Unsupervised Temporal Segmentation of Skeletal Motion Data using Joint Distance Representation

Christian Lins[1], Sebastian M. Müller[2], Max Pfingsthorn[2], Marco Eichelberg[2], Alexander Gerka[2], and Andreas Hein[1]

[1]*Carl von Ossietzky University, Ammerländer Heerstr. 140, Oldenburg, Germany*

[2]*OFFIS - Institute for Information Technology, Escherweg 2, Oldenburg, Germany*

{christian.lins, andreas.hein}@uni-oldenburg.de, {sebastian.mueller, max.pfingsthorn, eichelberg, alexander.gerka}@offis.de

Keywords: Human Motion Analysis, Temporal Segmentation, Joint Distance Matrices, Musculoskeletal Disorders, Ergonomics Assessment

Abstract: In this paper, we present an online method for the unsupervised segmentation of skeletal motion capture data for the assessment of unfavorable or harmful postures in the context of musculoskeletal disorders. The long-time motion capture data is segmented into short motion sequences using joint distances of the captured skeleton. We use the difference between joint distance matrices to detect variances in motion dynamics in which the motion is separated into either a dynamic motion or a static posture. Then, the static posture can be evaluated using well-known posture assessment methods such as the Ovako Working postures Analysing System (OWAS) to derive risk factors for musculoskeletal disorders. The algorithm works in real-time so that it can be incorporated in live warning systems for unfavorable or harmful postures. We evaluated the segmentation algorithm by comparing it with results from state-of-the-art offline motion segmentation algorithms as gold standard. Results show that the algorithm approaches the performance of state-of-the-art offline segmentation algorithms.

1 INTRODUCTION

The analysis and classification of human motion have been an active research topic for a long time in various disciplines (Wang et al., 2003; Aggarwal and Cai, 1997). Applications include but are not limited to media animation, biometrics, ergonomics, sports or computer sciences. In the last years, motion capture (MC or MoCap) systems, especially such with inertial sensors (gyroscope, accelerometer, and magnetometer, combined as IMU, short for Inertial Measurement Unit), have become available to a greater audience. Such systems, e.g. Xsens MVN (Roetenberg et al., 2013) or SIRKA (Lins et al., 2015), make it possible to capture the human motion continuously for hours or even days. Additionally, such small and embeddable motion capture suits make occupational in-situ observations possible (see Figure 1).

One application is the risk factor assessment for work-related musculoskeletal disorders (WMSD) (Wang et al., 2015). Industry workers regularly perform unfavorable, harmful or even dangerous postures during their work shifts. These postures can – together with other risk factors – lead to musculoskeletal disorders (MSD) such as back pain. Besides the personal

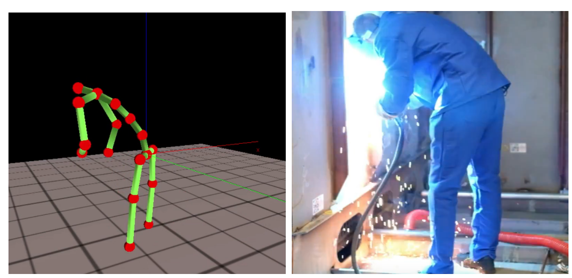

Figure 1: Welder's digitalized posture and reference footage.

inconveniences, MSDs are a primary cause for sick leave and early retirement in physically demanding occupations (Ellegast, 2013). According to (Punnett and Wegman, 2004), MSDs are the largest category of work-related illnesses in many countries. The treatment of MSDs causes high costs in the public health systems of various countries (for example the total economic costs of MSDs in Canada for the year 1994 equals 3.4% of the gross domestic product) (Coyte et al., 1998; Walker et al., 2003).

However, the emergence of MSDs can be delayed or even prevented, if appropriate and timely preventive measures are applied (Armstrong et al., 1996). Even if the causes of MSDs are not always occupational causes, heavy physical work such as manual handling and lifting is often considered a risk factor for the emergence of MSD (Amell and Kumar, 2001; Hoy et al., 2010; Matsui et al., 1997). Thus, prevention measurements become a necessity, e.g. as part of the corporate health management in industrial companies with physically hard-working employees. For the prevention, accompanying physiotherapists will support affected employees, i.e. those with risk factors, to improve working processes. This improvement can already prevent the emergence of occupational diseases, which usually occur in the second half of life. At the organizational level, an analysis of the workplace or work planning could be used as starting point for new managements concepts that already consider recovery phases in the work plan.

