Analysis of Facial Features Using Gaussian Steerable Filters

Arnd Gebhard[1], Weichuan Yu[2], Gerald Sommer[2], and Heinrich Niemann[1]

[1] Universität Erlangen–Nürnberg,
Lehrstuhl für Mustererkennung,
Martensstraße 3,
91058 Erlangen, Germany
{gebhard,niemann}@informatik.uni-erlangen.de,
http://www5.informatik.uni-erlangen.de
Tel.: +49 (0) 9131–8527799
Fax: +49 (0) 9131–303811

[2] Christian–Albrechts–Universität
Lehrstuhl für Kognitive Systeme
Preußerstraße 1–9,
24105 Kiel, Germany
{wy,gs}@ks.informatik.uni-kiel.de,
http://www.ks.informatik.uni-kiel.de
Tel.: +49 (0) 431–560496
Fax: +49 (0) 431–560481

Abstract. A novel approach for the analysis of facial features is presented. Instead of analyzing facial features with global models (such as the popular deformable templates) we perform a local analysis of the facial features. This local analysis is integrated in a system for diagnosis support of patient afflicted with single–sided facial paresis. This kind of paresis produces asymmetries in the patient's face, specially in the eye and mouth regions. Moreover, the asymmetries are amplified when special mimic exercises are performed.
In this system we apply Gaussian averaging masks, which are used as a steerable filter, to extract information about facial features with high precision to detect asymmetries inside the face. The system is composed of three successive functions: facial feature localization, local orientation analysis, and facial asymmetry evaluation. In an initial step the eye and mouth regions are localized. Then Gaussian averaging masks are applied to characterize the orientation information in the neighborhood of the corners of the eyes and mouth. This orientation information is analyzed and used for the diagnosis supporting task. Real experimental results show that this technique is promising.

Key words: medical image understanding, facial feature analysis, steerable filter, facial paresis, diagnosis supporting system

Analysis of Facial Features Using Gaussian Steerable Filters

Abstract. A novel approach for the analysis of facial features is presented. Instead of analyzing facial features with global models (such as the popular deformable templates) we perform a local analysis of the facial features. This local analysis is integrated in a system for diagnosis support of patient afflicted with single–sided facial paresis. This kind of paresis produces asymmetries in the patient's face, specially in the eye and mouth regions. Moreover, the asymmetries are amplified when special mimic exercises are performed.
In this system we apply Gaussian averaging masks, which are used as a steerable filter, to extract information about facial features with high precision to detect asymmetries inside the face. The system is composed of three successive functions: facial feature localization, local orientation analysis, and facial asymmetry evaluation. In an initial step the eye and mouth regions are localized. Then Gaussian averaging masks are applied to characterize the orientation information in the neighborhood of the corners of the eyes and mouth. This orientation information is analyzed and used for the diagnosis supporting task. Real experimental results show that this technique is promising.

Key words: medical image understanding, facial feature analysis, steerable filter, facial paresis, diagnosis supporting system

1 Introduction

The automatic analysis of human faces and facial features is a widespread field of current research. A very common way for the analysis was presented by Yuille et al. in [14]. There facial features were localized and analyzed by means of deformable templates. A human eye was approximated by a parametric model consisting e.g. of two parables for the lids and a circle for the iris. In other papers [1, 2] the whole face and facial features were described by parametric models.

The mentioned approaches have in common that the modelling of faces and facial features is performed in a global, but coarse way. A face modelled as an ellipse or eye lids and lips as parables are just approximations to the actual form of those parts. A close analysis of the segmentation results shows systematical variations between the model and the real appearance (e.g. at the corners of the eyes and the mouth). The idea of using more complex geometric models (e.g. polygons of higher orders) is not feasible because of the raised calculation times and/or the insufficient information inside the image. Additionally, it can be quite difficult to find an adequate global model to express real facts (cf. Fig. 5 in Sect. 4).

For this reason we are interested in an exact local analysis of facial features.

In this contribution we propose a novel approach for the analysis of facial features (here: the eyes and the mouth). A local orientation analysis at the corners of the facial features extracts structural information, e.g. the opening angle of the eyes and the mouth. The orientation analysis is performed by means of Gaussian averaging mask, a special kind of steerable filter. The orientation information is used to analyze the asymmetry of faces. Asymmetry is a symptom of facial paresis. This method is applied in a diagnosis system of patients afflicted with facial paresis.

The work is organized as follows: In Sect. 2 a module for the localization of human faces and facial features is presented. The main task is to localize the corners of the eyes and the mouth. The extraction of structural information is described in Sect. 3. Steerable filters are applied for the extraction of local orientation information at the corners of the eyes and mouth. An application of the described

modules in a diagnosis support system of patients with facial paresis is presented in Sect. 4. Experimental results are shown in Sect. 5. We conclude the paper in Sect. 6.

