

# Motif Discovery in OR Sensor Data with Application to Surgical Workflow Analysis and Activity Detection

Seyed-Ahmad Ahmadi<sup>1</sup>, Nicolas Padoy<sup>1</sup>, Kateryna Rybachuk<sup>1</sup>, Hubertus Feussner<sup>2</sup>, Sandro Michael Heining<sup>3</sup>, and Nassir Navab<sup>1</sup>

<sup>1</sup> Chair for Computer Aided Medical Procedures (CAMP), TU Munich, Germany

<sup>2</sup> Chirurgische Klinik und Poliklinik, Klinikum Rechts der Isar, TU Munich, Germany

<sup>3</sup> Trauma Surgery Department, Klinikum Innenstadt, LMU Munich, Germany

**Abstract.** The computer aided analysis of surgical activity and workflow in the operating theatre has gained much interest in the past few years. Many of these works deal with or depend on detection and classification of surgical activity which is represented by multi-dimensional, continuous signal data recorded from the Operating Room (OR). In this work, we propose a complementary approach directed towards intelligent intermediate processing of raw sensor data. We adopt a technique from data mining called motif discovery, which allows the unsupervised discovery of recurrent and semantically important patterns in otherwise unstructured data. Using data recorded by accelerometers placed on the operator, we discover an objective alphabet of surgical motions performed during simulated percutaneous vertebroplasties and autonomously identify the motion pattern for surgical tool changes in laparoscopic cholecystectomy. The results indicate the usability of motif discovery for autonomous pre-processing and mining of unstructured OR sensor data.

## 1 Introduction

The formal analysis of surgical activity and workflow in the operating theatre (OR) has gained much interest in the past few years. Research groups have put effort into evaluation of surgical skill [1, 2], formalization of surgical workflow [3, 4], monitoring of surgical activity [5] and formulation of languages for surgical gesture description [1]. The aforementioned research areas are all dependent on sensor data obtained from the OR.

In this work, we propose to use methods from the data mining and human activity detection communities for the unsupervised discovery of motifs in multivariate, continuous surgical data. In general, a motif can be defined as a signal pattern that is numerously encountered with only little structural and temporal variation in sequences of much longer duration than the motif itself. In the context of surgical workflow, motifs can represent events or activities with semantic meaning in sensor data that is obtained in the OR.

There are several motivations for the utilization of motif discovery algorithms. First, motif discovery can be seen as a form of data pre-processing that helps in filtering semantically important patterns out of otherwise unstructured data.

Second, the identification and learning of important segments in data can potentially reduce the data dimensionality and the amount of data necessary for pattern recognition algorithms, since data segments in the learning set are focussed onto the motif area and hence include less noise or distracting patterns. Third, higher-level processes can be broken down into a vocabulary of semantically important sub-processes. This suggestion of a language or vocabulary is performed automatically and can rely on no or only little user interaction.

There are several examples in literature where motif discovery could be beneficial. Lin et al. [1] use Hidden-Markov-Models (HMMs) to automatically segment motion data during a suturing task performed with a DaVinci surgical robot. They formulate a "language of surgery", where surgical gestures are described as a sequence of tasks or words called "surgemes". The borders of task segments were defined manually. Motif discovery can be seen as one way for unsupervised definition of a vocabulary of "surgemes". This might seem suboptimal compared to expert knowledge. However, one can imagine situations where detection of repetitive patterns is challenging even for an expert, e.g. in complete procedures instead of short training tasks, or in multivariate brain data. Similarly, Rosen et al. [2] model the surgical skills of an operator with an HMM. The HMM states are constructed manually from the possible combinations of tools position and orientation. Defining a subtask vocabulary on these signals has the potential to automatize this tedious task. Jannin et al. [3] and Burgert et al. [6] define ontologies for a formal description and analysis of surgical workflow. The automatic detection of surgical activity based on sensor data could alleviate the model generation. Certain ontology states could be represented by the presence or absence of particular motifs. An indication for the usability of motifs can be seen in the work of Padoy et al. [5]. The usage of laparoscopic tools is represented in binary form and used to recognize the current phase of the workflow, in which the surgery resides. If a vocabulary of motifs is automatically detected, unstructured multivariate sensor data can be translated into further discrete signals. This could yield higher recognition rates than with raw data, due to dimensionality reduction. In the following, we will demonstrate the usage of motif discovery to define motion vocabularies from inertial data acquired in two medical applications.

