**Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006**

An Adaptive Cusum Test Based on a Hidden Semi-Markov Model for Change Detection in Non-invasive Mean Blood Pressure Trend

Ping Yang* and Guy Dumont* and J Mark Ansermino**

Abstract— **A Hidden Semi-Markov Model is proposed to describe a trend signal for non-invasive mean blood pressure. Based on the model, a Viterbi beam search algorithm working with an adaptive Cumulative Sum test is proposed to detect change points and recognize change patterns online. Testing results on the simulated signals and clinical signals demonstrate that the algorithm has improved performance over the standard Cumulative Sum test for change detection, and has the potential to provide a clinically relevant description of trend changes.**

I. INTRODUCTION

Blood pressure is measured during anesthesia either by direct cannulation of a peripheral artery or non-invasive occlusion of the brachial artery with a cuff. During automated non-invasive oscillotonometric blood pressure measurement, the systolic pressure is first detected by slowly releasing the cuff pressure until a distal pulse occurs. Diastolic and mean blood pressures are then measured by processing the turbulence in the artery distal to the cuff. A measurement cycle takes 30 seconds. However, since frequent inflation leads to congestion of the arm, non-invasive blood pressure is usually measured every 3 to 5 minutes during surgery.

Anesthesiologists perform multiple tasks and cannot maintain continuous observation of the monitoring system. Alarm systems commonly test the non-invasive mean blood pressure (NIBPmean) against fixed thresholds, which do not adjust to the patient's current state, resulting in poor sensitivity and a high false alarm rate. We have previously described an algorithm for detecting trend changes in respiratory variables [1]; however, it is not suitable for the infrequent measurements seen with NIBPmean.

Changes in the NIBPmean trend have large amplitudes or sustained durations, whereas disturbances caused by physiological variations and movements are relatively small or short. The Cusum test is therefore an intuitive choice for change point detection.

Changes in the NIBPmean trend are characterized by different patterns. In Fig. 1(a), segments might be described as *stable*, *short abrupt increase* and *short gradual decrease*. Similar patterns have different significance in different contexts. For instance, a change of similar amplitude and duration to the significant change in Fig. 1(a) might be insignificant in Fig. 1(b), given the previous *long abrupt decrease*. The standard Cusum test uses fixed criteria for evaluating variations and would therefore introduce a false change point in this situation.

Department of Electrical and Computer Engineering, the University of British Columbia, Vancouver BC, CANADA*

Anesthesiology, Pharmacology & Therapeutics Department, the University of British Columbia, Vancouver BC, CANADA**

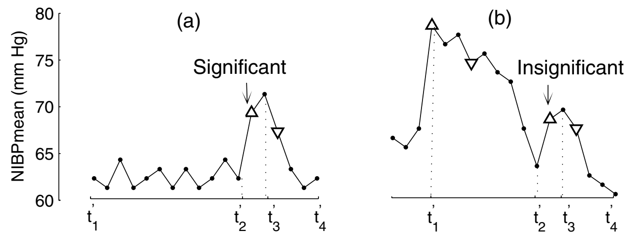

Fig. 1. Two example trends: similar patterns have different significance

TABLE I
SEGMENTAL STATES OF NIBPMEAN

A	Increase				Decrease				Stable
B	Abrupt		Gradual		Abrupt		Gradual		
C	lg	sht	lg	sht	lg	sht	lg	sht	
States	S^1	S^2	S^3	S^4	S^5	S^6	S^7	S^8	S^9

sht – short, lg – long

To increase the performance of a Cusum test, our proposed algorithm imitates the detection process adopted by anesthesiologists by recognizing change patterns and then using expert knowledge about the transition relationship between these patterns to evaluate upcoming data.

II. METHODS

A. Signal Model: Segmental Hidden Semi-Markov Model

The segmental states of NIBPmean trend are hierarchically classified into nine types according to their durations and shapes (see Table I). We assume that the likelihood of occurrence for each segmental state depends only on the previous segment, independent of other historical states. The trend signal can therefore be modeled by a first-order segmental Hidden Semi-Markov Model (HSMM).

