Optimizing Contextual Ergonomics Models in Human-Robot Interaction

Antonio Gonzales Marin\textsuperscript{1}, Mohammad S. Shourijeh\textsuperscript{2}, Pavel E. Galibarov\textsuperscript{2}, Michael Damsgaard\textsuperscript{2}, Lars Fritzsche\textsuperscript{3}, Freek Stulp\textsuperscript{1}

Abstract—Current ergonomic assessment procedures require observation and manual annotation of postures by an expert, after which ergonomic scores are inferred from these annotations. Our aim is to automate this procedure, and to enable robots to optimize their behavior with respect to such scores. A particular challenge is that ergonomic scoring requires accurate biomechanical simulations which are computationally too expensive to use in robot control loops or optimization.

To address this, we learn Contextual Ergonomics Models, which are Gaussian Process Latent Variable Models that have been trained with full musculoskeletal simulations for specific tasks contexts. Contextual Ergonomics Models enable search in a low-dimensional latent space, whilst the cost function can be defined in terms of the full high-dimensional musculoskeletal model, which can be quickly reconstructed from the latent space. We demonstrate how optimizing Contextual Ergonomics Models leads to significantly reduced muscle activation in an experiment with eight subjects performing a drilling task.

I. INTRODUCTION

Musculoskeletal diseases and disorders (MSDs) are a major problem for industry workforces worldwide. The Health and Safety Executive in the UK reports that between 2001 and 2002, 5.7 million working days were lost due to back pain related disorders that arose or deteriorated at work. Upper limb and neck related disorders accounted for an estimate of 1.4 million lost working days. Combined, this amounted to £5.7 billion in production losses [1]. In Germany, MSDs were responsible for 154 million days of absenteeism and costed over €30 billion gross value losses in 2016 [3].

The prevention of MSDs therefore has a high priority in the industrial world. To perform preventative ergonomic assessments by analyzing and identifying risk factors, several standards exist, such as the Ergonomics Assessment Worksheet (EAWS) [5], OCRA [6] and Rapid Upper Limb Assessment (RULA) [7]. The procedure for these assessments is that an ergonomics expert observes a human performing the task to be evaluated, and annotates the required human postures on a paper form, such as the RULA form shown in Figure 2. These annotations are then mapped to an ergonomics score for the task. These scores make clear which parts of the process should be re-designed for better ergonomics and to reduce injury risk. This approach has several disadvantages: 1) the form is filled manually and thus labor-intensive, and it requires expertise 2) the assessment metric is not continuous 3) the underlying risk factors, e.g. high muscle activation, are not identified in the form.

Our aim is to develop an ergonomic assessment that enables automatic assessment, that takes all musculoskeletal variables into account, and that can be performed continually over time. Furthermore, queries should be very fast, so that robots are able use such models to optimize their behavior, ideally integrating ergonomic knowledge in the control loop of the robot. Robotic safety, which focuses on immediate rather than long-term injuries, has seen a similar development, i.e. from off-line static assessments to safety-aware controllers [8].

The contributions of this paper are the following:

- Optimizing robotic behavior with respect to ergonomic scores based on full musculoskeletal reconstructions.
- Enabling fast optimization by dramatically reducing query times by learning a Contextual Ergonomics Model (CEM). A CEM is a surrogate model which is implemented as Gaussian Process Latent Variable Model. It is trained off-line with data from highly detailed musculoskeletal simulations [11].

\begin{itemize}
\item 1 German Aerospace Center (DLR), Robotics and Mechatronics Center (RMC), Münchener Str. 20, Wessling, Germany
\item 2 AnyBody Technology A/S, Niels Jernes Vej 10, Aalborg, Denmark
\item 3 IMK automotive GmbH, Amselgrund 30, Chemnitz, Germany
\end{itemize}
• Showing that optimization of Contextual Ergonomics Model leads to significantly reduced muscular activations in an experiment where 8 subjects perform a drilling task.

II. RELATED WORK

In this section, we discuss related work on ergonomic assessment through manual annotation by experts, or automated assessments using simulation and biomechanical models. Finally, we discuss approaches where ergonomic measures are optimized explicitly, as in our work.

