TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Lauri Palmroth

Performance Monitoring and Operator Assistance Systems in Mobile Machines

Julkaisu 955 • Publication 955

Tampere 2011

Tampereen teknillinen yliopisto. Julkaisu 955
Tampere University of Technology. Publication 955

Lauri Palmroth

Performance Monitoring and Operator Assistance Systems in Mobile Machines

Thesis for the degree of Doctor of Science in Technology to be presented with due permission for public examination and criticism in Konetalo Building, Auditorium K1702, at Tampere University of Technology, on the 24th of February 2011, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2011

ISBN 978-952-15-2533-9 (printed)
ISBN 978-952-15-2549-0 (PDF)
ISSN 1459-2045

Abstract

Studies related to the productivity of machine work often show that there is a large amount of unexplained variation in the productivity, which counts in most of the performance differences between the machines. This unexplained variation includes the skill and actions of the machine operators. It is clear that in machine work, the effect of the human operator is very significant to the overall performance. These differences between the operators possess a substantial potential for improvement. If novice machine operators would be provided with individual feedback and training based on their actual work performance and areas of improvement, right from the beginning, the chances of reaching good skills could be significantly increased. However, any practical and wide-spread means to provide objective and useful real-time feedback to the operators have not been available before. Measuring and accurate modeling of a man-machine system that operates in varying conditions is very challenging, because the skills and work procedures of the human operators controlling the process are always individual.

This thesis presents a method of the recognition of machine work tasks and work cycles based on the combination of multivariate control signals generated by the operator. The recognition of work cycles and tasks is based on Hidden Markov Models (HMM). As the actions of the operator become recognizable, the operator's effect to the overall performance of machine work does not need to be regarded as an unknown disturbance. It also facilitates the evaluation of operators' skill at the task level and the analysis of work technique.

The thesis also presents a method of using intelligent coaching systems (ICS) for example to provide useful feedback to operator training or to support the operators in decision making. The ICS is based on qualitative expert knowledge related to the man-machine work process. It uses skill and performance measures, which are defined for each work task. The values of the performance measures are evaluated

using corresponding statistical reference. The ICS makes observations and gives suitable feedback to the operator in the form of linguistic suggestions. The expert knowledge is formulated as rules of a fuzzy inference system.

Significant performance and productivity improvement in man-machine systems could be gained by enhancing the abilities of the machine operators to perform the work tasks more successfully. Moreover, the methods presented in this thesis are based on the measurements and performance measures that are already available from the process. Thus, implementation of the methods does not increase the manufacturing cost and complexity of the system, since it is not necessary to mount additional measuring equipment. The presented methods for work task and work cycle recognition and skill evaluation of machine operator at work task level have been implemented in industrial applications.

Preface

The research work presented in this thesis was carried out at the Department of Automation Science and Engineering (ASE), at Tampere University of Technology (TUT), in cooperation projects with John Deere Forestry under several research projects. I would like to acknowledge the project funding provided by the Finnish Funding Agency for Technology and Innovation (Tekes), which made the continual and persistent research work possible. I am also grateful for the financial support from Neles Corporation 30-year Anniversary Foundation, Jenny and Antti Wihuri Foundation, Walter Ahlström Foundation, Henry Ford Foundation, and the Finnish Society for Automation.

I would like to thank my supervisor Prof. Pentti Lautala for the valuable encouragement during my research work and for regularly reminding me not to forget to do the writing part of the thesis work also. Many thanks are due to Prof. Raimo Ylinen and Heikki Ovaskainen for the pre-examination of the thesis manuscript and providing valuable comments and criticism, which significantly improved the quality of this thesis.

In the spring of 2001, almost ten year ago, I was a shy undergraduate student, writing one of my first CVs for a job application. After a job interview with Pauli Viljamaa, I got hired for a summer job at the TUT Institute of Automation and Control (ACI). Thanks to all the people at ACI/ASE I have worked with in these years.

I have worked in research projects with very talented and enthusiast people from Control Engineering research group at Aalto University. First of all, many thanks to the head of the group, Prof. Heikki Koivo for the encouragement, positivity and the very important role in successful leading of the research projects. Lots of thanks to Kalevi Tervo, I have been very fortunate to do my research with such a skilled and productive colleague like you. I would also like to thank Jani Kaartinen for the software implementations of the research results in forest machines.

I acknowledge the strong commitment of the industrial partner John Deere Forestry to long term research, which has been a key element for making the results of this thesis possible. Most of the thanks go to Arto Peltomaa and Aki Putkonen, but I also want to send thanks to all the other personnel at John Deere Forestry, who have participated in the practical implementations and test arrangements.

Finally, I owe my deepest gratitude to my wife, daughter, family and friends for the support not only during the thesis work, but throughout my whole life.

Tampere, February 2011

Lauri Palmroth

List of abbreviations

BM	Block maximum (or minimum)
CAN	Controller area network
CDF	Cumulative distribution function
CRT	Choice reaction time
CTL	Cut-to-length timber harvesting method
DCS	Distributed control system
DSS	Decision support systems
ECCS	Extended cooperative control synthesis
EV	Extreme value distribution
FDI	Fault detection and identification
FIS	Fuzzy inference system
FT	Full-tree timber harvesting method
GEV	Generalized extreme value distribution
GPS	Global positioning system
HAM	Human adaptive mechatronics
HAMC	Human adaptive mechatronics and coaching
HMI	Human machine interface
HMM	Hidden Markov model
HPM	Human performance model

ICS	Intelligent coaching system
ID	Task difficulty index
ITS	Intelligent tutoring system
MF	Membership function
MOCM	Modified optimal control model for human operator
OCM	Optimal control model for human operator
OEE	Overall equipment efficiency
PD	Proportional-derivative (controller)
PDF	Probability density function
TEEP	Total effective equipment performance
WS	Warning systems

List of symbols

Overall equipment efficiency

A	Availability
L	Loading
N_g	Number of good units (accepted quality) produced
N_p	Number of units produced
P	Performance
Q	Quality
T_C	Realized cycle time
T_{CS}	Realized cycle time on schedule
T_{Cth}	Theoretical ideal cycle time
T_{TA}	Total available time
T_{TS}	Total scheduled time
ζ	Operating point of machine

Performance indices

d_j, k_j	Scaling coefficients for j^{th} bin
I	Combined performance index
I_a	Adjusted performance index
α_j	Lower bound for j^{th} bin
β_j	Upper bound for j^{th} bin

Hidden Markov models

$\boldsymbol{A}$	State transition probability matrix, $\boldsymbol{A} = \{a_{ij}\}$
$\boldsymbol{B}$	Observation probability matrix, $\boldsymbol{B} = \{b_j(k)\}$
M	Number of (discrete) observation symbols
N	Number of hidden states
O	Observation sequence within $1 \le t \le T$ as $O = \{o_1, o_2, \ldots, o_T\}$
o_t	Observation at time t
P^*	Highest probability of a single path of states (Viterbi-algorithm)
Q	State sequence within $1 \le t \le T$ as $Q = \{q_1, q_2, \ldots, q_T\}$
q_t	State at time t
q_t^*	State at time t, calculated using Viterbi-algorithm
S	Set of possible hidden states, $S = \{S_1, S_2, \ldots, S_N\}$
T	Length of (discrete) observation and state sequences ($1 \le t \le T$)
v_k	Observation symbol, where $1 \le k \le M$
$\boldsymbol{V}_t(i)$	Incremental quantity of Viterbi-algorithm at t for states ($1 \le i \le N$)
$\delta_{i,j}$	Kronecker delta, $\delta_{i,j} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$
$\boldsymbol{\varphi}_t(i)$	Tracking array of Viterbi-algorithm at t for states ($1 \le i \le N$)
$\boldsymbol{\lambda}$	Hidden Markov model, $\boldsymbol{\lambda} = (\boldsymbol{A},\boldsymbol{B},\boldsymbol{\pi})$
$\boldsymbol{\pi}$	Prior probabilities of states

Human performance evaluation

a_{ID} , b_{ID}	Coefficients for index of difficulty I_D
a_{CRT} , b_{CRT}	Coefficients for choice reaction time T_{CR}
C	Execution capacity of a person performing a task
D	Distance to target (task difficulty)
E_i	Total efficiency of the task S_i

f_i	Number of state transitions to S_i from any other state
f_{tot}	Total number of state transitions in state sequence
I_D	Index of difficulty (Fitts' law)
R_t	Amount or rate that resources are consumed between t and $t + 1$
T_M	Average time to complete a task
T_{CR}	Choice reaction time
W	Width of target (task difficulty)

Distributions

γ	Euler's constant ($\gamma \approx 0.577$)
μ	Location parameter
ξ	Shape parameter
σ	Scale parameter

Contents

Chapter 1
Introduction

When the main features of a mobile working machine are evaluated, obvious measures are productivity, performance and fuel economy. Ultimately, the solutions for machine performance follow up can be derived from these rather generic values. A modern machine monitoring system should provide detailed information on machine condition and work performance. Primary points of interest are monitoring of long term deviations in performance and time distributions of each work task. The analysis of the machine's performance should be based on statistical data which consist of measurements gathered over a sufficiently long period of time to ensure reliability. Onboard vehicle computers with modern processors and ample memory and hard drive storage possess a computational and storage capacity that permits the use of advanced algorithms not only to serve in the machine control, but in various field of applications, from machine monitoring and performance assessment to operator assistance systems.

The assessment of the overall performance of the machine is nothing but straightforward, as the work tasks of the machine work are performed by human operator. For example, the productivity of different types of forest machines has been under research in many studies, [Nur06], [Bjo91], [Eli98], [Han00], [Har88], [Kar04], but often there is a large amount of unexplained variation which counts in most of the differences [Gul95]. This unexplained variation includes the actions of the machine operators.

The performance of the man-machine system is affected by the actual technical machine performance as well as the operator's skills, work techniques and varying operational conditions [Bar90]. In many work tasks the human operator closes the control loop of the system. Hence, the skills of the operator have a significant influence on the response times and performance of the overall system. For example in cut-to-length (CTL) forest harvesting, motor-sensory skill as well as cognitive skills and knowledge are required from the machine operator [Ova07]. The operator needs to master an efficient, safe, environmental friendly and economically profitable forest harvesting process. Also, the operator plans the work ahead, makes quick decisions and controls the machine functions at the same time. Modern machines contain a lot of automatic functions that ease the workload of the operator and speed up the process. But on the other hand, the relatively high level of automation also requires a lot of special skills, since the parameters of the machine control system have to be tuned properly in order to reach the best performance of the machine.

According to the recent studies, productivity differences between professional forest harvester operators using same machine and working in similar conditions, may be significant, up to 40% [Ova04]. This performance gap possesses a substantial potential for improvement. The operators learn their work procedures, strategies and techniques mostly on their own and by experience. As a result, there is a lot of variation and tacit knowledge involved with the work procedures and some of these procedures are much more efficient than others [Vaa05], [Vaa04]. If novice machine operators would be provided with individual feedback and training based on their actual work performance and areas of improvement, right from the beginning, the chances of reaching good skills could be significantly increased. However, any practical and wide-spread means to provide objective and useful real-time feedback to the operators have not been available before.

It is clear that the effect of the human operator is very significant to the overall performance of the system. However, measuring and modeling a man-machine system that operates in varying conditions is very challenging. Accurate modeling is difficult because the skills, knowledge and work procedures of the human operator controlling the process are always individual. The amount of available measurement information in a modern machine is exhaustive, but at the same time rather limited, as for example the cognitive actions of human operators are not measurable [Ter10a].

Human performance can be thought as a result of two distinct simultaneous stochastic processes: human cognitive planning and thinking process, and the resultant actions [Yan97]. Planning, the intent hidden behind the actions constitutes a stochastic process which cannot be measured. On the contrary, the resultant actions are observable and form a measurable stochastic process, which would suggest the use of stochastic and probabilistic modeling methods.

This thesis presents a method of the recognition of machine work tasks and work cycles based on the combination of multivariate control signals generated by the operator. Although single control signals look like arbitrary, the separate work tasks can be revealed based on underlying sequences of combined control commands. The recognition of work cycles and tasks is based on Hidden Markov Models (HMM), which have proven to be successful in pattern recognition of dynamic signals [Rab86], [Rab89].

The recognition of work tasks and cycles opens up a great deal of possibilities. As the actions of the operator become recognizable, the operator's effect to the overall performance of machine work does not need to be regarded as an unknown disturbance. Consequently, the monitoring of overall process performance and efficiency will become more accurate. The recognition of the work tasks also facilitates the performance and productivity improvement of the man-machine system by enhancing the abilities of the operator to learn to perform the work tasks more successfully.

The thesis also presents an intelligent coaching system (ICS) for operators of mobile work machines. The purpose of the system is to assist the operators in their work by providing useful information and feedback of the work process. It can be used for example to provide useful feedback for operator training or to support the operators in decision making. The ICS is based on qualitative expert knowledge related to the man-machine work process. It uses skill and performance measures, which are defined for separate work tasks. The values of the performance measures are evaluated using corresponding statistical reference. The ICS makes observations and gives suitable feedback to the operator in the form of linguistic suggestions. The expert knowledge is formulated as rules of a fuzzy inference system [Pal09].

The approach for operator skill evaluation considered in this thesis, is task level skill evaluation via task sequence and work cycle recognition. The approach has proven to be effective in evaluation of operators of mobile working machines [Pal06], [Pal09], [Ter10a]. The operators' skills are evaluated at the task level, because it has been recognized that the operators may possess different levels of skill in different work tasks [Ter09]. Since the overall performance of a man-machine system depends fundamentally on the machine technical performance as well as the skills and expertise of the machine operator [Bar90], the technical performance of the machine is evaluated first, before the operator skill evaluation and the work technique analysis.

