Style and Abstraction in Portrait Sketching

Itamar Berger[1,2] Ariel Shamir[1,2] Moshe Mahler[2] Elizabeth Carter[2,3] Jessica Hodgins[2,3]

[1]The Interdisciplinary Center Herzliya [2]Disney Research Pittsburgh [3]Carnegie Mellon University

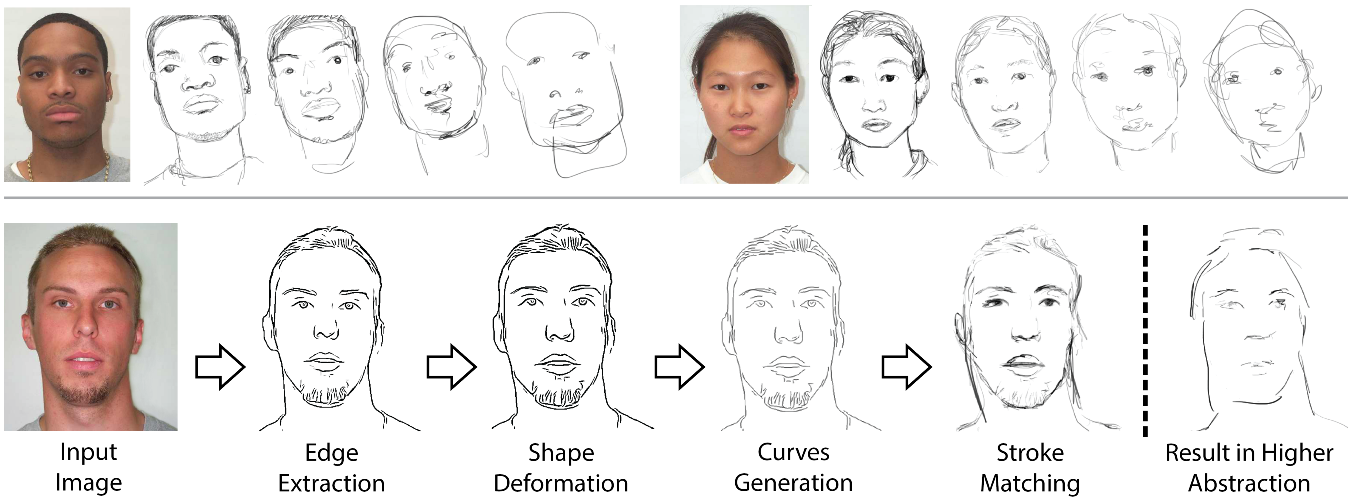


Figure 1: *By analyzing sketch data gathered from artists in various levels of abstraction (top: two examples from one artist), we can synthesize a sketch portrait from a photograph. Our analysis of both shape and strokes supports the process of generating a sketch in this artist's unique style at different levels of abstraction. We follow this artist's style both in terms of the stroke appearance and in terms of the shape – drawing larger jaws and moving the eyes higher.*

Abstract

We use a data-driven approach to study both style and abstraction in sketching of a human face. We gather and analyze data from a number of artists as they sketch a human face from a reference photograph. To achieve different levels of abstraction in the sketches, decreasing time limits were imposed – from four and a half minutes to fifteen seconds. We analyzed the data at two levels: strokes and geometric shape. In each, we create a model that captures both the style of the different artists and the process of abstraction. These models are then used for a portrait sketch synthesis application. Starting from a novel face photograph, we can synthesize a sketch in the various artistic styles and in different levels of abstraction.

Links: DL PDF WEB

1 Introduction

Visual abstraction has been used throughout history as a technique to communicate information more effectively and more efficiently – highlighting specific visual features while downplaying others. For example, in one of the most famous examples of abstraction, Pablo Picasso (1881-1973) created a suite named 'bull' containing eleven lithographs presenting gradual visual abstractions of a bull through progressive analysis of its form. Understanding the process of abstraction is not only interesting from an artistic point of view, but it can also assist in designing better artificial drawing tools and rendering programs by informing us about how information can be most effectively presented.

A general study of visual abstraction is too broad as every piece of art uses some level of abstraction to depict its subject, and there are endless methods and styles in art. We focus our study on a simple, yet important, domain: sketches of the human face. More specifically, we use a data-driven approach to study the process of abstraction, by gathering and analyzing sketches of faces at various levels of abstraction from seven artists. We asked them to sketch a portrait of a face from a reference photograph using time intervals decreasing from four and a half minutes to fifteen seconds.

As expected, the data gathered does convey a progression from more realistic to more abstract sketches as time decreases (Figure 1 and 2). However, the data also contains clear differences in the style of the different artists. In fact, the data we collected expresses a multi-dimensional space spanned by the abstraction level, the style of the artists, and the different subject faces (i.e. the 'content' itself). Using such data, we are able to study and build models describing both the process of abstraction and the elements of style. Although both are very intuitive to grasp perceptually, they are extremely difficult to define algorithmically.

To build models of abstraction and style, we analyze both the characteristics of the strokes and the differences between the shape of the faces and the reference photographs. This analysis reveals characteristic alterations that the artists make to the geometric shape of the face and not just their line depiction styles. Using our modeling of abstraction and style, we are able to synthesize new sketches from photographs at various levels of abstraction and with a style that approximates the stroke and shape interpretation of the individual artists whose drawings we captured. We also validate our results with a user study.

At the strokes level, we build a database of all strokes used by a specific artist and classify them to three major categories: shading strokes, complex strokes and simple strokes. We measure various curve characteristics such as spatial and temporal distribution, overlapping and length to analyze both style and abstraction. For synthesis purposes, we build a strokes library indexed by curvature and shape context descriptors [Belongie et al. 2002].

At the shape level, we fit a face mesh model to define the structure of the face and match its facial features on both the sketches and the input photographs. This procedure provides a correspondence between the true geometry of the face and the artists' interpretation in the sketches. We use statistical modeling akin to Active Shape Models (ASM) [Stegmann and Gomez 2002] to study the shape variations for a specific artist in the different levels of abstraction.

We demonstrate the use of these characterizations for portrait sketch synthesis: converting photographs to realistic sketches in a given style and abstraction levels. We also provide our sketch dataset for future research.

2 Previous Work

The human vision system is especially tuned to recognize and understand human faces [Sinha et al. 2006]. As a result, depiction of human faces has long been a fertile and challenging subject of research in graphics and image processing. Related topics include facial illustrations [Gooch et al. 2004], forensics [Zhang et al. 2010], portrait painting [Zhao and Zhu 2011; Tresset and Leymarie 2013], cartoonizing [Chen et al. 2002; Chen et al. 2004; Meng et al. 2010], and caricaturization [Chiang et al. 2004; Yu and Zhang 2010; Le et al. 2011; Liang et al. 2002]. In this section, we focus on the previous research in synthesizing portraits from photographs with an emphasis on work that has taken a data-driven approach.

Analysis of artist's sketches: Cole and colleagues [2008] analyzed where artists draw lines in sketches of inanimate objects such as bones, tools, and automobile parts, and found that artists are largely consistent in where they choose to draw lines, and that they focus on contours first and shading second. We also separate the strokes to contour strokes and shading strokes but our drawing task was less constrained and we found significant differences in the placement of the strokes, especially at the higher levels of abstraction.

Eitz and colleagues [2012] analyzed a much larger set of drawings by non-experts (20,000) and developed a visual feature descriptor for matching the sketches to the 250 object categories that inspired them. Because their goal was recognition rather than synthesis, they used the distribution of line orientation within a small local region as the primary feature. Limpaecher and colleagues [2013] collected and analyzed 13,000 drawings of faces using an iPhone game. Their goal is to allow auto-correcting of strokes for beginning artists. Unlike our work, they collect only registered faces, and do not study geometric distortion or artistic style.

