Artistic Rendering of Portraits

Mingtian Zhao and Song-Chun Zhu

Abstract This chapter presents an overview of the latest artistic portrait rendering techniques. Artistic rendering or portraits can be viewed as an image generating process controlled by two factors. The first one is the face fidelity, namely, a rendered portrait image should preserve a certain amount of the original face's information. The second factor is the artistic style, for example, sketch, painting, etc. In the literature, different portrait rendering algorithms either adopt different models and data structures to represent the facial information, or use different graphical elements and compositional methods to simulate various styles. These two factors essentially reveal the two principles in portraiture, namely the pursuit of likeness and aesthetic.

1 Introduction

Portraiture, which creates artistic representations of the appearances and expressions of human faces, is one of the oldest yet popular genres in visual arts. Generally there are two essential factors to consider in creating a portrait.

- The first factor is the face fidelity, namely, a portrait should preserve a certain amount of the original face's information, to ensure that not only can it be recognized as a face picture but there is an appropriate level of similarity in perception of the appearance or character between the person in the portrait and the person him/herself or in a photograph.

———————————————

Mingtian Zhao
University of California, Los Angeles, California 90095-1554, USA
e-mail: mtzhao@ucla.edu

Song-Chun Zhu
University of California, Los Angeles, California 90095-1554, USA
e-mail: sczhu@stat.ucla.edu

- The second factor is the artistic style of the portrait picture, for example, sketch, painting, paper-cut, caricature, etc. These styles/forms provide unique dictionaries of visual elements to express the various facial structures and appearances.

These two factors vary with different ages and genres of portraiture, revealing its two principles, namely the pursuit of likeness and aesthetic, respectively. For example, before the invention of photography in the 19th century, the mainstream artists used to pursue accurate likeness by studying the structure of bones and muscles beneath the facial skin, practicing their skills on depicting them, developing pigments made from various materials, and even using external tools such as mirrors and pinhole imaging to improve the fidelity. Nowadays, with the popularity of digital cameras, perfect fidelity is easily available, but many modern portraits usually depict only rough or even distorted likeness while resorting to new artistic styles and techniques to evoke strong psychological and emotional reactions in the audience, to demonstrate the sense of aesthetic.

These two factors/principles also apply to artistic rendering of portraits — the simulation of portraiture on the computer. From an image analysis and synthesis perspective, let W denote the facial information and Δ denote the elements to composite an image, then a natural image (photograph) $\mathbf{I}_N$ is generated with

$$\mathbf{I}_N = f(W_N; \Delta_N), \tag{1}$$

and in a similar way, an artistic portrait $\mathbf{I}_A$ can be synthesized with

$$\mathbf{I}_A = g(W_A; \Delta_A), \tag{2}$$

where f and g are image generating functions (rendering processes). Interestingly, Eqs. (1) and (2) differ in all their three aspects.

- $W_N \neq W_A$: The facial information W usually contains features such as geometry (2D or 3D), appearance, texture, color, and illumination. To generate a realistic photograph, W_N should usually approximate the truth very closely. In contrast, W_A often only captures part of the information interesting to artistic perception, which is regarded as the essense of a face by many artists.
- $\Delta_N \neq \Delta_A$: In the image analysis and computer vision literature, Δ_N is usually modeled with PCA, wavelets like Gabor bases, image patches, etc. Δ_A, however, is usually a dictionary of graphical elements used in creating artworks, for example, graphite sketches, paint brush strokes, etc.
- $f \neq g$: While f is usually a simple linear combination of the image elements, the portrait rendering process g can be a much complex process involving content-oriented algorithms for manipulating the sketches, strokes, etc.

In the non-photorealistic rendering (NPR) [7] literature, there are plenty of studies on artistic portrait rendering with different implementations of W_A and $g(\,\cdot\,; \Delta_A)$, corresponding to likeness and aesthetic, respectively.

- To preserve the facial fidelity, existing portrait rendering methods adopt different models and data structures to represent selected geometry and appearance features in W_A.
- To simulate different artistic styles, existing methods use different dictionaries of graphical elements which are maintained in Δ_A, and corresponding compositional algorithms g.

In the rest of this chapter, we review the latest artistic portrait rendering methods and their respective implementations of the two factors. We organize these methods by the four most studied types of portraits in NPR: sketch, paper-cut, oil-painting, and caricature.

2 Sketch

A sketch is a rapidly executed drawing demonstrating the basic shape and appearance features of objects. In this section, we review three types of portrait sketching methods. The first two types of sketches depict the boundaries and salient edges/curves in portraits with concise strokes (like stick drawings), the former using holistic models for the shape of face and the latter using part-based models with greater expressive power. The third type of portrait sketch pays more attention to the facial surface, including the appearance caused by illumination and shading effects.

2.1 Holistic Models

Li and Kobatake [10] made one of the earliest investigations in generating facial sketches from photographs. Their method consists of three steps:

1. Color coordinate transformation, in which an input image is first processed with the saturation component enhanced, and transformed to the YIQ color space. The Y channel represents the luma information, and the I and Q channels represent the chrominance information. They are used for extracting the face area and some facial parts: lips are red so they have relatively larger values in Q, the face area with skin color is generally larger in Q and smaller in I than the dark gray background, and black pupils of eyes are usually darker than other parts reflected in Y.
2. Facial components detection. In addition to lips and pupils, facial parts such as eyes, mouth, nose and chin are located with rough edges detected using the Y channel of the image.
3. Approximation of edges with feature points and feature curves. The method takes advantage of a facial sketch representation with 35 feature points connected by feature curves as shown in Figure 1a, in which spliced second-order polynomials
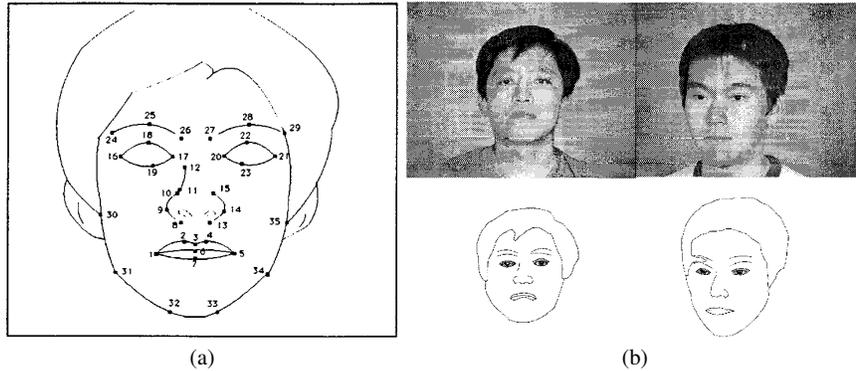
(a)   (b)

Fig. 1 (a) Feature points and feature curves used for extracting facial sketch images by Li and Kobatake [10, 11]. (b) Example results of facial sketch extraction using their method, from [11].

are used for approximating the edges of mouth, eyes, nose and chin. The 35 feature points are then detected as characteristic points on the curves.

