Automatic and continuous discomfort detection for premature infants in a NICU using video-based motion analysis

Yue Sun[1], Deedee Kommers[1,2], Wenjin Wang[3], Rohan Joshi[3], Caifeng Shan[3], Tao Tan[1], Ronald M. Aarts[1], Carola van Pul[1,2], Peter Andriessen[2], and Peter H.N. de With[1]

***Abstract*— Frequent pain and discomfort in premature infants can lead to long-term adverse neurodevelopmental outcomes. Video-based monitoring is considered to be a promising contactless method for identification of discomfort moments. In this study, we propose a video-based method for automated detection of infant discomfort. The method is based on analyzing facial and body motion. Therefore, motion trajectories are estimated from frame to frame using optical flow. For each video segment, we further calculate the motion acceleration rate and extract 18 time- and frequency-domain features characterizing motion patterns. A support vector machine (SVM) classifier is then applied to video sequences to recognize infant status of comfort or discomfort. The method is evaluated using 183 video segments for 11 infants from 17 heel prick events. Experimental results show an AUC of 0.94 for discomfort detection and the average accuracy of 0.86 when combining all proposed features, which is promising for clinical use.**

I. INTRODUCTION

Preterm birth is when an infant is born too early, before 37 weeks of gestation [1]. Based on statistics for 184 countries, the global average preterm birth rate in 2010 was 11.1%, giving a worldwide total of 14.9 million infants [2]. Neurobiological vulnerability to pain in preterm infants is well established, due to their lower pain threshold, sensitization from repeated pain, and immature systems for maintaining homeostasis [3][4]. Frequent and/or recurrent pain/discomfort is one of the greatest burdens for preterm infants, who undergo prolonged hospitalization at a time of physiological immaturity. At the same time of rapid brain development discomfort can cause complications, such as delay in cognitive and motor development [5]. More importantly, cumulative pain-related discomfort may also contribute to abnormal brain development, which thereby yields long-term adverse neurodevelopmental outcomes [6][7][8][9]. These findings suggest that continuous monitoring of preterm infants for discomfort or pain would be of high value to help develop appropriate treatments.

Preterm infants receive special care in the neonatal intensive care unit (NICU), where their vital signs are continuously monitored. However, there is currently no system for monitoring and detecting their stress/discomfort. Self-reporting is recognized as the standard approach for monitoring pain in clinical practice and is the current definitive indicator of presence and intensity of pain and discomfort [10]. However, we cannot obtain a verbal (or nonverbal, deliberate hand gestures, head nods/head shakes) self-report of pain/discomfort from preterm infants. Monitoring by healthcare professionals invokes high cost, and is time-consuming and subjective in assessment. Because of these limitations, infants are only observed during short intervals a few times a day, which likely leaves many discomfort moments unnoticed. Therefore, we propose an automatic and continuous discomfort detection system for preterm infants by analyzing their motion patterns using video monitoring.

II. RELATED WORK

In the past several years, there has been an increasing interest in stress/pain assessment [11]. Various approaches have been developed to assess pain based on physiological indicators, for instance vital signs such as heart rate (HR), heart rate variability (HRV), respiratory rate (RR), oxygen saturation (SpO2), body temperature, and blood pressure. These signs can be measured and used to assess the physical functioning level of a person. Acharya *et al.* [12] detected cardiac abnormalities by classifying cardiac rhythms using an artificial neural network and fuzzy relationships, which achieved an accuracy level of 80-85%. However, vital signs such as HR and RR are currently measured by electrocardiograms (ECG) and photoplethysmograph (PPG), which require contact with the patient's skin. Attaching the sensors to infant skin adds an extra burden to infants compared to a contactless method, for instance, using video monitoring.

Besides the physiology-based approaches, there is another category of methods that assess pain/discomfort based on behavior analysis. Existing behavioral-based approaches to evaluate infant pain can be based on facial expression and crying sound. Infants' cry is a common sign of discomfort, hunger, or pain. For classifying crying sound, Mima *et al.* [13] presented a method that analyzes baby cries in spectrography, and classifies them as cries due to pain, sleeping, hunger, etc. The obtained overall accuracy of the proposed method was 85%.

