Skill Evaluation of Human Operators in Partly Automated Mobile Working Machines

Kalevi Tervo, Lauri Palmroth, and Heikki Koivo, Senior Member, IEEE

Abstract—The performance of a mobile working machine is strongly dependent on the skills of the operator. It is commonly accepted that among operators, productivity may vary tens of percents. This research considers the human skill evaluation in working machines during normal work. A general framework for skill evaluation via task sequence recognition by a hidden Markov model (HMM) is described. The definitions of skill metrics, based on task resource consumption, task completion frequency, task difficulty, and ability to plan and make decisions, are given and justified through an example. The skill evaluation methods are illustrated by utilizing them in two industrial applications.

Note to Practitioners—The research results of skill evaluation of working machine operators during their normal work enable novel possibilities to improve the efficiency of the man-machine process. There exists significant untapped potential in productivity, quality and fuel economy which can be exploited only by considering both technical and human factors together. The skill evaluation approach proposed in this research can be used as a tool to assist the operators in their education phase to learn to perform the work more efficiently. Furthermore, the proposed method provides tools also for the professional operators to receive detailed feedback about their work, which has not been possible before.

Index Terms—Hidden Markov models (HMMs), human adaptive mechatronics, human factors, human performance, man–machine systems, skill evaluation.

I. INTRODUCTION

In industrial processes, the operator plays a significant role in terms of productivity, fuel economy, and quality of the end product. Instead of mere technical performance of the process, one needs to also consider the human factors to optimize the overall performance. In industry, the performance optimization has traditionally been viewed from the perspective of the machine’s performance. The tools range from performance assessment to fault detection and isolation (FDI) and fault tolerant control (FTC). However, currently there are several branches of research aiming for flexible, safe, and efficient operation in human–machine systems. One of the branches is the collaborative robot (Cobot) architecture for increasing safety and productivity in automotive assembly [1]. Several researchers have worked in the area of human machine collaborative systems (HMCSs), which have been used to prevent ergonomic injuries and operator wear in human operated processes [2], [3]. The emerging discipline of human adaptive mechatronics (HAM) covers more or less the aforementioned objectives while simultaneously going significantly deeper in psychological issues related to human skill and performance [4]–[7]. A key issue in HAM is in quantification of the operator’s skills.

It has been recognized and accepted that the actions of the human operators have a tremendous effect on the safety and profitability of an industrial process. In the chemical industry, up to 40% of abnormal operations can be results of human errors [8]. In mechanized timber harvesting, the variance of performance between operators has been studied in several specifically arranged field tests. According to these studies, the productivity difference between two professional operators working in similar operating conditions with similar machines can be over 40%. In addition to the productivity, the quality of the work is reported to vary between the operators [9].

Although being studied in the field tests with additional sensors and measurements, there are no reported results of evaluating the skills of operators of mobile working machines during the normal work using only the measurements readily available. One reason for this is the sensitivity of the subject, an unpleasant and wary feeling that often arises with an individual being monitored [10]. Another, perhaps more important reason has been the lack of proper data manipulation methods which could reveal the skill differences.

This paper proposes a novel method for skill evaluation of human operators in mobile working machines using the control signals given by the operator as inputs. The method is based on modeling the work by intuitively understandable tasks. These work tasks, that produce work cycles, are recognized using a probabilistic method, hidden Markov models (HMMs). The skill evaluation is based on efficiency and complexity of the recognized task sequence, or resource consumption during the tasks. In addition, methods based on task difficulty and the ability to plan and make decisions, are proposed. These methods are utilized to evaluate the skills of operators of cut-to-length forestry machinery.

The paper is organized as follows. A literature review of the related work is presented in Section II. Section III clarifies the skill evaluation methods used in this research. The HMMs and the work cycle recognition method are introduced in Section IV. As an example, the mechanized timber harvesting method and the machines, harvester and forwarder, are introduced in Section V. Furthermore, the section illustrates the...
operating conditions and the work cycles of the machines. Section VI presents the experimental setup and data used in the skill evaluation of forwarder operators as well as the obtained results. Section VII deals with the experimental results obtained with the harvester. Conclusions are presented in Section VIII.

II. HUMAN SKILL EVALUATION

The human performance and skill evaluation have been researched widely in various communities. In robotics the human skill is used, for example, in teaching the robot. Other main application areas have been in medical engineering, especially in surgery, commercial and military flight systems, robotics, and nuclear power plants. It is obvious that a high level of human skill is a necessity in these applications. In the following, related work around human skill evaluation is reviewed. A special subsection is devoted to the state-of-the-art of skill evaluation via task sequence recognition, which is the core approach in this paper.

A. Related Work Around Human Skill Evaluation

An important concept in human skill evaluation considering the evaluation of simple task execution is the Fitts’ law. It describes the relationship between a task performance metric and task difficulty index [11]. The task difficulty index is basically given by the width of the target. The Fitts’ law gives the execution capacity of a human. An extension for Fitts’ law to evaluate skills characterized by more than one variable is proposed in [12]. The problem with this method is that it applies only for a simple task execution. In mobile working machines, each task is unique: the distances, positions, and sizes of the targets vary and often the variables cannot be measured.

