Automatic classification of resuscitation activities on birth-asphyxiated newborns using acceleration and ECG signals

Huyen Vu, a,*, Kjersti Engan a, Trygve Eftestøl a, Aggelos Katsaggelos b, Samwel Jatosh c, Simeon Kusulla c, Estomih Mduua c, Hussein Kidanto d, Hege Ersdal e, f

a Department of Electrical Engineering and Computer Science, University of Stavanger, Stavanger, Norway
b Department of Electrical Engineering and Computer Science, Northwestern University, IL, United States
c Research Institute, Haydom Lutheran Hospital, Haydom, Manyara, Tanzania
d Department of Obstetrics and Gynaecology, Muhimbili National Hospital/MUHAS, Dar Es Salaam, Tanzania
e Department of Anaesthesiology and Intensive Care, Stavanger University Hospital, Norway
f Department of Health Sciences, University of Stavanger, Stavanger, Norway

ARTICLE INFO

Article history:
Received 23 October 2016
Received in revised form 11 February 2017
Accepted 19 March 2017

Keywords:
Newborns
Birth asphyxia
Acceleration signal
ECG signal
Short time energy
Wavelet packet decomposition
Linear discriminant analysis
Decision tree

ABSTRACT

Objectives: Newborn deaths are reported to be caused mainly by birth asphyxia. Information learned from ventilation and other treatment could help increase survival rate of newborns in need of resuscitation. Characteristics of manual bag-mask ventilation have been studied in our previous works. However, other resuscitation activities could have important impacts as well. This paper illustrates the classification of several predefined resuscitation activities using information from acceleration and ECG signal.

Methods: Time and frequency domain features were extracted from the acceleration and ECG signals. A 2-stage classifier was trained on data of manually annotated activities by observing videos of 30 resuscitation babies. Leave-one-out cross validation was used: for each fold, the classifier was trained on activities of 29 patients and tested on activities of 1 patient.

Results: The average accuracy of the classification of activities is 79%.

Conclusions: The performance of the classification algorithms indicates that it is possible to use ECG and acceleration signals to automatically derive useful information regarding resuscitation activities.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Complication at birth and/or birth asphyxia lead to the deaths of nearly a quarter of the estimated 3 million newborn deaths according to a recent report from the Save the Children organization [1,2]. Birth asphyxia is usually referred to as a failure to initiate spontaneous breathing and/or to have a 5-min Apgar score <7. Initiation of positive pressure ventilation within the first minute of life – “The Golden Minute”, is recommended as a treatment to increase survival chance. The delay in executing basic resuscitation might result in a progressive decrease in heart rate, blood pressure, brain injury or even death [3]. Guidelines for neonatal resuscitation are currently applied in practical use [4,5]. However, the key factors of effective treatment remain unrevealed.

Safer Births is a research collaboration between Haydom Lutheran Hospital in Tanzania and Stavanger University Hospital, University of Stavanger, and Laerdal Global Health in Norway. The mission of Safer Births is to establish new knowledge and develop innovative products in order to provide better equipment and improve competence of health workers. The project has collected resuscitation data of more than 500 newborns, using sensors measuring various signals: flow, pressure, CO2, ECG, and acceleration signals.

The detection and parameterization of events characterizing manual bag-mask ventilation have been described in our previous works [6]. We have also investigated the possible correlation between ventilation parameters and different patient groups by developing an exploratory analysis framework [7–9]. The results demonstrate that the interaction between clinical treatment and human physiological conditions are highly complicated. Ventilation is important yet not the only factor. During ventilation, there

* Corresponding author.
E-mail addresses: vtphuyen1987@yahoo.com, vu.huyen@uis.no (H. Vu).
1 www.savethechildren.org.

http://dx.doi.org/10.1016/j.bspc.2017.03.004
1746-8094/© 2017 Elsevier Ltd. All rights reserved.
could be other resuscitation activities applied simultaneously, such
as drying thoroughly, stimulation, suction, chest compression, etc.,
and those activities could have important impacts on the responses
of the babies.