2 Localization of Facial Features

Many localization modules in the literature are based on the assumptions of homogenous background (e.g. [3]) or constant face color distribution(e.g. [9]). These assumptions cannot be made in our system. The patients do not sit in front of homogenous backgrounds but inside a clinical environment or at their homes, and there is no constant face color distribution because of changing light conditions (intensity *and* color) and cameras.

We develop a localization module that handles the stated conditions. The localization is divided into several steps, which we will describe in the following. We assume that a patient is sitting toward the camera in front of a known, heterogenous background. The camera is configured so that portraits (images with whole head and shoulders) of the patient can be generated.

2.1 Head Localization

The first step is to detect and localize the patient's head inside the portrait images (Fig. 1). At first we take a color image $\boldsymbol{b}$ (RGB) of the background (Fig. 1a). After the patient sits down toward the camera we take five (cf. Sect. 4) different portrait images ${}_i\boldsymbol{f}$ with $i = 1, ..., 5$ (Fig. 1b). Since color information is sensitive to illumination changes we conduct a color normalization to obtain a stable background image $\hat{\boldsymbol{b}}$. For this purpose we define an upper left A_l and upper right A_r area inside the images that contain no parts of the patient's head (Fig. 1a and 1b). This has to be taken into account during the generation of the portraits ${}_i\boldsymbol{f}$. In equations (1) to (3) we show the calculation steps of the R(ed) channel. The same steps are evaluated in the G and B channels.

The average values of the color channels (here the Red channel) of A_l and A_r of ${}_1\boldsymbol{f}^R$, which is called m^R, and $\boldsymbol{b}^R$, which is called n^R,

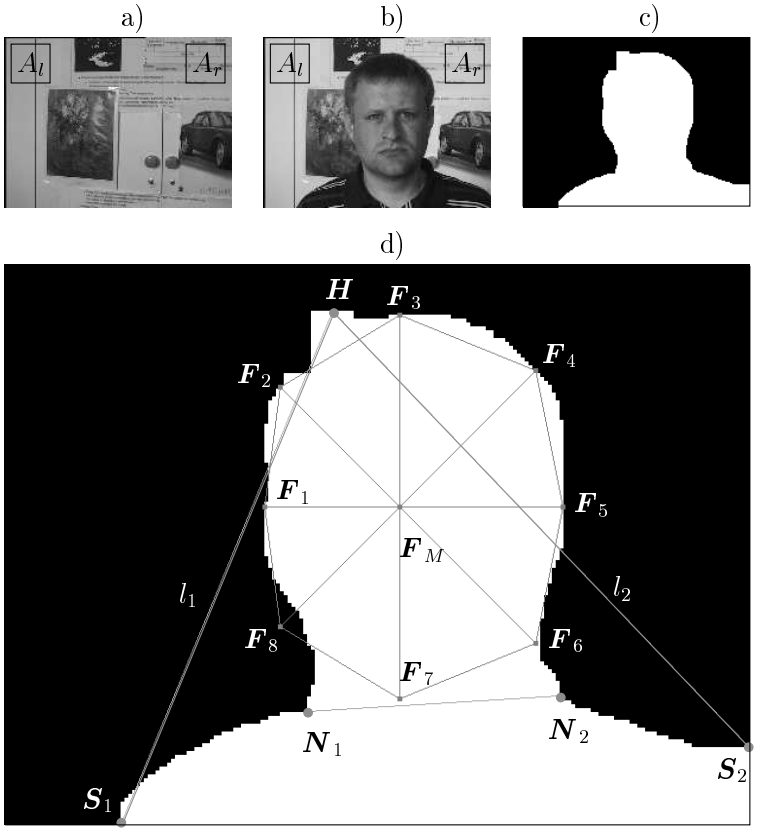

Fig. 1. Localization of faces. a) background image. b) Person in the foreground. Regions A_l and A_r remain background. c) Difference image. Foreground and background are separated. d) The result of a face localization. The face is estimated to be inside the contour of F_1 to F_8.

are calculated:

$$m^R = \frac{1}{|A_l| + |A_r|} \left(\sum_{(x,y)\in A_l} {}_1\boldsymbol{f}^R(x,y) + \sum_{(x,y)\in A_r} {}_1\boldsymbol{f}^R(x,y) \right) \quad (1)$$

$$n^R = \frac{1}{|A_l| + |A_r|} \left(\sum_{(x,y)\in A_l} \boldsymbol{b}^R(x,y) + \sum_{(x,y)\in A_r} \boldsymbol{b}^R(x,y) \right), \quad (2)$$

where $|A_l|$ and $|A_r|$ are the number of pixels inside A_l and A_r. The normalized red channel of image $\hat{\boldsymbol{b}}$ is given by

$$\hat{\boldsymbol{b}}^R(x,y) = \frac{m^R}{n^R} \boldsymbol{b}^R(x,y). \quad (3)$$

