

Behavioral State Detection of Newborns Based on Facial Expression Analysis

Lykele Hazelhoff^{1,2}, Jungong Han¹,
Sidarto Bambang-Oetomo^{1,3}, and Peter H.N. de With^{1,2}

¹ University of Technology Eindhoven, The Netherlands

² CycloMedia Technology B.V., Waardenburg, The Netherlands

³ Maxima Medisch Centrum, Veldhoven, The Netherlands

Abstract. Prematurely born infants are observed at a Neonatal Intensive Care Unit (NICU) for medical treatment. Whereas vital body functions are continuously monitored, their incubator is covered by a blanket for medical reasons. This prevents visual observation of the newborns during most time of the day, while it is known that the facial expression can give valuable information about the presence of discomfort.

This prompted the authors to develop a prototype of an automated video survey system for the detection of discomfort in newborn babies by analysis of their facial expression. Since only a reliable and situation-independent system is useful, we focus at robustness against non-ideal viewpoints and lighting conditions. Our proposed algorithm automatically segments the face from the background and localizes the eye, eyebrow and mouth regions. Based upon measurements in these regions, a hierarchical classifier is employed to discriminate between the behavioral states sleep, awake and cry.

We have evaluated the described prototype system on recordings of three healthy newborns, and we show that our algorithm operates with approximately 95% accuracy. Small changes in viewpoint and lighting conditions are allowed, but when there is a major reduction in light, or when the viewpoint is far from frontal, the algorithm fails.

1 Introduction

Prematurely born infants are observed in a Neonatal Intensive Care Unit (NICU) for medical treatment. These babies are nursed in an incubator, where their vital body functions such as heart rate, respiration, blood pressure, oxygen saturation and temperature are continuously monitored. To avoid excessive light exposure which may disturb sleep and thereby lead to discomfort, the incubator is covered by a blanket to reduce the intensity of light. The disadvantage of this practice is that visual observation of the newborn during most of the time is impaired. In this respect, pain and discomfort of the patient that cannot be observed by the monitoring of vital functions may pass unnoticed. The awareness that early treatment of pain and discomfort is important for the future development, is presently growing and this motivated us to develop a prototype for an automated video observation system that can detect pain and discomfort in newborn babies.

Discomfort on newborns can be detected by analysis of the facial expression [1]-[3]. Especially the appearances of the eyes, eyebrows and mouth are reported to be important facial features for detecting the presence of discomfort and pain. This has resulted in the development of scoring systems to assess the level of discomfort, based on the facial expression and physiological parameters. The scoring systems allow early alerting of caretakers when a patient suffers from pain or discomfort, so that appropriate actions can be taken in time.

So far, only one automatic video-surveillance system [4]-[5] for sick newborn infants has been developed. In this system, images of a large number of newborns are photographed during different situations: during a painful procedure (heel lance) and during other, non-painful, situations. Then, after manual rotation and scaling, pixel-based classifiers, such as Linear Discriminant Analysis and Support Vector Machines, are applied for classification of the facial expression. Although the results are acceptable, we consider this approach not robust enough to handle fluctuating, non-ideal situations, including varying lighting conditions and situations where the baby face is partly covered by plasters or tubing.

Because of this lack of robustness, we explore the design of a discomfort detection system, and describe a pilot system with the following properties. First, the detection of discomfort will be based on analyzing important facial regions, such as eyes, eyebrows and mouth, in an automated way. With this information, we detect the behavioral state of the newborn. Second, varying, non-ideal data-capturing situations are an integral part of the design considerations of this system in order to realize sufficient robustness. Third, for the time being, we omit practical conditions such as the visibility of plasters and tubes in the data, and use recordings of young infants under regular conditions. However, we do incorporate other practical situations, such as changes in lighting conditions and viewpoint, which typically lead to suboptimal conditions for video analysis.

The remainder of this paper is organized as follows. In Sect. 2, the requirements and algorithm overview are given, Sect. 3 explains the algorithmic steps, in Sect. 4 the experiments and results are described, followed by the conclusions in Sect. 5.

2 Requirements and Algorithm Overview

Let us first discuss some system requirements, which have an impact on the algorithm design. Experts indicate that an automatic system is desired. The system should be robust against varying capturing conditions and the camera should be of limited size, preferably placed outside the incubator. Furthermore, we want the system to be extensible for the type of computing tasks to be carried out. The computational part of the system should be designed with reasonable efficiency, both in size and computation, as it is placed in the vicinity of the incubator.

