

Automatic Detection and Recognition of Human Movement Patterns in Manipulation Tasks

Lisa Gutzeit¹ and Elsa Andrea Kirchner^{2,1}

¹AG Robotik, Universität Bremen, Robert-Hooke-Str.1, 28359 Bremen, Germany

²German Research Center for Artificial Intelligence (DFKI), Robotics Innovation Center, Robert-Hooke-Str.1, 28359 Bremen, Germany

Keywords: Human Movement Analysis, Behavior Segmentation, Behavior Recognition, Manipulation, Motion Tracking.

Abstract: Understanding human behavior is an active research area which plays an important role in robotic learning and human-computer interaction. The identification and recognition of behaviors is important in learning from demonstration scenarios to determine behavior sequences that should be learned by the system. Furthermore, behaviors need to be identified which are already available to the system and therefore do not need to be learned. Beside this, the determination of the current state of a human is needed in interaction tasks in order that a system can react to the human in an appropriate way. In this paper, characteristic movement patterns in human manipulation behavior are identified by decomposing the movement into its elementary building blocks using a fully automatic segmentation algorithm. Afterwards, the identified movement segments are assigned to known behaviors using k-Nearest Neighbor classification. The proposed approach is applied to pick-and-place and ball-throwing movements recorded by using a motion tracking system. It is shown that the proposed classification method outperforms the widely used Hidden Markov Model-based approaches in case of a small number of labeled training examples which considerably minimizes manual efforts.

1 INTRODUCTION

In future, robots and humans must interact very closely and even physically to satisfy the requirements of novel approaches in industry, production, personal services, health care, or medical applications. To facilitate this, not only the robotic systems must be equipped with enlarged dexterities and mechanisms that allow intuitive and safe interaction, but also the human intention, behavior and habits have to be better understood (Kirchner et al., 2015). To allow this, novel methods have to be developed which are easy to apply.

One highly relevant factor in human-computer interaction is an understanding of human behaviors. For example, the knowledge of the current state of the human is necessary to realize an intuitive interaction. Based on this knowledge, systems can interact with humans in an appropriate manner. To obtain this knowledge, the identification of the important parts of the human behavior and the assignment of the identified behaviors into categories which induce different reactions of the system are necessary. Only if the state of the human and the context which is described by

this state are known, the system can follow the working steps that are required in this situation or can support the human if desired.

Another example is imitation of human behaviors by a robotic system which is a current issue in robot learning approaches and has intensively been investigated, see for example (Metzen et al., 2013; Mülling et al., 2013; Pastor et al., 2009). Especially, Learning from Demonstration (LfD) is a relevant issue in this research area, in which learning algorithms are used to transfer human demonstrations of behavior to a robot (Argall et al., 2009). Because learning of complex behavior can be very time-consuming or even impossible, the behavior should be segmented into its main building blocks to be learned more efficiently. By grouping segments that belong to the same behavior and by recognizing these behaviors, it can be determined which segments are needed to be learned for a certain situation. Beyond that, movements can be identified that can already be executed by the system and thus do not need to be learned.

The hypothesis of the composition of human movement into building blocks is shown in several behavioral studies, e.g., in a study on infants (Adi-

Japha et al., 2008). These studies show that complex human behaviors are learned incrementally, starting with simple individual building blocks that are chunked together to a more complex behavior (Graybiel, 1998). If these building blocks should be detected by an artificial system, characteristics in the movement patterns have to be identified. In manipulation behaviors, bell-shaped velocity profiles have found to be a suitable pattern (Morasso, 1981). In this work, a velocity-based behavior segmentation algorithm presented by Senger et al. (Senger et al., 2014) is used to segment recorded human movement. The applied algorithm detects reliably and fully automatic movement sequences that show a bell-shaped velocity profile and are therefore assumed to be building blocks of human behavior.

As stated above, identified building blocks of human movement have also to be classified according to the actual behavior they belong to. By assigning suitable annotations to the recognized movement classes, the selection as well as the detection of the required behavior becomes intuitive and easy to use in different interaction scenarios. For supervised movement classification approaches, training data is needed that has to be manually pre-labeled. To keep the manual input low, it is desirable that the classification works with small sets of training data. We propose to classify detected building blocks by using simple k-Nearest Neighbor (kNN) classification. With suitable features extracted from the movements, kNN satisfies this condition.

This paper is organized as follows: In Section 2, different state-of-the-art approaches for segmentation and recognition of human movements are summarized. Our approach is described in Section 3. Afterwards in Section 4, the approach is evaluated on real human manipulation movements and compared to Hidden Markov Model (HMM)-based approaches which are widely used in the literature to represent and recognize movements. At the end of this paper, a conclusion is given.