The segmental Hidden Semi-Markov Model is an extension of the Hidden Markov Model (HMM) [2]. As in the HMM, transitions between segmental states in the HSMM follow a first-order Markov process. Each element in the state transition matrix A_{ji} specifies the conditional probability of changing from one state S^j to another S^i. However, unlike the HMM, where one signal state generates one measurement, each segmental state in the HSMM corresponds to a series of measurements. For example, the first segment in Fig. 1(a) is composed of measurements over $t_1^{'}... t_2^{'}$.

The segmental state s_k is not directly observable. It only determines the pattern parameters that guide the shape and duration of the signal within the segment. True signal values x then vary according to the pattern, and eventually generate measurements y. The process is modeled as:

$$\begin{aligned}&\underline{Segmental\ State}: s_k^i, t_k^{'}\\&\underline{Duration}:\ d_k^i = t_k^{'} - t_{k-1}^{'}\\&for\ t = t_{k-1}^{'}+1...t_k^{'}\\&\underline{Predictions}: e_t = \lambda e_{t-1} + (1-\lambda) y_{t-1}\\&\underline{Shape}: x_t = \begin{cases} min(e_t+\theta_k^i, limit_k^i),\ s_k^i \in \{S^1 \dots S^4\}\\ max(e_t+\theta_k^i, limit_k^i),\ s_k^i \in \{S^5 \dots S^8\}\\ e_t, \quad s_k^i = S^9\end{cases}\\&\underline{Measurements}: y_t = x_t + w\end{aligned} \quad (1)$$

In (1), s_k^i indicates the state of the kth segment is S^i. The distributions of the duration parameter d^i, the shape parameter θ^i and the $limit^i$ are all conditional on the state S^i, and are independent of each other. The additive white noise w follows a constant truncated Gaussian distribution.

The shape parameter θ^i is the magnitude by which true signal values vary from the predictions at one step, i.e., θ^i influences the change speed; $limit^i$ determines the amplitude. For abrupt changes, θ^i and $limit^i$ tend to distribute around larger mean values than gradual changes (see Fig. 2). For the duration parameters, it is natural to assume the segment of a long duration has a larger d^i (see Section II-D).

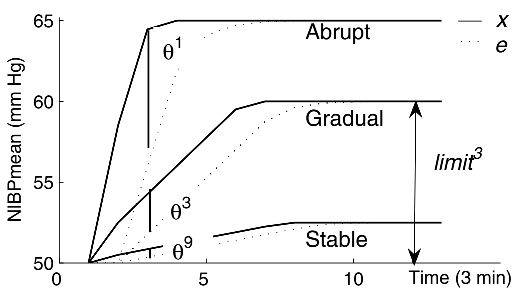

Fig. 2. Segmental shapes resulting from parameters of different values

With the signal model, the problem of change point detection and pattern recognition is translated to searching for a sequence of $(s, t^{'})$ given the measurements $y_{1...t}$.

B. Offline Solution: Extended Viterbi Algorithm

The extended Viterbi algorithm finds the segmental state sequence that has the maximum posterior probability. For each possible segmental state at time t, conditional on $y_{1...t}$, the algorithm calculates the probability of the most likely path leading to the node (s_k^i, $t_k^{'}$=t). A likelihood function $\mathbf{P}(i,t)$ is defined as [3]:

$$\mathbf{P}(i,t) = \max_{(s,t^{'})_{1...k-1}} P\{(s,t^{'})_{1...k-1}, y_{1...t} | s_k = i, t_k^{'} = t\} \quad (2)$$

The magnitude of $\mathbf{P}$ is determined by the degree of measurements fitting the shape description (Ps), the transition probability from the last segment (Pt), and the likelihood of the segment having a corresponding duration (Pd). The recursive updating process of $\mathbf{P}(i,t)$ is:

$$\begin{cases} Ps(i,t,t_{k-1}^{'}) = P\{y_{(t_{k-1}^{'}+1)...t} | \theta^i, limit^i\}\\ Pt(i,j,t_{k-1}^{'}) = \mathbf{P}(j, t_{k-1}^{'}) A_{ji}\\ Pd(i,t,t_{k-1}^{'}) = P\{d_k = t - t_{k-1}^{'} | d^i\}\\ \mathbf{P}(i,t) = \max_{(s_{k-1}^j, t_{k-1}^{'})}\{PdPtPs\}\end{cases} \quad (3)$$