A control theoretic approach for human skill evaluation in task execution is to identify the human controller. This means the input-output relationship between the outputs of the controlled dynamical system and the control signals applied by the human operator. The skill level is quantified based on the parameters of the identified controller, frequency characteristics of the control signals, or comparison of the optimal control trajectory to the realized trajectory [5], [13], [15]. Utilization of this approach requires good quality measurements sampled at high enough frequency to identify dynamic models or the optimal control law. In addition, the simplest controllers, which are proportional derivative (PD)-type, are suitable only for simple tracking task execution modeling. In case of a more complex task sequence, a hybrid type of controller should be utilized. Furthermore, in real-life systems where a human operator controls the process, e.g., by using joysticks, the control signal usually saturates. This makes the modeling even more challenging.

A popular approach for quantification of human skill is to train dynamic statistical models such as HMMs to represent different skill levels. The evaluation is then based on the posteriori probability of the models or, if there is only one model to represent an expert’s skill, the likelihood is used to give the skill level [16]–[19]. This approach gives the skill level with respect to predefined skill levels, but it does not provide insight of why the skill of the person is not at a desirable level or what are the most important areas of improving the performance. That would require models for all levels of performances. Furthermore, if the work objective can be achieved in multiple equally good ways, it may be impossible or impractical to define models to represent the skill level.

A framework for analysis and transfer of human skills from human to robot and human to human was proposed in [20]. Human control and motion skills were analyzed, and a computer simulator system was introduced to transfer the skills. Though the proposed system increased the skills of the test subjects, the implementation of the method in an online use in a human operated industrial machine is not yet solved.

B. Skill Evaluation via Task Sequence Recognition

This paper considers operator skill evaluation via a task sequence recognition because the approach has proven to be effective in the case of mobile working machine operators.

Skill evaluation of dynamic tasks in surgery was studied in [21]. Separate HMMs were trained for each motion based on gestures extracted from the dynamic state of the system. The motions constituted a network of HMMs. The most likely motion, at each time instant, is recognized by evaluating the likelihood of the gesture sequence for each motion model. The skill evaluation was based on the total number of motions used to perform the task, and the percentage of time used for each motion. References [22] and [23] modeled surgeons’ performances by defining the work as task sequences which were modeled by HMMs. Tool forces and torques were measured in three dimensions and the HMM states were defined to correspond to the tasks the surgeon performs during the operation. The skill evaluation was based on statistical distance to an expert model, task execution times, task execution frequencies, and force/torque measurements during tasks.

A skill evaluation system based on a deterministic state transition model was utilized in [6]. The skill evaluation was based on state transition frequency patterns recorded during the operation. The aim of the skill evaluation research was to realize a HAM system. In [24], a peg-in-hole task was modeled by using HMMs and the method was applied to the quantitative comparison of human skills between three workers. The comparison was done with a similarity measure between the HMMs trained separately for each worker. The peg-in-hole task was divided into task states. Some of the task states were described by only one state of a HMM whereas some others used more than one state. However, the sequence of states was not used to evaluate the workers’ skills.

III. SKILL EVALUATION METHOD

A. Overall Description of the Method

The human performance can be defined as person’s most likely outcome when performing a given task [16]. Thus, the skill is interpreted as the average performance from several repetitions of a task. In addition, the skill level can be measured based on how smoothly a sequential process is executed [6]. One can define the human skill in machine operation as “an ability to manipulate machines accurately, fast, with high repeatability, and to cope with emergency circumstances” [25]. In order to evaluate operators’ skills to perform complex work with advanced machines, a broader definition for a skill is needed.

Definition 1: Human operator’s skill in work performed with a mobile working machine consists of the following components: i) an ability to control the machine, ii) an ability to tune
the control parameters of the machine appropriately to suit the operator’s machine controlling skills, iii) a knowledge of the work technique and strategy, and iv) an ability to plan and make decisions.

As an example, one can consider the difference of an expert and a beginner when driving a car with manual transmission. This example is modified from the one presented in [26]. The manual gear shifting process is described by a task sequence in Fig. 1. For a person who has learnt to drive a car with automatic transmission, it can be very difficult to learn to shift gears manually. At the very early stage of learning, the driver needs to carefully perform the sequence of tasks. Each task needs to be cognitively rehearsed before completing it. When changing the gear, the beginner needs to check the location of the next gear before engaging the clutch. The driver shifts the stick to the middle position. Once the shift stick is in middle position, the location of the next gear may have been forgotten, so it needs to be rechecked. The driver might attempt to change to a wrong, say, too high a gear. The beginner might face some other problems as well, whereas the expert driver accomplishes the task sequence of gear shifting more smoothly, autonomously, faster and without hesitation.