A. Manual Annotation

In the current standard procedure for ergonomic assessment, an expert observes a human performing the task to be assessed, and records 'uncomfortable' and 'strange' postures with given prototypical poses and tables as in Figure 2. Such forms may also take into account risk factors such as excessive force exertion, long task duration, and frequent repetitions. After classifying problematic key postures, a score is given to each of them.

![Fig. 2: First step of the ergonomics assessment process for the Rapid entire body assessment (REBA), a modified version of Rapid Upper Limb Assessment (RULA)](image)

The scores are summed and the overall result dictates the level of risk within a given scale as in Figure 3. After this stage, it is clear which parts of the process should be redesigned for better ergonomics and to reduce injury risk.

![Fig. 3: EAWS traffic light risk assessment in accordance to Machinery Directive 2006/42/EC (EN614)](image)

As discussed in Section I, the disadvantages of this approach is that ergonomic expertise is required, the assessment is not continual, and the procedure must be carried out every time the working space is relocated, redesigned or altered in some way. Furthermore, this technique cannot isolate the underlying risk factor of why a posture is ergonomically inadequate, e.g. high muscle activation.

B. Automated Assessment with (Biomechanical) Simulations

To address the issues of manual annotation, several approaches simulate humans performing tasks, and perform the assessment based on this simulation. The challenge here is thus not the observation – one has full access to the simulation – but rather the simulation itself.

For instance, the Editor for Manual work Activities (EMA) [9] integrates a manufacturing planning tool with the semi-automatic generation of human behaviors. Once the simulation is set up, the EAWS worksheet can be automatically filled out, based on the simulated kinematics of the human operators acting in the environment, e.g. joint angles.

In [10] the analysis of a workplace is done with Tecnomatix Jack and Catia-Delmia Human, both tools have to be setup off-line and provide only off-line ergonomics assessment, i.e. values for spinal cord loading are given and one must compare them with a given metric, for example NIOSH lifting equation. Most of this software uses biomechanical approximations to calculate possible values of the muscle activation, reaction forces and joint moments, which might based on human performance characteristics gathered statically. However, a measure of how accurate those values really are for each evaluation is missing. Moreover, the analysis is driven by inverse kinematic approaches that might not represent a good range of human motion.

In contrast to kinematic-based methods, musculoskeletal modeling offers insight into the internal body loading. Many software packages have been developed for musculoskeletal modeling, that allow for complex motion simulations and their biomechanical analysis. Commonly used packages are the AnyBody Modeling System (AMS) [11], LifeModeler [12], VIMS [13], and OpenSim [14]. OpenSim is an open source software, where ergonomics analysis can also be performed through a plugin, also here the results are not available offline.

AMS [11] is distributed with a very comprehensive full body model. It can resolve configurations with closed kinematic loops, which are frequent at workplaces, e.g. two-handed tool holding, double-stance postures. Such musculoskeletal simulations have been used for ergonomic analysis, [15], but do not systematically automate the process.

C. Optimizing Robot Behavior with respect to Ergonomic Measures

Recently, the idea of optimizing ergonomics for HRC has been analyzed in [16]. There, a simplified human model is developed and used for optimizing the ergonomic score for kinematic poses based on REBA - Rapid Entire Body Assessment [17]. The authors derive a continuous cost function $C_{\text{posture}}(q) = \sum_{i=1}^{n} w_i Q_i(q)$, using a 2nd order polynomial $Q_i(q)$ of the model’s joint configuration $q$. The weights $w_i$ are learned by minimizing the difference of the REBA score and $C_{\text{posture}}$ for various random postures of the human model. The framework optimizes kinematic postures and, therefore, musculoskeletal values are not taking into account.
The validity of the approach is measured by questionnaires and the corresponding REBA scores for the task.

In [18], the authors developed a whole-body dynamical human model that describes the relation between contact points and a resultant contact force via a Lagrangian formulation, and use it on-line for robot control. The human model structure follows the statically equivalent serial chain method. The method is validated with hand-over experiment between a light-weight robot and a human wearing an Xsens suit [22]. This work not only minimizes the joint load, but also maximizes the manipulability measure of the human.