Synthesizing sketches of faces: Gooch and colleagues [2004] created sketches from photographs by computing the brightness and the luminance, thresholding each and combining them back into a single image. This approach created sketches that were a close match to the input photograph and they verified that the speed and accuracy of recognition and learning was not degraded for the sketches. Chen and colleagues [2002] used an example-based approach to generate sketches from photographs by building a non-parametric matching between the photographs and the example sketches. The parameters of the strokes that formed the sketch could then be controlled by the user. Wang and colleagues [2009] created sketches from photos by building multi-scale MRF models to map between pairs of photographs and sketches. In contrast to Chen's approach as well as ours, Wang's approach is based on textures rather than strokes, and thus creates a sketch that very closely resembles the details of the photograph. There are a number of competing approaches for performing this matching including partial least-squares (from photo/caricature pairs) [Liang et al. 2002], semi-coupled dictionary learning (from photo/sketch pairs) [Wang et al. 2012], and feature-level nearest neighbor approach [Chang and Cheng 2011; Liu et al. 2010]. Lastly, Aikon is a robotic portrait sketching system [Tresset and Leymarie 2013] that converts photographs to real sketches but it is still described as a "naïve drawer" not capable of learning different styles or abstraction.

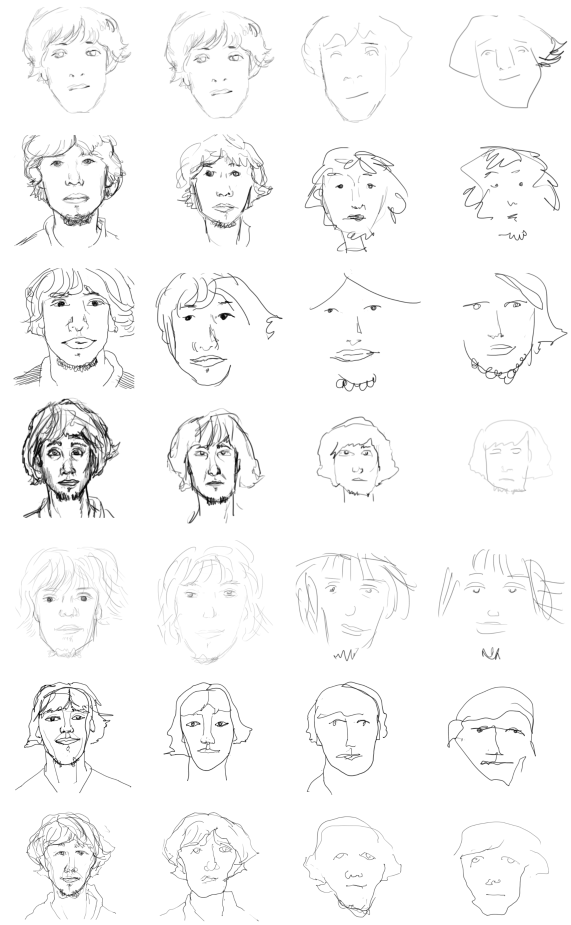


Figure 2: *A slice in our input dataset: each artist sketched face portraits of the same subject at increasing abstraction levels (in the electronic version of this paper you can zoom in to this and other figures for better view).*

Mimicking a particular style: Much work in NPR has focused on mimicking a particular style (see [Kyprianidis et al. 2013] for a survey). Style is most often described as being composed of two parts: geometry/shape and rendering (textures or strokes).

For strokes, Hertzmann and colleagues [2002] learn a statistical model of a 2D curve, while [Freeman et al. 2003] use linear combination of a set of given curves in a specific style. “The painting fool” program also presents several simulated paint/pencil/pastel strokes styles [Colton 2010]. More recently [Kalogerakis et al. 2012] learn the hatching style of a given example and are able to synthesize new illustrations in this style. In our approach, we do not learn a parametric model for stroke styles but directly use (modified) strokes from the input set. Moreover, we also address different abstraction levels which have not been dealt with in these works. A similar approach of transferring strokes from an input portrait drawing for synthesis was presented in [Zhao and Zhu 2011]. They use templates to fit a mesh to the face and transfer a set of strokes from a specific painting to a given photograph. However, their objective is not to learn a model of the style of an artist – not at the geometric nor the strokes level, and they do not tackle abstraction.

For geometry, Liang and colleagues [2002] learn a model of geometric exaggeration from pairs of photos and caricatures, and Lu and colleagues [2012a] use strokes and tone to represent shape and shading respectively. The style of an artist’s strokes is mimicked using shape context and filtered velocities as the features for matching in work by Lu and colleagues [2012b].

Synthesizing abstracted drawings: Abstraction is required to represent a photo in a different media such as brush-strokes [Kyprianidis and Kang 2011] or with shape simplifications [Kang and Lee 2008; DeCarlo and Santella 2002]. Geometric models have been abstracted with individual identifying features preserved [Yumer and Kara 2012] and with the characteristic curves preserved [Mehra et al. 2009]. Abstraction of architectural drawings has been performed by summarizing and abstracting groups of objects according to Gestalt rules [Nan et al. 2011]. We are not aware of work which has built on a database such as the one we have assembled representing abstraction of the human face as conceived by a set of artists.

3 Data Gathering

Picasso’s ‘bull’ suite can be seen as a master class in abstraction but different artists may have very different interpretations of the process of abstraction and the means to achieve it. For our analysis of abstraction, we needed a data-set from multiple artists drawing at multiple levels of abstraction where the drawings were sufficiently similar to permit creating a correspondence for analysis.

We forced our artists to abstract by limiting the amount of time they had to sketch a face. Drawing under a time limit is a common exercise in drawing classes and therefore, is a constraint that artists are accustomed to. The artists were instructed to draw the complete face within the allotted time so that they had to concentrate on the key features of the face. An alternative would have been to limit the detail (number of strokes) used in a sketch but our preliminary experiments indicated that this approach was not intuitive to the artists and therefore too disruptive.

We collected a database of portrait sketches from seven artists (art students, art professors, and animators) with extensive drawing experience, albeit with varying levels of skill. In each data-gathering session, we displayed a reference photograph of a face to the artists and asked them to sketch a portrait digitally using a stylus pen. We used photographs of 24 faces of both male and female subjects from the face database of the Center for Vital Longevity [Minear and Park 2004]. All sketches were captured using a Wacom pen, allowing the artist to modify the brush parameters but preventing them from erasing or undoing strokes. We capture each stroke as a parameterized directed curve along with the pen parameters (tilt, pressure, location). We also store each stroke as a transparent bitmap

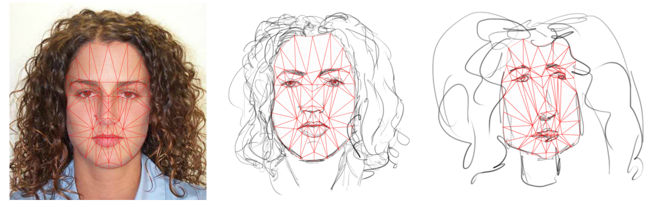


Figure 3: *Examples of fitting the mesh to the photograph and sketches in various levels of abstraction. The mesh includes eyes, mouth, nose and eyebrows as the important features for face recognition [Sinha et al. 2006].*

for later use in the synthesis of new sketches.

We use four time intervals in decreasing order (270, 90, 30 and 15 seconds) to allow the artists time to observe the face before attempting the quick abstractions. We asked them to accurately depict the face and avoid caricatures or exaggerated features. Our final dataset is composed of 672 sketches from seven artists, at four abstraction levels, containing around 8000 strokes for each artist (see Figure 2).

In a post-processing stage, we manually fit a template of a 2D triangulated face model to each of the sketches, as well as to the reference photographs. Although there are several automatic procedures that can fit a mesh to photographs, as well as to less abstract sketches, this process is more challenging for abstract ones. We wanted as accurate fit as possible and devised a simple user interface where mesh fitting onto a sketch takes less than a minute. Note that this was needed only for the training data. The resulting mesh is composed of 90 points and includes all important facial features (Figure 3).

We conduct the analysis of the data in two levels: strokes analysis (Section 4), and geometric shape analysis (Section 5). This decomposition is important as both properties affect the final visual characteristics of the sketch in both abstraction and style .

4 Strokes Analysis

The strokes recorded from all sketches of each artist are gathered together to create his/her strokes library. Our assumption is that characterizing the differences of various attributes of the strokes between artists, and sometimes also within the different levels of abstractions, can capture some of the dimensions that define a specific style, as well as the process of abstraction. Accordingly, we analyze and build models from the strokes library either separating levels of abstraction or merging them, depending on the trends we find in the data. We search for the following characteristics:

- The spatial distribution – *where* do artists draw strokes?
- The temporal distribution – *when* do artists draw strokes?
- Stroke statistics (length, amount, overlap etc.) – *how* do artists draw strokes?
- Stroke classification (contour strokes, shading strokes) – *what* types of strokes are used?

