

---

# Facial Image Processing

Xiaoyi Jiang<sup>1</sup> and Yung-Fu Chen<sup>2</sup>

<sup>1</sup> Department of Mathematics and Computer Science  
University of Münster, Germany  
[xjiang@math.uni-muenster.de](mailto:xjiang@math.uni-muenster.de)

<sup>2</sup> Department of Health Services Administration  
China Medical University, Taichung 404, Taiwan  
[yungfu@mail.cmu.edu.tw](mailto:yungfu@mail.cmu.edu.tw)

**Summary.** Faces are among the most important classes of objects computers have to deal with. Consequently, automatic processing and recognition of facial images have attracted considerable attention in the last decades. In this chapter we focus on a strict view of facial image processing, i.e. transforming an input facial image into another and involving no high-level semantic classification like face recognition. A brief overview of facial image processing techniques is presented. Typical applications include removal of eyeglasses, facial expression synthesis, red-eye removal, strabismus simulation, facial weight-change simulation, caricature generation, and restoration of facial images.

## 1 Introduction

In communication between people, faces play one of the most important roles. Faces allow to recognize a person's identity but moreover they carry rich information about a person's emotional state. We are able to sense the smallest differences in the appearance of a face and easily sense irony, the slightest disagreements or understanding. Thus, faces serve as an information carrier in a much more subtle manner than identity or emotions and help us to adapt our behavior. Not surprisingly, faces have been long a research topic in psychology.

Automatic processing and interpretation of facial images has attracted much attention in the last decades. Without doubt the most prominent research topic in this context is face recognition and verification in still images and videos [8, 26, 30, 65]. A fundamental issue in any face analysis system is to detect the locations in images where faces are present. A lot of work has been done on face localization which is often a preprocessing step within recognition systems [19, 61]. Moreover, many methods have been developed to detect facial features such as eyes, nose, nostrils, eyebrows, mouth, lips, ears, etc. [65]. These features can either be used for a direct recognition approach or

to normalize facial images for holistic recognition approaches like eigenfaces and Fisherfaces.

In this chapter we will take a strict view of facial image processing only, i.e. transforming an input facial image into another. Several tasks fall into this category, including removal of eyeglasses, facial expression synthesis, red-eye removal, strabismus simulation, facial weight-change simulation, and caricature generation. They are all image processing tasks and only involve very limited high-level semantic reasoning.

The motivation for these facial image processing tasks is manifold:

- Improvement of facial image quality  
A typical representative of this class is red-eye removal.
- Preprocessing for face recognition  
One crucial requirement on successful face recognition is robustness to variations arising from different lighting conditions, poses, scales, or occlusion by other objects. Glasses belong to the most common occluding objects and have a significant effect on the recognition performance. While it is possible to identify and use non-occluded local regions only, we may want to estimate the non-occluded facial image by removing the occluding objects, e.g. glasses.
- Simulation of effects  
Given a facial image, the task is to simulate expressions, strabismus, or to draw caricatures.

Facial image processing operations can be found as a part of various application scenarios. For instance, facial expression synthesis may serve to enrich face databases for improved recognition performance.

Compared to other images, processing of facial images is particularly delicate. We see and interpret many faces everyday and thus are continuously trained to successfully distinguish between a large number of different faces from the childhood on. As a consequence, we are very sensitive to small distortions and changes in the appearance of faces. The so-called Thatcher illusion [53] is an excellent example to illustrate this point, see Figure 1. In this illusion, the eyes and mouth of a face are inverted. The result looks grotesque in an upright face. However, when shown inverted, the face looks fairly normal. Although we do observe some distortions between the two facial images, the perceptual divergence is much smaller than seen upright. This and other related observations indicate that our brain tends to perceptually amplify the quantitatively measurable differences when seeing faces. In fact, it is the wonderful ability of the human's visual system in face perception that makes the automatic facial image processing so difficult.

While considerable survey work has been done in the past for face detection and face recognition, the goal of this chapter is to give a brief overview of facial image processing techniques. It is not our intention to provide a thorough review. Instead, we focus on the tasks mentioned before. For these tasks we



**Fig. 1.** Thatcher illusion. A normal face and a manipulated version of Thatcher (left); the same for Clinton

give a motivation, present the most recent and important methods, summarize the current state of research, and discuss the directions of future research.

## 2 Removal of glasses

The first publications concerning glasses in facial images are limited to the existence decision and position detection of glasses only. Jiang et al. [22] define six measures for classifying the presence of glasses in a facial image without suggesting a classifier. Wu et al. [58] go a step further and devise a sophisticated classifier for the existence of glasses based on support vector machines. For this classifier a recognition performance of 90 percent is reported.

Besides the existence decision similar to [22], Jing and Mariani [24] also extract the contour of glasses using a deformable model combining edge features and geometric properties. In [60] the extraction is done by decomposing face shape using the Delaunay triangulation. A 3D technique is reported in [59]. The authors perform a 3D Hough transform in trinocular stereo facial images based on the determination of a 3D plane passing through the rims of the glasses, without any assumption of the facial pose or the shape of glasses. Certainly, 3D information can make a substantial contribution to glasses localization. However, we pay a price of multiple cameras and the computational costs for 3D reconstruction.

The work [46] is among the first ones on the removal of glasses. The fundamental assumption is the availability of a set of  $M$  glasses-free images. Implicitly, it is further assumed that these prototype images are normalized to the same size and properly aligned. Let  $F_0$  be the average of all glasses-free images and  $F_1, F_2, \dots, F_{M-1}$  the eigenfaces from a principal component analysis (PCA). In some sense the  $M - 1$  eigenfaces span the space of glasses-free images. Given an input image  $I_g$  with glasses, then the corresponding glasses-free image  $I_f$  is computed by a PCA reconstruction method

$$I_f = F_0 + \sum_{i=1}^{M-1} ((I_g - F_0) \cdot F_i) \cdot F_i \quad (1)$$

It is not untypical that this simple PCA reconstruction cannot remove all traces of glasses.