Our approach also optimizes robot behavior with respect to ergonomic measures, the difference being that we reconstruct highly accurate and complete biomechanical simulations. The fast reconstruction of these simulations is achieved by off-line training of surrogate models, the so-called Contextual Ergonomics Models. The biomechanical reconstructions allow many different ergonomic scores to be computed and optimized, rather than only kinematic ones. In this paper for instance, we minimize muscle activations, as electromyography (EMG) can be readily used to empirically verify if muscle activations did indeed decrease.

## III. Training Contextual Ergonomics Models

The aim of this work is to develop models which on the one hand simulate the musculoskeletal system as accurately as possible – i.e. as accurate and complete as the AnyBody model – whilst having very low query times – i.e. fast enough to be run in an ergonomic optimization or robot control loop.

To achieve this aim, we learn surrogate models – called Contextual Ergonomics Models (CEMs) – from the full-scale musculoskeletal simulations for specific task contexts, for instance a drilling task. Most tasks place strong constraints on the possible body configurations, and the variations in these configurations can therefore be mapped to a latent space. The surrogate models are able to map the latent space to the full-scale musculoskeletal simulation with which it was trained. Therefore, it is both complete (full reconstructions) and fast (queries are made in the low-dimensional latent space). Next, we describe how data for the Contextual Ergonomics Model (CEM) is generated using the AnyBody Modeling System (AMS) [11], and how the CEM is learned using a Gaussian Process Latent Variable Model [21].

### A. Data Generation and Model Parameters

The AnyBody model is a full human body, comprised of all major body parts such as upper and lower extremities, a trunk model including detailed models of the lumbar spine and the abdominal pressure. Each part is based on cadaver measurements and/or medical imaging studies from an individual subject. This model includes 63 segments and nearly 1000 muscle branches that represents all main physiological muscle complexes of the human body. The generic model is 75.6 kg and 1.80 m, and can be scaled to different body postures [19].

Outputs variables from the musculoskeletal simulations in AMS include joint angles, joint moments, joint reaction loads, and muscle activation. AMS has been applied in many different fields of applications, including orthopedics, automotive, assistive devices, aerospace, ergonomics, and sports [20].

The training data for the Contextual Ergonomics Model is generated as follows. First the task and task parameter space is determined. In this paper, we chose a two-handed drilling task, where the task parameter vector is the 3D drill hole location, the mass of the drill, and the force applied to the drill. Uniformly sampling this space lead to 7632 kinematically feasible drill locations, which are visualized in Figure 4.

![Figure 4: Modeling a drilling task in AnyBody Modeling System. Green dots represent a reachable point of the drilling location. Virtual body markers can be seen on the left image.](Image)

For each drill location, the following ergonomics-relevant musculoskeletal variables are recorded: joint moments and force reactions, five muscle activation envelopes, i.e., for trunk, right and left arms and hands. Furthermore, for visualization purposes, we attach 117 virtual markers to the human model and record their \((x, y, z)\) global coordinates.

### B. Gaussian Process Latent Variable Models (GPLVM)

As is evident in Figure 4, holding the drill with two hands in front of the body dramatically reduces the number of possible postures. For this reason, we train a latent variable model with this data, so that the low-dimensional latent space that captures the variation in this reduced posture space is automatically determined.

In general, a latent variable model relates a set of latent variables \(X \in \mathbb{R}^{N \times q}\) to a set of observed variables \(Y \in \mathbb{R}^{N \times D}\) via some parameters \(W \in \mathbb{R}^{D}\). The parameters are found by maximizing the likelihood given the data. As a standard approach, the latent variables are marginalized and the parameters obtained by likelihood maximization. In GPLVM the parameters are marginalized and the latent variables optimized instead [21]. Unlike PCA and Factor Analysis (FA), GPLVM can handle nonlinear mappings.