In this section, four rather general frameworks for the skill evaluation of human operators using a HMM with explicitly defined states are proposed. The role of the HMM structure is to recognize the intent of the operator behind the resultant actions. In this case the resultant actions are the visible and measurable control actions of the machine operator, such as button presses and control lever motions. We define the states of an HMM to correspond to the individual tasks and operational phases of the work, e.g., “Engage the clutch.” We call the \( q^\text{th} \) state \( S_i \), where \( i \) denotes the index of the state. Furthermore, a state sequence \( Q = \{q_1, q_2, \ldots, q_T\} \) refers to a sequence of tasks that is needed to achieve the goal. For example, a state sequence for gear shifting could be:

\[
\begin{align*}
\text{Beginner} & \quad \text{Expert} \\
& \quad \text{Check the location of next gear} \\
\quad \text{Engage the clutch} \\
\quad \text{Move the shift stick to middle} \\
\quad \text{Recall the position of next gear} \\
\quad \text{Move the shift stick to correct gear} \\
\quad \text{Disengage the clutch} \\
\quad \text{Engage the clutch} \\
\quad \text{Move the shift stick to middle} \\
\quad \text{Move the shift stick to correct gear} \\
\quad \text{Disengage the clutch}
\end{align*}
\]

“Engage the clutch,” “Move the shift stick to middle position,” “Move the shift stick to the correct gear,” “Disengage the clutch.” With these definitions, the whole work process is divided into smaller and more understandable pieces, the tasks. Moreover, using HMMs in the task recognition, task, or task sequence dependent skill evaluation of a machine operator can be performed without additional sensors using the visible and measurable control actions of the operator.

The structure of the skill evaluation system is shown in Fig. 2. The measurements obtained from the process are used to determine the task sequence of the work. Once the tasks are determined, the process measurements can be associated with the tasks. This is important since good values for the measurements often depend on the task. For example, good value for the number of simultaneous motions in operation of a crane with several degrees of freedom is different in a picking task than in a moving task. Within the proposed skill evaluation system structure, the skill metrics can be defined based on task dependent measurements. Moreover, there can be several skill metrics for each of the skill components i)-iv) (Definition 1).

This paper focuses on how to extract the skill metrics from operational data. The advantage of the method is that the same system can be used to evaluate operators with any skill level because the skill level is determined based on the performance indices, i.e., the skill metrics. Consequently, since the skill metrics can be defined as task specific, and since in each task the different skill components can be evaluated (Definition 1) the reasons for low performance can be analyzed more easily.

It is important to note that the skill metrics obtained by the proposed evaluation system might depend on the prevailing operating conditions where the work is executed. Thus, the evaluations should always be done in respect with the work conditions.

### B. Four Frameworks for Skill Metric Development

In the following, the concept of Kronecker delta is used to denote whether two arguments are equal or not. It is defined as

\[
\delta_{ij} = \begin{cases} 
1, & i = j \\
0, & i \neq j
\end{cases}
\]
It is a useful notation when computing state transition frequencies. Moreover, the notation \( q_k \) describes the state (task) at time \( t \).

1) Task Efficiency as Skill Measure: The premise of this approach is to assume that a skilled operator accomplishes a single task or a sequence of tasks efficiently without wasting resources. That is, when comparing a beginner and an expert it is likely that the expert consumes fewer resources than the beginner, at least in the long run. Thus, it can be said that the expert is more efficient than the beginner. In [21]–[23], the task completion times are used to measure operator skills. In general, instead of time, any resource can be considered.

Let \( R_t \) denote the amount or rate the resources are consumed between observation time instants \( t \) and \( t + 1 \). Thus, we can define \( E_i \), the total efficiency of the task \( S_i \) during the task sequence within time window \( T \). It is obtained by summing over all the observations. That is

\[
E_i = \sum_{t=1}^{T} \delta_{i,q_t} R_t
\]

where \( \delta_{i,q_t} \) is the Kronecker delta. Furthermore, if the interest is to study the mean resource consumption of a single task, one can compute

\[
E_i = \frac{\sum_{t=1}^{T} \delta_{i,q_t} R_t}{\sum_{t=1}^{T} \delta_{i,q_t}}.
\]

Here, however, it is important to note that in order for this quantity to be well defined, the denominator cannot be zero. Thus, it can be computed only when at least one state transition to \( S_i \) appears.

The definition of the resource consumption factor \( R_t \) depends on the measurements available. The simplest way is to define \( R_t \) as the time consumed between time instants \( t \) and \( t + 1 \). Thus, the efficiency of task \( S_i \) is directly proportional to the time spent in that state. In the manual gear shifting example, the beginner most likely consumes considerably more time with each task compared to the expert. The task efficiency as a skill measure measures mainly the skill component \( iii \), but can reflect to all skill components \( i-iv \) in Definition 1, depending on the tasks of interest and the definition of \( R_t \).

2) Complexity of the Task Sequence as Skill Measure: This approach assumes that a skilled operator achieves the goal smoothly with a fairly simple sequence of tasks. As an illustration, consider the gear shifting example, where an expert completes the job in only four logical operational phases, whereas a beginner might need more phases before the new gear is successfully found. A beginner also might have to perform some phases more than once after an unsuccessful attempt.

Let \( f_i \) denote the number of state transitions to \( S_i \) from any other state. Thus, a simple skill measure for the task \( S_i \) is obtained by computing

\[
f_i = \sum_{t=2}^{T} \delta_{i,q_t} (1 - \delta_{i,q_{t-1}})
\]

where the quantity inside the summation is one only if the current state is \( S_i \) and the previous state is not \( S_i \). Consider again the gear shifting example. The expert moves the shift stick to the correct gear on the first attempt. Thus, the value of frequency \( f_{\text{move}} \) the shift stick to middle is one for the expert. For a beginner the value is often two or more, because the stick is moved to the middle position more than once.