Enhancing the abilities of the machine operators to perform the work tasks more successfully may result in significant performance and productivity improvement in man-machine systems. Moreover, the methods presented in this thesis for task and work cycle recognition, task level skill evaluation of machine operator, work technique analysis and operator assistance are based on the measurements and performance measures that are already available from the process. Thus, implementation of the methods does not increase the cost and complexity of the system, since it is not necessary to mount additional measuring equipment.

The main contributions of this thesis and the author can be summarized as follows:

- The author developed a HMM based work task and cycle recognition method for mobile work machines [Pal06], presented in Chapter 4.
- The author developed an intelligent coaching system (ICS) for operators of mobile work machines [Pal09], presented in Chapter 6.
- The author contributed to the development of work task level skill evaluation of mobile machine operators [Ter10a].
- The author contributed to the development of data-driven machine performance evaluation method [Hol05] and developed the soft-sensor algorithms for the CAN network based measurements.
- The author has contributed in the implementation of the presented methods in industrial work processes.

This thesis is organized as follows. Chapter 2 presents a brief review to models and metrics of process and work performance used in industry. Evaluation of machine performance is discussed in Chapter 3. Models of human performance are discussed in Chapter 4. It also presents the HMM based work cycle recognition method for mobile work machines. Evaluation of human skill and performance is discussed in Chapter 5 and operator assistance by decision support and coaching in Chapter 6. Chapter 7 presents a case example of an analysis of a group of operators in CTL forest harvesting. The performance of the machines and that of the operators are analyzed. The aim the extensive but at the same time rather simplified case example is to summarize in a concise manner, what kind of useful information can be discovered about the machine work by applying the methods proposed in the thesis. Finally, the conclusions are presented in Chapter 8.

Chapter 2

Review of process and work performance

This chapter presents a brief review of models and metrics of process and work performance used in industry.

2.1 Ideal performance in industry

A well-known concept in the manufacturing industry is the Overall Equipment Effectiveness (OEE). It is a hierarchy of metrics[1] which focuses on how effectively a manufacturing operation is utilized. OEE index presents ideal, 100% performance when a plant is running at full designed capacity, the full scheduled time and producing perfect quality. Naturally, due to unexpected errors, stoppages and quality losses, it is very unlikely that any manufacturing process can run at 100% OEE. More commonly the target may be around 85%, which is still quite challenging for many processes. [Han05]

The OEE index is obtained by computing

$$OEE = A \cdot P \cdot Q\,, \tag{2.1}$$

where A stands for availability. Availability means the ratio of realized uptime to expected uptime property. This is the portion of the OEE metric that represents the percentage of scheduled time that the operation is available to operate. Availability is often referred to as uptime. P stands for performance, which is the ratio of production to expected production. This is the portion of the OEE metric that represents the speed

[1] In this thesis, quantitative assessments of some features of interest of a system are referred as metrics. Metrics are defined explicitly and they should be objectively measurable, directly or indirectly from the system.

at which the manufacturing process runs as a percentage of its designed speed. Q stands for quality, which is defined as the ratio of non-defective production to total production. Thus, overall equipment effectiveness quantifies how well a system performs relative to its designed capacity, during the periods when it is scheduled to run. [Han05]

Total effective equipment performance (TEEP) measures OEE effectiveness against calendar hours. The TEEP index is obtained by computing

$$TEEP = L \cdot A \cdot P \cdot Q \ , \tag{2.2}$$

where L stands for loading. Loading is the portion of total calendar time scheduled for operation. Thus, the loading metric is a pure measurement of the effectiveness of the work schedule.

With OEE, manufacturing performance is expressed with three measurable components: availability, performance and quality. However, the areas of improvement in the manufacturing performance can also be targeted. Operation related factors are: planned down time, unplanned downtime (equipment failures), set up adjustments, performance reduction due to wear of parts, minor stoppages and idling, reduced speed of production and quality losses.

However, concerning mobile working machines, the human operator has a significant role in the performance, which means that each component depends more or less on the operator. Also the operating conditions have to be taken into account.

The performance portion of the OEE metric is the most challenging to define and measure with mobile working machines. It can be defined using the ratio of realized performance to the designed performance.

$$P = \frac{N_p T_{Cth}}{T_{TA}} \tag{2.3}$$

In (2.3) N_p is the number of units produced, T_{Cth} theoretical ideal cycle time, T_{TA} total available time. Taking into account, that theoretical ideal cycle time may be a function of changing operating conditions, the productivity is defined as

$$P(\zeta) = \frac{N_p T_{Cth}(\zeta)}{T_{TA}} \ , \tag{2.4}$$

where ζ represents the effect of all the measurable operating environment variables, referred as the *operating point* of the machine.

> **Definition 2.1:** All measurable operating environment and production related variables which affect the performance of a mobile working machine are referred as the *operating point* of the machine.

Performance is a pure metric of speed of production and does not take into account the effects of quality and availability losses.

Taking into account the operating point, OEE can be expressed

$$OEE(\zeta) = A \cdot P \cdot Q = \frac{T_{TA}}{T_{TS}} \cdot \frac{N_p T_{Cth}(\zeta)}{T_{TA}} \cdot \frac{N_g}{N_p} = \frac{N_g T_{Cth}(\zeta)}{T_{TS}} = \frac{N_g T_{Cth}(\zeta)}{T_{TS}} \quad , \qquad (2.5)$$

where N_g is the number of good units (accepted quality) produced and T_{TS} the total scheduled time.

The ratio of the number of good units produced per total scheduled time is usually referred as the productivity. OEE can also be expressed in terms of theoretical ideal cycle T_{Cth} time and the realized cycle time on the schedule T_{CS}.

$$OEE(\zeta) = \frac{N_g T_{Cth}(\zeta)}{T_{TS}} = \frac{1}{\frac{T_{TS}}{N_g}} T_{Cth}(\zeta) = \frac{T_{Cth}(\zeta)}{T_{CS}} \qquad (2.6)$$

Respectively,

$$P(\zeta) = \frac{N_p T_{Cth}(\zeta)}{T_{TA}} = \frac{1}{\frac{T_{TA}}{N_p}} T_{Cth}(\zeta) = \frac{T_{Cth}(\zeta)}{T_C} \qquad (2.7)$$

is the similarly defined OEE performance metric, where T_C is the realized work cycle time. It is important to note that in mobile working machines, the actions of the human operator affect considerably, what actually is the realized work cycle time. However, the OEE performance metrics are unable to separate the effects of machine performance and the actions of human operator to the realized work cycle time and thus to overall performance of the system.

OEE is mostly used in manufacturing industry, but the metrics used are rather generic, and they can also be applied to other industrial process, such as machine work. Using the OEE performance metrics, performance of a mobile working machine can be expressed as the ratio of theoretical ideal cycle time (as a function of operating point) and the realized work cycle time. To be able to measure the realized work cycle time, it is required that the separate work cycles can be recognized from the work. Then, the theoretical ideal cycle time can be obtained for example by using statistical data measured from sufficiently large number of machines and operators.

2.2 Human work performance and efficiency

Study of human work performance and work efficiency started in the late 19^{th} century, when industry turned from work of craftsmen to mass production. One of the most well known theories, scientific management is a theory of management that analyzes and synthesizes workflows, with the objective of improving labor productivity. The core ideas of the theory were developed by Frederick Winslow Taylor [Tay11]. The aim of scientific management is to produce knowledge about how to improve work processes. Taylor believed that decisions based upon tradition and rules of thumb should be replaced by precise procedures developed after careful study of an individual at work.

The general approach in scientific management was to develop a standard for performing each job, select workers with appropriate abilities and train them in the developed standard method. The approach also incorporates the shifting of decision making from employees to managers and supporting workers by planning their work in order to avoid interruptions. Taylor also observed that when paid the same amount, workers will tend to do the amount of work than the slowest among them does. This lead to the idea that work should arranged such that workers who are able to produce more also get paid more. Taylor believed that the productivity of workers does not increase unless the workers benefit. [Tay11]

Applications of scientific management have faced many difficulties throughout the history of industrialization. Common difficulties related to over-emphasizing the routine procedures, taking account the individuality of workers and combining the economic interests of employers and employees. It has also been blamed, often by labor unions, that the approach that aims to make work more efficient by removing unnecessary or wasted effort and to increase automation, is responsible for people losing their jobs. However, scientific management was an early attempt to improve processes by treating the process improvement as a scientific problem and studying carefully the work tasks and jobs. The core principles of scientific management also reflect without a question in today's sports, where stop watches and extensive studying of the elements of the athletes' performances play a significant role in the aim for success.

Throughout the history, the development and studies related to improvement of labor productivity have been controversial and sensitive subjects. Without a doubt, more advanced machines, work procedures and the increasing level of automation have reduced the need of manual work, and resulted structural changes in employment. But at the same time the productivity of work has increased significantly, and the freeing up of the labor force has allowed more people to enter higher more skilled or specialized jobs. This has improved well-being and quality of life of people in general.

Although the level of automation in industrial processes is increasing, human operators are present in many tasks and they will also continue to be present, at least in the near future. Many roles for humans in industrial processes currently lie beyond the scope of automation. Tasks requiring subjective assessment or synthesis of complex sensory data as well as high-level tasks such as strategic planning currently require human expertise. In many cases, where automation of industrial tasks would be possible, human operators are still more comprehensive, flexible and cost-effective than mechanical approaches.

Human operators possess a lot of tacit knowledge [Pol66] and different kind of skills. Some operators may also have learned by experience very effective ad hoc procedures to solve problems. There is a lot of potential in increasing the productivity of operators in general by study and transfer of the variety of operators' skills and tacit knowledge. At the moment, that potential has not been fully exploited.

Chapter 3
Evaluation of machine performance

Evaluation of machine performance is discussed in this chapter. Since the overall performance of a man-machine system depends on the machine technical performance as well as the skills and expertise of the machine operator [Bar90], it is extremely important to be able to evaluate the technical performance of the machine before the skills and work technique of the machine operator are being evaluated.

The objective of the evaluation of machine performance is to be able to assess the performance using only the measurements and performance measures that are already available from the process, without adding new measuring instrumentation. This objective rises from two important issues. First, adding new measuring equipment adds the cost and complexity of the machine and may lead to commercially unprofitable solutions. Secondly, the modern machines are often equipped with digital control and communication modules. These systems contain a lot of potentially useful information about the machine technical performance, which could be exploited with sophisticated manipulation of data.

Performance, condition monitoring and fault detection methods are often divided into three main categories: model-based, knowledge-based and data-driven systems. In model-based methods a mathematical model (usually a differential equation system) is built for the monitored system from physical principles and the predictions given by this model are compared with the actual measurements from the system to detect faults. The main limitation with this approach is that many systems are too complex to allow any sufficiently accurate model to be built. Often fault situations are handled by ad-hoc procedures developed by human experts. Knowledge-based methods (e.g. expert systems, fuzzy logic) try to automate the use of this knowledge. The problem with these methods is that it has proven out to be very tedious to gather the knowledge from the experts and to maintain the database as products evolve. [Hol05]

The presented method for machine performance evaluation is data-driven. Data-driven methods try to deduce the properties of the system more or less directly from the available measurement data, which may be readily available in modern production plants and working machines. The method uses dimensionless indices in the interval 0…100 which are purposed to summarize and visualize in a concise and easily comprehensible manner the relevant changes in machine performance. The performance evaluation method takes into account varying load and operating conditions, which often may have nonlinear effects on performance variables. The main advantage of the presented method is that with the use of sophisticated manipulation of the existing measurement data it is possible to produce a useful visualization of the relevant changes in the performance of a complex system. [Hol05]

3.1 Machine performance measures

The chosen approach in machine overall performance evaluation is to use only the available data from the process and not to add any sensors that are not required for normal operation. The approach is suited for applications, where adding new sensors solely for evaluating performance is not feasible. Instead, the indices are computed using the data and measurements the control modules share via for example the CAN bus of the machine. This usually sets a severe restriction to the available data, so process data metrics to describe the features of interest of the system have to be developed. There are several types of metrics that can be used

- Metrics based on measurement levels
- Metrics based on variances of measurements [Har99]
- Metrics computed from one or several measurements using a specialized software (Soft-sensor metrics)
- Time-based metrics (Soft-sensor or physical sensor)

The difference between a soft sensor and physical sensor is that the soft sensor variable is not measured directly, but it is obtained computationally using other known variables as inputs. A soft sensor can enhance variable monitoring by increasing sampling frequency or replacing inaccurate measurements. A soft sensor can also offer a low cost choice for a physical sensor. Soft sensors are typically used to replace sensors with high cost, inaccurate sensors, and for variables where evaluation involves manual sampling and laboratory analyses. The real-time predicted value from the soft sensor model can be used for control purposes. Implementing a soft sensor solution requires a good understanding of the process as a whole, and especially the known, measured variables. [For07]

The machine performance measures used in thesis for machine performance follow-up are mostly soft-sensor metrics based on the data and measurements the control modules share via the CAN bus of the machine. The metrics are derived from this multidimensional data, which have relatively high variance and strong nonlinear correlations in certain measurement variables. Therefore performance evaluation cannot be based on single measurement value. Instead, machine performance evaluation is performed using sufficiently large number of samples.

3.2 Performance indices

The performances of the different parts of the machine are evaluated using several metrics. The measurement variables used in the computation of the index of a particular subsystem must be chosen such that the variables have relevant significance in assessing the performance of the particular subsystem. Usually the metrics are in different scales and variances, which makes difficult to compare their values. Additionally, it is not always evident how to take into account changing operating points of the process.