2 RELATED WORK

Action recognition is an active research area which plays an important role in many applications. One main focus lies in the automatic annotation of human movements in videos, which can be used, e.g., to find tackles in soccer games, to support elderly in their homes or for gesture recognition in, e.g. video games (Poppe, 2010). Besides the detection of humans in video sequences, the classification of their movements is an important part in video-based ac-

tion recognition. Algorithms like Support Vector Machines, or their probabilistic variant the Relevance Vector Machines, Hidden Markov Models, k-Nearest Neighbors or Dynamic Time Warping-based classification are used to classify the observed actions. A more detailed overview is given in (Poppe, 2010).

But also in other areas, where the human is not observed by a camera but recorded with other modalities like markers fixed on the body, human action recognition is tackled. In this non-image-based movement recordings, the segmentation of the recorded movements is next to the classification of high interest. For example in (Fod et al., 2002), human arm movements were tracked and segmented into so-called movement primitives at time points where the angular velocity of a certain number of degrees of freedom crosses zero. After a PCA-based dimensionality reduction, the identified movements were clustered using k-Means. However, this approach is very sensitive to noise in the input data which results in over-segmentation of the data. Gong et al., on the other hand, propose Kernelized Temporal Cut to segment full body motions, which is based on Hilbert space embedding of distributions (Gong et al., 2013). In their work, different actions are recognized using Dynamic Manifold Warping as similarity measure. In contrast to the analysis of full body motions, we focus on the identification and recognition of manipulation movements which show special patterns in the velocity which should be considered for segmentation.

Beyond that, HMM-based approaches are often used in the literature, both for movement segmentation as well as for movement recognition. For example, Kulic et al. stochastically determine motion segments which are then represented using HMMs (Kulić et al., 2012). The derived segments are incrementally clustered using a tree structure and the Kullback-Leibler distance as segment distance measure. In a similar fashion, Gräve and Behnke represent probabilistically derived segments with HMMs, where segments that belong to the same movement are simultaneously classified into the same class if they can be represented by the same HMM (Gräve and Behnke, 2012). Besides these approaches, solely training-based movement classification with HMMs is widely used, e.g. in (Stefanov et al., 2010; Aarno and Kragic, 2008). Because HMMs are expected to perform not well when few training data is available, we propose to use kNN instead and compare it with the HMM approach.

3 METHODS

In this section we describe the velocity-based movement segmentation algorithm to identify building blocks in human manipulation behavior as well as our approach to recognize different known movement segments in an observed behavior.

3.1 Segmentation of Human Movement into Building Blocks

We aim to find sequences in human movement that correspond to elementary building blocks characterized by bell-shaped velocity profiles as shown in (Morasso, 1981). Therefore, we need a segmentation algorithm that identifies these building blocks. A second important property of the algorithm should be the ability to handle variations in the movements. Human movement shows a lot of variations both during the execution by different persons as well as by the same person. For this reason, it is important that the algorithm for human movement segmentation finds sequences that correspond to the same behavior despite differences in their execution.

An algorithm that tackles these issues is the velocity-based Multiple Change-point Inference (vMCI) algorithm (Senger et al., 2014). This algorithm fully automatically detects building blocks in human manipulation movements. It is based on the Multiple Change-point Inference (MCI) algorithm (Fearnhead and Liu, 2007) in which segments are found in time series data using Bayesian Inference. Each segment $y_{i+1:j}$, starting at time point i and ending at j , is represented with a linear regression model (LRM) with q predefined basis functions ϕ_k :

$$y_{i+1:j} = \sum_{k=1}^q \beta_k \phi_k + \varepsilon, \quad (1)$$

where ε models the noise that is assumed in the data and $\beta = (\beta_1, \dots, \beta_q)$ are the model parameters. It is assumed that a new segment starts if the underlying LRM changes. This modeling of the observed data allows to handle technical noise in the data as well as variation in the execution of the same movement. To determine the segments online, the segmentation points are modeled via a Markov process in order that an online Viterbi algorithm can be used to determine their positions (Fearnhead and Liu, 2007).