Furthermore, a measure for the complexity of the whole task sequence can be obtained adding up all state changes. That is

\[
f_{\text{tot}} = \sum_{i=1}^{N} f_i
\]

where \( N \) denotes the number of possible states. The complexity of the task sequence measures the skill components \( iv \) and \( iii \) of Definition 1, depending on the tasks of interest and the definition of \( f_i \).

3) Ability to Plan and Make Decisions as Skill Measure: When working in a complex and dynamic environment with time pressures about the work completion, an expert chooses the first available solution by intuition instead of carefully considering several solution candidates first and then choosing the optimum [27]. Once a decision is made, the task is executed determinedly without hesitation. For example, a beginner often performs tasks related to decision making several times upon completion while the expert can quickly make a successful decision and needs to perform those tasks only once. Several direction changes or stoppages in the trajectory of the movement during a task may also suggest that the operator has poor decision making skills.

The ability to plan and make decisions as a skill measure can be described mathematically analogously to the complexity of the task sequence using (3), where \( f_i \) now describes the state transitions to a particular decision making related task \( S_i \). If the number of direction changes or stoppages during task \( S_i \) is chosen to be the skill measure, it is obtained using (1), where \( R_t \) is the particular value of interest measured during the task. An ability to plan and make decisions measures essentially the skill component \( iv \) of the Definition 1.

4) Task Difficulty as Skill Measure: In addition to mere efficiency of a task or resource consumption, the way each task is executed provides valuable information about the operator skills. In this approach it is assumed that it is not possible to define a task difficulty index directly from the available measurements. However, it can be assumed that since the income of professional operators may often be directly related to the work outcome, they want to fully utilize their execution capacity. Thus, an expert operator can accomplish a task or a sequence of tasks in the most efficient way but with the cost of high difficulty, whereas the beginner has to seek for an easier and less efficient solution. As an interpretation of the Fitts’ law, the resulting average difficulty of tasks performed by a skilled professional operator can generally be assumed higher than with a less proficient one, and the task difficulty can be used as an indirect measure of the operator skill.

With mobile work machines, the HMM provides a way to study the characteristics of the task difficulty by observing the control commands given by the operator. According to reference [28], handling of control levers of a forest machine while moving the boom follows an autonomous over-trained motor-
sensory scheme. In comparison to a beginner, an expert does not need such a cognitive effort in performing the task. To complete the task as efficiently as possible, multiple machine functions are controlled simultaneously at a high rate and in overlapping sequences instead of controlling each function separately. Thus, the total number of parallel controlled machine functions and the mean control rate during a task can be used as measures of the task difficulty.

The skill measure of task $S_i$ difficulty can be obtained using (2), where $R_i$ is defined as the value of interest, the mean control rate or the mean number of parallel controlled functions during the task $S_i$. The task difficulty as a skill measure describes essentially the machine controlling skills, that is, skill component $i$) of Definition 1. An operator with high machine controlling skills can execute more difficult tasks than a low-skilled operator. Together with efficiency metrics and knowledge of the current parameter setup it can be used to measure component $i$) as well.

IV. WORK CYCLE RECOGNITION

A. Hidden Markov Models (HMMs)

Since late 1980s, the research of HMMs has escalated. The application areas have widened from speech recognition [29] to fault diagnostics [30], [31], prognostics [32], condition monitoring [33], communications engineering [34], human intent recognition [35], and human behavior modeling [36].

The performance of a human operator of completing a given task is a time-varying stochastic process where the result depends on several factors: the mental agility of the operator, the disturbance of circumstances, the imperfect and noisy nature of the human sensory processes and the skill associated with the task. Reference [16] describes the human performance as a result of two distinct simultaneous stochastic processes: human mental state and resultant actions. The mental states, including the intent hidden behind the actions, constitute a stochastic process which cannot be measured. On the contrary, the resultant actions are observable and form a measurable stochastic process. The goal of the work cycle recognition is to recognize the operator’s intent based on measurable actions.

A HMM consists of two simultaneous stochastic processes. The first, underlying stochastic process constitutes a Markov chain, but unlike with ordinary Markov models, the states cannot be observed. The second stochastic process produces a sequence of observations. Each state has a probability distribution for the observations to appear. Thus, based on the sequence of observations the most probable respective state sequence can be deduced [33]. Due to the stochastic and dualistic nature of HMMs, they are very often used in modeling the human performance.

A HMM with 1-D discrete observation probability distributions can be defined as follows [37].

1) A set of possible hidden states $S = \{S_1, S_2, \ldots, S_N\}$. State at time $t$ is denoted by $q_t$ and the state sequence within $1 \leq t \leq T$ by $Q = \{q_1, q_2, \ldots, q_T\}$.
2) Observation symbols $V_k$, where $1 \leq k \leq M$. Observation at time $t$ is denoted by $v_t$ and the observation sequence within $1 \leq t \leq T$ by $O = \{o_1, o_2, \ldots, o_T\}$.
3) State transition probabilities $A = \{a_{ij}\}$, where $1 \leq i, j \leq N$ and $a_{ij} = P(q_{t+1} = S_j | q_t = S_i)$.
4) Observation probabilities $B = \{b_j(k)\}$, where $b_j(k) = P(o_t = v_k | q_t = S_j)$.
5) Prior probabilities $\pi = \{\pi_i\}$, where $\pi_i = P(q_1 = S_i)$.

