Camera based Respiration Rate of Neonates by Modeling Movement of Chest and Abdomen Region

Vishnu Makkapati
Myntra Designs Pvt. Ltd.
Bangalore - 560 068, India
Email: vishnu.makkapati@myntra.com

Pujita Raman
Qualcomm Technologies, Inc.
San Diego, CA 92121
Email: praman@qti.qualcomm.com

Gautam Pai
Technion - Israel Institute of Technology
Haifa 3200003
Email: paigautam@techunix.technion.ac.il

***Abstract*—Respiration Rate (RR) is one of the important parameters used for monitoring of neonates in a Neonatal Intensive Care Unit (NICU). In this paper, we propose a contactless method of measuring RR using a camera. Our algorithm exploits the motion of the chest and abdomen during the breathing cycle of a neonate. The chest and abdomen regions will typically expand and contract during inspiration and expiration phases respectively. We model this natural movement by capturing the motion of these regions using optical flow. The optical flow vectors converge and diverge during expiration and inspiration respectively and we capture this pattern by using a convergence and divergence filter. We use Independent Component Analysis (ICA) to separate the natural motion of the chest and abdomen from noise. To reduce the computational complexity, we use the convergence and divergence filter in Clifford Fourier Transform (CFT) of the vector field. We show the efficacy of our method by comparing the results with the ground truth RR collected from a patient monitor.**

I. Introduction

Contactless monitoring of neonates in a Neonatal Intensive Care Unit (NICU) is highly desired due to their fragile skin. In a NICU setup, the neonates are monitored for vital signs such as Heart Rate (HR), Respiration Rate (RR) and Oxygen Saturation (SPO2) by using sensors that are in contact with the neonate e.g., Electrocardiography (ECG) leads and pulse oximeter. However the use of such probes in direct contact with the neonate causes discomfort and in some cases damage to the fragile skin. Camera lends itself naturally for contactless monitoring of the neonates.

Recently, there have been several attempts to capture RR using a camera [1]. A thermal imaging system has been proposed to monitor respiration rate in [2]. Temperature extremities are used to locate head and face. RGB values from this region are plotted to obtain a respiration signal. The method in [3] analyzes the distortions in reflections of structured illumination patterns and uses the 3D time-series data derived from it to estimate respiration rate. A Kinect sensor is used to monitor breathing rate in [4]. An infrared (IR) dot pattern is projected on the subject and the dots are detected using a camera with a matching IR filter. The dots are tracked and the trajectories are merged using Principal Component Analysis and Autoregressive Spectral Analysis to estimate RR. However, all these solutions require special purpose cameras and/or dedicated illumination.

The method in [5] tries to uncover the temporal dynamics of breathing rate using delay-coordinate transformation and independent component analysis (ICA). The average brightness in the upper body is calculated for each frame to form a single observed time series over time. The method, however, does not utilize the motion of the chest/abdomen regions to derive RR. In [6], the chest motion is encapsulated by performing a profile correlation of the chest over time. However, the approach suffers from the inherent drawback of the angle of view of the camera with respect to the subject where the vertical movement of the chest is expected to be the dominant motion corresponding to respiration.

Our main contribution in this paper is to show how we can model breathing movement using computer vision techniques by exploiting the natural motion of chest/abdomen region. We analyze the pattern of optic flow vectors computed from the chest/abdomen region of the subject [7], and show that the inhalatory and expiratory breathing movements can be very effectively modeled by computing the convergence and divergence of the optic flow field. We employ Clifford Fourier Transform to efficiently compute them. To separate breathing motion from other artifacts we employ ICA.

Patent pending. This work was performed when V. Makkapati and G. Pai were with Philips Research India and P. Raman was an intern there.