After obtaining $\hat{\boldsymbol{b}}$, which is composed of $\hat{\boldsymbol{b}}^R$, $\hat{\boldsymbol{b}}^G$, and $\hat{\boldsymbol{b}}^B$, we segment the foreground (patient with head and shoulders) from the background by calculating binary difference images ${}_id$ (Fig.1c):

$$ {}_id(x,y) = \begin{cases} 1 & : \ \|{}_i\boldsymbol{f}(x,y) - \hat{\boldsymbol{b}}(x,y)\| \geq \theta_{\mathrm{bf}} \\ 0 & : \ \text{else}, \end{cases} \tag{4} $$

with threshold θ_{bf} set to 20, which was determined experimentally. As parts of the analysis methods are implemented to handle gray value input images, the color images ${}_i\boldsymbol{f}$ are transformed to gray value images ${}_ig$ with the following equation:

$$ {}_ig(x,y) = \frac{{}_i\boldsymbol{f}^R(x,y) + {}_i\boldsymbol{f}^G(x,y) + {}_i\boldsymbol{f}^B(x,y)}{3}. \tag{5} $$

The next step is to find the face region inside the silhouette of the patient (Fig. 1d). The uppermost point of the head $\boldsymbol{H} = ({}_x\boldsymbol{H}, {}_y\boldsymbol{H})^T$ is found by searching the rows of ${}_id$ from top to bottom. $\boldsymbol{H}$ is the midpoint of the first coherent part in the first row of the foreground. The intersection points of the shoulders and the image edge $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are found by tracking the margins of the ${}_id$. Let l_1 be the connection line of $\boldsymbol{H}$ and $\boldsymbol{S}_1$. In every image row below the row containing $\boldsymbol{H}$ the difference between the x–coordinates of l_1 and the left part of the patient's silhouette is calculated. In the row with the biggest difference the silhouette point is considered to be a neck point $\boldsymbol{N}_1$. $\boldsymbol{N}_2$ is found similarly with line l_2 and the right part of the silhouette. With $\boldsymbol{H}$, $\boldsymbol{N}_1$, and $\boldsymbol{N}_2$ the mid point of the face $\boldsymbol{F}_M$ can be calculated:

$$ {}_x\boldsymbol{F}_M = \frac{{}_x\boldsymbol{H} + {}_x\boldsymbol{N}_1 + {}_x\boldsymbol{N}_2}{3} \tag{6} $$

$$ {}_y\boldsymbol{F}_M = \frac{2{}_y\boldsymbol{H} + {}_y\boldsymbol{N}_1 + {}_y\boldsymbol{N}_2}{4} \tag{7} $$

Starting from $\boldsymbol{F}_M$ five rays search the upper head silhouette in W(est), NW, N, NE and E direction. Every last point inside the foreground part ($d(x,y) = 1$) is assumed to be a sampling point of the head outline. That gives the upper points $\boldsymbol{F}_1$ to $\boldsymbol{F}_5$ of the head silhouette. $\boldsymbol{F}_2$ to $\boldsymbol{F}_4$ are mirrored on the connection line of $\boldsymbol{F}_1$ and $\boldsymbol{F}_5$. The resulting points are $\boldsymbol{F}_8$, $\boldsymbol{F}_7$, and $\boldsymbol{F}_6$, which are estimates of the lower points of the head contour.

2.2 Localization of Facial Features

In order to localize the eyes and mouth/nose region inside the estimated face region we use a parametric model shown in Fig. 2b. The

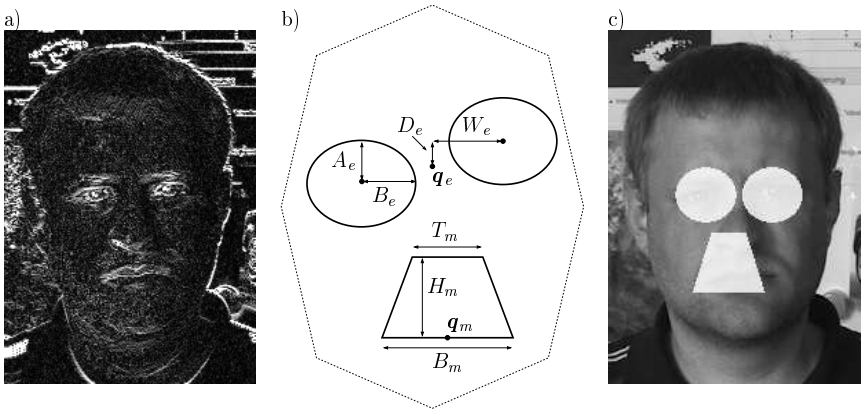

Fig. 2. Localization of facial features. a) Vertical edge strength representation of a face image. b) Parametric model for facial features. c) Localization results.

facial features are localized by optimizing the model parameters.