We have translated the above system aspects into an efficient processing architecture, from which the system overview is displayed in Fig. 1. An image containing the face region in front of a background is the input and the recognized state

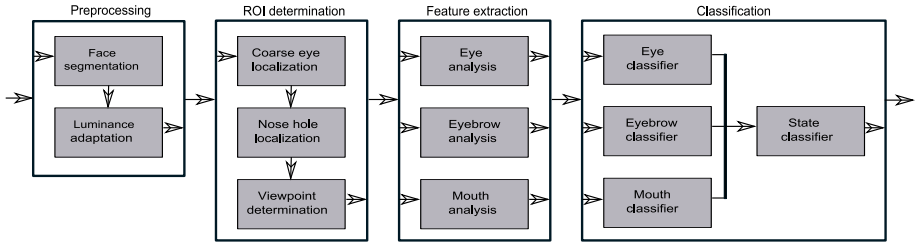


Fig. 1. System overview of the discomfort detection system

represents the output of the system. The four primary modules of the system are briefly described here and further detailed in the next section.

1. *Preprocessing*: From the input image, the face region is extracted using skin-color information. Afterwards, a lighting-compensation step is applied.
2. *Region-of-interest determination*: Within the face region, the eye and nose-hole positions are estimated. Based on the obtained locations, the Region Of Interest (ROI) is extracted automatically around all important facial components. Besides this, also the viewpoint is determined.
3. *Feature extraction*: Within the determined ROI, features are extracted from each important facial component.
4. *Classification*: A hierarchical classifier is employed to detect the behavioral state, based on the extracted features.

3 Algorithm Description

3.1 Preprocessing

From the input image, the face region is segmented from the background using skin-color information. To this end, a Gaussian skin-color model is applied in a loose way and the largest skin-colored blob is selected as face region. False face pixels are avoided by constraining the background during the recordings.

The lighting conditions are mainly non-uniform, especially in the hospital environment. To allow optimal detection results, all skin pixels should have about the same luminance value. Therefore, a lighting-compensation step is applied, aiming at removal of luminance differences within the extracted face region.

Fig. 2 illustrates both applied preprocessing steps.

3.2 Region of Interest (ROI) Determination

The ROI around the target facial components eyes, eyebrows and mouth can be defined in case the face orientation is known. This orientation is fixed in case the eyes and nostrils are localized, a process that is described in this subsection.

In literature, eyes are found using different methods. It is reported that the eyes have relatively low luminance and red intensity values [6]-[7], are elongated

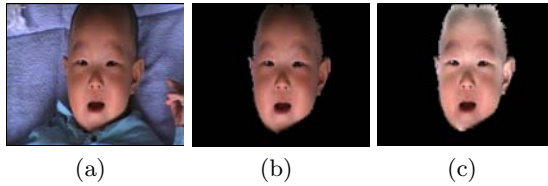


Fig. 2. Illustration of the consecutive preprocessing steps. (a): input image; (b): face segmentation; (c): result after luminance adaptation.

and are located in edge-filled areas [8]. Due to the large variability in the luminance channel Y around the eyes, the absolute gradient sum will also be large. This gradient sum G , operating in a region S around a pixel (x, y) , is defined as:

$$G(x, y) = \sum_{x', y' \in S} |Y(x, y) - Y(x', y')|. \quad (1)$$

Combination of all above-mentioned clues with the assumption that the face is approximately upright, results in multiple eye candidates. The candidates are matched pair-wise to obtain the coarse position of both eyes. Afterwards, a coordinate system is defined, with the y -axis intersecting both eye centers. The described steps are portrayed by Fig. 3 (a)-(c).

The nostrils are identified in a predefined search region along the x -axis. The large luminance differences around the nostrils result in large absolute gradient sum values. By applying an adaptive thresholding technique, the nose line, defined as the line just below the nostrils and parallel to the y -axis, is found. The ROI around the eyes, eyebrows and mouth can be identified based on both the coarse eye positions and the nose line. This is shown in Fig. 3 (d)-(e).