As described above, a Gaussian process (GP) is a non-parametric and probabilistic supervised learning approach that is defined within a Bayesian framework. A GPLVM is
also a non-linear method for dimensionality reduction [21]. It serves to think of it as an extension of PCA to non-linear mappings with a qualitative measure: the variance. As with PCA, by compressing the data we don’t lose important information since the latent components that are chosen are the most representative of the data and maximize its likelihood [21]. Under the assumption that \( y_i \) are i.i.d., for \( Y \in \mathbb{R}^{D \times N} \), we have the marginalized likelihood w.r.t. the latent variables of all the data as

\[
p(Y | W, \sigma^2) = \prod_{i=1}^{N} p(y_i | W, \sigma^2),
\]

from which we can find the parameters \( W \) through maximization. If instead the parameters are marginalized we get,

\[
p(Y | X, \sigma^2) = \prod_{i=1}^{D} N(y_i | 0, XX^T + \sigma^2 I).
\]

The \( X \) which maximize the likelihood are of the following form

\[
X = U_q L V^T.
\]

We can see that Equation (3) is a projection of the matrices \( L V^T \), associated with the eigenvalues of \( D^{-1} Y Y^T \) and their rotations, to a space \( X \) of dimension \( q \) through \( U \). In this sense, our data \( Y \) can be projected into an embedded lower-dimensional space \( X \).

We have chosen GPLVM over, for instance, deep learning methods, because GPLVM is a generative model, because it has an explicit measure of variance in the latent space, and can be trained with only very limited amounts of data.

**C. Training the GPLVM with the AMS Data**

The data acquired, as described in Section III-A, are used for the training of a GPLVM. In our case, a data point consists of a high dimensional vector of 117 position markers, that represent the posture of the human model, and the associated dynamic values\(^1\) such as joint moments, muscle force reactions and muscle activations of various muscle groups.

\[
y_i = [Kin, Dyn],
\]

The input vector \( y_i \) consists of kinematics and dynamic values, \( Kin, Dyn \). Joint marker global positions \( x, y, z \) are stored in \( Kin \in \mathbb{R}^{117} \), whereas the vector \( Dyn \) holds joint moments \( \in \mathbb{R}^{30} \), joint force reactions \( \in \mathbb{R}^{120} \) and envelope muscle activations \( \in \mathbb{R}^{6} \).

We chose a 2-dimensional latent space for the projection of the data, \( q = 2 \). Optimization of the latent variables is achieved through maximization of the log-likelihood and its derivatives via stochastic gradient descent. This algorithm is a probabilistic non-parametric approach to regression. The algorithm also provides dimensionality reduction as explained above, which can be helpful for visualizing the data or embedding the learned model into control methods for pHRI, as done with the so called safety map in [8].

Figure 5 shows the latent space visualization. Here, each red cross represents the training values, e.g. a 147-dimensional vector since we use joint moments. The uncertainty is represented on gray scale, with certain areas (white) around the observed data points.

In summary, we have used an accurate but computationally expensive musculoskeletal simulation to generate training data (147-D vectors). This training data is mapped to a 2D latent space using GPLVM. The trained GP is the Contextual Ergonomics Model. “Contextual” because it is specific to a task, in this case drilling. The constraints the task introduces enable the full 147-D to be mapped to a latent space. “Ergonomics”, because it is able to very quickly reconstruct the 147-D vector of musculoskeletal variables, from which different ergonomic scores can be computed. In the next section, the CEM is used to optimize robotic behavior.

**IV. EXPERIMENTAL VALIDATION**

The aim of this experiment is to demonstrate that 1) the CEM can be used to optimized robotic behavior 2) the optimized behavior leads to significantly lower muscle activation, an indicator of more ergonomic working conditions.

Our aim here is not to advocate muscle activation as the best or sole ergonomic measure to be take into account. Rather, it is to show that, because the CEM reconstructs the full musculoskeletal variables simulated in AnyBody Modeling System (AMS), after training, we are able to optimize with respect to muscle activation and calculate a posture that lowers strain on working muscles. Moreover, we can easily these signals with the help of EMG sensors while performing the task.