The proposed approach in [Hol09, Hol05] is to use indices that are dimensionless numbers in the closed interval 0…100 each of which represents the performance of a subsystem of function. Index values near 100 signify good performance. Varying load and operating conditions are taken into account by binning the measurements and computing index values separately for each bin. The essential and measurable load and operating conditions are referred in this thesis as the *operating points* of the machine (Definition 2.1).

Many existing condition monitoring methods require comprehensive teaching data that are labeled with fault information [Wan03], [Chi01], [Qiu05]. When the indices are used, labeled data are not needed, because the indices are related to performance in general, and not to for example certain faults.

The calculation of the indices can be summarized as follows:

I. Removing outliers from the metrics (or measurements)
II. Classification of the data (with respect to operating points)
III. Binning each measurement according to the operating points
IV. Computing the index value for each bin
V. Combining the index values of the bins

There are obviously several possible methods to scale the values measurements to the interval 0…100. The following piecewise linear scaling method was proposed in [Hol05]. First lower and upper bounds, α_j and β_j, respectively, are defined for the variables for each bin j.

In the case where larger performance metric values represent better performance, the following coefficients in (3.1) and (3.2) are computed for each measurement bin.

$$k_j = \frac{100}{\beta_j - \alpha_j} \tag{3.1}$$

$$d_j = -\frac{100\alpha_j}{\beta_j - \alpha_j} \tag{3.2}$$

If smaller measured values are better, the coefficients are

$$k_j = -\frac{100}{\beta_j - \alpha_j} \tag{3.3}$$

$$d_j = \frac{100\beta_j}{\beta_j - \alpha_j} \tag{3.4}$$

With the coefficients k_j and d_j the index for the j^{th} bin can be computed as

$$I_j = \min\left(\max\left(k_j \bar{x}_j + d_j ;0\right);100\right) , \tag{3.5}$$

where $\bar{x}_j$ is the arithmetic mean of the measurement variable for bin j. As the values of this quantity may be less than 0 or greater than 100, the obtained index value is then limited to the interval 0…100 with a hard limit.

The combined index value I for all M measurement bins is the weighted arithmetic mean over all bins

$$I = \frac{1}{N}\sum_{j=1}^{M} N_j I_j , \quad N = \sum_{j=1}^{M} N_j , \tag{3.6}$$

where the weights N_j are the numbers of measurements in each bin and N the total number of measurements.

Obviously, defining the scaling function this way, the selection of the upper and lower bounds, α_j and β_j plays a significant role in the visual sensitivity of the indices. The desired level of sensitivity can be achieved by optimizing the bounds, α_j and β_j, against identification data such that a known decrease of the measurement variables results in desired decrease in the indices.

Even if the compensation bins and bounds are selected correctly, the measurement variables may still contain a lot variance, which is regarded as normal operation. To reduce this tolerable variance in the upper end of the index value range and to accentuate the decrease in the index values, an adjusted index I_a can be computed.

$$I_a = \begin{cases} \frac{8}{5} I, & 0 \le I < 50 \\ \frac{2}{5} I + 60, & 50 \le I \le 100 \end{cases} \tag{3.7}$$

The adjustment of the indices as well as selection of the upper and lower bounds are strongly application dependent and the need for adjustment should be considered carefully. In this adjustment the value 50 of the original index is mapped to 80 in the adjusted index. In the application example presented later in this thesis the upper and lower bounds are chosen such that the values of the original index above 80 in normal operation. An example of this kind of piecewise linear scaling function is shown in Figure 3.1

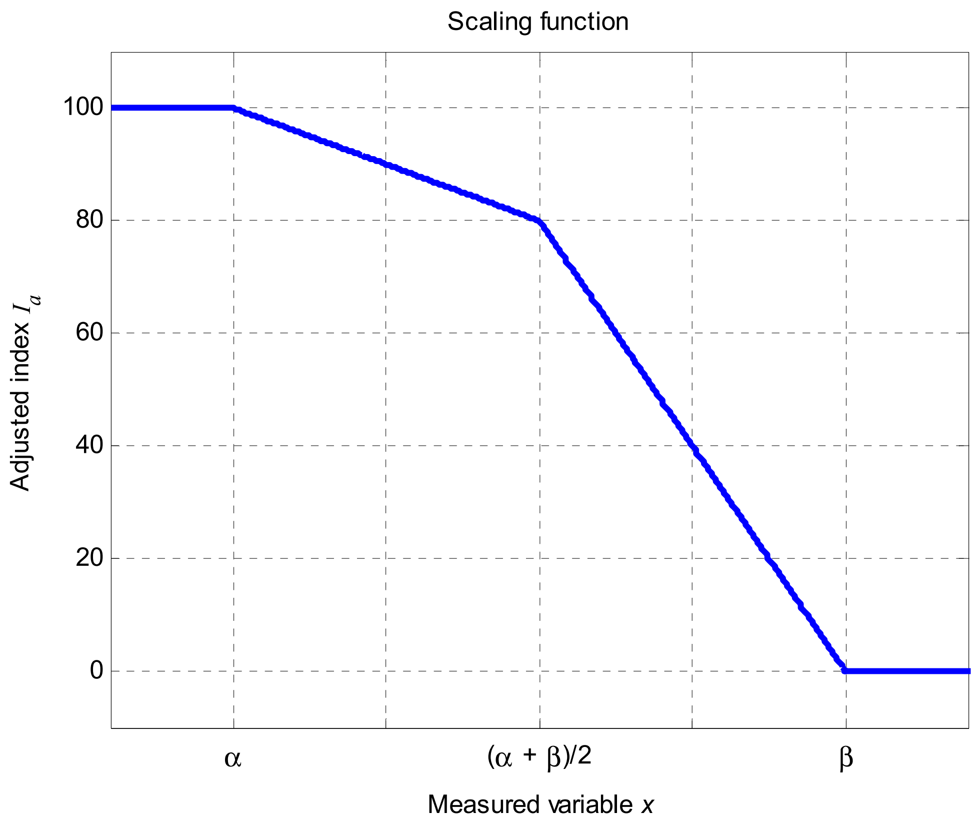

Figure 3.1: Piecewise linear scaling function in case where smaller measured values represent good performance.

As pointed out in [Hol09], there are many ways to scale a measurement value to the interval 0…100. Also, the scaling function can be interpreted to be formed from a cumulative distribution function (CDF), and in general, the formulas of any suitable probability distribution functions can be used instead. For example, it is possible to construct the piecewise linear scaling function in Figure 3.1 from piecewise uniformly distributed probability function. Ultimately, the purpose of the performance indices is to serve as a visualization tool for observing the changes in the plant performance data. This means that the choice of the scaling function and its parameters is subjective and depends much on the properties of the underlying process.

The performance indices can also be used to serve as a tool for fault detection and identification (FDI). Although the performance indices do not necessarily directly point out faults and faulty components, the main advantage is that the performance of the machine can be monitored during normal work. If the performance of a certain subsystem is seen to deteriorate, it is possible to detect faulty components or incipient faults by using specialized test sequences. This way it is possible to reduce mobile machine downtime which is caused by component faults. [Chi01], [Pal04]

Figure 3.2 shows an example of some of the indices related to cut-to-length forest harvester. During the data gathering period, the harvester felled and processed about 16 000 trees. Each one of indices 1-6 represents a different subsystem or function in the harvester. There are several interesting features in the figures. First of all, most indices are close to 90 for most of the time, thus indicating good overall performance. The variance of the indices is slightly different for different indices, because the effect of the random variation caused by the changes in the operating conditions is different for each index. In some cases, like for index 5 in the beginning of the data gathering period, the value of the index starts to decrease steadily and finally collapses to a low level. After this the index returns to a good level, indicating that the cause of low performance in the particular subsystem has been identified and corrected.

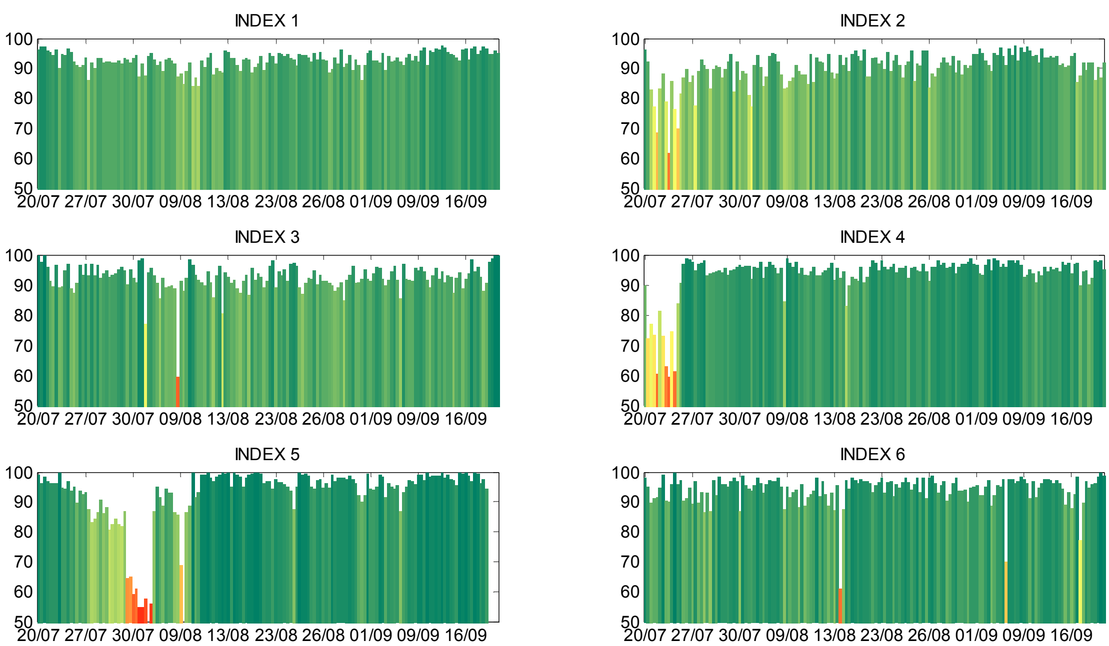

Figure 3.2: Performance indices of forest harvester. [Hol05]

The indices describing different subsystems in one machine should not correlate with each other if the subsystems are physically and functionally distinct. This is indeed the case in the abovementioned example where index 5 decreases while the other indices remain unchanged. Measurement data were gathered over a relatively long period of time, which means that the operating conditions have changed several times. However, the indices show only occasional, but not regular and repeated changes, indicating that compensation of disturbances is working properly.

Since no additional sensors are required, the approach is suited for applications, where adding new sensors solely for evaluating performance is not feasible. Moreover, data-driven approach is well-suited for complex systems, since there is no need for an accurate mathematical model of the system. The concept can be applied to other types of systems, especially when there is a need to compensate the changes in the data due to different operating conditions. The measurement variables used in the indices can be selected using statistical analysis or application experts' knowledge. Although the indices describe primarily the technical performance of the machine, they have proven to be valuable also in terms of condition monitoring of various subsystems and functions of mobile working machines.

The main advantage of the presented method is that sophisticated manipulation of the existing measurement data can lead to useful representation of the performance of a complex system. The method is also able to take into account varying load and operating conditions, which may often have strong nonlinear effects on the performance that is difficult model in practice. The dimensionless indices in the interval 0…100 summarize and visualize in a concise and easily comprehensible manner changes in machine performance. The role of the performance indices is also to ensure that the condition of the machine technical performance is acceptable. Since it is very difficult to separate the operator performance deterministically from the machine technical performance, the technical performance of the machine must be ensured first, before one can draw conclusions about the operator's performance.

Chapter 4
Modeling of man-machine process

This chapter discusses modeling of man-machine process. The approach of this thesis in modeling of man-machine process is essentially the modeling of human actions, decisions and performance while performing machine work. The actions are modeled as a network of tasks, which are consequently modeled as states of a Hidden Markov model (HMM). Different types of human performance models are presented. The chapter also discusses the typical properties of man-machine systems that need to be considered in modeling.

One of the main contributions of the thesis and the author, work task and work cycle recognition method using HMMs, is presented in this chapter. A brief introduction to HMMs is presented, and the use of HMM as a task network model for work and process modeling in a man-machine system. The strengths and limitations of the method are being discussed. An example of HMM as task network model of human actions is shown. In the task recognition of a simple example task network model, task classification success of the HMM based work task recognition method is compared to classification using Bayes decision.

4.1 Human performance models

Human performance models (HPMs) are used in modeling humans performing given tasks. For example system operators, performing supervisory and manual control tasks can be modeled as HPMs. Human performance models are potentially used in system design, development and evaluation [Bar90]. The types of HPMs considered in this thesis are quantitive models (analytic or computer based) of human operators in complex systems. In [Bar90], four types of human performance models are described:

1. *Information-processing model* is based on psychological theories for human information processing and decision making.
2. *Control theoretic model* is a continuous time description of man-machine system performance and dynamic responses. Human operator is modeled as the controller of the system.
3. *Task network model* is a model of human performance and human actions that uses a set of discrete tasks arranged as a network. Models based on this approach typically focus on the time required to complete individual or a collection tasks.
4. *Knowledge-based models* model human cognitive tasks such as planning, decision making and problem solving. Expert systems are an example of this type of human performance models.