## 2 Related Work

Motif discovery originated in the field of biosequence analysis, e.g. for unsupervised discovery of genomes or protein sequences. D'Haeseleer [7] gives an informative introduction to motif discovery in biosequences. This includes a basic overview over various detection methods, methods for measuring motif similarity and single motif quality as well as practical guidelines for motif discovery in biosequences. A survey by Sandve et al. gives a broader overview over methods employed in DNA motif discovery [8]. In video data mining, Xie and Chang [9] achieved notable results for unsupervised detection of repetitive segments in a video stream with consistent statistical characteristics which they call structure,

but which essentially means the detection of motifs in video. They achieve this by employing hierarchical hidden Markov models (HHMM) based on a multi-level Markov assumption. Hamid et al. [10] propose a different approach, based on the representation of activities as bags of n-grams. Different classes of activities are found by searching the graph for maximal cliques. The proposed methods, however, rely on the usage of a discrete set of high-level events, rather than raw sensor data. In this work, we intend to find motifs in multidimensional continuous sensor data, namely motion data recorded with accelerometers. The group around Keogh et al. has investigated this topic for large databases. In [11], Keogh et al. denote the pitfalls in simple clustering of time series subsequences and formulate requirements for motif discovery. An approach for parameter-free motif discovery based on SAX discretization and DTW distance computation is given in [12]. A probabilistic approach based on Random Projection is given in [13].

Minnen et al. use another approach called Subsequence Density Estimation (SDE) [14] for motif discovery, e.g. in multi-sensor data recorded during soldier activity in the field [15]. Since they use SDE for motif discovery in acceleration data, we adopt their algorithm for our problem setting.

### 3 Materials and Methods

#### 3.1 Data Acquisition

We recorded surgical motion during simulated and live surgeries of vertebroplasty and cholecystectomy, using three actibelt<sup>©</sup> accelerometers (3D, 100Hz, 12bit) attached to the waist and wrists of the surgeon. All recordings were performed with approval of the medical directors in charge as well as the sterilization departments of our medical partner institutions, using a sterile attachment method presented in [16]. Additionally, we recorded up to three videos of the OR with different perspectives, for later labeling of activity groundtruth and validation of the results. Videos had a framerate of 25Hz. It should be noted that no video information was used for motif discovery.

#### 3.2 Motif Discovery Algorithm

Minnen et al. have proposed an algorithm for motif discovery called Subsequence Density Estimation (SDE) and have shown its usability for human motion data recorded with accelerometers. We adopt this framework for our experiments and explain it in the following, in order to keep this paper self-contained.

The basic principle of SDE is to first split a dataset of  $N$  curves  $\mathbf{S} = \{S_1, \dots, S_N\}$  with length  $L_1, \dots, L_N$  and total length  $P$  into subsequences  $\mathbf{X} = \{x_1, \dots, x_n\}$  of length  $W$  and overlap  $o_{rel}$  according to  $n = \left\lfloor \frac{P}{W(1-o_{rel})} \right\rfloor$ .

The aim of subsequence density estimation (SDE) is then to find high-density regions in the subsequence space, since these local maxima are assumed to represent patterns with semantic meaning. These local maxima will be called motif seeds in the following.