In their follow-up work [11], detailed algorithms of the method were improved, including the symmetry measure, a novel rectangle filter, a geometric template, and morphological processing, which leads to more robust detections of the positions and edges of facial parts. Figure 1b shows two example facial sketches extracted using the method.

In Li and Tobatake's methods, the two factors introduced in Section 1 are implemented in a very straight-forward way. The facial information W_A is represented with the shape model shown in Figure 1a, which is extracted from the color and gradient features of the input image. This ensures that the portrait sketch looks similar to the photograph. As for the second factor, Δ_A simply defines a line drawing style, with the feature curves fitted using spliced second-order polynomials.

To further improve the results of Li and Tobatake's method, especially on the aesthetic aspect, many studies have been carried out recently. Chen et al. [2] developed an example-based facial sketch generation system whose pipeline is shown in Figure 2. Observing that the artistic styles of sketches vary among different artists and cannot be easily summarized by precise rules such as polynomial curves up to a specific order or curvatures of certain degrees, their system refers to a set of training examples for obtaining their styles. Each training example is a pair of portrait photograph and its corresponding sketch image created by artists, as shown in the top-left of Figure 2. An active shape model (ASM) [5] for face is also attached to each example, with the landmark points manually labeled for better accuracy, as shown by the dots in the middle-left of Figure 2.

- In the training phase, a mean shape of face is computed by averaging the ASM landmarks of all training examples, then a geometric transformation is performed on each training example to warp the image and the sketch to match the mean shape. After that, a prior probability model of the sketches is learned to cover
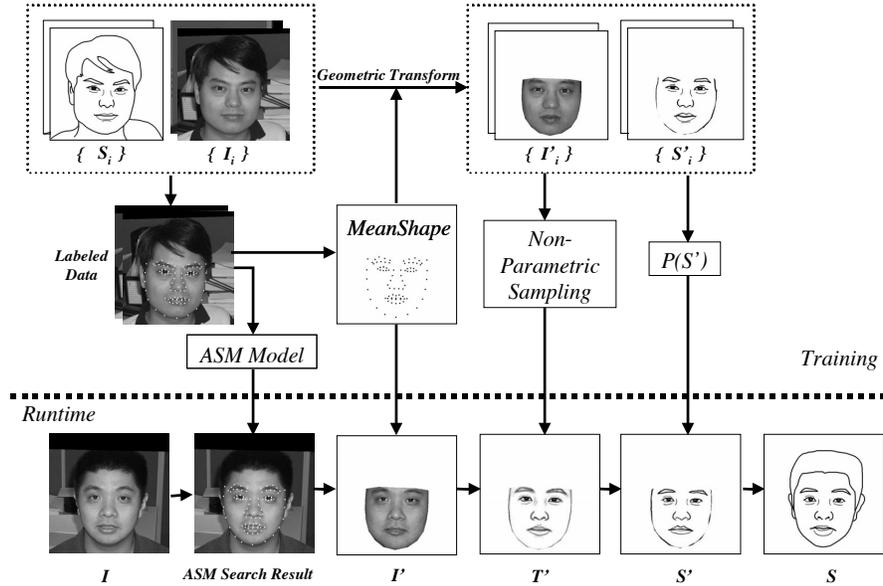
Fig. 2 Pipeline of the example-based facial sketch generation system developed by Chen et al. [2]

three types of curves: those always appear, those probably appear but are independent of other curves, and those depend on other curves.

- At runtime, given an input face photograph, ASM is first applied to extract the landmark points, with which a geometric transformation is defined between these landmarks and the mean shape. After applying this geometric transformation to warp the input image, non-parametric sampling is used for producing a sketch image for the input, which is then warped back to the original shape using an inverse geometric transformation, for the final sketch.

Compared with Li and Kobatake's method, Chen et al.'s system improved in both factors through W_A and Δ_A. For facial information W_A, the ASM model provides a more robust way to capture the shape of the face than the model in Figure 1a which relies on local edge detection and curve fitting. For a higher level of aesthetic, the sketches used as graphical elements in the style set Δ_A come essentially from the training examples created by artists instead of naive polynomial curves, and are encoded within a prior distribution of the sketch curves as we introduced above.

2.2 Part-Based Models

An disadvantage of the holistic methods is that the rendered sketches only contains stiff lines and curves, while sketches created by artists usually have various curve styles and levels of darkness and thickness for different facial components, as well

as the hair. To address this problem, part-based methods were introduced for processing different parts of the face separately in order for greater expressive power. Generally, these part-based models allow richer representations of W_A and larger varierty of elements in Δ_A.