As the facial features are mainly characterized by horizontal components we find them in the vertical gradient representation ${}_ig_v$ of the image ${}_ig$ (cf. Fig. 2a. The eyes, nose and mouth appear lighter as they contain a higher amount of vertical gradient). In the case of the eyes two elliptical regions with maximal ratios of vertical gradient energy to area are found. Among the model parameters $\boldsymbol{q}_e$ denotes the middle point between two eyes. The axes of the ellipses are denoted with A_e and B_e. The horizontal and vertical distances between two eyes are represented with $2W_e$ and $2D_e$, respectively.

The vertical edge energy in the eye region can be determined with the equation:

$$E_e^v = \int_0^{2\pi}\int_0^1 {}_ig_v({}_x\boldsymbol{q}_e + B_e r\cos\phi + W_e\,,\, {}_y\boldsymbol{q}_e + A_e r\sin\phi + D_e) + \\ + {}_ig_v({}_x\boldsymbol{q}_e + B_e r\cos\phi - W_e\,,\, {}_y\boldsymbol{q}_e + A_e r\sin\phi - D_e)\,dr\,d\phi. \quad (8)$$

The parameter values are determined by optimizing

$$(\boldsymbol{q}_e^*, A_e^*, B_e^*, W_e^*, D_e^*) = \underset{(\boldsymbol{q}_e, A_e, B_e, W_e, D_e)}{\operatorname{argmax}} \frac{E_e^v}{A_e\, B_e} \tag{9}$$

using a simplex method combined with an adaptive random search initialization (cf. [4]).

The nose/mouth region is modelled as an isosceles trapezoid. The position $\boldsymbol{q}_m$ is the origin of the mouth model. H_m is the height, T_m and B_m are the widths of the top and of the bottom. The localization is done by another optimization process where the five parameters are determined, similar to the eyes, by the search of a region with maximal ratio of vertical edge energy to area. The result of a facial feature localization is shown in Fig. 2c.

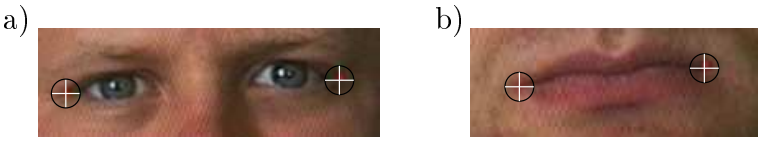

Fig. 3. Localization results: a) Corners of the eyes and b) corners of the mouth

2.3 Localization of Corners of Facial Features

The last step is to localize the corners of the eyes and the mouth (cf. Fig. 3). We observe that those corners produce dark areas in the image compared with their surrounding. Based on this observation we use a search method to localize the corners of facial features: Starting at the outer columns we calculate the column's average gray value and compare it with the minimal gray value of the column. If the ratio of both values is lower than a given threshold, we assume that the position of the minimal gray value is the position of the angle of the facial feature. As the localization of the facial features can be incorrect (e.g. the right side of the mouth in Fig. 1), the search areas are horizontally enlarged by 50%. The determined positions are used in the following as the *keypoint* for the local orientation analysis.

3 Gaussian Steerability for Orientation Analysis

In order to analyze facial features and to detect asymmetries of faces containing pareses exactly we need local orientation analysis with high precision. We used to apply a set of filters at every orientation to obtain the orientation information. This, however, led to enormous computational load. In order to reduce such effort, the concept of steerability was introduced [5, 10]. Denoting with θ ($\theta \in I\!R$) the deformation parameter we define a filter $F(\boldsymbol{x})$ with $\boldsymbol{x} \in I\!R^n$ as a steerable filter if its deformed versions $F_\theta(\boldsymbol{x})$ can be expressed as [8]:

$$F_\theta(\boldsymbol{x}) = \sum_{k=1}^{N} b_k(\theta) A_k(\boldsymbol{x}) \tag{10}$$

where $A_k(\boldsymbol{x})$ and $b_k(\theta)$ are referred to as basis filters and interpolation functions, respectively.

The signature $S(\theta)$ of an image point $g(x, y)$ can be obtained by applying a 2–dimensional steerable filter to the image g:

$$S(\theta) \stackrel{def}{=} \langle F_\theta(x, y) | g(x, y) \rangle = \sum_{k=1}^{N} b_k(\theta) \langle A_k(x, y) | g(x, y) \rangle. \tag{11}$$

Here $\langle \cdot | \cdot \rangle$ denotes the usual inner product for two real functions. We see the motivation of steerability clearly in equation (11): The responses of a given family of the filter $F_\theta(x, y)$ with $\theta \in I\!R$ are reduced to a linear combination of N basis filter responses. Since the number of basis filters is much less than that of the deformed versions of the original filter, the computational load is strongly reduced.