Significant attention was paid to facial expressions in adults. Shan *et al.* [14] empirically evaluated facial representation based on statistical local features, local binary patterns (LBP), for person-independent facial expression recognition and illustrated that LBP features are effective and efficient

[1]Yue Sun, Deedee Kommers, Tao Tan, Ronald M. Aarts, Carola van Pul and Peter H.N. de With are with Eindhoven University of Technology, 5612 WH Eindhoven, The Netherlands

[2]Deedee Kommers, Carola van Pul and Peter Andriessen are with Máxima Medical Center, 5504 DB, Veldhoven, The Netherlands

[3]Wenjin Wang, Rohan Joshi and Caifeng Shan are with Philips Research, Philips Research, High Tech Campus 34, 5656AE Eindhoven, The Netherlands

✉ caifeng.shan@philips.com

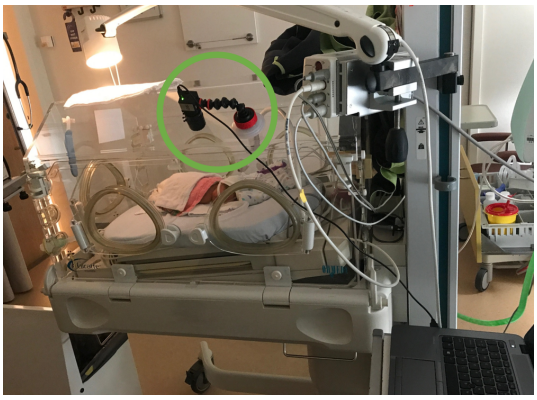

Fig. 1. Video acquisition system. The camera is outlined by a green circle.

for facial expression recognition. Kotsia *et al.* [15] achieved a recognition accuracy of 99.7% for facial expression recognition using the proposed multiclass SVMs and 95.1% for facial expression recognition based on a set of chosen facial action units.

Since the facial expression in premature infants in the NICU is usually occluded by breathing masks/feeding tubes, it is not suitable to fully rely on a facial expression method for discomfort monitoring. A good alternative is video-based body motion analysis. So far, very few studies reported pain assessment for premature infants based on body movements. Cattani *et al.* characterize clonic seizures and apneas by the presence or absence of periodic movements of parts of the body [16]. However, the method is intensity-based (in the paper called luminance) and thus is affected by the lighting condition.

In our work, we propose a video-based automatic system for detecting discomfort moments in premature newborns in the NICU. Our work shows that motion-related features can differentiate discomfort status from comfort.

III. METHOD

To detect the body motion of the infants, we employ optical flow to estimate pixel motion vectors between frames, which is followed by feature extraction for discomfort/comfort classification. The choice for optical flow is motivated by its capability to capture the motion of individual body parts more accurately [17].

A. Study design and population

The study was conducted with videos recorded at the Máxima Medical Center in Veldhoven, the Netherlands, by a fixed-position high-definition camera (uEye UI-222x) filming the infant's face and upper body (See Fig. 1 for the video acquisition system). For all infants in the database, written consent was obtained from the parents.

Heel lance procedure, a well-known pain stimulus, was part of regular care for collecting blood samples to monitor glucose, bilirubin, etc. In our work, it served as a recurring stimulus to study the infant's response to pain. The video recording started approximately 10 minutes prior to the heel lance procedure. Once the heel lance procedure was finished, the video recording continued for an observation time to return to baseline (10 minutes). We define the baseline as a period when there is no observable discomfort motion pattern/facial expression. The start and end time of the heel lance procedure was simultaneously noted by a research assistant.

Discomfort and comfort video segments were annotated by a researcher according to the timeline relative to the heel prick intervention. Discomfort video segments were labeled from the start point of heel prick to several minutes after the heel prick was done based on the researcher's observation. Comfort video segments were labeled from the baseline prior to the prick and the moment when the infant returned to baseline after the heel prick. Each video segment contains only one state (comfort or discomfort). All the moments that show interruption or occlusion from caregivers in the videos were excluded. The video segments lasting less than 10 seconds were removed from consideration. Eleven infants with an average gestational age of 31 weeks were filmed. We totally achieved 99 discomfort (2,738 seconds in total) and 84 comfort (3,429 seconds in total) video segments for 17 heel prick events. The duration of each video segment varies from less than 1 minute to several minutes with an overall median length of 21.2 seconds (interquartile range [IQR] 12.8 - 39 seconds).