Activity recognition based on accelerometer measurements has
been increasingly investigated during the past decades [10–16]. An
accelerometer is a device measuring acceleration forces in three
orthogonal directions (X, Y and Z axes). Wearable devices (e.g.,
smart phones, health status monitors, etc.) have made use of inte-
grated accelerometer sensors in detecting and classifying daily
human activities efficiently. Various features, e.g., features in time
domain [10–13] or frequency domain [15] or both [16,14], were
extracted from acceleration signals measured by the devices worn
tightly on the body and the activities typically categorized were
running, walking, standing, sitting, and jumping.

To identify activities having significant impacts on the improve-
ment of asphyxiated newborns during resuscitation, it is necessary
to study videos recorded in the delivery room [17,18]. However,
in practical situations, video recordings are not likely to be avail-
able, and video reviewing is very time consuming. This paper
illustrates the classification of several predefined resuscitation
activities using information extracted from acceleration and ECG
signals. The objective of this paper is to investigate the possibility
of studying the activities automatically. The ultimate aim is to de-
velop an automatic system to support clinicians in retrieving information
about activities not only more quickly, but also when videos are
not available. Fig. 1 shows an overview of the activity classification
system.

2. Dataset

Laerdal Global Health3 in Stavanger, Norway has developed the
Laerdal Newborn Resuscitation Monitor (LNRM) for research in a
limited resource setting. There are various types of information
measured by LNRM: the ventilation (airway pressure, flow, CO2),
ECG and acceleration signals. In this paper, we used information
from the ECG and accelerometer sensors. Fig. 2 shows the ECG sen-
 sor placed gently over the baby's abdomen. The acceleration signal
is measured by an accelerometer attached to the ECG-sensor. The
ECG and acceleration signals are sampled at 500 Hz and 100 Hz,
respectively. All sensor signals are synchronized. The signal data is
saved in a memory unit integrated in LNRM and can be transmitted
to a computer via a USB port.

In this paper, the data includes physiological signals and videos of
30 patients recorded at the Haydom Lutheran Hospital from
October 2013 to January 2014. The videos were manually annotated
using the ELAN4 tool to provide ground truth labels used for training
and testing the automatic classification of activities. A medical doc-
tor defined the resuscitation activity labels, and instructed how to
distinguish the activities when studying the resuscitation videos.
The annotations in this feasibility study were performed by the
main author after receiving such instructions.

Defined by clinicians, the manual annotations include seven
different types of resuscitation activities: chest compression (the
birth attendant places fingers on the chest of newborns to per-
form heart compression), back stimulation (the birth attendant is
moving a hand up and down the back to stimulate the column
and spinal cord), tactile stimulation (usually a more gentle stimula-
tion on chest or feet), drying thoroughly (the birth attendant uses a
cloth to wipe the baby), moving baby, moving ECG-sensor, and uncat-
ergorized movements (which also affect the ECG and acceleration
signal). The manual annotations may contain overlapped activi-
ties. Due to the similarity and therefore difficulty in distinguishing,
we exclude sequences of activities happening simultaneously. In
addition, we merge some classes into one group. Specifically, back
stimulation, tactile stimulation and drying thoroughly are grouped


as “stimulation” class. Moving baby, moving ECG-sensor, and uncategorized movements are grouped as “other”. And “chest compression” remains a separate class. For each patient, there could be sequences of different activities. The whole data set contains 446 annotated activity sequences with 21 sequences of chest compression (from 2 patients), 250 sequences of stimulation (from 30 patients) and 175 sequences of other (from 30 patients).

3. Method

3.1. Pre-processing

In the pre-processing step, the ECG signal was filtered using notch filter at 50 Hz to remove noise by power-line interference.

The acceleration signal is sampled at 100 Hz and consists of 3 orthogonal axes $A_x, A_y, A_z$. To isolate the low frequency component (due to the force of gravity) from the high frequency (caused by the body acceleration), an ILR low-pass filter is used as a pre-processing step, given by the difference equation:

$$a_1 \cdot A_{LP}(n) = b_1 \cdot A(n) - a_2 \cdot A_{LP}(n - 1)$$

where $A(n)$ and $A_{LP}(n)$ respectively denote the input and output of the filter and $a_1, a_2$ and $b_1$ are the filter coefficients.