For convenience, triplet $\lambda = (A, B, \pi)$ is usually used to represent a HMM.

With HMMs, there are three basic problems of interest [33]: First, the training problem, training the parameters of the model $\lambda$. Second, the evaluation problem, evaluating the probability of the observation sequence $O$ being produced by the model $\lambda$. Third, the decoding problem, decoding the most likely path of states $q$ that maximize the probability of the observation sequence $O$.

B. Work Cycles by a HMM

The work cycle recognition used in this study is based on explicitly defined HMM states, which is similar to approach [21]–[23]. However, due to the fact that in normal use of mobile working machines the inclusion of extra measurement devices, such as position, force or torque sensors, is not often feasible, no additional sensors are used. The control commands given by the operator are used as inputs.

The model is trained so that the states of the HMM correspond to the tasks needed to complete the work cycle. With the skill metric definitions given in Section III, a task could also consist of multiple HMM states instead of just one. Few of the most common exceptional situations which might appear during the process are also defined as model states. In forwarding work there are two different work cycles that are modeled by a HMM, one for loading and the other for unloading. The probabilities of measured observation sequences are calculated for both models and the model with the higher probability is chosen (evaluation problem). Given the observation sequence and the model, the most likely path of states (tasks) is calculated (decoding problem) using the Viterbi-algorithm [38]. In this paper, the HMM algorithms of the Matlab’s statistics toolbox are used for the evaluation and the decoding problem.

Due to the nature of available measurement information, a static classifier cannot separate the work phases to the extent desired. On the contrary, a dynamic classifier, such as a HMM performs better in classification of dynamic repeatable tasks. This is due to the ability of a HMM to take into account the temporal context of the current observation, that is, the observation history. Therefore, in order to utilize the method proposed, the work to be modeled needs to be repetitive and consist of a (relatively) small number of distinct tasks. If successive work cycles are similar enough, the use of dynamic classification is reasonable. The requirement of small number of classes comes from the fact that the generalization capability of a classifier decreases if the number of classes is increased too much [39].

C. Feature Extraction for Recognition

When the HMM is defined so that each state corresponds to a real work task, there is a significant challenge to preprocess the input data such that the classification is possible. The recognition system should also be robust to differences in the individual work styles and control parameter setups of the machine operators. Due to this, the features forming the codebook of the HMM, are quantized as one dimensional “characters.” These observation characters are extracted from the machine’s
distributed control system data by combining the information in the multidimensional control commands given by the operator and the other relevant data available. Quantization of the data is sparse in order to obtain as small a codebook as possible. Each observation character defines a different physical property derived from the control signals, such as estimated direction or velocity of the controlled actuator. [40]

The implementation of the task recognition should always be tailored to the specific application. The details of feature extraction are not discussed deeply in this paper, because the extracted, specific features depend strongly on the application and the work task definitions. In the end, the available measurements in each application determine which type of features can be extracted from the data.

V. MECHANIZED TIMBER HARVESTING

There are two main methods for mechanical timber harvesting depending on the wood utilization, transportation and the machine types needed. In North America the full-tree method is common whereas in Europe the cut-to-length method is dominant. In the cut-to-length method, two separate machines are used, a forest harvester and a forwarder. A harvester is used for felling, delimbing, and cutting trees. Once the harvester has processed the tree stems to logs, a forwarder picks up the logs and carries them to the roadside for further transportation. Pictures of a harvester and a forwarder are shown in Fig. 3.

A. Work Cycle of a Forest Harvester

The task of a harvester is to process the trees standing in the forest to logs. Harvester is equipped with a processing head (harvester head) that the operator uses to perform the multiple tasks to complete the work. The tasks to complete the tree processing can be expressed as follows.

1) Choose the tree to fell and move the harvester head to it.
2) Grab the tree with the harvester head and initiate chain saw to cut it down (fell cut).
3) Feed the tree to the next cutting point. During feeding the delimbing knives prune the stem and the harvester head measures its volume and length.
4) When the cutting point is reached, initiate the chain saw to cut the log (cross cut).
5) If the top of the tree is reached go back to step 1. Otherwise go back to step 3.

The third task forms another work cycle, namely, the stem feeding process. The aim of this step is to feed the stem to the next cutting point to get the best yield out of the stem, either in relation to a price list or a specific order. Harvesters are equipped with a color marking system to be able to separate the various timber grades during transportation. Most of the operations in the feeding process are automated but the operator is still a key factor in the performance of the process. The optimization system which assists in bucking (i.e., cutting the trees into logs) suggests the cutting points so that the value of the stem is maximized. However, the operator has to make the final decisions about the cutting points. The operator can also switch the automatic feeding off, which is typically done in an unexpected situation. In addition, the operator chooses the positions of the log piles by steering the crane to the desired position before cutting the stem into logs. The work tasks of a harvester operator are described in more detail in [28].