2.2.1 Flat Model

Chen et al. [3] proposed an example-based composite sketching approach. The main idea of this approach is to decompose the face into semantic parts (as shown in Figure 3) and also to use image-based instead of curve-based sketches (as shown in Figure 4). Their system also has a sub-system for hair rendering which contributes greatly to the visual quality of the portrait sketches.
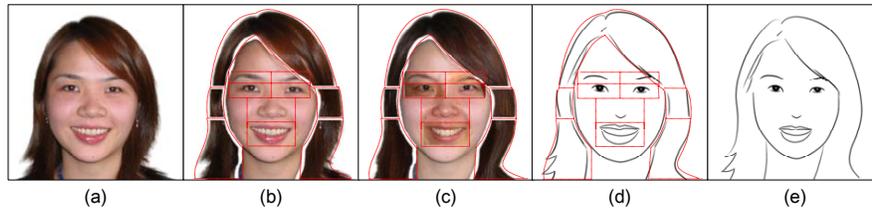

| (a) | (b) | (c) | (d) | (e) |

Fig. 3 Portrait sketching using example-based composite sketching [3]. The input image (a) is decomposed into facial parts shown in (b), then for each part the best match is found in the training examples, as shown in (c). Then sketches are drawn for each part in (d), which are composited in (e) considering both global and local features.
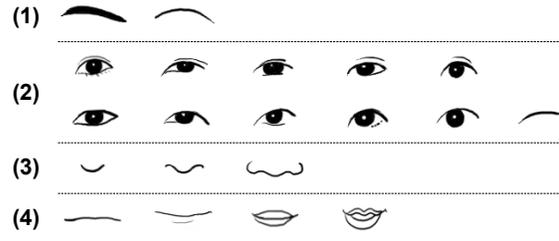

Fig. 4 Example sketches of facial parts used by Chen et al. [3] for Δ_A.

In the composite sketching system of Chen et al. [3], the sketches are splitted into two layers: a global layer and a local layer. The global layer captures the spatial placement of the facial parts, and the local layer captures how each facial part is depicted. For both layers, features are selected and learned using training image-sketch pairs provided by artists. At runtime, two processing steps are executed.

- Local Processing: The input image is first decomposed into facial parts using a refined ASM, then for each part the system finds the best match in shape (according

to ASM landmarks) from the training examples. After that, their corresponding sketches of the best matches are adopted.

- Global Processing: The sketches of the facial parts are composed according to the learned global spatial model, which adjusts the locations, sizes and orientations of the parts.

The hair rendering sub-system extracts both structural (boundary) and detail (streamline) components in the hair area, and fits them with curves. The two types of curves are then rendered using their respective example-based strokes (learned from training examples) to synthesize the hair sketch.

Regarding the two factors we have been discussing in this chapter, the global/local hybrid method of Chen et al. [3] works in a more flexible way for preserving face fidelity in W_A than the global geometric transformation based on ASM used in their early work [2]. For each facial part, the locally best matching sketch is selected and composed into a globally coherent image. But due to the selection of the best matches instead of a learned local sketch model [2], the system may lose certain degrees of likeness when the number of training examples is small, for example, the locally best matches may not be good enough in terms of similarity. On the aesthetic aspect, however, this method achieves much finer detailed appearances by maintaining image-based sketches instead of stiff curves in Δ_A, as shown in Figure 4, where different facial parts may be sketched using different techniques by artists.

2.2.2 Hierarchical Model

Since part-based portrait rendering methods have greater expressive power than global methods, more studies in this direction have been dedicated to better models and algorithms for both global and local processing. A powerful model for organizing the facial components is the hierarchical and compositional And-Or Graph (AOG) [23, 22, 24], as shown in Figure 5.

In this AOG, the And-nodes represent decompositions of the face or its parts into sub-components (e.g., decomposing the face into nose, mouth, etc.), and the Or-nodes represent alternative types of a part (e.g., there are multiple ways to sketch a nose, given by either models or examples). On the AOG, complex spatial constraints of the facial parts can be embedded at the And-nodes at multiple levels, for example, using Markov networks [22], and the photo-sketch similarity measure enforcing the likeness principle can be enforced at multiple resolutions thanks to the hierarchical structure, and optimized by switching at the Or-nodes during rendering.

To create the face AOG, a hierarchical structure is designed by hand with the nodes corresponding to semantic facial parts. Then a set of training sketch examples with their corresponding photographs are collected and manually decomposed corresponding to the structure of the AOG. The decomposed parts are then associated with the nodes in the AOG in order to construct Δ_N and Δ_A for multiple scales in the hierarchy as marked on the right side of Figure 5, with which models for constraining spatial configurations are learned at the And-nodes, and switching
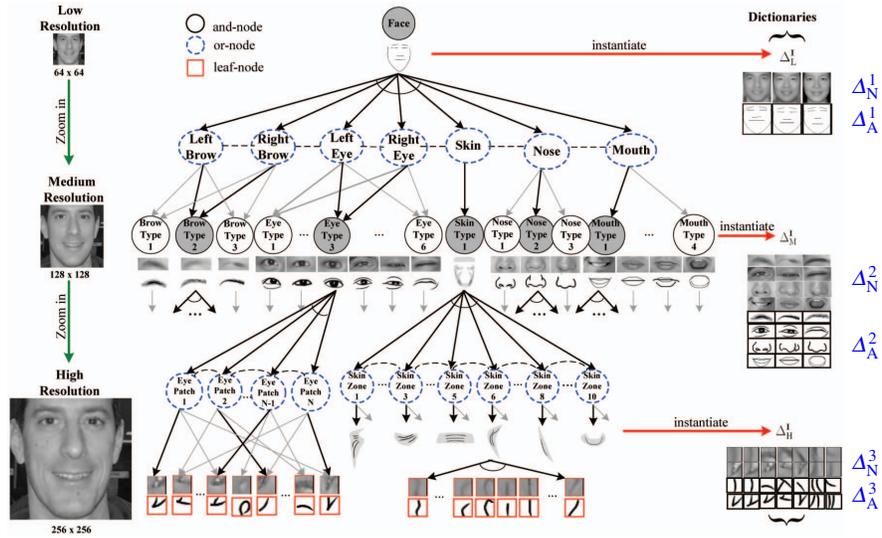
Fig. 5 A three-layer And-Or Graph (AOG) representation for face [24]. Dark arrows represent the paths for generating an instance from the AOG, by decomposing And-nodes (solid ellipses) into sub-components and choosing among alternative types at Or-nodes (dashed ellipses). As marked on the right side, examples in Δ_N and Δ_A of different scales can be embedded in the AOG.
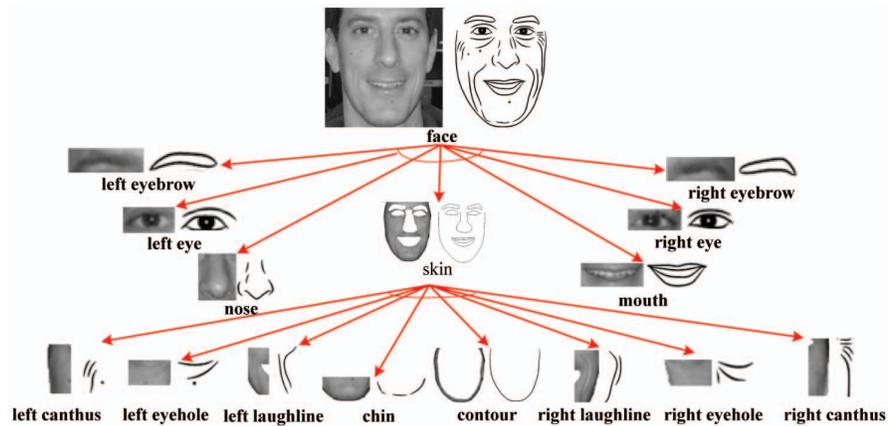