The SDE procedure is based on the following execution rules [14]:

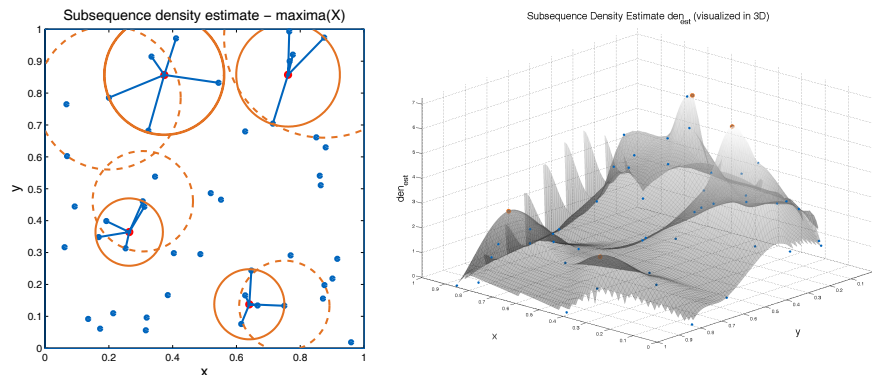
1. The DTW distances [17] between all non-overlapping subsequences  $dist(x_i, x_j)$  have to be calculated pair-wise, where  $i, j = 1 \dots n$ , and  $j \neq \left[ i - \left( \frac{1}{1-orel} - 1 \right), \dots, i + \left( \frac{1}{1-orel} - 1 \right) \right]$ . Overlapping subsequences have to be excluded in order to avoid trivial matches [14] which would render the SDE procedure meaningless [11].
2. The  $k$  nearest neighbors for each subsequence have to be found:  $knn(x_i) = x_{i,l}$ , where  $l = 1 \dots k$ .
3. The density estimation is calculated as the inverse of the distance to the  $k$ -th nearest neighbor:  $den_{est}(x_i) = \frac{1}{dist(x_i, x_{i,k})}$ .
4. Local maxima are those subsequences for which the local density estimate is higher than the density estimates of all its  $k$  nearest neighbors. More formally,  $maxima(X) = \{x_i \in X | \forall x_{i,j}, den_{est}(x_i) > den_{est}(x_{i,j}), j = 1 \dots k\}$ .

The approximate detection of density modes performed by SDE is illustrated in Figure 1, which displays 50 two-dimensional tuples drawn from a uniform distribution. Following the SDE procedure, we find density modes estimates. The figure shows the maxima and the distance between each local maximum  $x_i \in maxima(X)$  and its farthest neighbor  $dist(x_i, x_{i,k})$  shown as a solid circle. The distance from the farthest neighbor to its respective farthest neighbor is illustrated as a dashed circle. The density estimates are also displayed as a 3D in Fig. 1, where it is more obvious that SDE picked the local maxima as motif seeds.

In the experiments, we are varying the subsequence length  $W$  and observe its effect on the quality of motif discovery. The only parameter inherent to the SDE procedure, the number of nearest neighbors  $k$ , is set to one less than the number of curves in the investigated dataset,  $k^* = N - 1$ . Behind this choice lies the intuitive notion that we want to detect repetitive activities, i.e. activities which appear at least once in each training surgery. Therefore, such a one-time activity should have  $k^*$  closest neighbors. The change in subsequence duration has an effect on algorithm runtime according to the following formula for SDE complexity:  $\mathcal{O} \left( \left( \frac{P}{W(1-orel)} \right)^2 (W^2 + k) \right)$ .

### 3.3 Motif Redetection

The SDE algorithm finds a motif seed as a local maximum, based on an approximation deducted from the  $k$ -th neighbor. In the *VertSim10* dataset (cf. section 4.1), we investigated the possibility to use the motif seed and its  $k$  neighbours for training of Hidden Markov Models (HMMs) and subsequent re-detection of learned motifs in unknown data [14]. HMMs have been shown to be effective for modeling of time-series data in numerous cases. For an in-depth explanation of HMM theory and implementation, we refer the reader to Rabiner’s tutorial on HMMs [18]. For this work, we utilized the Bayesian Network Toolbox (BNT) implementation provided by Murphy et al. [19].



**Fig. 1.** The SDE algorithm approximates local maxima in density space with a  $k$ NN search. The density modes represent motifs in otherwise unstructured data.