Senger et al. expanded the MCI algorithm for the detection of movement sequences that correspond to building blocks characterized by a bell-shaped velocity profile. To realize this, the LRM of Equation 1 is split to model the velocity of the hand independent from its position with different basis functions, where

the basis function for the velocity dimension is chosen in a way that it has a bell-shaped profile. In detail this means that the velocity y^v of the observed data sequence is modeled by

$$y^v = \alpha_1 \phi_v + \alpha_2 + \varepsilon, \quad (2)$$

with weights $\alpha = (\alpha_1, \alpha_2)$ and noise ε . The model has two basis functions. First, the bell-shaped velocity curve is modeled using a single radial basis function:

$$\phi_v(x_t) = \exp \left\{ -\frac{(c-x_t)^2}{r^2} \right\}. \quad (3)$$

In order that the basis function can cover the whole segment, Senger et al. propose to choose half of the segment length for the width parameter r . The center c is determined automatically by the algorithm and regulates the alignment to velocity curves with peaks at different positions. Additionally, the basis function 1 weighted with α_2 accounts for velocities unequal to zero at start or end of the segment. As in the original MCI method, an online Viterbi algorithm can be used to detect the segment borders.

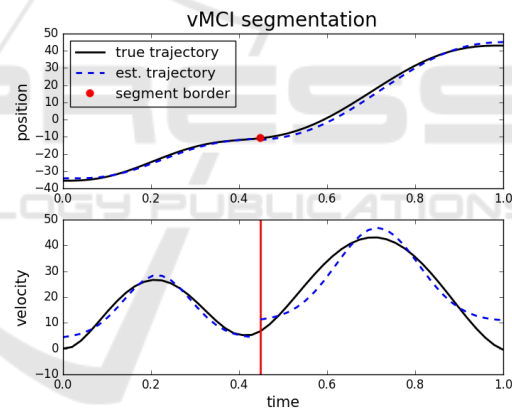


Figure 1: VMCI segmentation result on artificial data.

An example segmentation using the vMCI algorithm is shown in Figure 1. At the top, a one-dimensional simulated movement can be seen. The lower figure shows the corresponding velocity. To simulate two different behavior segments, the movement is slowed down at time point 0.4. For the position dimension, the algorithm fits LRMs to the data according to Equation 1 with pre-defined basis functions. In this case autoregressive basis functions are chosen. The velocity dimension is simultaneously fit with a LRM as introduced in Equation 2. The algorithm automatically selects the models which best fits parts of the data. In this case it is most likely that the data arises from two different underlying models, which results in a single segmentation point which

matches, within an acceptable margin, the true segmentation point. In contrast to other segmentation algorithms, as for example a segmentation based on the detection of local minima, vMCI is very robust against noise in the data, as shown in (Senger et al., 2014).

3.2 Recognition of Human Movement

There are many different possibilities to classify human movements, as reviewed in Section 2. In general, a movement classification algorithm which works with minimal need for parameter tuning is desirable to make the classification easily applicable on different data. Furthermore, manual efforts can be minimized if the algorithm reliably classifies movement segments in case that only a small training set is available. For this reasons we use the kNN classifier for movement recognition. It has only one parameter, k , and is able to classify manipulation movements with a high accuracy with a small training set, as shown in our experiments.

3.2.1 Feature Extraction

To classify the obtained movement sequences, features which reflect the differences between different behaviors have to be calculated. We use movement trajectories of certain positions on the demonstrator as features for the classification. The movements are recorded in Cartesian coordinates which results in different time series if the same movement is executed at a different position. Thus, we propose to transpose the data into a coordinate system which is not global but relative to the human demonstrator. As reference point, we use the position of the back (see Figure 2A) at the first time point of a segment, i.e. the data is transformed into a coordinate system centered at this point. Additionally, variances in the execution of the same movement are reduced by normalizing each movement segment to zero mean.

Next to the transformed and normalized tracking points of the demonstrator, additional features may be relevant to successfully classify movement segments. For example in manipulation movements, such as pick-and-place tasks, the involved objects and their spatial reference to the demonstrator are important features to distinguish between movement classes. Thus, the distance of the human hand to the manipulated object as well as the object speed are used in the pick-and-place experiment described in section 4.2 to classify manipulation segments into distinct movements. Depending on the recognition task additional features, like the rotation of the hand to distinguish between different grasping positions, can be relevant.

3.2.2 Movement Classification

We propose to use a kNN classifier to distinguish between different movements. In the kNN classification, an observed movement sequence is assigned to the movement class, which is the most common among its k closest neighbors of the training examples. We use the Euclidean distance as distance metric and account for segments of unequal length by applying an interpolation to bring all segments to the mean segment length. Alternatively, dynamic time warping (DTW) could be used as distance measure. This would have the benefit that using DTW the segments are additionally aligned to the same length. But in a preliminary analysis of kNN classification on manipulation behaviors our approach outperformed a DTW-based kNN. For the number of neighbors k , we take $k = 1$. That means we consider just the closest neighbor for classification because we want to classify with a small number of training examples. A bigger k could result in more classification errors due to the very low number of examples of each class.