In our experiment, the chosen cutoff frequency is 0.1 Hz [13]. The resulting filter coefficients are: $a_1 = 1$, $b_1 = 0.0063$, $a_2 = -0.9937$ [19]. The body acceleration is calculated by subtracting the low pass filtered data in each axis $(A_{LPx}, A_{LPy}, A_{Lpz})$ from the original signal:

$$A_x(n) = A_{original}(n) - A_{LPx}(n)$$

$$A_y(n) = A_{original}(n) - A_{LPy}(n)$$

$$A_z(n) = A_{original}(n) - A_{Lpz}(n)$$

The magnitude of the acceleration signal is computed as:

$$A(n) = \sqrt{A_x^2(n) + A_y^2(n) + A_z^2(n)}$$

3.2. Feature extraction

The accelerometer and ECG sensors attached on the baby’s abdomen provide information relating to performed resuscitation activities. Frequency and time domain features extracted from the acceleration signal have been used in previous research with promising results in recognizing daily activities of adults, e.g., jumping, running, standing, and walking [10–16]. Thus, in this work, we employed frequency and time domain features extracted from both acceleration and ECG signals, giving totally 46 features as the inputs for classification.

3.2.1. Time domain features

In time domain signals, we derive the following features:

- For $ith$ activity, the mean $(\mu^i)$, standard deviation $(\sigma^i)$, entropy $(H^i)$ and root mean square $(rms^i)$ features are computed as follows:

  From the acceleration signal:

  $$\mu^i = [\text{mean}(A^i(n)) \text{ mean}(A_x^i(n)) \text{ mean}(A_y^i(n)) \text{ mean}(A_z^i(n))]$$

  $$\sigma^i = [\sigma(A^i(n)) \sigma(A_x^i(n)) \sigma(A_y^i(n)) \sigma(A_z^i(n))]$$

  $$H^i = [H(A^i(n)) H(A_x^i(n)) H(A_y^i(n)) H(A_z^i(n))]$$

  $$\text{rms}^i = [\text{rms}(A^i(n)) \text{ rms}(A_x^i(n)) \text{ rms}(A_y^i(n)) \text{ rms}(A_z^i(n))]$$

  where $A^i(n)$ is the segment of the acceleration signal corresponding to the $ith$ activity. $0 < n < n_i$ is the activity duration. Each activity has different duration or $n_i \neq n_j$, where $i$ and $j$ are two different activities as illustrated in Fig. 3.

- Minimum value of the short time energy: $E_{\text{min}}^i$ which is the minimum value of short time energy signal of acceleration signal, where $n_{\text{start}}^i \leq n \leq n_{\text{stop}}^i$ are the indices of the signal where the activity starts and ends.

  $$E_{\text{min}}^i = \min(E_A(n)) \text{ for } n_{\text{start}}^i \leq n \leq n_{\text{stop}}^i$$

  From the ECG signal:

  $$\mu^i = \text{mean}(ECG^i(n))$$

  $$\sigma^i = \sigma(ECG^i(n))$$

  $$H^i = H(ECG^i(n))$$

  $$\text{rms}^i = \text{rms}(ECG^i(n))$$

  where ECG$^i(n)$ is the segment of the ECG signal corresponding to the $ith$ activity.

- The cross correlation pairs:

  $$\{A_x^i(n), A_x^j(n)\}, \{A_y^i(n), A_y^j(n)\} \text{ and } \{A_z^i(n), A_z^j(n)\}.$$  

  $$\text{corr}^i = [\text{corr}(A_x^i(n), A_x^j(n)) \text{ corr}(A_y^i(n), A_y^j(n)) \text{ corr}(A_z^i(n), A_z^j(n))]$$

- From both magnitude acceleration $A(n)$ and ECG$(n)$ signals:

  - Maximum value of the signal magnitude: $\max(|A(n)|)$, and $\max(|ECG(n)|)$.
  - Total short time energy: $\text{TSE}_A(n)$ and $\text{TSE}_{ECG}(n)$.
  - Maximum values of the short time energy: $\max(E_A(n))$, and $\max(E_{ECG}(n))$.
  - Energy of the auto-correlation signal: $\text{autocorr}(A(n))$, and $\text{autocorr}(ECG(n))$. 