The validity of the approach was evaluated with a human performing a simulated drilling task assisted by a manipu-
lator, as illustrated in Figure 1. In the following we present the experimental setup and the results.

A. Methods and Procedure

The setup consists of a LWR IV+ light-weight robot mounted on a linear axis, 8 DOF, and a plate that is to be drilled. A human operator stands in front of the robot in order to drill the plate. The plate is presented by the robot on different locations on a plane. All the experiments were carried out with the same tool, a driller, which represents a load of 1.8 [kg].

a) Task and Subjects: The drill was held with both hands by the participants. The participant pushes the plate with a certain force of 5, 10 or 15 [N], where visual feedback was given of the measured external force, i.e. the push, from the robot torque sensors. The readings at 5 [N] were considered as noise, both from the EMG and the robot sensors, as the force level was too low, therefore we present results gathered from the participants only for the experiments carried out at 10 and 15 [N] of exerted force. Each subject performed 24 drill tasks, i.e. 4 drill locations (one of them optimized), 2 repetitions per location, 3 push forces (5,10,15). The order in which these tasks were performed was randomized for each subject. In total, 8 subjects participated in this experiment.

b) Drill hole locations: The robot presented the plate at four different locations relative to the participant. We denote these locations up, as in the left picture of Figure 1, left, right, and optimized, as in Figure 1 right picture. The up, left and right positions where hard coded and scaled depending on the participant’s height and arm span, so that is reachable, which were measured before starting the experiment.

The optimized location is determined by minimizing maximum muscle activation, which are reconstructed from the Contextual Ergonomics Model. Each CEM query is very fast (12 [ms] ± 0.1), compared with 10 [s] for performing the simulation in AnyBody Modeling System for a point. And the search space is the latent space, which is only 2D. For these two reasons, a simple grid search was used to perform the optimization. The step of the grid can be varied, for these two reasons, a simple grid search was used to perform the optimization. The step of the grid can be varied, for example 10000 queries can be performed in 4 secs. For each query, the 2D latent point is projected to the observed space, the x, y, z coordinates of the joints and the high-dimensional musculoskeletal vector, 123-D for muscle activation, 147-D for joint moments and 237-D for joint force reactions. From this vector, we extracted the variables related to the optimized criteria and compute the cost associated with them, in this case muscle activation. As explained before, the inputs and the reconstructed high-dimensional vector consist of the kinematic data and the dynamic properties stacked in one vector. Thus, for calculating the cost, we just need to retrieve the values from indices 118 up to the end of the vector.

The optimized location was taken from the kinematic data of the optimal latent point, c.f. Figure 5. The (x, y, z) coordinates were then scaled using the participant’s height and arm span to calculate a desired end-effector Cartesian point w.r.t. the base frame of the robot. The participants were asked to stand on a position and not move until the end of all the experiments, all trials for 5[N], 10[N] and 15[N] push force with two repetitions per trial. In general, the required parameters, i.e. height, arm span and location of the participant, can be calculated from external sensors such as kinect or a motion capture device, e.g. Vicon.

c) Measuring muscle activation: Subjects wore non-invasive EMG surface sensors. We position 16 EMG sensors for gathering muscle activation data, 8 per each side, covering agonist antagonist muscles, as shown in Figure 6. Lower body muscles are not taking into account for these experiments. We use only EMG data that is recorded during the drilling motion.

B. Results

EMG readings from two EMG sensors are illustrated in Figure 1. The muscle shown on that figure is the left trapezius for the two different postures. For the non-optimized drill location, the left trapezius readings showcase a higher muscle activation than the pectoral in the optimized position.

To measure the effect of the optimization, we take the mean activation from each EMG sensor, as visualized in Figure 1. We perform a Student’s t-test between the one optimized drill location and the three non-optimized locations. Since not all muscle activation are equally affected and have different scaling, we perform these t-test separately for each EMG reading. In Table I, we list the p-value of this test for each of the 16 EMG sensors, and each of the three non-optimized drill hole locations. We see that using the optimized position leads to significant improvements in almost all muscles. On average the improvement is between 85 and 89% depending on which reference location is used (see left table).