4 EXPERIMENTS

In this section, the proposed segmentation and classification methods are evaluated on real human manipulation movements tracked by using a motion capturing system. First, the experimental setup including the evaluation technique used in two different experiments is described in section 4.1. Afterwards, the presented approach is applied and evaluated on pick-and-place movements. In a second experiment different human demonstrations of a ball-throwing movement are analyzed. For both experiments it is shown that the vMCI algorithm correctly detects segments in the recorded demonstrations which correspond to behavior building blocks with a bell-shaped velocity pattern. Furthermore, we evaluate the classification with kNN using small number of training data and compare the results with an HMM-based classification approach.

4.1 Experimental Setup

In the experiments conducted for this paper, human demonstrations of manipulation movements were tracked using 7 motion capture cameras. The motion capturing system measures the 3D positions of visual markers at a frequency of 500Hz , which was down-sampled to 25Hz . The markers were placed on the human demonstrator and in the pick-and-place experiment additionally on the manipulated object.

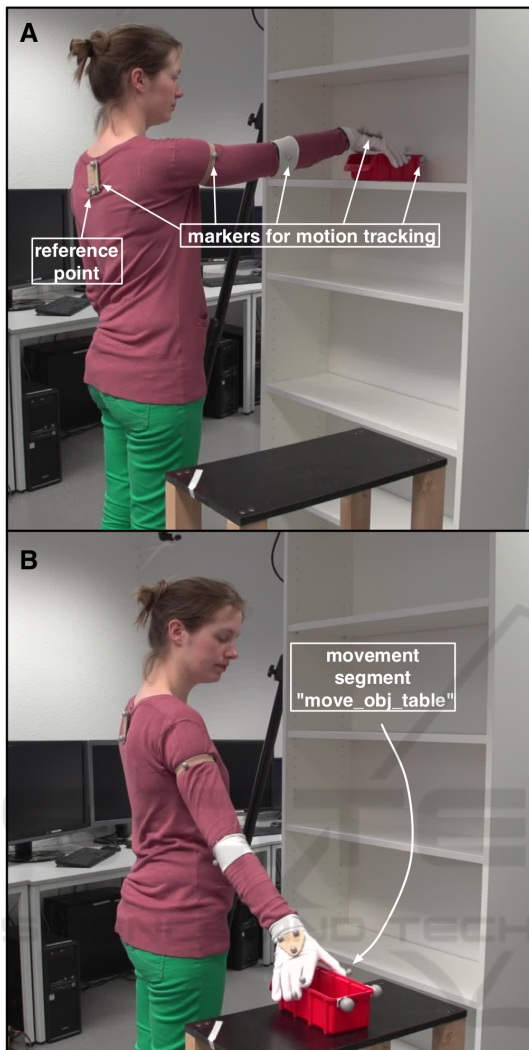


Figure 2: Snapshots of the pick-and-place task analyzed in this work. A: Markers for movement tracking are placed at the back, the arm and the hand of the demonstrator as well as on the manipulated object. The images show the grasping of the object from the shelf (A) which is then placed on a table standing on the right hand side (B). B: Movement segment `move_obj_table` is sketched.

The positions of the markers can be seen in Figure 2A and Figure 3. Three markers were placed on the back of the demonstrator to determine the position of the back and its orientation. This is used to transform the recorded data into the coordinate system relative to the back, as described in Section 3.2. To track the movement of the manipulating arm, markers were placed at the shoulder, the elbow, and the back of the hand. The orientation of the hand is determined by placing three markers instead of one on it. Grasping movements in the pick-and-place demonstrations were recorded by using additional markers which were placed at thumb, index, and middle fin-

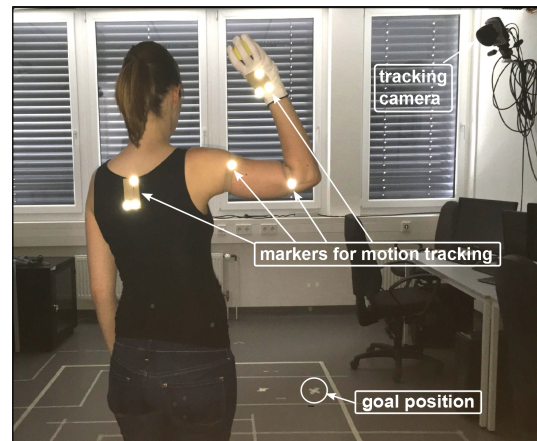


Figure 3: Snapshot of the ball-throwing task.

ger. Furthermore, two more markers were placed on the manipulated object in this experiment to determine its position and orientation. However, the tasks in our experiments required only basic manipulation movements (e.g., approaching the object or moving the object). Thus, just the position of the hand and the manipulated object were used for segmentation and recognition, but not their orientation.