Fig. 3. $N_i$ and $N_j$ are the lengths of manually annotated activity $i$ and $j$ respectively.
3.2.2. Frequency domain features

The 6-level wavelet decomposition is applied to the ECG and A signals. A Daubechies mother wavelet function is used in our experiment. The percentage of energy corresponding to the approximation (A6) and detail coefficients at each level are extracted as features in the frequency domain (D1, D2, D3, D4, D5, D6) [20,21].

3.3. Classification

The input features extracted in the preprocessing step are normalized (Z-score normalization) before feeding into the classifiers.

Chest compression is the activity when the health care provider uses two fingers and compress repeatedly on the chest. The magnitude of the acceleration signal is quite small. During stimulation or other activities (e.g., moving the baby or sensor), the energy of the acceleration signal (\(TSTE_{a(n)}\)), and sometimes of the ECG signal (\(TSTE_{ECC(n)}\)), would usually be quite large. Thus, \(TSTE_{a(n)}\) and \(TSTE_{ECC(n)}\) are the two most discriminative features. This information is illustrated in Fig. 4 and might indicate that the chest compression class is distinguishable from the stimulation and other classes. Thus, we design two stages of classification. In the first stage, the chest compression class is separated from the stimulation and other classes. In the second stage, the stimulation class is differentiated from the other class. The scheme of the two-stage classification is shown in Fig. 5.

The computation of the confusion matrices at each stage and after two stages is illustrated as follows. In stage 1, the classification result is presented by the confusion matrix depicted in Table 1(a). Activities classified as stimulation or other (b + d activities) go through to the second stage of classification, whereas misclassified stimulation and other (c activities) do not, as shown in Table 1(b). The final confusion matrix of all activities after two stages of classification is illustrated in Table 1(c). The final result in Table 1 is based on the following calculation: \(e + f = b, k + l = c, k + g + h =\) number of stimulation sequences, \(l + i + j =\) number of other sequences.

The classification accuracy is calculated as the ratio of the number of correctly classified activities over the total number of activities:

\[
\text{accuracy} = \frac{a + g + j}{a + e + f + k + g + h + l + i + j}
\]  

4. Experiment

Leave-one-out cross validation: The training and testing are repeated 30 times in this manner: resuscitation activities manually
annotated from 29 videos are used for training and the classifier is tested on the activities from the other video.

Classifiers: We experimented with several different supervised learning classifiers in the Matlab Machine Learning Toolbox: Decision tree, TreeBagger, Random Forest, Naive Bayes, Support Vector Machine, Discriminant Analysis, K_Nearest Neighbor and Neural Network [22].

We applied these classifiers to a 2-stage classification strategy and chose the classifiers giving the best classification accuracy, compared to the manual annotations, as the final classifiers for our system.

All combinations of these 8 different classifiers in the previous explained 2-stage classification strategy was tested giving $8 \times 7=56$ combinations. Because the number of combinations of the classifiers is large, we chose to only present the best result. The combination of classifiers and parameters giving the best classification accuracy, compared to the manual annotations, was chosen as the final proposed classifier strategy for our system. This gives the non-parametric decision tree classifier in the first stage and the linear discriminant analysis classifier in the second stage.

Because the number of samples in each class is imbalanced, the cost-sensitive parameter was used to specify the cost of misclassification of each class.

For the decision tree classifier in the first stage, the misclassification cost matrix was depicted in Table 2. For the discriminant analysis classifier in the second stage, the cost-sensitive parameters were also chosen as in Table 3.

The values of cost-sensitive classification were based on experiments with various values.

5. Results

The presented results are obtained by using all the features listed in Section 3.2 as input features. After using different classifiers, the highest classification accuracy is given by the decision tree classifier in the first stage and the discriminant analysis classifier in the second stage. The confusion matrices of the 2-stage classification are shown in Table 4.

The accuracy of the classification with 3 classes is computed as follows:

$$\text{accuracy} = \frac{16 + 208 + 127}{446} = 78.7\%$$ (7)

Chest compression and stimulation could both be considered as treatment whereas other as a non-treatment activity. Therefore, we could merge the confusion matrices of the 3 classes into 2 classes treatment and non-treatment as shown in Table 5. The accuracy of the classification is then calculated as:

$$\text{accuracy} = \frac{229 + 127}{446} = 79.8\%$$ (8)

6. Discussion

A method for automatic classification of activities during newborn resuscitation using the acceleration and ECG signals has been proposed in this paper. The objectives are not only to assist the clinicians in reviewing the videos but also to allow the possibility to investigate resuscitation activities when there is no video available.