In summary, with the Contextual Ergonomics Model (CEM) we are able to optimize drill hole locations, and this optimization leads to a significant decrease in muscle activation for most muscle groups, as confirmed by EMG readings.

V. Conclusion

In this work we proposed a novel approach for the fast optimization of ergonomic scores that are based on full
Table I: Muscle activation ratio of up/left/right position w.r.t. optimized position. For instance, 87% implies that the average muscle activations are 13% less when using the optimized position. White: difference not significant ($p > 0.05$). Green: significant improvement. Red: significant increase in muscle activations.

<table>
<thead>
<tr>
<th>Drill hole location</th>
<th>up</th>
<th>left</th>
<th>right</th>
<th>up</th>
<th>left</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.73</td>
<td>.73</td>
<td>.73</td>
</tr>
<tr>
<td>2</td>
<td>.00</td>
<td>.01</td>
<td>.00</td>
<td>.95</td>
<td>.93</td>
<td>1.12</td>
</tr>
<tr>
<td>3</td>
<td>.02</td>
<td>.13</td>
<td>.00</td>
<td>.95</td>
<td>.93</td>
<td>1.12</td>
</tr>
<tr>
<td>4</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.70</td>
<td>.79</td>
<td>.61</td>
</tr>
<tr>
<td>5</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.70</td>
<td>.79</td>
<td>.61</td>
</tr>
<tr>
<td>6</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.94</td>
<td>.82</td>
<td>.82</td>
</tr>
<tr>
<td>7</td>
<td>.19</td>
<td>.00</td>
<td>.00</td>
<td>.97</td>
<td>.87</td>
<td>.82</td>
</tr>
<tr>
<td>8</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.76</td>
<td>.66</td>
<td>.70</td>
</tr>
<tr>
<td>9</td>
<td>.70</td>
<td>.92</td>
<td>.05</td>
<td>.113</td>
<td>.129</td>
<td>1.48</td>
</tr>
<tr>
<td>10</td>
<td>.72</td>
<td>.02</td>
<td>.08</td>
<td>.100</td>
<td>.90</td>
<td>.93</td>
</tr>
<tr>
<td>11</td>
<td>.64</td>
<td>.02</td>
<td>.00</td>
<td>.106</td>
<td>.98</td>
<td>.70</td>
</tr>
<tr>
<td>12</td>
<td>.00</td>
<td>.01</td>
<td>.57</td>
<td>.87</td>
<td>.94</td>
<td>1.06</td>
</tr>
<tr>
<td>13</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.83</td>
<td>.81</td>
<td>.86</td>
</tr>
<tr>
<td>14</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.88</td>
<td>.79</td>
<td>.90</td>
</tr>
<tr>
<td>15</td>
<td>.01</td>
<td>.00</td>
<td>.00</td>
<td>.87</td>
<td>.87</td>
<td>.83</td>
</tr>
<tr>
<td>16</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.87</td>
<td>.69</td>
<td>.78</td>
</tr>
</tbody>
</table>

In our future work, we want to track humans continually over time, so that we may continually re-optimize the robot’s behavior by integrating Contextual Ergonomics Models in the control loop of the robot. Also, we will exploit the estimates of the variance made by the GPLVM to speed up the optimization; in higher-dimensional latent spaces grid searches may be unfeasible.

In our current experiment, 5 variables related to muscle activation were optimized. However, all musculoskeletal variables are easily reconstructed from the latent space, and any (or all) of these variables could be used in a cost function without slowing down the search. In our future work, we will investigate novel ergonomic measures which take into account this wealth of musculoskeletal information, which was hitherto not accessible for ergonomics assessments. Moreover, we plan to develop risk thresholds based on established ergonomic standards to not only determine relative changes in muscle activation through optimization but also report on risk levels (green – yellow – red, cf. Figure 2, Figure 3), analogous to traditional ergonomic methods.

Acknowledgments

This work was supported by the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No. 731540 (project An.Dy) and the HGF project “Reduced Complexity Models”.

References