career of Édouard Manet. This artist focuses primarily on just heads at the end of his career. Perhaps additional research will reveal the personal, social, and aesthetic forces that led to this development throughout his career, so clearly evident in the data.

A statistically significant change in the proportion of subjects' bodies in portraits of Paul Cézanne, as is clear in the right panel in Fig. 10. His earliest portraits, from around 1860, showed for the most part just the head

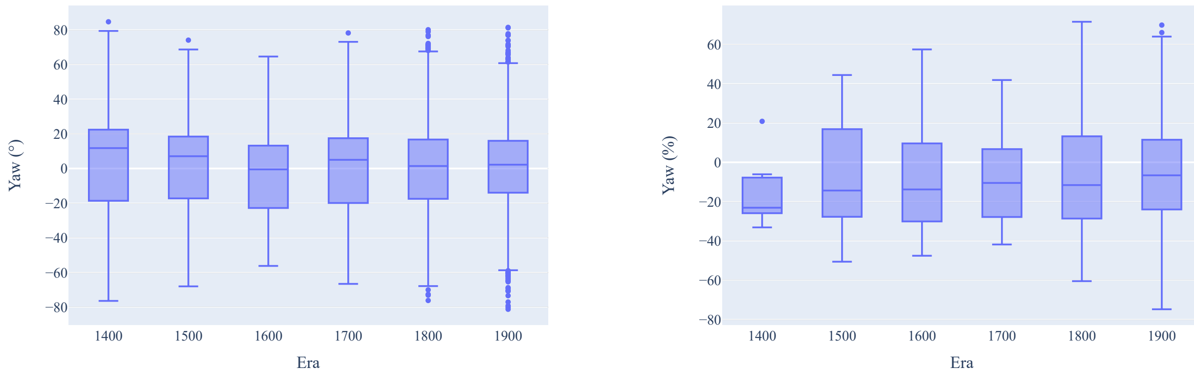


Figure 11. (L) Except for the 15th and 16h centuries, the yaw angle in portraits is centered around 0°, not showing any specific trend. (R) Yaw in self portraits versus time. Notice the overall negative yaw angle, which is consistent with the majority of artists being right-handed.

and possibly chest, while his latest portraits, from around 1900, typically showed nearly the full figure.

3.3 Trends and properties of self portraits

The above analyses only treated portraits excluding self portraits. Our trend analysis reveals, however, a few formal differences between these two sub-categories.

One noteworthy result in self portraits concerns yaw angle. In self portraits this yaw angle is statistically largely negative, while for regular portraits the yaw is statistically indistinguishable from neutral, that is, centered on 0°. It is natural to ascribe this statistical difference to hand dominance among artists. Right-handed artists would naturally paint themselves facing to the right in order to prevent their painting hand from obstructing their field of vision. The yaw of four artists—Francisco Goya, Fyodor Bronnikov, Henri Fantin-Latour, and Lucian Freud—are atypical in that their yaw angles are negative, which is consistent with them being left handed. Indeed, a nude self portrait by Lucian Freud *Painter working, reflection* of 1993—often considered his greatest self portrait—shows him holding his brush at the left, as viewed in a plane mirror reflection, which corresponds to his left hand. Likewise, self portraits by Francisco Goya and Fyodor Bronnikov show these artists were left-handed as well. Henri Fantin-Latour's sole self portrait suggests he is right-handed, however. Note, though, that this painting is from early in his career, perhaps when he was still learning and experimenting with his technique.

Figure 11 shows the yaw angle in portraits excluding self portraits versus period, grouped by century. There is no statistically significant change or anomalies evident in this data, a result that comports with the data found in self portraits, which we shall see. The right panel shows the yaw angle but in just self portraits. Here we see a slight difference between these distributions, possibly attributable to an offset in yaw due to the fact that most artists are right handed, as described above.

Figure 12 compares the distributions of yaw angles in portraits by four artists, Francesco Goya, Fyodor Bronnikov, Henri Fantin-Latour, and Lucian Freud. The distributions reveal a statistically significant positive yaw angle while the others we studied have a negative yaw angle. This is evidence consistent with these artists being left handed.