We have manually selected the list of features. Some experiments with reduced feature sets were conducted, but the best results were achieved using the full feature set.

Decision tree is a good choice for imbalanced dataset because the splitting rules using the class variable in the creation of the trees can force both classes to be addressed.

6.1. Clinical implication

Using various classifiers, the best results given for the 3-class problem (chest compression, stimulation and other) is an accuracy of 79% and for the 2 class problem (treatment and non-treatment) is 80%, approximately. The classification could be potentially improved, however, the results indicate that it is feasible to extract useful information about the activities by only using the acceleration and ECG signals. Such information could be helpful to develop a supporting tool for interpretation and annotation of video record-
ings, or for classification of activities when no video recordings are available.

### 6.2. Limitation

The number of chest compression sequences is remarkably small compared to the two other types of activity. This class imbalance is a known disadvantage for many classifiers [23]. In addition, the lack of data for the chest compression class might also be a problem. We have manually selected different features by trial and error method in each experiment. For example, removing features from ECG or acceleration signal, or using only time or frequency domain features. The results were not better than the set of features we presented in this paper. However, the dimension of the features vector is rather high, thus, the classification results could be affected by over-fitting. The linear discriminant analysis model assumes that the feature vectors have Gaussian mixture distributions. Thus, the results need to be validated with a larger and more balanced data set.

Additionally, the features are extracted from the segments of the acceleration and ECG signals corresponding to the manual annotation of activities. However, for later use of activity classification, the activities have to be detected automatically in order to provide the segments of the corresponding signals to extract features used as inputs to the classifiers.

### 7. Conclusions

Birth asphyxia is leading to a great number of newborn deaths every year world wide, and is especially a challenging issue in developing countries. Understanding the characteristics of bag-mask ventilation and other resuscitation activities might help find factors of effective intervention within the first minute after birth to reduce neonatal mortality worldwide.

### 7.1. Contribution

Reviewing videos of resuscitation in a delivery room is a tremendously time consuming task. In this work, we propose an approach to automatically classify the manually annotated activities based on information extracted from ECG and acceleration signals. This automatic classification would ease the work of clinicians when observing videos as well as providing some indications about resuscitation activities if no video is available.

### 7.2. Future work

Integrating information from other types of signals or implementing more sophisticated feature selection algorithms and classifiers could provide a higher and more reliable classification accuracy. In another work, we have described a method to detect the presence of activities [24]. The combination of the detection and classification of activities will be the next step in developing our system to support health care workers in making decisions. A proper feature selection test could be done to investigate if it is possible to reduce the feature set.

In an ongoing projects, experts are manually annotating an extended data set. This will be used in the further development and testing of the classification system.

### References


### Table 4
Confusion matrices of (a) first stage using decision tree, (b) second stage using discriminant analysis, (c) final result using combination of the two classifiers.

<table>
<thead>
<tr>
<th>Actual class</th>
<th>Chest compression</th>
<th>Stimulation &amp; other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest compression</td>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>Stimulation &amp; other</td>
<td>3</td>
<td>422</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Actual class</th>
<th>Chest compression</th>
<th>Stimulation</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest compression</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Stimulation</td>
<td>0</td>
<td>208</td>
<td>42</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>45</td>
<td>127</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Actual class</th>
<th>Chest compression</th>
<th>Stimulation</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest compression</td>
<td>16</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Stimulation</td>
<td>0</td>
<td>208</td>
<td>42</td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
<td>45</td>
<td>127</td>
</tr>
</tbody>
</table>

### Table 5

<table>
<thead>
<tr>
<th>Actual class</th>
<th>Treatment</th>
<th>Non-treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>16+5+208=229</td>
<td>42</td>
</tr>
<tr>
<td>Non-treatment</td>
<td>3+45=48</td>
<td>127</td>
</tr>
</tbody>
</table>