The problem is partly caused by the fact that the eigenfaces represent a space, which the input image  $I_g$  does not belong to. This gives rise to some modeling inaccuracy which remains visible after the reconstruction. An obvious solution lies in improving the modeling accuracy by using  $M$  prototype images with glasses instead. Denoting the average of these prototypes by  $G_0$  and their eigenfaces by  $G_1, G_2, \dots, G_{M-1}$ , an input image  $I_g$  can be well represented by:

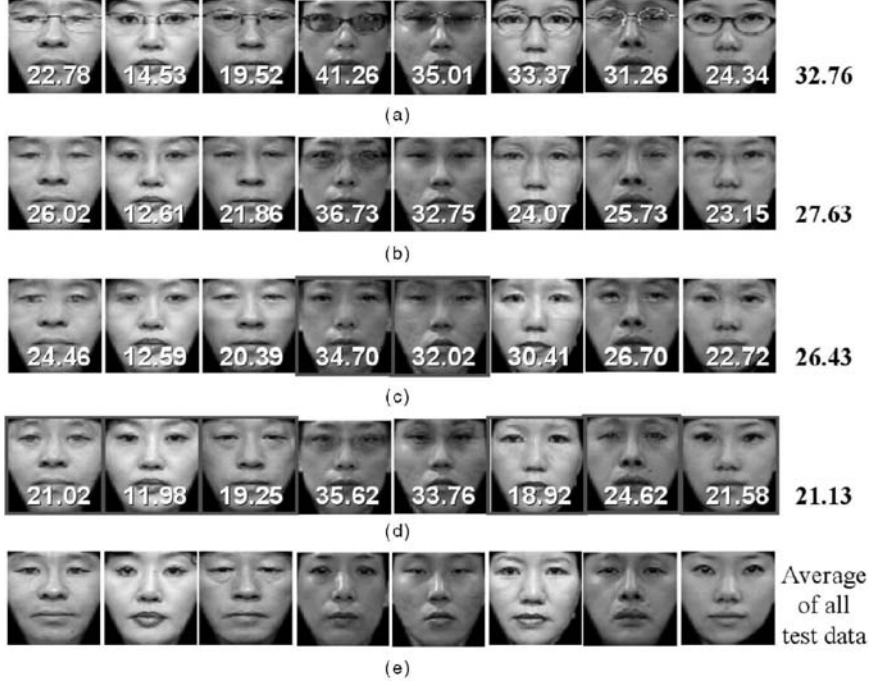
$$I_g = G_0 + \sum_{i=1}^{M-1} c_i \cdot G_i \quad (2)$$

in the space of facial images with glasses, where  $c_i = (I_g - G_0) \cdot G_i$  holds. Assume that a set of prototype pairs be available, i.e. there exist one glasses-free image and one glasses image for the same person in the example set. Then, we could construct an image  $I_f$  by retaining the coefficients  $c_i$  and replacing  $G_0$  and  $G_i$  in Eq. (4) by  $F_0$  and  $F_i$  respectively.

The example-based reconstruction method [42] follows this line. However, the PCA computation is done on  $M$  prototype images, each resulting from a concatenation of a glasses image and a glasses-free image of the same person. The single average image of these images can be decomposed into two separate average images  $G_0$  and  $F_0$  again. Similarly, the resulting eigenfaces can be decomposed into two separate sets of images  $G = \{G_1^*, G_2^*, \dots, G_{M-1}^*\}$  and  $F = \{F_1^*, F_2^*, \dots, F_{M-1}^*\}$ . Note that while  $G_0$  and  $F_0$  are identical to those before, the other images  $G_i^*$  and  $F_i^*$  differ from the eigenfaces of glasses images and glasses-free images, respectively, due to the different ways of PCA computation. Given an input image  $I_g$ , we look for its optimal representation in the space of facial images with glasses spanned by set  $G$  and then apply the representation coefficients to set  $F$  for constructing a glasses-free image  $I_f$ . In this case, however, the images  $G_i^*$  are no more orthogonal to each other and thus a least square minimization is required to obtain the optimal representation of  $I_g$  in terms of  $G$ .

A refined algorithm by using recursive error compensation is presented in [42]. It is based on the same idea of PCA reconstruction (in terms of glasses-free prototype images). An iterative scheme is defined, where in each iteration, the pixel values are adaptively computed as a linear combination of the input image and the previously reconstructed pixel values.

Some results for the three removal methods (simple PCA reconstruction [46], example-based reconstruction [42], and recursive error compensation [42]) are shown in Figure 2. Visually, the error compensation method seems to produce the best results; the images in Figure 2(d) have no traces of glasses and look seamless and natural. This impression is also confirmed by a quantitative error measure. The number on each facial image indicates the average pixel-wise distance to the corresponding original facial image without glasses. The numbers in the last column represent the mean of pixel-wise



**Fig. 2.** Examples of glasses removal: (a) input images with glasses; (b) simple PCA reconstruction method; (c) example-based reconstruction method; (d) recursive error compensation; (e) original faces without glasses. (Courtesy of J.-S. Park, © 2005 IEEE)

distances of a total of 100 test images. The recursive error compensation method tends to have the smallest error measure.

Another advanced glasses removal technique is presented in [59]. Given an image  $I_g$  with glasses, the glasses-free reconstruction is based on the maximum a posteriori (MAP) criterion:

$$I_f^* = \arg \max_{I_f} p(I_f | I_g) = \arg \max_{I_f} p(I_g, I_f)$$

The joint distribution  $p(I_g, I_f)$  is estimated by means of a set of aligned pairs of glasses and glasses-free images of the same person.

An assessment of glasses removal quality depends on applications. For image editing purpose the only criterion is simply to which degree the reconstructed image is seamless and natural looking. In the context of face recognition, however, the same question must be answered in a task-based evaluation manner. That is, we conduct face recognition experiments using both original glasses-free and reconstructed glasses-free images. The difference in recognition rate is the ultimate performance measure for glasses removal. Currently, very few work has been done following this line; only [42] reports a small-scale

study. More work is needed to demonstrate the potential of this preprocessing step in real-world face recognition systems.