## 4 Experiments and Results

### 4.1 Datasets

For evaluation of the SDE algorithm, we recorded three datasets for two different medical procedures, in our laboratory and together with our medical partners. The first procedure is percutaneous vertebroplasty. In this procedure, a metallic needle, or trocar, is inserted percutaneously into a broken or wedged vertebra and acrylic cement is injected which leads to stabilization and pain relief for the patient. Using the three accelerometers, we recorded surgical motions with varying relevance to the vertebroplasty workflow (cf. Table 1). The second procedure is laparoscopic cholecystectomy, a minimally-invasive abdominal surgery for removal of the gallbladder. The motions to be discovered by the motif discovery algorithm are *insert instrument* and *remove instrument*. This information is complementary to the binary instrument tool usage signals used by Padoy et al. [5] and could thus be used to improve the detection of surgical workflow phases.

We recorded three datasets:

1. **VertSim10:** The first dataset is a simulation of ten percutaneous vertebroplasties, performed by a single person (non-medical) and in a constrained environment. This dataset is deliberately held simple, in order to investigate and demonstrate the usability of SDE motif discovery in a medical background. A simplified percutaneous vertebroplasty on a spine phantom, with surgical tools was performed and care was taken to vary motion parameters such as periodicity during *hammering* or *steering*. However, random motions were avoided to a large extent, in accordance with related literature.
2. **VertSim5:** The second dataset consists of five simulated vertebroplasties, performed by a partner surgeon, on five different plastic spine phantoms. Motions and surgical activity were entirely unconstrained.

3. **LapChol:** The third dataset is the recording of a single laparoscopic cholecystectomy, in which the sensor-equipped surgeon had the leading role in almost the entire surgery.

	<b>Workflow-relevant motions</b>	<b>Workflow-irrelevant motions</b>
<b>VertSim10</b>	Silence, Walking, Hammering, Needle Position, Cement Stirring, Cement Injection	Clap Arm Synchr., Clap Waist Synchr., Put On Vest, Cut and Insertion, Syringe Cleaning, Trocar Retraction, Put Off Vest
<b>VertSim5</b>	Silence, Needle Position, Hammer Insertion, Twist Insertion, Cement Stirring, Fill Syringe, Apply Syringe	Random Move, Tool Change, Guidance Needle In, Guidance Needle Out, Apply Trocar Cap, Trocar Retraction, Apply Patch

**Table 1.** Motion alphabets used for labelling recorded for each vertebroplasty dataset, grouped by workflow relevance. The manual labeling is subjective and only used for comparison against objective and autonomous identification of motion vocabularies with motif discovery.

## 4.2 Quality Measures

Based on these datasets, we have performed experiments with the SDE motif discovery algorithm. We evaluate motif discovery with three measures:

**Motif Seed Diversity:** Motif seed diversity is calculated as the ratio of different activities among motif seed candidates to the size of the activity alphabet in the groundtruth data. The activity of a motif seed candidate is derived from groundtruth as the activity that accounts for most of the time frames of the motif seed plus its  $k$  neighbors. Motif seed diversity therefore gives an idea of how many different activities were detected in comparison to the number of manually labeled patterns known to exist in the data.

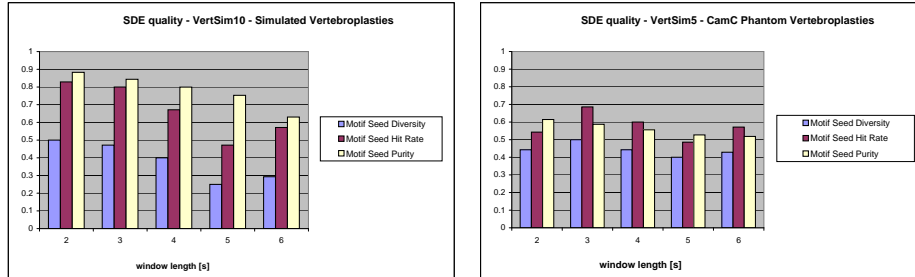
**Motif Seed Hit Rate:** Within the groundtruth alphabet of each vertebroplasty dataset, there are certain activities that are more relevant to the workflow of vertebroplasty than others. These activities are highlighted in Table 1, where the groundtruth motion alphabets for each vertebroplasty dataset are given. Motif seed hit rate is the percentage of how many of these activities were detected during SDE. Thus, motif seed hit rate measures whether SDE was actually able to capture interesting and workflow-relevant patterns within the data.