II. Background

A. Clifford Fourier Transform

The basis of 2D Clifford algebra [8] is given by $\mathbf{1}, \mathbf{e_1}, \mathbf{e_2}, \mathbf{e_{12}}$, where $\mathbf{e_1}$ and $\mathbf{e_2}$ are orthonormal vectors analogous to their respective 2D counterparts and $\mathbf{e_{12}}$ is the unit bi-vector which represents an oriented area. The elements of Clifford algebra are called multi-vectors. Each multi-vector is represented as a weighted summation of the basis vectors.

$$A = a_0 + a_1\mathbf{e_1} + a_2\mathbf{e_2} + a_3\mathbf{e_{12}} \tag{1}$$

The following axioms define $\mathbf{e_{12}}$ as well as multiplication in a 2D Clifford algebra:

$$\mathbf{e_1 e_2} = -\mathbf{e_2 e_1} = \mathbf{e_{12}} \tag{2}$$

$$\mathbf{1 e_i} = \mathbf{e_i} \quad i = 1, 2, 3.. \tag{3}$$

$$\mathbf{e_i e_i} = \mathbf{1} \quad i = 1, 2, 3.. \tag{4}$$

Fig. 1. Schematic showing the optimal mount position for a camera in a warmer (red dot)

Thus, we can see that $\mathbf{e_{12}e_{12}} = -\mathbf{1}$ and therefore we can define $\mathbf{j} = \mathbf{e_{12}}$ analogous to the complex number $\mathbf{i}$ (since $\mathbf{i}^2 = -1$). Therefore, the 2D Euclidean space $\mathbf{R}^2$ may be represented in the Complex space $\mathbf{C}$ by multiplying all elements in $\mathbf{R}^2$ with $\mathbf{e_1}$ or $\mathbf{e_2}$ from any side. The inverse operation is again multiplication with $\mathbf{e_1}$ from the left. Thus for
$A = a_1\mathbf{e_1} + a_2\mathbf{e_2} \in \mathbf{R}^2$, we define:

$$\tilde{A} = \mathbf{e_1}(a_1\mathbf{e_1} + a_2\mathbf{e_2}) = a_1 + a_2\mathbf{e_{12}} = a_1 + a_2\mathbf{j} \quad (5)$$

Similarly, a 2D vector field given by $\mathbf{v}(\mathbf{x}) = v_1(\mathbf{x})\mathbf{e_1} + v_2(\mathbf{x})\mathbf{e_2}$, where $\mathbf{x} = x_1\mathbf{e_1} + x_2\mathbf{e_2} \in \mathbf{R}^2$ can be written as $\mathbf{v}(\mathbf{x}) = v_1(\mathbf{x}) + v_2(\mathbf{x})\,\mathbf{j}$ in $\mathbf{C}$.

Analogous to the Fourier Transform for scalar fields, the Clifford Fourier Transform [9] is used to analyze vector fields in the frequency domain. Since the entire vector field can be mapped onto the complex space $\mathbf{C}$, the Clifford Fourier Transform $\mathcal{C}$ can be implemented using a classical complex Fourier transform F as:

$$\mathcal{C}\{\mathbf{v}(\mathbf{x})\}(\omega) = \mathbf{e_1}\left[\mathcal{F}\{v_1(x)\} + \mathcal{F}\{v_2(x)\}j(\omega)\right] \quad (6)$$
$$= \mathbf{e_1}\left[V_1(\omega) + V_2(\omega)j\right] \quad (7)$$

B. The Convergence Divergence Filter

Now, using the properties of Clifford algebra, the divergence operator applied in the frequency domain may be obtained as:

$$div(\mathbf{v}(\mathbf{x})) \overset{\mathcal{F}}{\leftrightarrow} \mathbf{j}\omega_1 V_1(\omega) + \mathbf{j}\omega_2 V_2(\omega) \quad (8)$$