It is an ongoing task of corporate health management to continuously assess psychological and physical risk factors of every workplace and every working individual. Risk factors based on physical activity such as postures and motions can be derived by human observers using pen-and-paper assessment methods such as Ovako Working postures Analysing System (OWAS) (Karhu et al., 1977), Rapid Entire Body Assessment (REBA) (Hignett and McAtamney, 2000), Rapid Upper Limb Assessment (RULA) (Corlett, 2003), the European assembly worksheet (EAWS) (Schaub et al., 2013), and others. Such assessments are tedious and time-consuming, i.e. costly, tasks which could be supported or replaced by automatized Motion Capture (MoCap) based systems.

When workers wear a MoCap system during their work shifts, the posture data captured by the MoCap system can be either stored for later analysis or preliminarily evaluated on a wearable device (for example as described by (Nath et al., 2017)). The analyzing software can identify critical postures to give the wearer an immediate feedback on her or his possibly harmful postures. Then the employee might be able to actively take a more ergonomic posture or interrupt the work for a moment to recover.

An evaluating device has to process the motion data in real-time, a task that requires appropriate algorithms that handle the continuous stream of motions. Consequently, we propose an approach using an algorithm suited for segmenting dynamic human motion data into short motion sequences (segments). Our contribution is an online algorithm that uses a time series of skeletal motions, which is transformed into a joint distance matrix representation that makes the skeletal representation rotation and translation invariant. The algorithm exploits the joint distance representation to derive a motion dynamics indication, i.e., whether a motion represents a static posture or a dynamic motion. The output of the algorithm is a set of segmented motion sequences with the associated motion dynamics indication that can be further processed by the assessment system (see also Figure 4).

In this paper, we first take a look at related work of segmentation algorithms (2). We then explain the segmentation algorithm in detail and compare the effectiveness of the new algorithm to the segmentation algorithms of (Krüger et al., 2016; Vögele et al., 2014) and (Zhou et al., 2013; Zhou et al., 2008). Finally, we discuss the results and conclude this paper in the last section.

2 RELATED WORK

Algorithms for unsupervised segmentation of motion data can be divided into two classes. First, algorithms that require *a priori* knowledge, i.e., the complete data set (or a large batch) to select meaningful segmentation (or cut) points (offline algorithms). Secondly, algorithms that can work without (or very little) *a priori* knowledge processing data incrementally (online algorithms).

Segmenting arbitrary time-series data into smaller parts is a well-known problem in computer science (Keogh et al., 2001), but there is some work specific to the problem of segmenting human motion data. Zhou et al. – as an example of an often mentioned offline method – use Dynamic Time Warping (DTW) to describe the problem of motion segmentation as a clustering problem which they address using kernel k-means (Zhou et al., 2008). The authors conclude that their method is computationally too expensive for larger motion data sets and provide an improved method that incorporates a hierarchical decomposition of motions at different temporal scales (Zhou et al., 2013). However, the improved version also has a time complexity of $\mathcal{O}(n^2 n_{max} t)$ with n number of input frames and t number of iterations. So it is presumably not usable for real-time segmentation.

Vögele et al. , Krüger et al., and Stollenwerk et al. order the skeletal input data using a k-d-tree to find neighbors of a specific posture within a search radius R (Vögele et al., 2014; Krüger et al., 2016; Stollenwerk et al., 2016). A set of neighbors is then represented as a sparse self-similarity matrix whose graphical structure is used to find distinct motion segments. The method requires the distance between different neighbored postures, which are not all available in an online system.

Kulic et al. (Kulic et al., 2009) – in contrast to the methods mentioned above – describe an approach for an unsupervised online segmentation and clustering algorithm. Here, a Hidden Markov Model (HMM) is defined on a set of sliding windows containing the motion data. A standard Gaussian density estimator is used as observation probability distribution. A modified Viterbi algorithm is used to find the optimum state sequence in the HMM, which represents the motion segment. The growing number of states is circumvented by limiting the number of windows, which weakens the online capabilities.