Operators of mobile work machines often perform a collection of tasks. These tasks are mostly performed in series, but may overlap in time, compete for the operator's attention and may require simultaneously both cognitive and motor-sensory skills. For example, a fighter pilot is engaged in control of the aircraft, navigation and execution of the mission. The type of human performance model primarily used in this thesis is the task network type. This means that the separate work tasks or work cycles that appear in machine work are modeled as discrete tasks (states) of a task network model. Knowledge-based human performance models are also used. The operator assistance system which constantly monitors and analyses the skills and work technique of the machine operator and provides suitable feedback for the operator, incorporates expert knowledge in the reasoning process.

A well-known model of human actions is the Norman's seven stage model [Nor88]. It divides human actions into processes of forming the goal, execution of the actions to achieve the goal (gulf of execution) and the evaluation of the results (gulf of evaluation). In the modeling the operator's actions and performance, the goals and intentions of the operator are not usually observable or obvious. However, the executed actions can be observed and used to determine the original goals and intentions behind the actions. This can be thought as a reverse process for the "gulf of execution" in the Norman's model [Ter10b].

4.2 Properties of man-machine systems

The man-machine systems related to mobile work machines often have some typical properties that make modeling of the man machine-system rather challenging. [Pal06], [Ter10a]

1. *Exceptions and variations in work cycles.* Machine work is often cyclic and repetitive by nature, but there may be a lot of exceptions and variations in the sequences of consecutive work tasks.
2. *Variations in machine controlling techniques.* Basically, individual operators may have different control techniques for controlling the machine. For example, a common technique in controlling the speed of the boom crane movement is simply to adjust the joystick position proportional to the desired speed. Some operators, however, may prefer a technique, where the joystick position is quickly altered at relatively high frequency in a manner that can be compared to pulse width modulation. Due to the dynamics of the control and hydraulic systems, the resulting actuator speed is smooth and proportional to the width and amplitude of the control "pulses".
3. *Variations of trajectories in work tasks.* The trajectories of work tasks cannot be assumed continuous, because the individual operators may have special boom maneuvers. For example in mechanized forestry, operators often have special tricks for sorting logs, balancing a bunch of logs in the grapple, removal of loose branches from the bunch or aligning the logs in the pile [Pal06]. Forest machine operators sometimes may have to avoid damage to the remaining trees in the forest while maneuvering the boom crane, which also adds variation to the trajectories.
4. *Limited measurement capability.* A common requirement for performance monitoring systems is that they should be based on the normal control data of the machine, without extra sensors. Adding new measuring equipment raises the cost and complexity if the machine and it may lead to commercially unprofitable solutions. Therefore the measurement data is often limited to the control system data generated by the human operator controlling the machine functions.

To overcome the difficulties, some assumptions are made in the process modeling [Pal06], [Ter10a].

- Between operators there can be a lot of variation in the skillfulness of handling the machine functions.
- Each operator has an individual performance level but a rather constant pace and style for controlling the machine.
- As a result of the varying control techniques, trajectories and exceptional situations, the measured control signals should be observed in a sufficiently wide time window. A too short set of samples of a measurement may easily lead to misinterpretations. On the other hand, the observation time window should not be too wide, in order to preserve the desired dynamic properties in the measurements.
- Samples should also include some form of context information such that the based on underlying sequences of combined control commands the separate work tasks can be revealed.

As a conclusion, the amount of variation in the process to be modeled makes it difficult to use deterministic models, which use cause and effect logic to describe the behavior of the man-machine system. The associated randomness suggests the use of stochastic and probabilistic models.

4.3 Hidden Markov models

A Hidden Markov Model, 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 [Bun00]. Due to the stochastic and dualistic nature of HMMs, they are very often used in modeling the human actions and performance [Rab89].

Since late 1980s, the research of HMMs has escalated. The application areas have widened from speech recognition [Lev83] to fault diagnostics [Yin00], [Ge08], prognostics [Qiu05], condition monitoring [Bun00], communications engineering [Bru92], human intent recognition [Zhu08], human gesture recognition [Man01], task recognition in surgery [Mur03], [Ros01], [Ros03] and human behavior modeling [Brd09].

A HMM with one dimensional discrete observation probability distributions can be defined as follows. A Hidden Markov Model λ consists of the following elements [Rab86], [Rab89]:

1. A set of possible hidden states, $S = \{S_1, S_2, \ldots, S_N\}$. The state at time t is denoted as q_t and the state sequence within $1 \leq t \leq T$ as $Q = \{q_1, q_2, \ldots, q_T\}$.
2. Observation symbols v_k, where $1 \leq k \leq M$. Observation at t is denoted as o_t and the observation sequence within $1 \leq t \leq T$ as $O = \{o_1, o_2, \ldots, o_T\}$.
3. State transition probabilities $\boldsymbol{A} = \{a_{ij}\}$, representing the probability of transition from state S_i to S_j, where $a_{ij} = P(q_{t+1} = S_j \mid q_t = S_i)$, $1 \leq i, j \leq N$.
4. Observation probabilities $\boldsymbol{B} = \{b_j(k)\}$, representing the probability of observing symbol v_k in state S_j, where $b_j(k) = P(o_t = v_k \mid q_t = S_j)$ and $1 \leq j \leq N$, $1 \leq k \leq M$. Note that if the observation v is a continuous set instead of discrete symbol v_k, b_j is a probability density function.
5. The prior probabilities of states $\boldsymbol{\pi} = \{\pi_i\}$, where $\pi_i = P(q_1 = S_i)$, $1 \leq i \leq N$.

For convenience A HMM is usually represented

$$\boldsymbol{\lambda} = (\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\pi}) \tag{4.1}$$

The state transition matrix A defines the type of the HMM. The model is called fully connected (ergodic), if all the elements of A are nonzero. Then every state of the model could be reached in a single step from every other state of the model. In practice, an ergodic model has the property that every state can be reached from every other state in a finite number of steps. For some applications, other types of HMMs have been found to account for observed properties of the signal being modeled better than the standard ergodic model. Other common types are for example left-right or cyclic structures.

A HMM model is called a left-right model if the underlying state sequence associated with the model has the property that as time increases, also the state index increases (or stays the same), i.e., the states proceed from left to right. Clearly the left-right type of HMM has the desirable property that it can readily model signals whose properties change over time e.g., speech. The cyclic structure of HMM differs from left-right such that, also the transition from last state to first state is allowed. In addition to these types, there are many other possible variations and combinations of state transitions that can be used to obtain the desired dynamical behavior of the model. [Rab89]

4.3.1 Basic inference problems associated with Hidden Markov models

With HMMs, there are three basic problems of interest [Rab89]:

I. *The estimation problem*, estimate the parameters in model λ, given the observation sequence O. The estimation problem is used in training of the model. If there is labeled training data available, where the hidden states are known, the model parameters can be trained using supervised learning procedure. Then the model parameters are adjusted to maximize the probability of the observation sequence and corresponding sequence of hidden states. If labeled state information is not available, the model parameters can be estimated using iterative learning procedures, such as the Baum-Welch method [Bau70].

II. *The evaluation problem*, evaluate the probability of the observation sequence O being produced by model λ, $P(O \mid \lambda)$. The solution of the evaluation problem is useful in classification, since it allows us to evaluate how well the given model matches a given observation sequence. The evaluation problem can be solved efficiently using forward-backward procedure [Rab86].

III. *The decoding problem*, decode the most likely sequence of states Q that maximizes the probability of the observation sequence O, given model λ. In the decoding problem, the most likely path of states that led to the observation sequence is calculated. Thus, the hidden states are uncovered. An often used recursive method for solving the decoding problem is called the Viterbi-algorithm [Vit67]. The solution to decoding problem is useful in recognition applications, where the hidden states represent the classified properties.

4.3.2 Estimation of the model parameters from training data

If there is labeled training data available (where the hidden states are known), the model parameters can be estimated directly from the training data sequences. In the following, the concept of Kronecker delta is used to denote whether two arguments are equal or not. The Kronecker delta is defined as

$$\delta_{i,j} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}, \tag{4.2}$$

which is a useful notation when computing state transition frequencies.

The parameters of the state transition matrix A in an HMM are estimated using the realized training sequence of (hidden) states $Q = \{q_1, q_2, \ldots, q_T\}$. First, the state transitions (from state $1 \leq i \leq N$ to state $1 \leq j \leq N$) in the training data are summed up. The elements of the state transition matrix A are obtained by [Rab89]

$$\hat{a}_{i,j} = \sum_{t=1}^{T-1} \delta_{i,q_t} \delta_{j,q_{t+1}} \tag{4.3}$$

To obtain a proper probability matrix, the sums in the rows a_i are normalized to 1.

$$a_i(j) = \frac{\hat{a}_i(j)}{\sum_{k=1}^{N} \hat{a}_i(k)} \tag{4.4}$$

Similarly, the observation probabilities are obtained using the realized training sequence of (hidden) states $Q = \{q_1, q_2, \ldots, q_T\}$ and the corresponding observation sequence $O = \{o_1, o_2, \ldots, o_T\}$. Here the number of discrete states is $1 \leq j \leq N$ and the number of discrete observations is $1 \leq k \leq M$. The elements of the observation probability matrix B are obtained by [Rab89]

$$\hat{b}_{j,k} = \sum_{t=1}^{T} \delta_{j,q_t} \delta_{k,o_t} \tag{4.5}$$

Again, the row sums are normalized to 1.

$$b_j(k) = \frac{\hat{b}_j(k)}{\sum_{i=1}^{M} \hat{b}_j(i)} \tag{4.6}$$

A drawback of the supervised estimation of model parameters is that if there is not enough training data to represent and generalize well the modeled phenomena, overfitting with the training data may occur. It is also possible that not all the relevant combinations of state transitions and observations appear in the data, which results zero probabilities in those cases. These practical difficulties can often be avoided by adding *pseudocounts* to the elements of matrices A and B according to prior belief of the frequencies, and normalizing the matrices again using (4.4) and (4.6).

If labeled training data is not available, the parameters of a HMM can be obtained using unsupervised training procedures, such as Baum-Welch method [Bau70] (or equivalently expectation-maximization, EM, [Dem77]), gradient techniques [Lev83] or using re-estimation procedures, that benefit the Viterbi algorithm for a recursive re-estimation of the model parameters starting from some initial values [Rab89]. However, the optimization surface may often be very complex and have many local maxima. Consequently, it is not guaranteed, that these methods yield practically usable models, especially if the goal of the training is to have the states of the HMM correspond to actual physical events, the actual work tasks and cycles performed by human operator. Therefore the use of supervised training procedure is recommended, or if re-estimation procedures are used, the initial values of the model parameters should be chosen very carefully.

4.3.3 Viterbi-algorithm

The Viterbi-algorithm [Vit67] is a dynamic programming algorithm for decoding the most likely sequence of hidden states given a sequence of observed events. The algorithm makes the following assumptions:

- At any time the system being modeled is in one of the finite number of states.
- The observations and hidden states must be in two aligned sequences of same length, such that each observation symbol within the observation sequence corresponds to exactly one hidden state in the resulting sequence of hidden states.
- Computing the most likely hidden sequence up to a certain point t depends only on the observation o_t at point t, and the calculated most likely state sequence up to point q_{t-1}.

- While multiple sequences of states (paths) can lead to a given state, at least one of them is a most likely path to that state. This is a fundamental assumption of the algorithm because the algorithm examines all possible paths leading to a state and keeps only one, the most likely.

These assumptions allow a recursive computation of the most likely sequence of hidden states. To be able to find the best matching sequence, an incremental quantity $V_t(i)$ has to be defined (for states $1 \leq i \leq N$):

$$V_t(i) = \max_{q_1, q_2 \cdots, q_{t-1}} P(\{q_1, q_2, \ldots, q_{t-1}, S_i\} \mid \{o_1, o_2, \ldots, o_t\}) \tag{4.7}$$

Thus, the quantity $V_t(i)$ is the best score (highest probability) along a single path which accounts for the first t observations and ends in state S_i at time t. This probability is maximized using induction.

$$V_{t+1}(j) = \max_i \left(V_t(i) a_{i,j}\right) b_j(k), \quad o_{t+1} = v_k \tag{4.8}$$

To actually retrieve the state sequence, it is also necessary to keep track of the argument (state) which maximized (4.8), for each t and j. This is done by defining a tracking array $\varphi_t(j)$ (or back pointer), which keeps track of the most likely state transitions.

The Viterbi algorithm has four steps, initialization, recursion, termination and backtracking of state sequence. At initialization, the values of the quantity $V_t(j)$ and the state sequence back tracking array φ are initialized (for states $1 \leq j \leq N$):

$$V_1(j) = b_j(k) \pi_j, \quad o_1 = v_k \tag{4.9}$$

$$\varphi_1(j) = 0 \tag{4.10}$$

Then the values are updated recursively using the sequence of observations $o_2, \ldots, o_T$, and maximizing the probability of the possible state transitions (from state $1 \leq i \leq N$ to state $1 \leq j \leq N$).

$$V_t(j) = \max_{1 \leq i \leq N} \left(V_{t-1}(i) a_{i,j}\right) b_j(k), \quad o_t = v_k \tag{4.11}$$

$$\varphi_t(j) = \arg\max_{1 \leq i \leq N} \left(V_{t-1}(i) a_{i,j}\right) \tag{4.12}$$

Note that after computing the incremental quantity $V_t(j)$ and state sequence back tracking array, only the best survives and all other paths are discarded. As a result, at each time t, only the value of $V_{t-1}(j)$ is required to be able to compute the new values and the older values do not need to be stored in memory.