Fig. 6 A parse graph instantiated from the AOG in Figure 5 [24]. This process essentially generates a portrait sketch by selecting sketches of different parts and composing them.

probabilities are summrized at Or-nodes (note: if all examples are preserved in the AOG without similar ones merged, the switching probabilities are uniform).

To generate a portrait sketch from an input photograph, it is equivalent to instantiate a parse graph from the AOG by decomposing at And-nodes and switching at Or-nodes, as shown in Figure 6. An instance optimizing the probabilities at the And- and Or-nodes as well as the similarity measures at multiple resolutions is expected to preserve the fidelity both globally and locally (c.f. the global/local hybrid model introduced by Chen et al. [3]).

Compared with the flat model described in the previous section, the hierarchical AOG has two advantages on the aspect of the face fidelity factor.

1. The AOG can encode spatial constraints more efficiently with the And-nodes at multiple levels on the graph. Furthermore, if multiple selections at Or-nodes are clustered with similar ones merged, the AOG compresses the storage of all training examples, which is important when the training set is large.
2. The hierarchical structure of AOG makes it easier to enforce likeness for different facial parts at multiple resolutions, especially when we prefer the likenesses at different parts/levels to have different weights. For example, the likeness of the eyes could be more important than that of the eyebrows, while the appearance of an entire eye might possibly be more important than those of the eyelids at an unnecessarily higher resolution. In this sense, the power of the face AOG has yet to be fully developed before comprehensive psychological studies on this topic are carried out.

As for the artistic style factor, Xu et al. [23, 22, 24] still used sketch examples created by artists. To improve the visual effects, Min et al. [15] collected more stylish training examples in Δ_A and added two sub-systems for processing the hair and clothes, respectively. Figure 7 includes an example generated by their system.
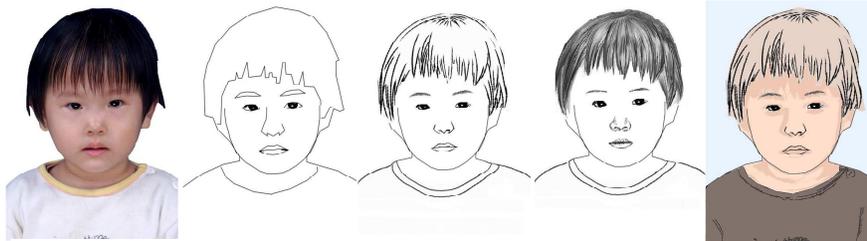

Fig. 7 An example portrait sketch generated by the system of Min et al. [15]

2.3 Sketching the Facial Surface

Besides the concise sketching styles discussed in Sections 2.1 and 2.2, there is another popular sketching style which pays more attention to the facial surface instead of the boundaries and salient edges/curves. This style depicts the appearance of facial surface affected by illumination and shading effects. There are a few studies on this type of portrait sketches in the NPR literature.

Wang and Tang [21] proposed an example-based method for synthesizing portrait sketches with surface strokes depicting the facial appearance. Figure 8 displays the framework of their method.
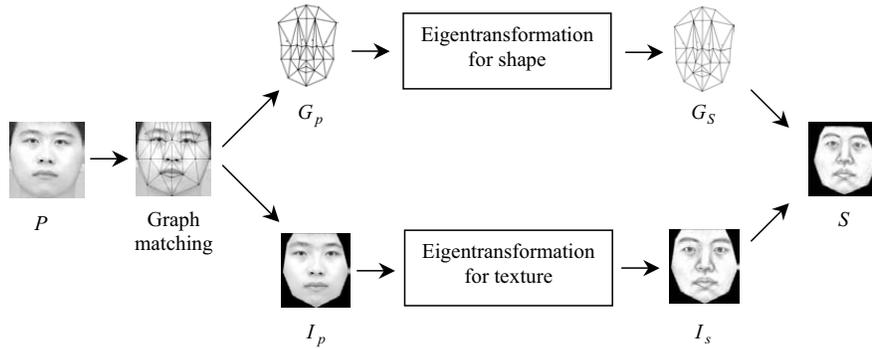

Fig. 8 Framework of the portrait sketch system developed by Wang and Tang [21].

In this method, a triangular mesh is attached to the face, whose vertices are located at the fiducial points on the face, for example, the eyeballs, the corners of the mouth, etc. Given a training set containing pairs of portrait photographs and their corresponding sketches created by artists, the fiducial points are located for each image (both photographs and sketches), and its corresponding triangular mesh is constructed. Then two eigentransformations from photograph to sketch are computed for shape (coordinates of mesh vertices) and texture (grayscale images), respectively. For rendering, given the input photograph, its fiducial points and triangular mesh are computed, with which the image is warped to the mean shape derived from the training set. This essentially separates the shape component G_p and the texture component I_p, which then pass through their respective eigentransformations we computed before to generate the sketch shape G_s and the sketch texture I_s. After warping I_s back to the original shape of the input, we obtain the final sketch.