The end point $t_{k-1}^{'}$ and state s_{k-1}^j of the previous segment that maximize $\mathbf{P}(i,t)$ are also recorded until the end of signal T. Starting from $argmax_i \mathbf{P}(i,T)$ and tracking back through these records, we can find the most likely state sequence.

The procedure of backward tracking not only results in extensive computation, but also makes this algorithm impractical for online change detection. A stopping rule is needed to decide when to start tracking backward and initiate a new segment.

C. Online Solution: Beam Search with Cusum Pruning

Beam search is widely used with the standard Viterbi algorithm and the HMM. It uses a heuristic function to estimate the promise of each node, and similar to how we search around a dark room following a flashlight beam, it only searches the promising nodes included in the beam [4].

The same idea is used in our algorithm. By evaluating the Cusum against the testing mask, a Cusum test determines whether a segment of a different state might have occurred, and tells of the possible new states. The Cusum test is therefore used to stop updating $\mathbf{P}(i,t)$, and define the conditions that nodes $(s_k, t_k^{'})$ have to fulfill to be included in the beam. The following process is carried out at time t:

1. The Cusum of prediction residuals is checked against the mask for the historical data belonging to the $beam_t$. A broken lower arm determines that t is an increasing change point, and $\{S^1 ... S^4\}$ are the states in the new beam $beam_{t+1}$; a broken upper arm determines a decreasing change point, and $\{S^5 ... S^8\}$ are the states in the new beam; an intact mask beyond the end of the mask suggests the segmental state is almost stable, and $\{S^3, S^7, S^9\}$ are the states in the new beam; other situations suggest no change has occurred, and the states in the old beam are copied to the new beam.

2. If no change is detected, t is added in the beam as a possible end point of the current segment. If a change point is detected, we search the old beam $beam_t$ to find the state and end point of the current segment $(s_k, t_k^{'})$ that maximizes the $\mathbf{P}$ function of the most likely states in the new beam as in (4); meanwhile, a new segment is initiated. $t_k^{'}+1...t$ are the possible end points for the new segment and therefore become the end points in the new beam.

$$(s_k, t_k^{'}) = \underset{(s_k^j, t_k^{'}) \in beam_t}{argmax}\{\max_{i \in beam_{t+1}} \mathbf{P}(i,t)\} \quad (4)$$

3. $\mathbf{P}(i, t+1)$ is updated as in (5) for the next iteration. The Cusum mask is also adapted to the recognized segment and the states in the new beam.

$$\mathbf{P}(i, t+1) = \max_{(s^j, t^{'}) \in beam_{t+1}}\{PdPtPs\} \quad (5)$$

The CUSUM testing mask is designed according to the signal model. The shape model in (1) indicates that before reaching the $limit$, prediction residuals follow:

$$res_t = y_t - e_t = w + \theta_i \quad (6)$$

For different change patterns, mean shifts of res_t have different values (see Fig. 2). Our mask design therefore aims at different shift values.

For instance, the targeted states of the upper arm are decreases $\{S^5...S^8\}$ and stable S^9 (see Fig. 3). Since S^5 is a sustained S^6, and S^7 a sustained S^8, only S^6, S^8 and S^9 are used. A standard arm is generated for each state by setting the slope as half of the mean (for k_6 and k_8) or half of the maximum value (for k_9) of the corresponding change speed θ, according to the design criteria proposed in [5]. The tail of each mask is flattened due to the constriction of $limit$. A small threshold h helps the mask capture change points quickly, whereas a large h lowers the false alarm rate. Anesthesiologists' experience in evaluating the significance of each state is adopted in dealing with the tradeoff.