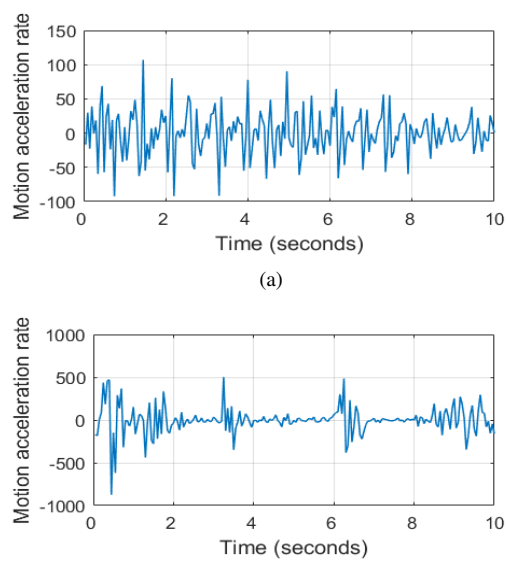

Fig. 2. Examples of extracted motion acceleration rate of (a) comfort, and (b) discomfort moments, where different motion patterns are shown in terms of movement intensity and periodicity.

B. Motion Estimation

Pixel-based motion vectors are first calculated for each video frame, with respect to the previous frame, using the optical flow proposed by Farnebäck *et al.* [17]. For the optical flow, we compute a motion matrix $\mathbf{M}$ of size $N \times W$, where N is the number of pixels in a video frame and W is the total number of frames in a video segment. The optical flow provides motion derivatives for each row of the matrix that represents the velocity magnitude of a pixel's trajectory.

We accumulate the magnitude values of all motion vectors for each frame. Hence, all the summed magnitude values comprise a one-dimensional (1D) velocity-estimating signal $\mathbf{V}$ (size $1 \times W$) for each video segment. We further estimate the motion acceleration rate $\mathbf{A}$ (size $1 \times W$) by taking the first derivative of the velocity-estimating signal $\mathbf{V}$. Fig. 2(a) and Fig. 2(b) show examples of the motion acceleration rate extracted from a comfort and a discomfort video segment, respectively.

C. Feature extraction

For each video segment, features are extracted from the 1D signal of the motion acceleration rate $\mathbf{A}$.

We calculate two groups of features: statistical and spectral, which are shown in Table I. We compute in total 18 features: mean, median, root mean square (see Eq. (1)), a group of 3 features from the autocorrelation function, and a group of 12 spectral peak features, which are discussed below.

- Mean, Median of the motion acceleration rate $\mathbf{A}$
- Root mean square

The root mean square (RMS) of the motion acceleration rate elements from $\mathbf{A}$ is utilized as an input to a classifier for infant status recognition. The RMS value is calculated according to:

$$RMS = \sqrt{\frac{1}{W}\sum_{i=1}^{W} {x_i}^2}, \tag{1}$$

where x_i is individual acceleration instance, and W the total number of frames of the video segment.

- Autocorrelation features

The autocorrelation characterizes different types of motion signals by indicating motion intensity and periodicity. We calculate the autocorrelation of motion acceleration rate $\mathbf{A}$, which is formulated as

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{W} (\mathbf{A}_t - \bar{\mathbf{A}}) \cdot (\mathbf{A}_{t-k} - \bar{\mathbf{A}})}{\sum_{t=1}^{W} (\mathbf{A}_t - \bar{\mathbf{A}})^2}, \tag{2}$$

where $\mathbf{A}_{t-k}$ is the motion acceleration rate shifted by k frames, and $\bar{\mathbf{A}}$ is the average of the motion acceleration rate. The numerator of Eq. (2) is essentially the covariance between the original acceleration rate and the k-frame lagged data. The denominator is the sum of the squared deviations of the original acceleration rate.

The peak height at zero-th lag (overall energy) is employed as an individual feature. The other two features are the height and location at the first peak, respectively, which identifies the dominant cyclic variation in motion.

- Spectral peak features

We further estimate the power spectral density on motion acceleration rate $\mathbf{A}$ using Welch's method [18]. From the derived spectrum, positions and power levels of the highest 6 peaks are taken as 12 spectral peak features.

TABLE I

FEATURE CATEGORIES WITH THEIR CORRESPONDING INTERPRETAION AND MEDIAN VALUES FOR ALL COMFORT (CMVAL) AND DISCOMFORT CASES (DMVAL).