In the termination step, the value of P^*, the highest probability of a single path of states and the final state q_T^*, are obtained by:

$$P^* = \max_{1 \le i \le N} \left(V_T(i) \right) \tag{4.13}$$

$$i = \arg\max_{1 \le i \le N} \left(V_T(i) \right), \quad q_T^* = S_i \tag{4.14}$$

Finally, the most likely sequence of states can be traced back from the final state q_T^*, using the state sequence back tracking array φ for time instants $t = T - 1, T - 2, \ldots, 1$.

$$i = \varphi_{t+1}(j), \quad q_t^* = S_i, \quad q_{t+1}^* = S_j \tag{4.15}$$

If the sequence of observations (and states) is long, there might be numerical difficulties calculating (4.11). To avoid this, implementations of the Viterbi algorithm often use logarithmic scales. This is done by taking logarithms of the model parameters (***A***,***B***,***π***), and using summation instead of product in (4.9) – (4.12).

4.3.4 Example of a Hidden Markov Model

An example of a simple HMM is shown in Figure 4.1. The parameters of the model are

$$
\begin{aligned}
A &= \begin{bmatrix} 0.65 & 0.34 & 0.01 \\ 0.02 & 0.58 & 0.40 \\ 0.19 & 0.01 & 0.80 \end{bmatrix} \\
B &= \begin{bmatrix} 0.59 & 0.05 & 0.25 & 0.01 & 0.10 \\ 0.02 & 0.27 & 0.01 & 0.50 & 0.20 \\ 0.01 & 0.15 & 0.04 & 0.20 & 0.60 \end{bmatrix} \\
\pi &= \{0.7, 0.2, 0.1\}
\end{aligned}
\tag{4.16}
$$

The model has three hidden states (N = 3) and there are five discrete observation symbols (M = 5). The model is fully connected, i.e. all the elements of the state transition matrix A are nonzero. However, the structure of the model has some cyclic features, for example the most probable transition from state 1 is to state 2, from state 2 to 3 and from state 3 to 1. Hence, the most probable state transitions form a cyclic path, although all other kinds of paths are allowed as well. The highest state transition probabilities appear on the diagonal elements of A, indicating that the model will more likely stay in each state (for more than one observation) than make transition to some other state. It must be noted, however, that at each time t, the state transition probabilities are independent of the state transition history, and therefore the sequence of observations actually determines if and when there will be a transition from a state to some other state.

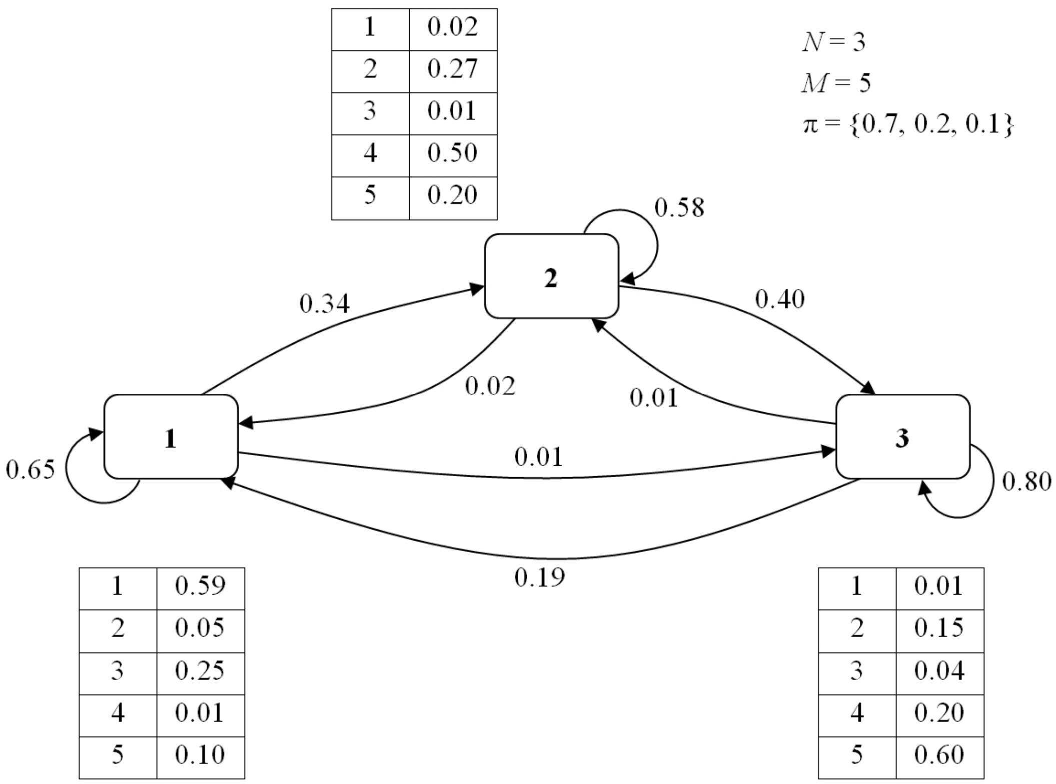

Figure 4.1: An example of a Hidden Markov model (4.16).

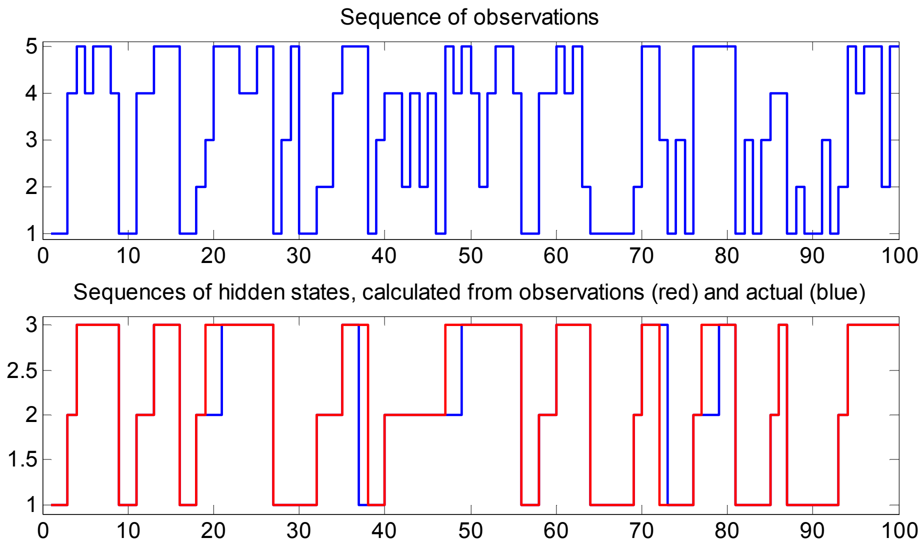

Figure 4.2: An example of an observation sequence produced by (4.16) and the most likely sequence of hidden states decoded using Viterbi-algorithm.

Figure 4.2 (upper) shows a random observation sequence produced using the probabilities of the model (4.16). The figure (lower) also shows the underlying sequence of hidden states (blue) and corresponding most likely sequence of hidden states decoded from the observations using Viterbi-algorithm (red). The purpose of this example figure is to show how the HMM and the Viterbi-algorithm work. Although a randomly generated sequence of observations may seem more or less arbitrary at first glance, the sequence of (hidden) states, revealed by the Viterbi-algorithm, clearly has a cyclic pattern, which is due to the state transition probabilities.

The recognition of the hidden states, using the Viterbi-algorithm, is based on the sequence of discrete observations, which can be measured from the process at constant sampling rate. However, it is important to note that if appropriate, it is not necessary to use constant sampling rate, since the Viterbi-algorithm only considers the mutual order of the discrete observations, not the sampling time in between. This feature is very useful and makes modeling with HMMs very flexible, especially if the dynamics of the process is time variant. Hence, the observations do not need to be tied to a constant sampling rate, they can be observed from the process each and every time something interesting happens.

4.4 Work and process modeling with Hidden Markov models

The performance of a human operator, completing a given task is a time-varying stochastic process. The performance of a human operator depends on several factors: the mental agility and workload of the operator, situation awareness, the disturbance of circumstances, the imperfect and noisy nature of the human sensory processes and the operator's skill associated with the task [She02]. Yang et al. [Yan97] described the human performance as a result of two distinct simultaneous stochastic processes:

1) Human cognitive planning and thinking process
2) The resultant actions

Planning, the intent hidden behind the actions constitutes a stochastic process which cannot be measured. On the contrary, the resultant actions are observable in a man-machine system and form a measurable stochastic process. Consequently, the goal in the man-machine system modeling is to recognize the operator's intent based on measurable actions.

The proposed method for operator intent (or task) recognition, by recognizing the work tasks and cycles conducted by the human operator, is based on explicitly defined HMM states. In the proposed method, the model is trained such that the states of the HMM correspond to the actual work tasks of the operator [Pal06]. The work tasks form a *task network* model. All the tasks, which are needed to complete whole work cycle, are modeled. A task could also consist of multiple HMM states instead of just one. Also some common exceptional situations, which might appear during the work, can also be defined as model states.

Solutions to all the three basic HMM problems may be needed in work and process modeling (Chapter 4.3). At first, the model parameters are estimated using supervised learning procedure (estimation problem). The set of observations O are features extracted from the available measurement data and the open loop control signals given by the operator. In some cases the work may be structured such that the work tasks form different, recognizable work cycles that are modeled by a set of separate HMMs instead of just one. In that case, the probabilities of measured observation sequences are calculated for all models, and the model that best matches the observations is chosen (evaluation problem). Finally, given the observation sequence and the HMM, the most likely path of states (tasks) is calculated using the Viterbi-algorithm [Vit67] (decoding problem). [Pal06]

An example of task network human performance model, modeled with HMM is shown in Figure 4.3. This example model of human actions covers some activities that humans may typically perform during the morning, starting from the state of sleep and reaching to the state of full readiness for the day's productive activity at work. The HMM states that model the human activities in the "*Morning*" model are defined as:

[S1] *Turn off alarm*

[S2] *Hit the snooze button and go back to sleep*

[S3] *Take a shower*

[S4] *Brush teeth*

[S5] *Get dressed*

[S6] *Have a cup of coffee*

[S7] *Eat breakfast*

[S8] *Read the morning newspaper*

[S9] *Go to work*

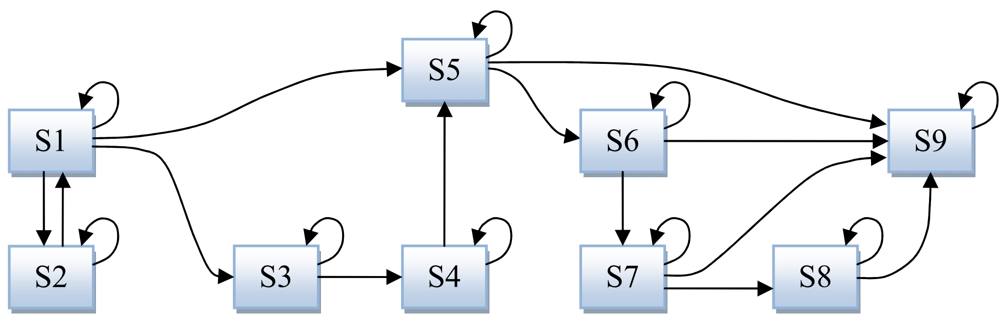

Figure 4.3: Task network model "*Morning*".

The task network structure, in principle, may allow state transitions from any state to all states, but some of the state transitions are more likely than other. Some typical state transitions are drawn in the figure. Many different combinations of paths through the states (activities) are also possible. For example, in the "Morning" model, let's say on Monday morning, one might do all the tasks (visit all states). But on another morning, let's say on Friday morning, one might feel tired and repeated transitions between states "*Turn off alarm*" and "*Hit the snooze button and go back to sleep*" may lead to such hurry that there is only time to visit the state "*Get dressed*" before finally reaching the state "*Go to work*". Thus, in a task network model, the resulted path of states (tasks) may be slightly different for each individual person and morning.

4.5 Discussion

4.5.1 Strengths and limitations of HMMs

A basic theoretical strength of HMMs is the combination of stationary processes and the temporal relationship among the processes in a well-defined probability space [Jua91]. This combination allows studying these two separate aspects of modeling using one consistent framework. Modeling with HMMs is also flexible, because a HMM can represent arbitrary distributions for the observations and the next values of the state variables. The implementation of the method is also quite straightforward and easy.

Another important point to note is the ability of a HMM to take into account the temporal context of the current observation, that is, the observation history. In first-order Markov models, the probability of being in a given state at time t depends simply on the state at time t - 1. However, the Viterbi algorithm takes a decision on the states based on the whole sequence and does not simple-mindedly accept the most likely state for a given time instant. It means that single particularly unlikely observations in an otherwise reasonable sequence do not cause falsely classified states, which would happen using static classifiers that make a decision separately at each time instant. Despite of the limitation of first-order Markov process, the HMMs have been found to be very usable for this kind of practical pattern recognition problem of (multivariate) dynamical signals, with discrete valued (hidden) state variables.

As an illustration of the dynamic classification properties, in the following, the classification of a given sequence of observations is classified using two different methods, HMM and Bayes decision. In Bayes decision, the probabilities of each state S_j ($1 \leq j \leq N$) in the state sequence are evaluated independently of each other and the most likely is chosen on each time instant t. The posterior probability of q_t being state S_j given the observation o_t is proportional to

$$P\left(q_t = S_j \mid o_t = v_k\right) \propto P\left(S_j\right) P\left(o_t = v_k \mid q_t = S_j\right) , \tag{4.17}$$

where $P(S_j)$ is the prior probability of state S_j and $P(o_t = v_k \mid q_t = S_j)$ the conditional probability of observing symbol v_k at state S_j. The prior probabilities $P(S_j)$ are obtained from the frequencies that each state appears in the training data. The conditional probability of observing symbol v_k at state S_j is actually the same as $b_j(k)$ in the HMM.