### 3 Facial expression synthesis

Facial expressions play a major role in how people communicate. They serve as a window to one's emotional state, make behavior more understandable to others, and they support verbal communication. A computer that is able to interact with humans through facial expressions (in addition to other modalities) would greatly advance human-computer interfaces. This ability includes both understanding of facial expressions and their synthesis. The semantic analysis of facial expressions is not topic of this chapter and the readers are referred to [14, 15, 28].

Several algorithmic paradigms for facial expression synthesis can be found in the literature [28, 43]. One class of methods are variants of the morph-based approaches [6, 49]. They can only generate expressions between two given images of the same person and their ability of generating arbitrary expressions is thus more than limited. If merely one image of a person is available, these approaches are not applicable at all.

Another popular class of techniques is known as expression mapping (performance-driven animation) [32, 43]. Its principle is quite simple: Given an image  $A$  of a person's neutral face and another image  $A'$  of the same person's face with a desired expression, the movement of facial features from  $A$  to  $A'$  is geometrically performed on a second person's neutral image  $B$  to produce its facial image  $B'$  with the expression. The major algorithmic steps are:

- Find the facial features (eyes, eyebrows, mouth, etc.) in  $A$ ,  $A'$ , and  $B$ , either manually or by some automatic method.
- Compute the difference vectors between the feature positions of  $A$  and  $A'$ .
- Move the features of  $B$  along the difference vector of corresponding feature of  $A$  and warp the image to a new one  $B'$  accordingly.

This technique is a geometry-driven feature mapping combined with image warping. As such, its applicability includes animation of 2D drawings and images far beyond facial expression synthesis.

The image  $B'$  produced by this expression mapping technique usually looks reasonable, since the transformation from  $B$  to  $B'$  captures all necessary geometric changes that are needed to copy the facial expression as exemplified by  $A$  and  $A'$ . However, the algorithm totally ignores the photometric changes and thus the result image lacks details such as the wrinkles on the forehead. An attempt of including photometric changes is made in [34]. Assume that all images  $A$ ,  $A'$ , and  $B$  are geometrically aligned. Using the Lambertian reflectance model, the authors show that the relationship:

$$B'(u, v) = B(u, v) \cdot \frac{A'(u, v)}{A(u, v)}$$

holds for each image position  $(u, v)$ . It tells us how  $B'$  is photometrically related to  $B$  in a way so that the same geometric changes between  $A$  and  $A'$  are done between  $B$  and  $B'$ . This relationship can be used to extend the traditional geometric warping method to the following algorithm:

- Perform traditional geometric warping on images  $A$ ,  $A'$ , and  $B$  to compute an image  $B'$ .
- Align  $A$  and  $A'$  with  $B'$  through image warping and denote the warped images by  $\tilde{A}$  and  $\tilde{A}'$ , respectively.
- Compute the photometrically corrected final result  $B^*(u, v) = B'(u, v) \cdot \frac{\tilde{A}'(u, v)}{\tilde{A}(u, v)}$  for each pixel  $(u, v)$ .

The result image  $B^*$  has exactly the same geometry as the initial result  $B'$ . It also respects the photometric changes from  $A$  to  $A'$  in addition to the geometric changes. Among others, it is very effective in generating wrinkles on the forehead.

In the recent work [63] quite different assumptions are made. The authors assume that a set of images of the same person is available. If these images cover enough different facial expressions, then they span in some sense the space of expressions (of that particular person) and any other expression can be represented as their convex combination [45]. Let  $G_i$ ,  $i = 1, 2, \dots, m$ , denote the geometry of the  $i$ -th example expression image  $I_i$  (vector of positions of feature points). Then, the geometry  $G$  of an arbitrary new expression image is represented as a convex combination:

$$G = \sum_{i=1}^m c_i \cdot G_i$$

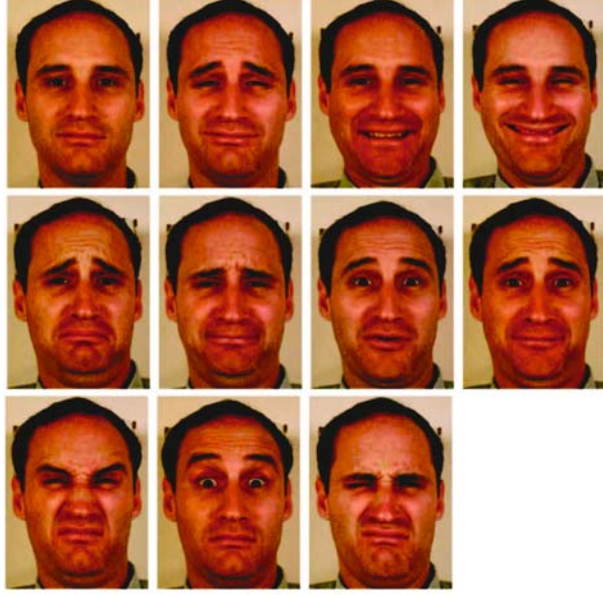
where the coefficients  $c_i$  satisfy the condition  $\sum_{i=1}^m c_i = 1$ . The optimal coefficients are determined by solving the optimization problem:

$$\text{minimize: } \|G - \sum_{i=1}^m c_i \cdot G_i\|^2, \quad \text{subject to: } \sum_{i=1}^m c_i = 1, \quad c_i \geq 0$$

which is a quadratic programming problem. If the  $m$  example images themselves are  $I_i$ ,  $i = 1, 2, \dots, m$ , then an image  $I$  corresponding to the desired expression geometry  $G$  can be composed by:

$$I = \sum_{i=1}^m c_i \cdot I_i$$

Since the example images are assumed to be aligned, this step is simply a pixel-wise color blending. This algorithm can be used for expression editing,



**Fig. 3.** Eleven example images of a male person. (Courtesy of Z. Liu, © 2006 IEEE)

where the geometry of the desired expression is, for instance, constructed using some editing tool. Figure 3 shows the set of example images. The six expression editing results in Figure 4 indicate the algorithm’s ability of generating expression images of a different geometry than any of the example images. A second application scenario is expression mapping, where another person’s facial image serves as the source of the geometry  $G$  and the algorithm mimics the expression of that person. In Figure 4 the right column shows image pairs of a female and a male face. The male faces are synthetic images generated by taking over the geometry of the female image to synthesize a similar expression. Interestingly, this technique has been extended in [63] to synthesize facial expressions of 3D head models.