Koenig and Matarić (Koenig and Matarić, 2006) use a sliding window to search for local maxima in feature space variances of demonstrated tasks in the domain of robot learning from demonstration. Features can vary depending on the behavior that is demonstrated to the robot. The authors recognize the importance of online capable algorithms for such tasks and propose a fast and intuitive method for segmenting tasks into behaviors. Our approach is similar to this one but specifically adapted to the purpose of human motion segmentation. We use distances of concurrent joint distance matrices as features. Additionally, our algorithm does not use a sliding but dynamically growing (up to a maximum for every sequence) window approach. Koenig and Matarić's algorithm emits a segmentation point when a specific variance threshold is recognized, our method, on the other hand, uses the ratio between the variances of two sequences within the window as the threshold. We explain our approach in detail in the next section.

3 MOTION SEGMENTATION ALGORITHM

3.1 Definitions and Preconditions

A motion capture system provides a sequence M of m skeletal postures S ordered in time, so $M = (S_1, S_2, ..., S_m)$. Every skeletal posture S_i is an n-tuple $S = (J_1, J_2, ..., J_n)$ where every J denotes a joint position of the skeleton in $\mathbb{R}^3$.

3.2 Transformation to Distance Matrices

The use of joint distance matrices as features for classifying motion capture data was first shown by (Vieira et al., 2012). The representation of a posture as distance matrix has one major advantage: the representation is invariant to rotation or translation of the whole body or the point of view of the observer. That means that it does not matter whether the skeleton comes from an inertial sensor suite with its coordinate origin at the body center or from an optical camera based system such as Kinect with its coordinate origin at the observing camera.

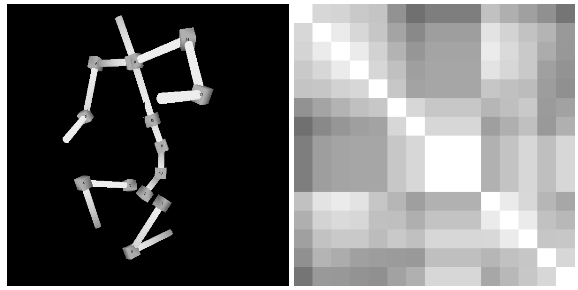

Figure 2: An example of a kneeling skeleton posture and a joint distance matrix visualization as grayscale heatmap (white pixel 0.0, black 1.0.)

The joint distance matrix is also a self-similarity matrix but it differs from the ones used in the other mentioned algorithms (Vögele et al., 2014; Krüger et al., 2016; Stollenwerk et al., 2016) as it is for the multiple joint coordinates of one pose and not for all poses of the motion sequence.

The skeletons provided for processing are in most cases not normalized, i.e., they are not fitted to a standard skeleton size to make postures comparable. For this reason, the algorithm can be provided with a normalization factor to transform the skeleton before further processing. The normalization factor can be derived from the length of a rigid skeleton segment, e.g., shoulder-elbow or a femur bone. A normalization makes the distance matrices comparable between different motion recordings and subjects. However, for segmenting motion data into motion sequences, normalization is not necessary.

The distance matrix D_S of a posture S can be defined as

$$D_S = [||J_k - J_l||]_{k,l} \qquad (1)$$

where l, k are the joint indices of the skeleton. In other words, D_S is a $n \times n$ matrix denoting the absolute distances (a distance metric, e.g. Euclidean) between every joint in posture S. As mentioned before, the distances between joints with a rigid connection do not change during motion so that these values can be discarded (e.g. set to 0). Figure 2 shows an example posture and the corresponding joint distance matrix as grayscale heatmap image.

3.3 Distance Measure of Joint Distance Matrices

We define a distance or similarity function $d_{ij} = s(D_i, D_j)$ with $d \in \mathbb{R}^+$. The similarity measure of ev-

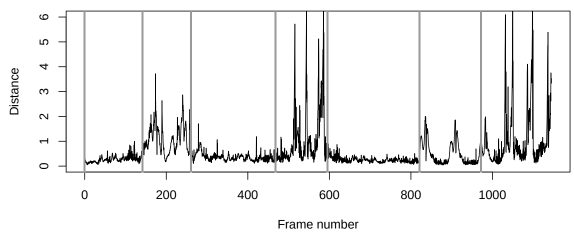

Figure 3: Distance Measures of Trial 01 with cut points (gray vertical lines) of our algorithm ($l_{min} = 3.5$).

ery distance matrix pair (D_i, D_{i+1}) is calculated for $0 \leq i \leq m-1$. That means that we determine the amount of change between two distance matrices each representing a posture. A plot of a distance change of one joint over time is shown in Figure 3.