As the duration ranges for $\{S^5...S^8\}$ and S^9 are connected head to tail with some overlapping, the outline of these arms is used as the Cusum testing mask. The mask can be further smoothed by averaging over the overlapping areas.

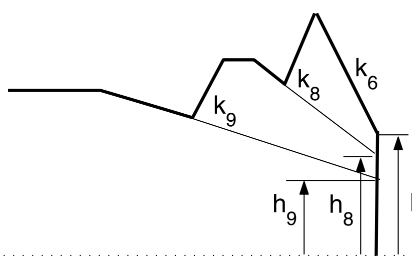

Fig. 3. Mask design of the upper arm

The mask generated above is only optimal for the initial iteration since anesthesiologists' knowledge about the significance of each state is based on an initial prior probability p_0 for each state. As mentioned in the beam search process, the prior probability p^i_{t+1} of the upcoming segment being S^i is conditional on the recognized $(s_k, t_k^{'})$ and the new possible states in $beam_{t+1}$, and updated as below:

$$p^i_{t+1} = \mathbf{I}\{\sum_{j \in beam_{t+1}} A_{s_k,j} Ps(j,t,t_k^{'}) Pd(j,t,t_k^{'}) A_{j,i}\} \qquad (7)$$

$\mathbf{I}$ means normalizing. h^i is multiplied by an adapting parameter $\lambda^i_{t+1} = 1 + \log(p^i_{t+1}/p^i_0)$, so that each part of the mask gets updated online.

D. Parameter Setting

An experienced anesthesiologist who is involved in this study and familiar with the concept of this algorithm was interviewed with a clinician-friendly questionnaire. The knowledge collected was used to generate the transition matrix and the distribution of each parameter. Below $\overline{n}(a,b,c,d)$ denotes a Gaussian distribution $N(a,b)$ truncated beyond $[c,d]$.

$$A = \begin{bmatrix} .100 & .100 & .050 & .050 & .200 & .050 & .200 & .050 & .200 \\ .050 & .050 & .0 & .0 & .250 & .200 & .250 & .200 & .0 \\ .150 & .100 & .0 & .0 & .225 & .100 & .225 & .050 & .150 \\ .100 & .100 & .0 & .0 & .225 & .200 & .200 & .175 & .0 \\ .200 & .050 & .200 & .050 & .100 & .100 & .050 & .050 & .200 \\ .250 & .200 & .250 & .200 & .050 & .050 & .0 & .0 & .0 \\ .225 & .100 & .225 & .050 & .150 & .100 & .0 & .0 & .150 \\ .225 & .200 & .200 & .175 & .100 & .100 & .0 & .0 & .0 \\ .125 & .125 & .125 & .125 & .125 & .125 & .125 & .125 & .0 \end{bmatrix}$$

Noise in a unit of 1 mm Hg: $w \sim \overline{n}(0, 6, -9, 9)$

Duration d^i in a unit of three minutes:

$d^1 \sim \overline{n}(8,4,4,8)$ $d^2 \sim \overline{n}(3,3,2,4)$ $d^3 \sim \overline{n}(11,4,5,11)$
$d^4 \sim (4,3,3,5)$ $d^5 \sim \overline{n}(9,4,5,9)$ $d^6 \sim \overline{n}(4,3,3,5)$
$d^7 \sim (13,4,6,13)$ $d^8 \sim \overline{n}(5,3,4,6)$ $d^9 \sim \overline{n}(20,8,10,20)$

θ^i in percentage of the end value of last segment:

$\theta^{1,2} \sim \overline{n}(20,3,15,28)$ $\theta^{3,4} \sim \overline{n}(10,2,8,15)$
$\theta^{5,6} \sim \overline{n}(-18,3,-26,-13)$ $\theta^{7,8} \sim \overline{n}(-8,2,-13,-6)$
$\theta^9 \sim \overline{n}(0,3,-8,8)$

$limit^i$ in percentage of the end value of last segment:

$limit^{1\cdots4} \sim \overline{n}(140,9,115,140)$ $limit^{5\cdots8} \sim \overline{n}(60,9,60,85)$
$limit^9 \sim \overline{n}(0,4,-8,8)$

III. Results

A. Simulation Testing

The extended Viterbi, the standard Cusum test and the proposed algorithm were tested on the same set of 10 simulated signals, each containing three segments. Half of these simulated signals were randomly generated using the proposed signal model and parameter setting, and the remaining signals each were composed of two randomly generated segments and one manually added segment. The manually added segment has an above-average probability $PsPd$ of being a desired pattern, but a minimum transition likelihood. This means the segment fits the state pattern description well; however, given the previous state, the change is insignificant.