The classification performance of these two methods for an artificial test data produced by the task network model "*Morning*" is illustrated in the Figures 4.4 and 4.5. Figure 4.4 shows the data, the sequence of observations (top) and the sequence of (hidden) states in the "*Morning*" model (bottom). Figure 4.5 shows the sequences of states calculated using HMM (top) and Bayes decision (bottom).

The data of the "*Morning*" model, seen from the Figure 4.4, consists of five mornings. It seems that each morning is slightly different from each other, but there is a clear pattern in the sequence of the states. The observations, however, seem more or less arbitrary at first glance. Nevertheless, a reasonably good classification of the sequences of states can be obtained by a HMM, as seen from the Figure 4.5. The sequences of the actual states of mornings from Monday to Thursday were used as the training data for both methods. The data of Friday morning, which has a slightly different pattern of states, was left as an additional new data for validation of the classification performance.

The HMM was trained using (4.3) – (4.6). To avoid zero probabilities in the HMM, pseudocounts were used, by setting the initial values of each element of A and B matrices in (4.3) and (4.5) to one. The observation sequences, the sequences of known states and the Matlab code for the HMM calculations are presented in Appendix A.

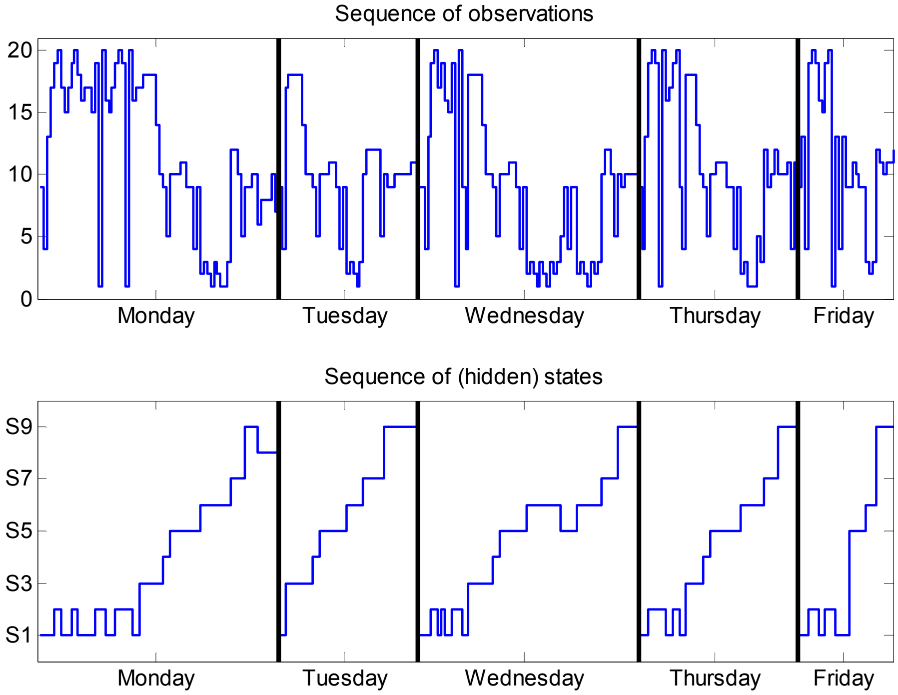

Figure 4.4: Sequence of observations and the sequence of actual (hidden) states in the "Morning" model.

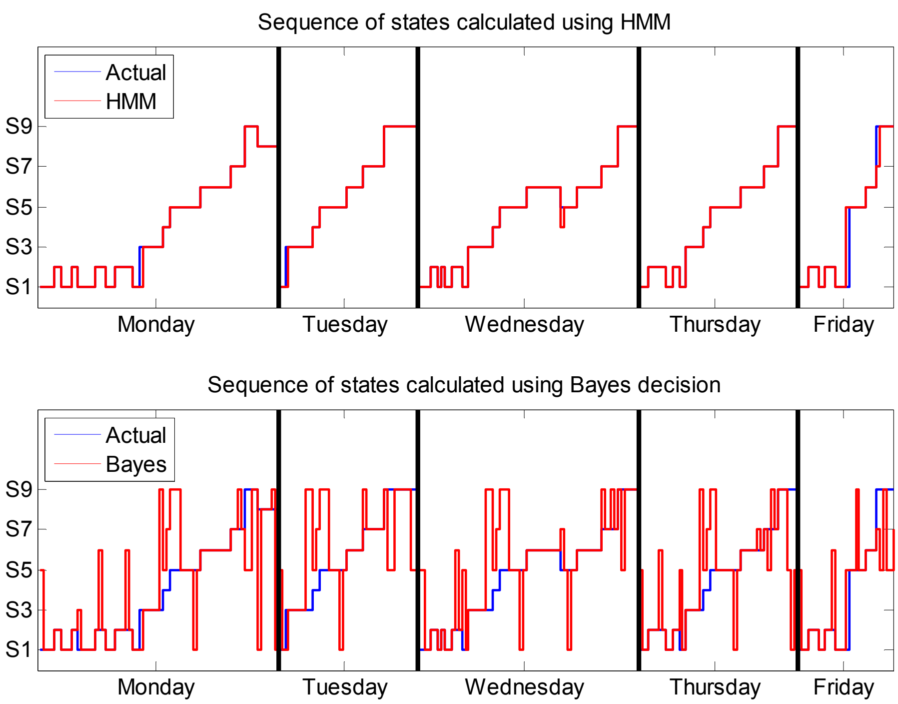

Figure 4.5: Sequence of states calculated from the observations using HMM and Bayes decision.

The classification performance, shown in Figure 4.5 was significantly different on these two separate methods. Since some of the individual observations contain contradictory information, it is natural that the Bayes decision, applied to each individual observation of the sequence, cannot recognize the correct pattern of the state sequence. Only 71% of the states were recognized correctly. Furthermore, the Bayes classifier was unable to detect state S4 at all. On the other hand, the HMM that takes into account the temporal context of the observations is able to detect the states correctly for 98 % of the observations. The pattern of state transitions is also recognized quite accurately.

4.5.2 Feature extraction

This thesis is focused to the situation, when it is not possible to extract the state information directly from the measurements, which is usually the case with mobile working machines. Another initial assumption is that the inclusion of extra measurement devices is not feasible and therefore the measurement capability may be rather limited. Thus, the data acquisition and intent (or plan) recognition systems are based mainly on the control signals given by the human operator. If there are sensors that provide information, which improves the accuracy of the classification, they are also used.

Feature extraction is a crucial phase in many practical pattern recognition problems. For successful pattern recognition, it is important to learn the application domain to obtain relevant prior knowledge and then set realistic goals for the application. Data cleaning and preprocessing should be performed thoroughly, in order to obtain consistent data. Data reduction and transformation is also important: One should try to find the most useful features and to reduce the number of variables or dimensions, instead of using all possible features. It has frequently been observed in [Pal06] and [Pal09] that in practice, beyond a certain point, the inclusion of additional features leads to worse rather than better performance. [Dud00]

When a HMM is defined such that each state corresponds to a real work task, it is very important to preprocess the input data such that the classification of the work tasks can be done as reliably as possible. The set of observations for HMMs should provide consistent information, such that the properties that are intended be classified actually can be revealed. [Dud00] For example, instead of using all the control signals separately, the set of observations should be created by extracting features that combine the relevant information of the control signals. In the machine work, where operator is controlling boom crane using joystick, this means that the important features can possibly be the direction and average displacement of the joystick in a selected time window. It is also possible to combine several features from different joystick axles in order to reduce the number of the discrete observations. [Pal06]

A practical recognition system for man-machine systems should also deal with the differences of the individual controlling techniques, trajectories and control parameter setups, among the different machine operators. Due to these reasons, in this thesis, the features that constitute the set of observations of the HMMs, are formed such that the multivariate measurements (continuous or discrete) are quantized as one dimensional discrete observation symbols v_k (where $1 \leq k \leq M$). Also, the quantization of the data is sparse in order to obtain as small a set of observations as possible (small M). Ultimately, the selected set of features (or observations) is always highly application specific. [Pal06]

4.5.3 Conclusion

The operational phases and work cycles that appear during machine work, can be defined as a network of tasks, which consequently can be modeled using HMMs. As a result of the recognition of the work tasks and cycles, a lot of other advantages are provided. The method offers valuable information for machine performance assessment and condition monitoring. Time distributions, work task and cycle durations and longer term consequences can be observed. These can be used to assist the operator to perform better as well as to indirectly monitor the performance and condition of the machine.

In [Pal06], it was observed from the measurements that the machine operators often have the ability to compensate adverse gradual degradation in machine technical performance without being aware of the actual change. Depending on the type of the work cycle and the operational phase, certain sub-functions of the machine are more dominant. This allows indirect estimation and isolation of possible failures and also gives feedback for the system parameter adjustments and machine component setup. If the degradation of machine performance has been observed by the changes in the work task time distributions, but the available real-time measurements do not enable the isolation of the fault, it can be detected by launching special test sequences [Pal04].

Regarding the man-machine interface, the recognition of operational phases provides valuable information for the evaluation of the operators' skills and how well the operator can handle the machine. Thus, the method gives valuable tools for the machine usability development and in operator training. [Pal09]

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 [Ter10a]. 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. It has frequently been found in [Pal06], [Ter06], [Ter07], [Aul09], [Pal09], [Ter09], [Aul10] and [Ter10a] that this kind of HMM based recognition of work tasks and cycles can be implemented successfully in many kind of applications related to mobile machines.

Chapter 5

Human performance evaluation

Human performance and evaluation of it are discussed in this chapter. A skill evaluation method for mobile working machine operators is presented. Skill evaluation is performed at the task level and the defined skill metrics are based on task and task sequence recognition. Statistical reference is used to be able to obtain relative performance levels. The relative human performance level is obtained using cumulative distribution functions that are approximated from the reference data. Statistics of human performance measures are modeled using generalized extreme value (GEV) distributions.

5.1 Skill evaluation

5.1.1 Related work around 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.

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 (index of difficulty), I_D [Fit52]. Difficulty index is defined as

$$I_D = \log_2\left(\frac{2D}{W}\right), \tag{5.1}$$

where D is the distance to target and W the width of the target. A common formulation of the original form of I_D by Fitts, is the Shannon formulation. The Shannon formulation has origins in electronic communications systems [Sha49] and interprets the task difficulty in the sense of human effective information processing (or execution) capacity [Mac91].

$$I_D = \log_2\left(\frac{D}{W}+1\right) \tag{5.2}$$

Using the difficulty index I_D, Fitts' Law defines the average time to complete a task, T_M, as

$$T_M = a_{ID} + b_{ID}\cdot\log_2\left(\frac{D}{W}+1\right), \tag{5.3}$$

where, a_{ID} and b_{ID} are constants. In (5.3), the Shannon formulation of I_D is used. Using I_D and T_M, execution capacity C of a person performing a task is defined as the ratio

$$C = \frac{I_D}{T_M}. \tag{5.4}$$

Another important concept related to human cognitive skill is choice reaction time, CRT. It is formally very similar to Fitts' law and describes the capacity of a person to process cognitive information, related to multiple choices. CRT is often referred as Hick's law or Hick-Hyman law [Hic52], [Hym53]. The choice reaction time T_{CR} is defined as

$$T_{CR} = a_{CRT} + b_{CRT}\cdot\log_2(n+1), \tag{5.5}$$

where n is the number of choices, a_{CRT} and b_{CRT} are constants.

An extension for Fitts' law for the skill evaluation of human operators, characterized by more than one variable using a probabilistic skill index was proposed in [Koi97]. It considers the execution capacity (5.4) as a random variable, which is characterized by the statistical mean and variance. These parameters are introduced for skill evaluation by calculating the statistical execution capacity of a person and that of a group of persons. In this way, the person's task performance can be evaluated relative to that of the average performance of the entire group. A probabilistic skill index, I_S, is defined as [Koi97]

$$I_S = P(T_{MD} < \alpha \mid d_1 \leq ID < d_2), \tag{5.6}$$

where the probabilistic skill index is the conditional probability of T_{MD}, the movement time being less than constant α, given the task difficulty index is between d_1 and d_2. In (5.8) it is assumed that the probability density functions related to T_{MD} for an individual or for a group of individuals are known. In [Koi97] the probability distributions of the measurement data were approximated using gamma distribution and the index for multiple skills was defined using multivariate distribution.

However, a problem with methods based on Fitt's Law is that they apply only for a simple task execution. Often the widths of the targets and distances cannot be defined. Hence, it is very hard to define the index of difficulty to apply for various kinds of tasks performed by humans.

A control theoretic approach human skill evaluation in task execution is to model human as a controller. That is, model the input-output relationship between the states of the controlled dynamical system and the control signals given 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. [Kur04], [Hid06], [Yu04] 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 that are used as a model of human controller, are PD-type and suitable only for modeling of simple tracking tasks. 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 often saturates (operator uses "full throttle"). This makes the modeling even more challenging.

Several analytic approaches for evaluation and modeling of human operator's performance have been proposed. Some of the most promising ones are based on the optimal control model for human operator (OCM) [Bar70] and the modified version of it, that is, the modified OCM (MOCM) [Dav92]. The models have been used to design optimal control laws for systems with human control. Extended cooperative control synthesis (ECCS) [Dav94] provides theoretical framework for designing assist control laws for human operated machines.