The global design of the method by Wang and Tang [21] is essentially very similar to that of Chen et al. [3], although the two differ in detailed models and algorithms for shape and texture. They both try to separate the shape and texture information, and define certain mapping relationships for both components to transfer a photograph into a sketch image.

- For shape, Chen et al. adopted ASM, while Wang and Tang used triangular mesh.

- For texture, Chen et al. used an non-parametric sketch model depending on the photograph, while Wang and Tang used an eigentransformation in the image space.

Therefore, the face fidelity and artistic style factors are implemented in a similar way as in the flat model of Chen et al. with different details. The fidelity is enforced using the triangular mesh attached to the face, with shape and texture eigentransformations, and the artistic style is defined by the sketch examples created by artists.

Another interesting work on sketching the facial surface was presented by Tresset and Leymarie [20], as illustrated in Figure 9. In this work, sketches are drawn randomly as Bezier curves at different densities for each color-clustered region in the segmented face area. In this way, the face fidelity is only roughly preserved at a low-resolution level in terms of approximate grayscale levels of different regions, but dropped for most details. Meanwhile, this random sketching process defined in its style Δ_A produces a unique appearance which may appear aesthetic to people.

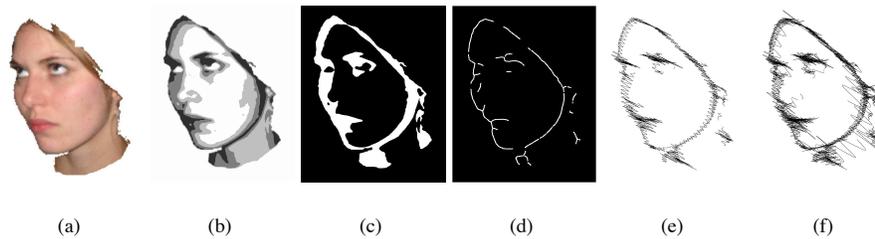| (a) | (b) | (c) | (d) | (e) | (f) |

Fig. 9 Portrait sketch generation using the method of Tresset and Leymarie [20]. The segmented face region in (a) is blurred and the pixel colors are clustered in (b). Then a binary map (c) representing two grayscale levels are generated from (b), and its medial axis is computed in (d). For (c) and (d) the system sketches randomly at different densities using Bezier curves to obtain (e). If replacing (c) and (d) with four levels and sketching correspondingly, we may obtain (f).

3 Paper-Cut

The second portrait genre we review in this chapter is paper-cut, which is essentially a binary image called Mooney images in the psychological literature [8]. But the binarization process is inhomogeneous for different facial areas which usually does not correspond strictly to the grayscale levels. While there have been a few studies on paper-cut in the NPR literature, to the best of our knowledge, the only dedicated portrait paper-cut work is by Meng et al. [14].

The method of Meng et al. adopts the hierarchical AOG face representation we introduced in Section 2.2.2, as shown in Figure 10. Its global design is very similar to that of Min et al. [15], except a few differences in detailed algorithms, and that sketch examples in the latter are replaced by paper-cut versions in the former, as

shown in Figure 11. Beside, Meng et al. also used separated sub-systems for face, hair and the clothes.

Figure 12 includes two results generated by the system of Meng et al. Regarding the fidelity and style factors, the structural information of the face is represented by the AOG model, and the artistic style and the aesthetic rely on the paper-cut examples in Δ_A created by artists.
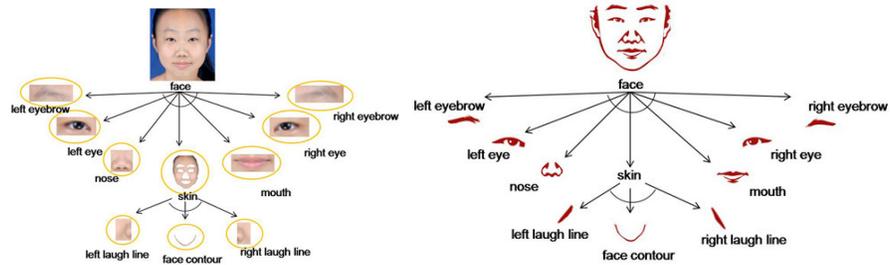

Fig. 10 Parse graph or a face image and its corresponding paper-cut [14].
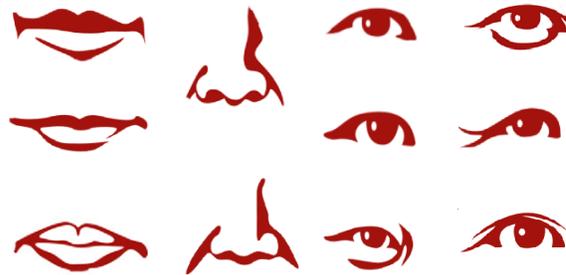

Fig. 11 Paper-cut graphical elements in Δ_A for facial parts used in the system of Meng et al. [14]

Fig. 12 Example paper-cuts generated using the method of Meng et al. [14]

Meng et al. [14] is so far the only work that studied the tradeoff between fidelity and aesthetic in the rendering process. In this work, the fidelity is achieved by a binary proposal generated from the source image using dynamic thresholding, as shown in Figure 13, and the aesthetic level is controlled by the compatibility level of the facial parts in the template dictionary Δ_A. To obtain a tradeoff between the two, a weighted sum of both cost functions

$$d + \lambda c \tag{3}$$

is measured, where d is the difference between the template image and the binary proposal, and c is the number of template instances that the facial parts are selected from, which is a metric of compatibility (assuming parts are more compatible if they are from the same template). By tuning λ it is possible to obtain a continuous spectrum from pursuing only likeness, to considering both with weights, then to pursuing only aesthetic.

Fig. 13 The dynamic thresholding algorithm for computing the binary proposal used by Meng et al. [14]

4 Oil-Painting

Besides sketch and paper-cut, another type of visual art in which portraiture plays an important role is oil-painting. Portrait painting has also been studied in the artistic rendering literature.

Zhao and Zhu [25] developed a system for rendering portrait paintings from photographs using active templates. The main idea of this system is similar to many methods introduced above on sketching. However, due to the much more detailed appearance in paintings than in sketches, the algorithms for depicting these details are especially important for both likeness and aesthetic.