Spatial distributions Using the correspondence created by the mesh model on each line drawing, we can deform the sketch back to fit the base template and compare the spatial distribution of various parameters of the strokes in all sketches. For instance, if we create an intensity map by averaging the intensities of strokes in Figure 4, we see differences in the style of artists as to where they draw more intensely, as well as differences that depend on the level of abstraction, details are lost when abstracting. The intensity map is affected by the pen pressure, the number of strokes and the overlap between

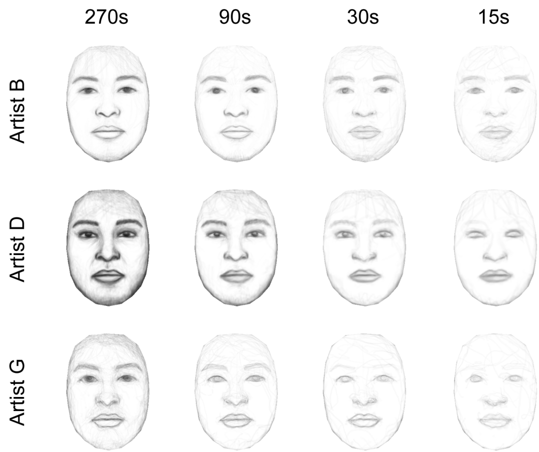


Figure 4: *Examples of the average distribution of strokes from white (very low) to black (very high) for different styles and abstraction levels. Each row represents a different style (artist) and each column a different abstraction, from most detailed (left) to most abstract (right). Details are lost as the abstraction increases: eyes lose intensity, nose outlines disappear, and the lips are simplified.*

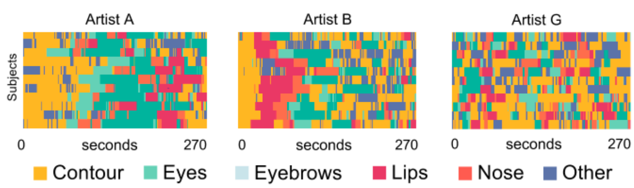
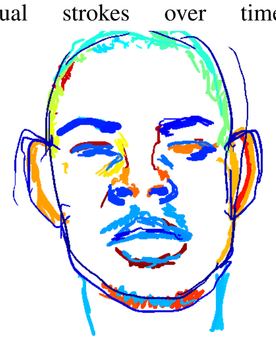


Figure 5: *Timeline of artists' sketching (270s): each color represents a different facial feature, and each row in the matrix represents a different subject. Artists usually start with the face contour but then their style diverges. For instance, artist A starts with the eyes while artist B with the lips. Both of them are more consistent than artist G.*

strokes (see the supplemental material for this separation). This information can guide us when synthesizing new sketches to better compute where strokes should appear, and what intensity to assign to them.

Temporal distributions Each sketch is created by combining many individual strokes over time. On the right is an example of a sketch where the colors of the strokes represent the time of drawing from blue (first) to red (last). To combine the statistics of a group of strokes, we classify each stroke according to its closest facial feature. Then, we can plot a timeline displaying the development of each sketch based on facial features (Figure 5). This information can be used when animating a specific drawing style.



Aggregated Strokes Statistics Calculating statistics for the recorded strokes reveals a unique signature for different artists as well as information regarding the process of abstraction. For each

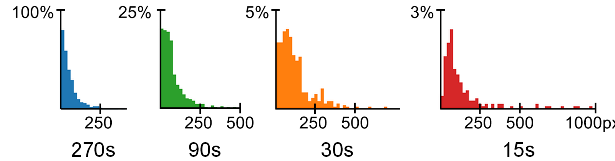
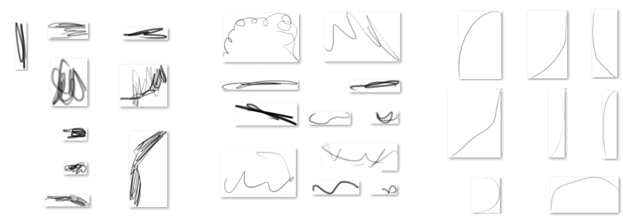


Figure 6: *The distribution of stroke lengths for four levels of abstractions averaged over all artists. Longer strokes are used for more abstract sketches (going from left to right). This trend is similar for each individual artist as well.*



(a) Shading (b) Complex (c) Simple

	270	90	30	15
ArtA	26.0%	28.0%	28.0%	23.0%
ArtB	37.0%	38.0%	40.0%	38.0%
ArtC	45.0%	52.0%	53.0%	52.0%
ArtD	7.0%	6.0%	11.0%	14.0%
ArtE	37.0%	35.0%	25.0%	28.0%
ArtF	66.0%	67.0%	63.0%	66.0%
ArtG	45.0%	69.0%	67.0%	52.0%

(d) Complexity

	270	90	30	15
ArtA	11.1%	6.4%	2.0%	0.7%
ArtB	17.2%	9.0%	2.6%	1.2%
ArtC	5.1%	2.0%	1.0%	0.2%
ArtD	12.7%	5.6%	2.0%	0.8%
ArtE	15.4%	9.4%	3.9%	1.2%
ArtF	5.1%	3.4%	1.3%	1.0%
ArtG	4.9%	2.0%	0.6%	0.4%

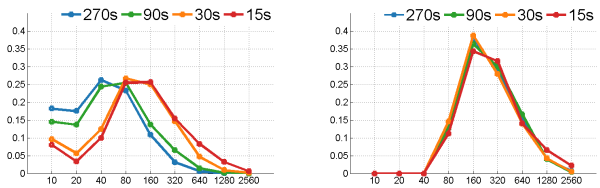
(e) Overlap

Figure 7: *Stroke types, style and abstraction: the percent of complex strokes (left) is clearly a part of the artist's style, while the amount of overlap between strokes is linked more to the process of abstraction. The shading graphically represents the percentage.*

artist in each abstraction level, we create a histogram of various parameters: stroke length, strokes overlap, curvature, pen pressure, and stroke speed. For instance, Figure 6 illustrates that as the sketches get more abstract, less strokes are used by all artists. Other parameters show (see supplemental material) that the strokes become longer and stronger (artists use more pen pressure), and the amount of overlap between strokes reduces (see Figure 7). This information can guide the synthesis of different abstraction levels.

Strokes Classification In general, strokes in a sketch are used for two purposes: to depict the shape and details of the subject (contour strokes) and for shading. We separate shading strokes from contour strokes and further classify contour strokes into complex strokes and simple strokes (see Figure 7). Shading strokes are defined as strokes where the ratio of drawn to non-drawn pixels inside the tight bounding box of the stroke is above a given threshold (75%), and the aspect ratio of the bounding box is above a threshold ($1:3$). All non-shading strokes are considered contour strokes. Complex strokes are classified as contour strokes that have more than four maximum curvature points above a threshold (0.1). Any contour stroke whose length is below a given threshold (5 pixels) is discarded. We have found that all types of strokes are used in all levels of abstraction (although in a different relative amount) and therefore classify all strokes of an artist together.