After data acquisition, the individual movement parts of the demonstrations were identified using the vMCI algorithm described in Section 3.1. The segmentation algorithm was applied on the position and the velocity of the recorded hand movements. For this, the recorded positions of each demonstration were pre-processed to a zero mean and such that the variance of the first order differences of each dimension is equal to one, as proposed in (Senger et al., 2014).

The resulting movement segments were manually labeled into one of the movement classes defined for each experiment. However, some of the obtained segments could not be assigned to one of the movement classes because they contain only parts of the movement. This could result from errors in the segmentation as well as from demonstrations where a movement is slowed down before the movement class ends, e.g. because the subject thought about the exact position to grasp the object. An example can be seen in the top plot of Figure 4. The concatenation of the first two detected segments belong to the class `approach_forward`. Nonetheless, the vMCI algorithm detected two segments because the subject slowed down the movement right before reaching the object. These incomplete movement segments were discarded for the evaluation of the classification approach. Furthermore, some of the identified movement segments do not belong to one of the pre-defined movement classes of the experiment. Usually, these

nonassignable segments belong to small extra movements, that are not part of the main movement task and thus are not considered in the defined movement classes. These movement segments were as well not used for the evaluation of the classification.

Before classification, the original recorded marker positions of each obtained segment were pre-processed as described in Section 3.2. Depending on the manipulation task, additional features were calculated. As proposed in section 3.2, we classify the obtained segments using the 1NN algorithm. For each of the two experiments, we evaluate the accuracy of the 1NN classification using a stratified 2-fold cross-validation with a fixed number of examples per class in the training data. The training set sizes are varied from 1 example per class to 20 examples per class and the remaining data is used for testing. Since we want to show the performance of the classification with small training set sizes, the maximal number of training examples per class is kept low. For each number of examples per class in the training data, the cross-validation was performed with 100 iterations.

For comparison, the data was also classified using a HMM-based approach, which is a standard representation method for movements in the literature, see Section 2. In the HMM-based classification, one single HMM was trained for each movement class. To classify a test segment, the probability of the segment to be generated by each of the trained HMMs is calculated. The label of the most likely underlying HMM is assigned to the segment. The number of states in the HMMs was determined with a stratified 2-fold cross-validation repeated 50 times with equally sized training and test sets. As a result, we trained each HMM with one hidden state. The accuracy of the HMM-based classification with 1 hidden state per trained HMM was evaluated like the 1NN classification with a stratified 2-fold cross-validation with fixed numbers of training examples for each class.

4.2 Segmentation and Recognition of Pick-and-Place Movements

In our first experiment, we evaluated the presented approach on pick-and-place movements. The task of the human demonstrator, partly shown in Figure 2, contained 6 different movements. First, a box placed on a shelf should be grasped (movement class: `approach_forward`) and placed on a table standing at the right hand side of the demonstrator (`move_obj_table`). After reaching a rest position of the hand (`move_to_rest_right`), the object had to be grasped again from the table (`approach_right`) to move it back to the shelf (`move_obj_shelf`). At the

end, the arm should be moved into a final position in which it loosely hangs down (`move_to_rest_down`). Beyond that, short periods of time in which the demonstrator did not move his arm can be assigned to the class `idle`.

Overall, the pick-and-place task was performed by three different subjects, repeated 6 times by each. Two of these subjects performed the task again with 4 repetitions while their movements were recorded with slightly different camera positions and a different global coordinate system. This resulted in different positions of the person and the manipulating object in the scene which should be handled by the presented movement segmentation and recognition methods. Thus, 26 different demonstrations from different subjects and with varying coordinate systems were available to evaluate the proposed approaches.

4.2.1 Results

The vMCI algorithm successfully segmented the trajectories of the pick-and-place demonstrations into movement parts with a bell-shaped velocity profile. Three examples of the segmentation results can be seen in Figure 4. The resulting movement segments were manually labeled into one of the 7 movement classes that are present in the pick-and-place task. This resulted in 155 labeled movement segments with different occurrences of each class, as summarized in Table 1.

Table 1: Occurrences of each class in the recorded pick-and-place data.

movement class	num. examples
<code>approach_forward</code>	20
<code>move_obj_table</code>	26
<code>move_to_rest_right</code>	25
<code>approach_right</code>	23
<code>move_obj_shelf</code>	26
<code>move_to_rest_down</code>	24
<code>idle</code>	11

As described in section 4.1, next to the positions of the markers attached on the subject, the distance from the hand to the object and the object velocity were calculated as additional features in this experiment. An example result of the classification using 1NN is shown in Figure 5. For this example demonstration of the pick-and-place task, all segments have been labeled with the correct annotation using a training set with 5 examples for each class.