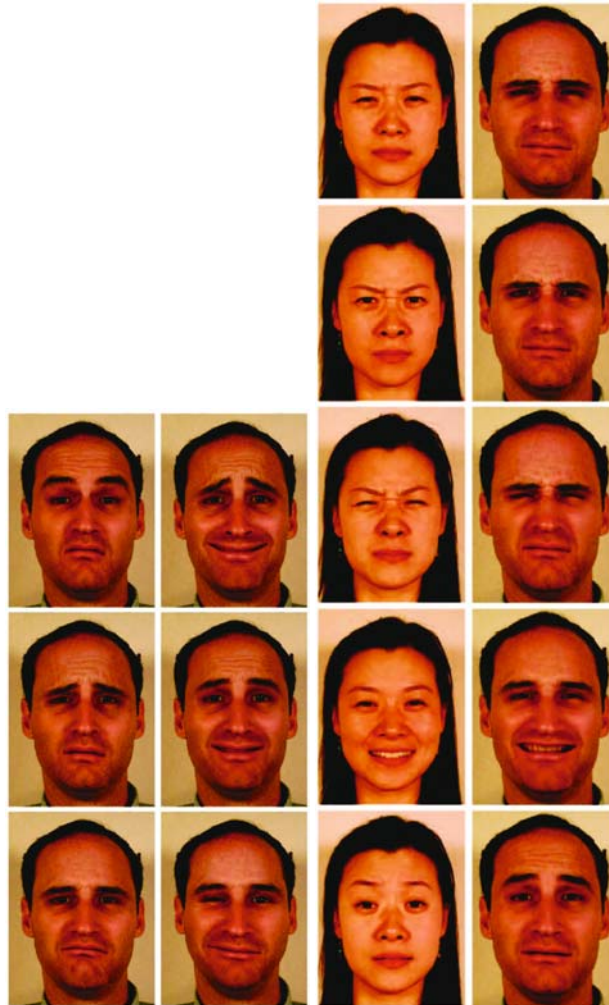
Facial expression synthesis has attracted many researchers in the last years [1, 2, 7, 13, 18, 62, 66]. It is expected to have substantial impact on facial image editing and human-computer interfaces.

#### 4 Eye synthesis

The most important perceptual feature of a face is probably the eye as the eye appearance plays a central role in human communication. Consequently, eye synthesis is of interest in several contexts.

Face recognition is confronted by significant natural variations, e.g. lighting conditions, image size, etc. Although some factors like image size can be





**Fig. 4.** Expression editing: Six synthetic images based on editing the geometry (left). Expression mapping: The male's synthetic image is generated by taking over the geometry of the female (right). (Courtesy of Z. Liu, © 2006 IEEE)

alleviated by geometric and photometric normalization, the availability of a collection of training images covering other variations and appearances is of great help. Since it is hardly possible to capture all possible variations, one solution lies in their synthesis. In [25] the authors investigated a number of operators for changing the eye shape such as:

- Pull down or push up of the eyebrow (without any change to eye location)
- Rotation of the entire eyebrow around its middle

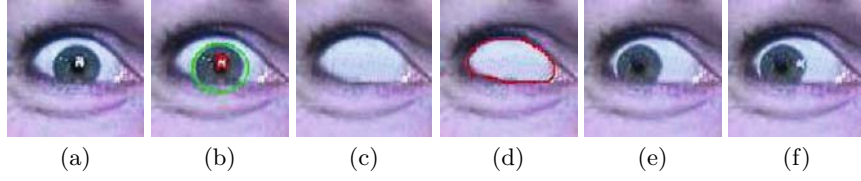
- Pull down or push up of the upper eyelid so that the eye appears less or more open
- Pull down or push up of the lower eyelid so that the eye appears more or less open

Although these operations do not produce precise anatomical changes, they seem to be adequate for appearance-based systems. The synthesized images artificially enlarge the training set of a face recognition system in order to represent the entire space of possible variations more completely. In [25] the authors show that this enrichment indeed improves the recognition performance.

Gaze redirection is another important application of eye synthesis. In video-conferencing and other visual communication scenarios, the participant watches the display rather than the camera. This turns the gaze direction away from the conversation partner and impairs the desired eye contact. Several hardware solutions have been proposed to alleviate the problem. Sellen [48] places the display/camera unit sufficiently far away from the user, so that gazing at the screen becomes indistinguishable from gazing at the camera. A more sophisticated solution is suggested in [40], where the camera is positioned behind a semi-transparent display. Due to the need of special hardware, however, this kind of solutions has found limited acceptance so far. The authors of [56] propose to artificially redirect the gaze direction. The eye-white and the iris are detected in a facial image and replaced by corresponding parts from a real eye image at desired position.