Steerable filters have been used in local orientation analysis [7, 11, 13]. Among them the Gaussian steerability [13] has superior performance in achieving high orientation selectivity with low computational load. Thus, we apply the Gaussian steerability in our approach. In the Gaussian steerability we apply the Gaussian function as interpolation function:

$$b_k(\theta) = G_0(\mathcal{D}(\theta, \theta_k)) \tag{12}$$

where θ denotes the angular variable and $\theta_k (k = 1, \cdots, N)$ are evenly distributed angular values. Since θ and θ_k are circular angles ($\theta, \theta_k \in$

$[0, 2\pi]$), we define a $\mathcal{D}(\cdot)$ to represent the minimal circular difference between θ and θ_k

$$\mathcal{D}(\theta, \theta_k) = \min(|\theta - \theta_k|, |\theta - \theta_k + 2\pi|, |\theta - \theta_k - 2\pi|), \tag{13}$$

and

$$G_0(\mathcal{D}(\theta, \theta_k)) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\mathcal{D}(\theta, \theta_k))^2}{2\sigma^2}}. \tag{14}$$

To treat the variables naturally we first conduct a local polar transformation: $g(x, y) \rightarrow g(r, \theta)$ (see Fig. 4). We set $R_{\min} > 0$ to avoid the confusion close to the keypoint [7]. In order to obtain angular samples A_k robustly after the non-uniform polar mapping, we enlarge sampling masks and calculate averaging values along the angular directions as sampling outputs. Intuitively, the pixels in the middle of the mask should be more weighted than the pixels at the boundary. Therefore, we use a mask of Gaussian shape along the angular direction as sampling mask.

Taking these issues into account the angular samples $A_k (k = 1, \cdots, N)$ read:

$$A_k = \sum_{r=R_{\min}}^{R_{\max}} \sum_{\theta=\theta_k-\frac{W}{2}}^{\theta_k+\frac{W}{2}} g(r, \theta) \frac{G_0(\mathcal{D}(\theta, \theta_k))}{\mathcal{N}(R_{\min}, R_{\max}, \theta_k)} \tag{15}$$

where $\mathcal{N}(R_{\min}, R_{\max}, \theta_k)$ is an averaging factor along the radial direction which is the sum of discrete weights inside the sampling mask centered at θ_k. Since sampling masks are rotation-variant, it is therefore dependent on the parameters $R_{\min}$,$R_{\max}$, and θ_k. Theoretically, a Gaussian function is not compactly supported. In practice we only consider one part of $G_0(\mathcal{D}(\theta, \theta_k))$ whose variable varies from $\theta_k - \frac{W}{2}$ to $\theta_k + \frac{W}{2}$ (see equation (15)). Therefore we apply a filter with compact support. Here W denotes the angular width of the sampling mask. It is easy to show that in order to keep the energy of the cut-off area below 1% of the total energy the width of the mask must be at least 5σ. In this paper, we set $W = 6\sigma$. In Fig. 4 we show a sampling mask centered at θ_k.

4 Diagnosis Supporting System

Facial paresis occurs in the most cases (approximately 99%) single–sided. I.e., the physical defects appear only on one side of the pa-

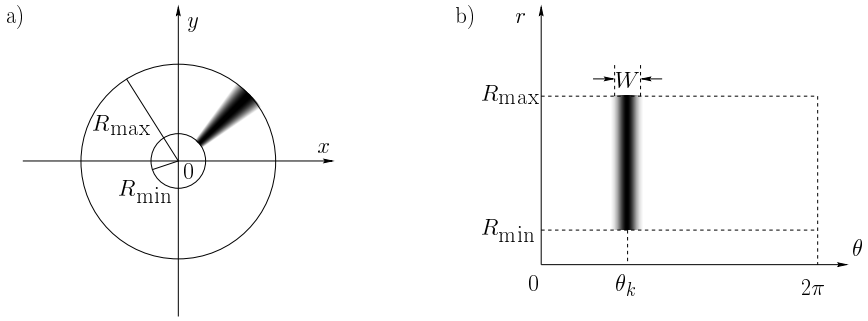

Fig. 4. A Gaussian averaging mask centered at angle θ_k. Darker pixels represent larger mask weights. a) The mask in the Cartesian coordinate system. The keypoint is at the center of the circle. b) The mask with r and θ as coordinates, $R_{\max}$ and $R_{\min}$ are radial boundaries of the mask, W is the angle width of the mask. We set $R_{\min} > 0$ to avoid the confusion close to the keypoint.

tient's face. These defects result in asymmetries in the faces. Fig. 5a and 5b show asymmetries in the eye region of a patient, Fig. 5c and 5d show those in the mouth region.