Self-similarity matrices (such as joint distance matrices) have some characteristics that are useful here. First, the matrix is symmetric on its diagonal. That is because of the definition of the distance function that is used to fill the matrix. Then, if we assume that a bone connecting two joints is rigid, the distance between those two does not change throughout the recording. It will only vary below the noise threshold. As a result, only the distances of joints not directly connected are relevant for distance measure between two data frames. A indicator function $c(J_l, J_k)$ is derived from the skeleton definition returning 0 if the connection is rigid or $l = k$ and 1 if the connection is not rigid.

With this indicator function c we can define the distance function as modified L^1 norm:

$$s(X,Y) = \sum_{k=0}^{n} \sum_{l=k}^{n} (|x_k - y_l| \cdot c(J_k, J_l)) \qquad (2)$$

with X, Y being two distance matrices as defined by Equation 1. Other distance measures for matrices could be adapted as well.

3.4 Maximize variance ratio

Differences between distance matrices (see Equation 2) can be seen as a continuous time series. One problem with sensor-based time series is the handling of noise. If the data is noisy, a non-robust segmentation algorithm will return many extrema on the noisy data. So the challenge is to find an algorithm that properly filters noise, handles the peculiarities of human motion and returns timestamps that can be used as cut points for motion data.

Our algorithm finds a frame in a dynamic window W that separates the window into two segments, one with high variance and one with low variance. Such frame can be seen as start or end point of a motion segment, e.g. a constrained posture with little dynamic and a change to high dynamic (meaning high variance) when the subject starts to move.

Algorithm 1 Segmentation algorithm

Require: W is a n-element window of distances d_0 to d_n

Require: l_{min}, l_{max} is the minimum/maximum segment length (default: $l_{max} = 4 \cdot l_{min}$)

Require: T is the min. threshold ratio that is required for segmenting a window (default: $T = 10$)

function SEGMENT(W, l_{min}, l_{max}, T)
    $r \leftarrow 0$
    for $p \leftarrow l_{min}$ to $p \leftarrow n - l_{min}$ **do**
        $\mu_1 \leftarrow \text{mean}(W, 0, p)$ ▷ Sample mean
        $\mu_2 \leftarrow \text{mean}(W, p+1, n)$
        $\sigma_1 \leftarrow \text{var}(W, \mu_1, 0, p)$ ▷ Sample variance
        $\sigma_2 \leftarrow \text{var}(W, \mu_2, p+1, n)$
        $r \leftarrow \max\left(r, \left(\frac{\max(\sigma_1, \sigma_2)}{\min(\sigma_1, \sigma_2)}\right)\right)$
        if $r > T$ **or** length(W) $> l_{max}$ **then**
            $p \leftarrow \text{findMin}(p, W, \sigma_1, \sigma_2, l_{min})$
            return p and σ_1
    return ▷ No proper cut point found yet

The algorithm is outlined in pseudocode as Algorithm 1. We assume a minimum segment length l_{min} to avoid segmentation into very small motion fragments. So the first possible segmentation (or cut) point is the index at l_{min}.

The index p for which the ratio $r = \max$ is a plausible candidate for a cut point. To ensure that d_p is a local minimum, which can be seen as a rest pose, we search $\frac{l}{c}c$ steps for a local minimum in the part of the window with lower variance. $\frac{l}{c}$ must be a small fraction of l. The function *findMin* (see Algorithm 1) implements this linear search within the surrounding data points (e.g. $p - \frac{l}{2c} \rightarrow p + \frac{l}{2c}$).

3.5 Detection of Constrained Postures

Constrained postures are often performed by industry workers during their duties, e.g. holding a tool in

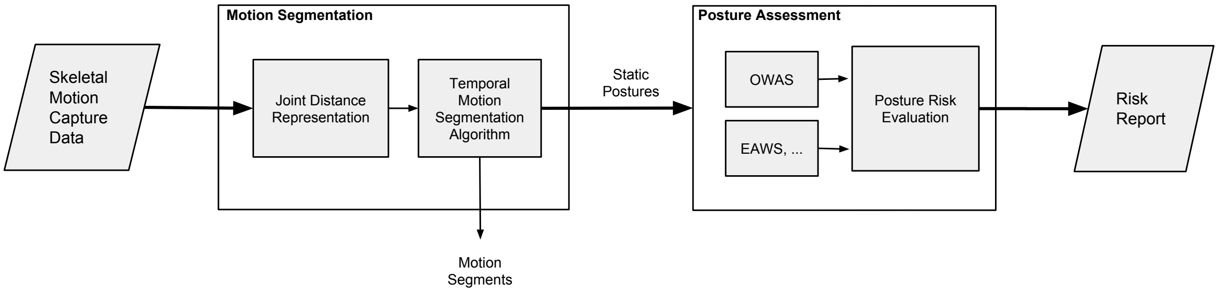

Figure 4: Process of WMSD Risk Factor Identification as enhancement to the model of (Lins et al., 2016).