Synthesizing strabismic face images, each of a different angle, based on a normal frontal face image is needed for conducting studies in psychosocial and vocational implications of strabismus and strabismus surgery. Strabismus can have a negative impact on an individual's employment opportunity, school performance, and self-perception. A recent study [11], for instance, indicates that large-angle horizontal strabismus appeared to be vocationally significant particularly for female applicants, reducing their ability to obtain employment. One possibility of synthesizing strabismic face images lies in manual image editing using standard image processing packages [11]. Alternatively, one may ask a person to simulate different strabismic angles. Both methods are tedious and time-consuming, thus not applicable if many face images and various strabismic angles are needed. The latter approach has an additional problem that it is hardly possible for a person to precisely simulate a particular desired angle. The work [23] proposes an algorithm for synthesizing strabismic face images of an arbitrary angle. It consists of the following main steps (see Figure 5):

- Detection of the contour of the iris and the reflection point
- Removal of the iris
- Detection of the contour of the eye
- Rotation of the eye, i.e. re-insertion of the iris
- Placement of the reflection point



**Fig. 5.** Main steps of strabismus synthesis algorithm: (a) (part of) input image; (b) detected contour of iris and reflection point; (c) removal of iris; (d) detected eye contour; (e) rotated eye; (f) embedded reflection point. The final result is given in (f)



**Fig. 6.** Results of strabismus simulation. Top: input image. Bottom: 20° and 40° to the right, 20° and 40° to the left, and strabismus with a vertical angle (from left to right). Only the right eye of the person is processed

All three contour detection subtasks (iris, reflection point, and eye) are solved by a dynamic programming technique. The iris removal should be done with care. Typically, the eyelashes partially interlay with the iris so that a straightforward removal and subsequently filling this area by the eye background would produce unnatural appearance in the eyelash. Instead, one has to fill the missing background and continue the missing eyelashes in a natural way simultaneously. In [23] this step is performed by means of an image inpainting algorithm. Despite of strabismus the reflection point should remain unchanged. Therefore, as the last operation the detected reflection point is embedded exactly at the same coordinates as in the input image. Accordingly, the reflection point has to be removed from the source image before it is re-inserted into the inpainted background without iris. Figure 6 shows some results of strabismus simulation.

The reversed direction of correction of strabismic face images is of great interest in plastic surgery applications. This gives the patient an approximate

post-operation look of strabismus surgery in the pre-operation phase. In fact, both strabismus simulation and correction are closely related to gaze redirection. The main difference lies in the required image quality. Psychosocial studies tend to need static pictures of higher quality than typically in communication scenarios.

## 5 Redeye removal

Redeye is a common problem in consumer photography. When a flash is needed to illuminate the scene, the ambient illumination is usually low and a person's pupils will be dilated. Light from the flash can thus reflect off the blood vessels in the person's retina. In this case it appears red in color and this reddish light is recorded by the camera.

One possible hardware technique to avoid redeye is to increase the distance between the flash unit and the camera lens. Another popular solution is the use of pre-exposure flashes. A pre-exposure flash will contract the person's pupil and thus reduce the chance that light reflected off the retina will reach the lens. The drawback of this approach is that people will sometimes close their eyes by reflex and the substantial need of power; additional flashes further lower the battery life. In addition the red-eye artifacts are reduced, but not completely eliminated.

With the advent of digital photography, software solutions become popular. While several photo editing software packages allow for manual correction of redeye, they tend to be semiautomatic and do not always give satisfactory results. Today, there exist several patents for detecting and removing redeyes (see listings in [38, 47, 51]).

All research work published in the literature has a modular structure of redeye detection followed by a removal operation. As for redeye detection, the two papers [16, 17] are based on a face detection. This narrows the search area for redeye artifacts to facial regions only. In addition the face information can be used to infer properties like the location and size of possible redeye artifacts. Within a face region, features such as redness and changes in luminance and redness are then applied to find redeye parts.

In [55] an active appearance model is trained using a set of typical redeye subimages. Given an input image, color cues are first used to locate potential redeye regions. Then, a matching is performed to minimize the difference between a candidate image part and one synthesized by the appearance model.

Several algorithms do not perform a complete face detection prior to the redeye detection. Although the prior face detection could provide extremely useful information it is avoided as it is a challenging task itself [19, 61]. Instead, the redeye detection is performed by a framework for pattern classification where image subparts which serve as candidates for redeye artifacts are found by relatively easy methods. Then, a classifier trained by a set of typical redeye subimages throws away many false positives. The references [37, 64] follow

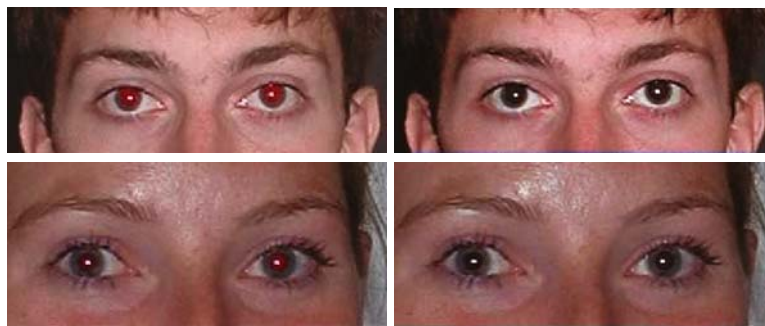
this line. For instance, red oval regions are found by simple image processing operations and regarded as redevye candidates [37]. In contrast, [36] trains a classifier to decide whether a region of fixed size  $11 \times 11$  is a redevye subimage. This classifier is applied to image patches of the same size at all possible positions. To handle different sizes of redevyes, the same operation is also carried out on scaled versions of the image.

The approach in [51] starts with a skin detection, followed by a morphological postprocessing of the skin-like regions. Then, the color is converted to a greylevel image, in which redevye artifacts in skin-like regions are highlighted as bright spots. Finally, convolutions are done to find pupils of circular shape.

After the detection of redevye parts, the correction can be performed in many different ways. The simplest solution of correction is to replace the red value of a pixel in a redevye part by its other color channel values such as the average of green and blue [51]. The reference [54] discusses a means to correct redevyes in a perceptually pleasing manner. Figure 7 shows two examples of redevye removal.

An unconventional hard/software solution is suggested in [38]. For the same scene two images are acquired, one with flash and one without flash. This conceptual idea could be implemented on today's digital cameras since many of them support "continuous shooting". If two images of the same scene are shot within a very short time interval, it can be assumed that no motion occurred and the images are well aligned. A flash and a non-flash image should have similar chromatic properties except in the redevye artifacts. This fact gives the fundamental of redevye detection in [38]. Furthermore, the chromatic information in the non-flash image is also used to restore the original iris color for redevye correction.