Stroke classification also provides insight into the abstraction pro-



(a) Simple strokes (b) Complex strokes

Figure 8: *The distribution of the length in pixels (x-axis) of simple and complex strokes for the four levels of abstraction. See text for details.*

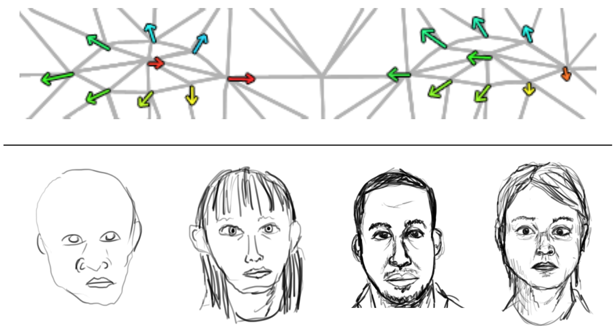


Figure 9: *A close-up example of the offset vectors created from the data of one artist for the eyes (color represents direction). This data indicates that the artist has a general tendency to draw pupils too close to each other and to draw large eyes (note the red and green arrows of the pupil mesh points, and the arrow directions around the eyes). This observation is supported by examining the artist's sketches themselves. By recognizing typical shape variations in the sketches, our geometric shape analysis provides artists with a tool to increase their proficiency in drawing.*

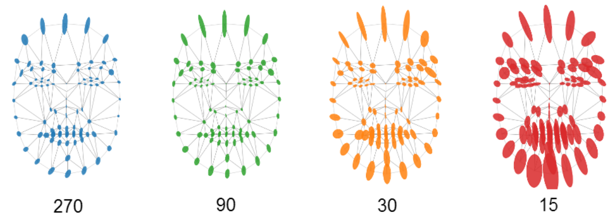
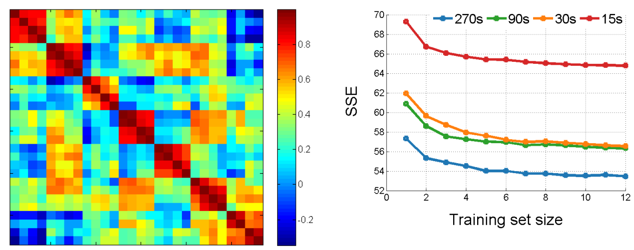


Figure 10: *The variance of the offset vectors of each point averaged across artists in each abstraction level: the higher the abstraction the larger the variance.*



(a) (b)

Figure 11: *(a) The correlation matrix of the shape variation representation vectors (combined* 180*-dimensional offset vectors). High correlation can be seen within the artist sketches (the inner* 4×4 *squares ordered from abstract to detailed), even between the abstract drawings and the more detailed ones of the same artist. (b) Using more than ten sketches, the average error of the vectors converge to a stable offset, which is larger as the abstraction is higher.*

cess. Figure 8 shows the aggregated distribution of the length of the contour strokes in all drawings, according to the level of abstraction. It is clear that longer strokes are used more in abstract sketching than short ones. However, complex strokes tend to have the same distribution in all abstraction levels, while simple strokes are longer in abstract sketches. The fact that the two longer and two shorter periods of sketching are coupled suggests there are two modes of sketching: one for short-time abstract sketches, and one for detailed realistic ones. This distribution can be used to guide the synthesis of sketches by defining the mixture ratios of the stroke types.

5 Shape Analysis

We want to measure variations of the artists' interpretation of a face. Like many others working in the area, we base our approach on Active Shape Models (ASM) [Stegmann and Gomez 2002; Truong 2011] and use it to bring a set of shapes into a frame of reference, and describe the variation within that frame. Our geometric shape analysis is performed using the correspondence between the portrait sketches and the photograph face shape created by the mesh fitted on both. ASM models use principal component analysis (PCA) to define a reduced sub-space that encodes most of the variations of the original space of shapes. Previously, such models examined the changes of the mesh vertex positions compared to some average position (e.g. "average face"), capturing face variations such as tilt, rotation and facial feature variations. In contrast, we do not use an average shape, and apply spectral analysis to the pairwise *differences* between each sketch and its ground truth face in the photograph. This analysis captures the differences between how the artist depicts a face and the true geometric shape of the face. These differences define, in effect, the artist's general stylistic interpretation (be it intentional or not) in terms of the shape of the face.

Let M_p be the mesh model fitted to the photograph of a given subject, and M_s the mesh model fitted to a sketch of the same subject. We uniformly normalize M_p, M_s to fit the scale of the template mesh without changing proportion or rotating, and align their center mesh point (a point on the nose). In this position, we compare all pairs of matching mesh points and define an offset vector by measuring the difference in their positions:

$$v_i = (p_i - s_i), p_i \in M_p, s_i \in M_s$$

We average the offset vectors of each artist in each level of abstraction, combining 24 sketches in total, to create a set of vectors representing the artist's local shape variations in any level of abstraction. Figure 9 shows an example of these vectors and the information they encode.

Figure 10 illustrates the variance of the offset vectors in the two principal directions of each point in each mesh. These values were averaged over all sketches of all artists, in four levels of abstraction. In addition, (see supplemental material for details) each artist has his/her own average variance for the offsets that defines his/her shape variation style. The total shape variation of a specific sketch can therefore be encoded using the 90×2 high dimensional vector of all offset vectors. Figure 11(a) shows the correlation of these vector representations for every artist in any abstraction level. Using PCA, we define a new basis for these shape variations and reduce the 180-dimensional representation by using only the first few

(3-10) principal components (PC). These capture between 70% to 98% of the variation.

We ran a test to measure the stability of our shape model. We used n out of 12 sketches for training and another set of 12 for testing. We used 10-fold cross validation test and measured the average difference of our shape model built from n sketches using 3 PC's, and the true sketches in the test set. Figure 11 shows the plot of the average difference as n increases. We found that ten sketches suffice to stably determine the model, but the error of the model is larger for more abstract sketches.

To illustrate how this analysis captures the geometric-shape style of an artist, we construct our shape model for all artists in the most detailed sketch and apply it by modifying the shape geometry of the true face mesh. We use 3 PC's in our model and modify the geometry by moving the points in a random direction from the artist's mean position by one standard deviation. A visual comparison of the resulting shape geometry and the shape of the artist's true sketch is given in Figure 12. This demonstrates that we capture the style of the major face deformations per artist even with a very small number of PCs.

6 Sketch Synthesis

We demonstrate the use of our analysis in synthesizing realistic sketches from photographs in various styles and abstraction levels. Synthesizing abstract (or detailed) stylistic sketches from a given photograph requires converting the photograph to a line drawing and applying stylization and abstraction.

To convert a photograph to a sketch we follow the outline in Figure 1. First, edges are extracted from the photograph and filtered. Next, the shape deformation model is applied to the face according to the desired style and abstraction level. Then the edges are converted to curve segments, yet again using information from the desired style and abstraction. Lastly the curves are replaced by strokes from the strokes database of the given artist. Note that naively extracting the edges from the image and replacing them with strokes from the library of an artist will not work well as the results will contain erroneous edges, wrong intensity and contrast and other noticeable artifacts (see Figure 13). Similarly, simple approaches for edge detection will not work well as some edges such as highlights should be removed while others should be strengthened (See Figure 14, left). In fact, throughout the synthesis process we rely on information extracted during analysis, which is key for creating a good approximation of style and abstraction. We elaborate each step in the following sections.

6.1 Edges Extraction

Lines extracted using simple edge detection from a photograph do not represent correctly the lines an artist would draw as a sketch of a face. Moreover, real abstraction demands a change in the style of the lines (e.g. merging of strokes), as well as accuracy reduction, which are both impossible to achieve using simple edge detection. We use an advanced edge detector that is based on Difference of Gaussian (FDOG) [Kang et al. 2007]. This NPAR technique is better suited to our needs as it extracts a set of coherent, smooth lines that better convey important facial features (see Figure 14, left). Furthermore, this method allows control of the amount of detail in the resulting edge image by employing different levels of bilateral filtering [Durand and Dorsey 2002] on the original photograph before extracting the edges.