**Motif Seed Purity:** As mentioned in the definition of motif seed diversity, each motif seed candidate can be associated with an activity from the groundtruth’s alphabet. The confidence of a motif seed is the ratio of data frames associated with the motif activity to the total number of frames of the motif seed and its  $k$  neighbors, measured for all detected motif seeds across all folds. Motif seed purity is therefore a measure for the purity of the data that a motif HMM is trained on, i.e. a measure for how much a motif seed is based on data of a single activity rather than a mixture of activities.

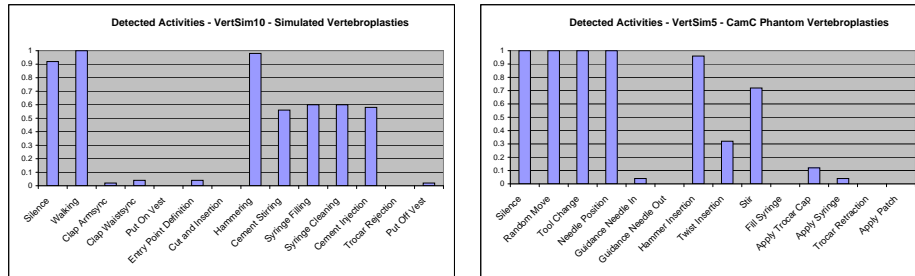
Additionally, we experimented with motif re-detection using HMMs in the *VertSim10* dataset. Similar to Minnen et al. [15], we evaluate classification accuracy in terms of correct detections (true positive rate, TPR), correct non-

occurrences (true negative rate, TNR), missed occurrences (false negative rate, FNR) and false detections (false positive rate, FPR). The results for the selected motion alphabets on the *VertSim10* dataset are given in figure 4. The most important value in these plots is the true positive rate (TPR) which is equivalent to the accuracy of detection.

### 4.3 Motif Discovery Results



**Fig. 2.** Motif seed quality measures for the VertSim10 and VertSim5 datasets for different window lengths and averaged across all folds.



**Fig. 3.** Detected activities in the VertSim10 and VertSim5 datasets across all subsequence lengths and folds. A value of 1 indicates that an activity was detected as a motif seed for all subsequence lengths and all folds. The results indicate which activities were often and which activities were rarely detected as a motif seed.

The main results are summarized in Figures 2 and 3. Several insights can be drawn from figure 2.

One of the results is that motif seed diversity is tendentially higher for shorter subsequence lengths. This seems intuitive because shorter segments of accelerometer data allow even short occurrences of activities to be detected. Such short occurrences might be overseen by subsequence density estimation if window lengths are too long. Accordingly, motif seed purity, which is a measure for the purity of motif seeds, decreases with increasing window lengths. Again, this seems intuitive since shorter subsequences are more likely to contain only data frames from a single activity rather than a mixture. Motif seed hit rate, which measures

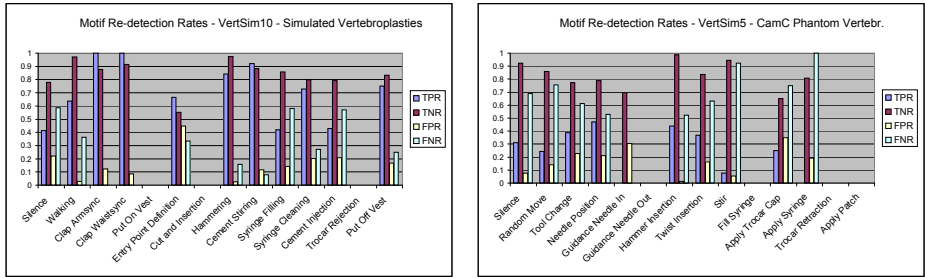


Fig. 4. Quality of HMM redetection of motifs in *VertSim10* dataset.

whether workflow-relevant activities were captured by SDE, behaves in a similar way as motion seed diversity. It is especially high for the *VertSim10* dataset. This indicates that motions are detected with high probability if they are executed in a systematic manner, which is in line with related literature [14, 20, 21]. Movements that are executed in an unconstrained environment, as it is the case in the *VertSim5* dataset, yield motif seeds with much lower purity and hit rate. In other words, motif discovery faces limitations when movements are executed in increasingly less-constrained environments.