The test results of each method were compared in Table II. For the Viterbi algorithm and our algorithm, the falsely detected patterns were classified into three types, based on the error level in the hierarchy (see Table I). For example, if a *short gradual increase* was detected at the location where the true pattern was a *short abrupt increase*, the error was classified as *Type B*. A *Type A* error is the most severe, indicating the direction of a trend is falsely detected; a *Type B* error is more tolerable to anesthesiologists, indicating an inaccurate description of the speed of change; a *Type C* error is considered the least severe, indicating an inaccurate description of the duration. Due to the lack of pattern recognition, false Cusum results were only grouped into false positives, i.e., *Type A* errors, and false negatives.

For the 25 significant segments, the Viterbi algorithm generated fewer *Type B* and *C* errors and fewer false negatives. However, similar to the Cusum test, it detected more insignificant segments, resulting in more *Type A* errors.

The deviations between the identified $t^{'}$ and the true end points were calculated for the recognized segments excluding *Type A* errors. The bias of our algorithm has a similar distribution to the Viterbi algorithm, and a majority of points were located at their true locations with zero bias (see Fig. 4).

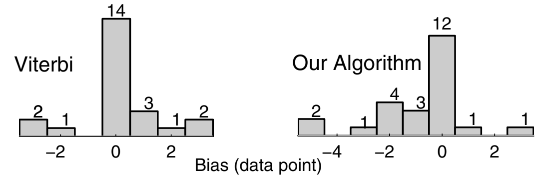

Fig. 4. Histogram: bias of the detected end points

TABLE II

RESULTS ON SIMULATED DATA (PERCENTAGE/NUMBER)

Error Type	A: FP	B	C	FN
Extended Viterbi	**24.0/6**[§]	4.0/1	8.0/2	**0.0/0**
Our Algorithm	**8.0/2**[†]	8.0/2	24.0/6	**4.0/1**
Standard Cusum	**16.0/4**[‡]			**12.0/3**

(1) FP-False Positive, FN-False Negative (2) 25 significant and 5 insignificant segments (3) [§]4, [†]1 and [‡]4 are at the insignificant segments.

TABLE III

RESULTS ON CLINICAL DATA (PERCENTAGE/NUMBER)

Error Type	A: FP	B	C	FN
Extended Viterbi	**38.9/21**	5.6/3	27.8/15	**1.9/1**
Our Algorithm	**7.4/4**	24.1/13	29.6/16	**3.7/2**
Standard Cusum	**14.8/8**			**5.6/3**

(1) FP-False Positive, FN-False Negative (2) 54 annotated segments

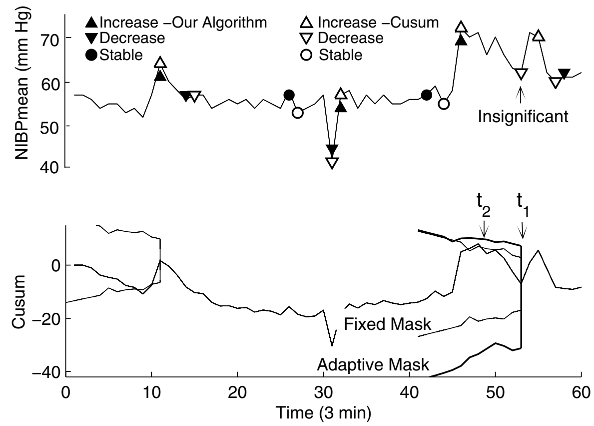

Fig. 5. Proposed algorithm vs. standard Cusum: an example

B. Clinical Data Testing

Following ethical approval, NIBPmean trend signals were collected from 10 day case surgeries at British Columbia Children's Hospital. Since the data collection software sampled the readings of blood pressure monitors at a higher rate (every 5 seconds) than the monitors' cycling rate (every 3 minutes), multiple samples were collected for one measurement. The signals were resampled down before testing. An anesthesiologist annotated the end points and states of the clinically relevant segments after surgery, and the annotations were used as a reference to evaluate the test results.