In order to gain robustness against viewpoint deviations, information about the horizontal viewpoint is extracted. This is important for cases where the viewpoint is not frontal, but slightly panned, resulting in an asymmetrical view where one face half is better visible than the other. Therefore, we extract the widths w_l , w_r of the left and right eyebrow area, respectively, as viewpoint indicator and transform these widths to weights η_l , η_r for both face halves, defined as:

$$\begin{aligned} \eta_l &= \begin{cases} 0 & \frac{w_l}{w_l + w_r} < 0.3, \\ 1 & \frac{w_l}{w_l + w_r} > 0.7, \\ 2.5 \cdot \left(\frac{w_l}{w_l + w_r} - 0.3 \right) & \text{else,} \end{cases} \\ \eta_r &= 1 - \eta_l. \end{aligned} \quad (2)$$

These weights can be used during the classification phase. It is experimentally found that in case one face side has a width ratio of about 30%, classification on that side becomes unreliable.

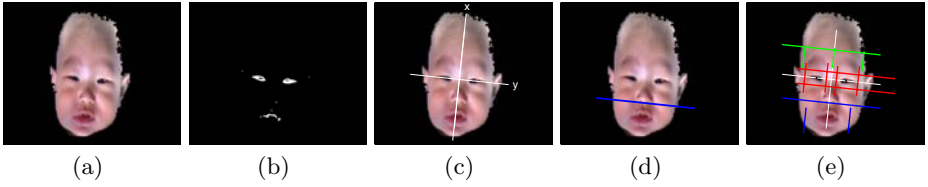


Fig. 3. Steps for ROI determination. (a): input image; (b): eye candidates; (c): coordinate system defined on obtained eye positions. (d): nose line; (e): determined ROIs.

3.3 Feature Extraction

The feature-extraction process is performed on the extracted ROIs, and is designed for robustness against viewpoint, lighting conditions and occlusions.

Eye Region. Within the eye region, a gradient-like operator is applied and the eye mask is obtained using an adaptive thresholding technique. This technique aims at selecting pixels with a ‘large’ gradient, where ‘large’ is defined based on the mean and maximum gradient at the same image row. Small noise regions are removed and a filling operation is performed to obtain the final eye mask. From this mask, the eye mask heights at $\frac{1}{3}$ and $\frac{2}{3}$ of the computed mask width are extracted as indication for the eye aperture size. Because the mask can be too wide under certain lighting conditions, the found height is decreased until the luminance value differs at least 20% from the vertical line-average. Fig. 4 illustrates the eye-region analysis process.

Eyebrow Region. In literature, the eyebrows are usually identified by their low luminance value [13]. However, we have found that the contrast between brow and skin is rather small for newborns, making the luminance value a poor discriminator under various lighting conditions. Therefore, we obtain the eyebrow mask in a different way. First, we apply the following empirically determined gradient operator:

$$X(x, y) = (R(x, y) - R(x + \delta, y)) - 0.5 \cdot (G(x, y) - G(x - \delta, y)) - 0.5 \cdot (G(x, y) - G(x - \delta, y)). \quad (3)$$

In this equation, x and y correspond with the coordinate system defined above; parameter δ is used with a value of 10. Afterwards, the same adaptive thresholding technique as used within the eye region is applied. On the acquired eyebrow mask, a three-point ‘gravity’ model is fitted, similar to [12]. The model points a , b and c (see Fig. 5 (c)) are extracted from the mask as the vertical center points at $\frac{1}{6}$, $\frac{1}{2}$ and $\frac{5}{6}$ of the mask width, where a is defined as the point closest to the nose. From these points, the cosine rule is applied to extract the angles between the vertices and the y -axis and between both vertices. As additional feature, a linear line piece is fitted through the eyebrow mask to estimate the dominant orientation of the eyebrow. Fig. 5 visualizes this process (without the latter line model).

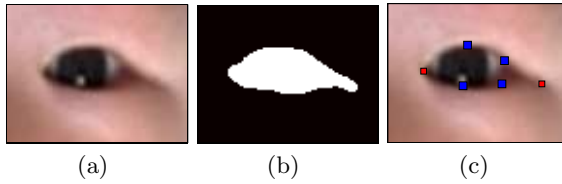


Fig. 4. Illustration of the eye investigation process. (a): input ROI; (b): eye mask; (c): extracted horizontal eye borders (red) and eye heights at $\frac{1}{3}$, $\frac{2}{3}$ of the eye width (blue).