B. Work Cycle of a Forwarder

The task of a forwarder is to load the processed logs from the forest into the load space of the machine, carry them to the roadside and then unload the logs to a pile. In forwarder work, the boom operation plays the most significant role in terms of the overall performance of the machine. Unlike the harvester work, there are no automated functions. Thus, the work of the operator is highly demanding regarding the sensory-motor skill associated with the manual control of the boom. A skilled operator is able to control multiple boom joints simultaneously and to maneuver the payload fast and precisely, despite that the boom crane has multiple oscillating modes that change significantly over position and load.

A work cycle of loading or unloading operation consists mainly of the following basic tasks.

1) Move the grapple to payload.
2) Grab the payload.
3) Move the payload into the load space or to the pile at roadside.
4) Position the payload.

A requirement for an efficient forwarder work is to maximize the payload in the load space that is carried to the roadside and to minimize the required driving distance. Usually separate timber grades are transported to separate locations. Sometimes various timber grades may be loaded to the load space which introduces more sorting and arrangement of logs into separate piles.

VI. EXPERIMENTAL RESULTS WITH FORWARDER

The main goal of the forwarder operator is to complete the whole work task, loading logs into load space and unloading them to the roadside as efficiently as possible. To achieve the best possible overall result there exist variations between the operator’s work strategies, for example concentrating more on a particular task than others. This could mean that an operator
TABLE I
DESCRIPTION OF THE HARVESTER OPERATOR SKILL METRICS, RELATED EQUATIONS (Eq) AND ASSOCIATED SKILL COMPONENTS (SC)

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Eq</th>
<th>Sc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metric1</td>
<td>Work cycle count</td>
<td>(3)</td>
<td>iii</td>
</tr>
<tr>
<td>Metric2</td>
<td>Complexity of task sequence</td>
<td>(4)</td>
<td>i,i',iv</td>
</tr>
<tr>
<td>Metric3</td>
<td>Control rate 1</td>
<td>(2)</td>
<td>i</td>
</tr>
<tr>
<td>Metric4</td>
<td>Control rate 2</td>
<td>(2)</td>
<td>i</td>
</tr>
<tr>
<td>Metric5</td>
<td>Ability make decisions</td>
<td>(1)</td>
<td>iv</td>
</tr>
<tr>
<td>Metric6</td>
<td>Parallel control of machine functions 1</td>
<td>(2)</td>
<td>i</td>
</tr>
<tr>
<td>Metric7</td>
<td>Parallel control of machine functions 2</td>
<td>(2)</td>
<td>i</td>
</tr>
</tbody>
</table>

TABLE II
RELATIVE PRODUCTIVITIES (RP) AND AVERAGE VALUES OF THE SKILL METRICS OF FORWARDER OPERATORS

<table>
<thead>
<tr>
<th></th>
<th>FOP1</th>
<th>FOP2</th>
<th>FOP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP</td>
<td>1.072</td>
<td>1.212</td>
<td>0.714</td>
</tr>
<tr>
<td>Metric1</td>
<td>0.700</td>
<td>0.727</td>
<td>0.272</td>
</tr>
<tr>
<td>Metric2</td>
<td>0.495</td>
<td>0.475</td>
<td>0.524</td>
</tr>
<tr>
<td>Metric3</td>
<td>0.633</td>
<td>0.790</td>
<td>0.209</td>
</tr>
<tr>
<td>Metric4</td>
<td>0.392</td>
<td>0.843</td>
<td>0.156</td>
</tr>
<tr>
<td>Metric5</td>
<td>0.676</td>
<td>0.643</td>
<td>0.323</td>
</tr>
<tr>
<td>Metric6</td>
<td>0.447</td>
<td>0.751</td>
<td>0.248</td>
</tr>
<tr>
<td>Metric7</td>
<td>0.483</td>
<td>0.816</td>
<td>0.183</td>
</tr>
</tbody>
</table>

TABLE III
STANDARD DEVIATIONS OF THE FORWARDER OPERATOR SKILL METRICS

<table>
<thead>
<tr>
<th></th>
<th>FOP1</th>
<th>FOP2</th>
<th>FOP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ1</td>
<td>0.213</td>
<td>0.351</td>
<td>0.252</td>
</tr>
<tr>
<td>σ2</td>
<td>0.194</td>
<td>0.816</td>
<td>0.416</td>
</tr>
<tr>
<td>σ3</td>
<td>0.229</td>
<td>0.282</td>
<td>0.115</td>
</tr>
<tr>
<td>σ4</td>
<td>0.120</td>
<td>0.242</td>
<td>0.106</td>
</tr>
<tr>
<td>σ5</td>
<td>0.224</td>
<td>0.300</td>
<td>0.444</td>
</tr>
<tr>
<td>σ6</td>
<td>0.344</td>
<td>0.231</td>
<td>0.168</td>
</tr>
<tr>
<td>σ7</td>
<td>0.262</td>
<td>0.188</td>
<td>0.099</td>
</tr>
</tbody>
</table>

may spend much time on collecting many logs in the grapple and still achieve a good overall result if the work is completed with fewer work cycles. Therefore measuring only the durations of the phases would not be the best way to evaluate the operator skills.