The results of the cross-validation using 1NN and HMM-based classification are shown in Figure 6. Because the data contains 7 different classes, an accuracy of 14.3% can be achieved by guessing. The

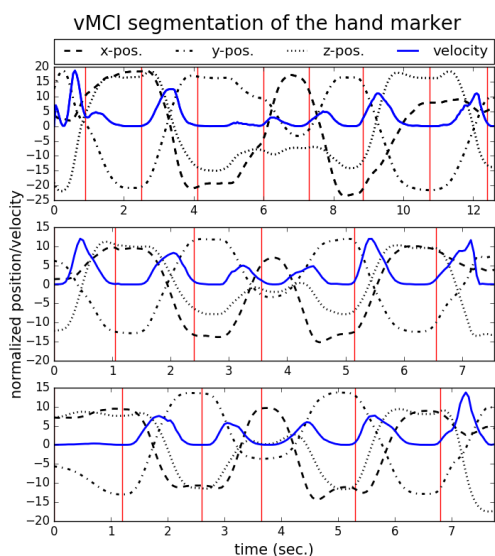


Figure 4: Segmentation results of three different demonstrations. The black lines are the x-, y- and z-position of the hand. The blue line corresponds to the velocity of the hand and the red vertical lines are the segment borders determined by the vMCI algorithm.

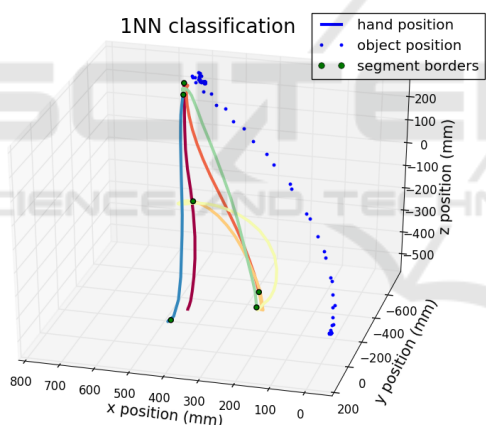


Figure 5: Classification result of a demonstration of the pick-and-place task with 1NN. The different movement classes of the task are indicated with different colors along the color spectrum starting with red for approach_forward and ending with blue for move_to_rest_down.

1NN classification clearly outperforms the HMM-based classification using training sets with occurrences of each class smaller or equal to 20. Already with 1 example per class an accuracy of nearly 80% can be achieved using 1NN. With 10 examples per class, the accuracy is 97.5% and with 20 examples per class 99.2%, which is very close to an error-less classification. In contrast, 14 examples per class are needed in the HMM-based classification to achieve an accuracy of 90% in this evaluation. With not more than 10 examples per class, the accuracy of the

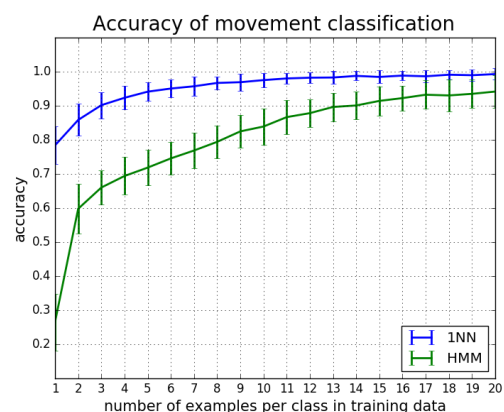


Figure 6: Comparison of the accuracy of the classification of manipulation movement segments using 1NN and HMM-based classification.

HMM-based classification is considerably below the achieved accuracy using 1NN.

These results show that with the proposed 1NN classification, manipulation movements can be assigned to known movement classes with a very small number of training examples. This means that with minimal need for manual training data labeling and no parameter tuning, very good classification results can be achieved using the proposed approach. Furthermore, the 1NN classification considerably outperforms the widely used HMM-based classification in case that only a small number of training examples is available.

4.3 Segmentation and Recognition of Ball-Throwing Movements

In a second experiment, the vMCI segmentation and the 1NN classification were evaluated on ball-throwing demonstrations. The task of the subject was to throw a ball to a goal position on the ground located approximately 1.5 m away. The numerous possibilities to throw the ball were limited by the restriction that the ball should be thrown from above, i.e. the hand has a position higher than the shoulder before the ball leaves the hand, see Figure 3. Before and after the throw, the subject had to move into a rest position, in which the arm loosely hangs down. The individual movement parts of each throw could be divided into four different main classes: strike_out, throw, swing_out and idle. In contrast to the pick-and-place task, only the movement of the arm was tracked in this task and not the position of the involved object, the ball. This is because in this experiment, the spatial distance of the ball to the demonstrator plays only a minor role and the movement of the arm has a much higher relevance to distinguish between move-