For the redevye detection part, a quantitative performance evaluation is straightforward. Some of the publications [16, 36, 37, 55, 64] mention this kind of performance measures, typically on hundreds of test images, while others do not. All the experiments were done on the authors' own images. Today, there exist still no common public test image databases with a well



**Fig. 7.** Examples of redevye removal. (Courtesy of S. Jirka and S. Rademacher)

accepted performance evaluation protocol like in other fields (face recognition [44], range image segmentation [20], etc.). In contrast, a (quantitative) evaluation of the redeye correction is much more subtle. Despite of the difficulties, performance evaluation on common data would certainly advance the research on redeye detection and correction.

## 6 Restoration of facial images

If some parts of a facial image, for instance the eyes and mouth, are damaged or missing, then we are faced with the task of facial image reconstruction. This problem is different from what image inpainting techniques [50] intend to solve, namely removing large objects and replacing them with visually plausible backgrounds. In contrast, face reconstruction recovers in some sense the “foreground” object (face) and as such some face modeling work is certainly needed.

Hwang and Lee [21] present an approach to face reconstruction. Face modeling is done by PCA on a set of prototype facial images so that the face space is spanned by the computed eigenfaces. A damaged facial image is considered as a point in a subspace spanned by reduced eigenfaces which only contain the pixels not damaged in the input image. Its optimal representation in terms of a linear combination of the reduced eigenfaces can be computed by a least square method. Then, using the same representation coefficients to linearly combine the original eigenfaces delivers a reconstructed facial image. In [21] this fundamental idea is separately applied to reconstruct the geometry and texture of a face. Afterwards they are fused to synthesize a facial image without damages. The same principle has also been used to enhance the resolution of facial images [41].

## 7 Artistic processing of facial images

A caricature is a humorous illustration that exaggerates the basic essence of a person to create an easily identifiable visual likeness. Artists have the amazing ability to capture distinguished facial features that make a subject different from others and to exaggerate these features. The central issue here is the question which facial features are significant from others and how to exaggerate.

There is only very few work on caricature generation [10, 27, 31]. In [10] the authors define a set of 119 nodes based on the MPEG-4 face definition parameters and face animation parameters to control the geometric shape of a face. They use 100 pictures of Asian female with manually labeled nodes to build an average shape. Given a new facial image, the corresponding mesh representation is computed by a chain of processing steps including locating

facial features like mouth and iris, and an iterative mesh optimization procedure. This mesh is then compared with the average face representation. Those nodes far away from the average are selected for exaggeration. The caricature generation itself is an image warping.

Related to caricature generation is the task of generating a line drawing from a facial image [9]. A further artistic processing of facial images is automatic generation of sketches [33, 52], which however has an important real application in law enforcement. An essential task there is automatic retrieval of suspects from a face database using a mug-shot only. In [52] facial images of the database are all transformed into a sketch, which are compared with the mug-shot. Based on a set of example photo-sketch pairs the photo-to-sketch transformation is performed by a technique similar to those for glasses removal using pairs of glasses and glasses-free example images.

## 8 Facial weight change

The task of weight change is to make a person's face to look fatter or thinner while maintaining the natural appearance of the face. Potential applications include beauty industry, security and police work. In [12] the authors report another use of weight change as a diagnostic tool in medicine. To diagnose, evaluate and treat eating disorders, especially Anorexia Nervosa, weight-changed facial images are needed for testing purpose.

The weight change simulator built in [12] is a straightforward image warping. The user selects thirteen landmarks to specify the shape of the neck and the cheeks. Their positions after the weight change are defined by a transformation controlled by a factor  $w$ . Then, the thin plate spline warping [5] is applied to transform other image points.

## 9 Conclusion

In this chapter we have given a brief overview of facial image processing tasks. The tasks considered include glasses removal, facial expression synthesis, eye synthesis, redeste removal, restoration of facial images, artistic processing of facial images, and facial weight change.

In addition to studying the algorithms for particular problems, it should be emphasized that several techniques we have seen in the last sections are of general nature and may find applications in other situations. In particular, the image mapping approach based on a set of pairs of images is despite of its simplicity powerful enough for challenging tasks like glasses removal [42, 46] and sketch generation [52].

A few topics are not discussed in this chapter. For instance, simulation of aging effects [29, 35] (making somebody look younger) is both a fascinating cosmetic image operation and a helpful tool in forensic medicines and law



**Fig. 8.** Redeye effect with dogs. (Courtesy of S. Jirka and S. Rademacher)

enforcement. In the latter case we may need to recognize a person, for which we only have older pictures in the database showing the person at a younger age. Handling varying imaging conditions like lighting is essential for robust image analysis systems. Face relighting [57] intends to re-render a facial image to arbitrary lighting conditions. Finally, the characteristic of facial images requires particular consideration in watermarking [39].

The facial image processing tasks considered in this chapter are mostly restricted to frontal images; the exception is redeye removal. Manipulating images of faces from arbitrary viewpoints clearly causes increased complexity. Some work based on 3D face modeling can be found in [3, 4].

Some of the discussed facial image processing operations could be extended to deal with animals. For instance, photographs of animals have similar effects as redeye, see Figure 8. Although in this case the artifacts may be green or blue, an automatic detection and correction still makes sense.

Tasks of facial image processing are mostly very challenging. As discussed in the introduction section, our brain with his intelligent visual processing ability in general and the very sensitive perception of faces increases the complexity. Nevertheless some powerful techniques have been developed so far and further development can be expected to produce powerful tools for facial image editing, communication, and recognition in a variety of application scenarios.

## Acknowledgments

The work was done when the first author was visiting Da-Yeh University, Taiwan. The financial support from both the Ministry of Education of Taiwan and Da-Yeh University is greatly appreciated. The authors want to thank S. Wachenfeld for proofreading the manuscript.