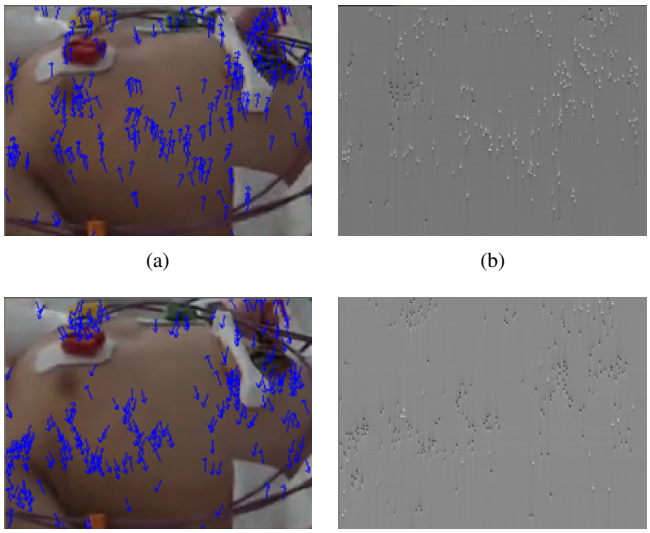
where $V_1(\omega)$ and $V_2(\omega)$ are the standard 2D fourier transforms of $v_1(\mathbf{x})$ and $v_2(\mathbf{x})$ respectively.

Similarly, the curl of the 2D vector field in this case, would be a vector pointing in the direction perpendicular to the 2D flow field and it's value is given by:

$$curl(\mathbf{v}(\mathbf{x}))_3 = -\frac{\partial v_2(\mathbf{x})}{\partial x_1} + \frac{\partial v_1(\mathbf{x})}{\partial x_2} \quad (9)$$

and therefore in the Clifford Fourier domain:

$$curl(\mathbf{v}(\mathbf{x})) \overset{\mathcal{F}}{\longleftrightarrow} -\mathbf{j}\omega_1 V_2(\omega) + \mathbf{j}\,\omega_2 V_1(\omega) \quad (10)$$



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

Fig. 2. Flow vectors and their CFT representation.

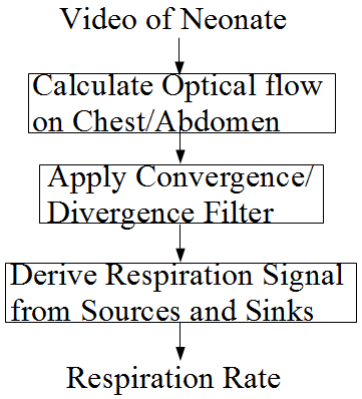


Fig. 3. Flowchart of proposed scheme

Note, that in general, $V_1(\omega)$ and $V_2(\omega)$ may be complex. We use the analysis and procedure mentioned in [10] to apply the joint convergence divergence filter in the Clifford Fourier domain, which can be easily realized with low computational complexity.

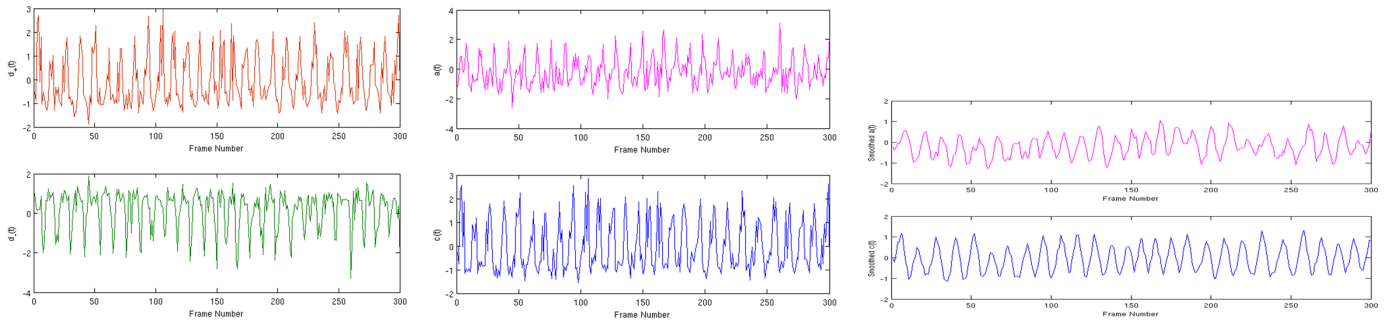
$$curl(\mathbf{v}(\mathbf{x})) + \mathbf{j}\; div(\mathbf{v}(\mathbf{x})) \overset{\mathcal{F}}{\leftrightarrow} (-\omega_1 + \mathbf{j}\omega_2) \quad (11)$$
$$(V_1(\omega) + \mathbf{j}V_2(\omega))$$
$$\overset{\mathcal{F}}{\leftrightarrow} H(\omega)Y(\omega)$$

where $H(\omega) = (-\omega_1 + \mathbf{j}\;\omega_2)$. The divergence of the field may be obtained by computing the imaginary part of the inverse Clifford Fourier transform of the filtered spectrum i.e.