Another research branch which has gained increasing attention recently is the human adaptive mechatronics (HAM) systems. A HAM system is an intelligent interactive system which compensates the skill deficiencies of a human operator by adapting itself to the human skill level aiming for increased performance and faster learning of the operator [Har06]. Several skill acquisition, assist control and schemes for the adaptation of the human-machine interface (HMI) have been proposed and they have been proved to accelerate the learning of a human operator and increasing the overall performance of the system [Iwa06], [Kur04], [Suz05a]. A skill evaluation system for HAM, based on a deterministic state transition model was utilized in [Suz04]. The skill evaluation was based on state transition frequency patterns recorded during the operation. A new type of a HAM system for mobile working machines called human adaptive mechatronics and coaching (HAMC) system was proposed in [Ter10b]. It was designed to also account for the challenges regarding to the measurement capability and the work complexity in the real-life machines.

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 [Yan97], [Yu04], [Meg06], [Sol08]. This approach gives the skill level in 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.

The cognitive skills related to fast decision making in a technical environment are very important for mobile work machine operators. [Ova07] For example in the work of forest harvester operator, the number of decisions is much greater than measurable control activities. [Har05] Most of the decisions that the operator makes are pre-trained and automatic. This is the reason why the work of forest harvester operators has sometimes been compared to that of fighter pilots [Har05]. However, most of the research related to decision making skills of human operators is based on simulator training. It is more difficult to measure and assess the decision making abilities and the success of the decisions of operators in normal working conditions and in reality instead of virtual reality, where the scenarios can be pre-selected and easily repeated.

5.1.2 Skill evaluation via task sequence and work cycle recognition

The approach for operator skill evaluation considered in this thesis is task level skill evaluation via task sequence and work cycle recognition. The approach has proven to be effective in evaluation of operators of mobile working machines [Pal06], [Pal09] [Ter10a]. Somewhat similar methods have also been used in different application areas.

Skill evaluation of dynamic tasks in surgery was studied in [Mur03]. 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 the total time used in each motion. References [Ros01] and [Ros03] 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. A surgeon's skill was evaluated by comparing statistical distances of the observation sequences of the test subject to expert and novice HMM models. These models were obtained by training HMMs using data from two groups of surgeons, expert and novice groups. Also task execution times, task execution frequencies, and force/torque measurements during tasks were used in skill evaluation.

In [Ita97] 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 likelihood-based 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.

5.1.3 Skill evaluation method in work performed with a mobile working machine

In this section, four rather general frameworks for the skill evaluation of human operators using a HMM with explicitly defined states are proposed. As described in Chapter 4, the role of the HMMs is to recognize the work tasks of the operator from the resultant, observable actions.

The human performance or skill can be defined as person's most likely outcome when performing a given task [Yan97]. Thus, the skill is interpreted as the average performance from several repetitions of a task. In addition, the skill level can be estimated based on how smoothly a sequential process is executed [Suz04]. 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" [Suz05b]. In order to evaluate operators' skills to perform complex work with advanced machines, a broader definition for a skill was defined in [Ter10a].

> **Definition 5.1:** 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*.

Moreover, using HMMs in the task recognition, task or task sequence dependent skill evaluation of a machine operator, using definition 5.1, can be performed without additional sensors. Instead, only the visible control actions of the operator are used.

The structure of the skill evaluation system is shown in Figure 5.1.

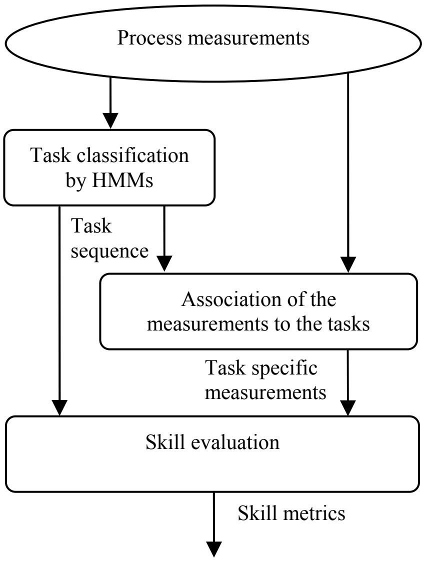

Figure 5.1: Structure of task dependent skill evaluation system. [Ter10a]

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 of the measurements often depend on the task. For example, a good value for the number of simultaneous motions in operation of a boom crane with several degrees of freedom is different in moving task than in positioning task. Within the proposed skill evaluation system structure, the skill metrics can be defined based on task dependent measurements and the task dependent metrics can be defined using for example control theoretic approaches. Furthermore, several skill metrics can be calculated for each of the skill components (*i*)-(*iv*) of definition 5.1.

5.1.4 Development of skill metrics in task level skill evaluation

Extracting skill measures and defining skill metrics from operational data is always a very application dependent procedure. The advantage of the proposed method is that the same skill evaluation system can be used to evaluate operators with any skill level or task difficulty, because the skill level is evaluated based on the skill metrics, which are defined separately for each task. Consequently, since the skill metrics are defined at the task level, and since in each task the different components of skill (definition 5.1) are evaluated separately, the reasons for low performance or poor skill can be analyzed more easily.

Skill metric development can be divided into four frameworks [Ter10a]

a) *Task efficiency* as skill measure.

b) *Complexity of the task sequence* as skill measure.

c) *Ability to plan and make decisions* as skill measure.

d) *Task difficulty* as skill measure.

The premise of using *task efficiency* as skill measure 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. The consumed resource can be any measureable quantity, for example time or fuel. 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. Using Kronecker delta $\delta_{i,j}$, it is obtained by summation over all the observations.

$$E_i = \sum_{t=1}^{T} \delta_{i,q_t} R_t \tag{5.7}$$

Furthermore, if the interest is to study the mean resource consumption of a single task, one can compute

$$\overline{E}_i = \frac{\sum_{t=1}^{T} \delta_{i,q_t} R_t}{\sum_{t=1}^{T} \delta_{i,q_t}} . \tag{5.8}$$

However, it is important to note that the efficiency of a task S_i 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.

When the *complexity of the task sequence* is used as skill measure, it is assumed that a skilled operator achieves the goal smoothly with fairly simple sequence of tasks. Hence, an expert completes the job in fewer operational phases than a beginner and without failed attempts. A beginner might have to perform some phases more than once after an unsuccessful attempt. Task sequence complexity is defined by the number of state transitions. 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

$$f_i = \sum_{t=2}^{T} \delta_{i,q_t} \left(1 - \delta_{i,q_{t-1}}\right), \tag{5.9}$$

where the quantity inside the summation is one only if the current state is S_i and the previous state is not S_i. Furthermore, a measure for the complexity of the whole task sequence can be obtained adding up all state changes. That is,

$$f_{tot} = \sum_{t=1}^{N} f_i \,, \tag{5.10}$$

where N denotes the number of possible states. The complexity of the task sequence measures the skill components (*iv*) and (*iii*) of the Definition 5.1, depending on the tasks of interest and the definition of f_i. In the *complexity of task sequence* as a skill measure, it is possible to for example treat state transitions to unwanted states differently or allow some particular state transitions to appear more often than others.

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 [Kle98]. 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 make a successful decision in significantly shorter time and needs to perform those tasks only once. Thus, the *ability to plan and make decisions* as skill measure can be defined by using (5.9) by calculating the number of state transitions to states related to decision making or calculating the time spent in those states. Several direction changes or stoppages in the trajectory of the movement during a particular task may also suggest that the operator has poor decision making skills. In this case the skill measure is obtained using (5.7), where R_t is the particular value of interest measured during the task. An *ability to plan and make decisions* is essentially the measure of skill component (*iv*) of the Definition 5.1.

In addition to mere efficiency of a task or resource consumption, the way each task is executed provides valuable information about the operator skills. The skill measure of *task difficulty* can be obtained using (5.8), where R_t is defined as the value of interest, the mean control rate or the mean number of parallel controlled functions during the task. The task difficulty as a skill measure describes essentially the machine controlling skills, that is, skill component *i*) of Definition 5.1. 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 *ii*) as well.

When *task difficulty* is used as skill measure, it is first assumed that the task is complex and it is not possible to define a task difficulty index using (5.2) directly from the available measurements. However, it can be assumed that professional operators want to fully utilize their execution capacity, since their financial income may often be related to the work outcome. Thus, an expert operator accomplishes a task or a sequence of tasks in the most efficient way, although it may be quite difficult, while 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. For example in (5.3) the task completion times (T_M) may be the same for different levels of difficulty index, if the task execution capacities (coefficients a_{ID} and b_{ID}) are different. Therefore, if task difficulty can be estimated indirectly, it can be used as skill measure.

With mobile work machines, the task level skill evaluation using HMMs provides a way to study the characteristics of the *task difficulty* by observing the control commands given by the operator during a given work task. According to reference [Gel02], handling of control levers of a forest machine while moving the boom follows an autonomous over-trained motor-sensory scheme. In a comparison with a beginner, an expert operator does not need such a cognitive effort in performing the task. To complete the task as efficiently as possible, multiple machine functions should be 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 frameworks of skill metrics can also be considered from the point of Rasmussen's Skill-Rule-Knowledge model for human performance [Ras83], [Ter10b]. In the Rasmussen's model, the human behavior is divided into three levels: knowledge-based, rule-based, and skill-based behaviors. The skill-based level represents the motor-sensory performance during actions, while the rule-based behavior is generally based on explicit know-how. The knowledge-based behavior level represents behavior in situations where the human does not directly have any rule or patter of how to act. Typically, while learning, human behavior shifts from knowledge to skill based behavior, which reflects to the human responses and should be observable from the skill metrics, as the skills of an individual learner develop.

5.2 Statistics of human performance

By examining the values of some given human performance measures alone, without being a specialized expert on the subject, it is quite difficult to say whether the performance is good or not. To be able to objectively evaluate the values of performance measures, a reasonable statistical reference is required before the values can be compared. In practice, there are two important questions concerning the measured performance: How good the performance is and how much it theoretically could be improved? Using the context of statistical reference values of a group of individuals, both of these questions can be answered.

The problem of evaluating the performance of individuals within a selected group is very common in education. All around the world, students are given grades based on exam points, and it is up to the teacher to create the mapping from the exam points into grades. A very common practice is to force the grades to fit the normal distribution. This tradition of "grading on the Gauss curve" however is originated from early social scientist studies, in which the mathematical assumptions behind the normal distribution often were not fully understood in the empirical test arrangements. The mean in any distribution of concerning human phenomena was commonly interpreted as "ideal", not as a purely descriptive tool as it should have. Extremes in both directions were interpreted as (undesirable) deviations. This leaded to the unjustified assumption that the human performance is normally distributed. [Goe81]

In reality, all measures which contain a good deal of random variation in the measuring arrangement will successfully fit the normal distribution, which is by definition, a measure designed to measure random variation. If a student taking a test answers a large number of items and receives a total score corresponding to the number of items that the student answers correctly, the measurement arrangement has an inherent bias toward the normal distribution. In that, it is essentially an averaging process, and the central limit theorem shows that distributions of means (average number of correct answers) tend to be normally distributed even if the underlying distribution is not. [Goe81]

On the contrary, there are many examples of studies related to human performance that do not make the assumption of normal distribution. Extreme values (EV) reflect rare events that have an unlikely occurrence. Modeling of extreme cases, that are sometimes treated as "outliers" in classical statistic methods, are approached in extreme value theory using block maxima (BM), where the maxima (or minima) of a variable are considered. Examples of the extreme value applications can be found in finance, risk analysis, nature phenomena and in sports [Kot00], [Col01], [Ein09], [Rob95]. Block maxima (or minima) follows generalized extreme value distribution (GEV) [Kot00], [Col01]. Generalized extreme value distribution has the probability density function

$$f(x) = \frac{1}{\sigma}(1+z)^{-(1/\xi+1)} \exp\left(-(1+z)^{-1/\xi}\right), \tag{5.11}$$

where

$$z = \xi\left(\frac{x-\mu}{\sigma}\right), \tag{5.12}$$

and ξ, μ and σ are the parameters that determine the shape of the distribution. Parameter μ determines the location and σ the scale of the distribution. The shape parameter ξ governs the tail behavior of the distribution. In the case when $\lim_{\xi \to 0}$, the distribution is called *type I* extreme value distribution (or Gumbel distribution). For values of $\xi > 0$, the distribution is called *type II* extreme value distribution (Fréchet) and for values of $\xi < 0$, the distribution is called *type III* extreme value distribution (Weibull).

The *type I* EV distribution, named after E. J. Gumbel, has been widely used in modeling of nature phenomena related to extreme values, such as floods, rainfalls, earthquake magnitudes, annual sea-level prediction and so on. [Kot00], [Col01] Gumbel distribution is derived to model the distribution of the block maximum (or the minimum) of a certain number of samples that can come from various distributions. For example, in [Gum41], Gumbel distribution is used as a model of the distribution of the annual maximum flow of a river (peak flood discharge), given that the flood peaks from previous years are known.