Because of both style and abstraction, artists may choose to downplay or leave out some details, and intensify others. We use the

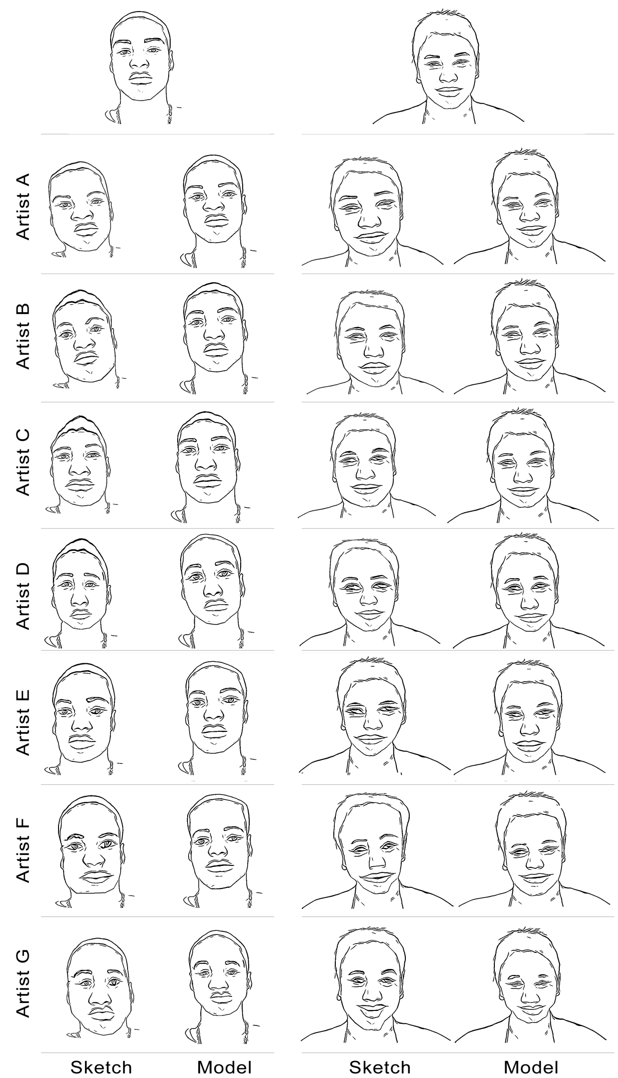


Figure 12: *Comparing our shape variation model (second and fourth columns) to the true artists' shape variations (first and third columns) for the same subject (shown at the top). Note that in these examples we use an edge map of the face to eliminate the effect of the strokes, and more clearly illustrate the shape variations.*

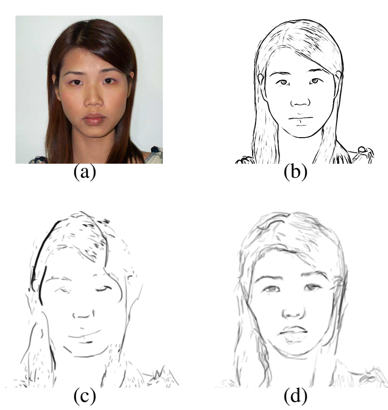


Figure 13: *Edge detection (b) or naïvely replacing the edges with strokes collected from an artist (c) do not resemble realistic sketches as our results (d).*

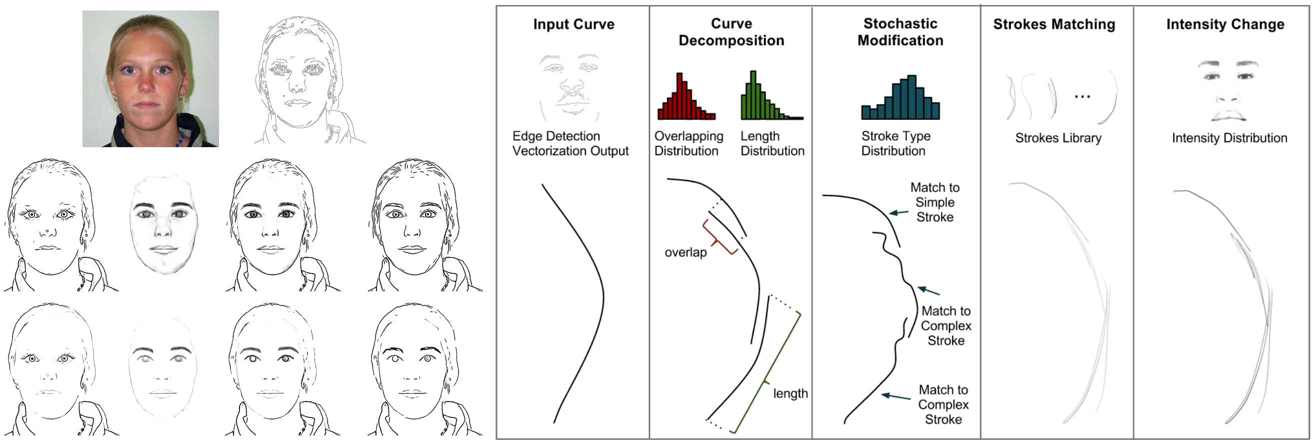


Figure 14: *From edges to strokes. Left shows the edge extraction: using a simple Canny edge detector (top row) does not produce‘ coherent smooth lines. We use the FDOG edge detection algorithm in different levels of abstraction (second vs. third row). Note that although the resulting edges are better (leftmost image in row), there are still incorrect edges (tip of the nose highlight) and missing edges (lower lip). We blend the intensity map of the artist with the edge image in the correct level of abstraction and extract edges from the results. This operation provides our base stroke edges for different abstraction levels (rightmost image in row). Right shows the process of converting edge curves to strokes. At all stages, we use information from our analysis (shown at the top) to preserve and guide the style and abstraction level including distributions and intensity maps (see text for details).*

normalized image of the stroke distribution map of an artist I_d (Section 4) at a specific level of abstraction to correct the edge detection results image I_e. Assuming both are in the range $[0, 1]$, in each pixel p, if the difference between the two is below a given threshold (0.35) we blend them using the following formula $I(p) = I_e(p) + I_d(p) - 1$. Then, we extract edges from the resulting image I with a lower threshold to produce an edge-image biased towards the artist's tendency of strokes (Figure 14, left).

6.2 Shape Deformation

To achieve a more realistic face shape, which matches a certain style and level of abstraction, we apply shape deformations to the edge image results from previous stage. We fit the mesh template to the face of the photograph, and choose the shape model trained on the desired style and abstraction level (Section 5). Then, we move the mesh points by a random amount of up to one standard deviation in the direction of the first n principal components of the chosen shape model. We warp the edge map of the image according to the mesh deformation achieving a deformed edge map resembling the artist's style at a given level of abstraction.

6.3 Curve Generation

Our goal is to convert the edges in the deformed edge map to strokes that capture the artist's stroke-level style. Towards this end, we employ Noris et al.'s method [2013] where image edges are converted to vector curves. As these curves are usually long and smooth, we used segmentation to better match the characteristics of the artist's strokes. There are three parameters that govern this process: the distribution of length of strokes, the amount of overlap of strokes and the amount of perturbation applied to edge curve points after segmentation to match more complex strokes. These measures are taken from the desired artist's style and abstraction level (Section 4). For each curve, we draw two random values from the desired distribution of strokes length (l), and amount of overlap (m). We segment each curve to sub-curves that have a length l (apart from the last sub-curve, that can be smaller), overlapping each other by m pixels. Next, we use a small perturbation value d that is proportional to the artist's strokes complexity, and perturb the x and y coordinates of a sample set of points of each sub-curve by a random amount in the interval $[-d, d]$. In general, the higher the abstraction level, the longer the strokes remain, the less they overlap and the more we apply perturbation to the curves (see Figure 14, right). Figure 15 demonstrates the effect of the different parameters on the style of the sketch and their use to match a specific artists' style.

6.4 Stroke Matching

Because of the diversity and complexity of the strokes in our data, especially at higher abstraction levels, we use a data-driven approach where we copy real artists' strokes to the image instead of using parametric pen-strokes for synthesis. To retain specific styles, we compose a strokes library for each artist guided by our analysis. We separate shading strokes from contour strokes, and separate the contour strokes to complex and simple. Within each category, we choose stroke descriptors that capture the shape compactly for accurate matching and fast querying.

The recorded strokes' raw representation is a list of sampled points that compose the stroke trajectory. Each sample point includes the recorded pen parameters: the pen tilt, pressure, location and time (to calculate speed). Our descriptor for each stroke is a vector representation based on three signatures: the histogram of shape context, the stroke's curvature and its length. We calculate the shape context descriptor for each sample point, we used 5 logarithmic bins for distances and 12 bins for orientation. Then, we combine all shape context descriptors of all points to a normalized 2D histogram. The length of the stroke is calculated simply by summing up the lengths of the segments between each two sample points. The curvature descriptor is a histogram with 10 buckets of the curvatures of each sample point. We use cascading nearest neighbor search to find a matching stroke for each curve segment from the previous step.