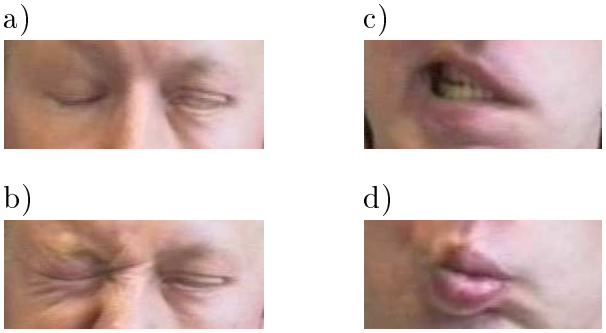

Fig. 5. Appearance of parts of a patient's face during the performance of different mimic exercises.

The mentioned asymmetries can appear in different peculiarities. If the face is in a relaxed state often just few or even no asymmetries can be observed. On the other hand, if certain physical actions are performed (see below) the asymmetries can be amplified and observed clearly. These facts, on which our system is based, are also used by approved medical indexing systems [6, 12].

There are four different mimic exercises that amplify the facial asymmetries in a canonical way:

1. Lift the eyebrows so that wrinkles occur on the forehead. We call this exercise "Frowning".
2. Close the eyes. Patients with heavy paresis in the eyes region are not able to close their eyes at all. As a result we observe only a roll up of the eye ball.
3. Show the teeth. The lips have to be opened with the lower jaw in highest position.
4. Point the mouth. The lips are closed and the tip of the mouth depicts a small ring.

The first two exercises intensify functional disturbances in the eyes' region, the last two exercises disturbances in the mouth region.

After we take a portrait image ${}_1\boldsymbol{f}$ of the face in a relaxed state, we generate four portrait images ${}_2\boldsymbol{f}, \cdots, {}_5\boldsymbol{f}$ when the patient is performing the four mentioned exercises. We assume that these five images contain adequate information to reach a decision concerning the presence of facial paresis. In the present system we use no further information such as an image sequence of the moving face.

The analysis process starts with the localization of the face, the facial features, and the corners of the facial features in the images ${}_1\boldsymbol{f}$ to ${}_5\boldsymbol{f}$ as shown in Sect. 2.

4.1 Aligned Angular Samples of Averaging Gaussian Masks

The analysis approach uses the angular samples (15) of the Gaussian averaging masks (Sect. 3). The localized corners (see Sect. 2) of the eyes and the mouth are used as keypoints. Different mask parameters are used for the analysis of the eyes and the mouth regions. The opening angle of the wedges is the same in both cases: $W = 6^o$. For the eyes region we use $R^e_{\min} = 3$ and $R^e_{\max} = 30$, and for the mouth region $R^m_{\min} = 5$ and $R^m_{\max} = 40$. All these values were determined experimentally.

We filter only the half of the neighborhood around the keypoints to extract orientation information (Fig. 6). In Fig. 6 we present also typical filter signatures of a healthy person. The number of Gaussian

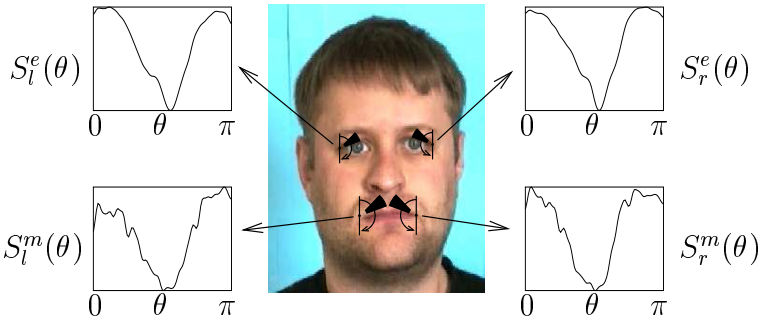

Fig. 6. Gaussian averaging masks used for orientation analysis of eyes and mouth corners.

masks in this range is $N = 91$. I.e.,the direction of the masks changes stepwise by 2^o.