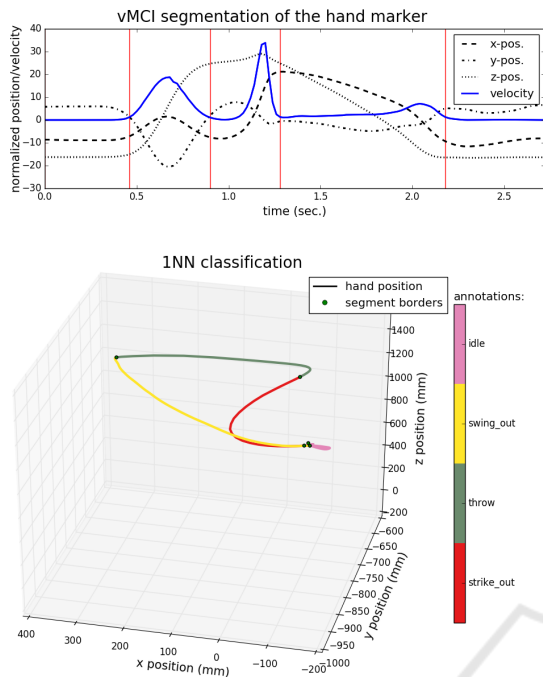


Figure 7: Segmentation and classification result of one demonstration of the ball-throwing task.

ment classes. Furthermore, it was not recorded if the goal position was actually hit by the ball.

The ball-throwing task was demonstrated by 10 different subjects, each performing 24 throws.

4.3.1 Results

As already shown in a similar ball-throwing experiment in (Senger et al., 2014), the vMCI algorithm is able to identify the different movement parts in the demonstrations based on the position and velocity of the hand. In Figure 7 a representative example of the segmentation result is shown.

To evaluate the classification of the ball-throwing movements, the resulting segments of all 240 demonstration were manually assigned to one of the four movement classes. Again, each class had a different occurrence in the available data, as summarized in Table 2.

Table 2: Occurences of each class in the ball-throwing data.

movement class	num. examples
strike_out	221
throw	227
swing_out	339
idle	208

In this experiment, only the positions of the markers attached to the subject, see Figure 3, were used

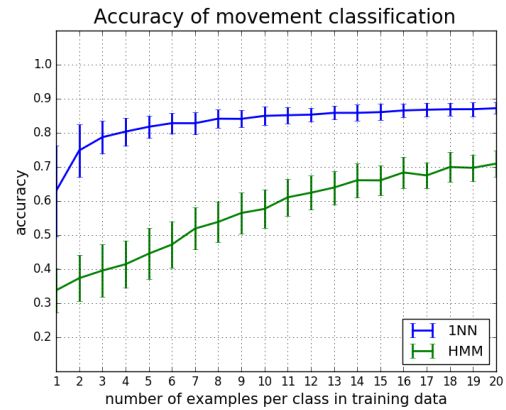


Figure 8: Comparison of the accuracy of the classification of ball throwing segments using 1NN and HMM-based classification.

as features for the automatic movement classification. Figure 7 shows an example classification result using 1NN and 5 examples per class in the training data. The 5 movement segments were correctly classified into one of the predefined classes.

The results of the cross-validation comparing 1NN with HMM-based classification are visualized in Figure 8. Like in the pick-and-place experiment, 1NN outperforms HMM-based classification in the case of small training data sets. In this experiment, containing considerably more demonstrated movements compared to the pick-and-place experiment, the difference between the classification algorithms is even more clear. With one example per class in the training data, an accuracy of 62.9% using 1NN can be achieved and only 33.8% by using HMM-based classification. This experiment contains 4 different classes, i.e. an accuracy of 25% can be achieved by guessing. Using 1NN, a classification accuracy of 80% is accomplished using 4 examples per class during training. In contrast to this, this accuracy is not reached using HMM-based classification in this evaluation. For comparison, the evaluation was additionally conducted using 100 examples per class during training. This resulted in an accuracy of 91.5% using 1NN, and 77.8% using HMM-based classification. This shows that even if more training data is available, the 1NN classification outperforms the HMM-based approach.

5 CONCLUSIONS

In this paper, we identified and recognized characteristic movement patterns in human manipulation behavior. We successfully segmented pick-and-

place and ball-throwing data into movement building blocks with a bell-shaped velocity profile using a probabilistic algorithm formerly presented in (Senger et al., 2014). Furthermore, we showed that using the simple 1NN classification, the obtained segments can be reliably classified into predefined categories. Especially, this can be done using a small set of training data. In comparison to HMM-based movement classification, a considerably higher accuracy can be achieved with small training sets.