The system of Zhao and Zhu has a few crucial components.

- A dictionary of portrait painting templates. Artists are asked to paint portraits on a screen with digitizer for given photographs, using image-example-based brush strokes as shown in Figure 14. The color statistics (mean and variance), geometry (control points of the backbone curve), and texture (example ID in the brush dictionary [26]) of each stroke for composing the portrait are recorded. Then for each portrait painting we have the complete information about the sequence of brush strokes, as shown in Figure 15b (only part of the strokes are visualized).

A dictionary of these portrait painting examples with their corresponding photographs is constructed, as shown in Figure 16, which are used as templates later in rendering.

- A representation of the spatial configuration of the face and the brush strokes. For each instance in the dictionary, the shape of the face is captured using an active appearance model (AAM) [6] with 83 landmarks. The positions of these landmarks are put in the same coordinate system as the control points describing the geometry of the brush strokes. Shape matching between two different faces (e.g., an input photograph to paint from and a reference example in the dictionary) are achieved by computing a Thin Plate Spline (TPS) transformation [1] between the coordinate pairs of their AAM landmarks.

- A brush stroke rendering algorithm. Given an input face photograph, in order to synthesize a corresponding portrait painting, we first select a reference example from the dictionary of templates. Then we compute the TPS transformation from the shape of the reference example to that of the input photograph. After that, this TPS transformation is applied to the control points of all brush strokes in the reference example, and the output control points with new coordinates defines the new geometry of the strokes for composing the portrait corresponding the input photograph, as shown in Figure 15c. Finally, the color of each brush stroke is transferred to match the color of the target photograph, and the brush strokes are superimposed to compose the result painting image, as shown in Figures 15d and 15e.
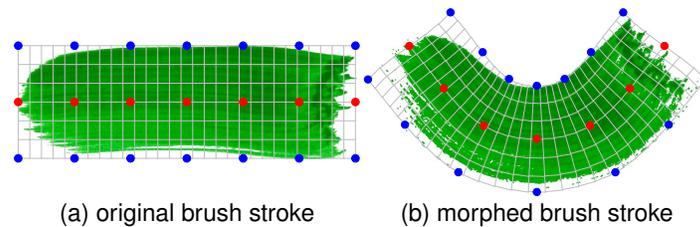

(a) original brush stroke   (b) morphed brush stroke

Fig. 14 The example-based brush stroke model used by Zhao and Zhu [25].

(a)   (b)   (c)   (d)   (e)
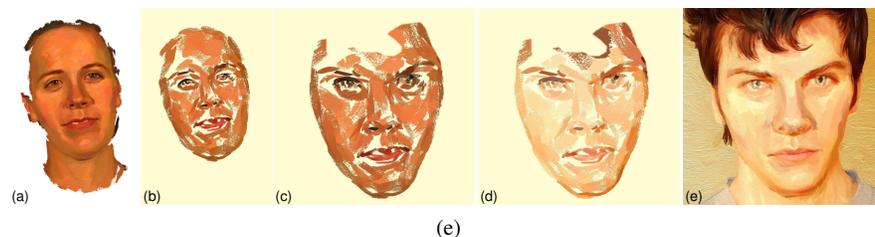
(e)

Fig. 15 Pipline of the portrait painting system of Zhao and Zhu [25]. See explanations in the text.

Fig. 16 Example portrait painting templates from the dictionary constructed by Zhao and Zhu [25].

Fig. 17 Example results generated using Zhao and Zhu's system [25].

Figure 17 displays example results generated using Zhao and Zhu's system. In this system, the face fidelity factor is taken care of by the AAM landmarks and a few detailed algorithms in the portrait painting system.

- Template selection. For a given input photograph, the top-10 best matched examples from the dictionary in terms of both shape and appearance are reported for selection. This avoids using templates differing too much from the target image, which may potentially cause problems in likeness and rendering.
- Shape matching and stroke deformation. Shapes of different faces are matched through the 83 AAM landmark points. Using stroke deformation defined by the matching between landmarks, it ensures the strokes are rendered at appropriate positions with correct curvatures to depict the facial surface and parts.
- Stroke color transfer, which maps the colors of all strokes to the target photograph in a coherent way. This contributes crucially to preserving the global appearance of the original photograph.

The artistic style factor and the aesthetic is supported by the portrait painting dictionary, including

- The sequence of sparse but decisive and colorful strokes which, while put together, convey an impression of 3D structures and vibrant contrasts, and
- The individual textured brush strokes, which deliver elegant oil-painting details with illumination and shading.

5 Caricature

Caricature is a very special type of portrait images. It differs from traditional sketch, paper-cut and painting in the way of manipulating the two factors we study in this chapter. Most caricatures *explicitly* trade certain degree of fidelity for unique aesthetic effects, for example, by exaggerating features in several parts of the face.

Tominaga et al. [18, 19] developed the PICASSO facial caricaturing system. In the PICASSO system, 445 characteristic points are located on the edges of facial parts, and various expressions are defined as offsets of these points from a mean face (without expression), as shown in Figure 18. The idea of delibrately manipulating the facial shape for caricaturing and further exaggeration is the basis of most later studies on portrait caricatures.
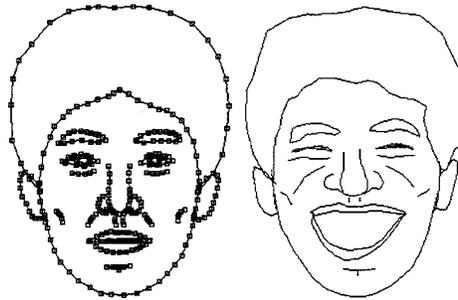

Fig. 18 In PICASSO [18, 19], 445 characteristic points with contour lines are used for representing the facial shape, and a static representation of expressions is by offseting these points.

As an improvement over PICASSO, Obaid et al. [17] also presented a system for rendering and animating expressive caricatures. Their system subdivides the face into 16 muscle-based regions, and models different expressions as second-degree rubber-sheet deformations of the 16 regions. In this way, the expressions can be exaggerated by manipulating the coefficients in the rubber-sheet transformation. An example is shown in Figure 19.