$$div(\mathbf{v}(\mathbf{x})) = Im(\mathcal{F}^{-1}\{\mathbf{e_1}H(\omega)Y(\omega)\}) \quad (12)$$

III. Proposed Scheme

We process the video data of neonates to derive the respiration rate. The respiration rate can be calculated from either the



(a) (b) (c)

Fig. 4. (a) Average source ($d_+(t)$) and sink ($d_-(t)$) values. (b) Components obtained after applying ICA on $d_+(t)$ and $d_-(t)$. (c) Smoothed versions of ICA components.

chest or abdomen regions or both. In this paper, we use both the regions to calculate RR since they are representative of the respiratory motion. We detect the movement of the chest or abdomen from the video using optical flow [11]. The flow vectors are analyzed to derive the respiration rate. The method exploits the fact that the optical flow vectors as seen by a camera mounted on top of vertical column of warmer (Fig. 1) will converge and diverge with breathing motion (Fig. 2). The flow chart of the method is given in Fig. 3.

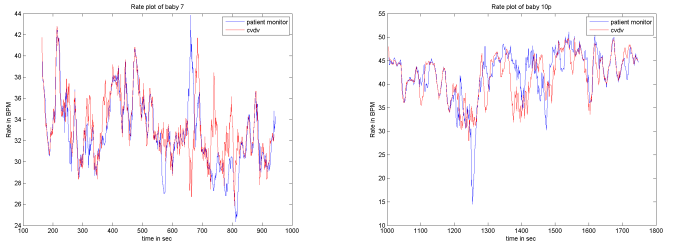
Figures 2(a) and 2(c) shows the optical flow vectors on the chest and abdomen region. It can be observed that the flow vectors converge and diverge during expiration and inspiration respectively. However, the vectors do not converge and diverge to a single point but to multiple points in the frame. It can be found that convergence and divergence is the predominant pattern but not the only pattern during these respiratory cycles. It can also be seen that the flow vectors are not of the same length. These variations are expected due to the shape of the chest and abdomen, the physiology of breathing and the orientation of neonate with respect to camera.

A question would naturally occur as to whether the convergence and divergence pattern will always be observed by the camera due to the variations in the distance of the camera and orientation of it with respect to the neonate. For the particular mount of the camera that we considered, the top of the vertical column of the warmer (Fig. 1) which suffers less from occlusions due to the care given by the clinical staff and the movement of the limbs of the neonate, convergence and divergence is the expected pattern of the flow vectors. In this position, the chest will also come closer to (farther away from) the camera during inspiration (expiration). This phenomenon is similar to zooming in/out of a camera which will also result in divergence/convergence of the flow vectors [12]. Convergence and divergence is the natural pattern resulting from the breathing of the neonate and is found to hold even when the camera is mounted at other positions as well.

We apply the convergence and divergence filter (see Sec. II-B) to enhance the pattern of the flow vectors within a frame. We are interested in the pattern of the flow vectors but not necessarily in the magnitude of the flow vectors. Hence we normalize the flow vectors to unitary magnitude before applying the filter. However, calculating divergence for all the flow vectors in every frame is a computationally intensive task and may not be feasible in real-time. To overcome this challenge, we utilize the properties of Clifford Algebra and Clifford Fourier Transform (CFT) to get a fast and accurate estimate of the convergence/divergence of the vector field.