Realistically, surgeons do face several constraints in their motion freedom. First, surgeons are highly concentrated while interacting with the body of the patients. This leads to very systematic and recurrent motion patterns. Furthermore, due to the sterility of their coats, surgeons are not allowed to touch surrounding objects which are non-sterile. This leads to a confined space for surgeons to move around. Furthermore, surgeons usually keep their hands to their body while being idle. This behavior was observed to be consistent across all three medical partners participating in data acquisitions.

Another observation in these experiments is that there are some activities that are consistently detected across all datasets while other movements are detected with less consistency. Figure 3 shows that the most "popular" motion patterns to be detected are silence and movements which involve the movement of the whole body, such as *walking* in *VertSim10* and *tool change* in *VertSim5*. Surprisingly, *random move* was detected as a motif for all window lengths and all folds for the *VertSim5* dataset. This might indicate that a repetitive movement that was not identified within the human labeling procedure might be hidden in the random movement data. Workflow-relevant motions such as *silence*, *hammering*, *walking* and *stirring* were discovered reliably in both datasets.

Concerning the *LapChol* dataset, the SDE motif discovery was able to identify the motions of *tool change*, *insert instrument* and *remove instrument*, as two separate motifs with an average motif seed purity of 63.8%. More data than the single recorded cholecystectomy could help improve this result. Training a HMM model on these two activities could be seen as pre-processing and translation of raw sensor data into binary signals with semantic meaning, similar to [5], thus aiding in improved recognition of surgical workflow.

## 5 Discussion and Conclusion

The experiments and results shown for motif discovery demonstrate that the SDE algorithm is able to recover recurrent and workflow-relevant patterns in unconstrained motion data. Important motions were identified and can then be summarized into a motion alphabet. However, there are also limitations to the SDE procedure. SDE faces limitations with motions that are performed only shortly or only rarely or maybe even only once during each procedure, such as *syringe filling* in the *VertSim10* and *VertSim5* datasets. Partly, these shortcomings may be cured by variations of the parameter  $k$ . However, as with every unsupervised algorithm, it cannot be guaranteed that SDE finds all patterns of interest to the user. Thus, it might be useful to incorporate a small user interaction step into the motif discovery. However, the autonomous SDE discovery can help to reduce user interaction significantly.

We also demonstrated that the motif discovery process can yield a small set of training data for HMM modeling of activities and subsequent re-detection in unknown curves. This sets the premise that motif discovery yielded motifs of high data purity, i.e. training data from a single activity rather than a mixture of activities. The improvement of the motif discovery algorithms could thus yield better re-detection accuracies as well.

Future work might include the incorporation of sub-dimensional motif discovery [22]. Also, SDE has a complexity of  $n^2$ , which might be reduced to  $n(\log n)$  by using spatial indexing techniques for accelerated  $k$ NN search [23]. We will also investigate a comparison with Random Projection motif discovery [13] in future.

In this paper, we have introduced motif discovery as a technique for autonomous identification of alphabets of surgical activity. The alphabets are more objective, require little to no user interaction and can aid in improved surgical workflow analysis and activity detection. Due to the promising results, we will further investigate into this important field of research, concerning unsupervised pre-processing and mining of unstructured OR sensor data.