Feature category	Interpretation	CMVAL	DMVAL
Mean	Motion intensity	0.015	-0.275
Median	Motion intensity	-0.091	0.296
RMS	Motion intensity	17.1	237
Autocorrelation - height at 0th lag	Motion intensity	2.16×10^5	2.45×10^7
Autocorrelation - 1st peak location	Motion periodicity	0.2	0.2
Autocorrelation - 1st peak height	Motion intensity	1.20×10^4	3.34×10^5
Spectrum - positions of highest 6 peaks	Movement frequency	3.63 4.67 5.41 6.14 6.86 7.59	1.88 3.07 3.93 5.16 6.22 7.28
Spectrum - power levels of highest 6 peaks	Motion level at each frequency	115 150 150 121 119 129	2.21×10^4 2.30×10^4 2.54×10^4 2.68×10^4 2.66×10^4 2.26×10^4

D. Classification

Finally, we adopt a linear support vector machine (SVM) classifier on video segments to recognize infant status of comfort or discomfort, using the 18 extracted features. We employ the SVM implementation in Matlab (Mathworks, Natick, MA, USA) for the two-class classification. Leave-one-infant-out cross-validation is used for the experiments. The receiver operating characteristic (ROC) is plotted to evaluate the performance with the value of the area under the curve (AUC). The classification accuracy is also measured and reported as an evaluation metric.

IV. EXPERIMENTAL RESULTS

Using the recorded infant dataset, leave-one-infant-out cross-validation was performed for evaluating our proposed method. The classification accuracy is summarized in Table II.

TABLE II

PERFORMANCE MEASURES INCLUDING CLASSIFICATION ACCURACY, AND AUCS OF DIFFERENT CATEGORIES OF FEATURES.

Feature category	Accuracy rate	AUC
Mean	0.50	0.72
Median	0.39	0.53
RMS	0.81	0.91
Autocorrelation	0.68	0.84
Spectrum	0.75	0.91
Combined	**0.86**	**0.94**

The ROC curves are plotted in Fig. 3. From the light-blue curve based on all features, we can detect 85% of discomfort

video segments at the cost of only 10% of false alarms. When all features are combined, the average accuracy for all infants is 0.86 (see Table II).

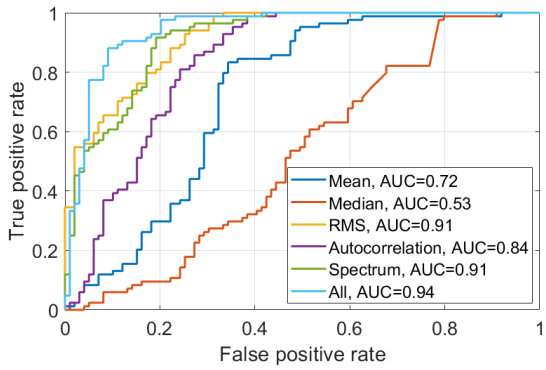

Fig. 3. ROC curves for classification using each individual category of features, and combining all features together.

V. Discussion and Conclusions

For the purpose of anticipating on pain and discomfort in premature infants, we propose an automated video-based system that can differentiate discomfort of infants from comfort status by analyzing motion patterns. The recognition process starts with the acquisition of the motion signals, which are subsequently estimated by optical flow. For each video segment, a vector of 18 features characterizing motion trajectories is extracted from the motion signals in the time and frequency domain. An AUC of 0.94 is achieved, which is promising for clinical practice. The highest AUC is achieved when combining all the proposed features, which proves that the features are contributing and complementary. We have extracted only 18 features from each video segment, which is particularly suitable for a real-time video-based application.

Due to incorporating autocorrelation and spectrum features, which can capture and describe periodic motion patterns, our system can recognize discomfort body motion and regular periodic movements (e.g. respiration). With the high sensitivity of 85%, the trade-off of the system is only 10% detecting false positives.

In reality, video interruptions may occur by care-handling the infant and, thereby, occluding the infant's face and/or body. Moreover, for most critically ill infants, no response to pain stimulus can be observed visually. In these respects, for further work, we may add features extracted from the monitored vital signs, such as HR, RR, and SpO2.

For practical application, further lowering the amount of false positives would be valuable. A majority of these false positives are segments with irregular comfortable movements (e.g. stretching an arm). With limited effort of human assistance (e.g. visual check), these false positives might be easily corrected.