For future work, an integrated algorithm for segmentation and classification should be developed, in which both motion analysis parts influence each other. Such an approach becomes for example relevant when extra segments are generated. Extra segments may be caused from not fluently executed movements of the demonstrator in situations in which he slowed down his movement to think about the exact position to place an object. Such extra segments could be merged by identifying that only their concatenation belongs to one of the known movement classes.

To gain a higher classification accuracy, more sophisticated feature extraction techniques may be of high interest. Mainly in the analysis of manipulation movements, features based on the joint angles should be evaluated.

Furthermore, it is desirable that the manual effort needed for classification is further minimized by classifying the movement segments using an unsupervised approach. Nonetheless, annotations, like `move_object`, are needed in many applications, e.g. to select segments that should be imitated by a robot. Ideally, this annotation is done without manual interference, e.g., by analyzing features of the movement arising from different modalities. Besides the analysis of motion data, psychological data like eye-tracking or electroencephalographic-data could be used for this annotation.

Simple approaches as the one presented here become highly relevant for the development of embedded multimodal interfaces. They allow to use miniaturized processing units with relatively low processing power and energy consumption. This is most relevant since in many robotic applications extra resources for interfacing are limited and will thus restrict the integration of interfaces into a robotic system. On the other hand, wearable assisting devices are also limited in size, energy and computing power. Hence, future approaches must not only focus on accuracy but also on simplicity. Apart from that, our results show that both, accuracy and simplicity can be accomplished.

REFERENCES

- Aarno, D. and Kragic, D. (2008). Motion intention recognition in robot assisted applications. *Robotics and Autonomous Systems*, 56:692–705.
- Adi-Japha, E., Karni, A., Parnes, A., Loewenschuss, I., and Vakil, E. (2008). A shift in task routines during the learning of a motor skill: Group-averaged data may mask critical phases in the individuals' acquisition of skilled performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24:1544–1551.
- Argall, B. D., Chernova, S., Veloso, M., and Browning, B. (2009). A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483.
- Fearnhead, P. and Liu, Z. (2007). On-line inference for multiple change point models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69:589–605.
- Fod, A., Matrić, M., and Jenkins, O. (2002). Automated derivation of primitives for movement classification. *Autonomous Robots*, 12:39–54.
- Gong, D., Medioni, G., and Zhao, X. (2013). Structured time series analysis for human action segmentation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7):1414–1427.
- Gräve, K. and Behnke, S. (2012). Incremental action recognition and generalizing motion generation based on goal-directed features. In *International Conference on Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ*, pages 751–757.
- Graybiel, A. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory*, 70:119–136.
- Kirchner, E. A., de Gea Fernández, J., Kampmann, P., Schröder, M., Metzner, J. H., and Kirchner, F. (2015). *Intuitive Interaction with Robots - Technical Approaches and Challenges*, pages 224–248. Springer Verlag GmbH Heidelberg.
- Kulić, D., Ott, C., Lee, D., Ishikawa, J., and Nakamura, Y. (2012). Incremental learning of full body motion primitives and their sequencing through human motion observation. *The International Journal of Robotics Research*, 31(3):330–345.
- Metzner, J. H., Fabisch, A., Senger, L., Gea Fernández, J., and Kirchner, E. A. (2013). Towards learning of generic skills for robotic manipulation. *KI - Künstliche Intelligenz*, 28(1):15–20.
- Morasso, P. (1981). Spatial control of arm movements. *Experimental Brain Research*, 42:223–227.
- Mülling, K., Kober, J., Koerner, O., and J.Peters (2013). Learning to select and generalize striking movements in robot table tennis. *The International Journal of Robotics Research*, 32:263–279.
- Pastor, P., Hoffmann, H., Asfour, T., and Schaal, S. (2009). Learning and generalization of motor skills by learning from demonstration. In *2009 IEEE International Conference on Robotics and Automation*, pages 763–768. Ieee.

- Poppe, R. (2010). A survey on vision-based human action recognition. *Image and Vision Computing*, 28(6):976–990.
- Senger, L., Schröder, M., Metzger, J. H., and Kirchner, E. A. (2014). Velocity-based multiple change-point inference for unsupervised segmentation of human movement behavior. In *Proceedings of the 22th International Conference on Pattern Recognition (ICPR2014)*, pages 4564–4569.
- Stefanov, N., Peer, A., and Buss, M. (2010). Online intention recognition in computer-assisted teleoperation systems. In *Haptics: Generating and Perceiving Tangible sensations*, pages 233–239. Springer Berlin Heidelberg.