In all five images ${}_ig(i = 1, \cdots, 5)$, the structural information at the corners of mouth and eyes is locally analyzed. Four vectors of angular samples are extracted:

$$\begin{aligned} {}_i\boldsymbol{A}^{e,l} &= ({}_iA_0^{e,l}, ..., {}_iA_{90}^{e,l})^T, \quad & {}_i\boldsymbol{A}^{e,r} &= ({}_iA_0^{e,r}, ..., {}_iA_{90}^{e,r})^T \\ {}_i\boldsymbol{A}^{m,l} &= ({}_iA_0^{m,l}, ..., {}_iA_{90}^{m,l})^T, \quad & {}_i\boldsymbol{A}^{m,r} &= ({}_iA_0^{m,r}, ..., {}_iA_{90}^{m,r})^T \end{aligned} \tag{16}$$

Different illuminations of the two halves of the face or a small rotation of the patient face will result in a translation or a scaling of the filter signature and the angular samples. These effects, which also will produce asymmetries in the face, can be restored by an optimal alignment (scaling and translation) of the angular samples of the two facial sides. The results are the features ${}_ic_1$ for the eyes and ${}_ic_2$ for the mouth regions:

$${}_ic_1 = \min_{s,t} \sum_{j=0}^{90} |{}_iA_j^{e,l} - s\,{}_iA_{j+t}^{e,r}| \tag{17}$$

$${}_ic_2 = \min_{s,t} \sum_{j=0}^{90} |{}_iA_j^{m,l} - s\,{}_iA_{j+t}^{m,r}| \tag{18}$$

4.2 Average Gray Value Difference

The approach presented in Sect. 4.1 is compared with the asymmetry analysis using gray value differences. I.e. the asymmetry is measured

by the sum of absolute differences of pixels in the left and right half of the image. To get a feature of the eye region, we take the gray values of the left eye, mirror the single lines and match the gray values with the right eye by varying the x– and y–coordinates to find minimal absolute sums ${}_iD_1^*$ of the pixel differences for every image ${}_ig, i = 1, ..., 5$.

$$ {}_iD_1^* = \min_{x_t, y_t} \iint\limits_{\text{right eye}} |({}_ig(x,y) - {}_ig(2({}_iB_e - {}_iW_e) - x + x_t, y - 2\,{}_iD_e + y_t)|\,dx\,dy \tag{19} $$

The absolute sums divided by the area of the right eye (in our face model both eyes have the same size) are taken as features ${}_ic_3$.

$$ {}_ic_3 = \frac{{}_iD_1^*}{{}_iA_e\;{}_iB_e} \tag{20} $$

which are approximately the average gray value difference in the pixels of the left and the right eye region.

To analyze the mouth regions we find row indices ${}_ir^*$ inside the mouth regions that will give minimal sums of absolute differences when matching the pixels on the left of row ${}_ir^*$ with those on the right side. The rows ${}_ir^*$ represent the vertical symmetric axes of the mouth which are the elongations of the nasal labial folds.

$$ {}_ir^* = \operatorname*{argmin}_r \int\limits_{\text{mouth left}}^{r} \int\limits_0^{H_m} |({}_ig(x,y) - {}_ig(2\,r - x, y)|\,dx\,dy \tag{21} $$

$$ {}_iD_2^* = \int\limits_{\text{mouth left}}^{{}_ir^*} \int\limits_0^{H_m} |({}_ig(x,y) - {}_ig(2\,{}_ir^* - x, y)|\,dx\,dy \tag{22} $$

The values ${}_iD_2^*$ are divided by the areas of the analyzed regions to get features for the asymmetry of the mouth regions.

$$ {}_ic_4 = \frac{{}_iD_2^*}{{}_ir^*\,{}_iH_m}. \tag{23} $$

The use of the extracted features for the diagnostic task is shown in Sect. 4.3. Classification results of all analysis approaches are presented in Sect. 5.

4.3 Detection of Facial Paresis

In this section we use the extracted information (cf. Sect. 4.1 and Sect. 4.2) to detect facial paresis. As mentioned before five portrait images of the patient were generated. In image ${}_1\boldsymbol{f}$ the patient keeps his face relaxed. In the other images the patient is performing different mimic exercises that intensify the facial asymmetries. We want to analyze the asymmetries which are produced by a missing functionality of the face. So asymmetries which occur because of facts like different illumination or natural variations in the two halves of the face have to be ignored. This is done by normalizing the features ${}_2c_n, {}_3c_n, {}_4c_n, {}_5c_n$, by ${}_1c_n (n = 1, ..., 4)$. In the following we will focus on $n = 1$ and $n = 2$. For the other cases all observations are equivalent.