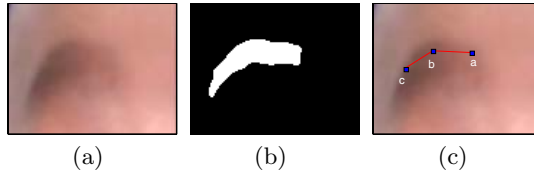


Fig. 5. Illustration of the eyebrow investigation procedure. (a): input ROI; (b): eyebrow mask; (c): fitted model, with a , b and c the model parameters.

Mouth Region. Lip-color information is often applied for mouth analysis and lip contour extraction [14]–[15]. However, for newborns, the lip color is not very clear, especially not for the upper lip. Therefore, we employ a gradient-based method for lip detection. First, to distinguish between lip and surrounding skin, a suitable color transformation is adopted from [14], which equals:

$$O = R + B - 6 \cdot G. \quad (4)$$

The same gradient-like operator is applied to signal O as within the eye region. A fixed threshold is used to remove small noise regions; the value of this threshold is not critical. Also regions with low edge densities are discarded. In the resulting mask, the blob corresponding to the upper lip has a negative value, while the blob corresponding with the lower lip is positive valued. The vertical projection histogram is employed to determine both vertical lip borders and the position of the line between both lips. The lip corners can be found on this line by combining luminance intensity, absolute gradient sum and parameter O . After this, a quadratic model is fitted on the lip blobs, using the determined lip corners and middle line. With A being the model height and w the relative distance between both lip corners, the model satisfies the relation:

$$M(w) = A \cdot \min(0.9, -4(w - 0.5)^2 + 1), \quad (5)$$

for $0 \leq w \leq 1$. The top and bottom borders of the lip mask are found by fitting model M (indicated as a white dotted contour in Fig. 6 (e)). The best fitting parameters A for both lips are extracted as features. Since the lip thickness can vary, these features cannot discriminate between a slightly open and a closed mouth. Therefore, we detect the mouth cavity, i.e. an area with both a low

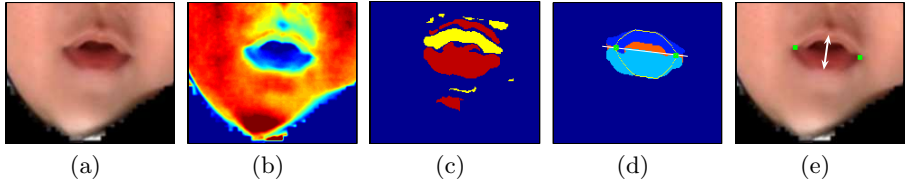


Fig. 6. Visualization of the mouth analysis process. (a): input ROI; (b): signal O ; (c): filtered gradient map, negative values are shown in yellow, positive values in red; (d): segmented lips, lip corners (green) and fitted model (yellow); (e): extracted mouth height (white) and mouth corner points (green).

luminance and a high saturation value [16]. We calculate the following function in the mouth area Q :

$$\epsilon(x) = \sum_{\substack{y \\ x, y \in Q}} \frac{S(x, y)}{Y(x, y)}, \quad (6)$$

where S , Y correspond to the saturation and luminance signals, respectively. Within the mouth area, the average of ϵ is computed at each line. The maximum average value is extracted as the third feature. The mouth-region analysis-process is displayed in Fig. 6.

Summarizing, from the determined ROIs, the following features are extracted. Within the eye regions, the eye height is extracted at two locations. From the eyebrow area, three angles are extracted from a fitted model, together with the linear regression coefficient. From the mouth region, the distance between lip and inter-lip line is estimated for both lips, together with a measure for the presence of the mouth cavity.

3.4 Behavioral State Classification

Determining the facial expression based on the above-mentioned features is similar to feature-based emotion classification, which is a major topic in literature. Different classification techniques are common, including k-Nearest Neighbors (kNN) [9], Support Vector Machines [10] and Neural Networks [11]-[12]. In the target environment, not all facial components are always visible and not all features can be extracted reliably at all time instants. Moreover, we have also extracted information about the visibility of both face sides. This makes classification based on the aforementioned techniques difficult and impractical. Therefore, instead of using one overall classifier, we propose a *hierarchical* classification method, as depicted in Fig. 7. This method is able to handle the visibility information and is extensible to both the use of additional facial features and the inclusion of other reliability information.