Acknowledgment

The authors would like to thank all the medical staff from the NICU department of Máxima Medical Center, Veldhoven for the cooperation and assistance in collecting the infant videos.

References

[1] World Health Organization, "Who: recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths. Modifications recommended by FIGO as amended October 14, 1976," *Acta Obstet Gynecol Scand*, vol. 56, no. 3, pp. 247–253, 1977.

[2] H. Blencowe, S. Cousens, M. Z. Oestergaard, D. Chou, A.-B. Moller, R. Narwal, A. Adler, C. V. Garcia, S. Rohde, L. Say, *et al.*, "National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications," *The Lancet*, vol. 379, no. 9832, pp. 2162–2172, 2012.

[3] M. Fitzgerald, "The development of nociceptive circuits," *Nature Reviews Neuroscience*, vol. 6, no. 7, pp. 507–520, 2005.

[4] M. Fitzgerald, C. Millard, and N. McIntosh, "Cutaneous hypersensitivity following peripheral tissue damage in newborn infants and its reversal with topical anaesthesia," *Pain*, vol. 39, no. 1, pp. 31–36, 1989.

[5] R. E. Grunau, M. F. Whitfield, J. Petrie-Thomas, A. R. Synnes, I. L. Cepeda, A. Keidar, M. Rogers, M. MacKay, P. Hubber-Richard, and D. Johannesen, "Neonatal pain, parenting stress and interaction, in relation to cognitive and motor development at 8 and 18 months in preterm infants," *Pain*, vol. 143, no. 1, pp. 138–146, 2009.

[6] J. R. Petrini, T. Dias, M. C. McCormick, M. L. Massolo, N. S. Green, and G. J. Escobar, "Increased risk of adverse neurological development for late preterm infants," *The Journal of pediatrics*, vol. 154, no. 2, pp. 169–176, 2009.

[7] R. Whit Hall and K. Anand, "Short-and long-term impact of neonatal pain and stress," *NeoReviews*, vol. 6, pp. 69–75, 2005.

[8] R. Behrman and A. Stith Butler, "Institute of medicine committee on understanding premature birth and assuring healthy outcomes board on health sciences outcomes: preterm birth: causes, consequences, and prevention," *Preterm birth: causes, consequences, and prevention, National Academies Press, Washington, DC*, 2007.

[9] T. N. Raju, R. D. Higgins, A. R. Stark, and K. J. Leveno, "Optimizing care and outcome for late-preterm (near-term) infants: a summary of the workshop sponsored by the national institute of child health and human development," *Pediatrics*, vol. 118, no. 3, pp. 1207–1214, 2006.

[10] U. D. of Health, H. Services, *et al.*, "Acute pain management in infants, children, and adolescents: Operative and medical procedures," *Agency for Health Care Policy and Research: Rockville, Maryland*, 1992.

[11] Z. Hammal and J. F. Cohn, "Towards multimodal pain assessment for research and clinical use," in *Workshop on Roadmapping the Future of Multimodal Interaction Research including Business Opportunities and Challenges*, vol. 3, 2014, pp. 13–17.

[12] R. Acharya, A. Kumar, P. Bhat, C. Lim, N. Kannathal, S. Krishnan, *et al.*, "Classification of cardiac abnormalities using heart rate signals," *Medical and Biological Engineering and Computing*, vol. 42, no. 3, pp. 288–293, 2004.

[13] Y. Mima and K. Arakawa, "Cause estimation of younger babies' cries from the frequency analyses of the voice-classification of hunger, sleepiness, and discomfort," in *Intelligent Signal Processing and Communications, 2006. ISPACS'06. International Symposium on*. IEEE, 2006, pp. 29–32.

[14] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803–816, 2009.

[15] I. Kotsia and I. Pitas, "Facial expression recognition in image sequences using geometric deformation features and support vector machines," *IEEE transactions on image processing*, vol. 16, no. 1, pp. 172–187, 2007.

[16] L. Cattani, D. Alinovi, G. Ferrari, R. Raheli, E. Pavlidis, C. Spagnoli, and F. Pisani, "Monitoring infants by automatic video processing: A unified approach to motion analysis," *Computers in biology and Medicine*, vol. 80, pp. 158–165, 2017.

[17] G. Farnebäck, "Two-frame motion estimation based on polynomial expansion," in *Scandinavian conference on Image analysis*. Springer, 2003, pp. 363–370.

[18] F. J. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, 1978.