Liang et al. [12] extended the sketching method of Chen et al. [2] with exaggeration for generating facial caricatures, whose framework is shown in Figure 20.

In its prototype-based shape exaggeration model, the training examples are analyzed and clustered into a few exaggeration prototypes. Each prototype represents a trend of exaggeration in some facial features, towards which the sketch image is warped using a local linear model at runtime to generate caricatures. Figure 21 displays two examples of exaggerated caricature generated by this system.
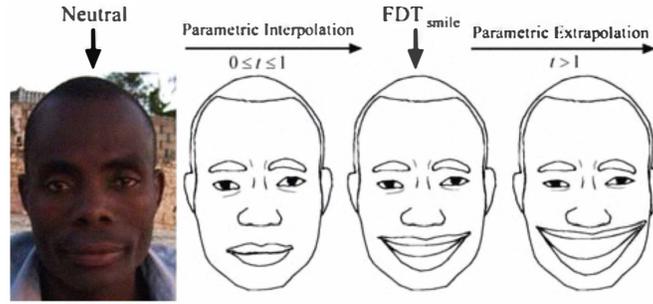
Fig. 19 Caricature exaggeration levels of the smile facial expression, by Obaid et al. [17]
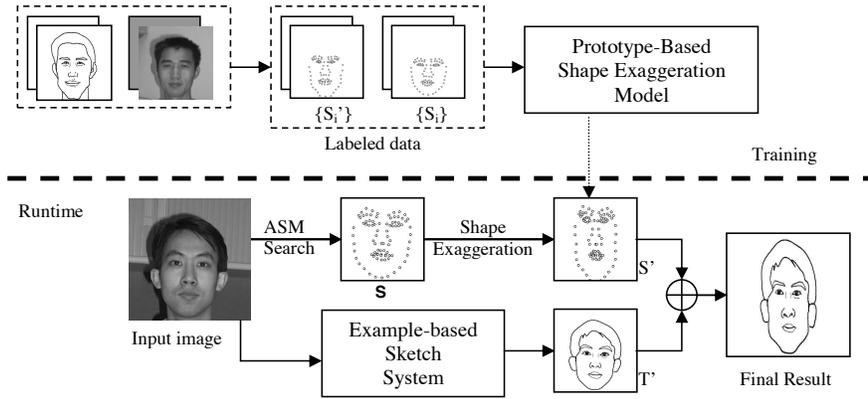

Fig. 20 Framework of the example-based caricature generation system of Liang et al. [12]

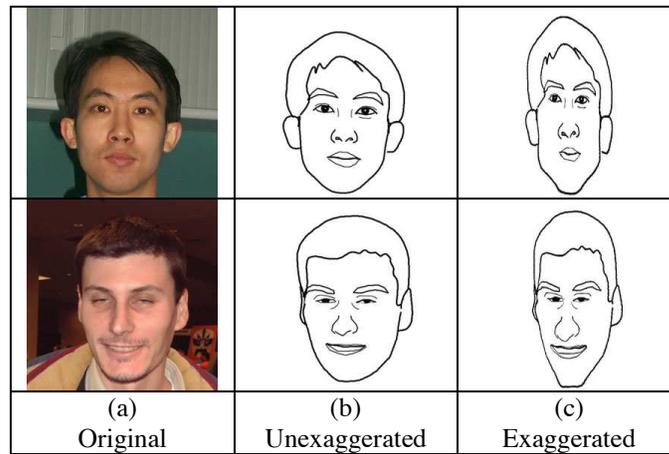| (a)
Original | (b)
Unexaggerated | (c)
Exaggerated |

Fig. 21 Two examples of exaggerated caricature generated by the system of Liang et al. [12]

The additional exaggeration module to the system of Chen et al. essentially reduces the level of face fidelity while enhancing features exaggerated in training examples which reflects a type of aesthetic.

Luo et al. [13] developed a semi-automatic caricature system which enables manually capturing much finer details of the structures of facial parts in W_A, such as the eyeballs and the bridge of the nose, as shown in Figure 22a. An example of their result is shown in Figure 22b.
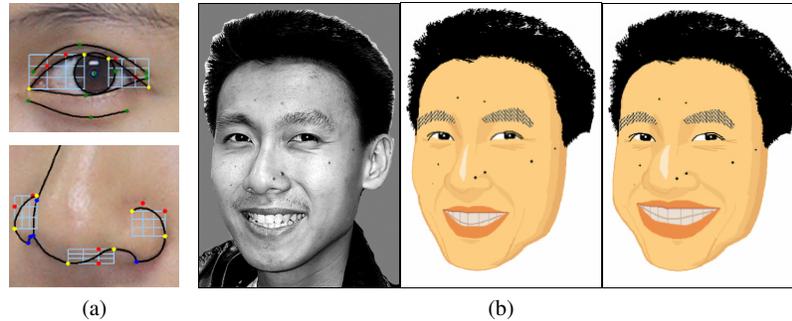

(a)   (b)

Fig. 22 (a) Structure models of the eye and nose captured by Luo et al. [13] (b) An example result generated by their system. From left to right: input photograph, unexaggerated caricature, and exaggerated caricature.

.

Mo et al. [16] noticed that the exaggeration level should depend not only on the absolute difference from the mean shape, but also the variance level of such differences among examples. This idea enables the comparison between the exaggeration levels of different facial parts which corresponds better to our perception. Figure 23 displays two of they results.

Fig. 23 Two example results generated using the method of Mo et al. [16] From left to right: input photographs, photorealistic caricatures, and stylized caricatures.

Fig. 24 An example result generated using the caricature algorithm of Chiang et al. [4]

Another interesting work on caricature was presented by Chiang et al. [4], which adds exaggeration upon a mesh-based representation for facial geometry (similar to Wang and Tang [21]) instead of the contour-based ones used by Tominaga et al. [18, 19], Obaid et al. [17], and Liang et al. [12]. This enables more details in color and texture in the rendered caricature, as shown in Figure 24.

6 Summary

In this chapter, by reviewing recent work on artistic rendering of portraits, we have studied its two essential factors, namely the face fidelity factor supported by certain face representations stored in W_A, and the artistic style with graphical elements and rendering details defined in Δ_A, as introduced in Section 1. In most methods, these two factors are implemented in a few different solutions.