For the individual components, we have defined a discrete number of states: the eyes can be closed, open or wide open, the eyebrows can be normal or lowered, and the mouth can be closed, open or wide open. At the lowest classification level (*component level*), a kNN classifier is used to detect the component states, where

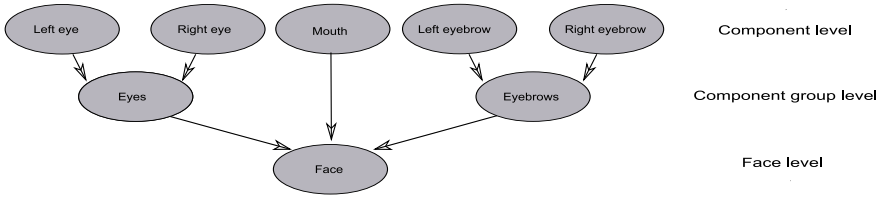


Fig. 7. Schematic overview of the hierarchical classifier

the k neighbors form evidence for a certain state. Each classifier is trained with 25 examples per state, all inferred from one newborn. Whereas a number of features consist of a measured distance on the face, these features are divided by the distance between both eye centers to obtain scale invariance.

At the medium classification level (*component group level*), the states of symmetric components (e.g. both eyes) are combined. In this step, the state for which most evidence is present is chosen. The extracted visibility fractions η_l , η_r are used to weight the evidence from the left and right components. This method assumes that both symmetric components have the same behavior; an assumption that is confirmed by medical personnel, which have experienced that for newborns the facial expression is always symmetric.

At the final classification level (*face level*), a rule-based classifier is applied to combine the component group states and detect the facial expression. Our prototype discriminates between the facial expressions corresponding to the behavioral states discomfort, awake and sleep. Note that this approach of using individual components which together contribute to a classification, is essentially the same as the pain-scale approaches reported in medical literature [2]-[3].

4 Experiments and Results

The above-described algorithm is tested on recordings of three different, healthy newborns. The age of the newborns varied from a few weeks up to a few months. A standard digital camera with at a resolution of 720×576 pixels is used, where the optical zoom is applied to focus on the infants' face. Since discomfort as defined in medical literature [1]-[3] is not common for healthy newborns, we aim at detection of a specific type of cry, with a facial expression close to the expression of medical discomfort. The target expression consists of closed or just opened eyes, a wide open mouth and lowered eyebrows.

The total of 40 sequences is grouped into two different classes. Class I contains normal situations, with a near low-frontal viewpoint and reasonable lighting conditions (21 sequences). Class II consists of sequences with non-frontal viewpoints and poor lighting conditions (19 sequences). Since two newborns already were aged a few months, they were able to make large movements. Due to the caused motion and compression artifacts, not all frames are suitable for analysis. Furthermore, if adjoining frames are too similar, a representative example

selection is made. Thus, from each shot of each sequence, a number of significantly differing key frames are evaluated.

4.1 Class I: Normal Lighting Conditions With a Frontal Viewpoint

The presented system is first evaluated for images with a near-low frontal viewpoint, where the nostrils are clearly visible, and with acceptable lighting conditions. We address the results of the feature extraction process, component-level classifiers and complete system independently.

The feature points are extracted reliably in most cases, although small shadows and reflections can cause the features to be detected a few pixels from the optimal location. In case the eyes are stiffly closed, the coarse eye localization process can fail, causing the algorithm to abort. Also, for one of the newborns, the eyebrows were too light, and a shadow is detected as brow. Doctors indicate that this is typical for some newborns. However, in case of discomfort the eyebrows bulge, causing dark horizontal shadow lines. This enables detection of discomfort even when the brows themselves are not visible. Furthermore, for the somewhat older newborn, hair and eyelashes interfere with the eyebrows and eyes. This caused a slightly higher eye-height for closed eyes, thereby leading to misclassifications.

The results of the component-level classifiers are shown in Table 1. For the eyebrows only the results for clearly visible brows are listed; undetected brows are marked as *undefined*. As a result of using a discrete number of states, borderline cases can cause classification errors, since the feature values corresponding to a borderline case vary a little per newborn. Whereas this difference especially occurs for the *wide open* states, we expect that this deviation can be corrected by using a newborn-dependent normalization-coefficient. At component group level, most of these errors are corrected for the eyes and eyebrows.