a static awkward position for several minutes. Such postures should be avoided as they can cause damage to the musculoskeletal system. As a side effect of the segmentation algorithm, the variance of the motion segment σ_1 is known and can be used to detect static postures, because the high variance in the joint distances is caused by rapid movement of limbs. A small variance in the joint distances of consecutive frames can be caused by no or little overall movement of the subject or by continuous movement with very little change over time (e.g. sneaking). As a result, a segment can be identified as static posture by the segmentation algorithm, but its classification as constrained posture must be done by other parts of the assessment system (see *Posture Assessment* in Figure 4).

3.6 Time Complexity

For every input frame that is added to the current window W the *segment* function (see Algorithm 1) is called. The function consists of a loop with nested calculations of mean and variance. The worst-case time complexity of that is $\mathcal{O}((n-2l_{min})(2n+5)+\frac{l_{min}}{c}+2)$, with n being size of the current window, l_{min} minimum motion length and c constant. This can be simplified to $\mathcal{O}(n^2)$. However, the window length is limited by l_{max}, so for a large number of data frames the complexity of the *segment* function is not relevant. As a result the overall time complexity for large n (with n: number of input frames) is $\mathcal{O}(l_{max} \cdot c \cdot n) \rightarrow \mathcal{O}(n)$, which is necessary for an online algorithm working in real-time.

4 EVALUATION OF THE SEGMENTATION ALGORITHM

To evaluate the quality of the algorithm we use the results from the offline algorithms of (Zhou et al., 2013) and (Krüger et al., 2016) as gold standard. Both papers used motion capture data of the publicly available CMU database (Carnegie Mellon University, 2016), in particular takes 01 to 14 from subject 86 (resampled from 120 Hz to 30 Hz). We expect better results from offline algorithms because they can exploit the full temporal spectrum of the data, so they represent the high bar of the results besides the ground truth. See Figure 5 for a comparative overview of the results (our algorithm with $l_{max} = 3.5$).

The algorithms mentioned above differentiate between motion segment and action class: the first can be seen as a more granular motion primitive whereas the latter can be seen as a broader view to activities of a human. The ground truth of the human observers only refers to action classes and as our algorithm does not recognize different action classes but simple segments, we only compare to the ground truth's action classes and do not use the segment boundaries reported by the other algorithms.

The output of the segmentation algorithms is a time-series of segment boundaries. As the algorithms may return time-series of different lengths (because of different numbers of segments), we use Dynamic Time Warping (DTW) (Müller, 2007) to calculate the path with the minimal total cost that is required to warp the time-series of an algorithm to the one of the ground truth. The minimal total cost can be seen as the minimal amount of frames that are required to shift every segment boundary to a ground truth segment boundary.

Be X a time-series length n, Y length m, then the distance DTW of the *optimal warping path* $p*$ is (Müller, 2007):

$$DTW(X,Y) = c_{p*}(X,Y) = \min \sum_{l=1}^{L} c(x_{nl}, y_{ml}) \quad (3)$$

In our case the cost function $c(x,y)$ simply means the distance $|x-y|$. Table 1 shows the normalized DTW distances of our method and two state-of-the-art algorithms to Ground Truth (GT) for every of the 14 trials.

Table 1: Normalized DTW distances of the three algorithms to Ground Truth

Trial	Our method ($l_{min} = 3.5$)	(Krüger et al., 2016)	HACA (Zhou et al., 2013)
01	20.59	11.38	26.68
02	25.21	19.49	30.40
03	39.21	18.42	27.25
04	32.71	17.28	25.88
05	31.00	17.33	15.97
06	65.50	15.13	18.19
07	27.27	10.38	12.48
08	26.06	11.11	21.71
09	44.41	10.90	17.75
10	47.18	25.75	22.00
11	49.15	17.83	17.05
12	53.18	35.83	26.60
13	39.42	32.63	31.35
14	32.62	21.39	21.33
∅	38.04 frames	18.92	22.47