4. BIAS, VARIANCE, AND ANALYSIS STRATEGIES

The above graphs and analyses are representative of the new views into art corpora offered by our computational methods. Of course, they are not exhaustive, and in Sect. 5 we shall mention some techniques for expanding and enhancing them. Nevertheless, the above are sufficient for us to explore more deeply methodological recommendations.

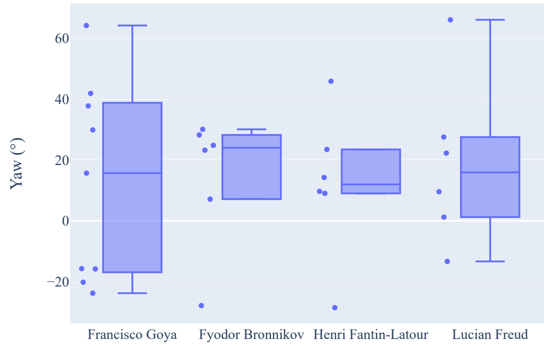


Figure 12. The yaw angle in the four artists in our database having positive average yaw. Such a result is consistent with these artists being left-handed. Indeed, the full self portraits by Francesco Goya, Henri Fantin-Latour, and Lucian Freud show they were indeed left handed.

An important issue is reducing bias in such statistical analyses of fine art.[14] Recall that bias is a systematic difference between the estimate of a value and its true value.[8] For instance, we saw in Fig. 5 evidence that the Perspective-n-point method might yield a bias in the estimation of pitch. One benefit of working with trends in large corpora is that even if there may be a bias associated with some property, that bias will rarely if ever effect slopes associated with *trends* over time. In short, our scholarly conclusions based on such trends and *differences* are likely unaffected by potential bias. Nevertheless, of course we must always be alert to sources of bias in our methods and reduce them to the extent possible.

A closely related benefit of the class of results produced by our methods concerns estimation variance and associated statistical significance. A deep result from statistics is the bias-variance tradeoff, which states that under very broad conditions, an estimation method that has low bias will generally have high variance, and vice versa. We choose estimation methods that, for our problem, will have low bias, in order to find the true value of the parameter in question. The resulting variance, which can often be read directly from our results, then reveals how confident we can be in our estimated value. In some cases, a putative result is not statistically significant, in which case we cannot make firm conclusions about the art historical question at hand.

5. CONCLUSION AND FUTURE DIRECTIONS

We have demonstrated that deep network based face analysis methods, trained using natural photographs, can reliably estimate pose in fine-art portraits despite the fact that they are in a wide range of styles. Trends and anomalies in such formal properties of portraits—represented graphically and described statistically—provide new views to art scholars, who can analyze such results in context. Our goal here has not been to make strong conclusions about the sources of these trends but instead to inform the community of art scholars of the power (and modest limitations) of our methods.

Our methods can be refined and augmented in a number of ways. For instance, our methods can be augmented with reliable computational gender-identification methods, so trends in pose for male and female subjects can be analyzed separately and compared.[15] An analogous extension would be to incorporate computational estimation of subjects' ages, in order to analyze trends in portraits as a function of age.[16] A rather straightforward extension to our methods would be to include the estimates of the location of the head within the picture plane together with estimates of the rotation angles. We note that there is strong statistical evidence that artists place an eye of portrait subjects on the vertical line bisecting the picture plane.[17] In this way we can infer whether a subject is looking toward the center of the picture plane, or instead outside the picture plane. This overall approach, applied to multi-figure works, might shed light on trends in gaze throughout the history of art.

Additional future directions include training systems with *art* images, rather than natural photographs—a technique that has improved performance on other art-analysis tasks, such as segmentation.[18,19] The challenge here is to get a sufficiently large corpus of labeled art images that can be expanded through semi-supervised learning. Such protocols have been demonstrated for problems in art analysis through non-expert human feedback through Mechanical Turk.[6]

After algorithmic refinement and extensions, our work might serve as a foundation for a standalone software tool for art scholars, complete with natural and intuitive interface to tailor image properties to be estimated, easy mounting of image databases, natural graphical and statistical measures of results, and so on. Such an effort would further demonstrate the value of rigorous computational image analysis in the study of fine art.