Once a stroke is matched, we still need to position and render it. To better fit the stroke to the query segment, we use the Iterative Closet Points (ICP) algorithm, matching the sample points on the two curves we find the best rigid transformation aligning the stroke to the curve edge. After replacing the edge with the stroke, we

Figure 15: *Parameters of stroke style: changing the stroke average length (first row), the amount of stroke overlap (second row), and the amount of complex strokes used (third row) can create various stylistic effects while matching the strokes of a given artist to the curve edges. In the analysis stage, we record the values of these parameters, along with stroke intensity, for a given artist. While synthesizing we use these values to define the stroke style of an artist in a given level of abstraction. The bottom row shows the settings of these parameters for our seven artists.*

modify the intensity of the stroke according to its position by using the stroke intensity map for the given style and abstraction (see Figure 14, right).

6.5 Animated Drawing

An optional step for synthesis is animating the creation of the drawing, similar to the way that the artist would draw the sketch. For this we use the temporal distribution of strokes (see Figure 5) according to facial features. We combine a smoothed version of the time-line of each separate feature in all sketches of the artist, and normalize it to build a probability density function (PDF) for when a feature is most probably drawn. While animating the sketch, we randomly sample all feature PDF's, and choose the most probable one to draw. The next stroke to draw would be taken from this feature (if such a stroke still remains). This process continues until all strokes are drawn.

Figure 16: *Examples of previous work (from top left to bottom right): input image, FDOG, PhotoToSketch (Commercial Application), Chen et al. 2004, Gooch et al. 2004, PhotoShop (sketch graphic pen effect), Pictoon, and our results (we chose a representative result when using the input image was not possible). Note the look-and-feel that our results convey as an approximation of a real sketch and not just an abstraction of an image.*

7 Results

There is a large body of work that proposes and examines stylistic sketching and abstraction effects that apply various image processing filters (see Figure 16 and Section 2). Less work has been done that tries to convey the specific style of individual artists, and, to our knowledge, none have focused on the process of abstraction of real artists. Our goal was not to produce visually aesthetic sketches or exaggerated caricatures, but to produce sketches that simulate the abstraction and style of individual artists.

The figures in this paper include several examples of our synthesized sketches. Synthesizing a sketch takes between 30 to 100 seconds, depending on the artist and the abstraction level. Although we can synthesize any input face from a photograph, we repeatedly synthesized the same 24 faces that were used for data gathering to allow comparison between real and synthetic results. When doing so, we omitted the strokes of the input image from the training set for learning. Figure 17 demonstrates a comparison between real and synthesized sketches of a single subject in the styles of the seven artists at two levels of abstraction. Figure 18 shows abstractions created by our methods to various faces. More results can be seen in Figure 19 and the supplementary material for this paper.

7.1 Perceptual Study

To assess our results, we conducted three perceptual studies. In all experiments, we removed the first drawings done by the artists because artists were still adjusting to the use of the tablet and the data gathering procedure. In Experiment 1, we examined whether style was being conveyed in the synthesized results as well as it is conveyed in human-created artwork. For Experiment 2, we determined how well viewers could align the synthesized results with the human artists' specific styles. For these experiments we selected the most detailed sketches (270 sec.) and corresponding synthesized sketches as it would be more difficult, even for a trained artist, to establish style on more abstract depictions. We used two groups of eight adult participants who were unfamiliar with this research for the two experiments.

In a third experiment, participants saw a series of real and synthesized images at either the lowest or highest level of abstraction and identified whether they were real or synthesized. We used selections from seven artists at both 270 sec. and 15 sec. data length. Twenty adults participated in Experiment 3 via Mechanical Turk.

Experiment 1. We randomly picked eight faces out of the 23 and selected the corresponding drawings from the seven artists. The eight sketches from each artist were composed into two images $R1, R2$, each containing four of the faces presented in seven rows, one row for each artist (see examples in Figure 19). We duplicated this procedure to select eight sketches of seven artists from the *synthesized* results, and created two more images $S1, S2$ with the row order held the same. The remaining 15 drawings and 15 synthesized results that had not been selected were used in trials for the comparison task.

Participants were instructed that each row had been created in a particular artist's style, and they were to match the single image to one of the seven rows based on which style was the most similar. For each trial, a single image was shown on the right half of the screen and the image collections on the left. The participants were asked "Which row does this image most resemble?" For $R1$ and $R2$, real sketches were presented on the right, and for $S1$ and $S2$, synthesized sketches were presented on the right. Each trial ended when the participant keyed in a response. The order of the single images for the trials was randomized. The experiment lasted around 45 minutes.

Overall, participants did not significantly differ in their abilities to classify real and synthesized images by style (mean drawings = 72.4%, stdv = 10.2%; mean synthesized = 76.4%, stdv = 12.6%; t = 2.0, p = .09), with a slight trend towards better classification of synthetic images. For both types of images, classification was significantly better than the 14.3% accuracy expected by chance (chi squared = 6276.8, $p < .0001$). These results indicate that our sketch generation method captures and differentiates artists' styles well.

Experiment 2. We used the same procedure and stimuli as in Experiment 1; however, we switched the role of $R1, R2$ and $S1, S2$, i.e. we had participants match synthesized sketches to a collection of real sketches and vice versa. Again, participant performance did not significantly differ when they were sorting single synthetic images using real drawings for the style and vice versa (mean synthetic sorting = 48.7%, stdv = 8.0%; mean real sorting = 50.1%, stdv = 5.2%, t = .48, p = .64). Participants performed significantly above chance (chi squared = 1690.1, $p < .0001$). These findings suggest that our sketch generation method accurately reflects the styles of the individual artists whose work was used as input. Participants commented that it was difficult to sort the images using seven style categories. In an earlier study run on an additional eight participants using only five artists the same pattern of results was found, but with higher accuracy. The t-test showed no significant difference in accuracy across sorting conditions (mean synthetic sorting = 76.4%, stdv = 10.8%; mean real sorting = 72.8%, stdv = 7.2%, t = .50, p = .63), and significantly better overall accuracy than would be expected by chance (chi squared = 4932.5, p $< .0001$).

Experiment 3. For the third experiment, we used SurveyGizmo.com to show participants a single image at a time and ask them, "Was this image created by hand or by a computer?" We used two sets in this experiment. One set for the most abstract sketches included three randomly selected images from each of the seven artists of both real and synthesized sets arriving at 42 trials. The other set for the least abstract sketches included two images from each of five artists of both real and synthesized sets arriving at 20 trials. Each set was carried out as a separate experiment and in each experiment the order of trials was randomized and the trials advanced only after an answer was entered.

In the most abstract set experiment, the twelve participants did not show significantly different levels of accuracy of identification of the real and synthesized images (mean drawings = 63.1%, stdv = 26.2%; mean synthesized = 48.0%, stdv = 22.5%; t = 1.21, p = .25). For the least abstract set experiment, the twelve participants did not show significantly different levels of accuracy of identification of the real and synthesized images (mean drawings = 60.8%, stdv = 16.4%; mean synthesized = 53.9%, stdv = 16.7%; t = .95, p = .36). Although preliminary due to the small number of participants, these data indicate that our method of synthesizing sketched images in various styles at these two levels of abstraction may be similar in perceived realism to hand drawn sketches.

These results demonstrate that our sketch generation method for portraits can produce multiple, distinct styles, that are similar to real hand-drawn sketches.

8 Discussion

Based on the analysis of the data-set we gathered, we can define a clearer model for the process of abstraction in line-drawings of portraits. Abstraction is composed of the following principles:

Prioritized reduction in details: less strokes are used, strokes are concentrated on more important facial features.
Merging of strokes: fewer, longer, and more complex-shape strokes are used instead of many short simple ones.
Stylistic accuracy reduction: larger errors are introduced both in terms of shape composition and accuracy of strokes positioning, but these are not random, and carry the style of the artist.

In terms of style, we found that both shape and stroke level characteristics are key players in defining an artistic style. Our analysis found consistent tendencies of artists that sometimes they did not know themselves. Assuming these shape adjustments are not intentional, recognizing such tendencies can also help artists increase the accuracy of their drawing and their proficiency.

Limitations & future directions There are several limitations to our analysis. First, we focus on a specific domain – face portraits.