## References

1. B. Abboud, F. Davoine, and M. Dang. Facial expression recognition and synthesis based on an appearance model. *Signal Processing: Image Communication*, 19(8):723–740, 2004
2. B. Abboud and F. Davoine. Bilinear factorisation for facial expression analysis and synthesis. *IEE Proceedings of the Vision Image and Signal Processing*, 152(3):327–333, 2005
3. V. Blanz, C. Basso, T. Poggio, and T. Vetter. Reanimating faces in images and video. *Computer Graphics Forum*, 22(3):641–650, 2003
4. V. Blanz, K. Scherbaum, T. Vetter, and H.-P. Seidel. Exchanging faces in images. *Computer Graphics Forum*, 23(3):669–676, 2004
5. F. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989
6. C. Bregler, M. Covell, and M. Slaney. Video rewrite: Driving visual speech with audio. In *Proceedings of SIGGRAPH*, pages 353–360, 1997
7. N.P. Chandrasiri, T. Naemura, and H. Harashima. Interactive analysis and synthesis of facial expressions based on personal facial expression space. In *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition*, pages 105–110, 2004
8. K. Chang, K. Bowyer, and P. Flynn. A survey of approaches and challenges in 3D and multi-modal 2D+3D face recognition. *Computer Vision and Image Understanding*, 101(1):1–15, 2006
9. H. Chen, Y.-Q. Xu, H.-Y. Shum, S.-C. Zhu, and N.-N. Zheng. Example-based facial sketch generation with non-parametric sampling. In *Proceedings of International Conference on Computer Vision*, volume 2, pages 433–438, 2001
10. P.-Y. Chiang, W.-H. Liao, and T.-Y. Li. Automatic caricature generation by analyzing facial features. In *Proceedings of Asian Conference on Computer Vision*, 2004
11. D.K. Coats, E.A. Paysse, A.J. Towler, and R.L. Dipboye. Impact of large angle horizontal strabismus on ability to obtain employment. *Ophthalmology*, 107(2):402–405, 2000
12. U. Danino, N. Kiryati, and M. Furst. Algorithm for facial weight-change. *Proceedings of International Conference on Electronics, Circuits and Systems*, pages 318–321, 2004
13. Y. Du and Y. Lin. Emotional facial expression model building. *Pattern Recognition Letters*, 24(16):2923–2934, 2003
14. P. Eisert. Analysis and synthesis of facial expressions. In N. Sarris and G.M. Strintzis, editors, *3D Modeling and Animation: Synthesis and Analysis Techniques for the Human Body*, pages 235–265, Idea Group Inc., 2004
15. B. Fasel and J. Luetttin. Automatic facial expression analysis: A survey. *Pattern Recognition*, 36(1):259–275, 2003
16. F. Gasparini, R. Schettini. Automatic redeye removal for smart enhancement of photos of unknown origin. In *Proceedings of the International Conference on Visual Information Systems*, 2005
17. M. Gaubatz and R. Ulichney. Automatic red-eye detection and correction. In *Proceedings of International Conference on Image Processing*, volume I, pages 804–807, 2002

18. J. Ghent and J. McDonald. Photo-realistic facial expression synthesis. *Image and Vision Computing*, 23(12):1041–1050, 2005
19. E. Hjelmas and B.K. Low. Face detection: A survey. *Computer Vision and Image Understanding*, 83:237–274, 2001
20. A. Hoover, G. Jean-Baptiste, X. Jiang, P.J. Flynn, H. Bunke, D. Goldgof, K. Bowyer, D. Eggert, A. Fitzgibbon, and R. Fisher. An experimental comparison of range image segmentation algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(7):673–689, 1996
21. B.-W. Hwang and S.-W. Lee. Reconstruction of partially damaged face images based on a morphable face model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(3):365–372, 2003
22. X. Jiang, M. Binkert, B. Achermann, and H. Bunke. Towards detection of glasses in facial images. *Pattern Analysis and Applications*, 3(1):9–18, 2000
23. X. Jiang, S. Rothaus, K. Rothaus, and D. Mojon. Synthesizing face images by iris replacement: Strabismus simulation. In *Proceedings of first International Conference on Computer Vision Theory and Applications*, pages 41–47, 2006
24. Z. Jing and R. Mariani. Glasses detection and extraction by deformable contour. In *Proceedings of International Conference on Pattern Recognition*, 933–936, 2000
25. B. Kamgar-Parsi and A.K. Jain. Synthetic eyes. In J. Kittler and M.S. Nixon, editors, *Audio- and Video-Based Biometric Person Authentication*, pages 412–420. Springer, Berlin Heidelberg New York, 2003
26. S. Kong, J. Heo, B. Abidi, J. Paik, and M. Abidi. Recent advances in visual and infrared face recognition – A review. *Computer Vision and Image Understanding*, 97(1), 103–135, 2005
27. H. Koshimizu, M. Tominaga, T. Fujiwara, and K. Murakami. On KANSEI facial image processing for computerized facialcaricaturing system PICASSO. In *Proceedings of Conference on Systems, Man, and Cybernetics*, volume 6, pages 294–299, 1999
28. S. Krinidis, I. Buciu, and I. Pitas. Facial expression analysis and synthesis: A survey. In *Proceedings of International Conference on Human–Machine Interaction*, pages 1432–1436, 2003
29. A. Lanitis, C.J. Taylor, and T.F. Cootes. Toward automatic simulation of aging effects on face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):442–455, 2002
30. S.Z. Li and A.K. Jain, editors. *Handbook of Face Recognition*. Springer, Berlin Heidelberg New York, 2005
31. L. Liang, H. Chen, Y.Q. Xu, and H.Y. Shum. Example-based caricature generation with exaggeration. In *Proceedings of 10th Pacific Conference on Computer Graphics and Applications*, 2002
32. P. Litwinowicz and L. Williams. Animating images with drawings. In *Proceedings of SIGGRAPH*, pages 409–412, 1994
33. Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma. A nonlinear approach for face sketch synthesis and recognition. In *Proceedings of Conference on Computer Vision and Pattern Recognition*, volume 1, pages 1005–1010, 2005
34. Z. Liu, Y. Shan, and Z. Zhang. Expressive expression mapping with ratio images. In *Proceedings of SIGGRAPH*, pages 271–276, 2001
35. Z. Liu, Z. Zhang, and Y. Shan. Image-based surface detail transfer. *IEEE Computer Graphics and Applications*, 24(3):30–35, 2004