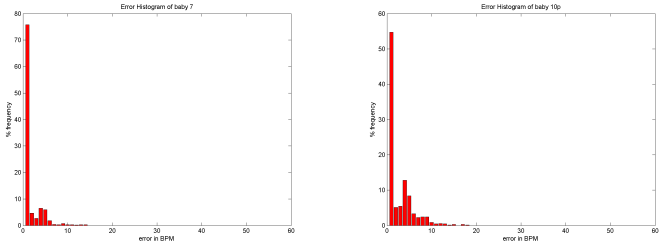
In the spatial domain, the convergence and divergence filter has to be convolved with the vector field. This translates to multiplication in the CFT domain and fast algorithms for calculating Fourier transform exist. However, the Fourier transform being discrete in nature, would need the flow vectors to be present at the specified grid locations. A fixed dense grid size would result in a dense optical flow but would be computationally intensive, without adding much of additional information. Further, forcing a tracker such as KanadeLucas-Tomasi (KLT) [13] to find matching points in frame $i+1$ for each grid point in frame i may not yield good result since they may not be the good features to track. Hence we take an approach where we find the good features points to be tracked using [14] and then assign them to the nearest grid location. This will slightly compromise the results but will not alter them significantly if the grid resolution is chosen judiciously.

Figures 2(b) and 2(d) show the output of the convergence/divergence filter applied on the flow vectors in Fig. 2(a) and Fig. 2(c) respectively. The black and white dots show the points of convergence and divergence respectively. It can be noticed that the black and white dots are predominant during expiration and inspiration respectively. The black and white dots actually correspond to values with positive and negative signs respectively and hence we take the sum of them independently to be representatives of the respiration signal from that frame. These values when plotted across time will form the respiration signal of the subject (Fig. 4(a)). It can be observed that there is a trough in the sink signal at the frame when the source signal has a peak and *vice versa*. The peaks and troughs of the source and sink signals respectively are of interest since they capture the inspiration and expiration breathing patterns of the neonate respectively. The respective troughs and peaks of the source and sink signals that occur



(a) (b)

Fig. 5. Comparison of RR for two babies in (a) supine and (b) prone positions.



(a) (b)

Fig. 6. Histograms for error in RR for two plots in Fig. 5.

at these frames are weak and this is expected. The inter-peak (trough) distance in the source (sink) signal can be used to calculate the RR. In principle, the RR can be calculated from any one of them. However, the presence of noise and/or motion of the neonate can corrupt these signals.

We use a source separation algorithm to cleanly extract the respiratory signal from the source and sink signals by isolating the noise and motion of the neonate. We use Independent Component Analysis (ICA), a powerful blind source separation method for this purpose [15]. Figure 4(b) shows the two source signals obtained when ICA is applied on the two signals in Fig. 4(a). It can be noticed that ICA has separated the respiration signal $c(t)$ and the noise $a(t)$ (Fig. 4(b)). There are some clinical conditions such as respiratory distress wherein a seesaw pattern of the raise/fall of chest and abdomen regions is observed (i.e., chest goes up while the abdomen falls and *vice versa*). In such a case, ICA will be helpful in faithfully separating the respiration signal from noise. The respiration signal will be periodic while the one riddled with noise will not be. Hence, we pick the source signal with the maximum peak in the spectrogram of the signals obtained after applying ICA.

IV. Performance Evaluation

We evaluated the performance of our scheme using several videos acquired from neonates in a NICU. The videos were recorded using a color camera UI-2220SE-C-HQ from IDS Imaging by mounting it on a Benro Digital Tripod at approximately the same height as that of the top of the vertical column of a warmer. The resolution of the videos was 576×768 pixels with a frame rate of 20 per second. The ground truth to compare the results was collected from a Philips IntelliVue MP40 patient monitor. The videos were of approximately 10 minute duration with the neonates lying in the prone and supine positions, which are the normally observed poses of neonates in a NICU setting. The methods were implemented in C++ using OpenCV [16].

We run a standard 10 point Chebyshev/Butterworth low pass filter to filter out frequencies beyond 2Hz so that only the breathing motion is extracted from the camera respiration signal (Fig. 4(c)). We run a peak detection algorithm on the filtered signal to calculate the rate. For the purpose of comparing the rate, we analyzed the peaks of signals coming from the patient monitor and our method by the observing local maxima. The peaks were observed over a buffer of 15 seconds in a sliding mode fashion, evaluated every second and the inter-peak distance was used to calculate the rate.