Seven skill metrics were defined using (1) – (4). The metrics were scaled between 0 and 1 so that a high value denotes a good performance related to productivity of the work. In this case, productivity is defined as the inverse of the total time needed to load and unload one full load of logs. Driving time is excluded, because there is a huge variation in the driving distances and conditions.

The metric descriptions are given in Table I. The equations used to obtain the values of the metrics and the associated skill components are also shown in the table. Metric1 describes how many work cycles are needed to load and unload a full load of logs. It is assumed that filling the load with fewer work cycles is efficient. However, this metric is strongly dependent on the size of the collected logs. It is much easier to fill the load space with large logs using fewer work cycles than with small logs. Metric2 describes the complexity of the task sequence and Metric5 the ability to make decisions. The other metrics are based on task difficulty. Metrics3 and -4 describe the average control rates during certain work tasks that require the most motor-sensory skill. Metrics6 and -7 describe the skill related to control the machine functions in parallel rather than in series.

A. Setting up the Experiment

In the forwarder experiment, data were collected during normal forwarding operation from three professional forwarder operators, who are referred to as FOP1, FOP2, and FOP3. The measurements consisted of the data available from the machine CAN-bus including joystick (control lever) motions and driving speed. No additional sensors were used. Part of the work was filmed on a video and the recorded data set was divided to two parts that were used in training and validation of the models.

The operators used the same machine, but they were allowed to set the boom control parameters according to their own preferences. These parameter values were also recorded. Operator FOP2 used a clearly different parameter setup than the others by setting all the boom functions to fastest possible settings or at least close to it. Working time in the measurements was one working shift per operator, except for operator FOP1 whose data were recorded from two work shifts. The operators did not work in exactly the same conditions, but they were forwarding the same timber grade. The average timber size for operator FOP1 was larger than for the others, which slightly increased the relative productivity compared to operators FOP2 and FOP3. Operator FOP3 was known to have a significantly shorter working experience than the other two operators.

The relative productivity of the operators and the average values of the skill metrics are shown in Table II. As explained above, the high values in the table signify good performance. Highest and lowest values are emphasized. The standard deviations of the skill metrics are shown in Table III. The values of the skill metrics and standard deviations are calculated from the observation sequences. The averages were weighted by time because the durations of the observation sequences were not constant.

B. Discussion of the Results

The forwarder operator skill metrics showed to have a clear correlation with productivity. The most productive operator, FOP2 was also the most successful regarding the skill metrics, having the highest values in five of the seven metrics, and close to the highest in the remaining two. On the contrary, the least productive operator FOP3 with the shortest working experience had the lowest values in six of the seven metrics.

The standard deviations of the skill metrics calculated from individual operators are generally smaller than the differences between the metric values of different operators. This suggests that the operators have an individual skill level and a work style and rate that remain relatively constant.

In forwarder work, the complexity of the task sequence (Metric2) does not seem to be a good skill measure, since the differences between the operators’ metric values are small but the standard deviations are high. Forwarding work cycles are basically simple consisting of picking, moving and positioning, which the professional operators master well. Rather than a skill measure, the complexity of task sequence in this case may...
be seen as a measure of the operating conditions or how often exceptional situations appear.

Based on the values of skill metrics, a short analysis of the forwarder operators can be done. Operator FOP2 is very skilled in controlling the machine functions. Despite the fact that he used extremely fast control parameter settings, he is able to control multiple functions in parallel using high control rates. Effective machine controlling combined with good decision making skills result in really high productivity. Operator FOP1 is not as skilled in controlling the machine functions, but has good decision making skills. The productivity level of the operator can be considered as average, though the working conditions regarding productivity were favorable. The least experienced operator FOP3 has low control rates and tends to control machine functions more in series rather than parallel. This results in low productivity that is down to almost half of the most productive operator FOP2.

VII. EXPERIMENTAL RESULTS WITH HARVESTER

The work cycle of the forest harvester is significantly more complicated than that of the forwarder. The results reported in this paper analyze only the stem feeding process. It is the most complicated subprocess of the harvester as it contains both automatic and manual operations. Harvester operator skill evaluation which covers the entire work cycle is left for future work.

Unlike in the forwarder experiment, the data in the harvester case were not collected under supervision. That is, the data consisted of a database recorded during a long period of forest harvester work. There were no descriptions available of the expertise or the experience of the operators. Furthermore, no detailed information about the operating conditions was available, only the physical dimensions and species of the trees were measured. Therefore, the goal of this analysis was to find out whether the skill metrics developed for the stem feeding process can explain the productivity variations between the harvester operators.

A. Setting up the Experiment

The data used in the analysis were gathered during several months period of normal work from two machines. The data consist of the measurements available on CAN-bus during the work. Basically the state of the processing head was measured. This includes the stem feeding speed, position, direction and diameter measurements. Also the control commands given by the operator in automatic and manual feeding modes were measured. The boom crane control lever motions were available on the CAN-bus, but were not measured in this data set. No hydraulic pressure measurements were available. The productivity of the stem feeding process is defined as the ratio between the stem volume (in cubic meters) and the time used to process the stem.

The stem feeding process is divided into work tasks and an algorithm for extracting the observations for the HMM from the measurements was developed. The model is trained and validated with a different data set than the set that is used in this analysis. Modeling the stem feeding process by a HMM is explained in detail in [41].