The classification results of the complete system are shown in Table 2. The classification result can be seen as promising: in each of the experiments, only a few situations (around 5%) are classified into a different state. Since the system operates on frame basis, misclassifications can be caused by transients such as

Table 1. Performance matrix of the component level classifier

Eyes	Found State:	Wide open	Open	Closed	Total
	Wide open	89	7	0	96
	Open	4	75	1	80
	Closed	0	12	54	66
Eyebrows	Found State:	Normal	Lowered	Undefined	Total
	Normal	86	4	16	106
	Lowered	1	27	0	28
Mouth	Found State:	Wide open	Open	Closed	Total
	Wide open	28	4	0	32
	Open	0	35	5	40
	Closed	0	5	44	49

Table 2. Performance matrix of the complete classification system

Found State:	Sleep	Awake	Discomfort	Total
Sleep	49	1	0	50
Awake	3	47	0	50
Discomfort	0	2	28	30

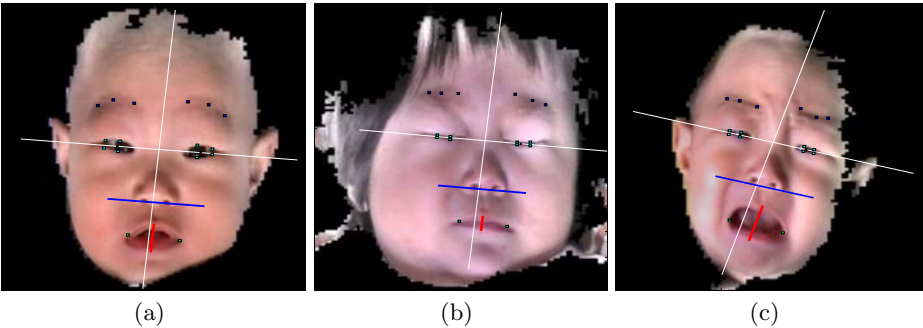


Fig. 8. Examples of extracted features for each of the three newborns. The white lines indicate the determined coordinate system; the horizontal line at the nose bottom corresponds to the obtained nose line. In (a) and (b), the newborns are shown during normal awake and sleep behavior, respectively. In (c), a discomfort situation is shown.

eye blinks. This can be solved by determining the state based on a window of e.g. 2 seconds, but since the newborns moved a lot especially during cry, this would reduce the amount of test data. With respect to cry, only one newborn showed the target type of cry during the recordings, but this situation occurred at various times in multiple sequences. The other newborns only showed normal behavior. Examples of investigated situations in Class I are displayed in Fig. 8.

4.2 Class II: Robustness to Lighting and Viewpoint Deviations

This class of test data has been explored to define difficult situations, classification problems etc., with the aim to outline future research directions. The tests focus on two major issues: lighting condition and viewpoint changes. Since this class only contains a few different samples per specific non-ideal situation, the test set contents would highly influence the results; therefore, no numerical results are given.

Various lighting conditions can be handled, and in many cases feature extraction for the eyes and mouth succeeds with small deviations from the expected values. However, it was found that feature extraction in the eyebrow area is sensitive to shadows and that reflections on the skin can cause improper skin-color segmentation and feature-extraction errors, e.g. when a part of the eyebrow is subject to the reflection. Fig. 9 shows investigated lighting conditions.

Viewpoint changes over the horizontal direction can be handled, unless there is no skin visible next to both eyes, disabling estimation of the eye positions.



Fig. 9. Screenshots from evaluated sequences with different lighting conditions. For all three examples, the feature points can be extracted and classification can be performed.



Fig. 10. Screenshots from investigated sequences with different viewpoints. For all three examples, the feature points on the most important face side can be extracted.

Viewpoint deviations over the vertical direction are harder to handle: when viewing more from above, nostril detection is difficult, resulting in inaccurate ROIs. For large vertical deviations, the mouth cavity cannot always be found, since then either the palate or tongue is visible instead of the cavity. Note that the performance sensitivity in classification to vertical changes in viewpoint is partly explained by the measurement of a number of vertical distances on the face. Fig. 10 displays examples of investigated viewpoints.

5 Conclusions and Future Work

This paper has described a prototype system for automatic detection of the facial expression of newborns to estimate their behavioral state. In contrast with earlier work, we have developed a component-based approach with increased robustness against non-ideal viewpoints and lighting conditions. The system extracts features from the eye, eyebrow and mouth regions; a hierarchical classifier is employed to distinguish between the states sleep, awake and discomfort.