Figures 5(a) and 5(b) show the rates from the patient monitor and that from our method. We observe that the rates from our algorithm match closely with those from the patient monitor. To achieve a comparable performance, both respiration signals were processed using identical filtering schemes using the same parameters for rate calculation. To evaluate the results quantitatively, we calculate the error in rates between our method and that from a patient monitor. Figures 6(a) and 6(b) show the error histograms for the plots in Fig. 5(a) and Fig. 5(b) respectively. It can be observed that a significant majority of the rate values match. Hence the method proposed by us provides RR that matches that coming from a patient monitor.

V. Conclusions

We propose a novel method to derive respiration rate of neonates by using a camera. We estimate the motion of chest and abdomen region using optical flow and process the flow vectors to compute the respiration signal. We exploit the natural pattern of the flow vectors that results in convergence and divergence of flow vectors during breathing. We separate the noise and the respiration signal by using ICA. Evaluation of our scheme on several videos taken from neonates shows that we match the respiration rate from a patient monitor faithfully in both prone and supine positions. Hence our method shows promise for use in clinical settings.

Acknowledgement

The authors would like to thank clinical partner Dr. Ranjan Kumar Pejaver of Arrow Medical Information Services, Bangalore, India for providing the data.

References

[1] V. Makkapati and S. S. Rambhatla, "Camera based estimation of respiration rate by analyzing shape and size variation of structured light," in *International Conference on Acoustics, Speech, and Signal Processing*. IEEE, March 2016.

[2] B. XU, L. K. MESTHA, and G. PENNINGTON, "Monitoring respiration with a thermal imaging system," U.S. Patent 20 120 289 850, May 9, 2011.

[3] L. K. MESTHA, E. A. BERNAL, and B. XU, "Processing a video for respiration rate estimation," U.S. Patent 20 130 324 875, June 21, 2012.
[4] M. Martinez and R. Stiefelhagen, "Breath rate monitoring during sleep using near-ir imagery and pca," in *Pattern Recognition (ICPR), 2012 21st International Conference on*, Nov 2012, pp. 3472–3475.
[5] F. , M. Li, Y. Qian, and J. Z. Tsien, "Remote measurements of heart and respiration rates for telemedicine," *PLoS ONE*, vol. 8, no. 10, p. e71384, 10 2013.
[6] M. Bartula, T. Tigges, and J. Muehlsteff, "Camera-based system for contactless monitoring of respiration," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE, 2013, pp. 2672–2675.
[7] A. Venkitaraman and V. Makkapati, "Motion-based segmentation of chest and abdomen region of neonates from videos," in *International Conference on Advances in Pattern Recognition*, Jan 2015, pp. 1–5.
[8] D. Garling, *Clifford algebras: an introduction*. Cambridge University Press, 2011, vol. 78.
[9] J. Ebling and G. Scheuermann, "Clifford fourier transform on vector fields," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 11, no. 4, pp. 469–479, 2005.
[10] H. Mohammadzade and L. T. Bruton, "A simultaneous div-curl 2d clifford fourier transform filter for enhancing vortices, sinks and sources in sampled 2d vector field images," in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*. IEEE, 2007, pp. 821–824.
[11] S. S. Beauchemin and J. L. Barron, "The computation of optical flow," *ACM Computing Surveys*, vol. 27, no. 3, pp. 433 – 466, September 1995.
[12] V. Makkapati, "Camera pan and zoom change detection using optical flow," in *National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*, January 2008, pp. 73–78.
[13] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: a factorization method," *International Journal of Computer Vision*, vol. 9, no. 2, pp. 137–154, 1992.
[14] J. Shi and C. Tomasi, "Good features to track," in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*. IEEE, 1994, pp. 593–600.
[15] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4, pp. 411–430, 2000.
[16] "Open source computer vision," http://opencv.org.