5 DISCUSSION

The results show that the algorithm can continuously segment motion data with robust although not superior quality compared to the offline segmentation algorithms that can fully exploit the whole temporal range of the data. In summary, the DTW distances of our algorithms' results are about twice as high as of the reference algorithms, although the results vary notably throughout the different trials. The quality of a segmentation algorithm is of course highly dependent on the tasks performed by the motion capture subject. It is not surprising that the offline working algorithms deliver results of higher quality, but on the other hand, they have roughly quadratic complexity and require *a priori* knowledge whereas our algorithm has linear complexity and requires very little *a priori* knowledge (maximum window size). In practice, the average DTW distance of our algorithm means that the segment is 1-2 seconds away from the ground truth, which is sufficient if the motion segment is significantly longer. As a plus, our method returns the motion dynamics variance of the segmented motion which is used as static posture indicator, i.e., if the algorithm returns a segment with low variance we assume a static posture that can be further processed by an ergonomic assessment method. In summary, the method is a practically usable approach that is easy to understand and implement although its accuracy falls behind the state-of-the-art offline methods.

6 CONCLUSION

We described and evaluated an online-capable (real-time) temporal segmentation algorithm for skeletal motion data. The algorithm can be used to detect static postures in a continuous stream of skeletal motion capture data. Together with a digitalized ergonomic assessment method such as OWAS (Karhu et al., 1977) the detected postures can be used to derive risk factors for (work-related) musculoskeletal disorders. Due to the online capabilities of the algorithm, it is possible to implement a live feedback system for users of MoCap suits/systems when they perform unfavorable or dangerous postures. (Yan et al., 2017; Ray and Teizer, 2012; Peppoloni et al., 2014) are examples for such systems and could possibly be used with our algorithm. Because of its simple computability, the algorithm works well on embedded hardware, since only simple floating point calculations are necessary.

ACKNOWLEDGEMENTS

This work was partly funded by the German Ministry for Education and Research (BMBF) within the joint research projects SIRKA (grant 16SV6243) and MeSiB (grant 16SV7723). The authors would like to thank Anna Vögele (Vögele et al., 2014) for kindly providing the reference data of Table 1 and Figure 5.

REFERENCES

Aggarwal, J. K. and Cai, Q. (1997). Human motion analysis: A review. In *IEEE Nonrigid and Articulated Motion Workshop Proceedings*, pages 90–102, Washington, DC, USA. IEEE, IEEE Computer Society.

Amell, T. and Kumar, S. (2001). Work-related musculoskeletal disorders: Design as a prevention strategy. a review. *Journal of Occupational Rehabilitation*, 11(4):255–265.

Armstrong, T., Buckle, P., Fine, L., Hagberg, M., Haring-Sweeney, M., Martin, B., Punnett, L., Silverstein, B., Sjøgaard, G., Theorell, T., et al. (1996). Musculoskeletal disorders: Work-related risk factors and prevention. *International Journal of Occupational and Environmental Health*, 2(3):239–246.

Carnegie Mellon University (2016). CMU Graphics Lab Motion Capture Database.

Corlett, E. N. (2003). Rapid Upper Limb Assessment (RULA). *Occupational Ergonomics: Principles of Work Design*, 1(June):9.

Coyte, P. C., Asche, C. V., Croxford, R., and Chan, B. (1998). The economic cost of musculoskeletal disorders in canada. *Arthritis & Rheumatism*, 11(5):315–325.

Ellegast, R. (2013). Gefährdungsbeurteilung am Arbeitsplatz. In Hartmann, Spallek, and Ellegast, editors, *Arbeitsbezogene Muskel-Skelett-Erkrankungen: Ursachen, Prävention, Ergonomie, Rehabilitation*, pages 107–130. Hüthig Jehle Rehm.

Hignett, S. and McAtamney, L. (2000). Rapid Entire Body Assessment (REBA).

Hoy, D., Brooks, P., Blyth, F., and Buchbinder, R. (2010). The epidemiology of low back pain. *Best practice & research Clinical rheumatology*, 24(6):769–781.

Karhu, O., Kansi, P., and Kuorinka, I. (1977). Correcting working postures in industry: A practical method for analysis. *Applied Ergonomics*, 8(4):199–201.

Keogh, E., Chu, S., Hart, D., and Pazzani, M. (2001). An online algorithm for segmenting time series. In *Proceedings 2001 IEEE International Conference on Data Mining*, pages 289–296. IEEE.