All extracted features ${}_ic_1$ and ${}_ic_2$ are normalized by the features ${}_1c_1$ and ${}_1c_2$ respectively, which were extracted from image ${}_1\mathbf{f}$. That gives four ratios for the eye regions and another four ratios for the mouth region:

$$\frac{{}_2c_1}{{}_1c_1}, \frac{{}_3c_1}{{}_1c_1}, \frac{{}_4c_1}{{}_1c_1}, \frac{{}_5c_1}{{}_1c_1}, \text{ and } \frac{{}_2c_2}{{}_1c_2}, \frac{{}_3c_2}{{}_1c_2}, \frac{{}_4c_2}{{}_1c_2}, \frac{{}_5c_2}{{}_1c_2}. \tag{24}$$

The ratios are thresholded with θ_e and θ_m, respectively. In the present system we use equal thresholds $\theta_e = \theta_m = 1.3$. If at least one ratio from the eye region is greater than the threshold θ_e a facial paresis in the eye region of the person is assumed. The diagnose for the mouth region is similar with the threshold θ_m. First results of the evaluation of the feature qualities are presented in Sect. 5.

5 Results

Table 1 presents the correct localization rates of the module presented in Sect. 2.

The analysis of faces is divided into patients and healthy persons. Table 2 presents the correct diagnosis response of an eye region and mouth regions. The format is x/y, where x is the number of correct analysis results and y the total number of analyses. The test data base contains images of 19 persons. Among these people ten have paresis in the eye region, and twelve have paresis in the mouth region. A comparison of the correct classification rates shows that the

Head/Face	95%
Eyes	85%
Mouth	80%
Eye Corners	65%
Mouth Corners	60%

Table 1. Localization results of the head, the facial features, and their corners.

Approach	Aligned Angular Samples		Gray Value Differences	
Region (Feature)	Eye (${}_ic_1$)	Mouth (${}_ic_2$)	Eye (${}_ic_3$)	Mouth (${}_ic_4$)
Patient	7/10	10/12	6/10	7/12
Healthy Person	8/9	7/7	7/9	7/7

Table 2. Analysis results.

approach based on aligned angular samples produces better results than the approach based on absolute gray value differences.

6 Conclusion

We presented a novel approach for the analysis of facial features. This approach was integrated into a system for diagnosis support of patients with facial paresis. The system was adapted to patients with single sided facial paresis. This allowed us to use the facial asymmetries of a patient as symptoms for the diagnose. These asymmetries were amplified in a canonical way by the execution of special mimic exercises.

After the patient face, the facial features, and the corners of these features are localized, the local structure information is analyzed. The local analysis is based on the application of steerable filters. Aligned angular samples of filter signatures are used to detect facial asymmetries. First experiments show that the aligned angular samples produce better results for the diagnosis task than the comparison of gray values in the eyes and mouth regions. The system is now installed at the Department of Otorhino–Larygology of the University Erlangen–Nuremburg.

References

1. R. Brunelli and T. Poggio. Face recognition: Features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10):1042–1052, 1993.
2. Ian Craw, David Tock, and Alan Bennett. Finding face features. In G. Sandini, editor, *Computer Vision — ECCV '92*, Lecture Notes in Computer Science, pages 92–96, Santa Margherita Ligure, 1992. Springer.
3. L. De Silva, K. Aizawa, and M. Hatori. Detection and tracking of facial features. In *SPIE Visual Communications and Image Processing'95 (VCIP'95)*, volume 2501, pages 2501/1161–2501/1172. The International Society for Optical Engineering, May 1995.
4. S. M. Ermakov and A. A. Zhiglyavskij. On random search of global extremum. *Probability Theory and Applications*, 28(1):129–136, 1983.
5. W.T. Freeman and E.H. Adelson. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:891–906, 1991.
6. J.W. House. Facial nerve grading systems. *Laryngoscope*, 93:1056–1069, 1983.
7. M. Michaelis and G. Sommer. Junction classification by multiple orientation detection. In *Proc. Third European Conference on Computer Vision*, volume I, pages 101–108, Stockholm, Sweden, May 2-6, J.O. Eklundh (Ed.), Springer LNCS 800, 1994.
8. M. Michaelis and G. Sommer. A Lie group approach to steerable filters. *Pattern Recognition Letters*, 16:1165–1174, 1995.
9. N. Oliver, A. Pentland, and F. Berard. LAFTER: Lips and face real time tracker. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 123–129, S. Juan, Puerto Rico, 1997.
10. P. Perona. Deformable kernels for early vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5):488–499, 1995.
11. E. P. Simoncelli and H. Farid. Steerable wedge filters for local orientation analysis. *IEEE Trans. Image Processing*, 5(9):1377–1382, 1996.
12. E. Stennert, C. H. Limberg, and K. P. Frentrup. Parese- und Defektheilungs-Index. *HNO*, 25:238–245, 1977.
13. W. Yu, K. Daniilidis, and G. Sommer. Approximate orientation steerability based on angular gaussians. Submitted to IEEE Trans. on Image Processing.
14. Alan L. Yuille, Peter W. Hallinan, and David C. Cohen. Feature extraction from faces using deformable templates. *International Journal of Computer Vision*, 8:99–111, 1992.