36. S. Loffe. Red eye detection with machine learning. In *Proceedings of International Conference on Image Processing*, volume II, pages 871–874, 2003
37. H. Luo, J. Yen, and D. Tretter. An efficient automatic redevye detection and correction algorithm. In *Proceedings of International Conference on Pattern Recognition*, volume 2, pages 883–886, 2004
38. X.P. Miao and T. Sim. Automatic red-eye detection and removal. In *Proceedings of International Conference on Multimedia and Expo*, pages 1195–1198, 2004
39. A. Nikolaidis and I. Pitas. Robust watermarking of facial images based on salient geometric pattern matching. *IEEE Transactions on Multimedia*, 2(3):172–184, 2000
40. K.-I. Okada, F. Maeda, Y. Ichikawaa, and Y. Matsushita. Multiparty video-conferencing at virtual social distance: MAJIC design. In *Proceedings of ACM Conference on Computer Supported Cooperative Work*, pages 385–393, 1994
41. J.-S. Park and S.-W. Lee. Reconstruction of high-resolution facial images for visual surveillance. In C.H. Chen and P.S.P. Wang, editors, *Handbook of Pattern Recognition and Computer Vision*, 3rd edn., pages 445–459. World Scientific, Singapore, 2005
42. J.-S. Park, Y.-H. Oh, S.-C. Ahn, and S.-W. Lee. Glasses removal from face image using recursive error compensation. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 27(5):805–811, 2005
43. F.I. Parke and K. Waters. *Computer Facial Animation*. A.K. Peters Ltd., 1996
44. P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss. The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000
45. F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, and D. Salesin. Synthesizing realistic facial expressions from photographs. In *Proceedings of SIGGRAPH*, pages 75–84, 1998
46. Y. Saito, Y. Kenmochi, and K. Kotani. Estimation of eyeglassless facial images using principal component analysis. In *Proceedings of International Conference on Image Processing*, volume 4, pages 197–201, 1999
47. J.S. Schildkraut and R.T. Gray. A fully automatic redevye detection and correction algorithm. In *Proceedings of International Conference on Image Processing*, volume I, pages 801–803, 2002
48. A. Sellen. Remote conversations: The effects of mediating talk with technology. *Human Computer Interaction*, 10(4):401–444, 1995
49. S.M. Seitz and C.R. Dyer. View morphing. In *Proceedings of SIGGRAPH*, pages 21–30, 1996
50. T.K. Shih and R.-C. Chang. Digital inpainting – survey and multilayer image inpainting algorithms. In *Proceedings of International Conference on Information Technology and Applications*, volume 1, pages 15–24, 2005
51. B. Smolka, K. Czubin, J.Y. Hardeberg, K.N. Plataniotis, M. Szczepanski, and K. Wojciechowski. Towards automatic redevye effect removal. *Pattern Recognition Letters*, 24:1767–1785, 2003
52. X. Tang and X. Wang. Face sketch recognition. *IEEE Transactions Circuits and Systems for Video Technology*, 14(1):50–57, 2004
53. P. Thompson. Margaret Thatcher: A new illusion. *Perception*, 9:483–484, 1980
54. R. Ulichney and M. Gaubatz. Perceptual-based correction of photo red-eye. In *Proceedings of IASTED International Conference on Signal and Image Processing*, 2005

55. J. Wan and X.P. Ren. Automatic red-eyes detection based on AAM. In *Proceedings of International Conference on Systems, Man and Cybernetics*, volume 7, pages 6337–6341, 2004
56. D. Weiner and N. Kiryati. Virtual gaze redirection in face images. In *Proceedings of International Conference on Image Analysis and Processing*, pages 76–81, 2003
57. Z. Wen, Z. Liu, and T.S. Huang. Face relighting with radiance environment maps. In *Proceedings of Conference on Computer Vision and Pattern Recognition*, volume 2, pages 158–165, 2003
58. C. Wu, C. Liu, and J. Zhou. Eyeglasses verification by support vector machine. In *Proceedings of IEEE Pacific Rim Conference on Multimedia*, pages 1126–1131, 2001
59. H. Wu, G. Yoshikawa, T. Shioyama, T. Lao, and T. Kawade. Glasses frame detection with 3D Hough transform. In *Proceedings of International Conference on Pattern Recognition*, volume 2, pages 346–349, 2002
60. Y. Xiao and H. Yan. Extraction of glasses in human face images. In *Proceedings of International Conference on Biometric Authentication*, pages 214–220, 2004
61. M.-H. Yang, D. Kriegman, and N. Ahuja. Detecting faces in images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1):34–58, 2002
62. L. Yin and A. Basu. Generating realistic facial expressions with wrinkles for model-based coding. *Computer Vision and Image Understanding*, 84(2):201–240, 2001
63. Q. Zhang, Z. Liu, B. Guo, D. Terzopoulos, and H.-Y. Shum. Geometry-driven photorealistic facial expression synthesis. *IEEE Transactions Visualization and Computer Graphics*, 12(1):48–60, 2006
64. L. Zhang, Y. Sun, M. Li, and H. Zhang. Automated red-eye detection and correction in digital photographs. In *Proceedings of International Conference on Image Processing*, volume IV, pages 2363–2366, 2004
65. W. Zhao, R. Chellappa, J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458, 2003
66. C. Zhou and X. Lin. Facial expressional image synthesis controlled by emotional parameters. *Pattern Recognition Letters*, 26(16):2611–2627, 2005

<http://www.springer.com/978-3-540-76830-2>

Applied Pattern Recognition

Bunke, H.; Kandel, A.; Last, M. (Eds.)

2008, XII, 246 p. 110 illus., 51 illus. in color., Hardcover

ISBN: 978-3-540-76830-2