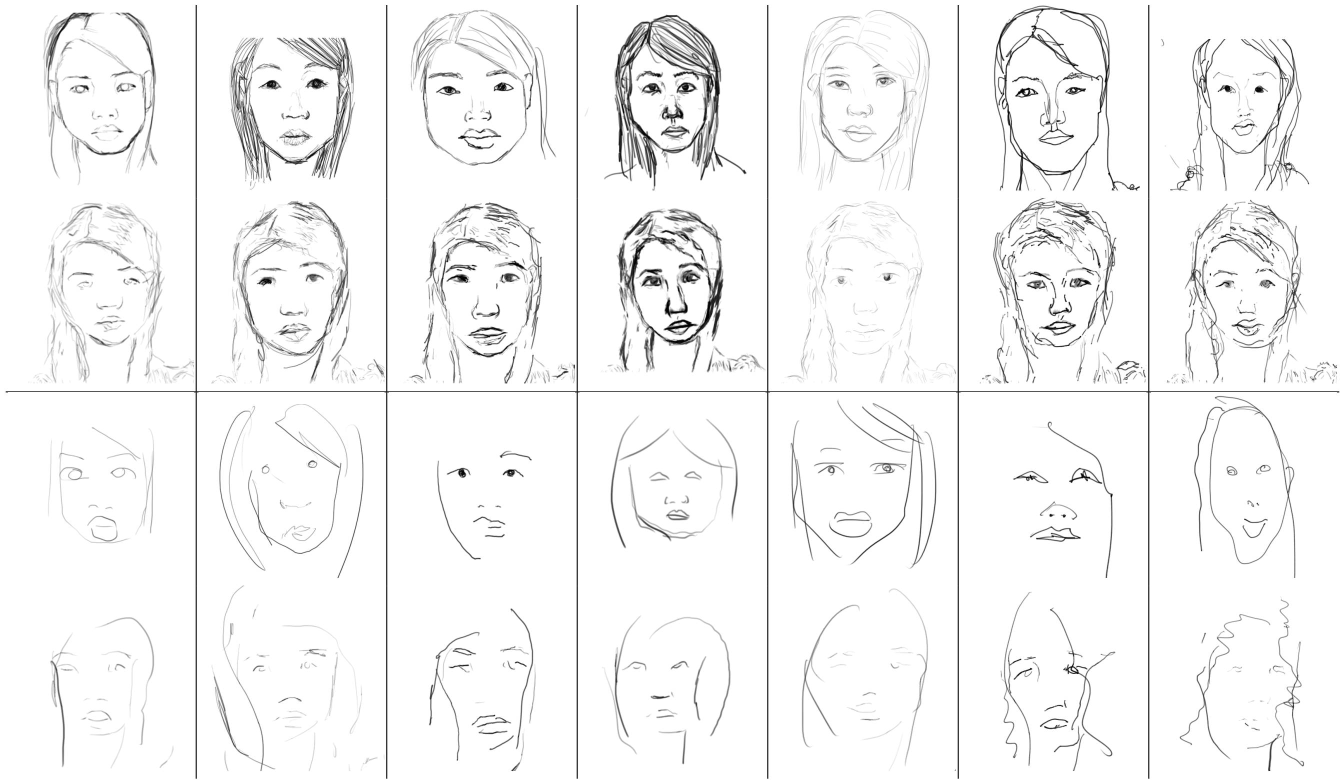


Figure 17: *Comparison of real and synthesized results of all seven artists' styles (columns) and in two levels of abstraction (top and bottom) of a single woman model (shown in Figure 13). The first and third rows are the real sketches of the artists at the least and most abstract levels respectively, while the second and fourth are our corresponding synthesized results. Note how each artist has his/her own way of drawing the eyebrows, nose, and mouth.*

Our shape analysis would be difficult to generalize to other sketch subjects, but we believe our strokes analysis could be utilized for general sketches as well. Second, we focus on a specific technique – sketching. It would be more difficult to carry over the strokes analysis to other painting techniques, although the shape analysis could be utilized for general portrait paintings. It would also be interesting to extend our perceptual study to measure the relative importance of the two component: shape and strokes, on capturing the style of an artist.

In terms of sketch portrait analysis, our key model fit the face of the subjects but did not model the subjects' hair. This limitation can sometimes be noticed in our synthesized results. We concentrated on using contour strokes and did not utilize shading strokes, and used only curve segmentation to match strokes. Utilizing shading strokes can enrich the sketch results, while merging curves can assist especially when synthesizing abstract sketches.

Another avenue for possible future investigation is building a deformation model based on individual facial features (eyes, nose etc.) and not the whole face. More generally, our abstraction model did not utilize semantic understanding except in how it was captured by the artist's drawings.

Summary We have presented a data-driven method to analyze the process of abstraction and to learn different styles in portrait sketching. Using two-levels: shape and strokes, we created models of both artistic traits and illustrated their use by building a sketch synthesis application that converts a photograph to a sketch. User validation showed that our synthesis can capture both style and abstraction.

Acknowledgements

We would like to thank all artists that participated in our data collection. Jenn Tam Hyde, Guy Hoffman and Ronit Slyper for helping in the user study. This project was partly supported by the Israel Science Foundation (grant no. 324/11).

References

BELONGIE, S., MALIK, J., AND PUZICHA, J. 2002. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24*, 4, 509–522.

CHANG, I.-C., AND CHENG, R.-M. 2011. Caricaturation for human face pictures. In *International Conference on Machine Learning and Cybernetics*, 1702–1707.

CHEN, H., ZHENG, N., LIANG, L., LI, Y., XU, Y., AND SHUM, H. 2002. Pictoon: a personalized image-based cartoon system. In *Proceedings of the Tenth ACM International Conference on Multimedia*, 171–178.

CHEN, H., LIU, Z., ROSE, C., XU, Y., SHUM, H., AND SALESIN, D. 2004. Example-based composite sketching of human portraits. In *Proceedings of the 3rd International Symposium on Non-photorealistic Animation and Rendering*, NPAR '04, ACM, 95–153.

CHIANG, P., LIAO, W., AND LI, T. 2004. Automatic caricature generation by analyzing facial features. In *Proceedings of the 2004 Asia Conference on Computer Vision*.

Figure 18: *Synthesized portrait abstractions in various styles. Drawings are adjusted for on-screen viewing*

COLE, F., GOLOVINSKIY, A., LIMPAECHER, A., BARROS, H. S., FINKELSTEIN, A., FUNKHOUSER, T., AND RUSINKIEWICZ, S. 2008. Where do people draw lines? *ACM Transactions on Graphics (Proc. SIGGRAPH) 27*, 3 (Aug.).

COLTON, S. 2010. Stroke matching for paint dances. In *Proceedings of the Sixth international conference on Computational Aesthetics in Graphics, Visualization and Imaging*, Computational Aesthetics'10, 67–74.

DECARLO, D., AND SANTELLA, A. 2002. Stylization and abstraction of photographs. *ACM Transactions on Graphics 21*, 3, 769–776.

DURAND, F., AND DORSEY, J. 2002. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics 21*, 3, 257–266.

EITZ, M., HAYS, J., AND ALEXA, M. 2012. How do humans sketch objects? *ACM Transactions on Graphics 31*, 4, 44.

FREEMAN, W. T., TENENBAUM, J. B., AND PASZTOR, E. C. 2003. Learning style translation for the lines of a drawing. *ACM Transactions on Graphics 22*, 1, 33–46.

GOOCH, B., REINHARD, E., AND GOOCH, A. 2004. Human facial illustrations: Creation and psychophysical evaluation. *ACM Transactions on Graphics 23*, 1, 27–44.

HERTZMANN, A., OLIVER, N., CURLESS, B., AND SEITZ, S. M. 2002. Curve analogies. In *Proceedings of the 13th Eurographics Workshop on Rendering*, 233–246.

KALOGERAKIS, E., NOWROUZEZAHRAI, D., BRESLAV, S., AND HERTZMANN, A. 2012. Learning Hatching for Pen-and-Ink Illustration of Surfaces. *ACM Transactions on Graphics 31*, 1.

KANG, H., AND LEE, S. 2008. Shape-simplifying image abstraction. *Computer Graphics Forum 27*, 7, 1773–1780.