## References

1. Lin, H.C., Shafran, I., Yuh, D., Hager, G.D.: Towards automatic skill evaluation: Detection and segmentation of robot-assisted surgical motions. *Computer Aided Surgery* **11**(5) (2006) 220–230
2. Rosen, J., Brown, J., Chang, L., Sinanan, M., Hannaford, B.: Generalized approach for modeling minimally invasive surgery as a stochastic process using a discrete markov model. *IEEE Trans. on Biomedical Engineering* **53**(3) (2006) 399–413
3. Jannin, P., Morandia, X.: Surgical models for computer-assisted neurosurgery. *NeuroImage* **3** (2007) 783–791
4. Neumuth, T., Strau, G., Meixensberger, J., Lemke, H.U., Burgert, O.: Acquisition of process descriptions from surgical interventions. In: *Lecture Notes in Computer Science DEXA*. Volume 4080. (2006) 602–611

5. Padoy, N., Blum, T., Feuner, H., Berger, M.O., Navab, N.: On-line recognition of surgical activity for monitoring in the operating room. In: Proc's of the 20th Conference on Innovative Applications of Artificial Intelligence (IAAI-08). (2008)
6. Burgert, O., Neumuth, T., Gessat, M., Jacobs, S., Lemke, H.U.: Deriving dicom surgical extensions from surgical workflows. In: SPIE Medical Imaging 2007, PACS and Imaging Informatics. Volume 6516. (2007) 651604
7. D'haeseleer, P.: How does dna sequence motif discovery work? *Nature Biotechnology* **24**(8) (2006) 959–961
8. Sandve, G.K., Drabls, F.: A survey of motif discovery methods in an integrated framework. *Biology Direct* **1** (2006) 11 ff.
9. Xie, L., Chang, S.F.: Unsupervised Mining of Statistical Temporal Structures in Video. In: Video Mining. Springer (2003) 279–307
10. Hamid, R., Johnson, A.Y., Batta, S., Bobick, A.F., Isbell, C., Coleman, G.: Detection and explanation of anomalous activities: Representing activities as bags of event n-grams. In: CVPR (1), IEEE Computer Society (2005) 1031–1038
11. Keogh, E., Lin, J., Truppel, W.: Clustering of time series subsequences is meaningless: Implications for previous and future research. *icdm* **00** (2003) 115
12. Lin, J., Keogh, E., Lonardi, S., Chiu, B.: A symbolic representation of time series, with implications for streaming algorithms. In: Proc's of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery. (2003)
13. Chiu, B., Keogh, E., Lonardi, S.: Probabilistic discovery of time series motifs. In: KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, New York, NY, USA, ACM (2003) 493–498
14. Minnen, D., Isbell, C., Essa, I.A., Starner, T.: Discovering multivariate motifs using subsequence density estimation and greedy mixture learning. In: AAAI, AAAI Press (2007) 615–620
15. Minnen, D., Westeyn, T., Presti, P., Ashbrook, D., Starner, T.: Recognizing soldier activities in the field. Proceedings of International IEEE Workshop on Wearable and Implantable Body Sensor Networks (BSN) **13** (2007) 236–241
16. Ahmadi, S.A., Padoy, N., Heining, S.M., Feussner, H., Daumer, M., Navab, N.: Introducing wearable accelerometers in the surgery room for activity detection. In: Computer-und Roboter-Assistierte Chirurgie (CURAC). (2008)
17. Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. *Acoustics, Speech, and Signal Processing* [see also *IEEE Transactions on Signal Processing*], *IEEE Transactions on* **26**(1) (1978) 43–49
18. Rabiner, L.R.: A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE* **77**(2) (1989) 257–286
19. Murphy, K.P.: *The bayes net toolbox for matlab* (2001)
20. Minnen, D., Starner, T., Essa, I.A., Isbell, C.: Discovering characteristic actions from on-body sensor data. In: *IEEE Int'l Symposium on Wearable Computers (ISWC)*. (2006) 11–18
21. Krause, A., Siewiorek, D.P., Smailagic, A., Farringdon, J.: Unsupervised, dynamic identification of physiological and activity context in wearable computing. In: *7th International Symposium on Wearable Computers (ISWC 2003)*. (2003) 88–97
22. Minnen, D., Isbell, C., Essa, I., Starner, T.: Detecting subdimensional motifs: An efficient algorithm for generalized multivariate pattern discovery. In: *Proceedings of the 7th IEEE International Conference on Data Mining (ICDM)*. (2007) 601–606
23. Berchtold, S., Keim, D.A., Kriegel, H.P.: The x-tree: An index structure for high-dimensional data. In: *Proceedingost the 22ndV LDB Conference*. (1996)