Koenig, N. and Matarić, M. J. (2006). Behavior-based segmentation of demonstrated tasks. In *Proceedings of the International Conference on Development and Learning*. Unknown.

Krüger, B., Vögele, A., Willig, T., Yao, A., Klein, R., and Weber, A. (2016). Efficient unsupervised temporal segmentation of motion data. *IEEE Transactions on Multimedia*, PP(99):1–1.

Kulic, D., Takano, W., and Nakamura, Y. (2009). Online segmentation and clustering from continuous observation of whole body motions. *IEEE Transactions on Robotics*, 25(5):1158–1166.

Lins, C., Eichelberg, M., Rölker-Denker, L., and Hein, A. (2015). SIRKA: Sensoranzug zur individuellen Rückmeldung körperlicher Aktivität. In *55. Wissenschaftliche Jahrestagung 2015 der Deutsche Gesellschaft für Arbeitsmedizin und Umweltmedizin e.V., München*, pages 301–303. Deutsche Gesellschaft für Arbeitsmedizin und Umweltmedizin (DGAUM) e.V.

Lins, C., Müller, S. M., and Hein, A. (2016). Model-based approach for posture and movement classification in working environments. In Wichert, R. and Klausing, H., editors, *Ambient Assisted Living: 8. AAL-Kongress 2015,Frankfurt/M, April 29-30. April, 2015*, pages 25–33. Springer International Publishing, Cham.

Matsui, H., Maeda, A., Tsuji, H., and Naruse, Y. (1997). Risk indicators of low back pain among workers in japan: association of familial and physical factors with low back pain. *Spine*, 22(11):1242–1247.

Müller, M. (2007). Dynamic time warping. In *Information Retrieval for Music and Motion*, pages 69–84. Springer Berlin Heidelberg, Berlin, Heidelberg.

Nath, N. D., Akhavian, R., and Behzadan, A. H. (2017). Ergonomic analysis of construction worker's body postures using wearable mobile sensors. *Applied Ergonomics*, 62.

Peppoloni, L., Filippeschi, A., Ruffaldi, E., and Avizzano, C. A. (2014). (WMSDs issue) A novel wearable system for the online assessment of risk for biomechanical load in repetitive efforts. *International Journal of Industrial Ergonomics*, 52:1–11.

Punnett, L. and Wegman, D. H. (2004). Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. *Journal of Electromyography and Kinesiology*, 14(1):13–23.

Ray, S. J. and Teizer, J. (2012). Real-time construction worker posture analysis for ergonomics training. *Advanced Engineering Informatics*, 26(2):439–455.

Roetenberg, D., Luinge, H., and Slycke, P. (2013). Xsens MVN : Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors. Technical report, Xsens Technologies B.V.

Schaub, K., Caragnano, G., Britzke, B., and Bruder, R. (2013). The European Assembly Worksheet. *Theoretical Issues in Ergonomics Science*, 14(6):616–639.

Stollenwerk, K., Vögele, A., Krüger, B., Hinkenjann, A., and Klein, R. (2016). Automatic temporal segmentation of articulated hand motion. In *International Conference on Computational Science and Its Applications*, pages 433–449. Springer.

Vieira, A. W., Lewiner, T., Schwartz, W. R., and Campos, M. (2012). Distance matrices as invariant features for classifying mocap data. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 2934–2937. IEEE.

Vögele, A., Krüger, B., and Klein, R. (2014). Efficient unsupervised temporal segmentation of human motion. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '14, pages 167–176, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

Walker, B., Muller, R., and Grant, W. (2003). Low back pain in australian adults: the economic burden. *Asia Pacific Journal of Public Health*, 15(2):79–87.

Wang, D., Dai, F., and Ning, X. (2015). Risk assessment of work-related musculoskeletal disorders in construction: State-of-the-art review. *Journal of Construction Engineering and Management*, 141(6):04015008.

Wang, L., Hu, W., and Tan, T. (2003). Recent developments in human motion analysis. *Pattern Recognition*, 36(3):585 – 601.

Yan, X., Li, H., Li, A. R., and Zhang, H. (2017). Wearable IMU-based real-time motion warning system for construction workers' musculoskeletal disorders prevention. *Automation in Construction*, 74:2–11.

Zhou, F., De la Torre, F., and Hodgins, J. K. (2008). Aligned cluster analysis for temporal segmentation of human motion. In *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on*, pages 1–7. IEEE.