Two operators from two machines were chosen to be analyzed in this experiment. These operators are referred as HOP1, HOP2, HOP3, and HOP4. The numbers of work shifts for each operator analyzed in this research are 77, 74, 46, and 241, respectively. Thus, the data set from operators HOP1 and HOP2 covers slightly less than four, HOP3 over two and HOP4 almost ten months of work. The performance variables described above are measured from each log, and the values of one work shift are averaged. In total, the analysis data set consists of 438 mean performance data points for each variable.

Based on the task efficiency and complexity approaches, seven skill metrics are defined. The descriptions are shown in Table IV. The metrics are based on the complexity of the task sequence needed to complete the work, the operator’s decision making ability, the success of the decisions made, the proportion of the efficient processing time used in automatic processing mode, and the efficiency of the automatic and manual processing. The equations used to obtain the values of the metrics are shown in the table. Since the data set did not include measurements of the handling of the control levers, the metrics describe mostly the operator’s knowledge of the work technique and strategy (iii) and the ability to plan and make decisions (iv).

The performance of a harvester is highly dependent on the operating point of the machine. In this context the operating point refers to the essential operating and load conditions related to the tree processing such as the tree species, timber grades, mass at hand when processing the tree, etc. Thus, the effect of the operating point needs to be compensated from the performance metrics. This is done by evaluating the metrics with respect to the operating point. Therefore, the compensated metrics in different operating points are comparable with each other. The compensated metrics were scaled so that high values signify good performance. Exceptionally, the low values of Metric6 do not necessarily signify low performance. Instead, Metric6 should be understood rather as a measure of the work style of the operator.

The relative productivities (RP) and the average values of the skill metrics are shown in Table V. As explained above, high values in the table signify good performance, except for Metric6. Highest and lowest values in the table are emphasized. The standard deviations of the skill metrics are shown in Table VI.

B. Discussion of the Results

One can notice from Table V that operator HOP1 performs more tasks to complete the work than the other operators studied here. In fact, the values of all metrics for HOP1 except Metric7 are considerably low when compared to the other operators.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Eq</th>
<th>Sc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metric1</td>
<td>Overall complexity of the task sequence</td>
<td>(4)</td>
<td>(iii, iv)</td>
</tr>
<tr>
<td>Metric2</td>
<td>Ability to make decisions</td>
<td>(1)</td>
<td>(v)</td>
</tr>
<tr>
<td>Metric3</td>
<td>Decision success 1</td>
<td>(3)</td>
<td>(v)</td>
</tr>
<tr>
<td>Metric4</td>
<td>Decision success 2</td>
<td>(3)</td>
<td>(v)</td>
</tr>
<tr>
<td>Metric5</td>
<td>Decision success 3</td>
<td>(3)</td>
<td>(v)</td>
</tr>
<tr>
<td>Metric6</td>
<td>Automatic processing proportion</td>
<td>(1)</td>
<td>(iii)</td>
</tr>
<tr>
<td>Metric7</td>
<td>Efficiency of the processing</td>
<td>(1)</td>
<td>(iii)</td>
</tr>
</tbody>
</table>
The skill evaluation of human operators in mobile working machines was studied using a HMM-based approach. The intent of the human operator, that is, the operational phases and tasks which appear during the work, were recognized by HMMs using only the modest available measurement information. The states of the HMM were defined explicitly to correspond to the actual work phases. The developed skill metrics provide a basis for the operator skill optimization, namely coaching, during normal work. Furthermore, the metrics can be used to form indicators to decide whether a suboptimal performance of the machine is due to the operator’s skills or a technical fault.

The method was applied to machines used in mechanized timber harvesting. Data were gathered during normal operation using the measurements readily available. Four harvester operators and three forwarder operators were compared against their performance in productivity. The experimental results show that the skill metrics derived from the presented approaches make a clear distinction between the operators. The variation of individual operator’s skill metrics during the work is smaller than the differences between the skill metrics of different operators. The skill metrics also clearly reflect on the productivity of the man-machine system.

The duration of the test period has been quite long, but the number of operators studied in this paper was rather small. However, since writing the paper further evidence has been accumulated from several other cases. No contradictory evidence to the paper’s results has been found. Another publication is under preparation, which will further discuss the reliability of the approach as well as the possibility to give feedback to the operator based on the skill metrics.

The most important characteristic of a skilled working machine operator is regarded to be a good total long term work outcome. A skilled operator with a good work technique is able to keep the performance at a high level for the whole work period and under fatigue. Therefore the skill metrics are calculated from a time period of at least one work shift. A long time window also reduces the effect of the error in the evaluated skill level caused by the immeasurable variation in the operating conditions.

The advantage of the task performance metrics as the basis of an operator assessment is that it is possible to directly point out the potential areas of improvement in the operator skills. In most cases this information is sufficient for a motivated operator to learn to achieve even better performance. The skill metrics answer to the question what kind of skills should be improved, but they do not give an answer how the high skill level could be reached. That would require guidance or at least some level of experience from the operator. Therefore, future research will focus on the development of an online intelligent coaching system to improve the performance and proficiency of the machine operators.

**References**