- The fidelity is usually enforced at two different levels. At the local level of details of facial parts, certain similarity measures are adopted between the photograph and the artistic depiction or implicitly applied by defining a mapping/transformation between them. At the global level of the face, certain shape models are adopt such as ASM, AAM or triangular mesh to constrain the spatial configuration of the facial parts. Besides, the And-Or Graph model has also been applied to integrate the two levels (or more levels) in a single hierarchy.
- The artistic style is usually applied by using training example art pieces or elements created by artists, in addition to some simple rule-based strategies for straightness, smoothness, etc. In caricature, its unique style is mainly attributed to the exaggerated features which are also learnable from training examples.

Guided by the two factors, past studies on artistic rendering of portraits took various ways to the pursuit of likeness and aesthetic in portraiture.

Despite the progress, a few key questions still remain unclear, which stand in our way to a systematic theory and solution to the portrait rendering problem.

- How should we select the important information from W_N to be used in W_A? In practice, this means to choose a facial model that captures features essential to artistic perception and ingores unimportant parts. Also, what is the minimum information we need in W_A to satisfy the likeness principle?
- Is there a principled method for balancing between likeness and aesthetic, or can these two be separated for independent manipulation under certain circumstances? The study of Leopold et al. [9] gives us some hints on this question, which tells that extrapolation against a mean face can preserve a person's identity without being confused with other people in the experimental dataset. This supports the caricature rendering methods based on shape exaggeration which improves the aesthetic without giving up the likeness. Meanwhile, a comprehensive investigation to this problem has yet to be conducted.

References

1. Barrodale, I., Skea, D., Berkley, M., Kuwahara, R., Poeckert, R.: Warping Digital Images Using Thin Plate Splines. Pattern Recog. 26(2):375–376 (1993)
2. Chen, H., Xu, Y.-Q., Shum, H.-Y., Zhu, S.-C., Zheng, N.-N.: Example-Based Facial Sketch Generation with Non-Parametric Sampling. In: Proceedings of ICCV (2001)
3. Chen, H., Liu, Z., Rose, C., Xu, Y., Shum, H.-Y., Salesin, D.: Example-Based Composite Sketching of Human Portraits. In: Proceedings of NPAR (2004)
4. Chiang, P.-Y., Liao, W.-H., Li, T.-Y.: Automatic Caricature Generation by Analyzing Facial Features. In: Proceedings of ACCV (2004)
5. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active Shape Model — Their Training and Application. Comput. Vis. Image Understand. 61:38–59 (1995)
6. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. 23(6):681–685 (2001)
7. Gooch, A., Gooch, B.: Non-Photorealistic Rendering. A K Peters (2001)
8. Hegdé, J., Thompson, S., Kersten, D.: Identifying Faces in Two-tone ('Mooney') Images: A Psychophysical and fMRI Study. J. Vision 7(9):Article 624 (2007)
9. Leopold, D.A., O'Toole, A.J., Vetter, T., Blanz, V.: Prototype-referenced Shape Encoding Revealed by High-level Aftereffects. Nat. Neurosci. 4(1):89–94 (2001)
10. Li, Y., Kobatake, H.: Extraction of Facial Sketch Images and Expression Transformation Based on FACS. In: Proceedings of ICIP (1995)
11. Li, Y., Kobatake, H.: Extraction of Facial Sketch Image Based On Morphological Processing. In: Proceedings of ICIP (1997)
12. Liang, L, Chen, H., Xu, Y.-Q., Shum, H.-Y.: Example-based Caricature Generation with Exaggeration. In: Proceedings of Pacific Graphics (2002)
13. Luo, W.C., Liu, P.C., Ouhyoung, M.: Exaggeration of Facial Features in Caricaturing. In: Proceedings of International Computer Symposium (2002)
14. Meng, M., Zhao, M., Zhu, S.-C.: Artistic Paper-Cut of Human Portraits. In: Proceedings of ACM Multimedia (2010)
15. Min, F., Suo, J.-L., Zhu, S.-C., Sang, N.: An Automatic Portrait System Based On And-Or Graph Representation. In: Proceedings of EMMCVPR (2007)
16. Mo, Z., Lewis, J.P., Neumann, U.: Improved Automatic Caricature by Feature Normalization and Exaggeration. In: SIGGRAPH Sketch (2004)
17. Obaid, M., Mukundan, R., Billinghurst, M.: Rendering and Animating Expressive Caricatures. In: Proceedings of ICCSIT (2010)

18. Tominaga, M., Fukuoka, S., Murakami, K., Koshimizu, H.: Facial Caricaturing with Motion Caricaturing in PICASSO system. In: Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics (1997)
19. Tominaga, M., Hayashi, J.-I., Murakami, K, Koshimizu, H.: Facial caricaturing system PICASSO with emotional motion deformation. In: Proceedings of the 2nd International Conference on Knowledge-Based Intelligent Electronic System (1998)
20. Tresset, P., Leymarie, F.F.: Generative Portrait Sketching. In: Proceedings of VSMM (2005)
21. Wang, X., Tang, X.: Face Sketch Synthesis and Recognition. In: Proceedings of ICCV (2003)
22. Xu, Z., Luo, J.: Accurate Dynamic Sketching of Faces from Video. In: Proceedings of CVPR (2007)
23. Xu, Z., Chen, H., Zhu, S.-C.: A High Resolution Grammatical Model for Face Representation and Sketching. In: Proceedings of CVPR (2005)
24. Xu, Z., Chen, H., Zhu, S.-C., Luo, J.: A Hierarchical Compositional Model for Face Representation and Sketching. In: IEEE Trans. Pattern Anal. Mach. Intell. 30(6):955–969 (2008)
25. Zhao, M., Zhu, S.-C.: Portrait Painting Using Active Templates. In: Proceedings of NPAR (2011)
26. Zeng, K., Zhao, M., Xiong, C., Zhu, S.-C.: From Image Parsing to Painterly Rendering. ACM Trans. Graph. 29(1):Article 2 (2009)