Acknowledgments

The first author acknowledges the support of Professor Maneesh Agrawala of the Stanford Computer Science Department, and the second author thanks the Getty Research Library and its staff, where some of this research was performed.

REFERENCES

1. R. Sablatnig, P. Kammerer, and E. Zolda, "Hierarchical classification of painted portraits using face- and brush stroke models," in *Proceedings of the 14th International Conference on Pattern Recognition*, A. K. Jain, S. Venkatesh, and B. C. Lovell, eds., **I**, pp. 172–174, IEEE Press, (Los Alamitos, CA), 1998.
2. R. Srinivasan, A. Roy-Chowdhury, C. Rudolph, and J. Kohl, "Quantitative modeling of artist styles in renaissance face portraiture," in *Proceedings of the 2nd International Workshop on Historical Document Imaging and Processing*, pp. 94—101, Association for Computing Machinery, (New York, NY, USA), 2013.
3. S. Liu, J. Yang, S. S. Agaian, and C. Yuan, "Novel features for arm movement classification of portrait paintings," *Image and vision computing* **108**, p. 104121, 2021.
4. B. Lee, M. K. Seo, D. Kim, I. seob Shin, M. Schich, H. Jeong, and S. K. Han, "Dissecting landscape art history with information theory," *Proceedings of the National Academy of Science* **117**(43), pp. 26580–26590, 2020.
5. A. Elgammal, B. Liu, D. Kim, M. Elhoseiny, and M. Mazzone, "The shape of art history in the eyes of the machine," in *AAAI Conference on Artificial Intelligence*, pp. 2183–2191, 2018.
6. P. Achlioptas, M. Ovsjanikov, K. Haydarov, M. Elhoseiny, and L. Guibas, "ArtEmis: Affective language for visual art," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11569–11579, 2021.
7. S. West, *Portraiture*, Oxford University Press, Oxford, UK, 2004.
8. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, John Wiley and Sons, New York, NY, 2 ed., 2001.
9. J. Yaniv, Y. Newman, and A. Shamir, "The face of art: Landmark detection and geometric style in portraits," *ACM Transactions on Graphics* **38**(4), 2019.
10. C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge," in *IEEE International Conference on Computer Vision Workshop*, pp. 397–403, IEEE, 2013.
11. T. Yang, Y. Chen, Y. Lin, and Y. Chuang, "FSA-Net: Learning fine-grained structure aggregation for head pose estimation from a single image," in *IEEE Convference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1087–1096, IEEE, 2019.
12. https://www.kaggle.com/c/painter-by-numbers, Kiri Nichol.
13. L. Wasserman, *All of statistics: A concise course in statistical inference*, Springer, New York, NY, 2010.
14. Z. Zhang, J. Li, D. G. Stork, E. Mansfield, J. Russell, C. Adams, and J. Z. Wang, "Reducing bias in AI-based analysis of artworks: Overview and tutorial," *IEEE BITS* **2**(1), pp. 36–48, 2022.
15. Y. Akbulut, A. Şengür, and S. Ekici, "Gender recognition from face images with deep learning," in *International Artificial Intelligence and Data Processing Symposium (IDAP)*, pp. 1–4, IEEE, 2017.
16. Y. Dong, Y. Liu, and S. Lian, "Automatic age estimation based on deep learning algorithm," *Neurocomputing* **187**, pp. 4–10, 2016.
17. C. W. Tyler, "Eye placement principles in portraits and figure studies over the past two millennia," in *Human Vision and Electronic Imaging III*, **3299**, pp. 431–438, 1998.

18. T. Heitzinger, M. Wödlinger, and D. G. Stork, “Improving semantic segmentation of fine art images using photographs rendered in a style learned from artworks,” in *Computer vision and analysis of art*, D. G. Stork and K. Heumiller, eds., 2022.
19. T. Heitzinger and D. G. Stork, “Improving semantic segmentation of fine art images using photographs rendered in a style learned from artworks,” in *Computer vision and analysis of art*, D. G. Stork and K. Heumiller, eds., IS&T, 2022.