Zhou, F., De la Torre, F., and Hodgins, J. K. (2013). Hierarchical aligned cluster analysis for temporal clustering of human motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(3):582–596.

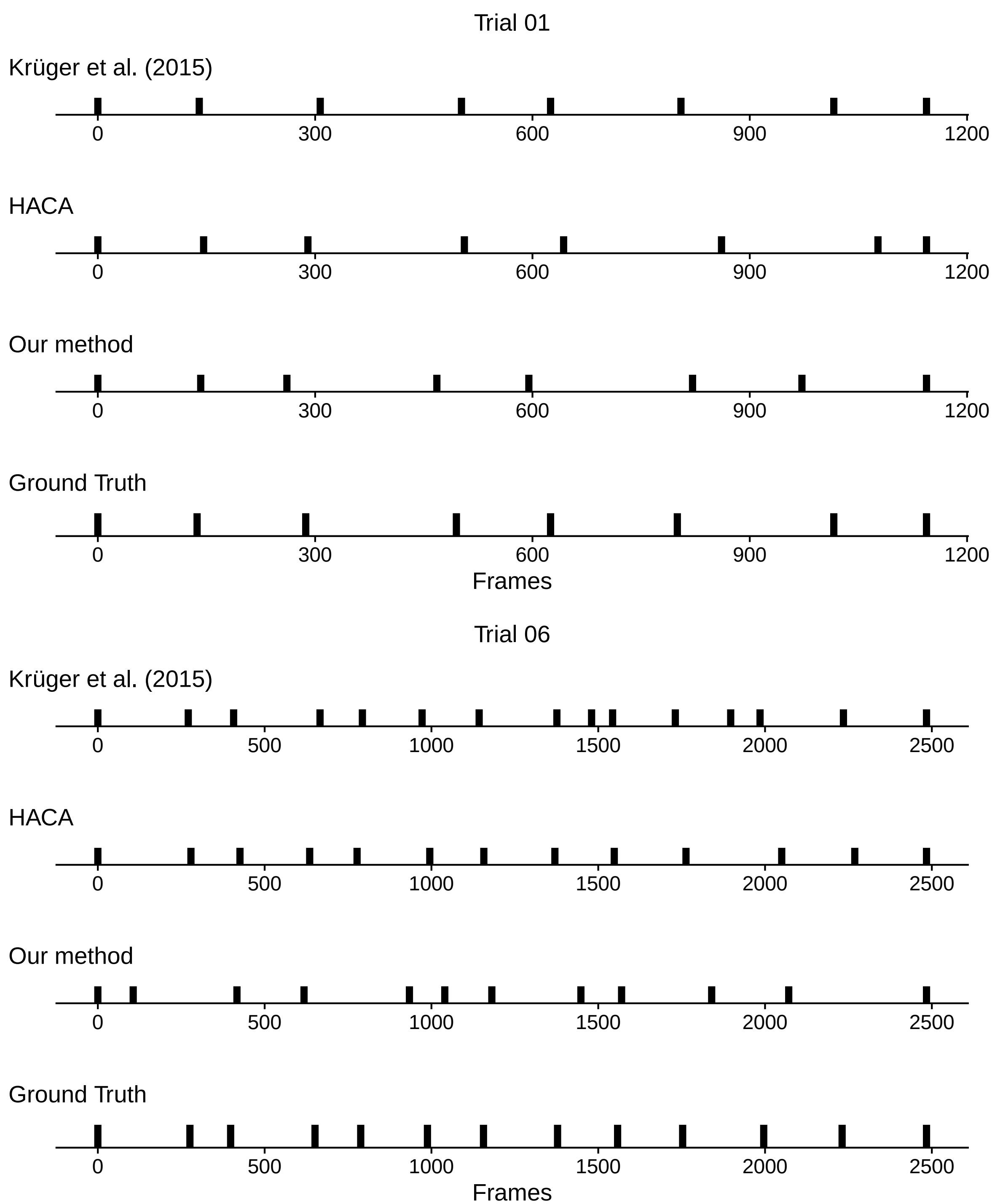

Figure 5: Graphical comparison of three segmentation algorithms with ground truth data for trials 1 (best case) and 6 (worst case) of subject 86. Black bars mark the segmentation points. In case of (Krüger et al., 2016) the center of uncertainty was chosen as the segmentation point.