Tests on three healthy newborns show that our prototype system can be used for determination of the states of the above-mentioned facial components with an acceptable accuracy (approximately 88,2%). The obtained states are successfully combined to detect present discomfort, where normal situations, including light cries, are rejected. The system is able to determine the behavioral state with about 95% accuracy. When non-ideal viewpoints and poor lighting conditions are present, the system can still detect the facial states in many cases. Therefore, the algorithm is relatively robust against varying situations. An exploration with more extreme conditions has shown that vertical viewpoint changes and poor lighting conditions such as reflections can decrease the score and reliability of

the system. However, since only one newborn showed all possible behavioral states and the total number of investigated newborns is small, we note that the reliability of the experiment is too limited for a real performance evaluation.

Application in a realistic hospital environment is not directly possible, since the system should then also be able to deal with skin-colored plasters, tubings in the nose and large horizontal viewpoint changes. These items are the main research topics for future work, together with newborn-independent classification and an extensive validation of our proposed approach.

References

1. Grunau, R.V.E., Craig, K.D.: Pain expression in neonates: facial action and cry. *Pain* 28(3), 395–410 (1987)
2. Stevens, B., Johnston, C., Pethryshen, P., Taddio, A.: Premature Infant Pain Profile: Development and Initial Validation. *Clin. J. Pain* 12(1), 13–22 (1996)
3. Chen, K., Chang, S., Hsiao, T., Chen, Y., Lin, C.: A neonatal facial image scoring system (NFISS) for pain response studies. *BME ABC*, 79–85 (2005)
4. Brahnham, S., Chuang, C., Shih, F.Y., Slack, M.R.: Machine recognition and representation of neonatal facial displays of acute pain. *Artificial Intelligence in Medicine* 36(3), 211–222 (2006)
5. Brahnham, S., Chuang, C., Sexton, R.S., Shih, F.Y.: Machine assessment of neonatal facial expressions of acute pain. *Special Issue on Decision Support in Medicine in Decision Support Systems* 43, 1247–1254 (2007)
6. Peng, K., Chen, L.: A Robust Algorithm for Eye Detection on Gray Intensity Face without Spectacles. *Journal of Computer Science and Technology*, 127–132 (2005)
7. Vezhnevets, V., Degtiareva, A.: Robust and Accurate Eye Contour Extraction. In: *Proc. Graphicon*, pp. 81–84 (2003)
8. Asteriadis, S., Nikolaidis, N., Hajdu, A., Pitas, I.: A novel eye-detection algorithm utilizing edge-related geometrical information. In: *EUSIPCO 2006* (2006)
9. Sohail, A.S.M., Bhattacharya, P.: Classification of facial expressions using k-nearest neighbor classifier. In: *Gagalowicz, A., Philips, W. (eds.) MIRAGE 2007. LNCS*, vol. 4418, pp. 555–566. Springer, Heidelberg (2007)
10. Michel, P., Kaliouby, R.: Real time facial expression recognition in video using support vector machines. In: *ICMI* (2003)
11. Ioannou, S.V., Raouzaoui, A.T., Tzouvaras, V.A., Mailis, T.P., Karpouzis, K.C., Kollias, S.D.: Emotion recognition through facial expression analysis based on a neurofuzzy network. *Neural Networks* 18, 423–435 (2005)
12. Tian, Y., Kanade, T., Cohn, J.F.: Recognizing action units for facial expression analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23(2), 97–115 (2001)
13. Chen, Q., Cham, W., Lee, K.: Extracting eyebrow contour and chin contour for face recognition. *Pattern Recognition* 40(8), 2292–2300 (2007)
14. Gomez, E., Travieso, C.M., Briceno, J.C., Ferrer, M.A.: Biometric identification system by lip shape, *Security Technology* (2002)
15. Leung, S., Wang, S., Lau, W.: Lip image segmentation using fuzzy clustering incorporating an elliptic shape function. *IEEE Transactions on Image Processing* 13(1), 51–61 (2004)
16. Gocke, R., Millar, J.B., Zelensky, A., Robert-Ribes, J.: Automatic extraction of lip feature points. In: *Proc. of ACRA 2000*, pp. 31–36 (2000)