As seen in Table III, compared to the standard Cusum, our algorithm detected noticeably fewer false positives and marginally fewer false negatives. For instance, in Fig. 5 at t_1 by tracking backward along the fixed mask, the standard Cusum detected the upper arm was broken from within at t_2 and therefore falsely detected this insignificant decrease. At the same location, the adaptive mask was adjusted by our algorithm to a very different shape. Since the last segment was likely a *long abrupt increase*, the adapting parameters λ for detecting the upcoming *short decreases* (S^6 or S^8) became relatively large, and made the corresponding parts of the upper arm larger than the fixed mask. Neither the duration nor the amplitude of this decrease was large enough to break the upper arm; it was therefore not detected.

Compared to the simulation results, the Viterbi algorithm again had the highest false positive rate.We also found higher *Type B* and *C* error rates for both Viterbi and our algorithm, indicating compromised pattern recognition performance.

IV. DISCUSSIONS AND CONCLUSION

The testing results on the simulated signals and the clinical signals demonstrate that by incorporating the knowledge of anesthesiologists, our algorithm provides improved performance over the standard Cusum test for change detection.

Compared to the Viterbi algorithm, our algorithm is able to work online and generates fewer false positives. Most of the false positives detected by the Viterbi algorithm were insignificant changes. The nature of the change point detection problem requires that the significance of the segment is evaluated first, and only then the patterns for the detected segments need to be recognized. However, the Viterbi algorithm maximizes the combined probability of a segment fitting the shape description and the state transition as well. A well fitted pattern can compensate a small transition likelihood and cause a false positive. In our algorithm the transition probability is used to adjust the testing mask before pattern recognition; insignificant segments are therefore avoided.

The pattern recognition accuracy of our algorithm on the clinical data was lower than that on the simulated data. This is not surprising since the parameter setting based on the experience of one anesthesiologist can be biased. If a large number of annotated cases are available, the Expectation Maximization algorithm [6] can be used to obtain more accurate estimates. An alternative solution would be to collect expert knowledge from multiple anesthesiologists and use the inter-rater agreed upon settings to alleviate the bias.

The pattern recognition accuracy is also influenced by the signal characteristics. For states that have overlap in the pattern parameter distributions, errors are inevitable.

Linguistic description of the clinically relevant trend changes can be readily generated using the output of our algorithm, and further used in a high level expert system.

V. ACKNOWLEDGMENT

The authors thank Natasha McCartney for her valuable writing and editing suggestions.

REFERENCES

[1] P. Yang, G. Dumont, J. Lim, and J. M. Ansermino, "Adaptive change point detection for respiratory variables," in *27th Annu. Int. Conf. IEEE EMBS*, (Shanghai, China), 2005.

[2] M. Ostendorf, V. V. Digalakis, and O. A. Kimball, "From HMMs to segment models: A unified view of stochastic modeling for speech recognition," *IEEE Trans. Speech and Audio Proc.*, 1996.

[3] X. Ge, *Segmental Semi-Markov Models and Applications to Sequence Analysis*. PhD thesis, Univ. of California, Irvine, December 2002.

[4] H. Ney, D. Mergel, A. Noll, and A. Paesler, "Data driven search organization for continuous speech recognition," *IEEE Trans. on Signal Processing*, vol. 40, no. 2, pp. 272–281, 1992.

[5] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Prentice Hall, 1993.

[6] K. Squire and S. Levinson, "Recursive maximum likelihood estimation for hidden semi-markov models," in *2005 IEEE Machine Learning for Signal Processing Workshop*, (Mystic, Connecticut), 2005.