KANG, H., LEE, S., AND CHUI, C. 2007. Coherent line drawing. In *Proceedings of the 5th International Symposium on Non-photorealistic Animation and Rendering*, ACM, 43–50.

KYPRIANIDIS, J., AND KANG, H. 2011. Image and video abstraction by coherence-enhancing filtering. *Computer Graphics Forum 30*, 2, 593–602.

KYPRIANIDIS, J. E., COLLOMOSSE, J., WANG, T., AND ISENBERG, T. 2013. State of the art: A taxonomy of artistic stylization techniques for images and video. *IEEE Transactions on Visualization and Computer Graphics 19*, 5, 866–885.

LE, N. K. H., WHY, Y. P., AND ASHRAF, G. 2011. Shape stylized face caricatures. In *Proceedings of the 17th International Conference on Advances in Multimedia Modeling - Volume Part I*, MMM'11, 536–547.

LIANG, L., CHEN, H., XU, Y.-Q., AND SHUM, H.-Y. 2002. Example-based caricature generation with exaggeration. In *Proceedings of the 10th IEEE Pacific Conference on Computer Graphics and Applications*, PG '02, 386–393.

LIMPAECHER, A., FELTMAN, N., TREUILLE, A., AND COHEN, M. 2013. Real-time drawing assistance through crowdsourcing. *ACM Transactions on Graphics 32*, 4 (Aug.).

LIU, Y., SU, Y., SHAO, Y., AND JIA, D. 2010. A parameterized representation for the cartoon sample space. In *Proceedings of the 16th International Conference on Advances in Multimedia Modeling*, 767–772.

LU, C., XU, L., AND JIA, J. 2012. Combining sketch and tone for pencil drawing production. In *Proceedings of the 10th Inter-*

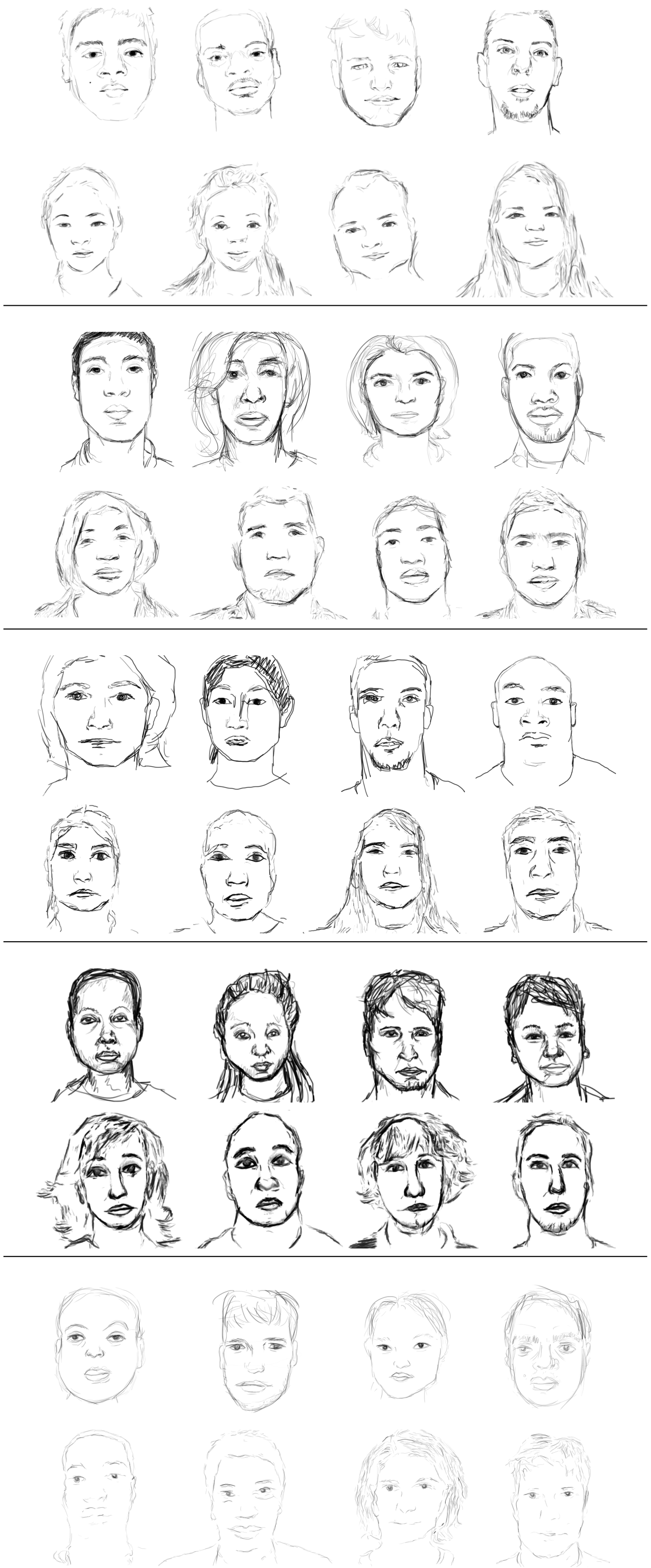


Figure 19: *From top to bottom in pairs: examples of the real (top row) and synthesized (bottom row) sketches of five Artists used in our perceptual study.*

national Symposium on Non-Photorealistic Animation and Rendering, NPAR '12, 65–73.

LU, J., YU, F., FINKELSTEIN, A., AND DIVERDI, S. 2012. HelpingHand: Example-based stroke stylization. *ACM Transactions on Graphics 31*, 4 (Aug.), 46:1–46:10.

MEHRA, R., ZHOU, Q., LONG, J., SHEFFER, A., GOOCH, A., AND MITRA, N. 2009. Abstraction of man-made shapes. *ACM Transactions on Graphics 28*, 5, 137.

MENG, M., ZHAO, M., AND ZHU, S. 2010. Artistic paper-cut of human portraits. In *Proceedings of the ACM International Conference on Multimedia*, 931–934.

MINEAR, M., AND PARK, D. 2004. A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers 36*, 630–633.

NAN, L., SHARF, A., XIE, K., WONG, T., DEUSSEN, O., COHEN-OR, D., AND CHEN, B. 2011. Conjoining gestalt rules for abstraction of architectural drawings. *ACM Transactions on Graphics 30*, 6, 185.

NORIS, G., HORNUNG, A., SUMNER, R. W., SIMMONS, M., AND GROSS, M. 2013. Topology-driven vectorization of clean line drawings. *ACM Transaction on Graphics 32*, 1, 4:1–4:11.

SINHA, P., BALAS, B., OSTROVSKY, Y., AND RUSSELL, R. 2006. Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE 94*, 11, 1948–1962.

STEGMANN, M., AND GOMEZ, D. 2002. A brief introduction to statistical shape analysis. *Informatics and Mathematical Modelling, Technical University of Denmark, DTU*, 15.

TRESSET, P. A., AND LEYMARIE, F. F. 2013. Portrait drawing by Paul the robot. *Computers and Graphics 37*, in press.

TRUONG, V. 2011. Face alignment using active shape model and support vector machine. *International Journal of Biometrics and Bioinformatics 4*, 6, 224.

WANG, X., AND TANG, X. 2009. Face photo-sketch synthesis and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence 31*, 11, 1955–1967.

WANG, S., ZHANG, L., LIANG, Y., AND PAN, Q. 2012. Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2216–2223.

YU, H., AND ZHANG, J. 2010. Caricature synthesis based on mean value coordinates. In *CASA 2010: Proceedings of the 23rd International Conference on Computer Animation and Social Agents*.

YUMER, M., AND KARA, L. 2012. Co-abstraction of shape collections. *ACM Transactions on Graphics 31*, 6, 166.

ZHANG, Y., MCCULLOUGH, C., SULLINS, J., AND ROSS, C. 2010. Hand-drawn face sketch recognition by humans and a pca-based algorithm for forensic applications. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans 40*, 3, 475–485.

ZHAO, M., AND ZHU, S. 2011. Portrait painting using active templates. In *Proceedings of the 11th International Symposium on Non-Photorealistic Animation and Rendering*, NPAR '11, 117–124.