When blocks of that consist of a certain number of samples from an underlying distribution are considered it follows that the distribution of maxima or minima is skewed. The Gumbel distribution for block maximum is skewed to the right. Block minimum has the same shape, except that it is mirrored. The shape (skewness) of Gumbel distribution is fixed and it has only the location μ and scale σ parameters. This restricts its flexibility. But on the other hand, having only two parameters makes it a safer choice, especially if the number of samples, which are used to determine the parameters of the distribution, is small. The probability density function of Gumbel distribution for block maximum (skewed to the right) is obtained from (5.11) as $\lim_{\xi \to 0}$

$$f(x) = \frac{1}{\sigma} \exp\left(-e^{-\frac{x-\mu}{\sigma}}\right) e^{-\frac{x-\mu}{\sigma}} . \tag{5.13}$$

Respectively, probability density for block minimum (skewed to the left) is

$$f(x) = \frac{1}{\sigma} \exp\left(-e^{\frac{x-\mu}{\sigma}}\right) e^{\frac{x-\mu}{\sigma}} . \tag{5.14}$$

The cumulative distribution function of Gumbel distribution has the form for block maximum (skewed to the right)

$$F(x) = \exp\left(-e^{-\frac{x-\mu}{\sigma}}\right), \tag{5.15}$$

and the cumulative distribution for block minimum is

$$F(x) = 1 - \exp\left(-e^{\frac{x-\mu}{\sigma}}\right). \tag{5.16}$$

The estimate for scale σ parameter of Gumbel distribution can be obtained using the standard deviation s of the samples

$$\sigma = \frac{s\sqrt{6}}{\pi} \ . \tag{5.17}$$

Then, the estimate for location μ parameter for block maximum (skewed to the right) can be calculated using the obtained scale parameter and the sample mean

$$\mu = \bar{x} - \gamma\sigma \ . \tag{5.18}$$

The location parameter for the case of block minimum is respectively

$$\mu = \bar{x} + \gamma\sigma \ , \tag{5.19}$$

where γ is the Euler's constant ($\gamma \approx 0.577$).

The use of Gumbel distribution as a model of nature phenomenon is illustrated in Figure 5.2, where the annual maximum flood discharge of river Rhone in France was modeled using EV distribution. The statistics of the flood discharges of the Figure 5.2 are obtained from the original publication of E. J. Gumbel [Gum41]. The distribution was estimated using recorded flood flows from over a century of time. The figure shows that the flood peak is usually above a certain value due to constraints to the river flow set by nature, climate and geography. For example the flood discharge of Rhone has been below 1000 m^3/s only once between the years 1826 and 1936, while the most probable flood (mode of the distribution) was 2177 m^3/s. The maximum value of the flood flow is not that restricted. The other tail of the distribution is much "fatter", and there are records of flood discharges of over 4000 m^3/s.

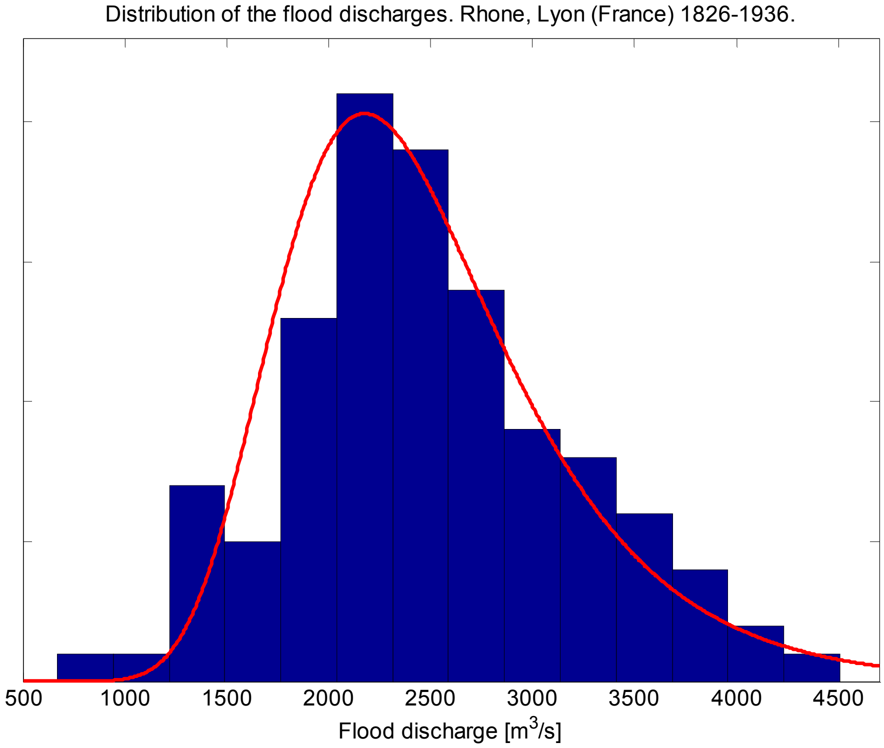

Figure 5.2: Distribution of flood discharges, modeled with type I extreme value distribution. Rhone, Lyon (France) 1826 – 1936. [Gum41]

Skill, expertise or performance levels of a group of individuals, that are selected and specialized for some specific job or task, can be interpreted as a sort of block maxima (or minima), and the performance of the group can be model using extreme value theory [Pal09]. There are also many examples of extreme value theory concerning human performance, especially in sports. Many of them are related to world records. In [Rob95] the statistics the current population of competing athletes was analyzed to find out if extreme value analysis supports the doubts that an inconsistently large world record break had been influenced by performance enhancing drugs. In another example, extreme-value theory was used to estimate the "ultimate" world records (how good the world record can be) for the 100m running, for both men and women [Ein09]. For this aim, the fastest personal best times of elite athletes from a certain, recent time period were collected. Although world records are always broken by individuals, whose performances are independent of statistics of the population of other athletes, a fairly reasonable prediction on the world record was found using the extreme value distributions.

It is evident that world records in sports can be modeled using extreme value theory, but regarding the evaluation of human operator performance, it is essentially about comparing the performances of an individual against the performances of a corresponding reference group. Due to the fact that results in sports are public, it is easy to collect data of the performances of different groups and individuals. Figure 5.3 shows the results of four different competitions in different sports, presented as histograms. The competition levels vary from national competition to world championship level, the distances from 200m to 90km and the number of competitors from 21 to 27886. The figure also shows the probability densities of Gumbel and normal distributions, estimated from the samples in each case. The figure shows that Gumbel distribution is clearly more reasonable approximation for the histograms than the normal distribution. Nevertheless, the number of samples is quite small in the two cases on the left.

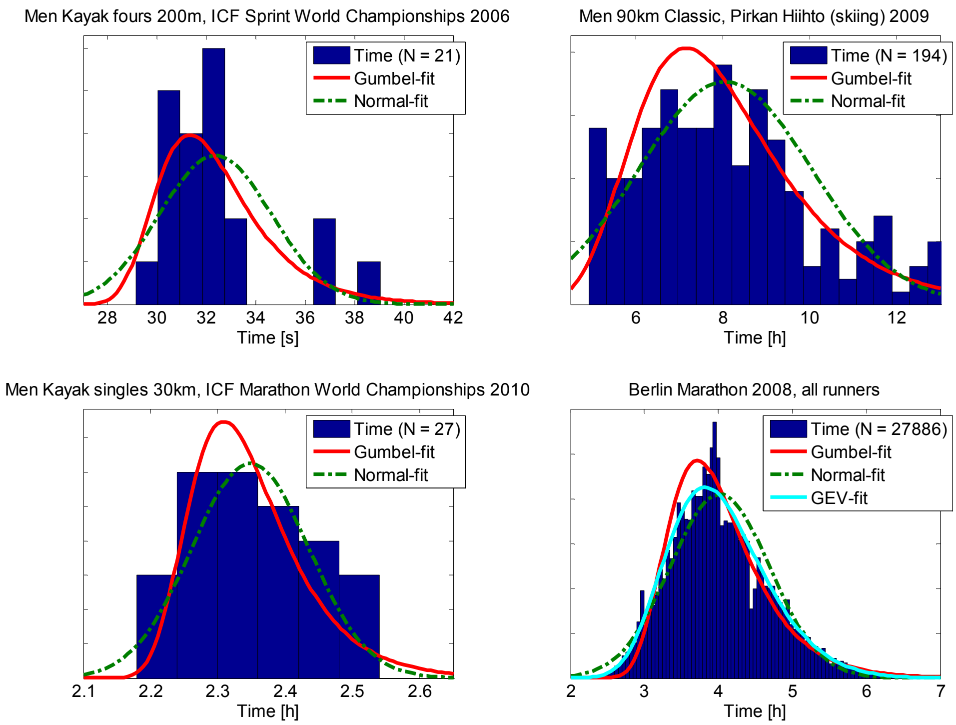

Figure 5.3: The results of four competitions in different sports presented as histograms.

As the number of samples increases, the histograms tend to be smoother, and the estimated probability density approximates fit better to the histograms. A large number of samples allows the use of three parameter GEV distributions, which may also make the distribution approximate more accurate against the samples. This is shown in Figure 5.3, bottom right, which shows the results of Berlin marathon 2008 that had 27886 runners, who reached to the finish line. In this case a three parameter GEV distribution is a fairly good estimate of the distribution of the results.

Unfortunately, the increase of the number of samples (individuals) in human performance related phenomena leads to worse comparability between the individuals in the group and probably introduces more random error. In the sense of statistics, it can be said that the samples are not drawn from the same population. Consequently, although the winning time (2.03:59 by Haile Gebreselassie) of the Berlin marathon 2008 in fact broke the world record, it would not have been possible to predict that time accurately using the samples all runners, because the variance of such a large population is also very large. On the contrary, the variance within a more specialized group is significantly smaller, as in the results of World Championship events (on the left), and it would be possible to construct a more accurate statistical estimate of the winning time or world record time, as it is done in [Ein09].

Hence, in statistical human performance evaluation, there is always a tradeoff between the number of samples and the comparability among the collected samples. A large number of samples provides a more reliable distribution approximation, but it may also introduce random error and cause inherent bias towards normal distribution. GEV distribution has been found to be suitable for modeling human performance levels within a group of individuals [Pal09]. Particularly, if the number of samples is small, *type I* GEV (Gumbel) distribution with only two parameters is suitable. However, the choice of the distribution eventually depends on the data, and other families of distributions, for example gamma distribution [Koi97], have been used as well.

An example of a time-based skill metric, measured from several (N = 104) operators from machine work, is shown in Figure 5.4. The Gumbel distribution approximates the histogram reasonably well. Usually there is some variation in measurements of human performance between consecutive repetitions of a task. Consequently, the skill (level) is calculated using several repetitions of the task. In this particular case the values are daily averages of each operator. However, in practice, much shorter period would often be enough to obtain a reasonably reliable estimate of the skill level.

To be able to answer the two questions concerning the measured performance, how good the performance is and how much it theoretically could be improved, a statistical *expert performance* level is defined.

> **Definition 5.2:** Statistical *expert level performance* is a good value of human operator's skill or performance metric. The value of expert level performance is defined as the p-th percentile value of respective performance metric levels, measured from reference group of operators such that the expert level value is better than p % of the values of the reference population.

A probabilistic performance index can be obtained using the cumulative distribution function. It is probability of an average value of a measured performance metric x being better than that of the reference group X (close to the expert level performance).

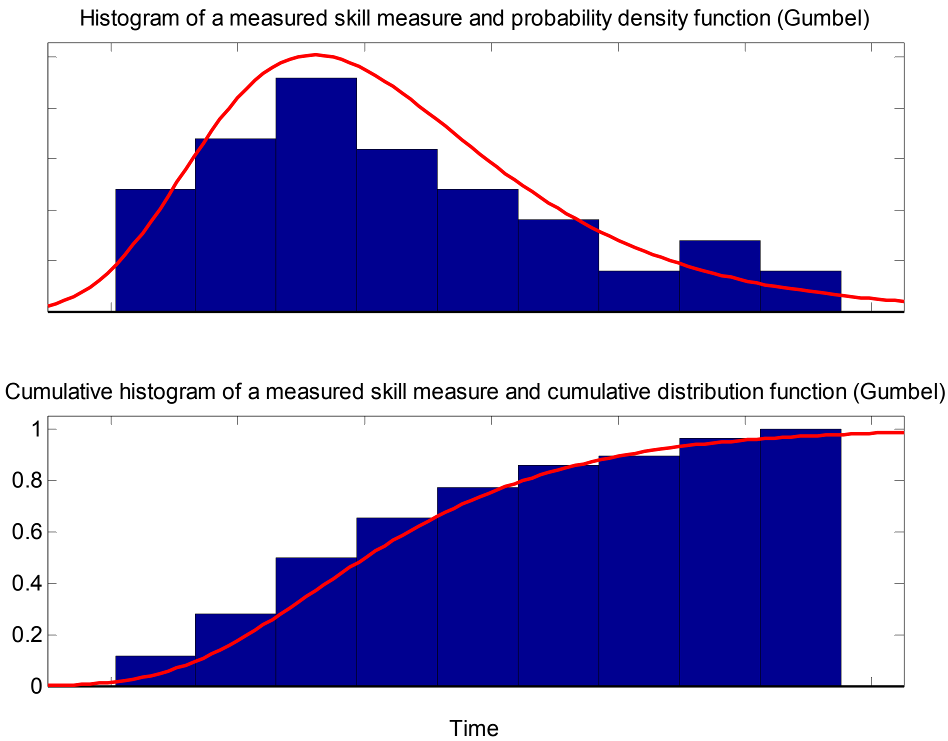

Figure 5.4: Histogram presentation and distribution estimate of a measured time-based skill metric. ($N = 104$)

In case where small values of a performance metric represent good performance (expert level) and Gumbel distribution (block maximum) is used, the probabilistic index is defined as

$$I = P(x \le X) = 1 - F(x) = 1 - \exp\left(-e^{-\frac{x-\mu}{\sigma}}\right), \tag{5.20}$$

where $P(x \le X)$ signifies the probability that the statistical reference values X are larger than the evaluated measure value x (i.e., when small values represent good performance, x is better than X).

Respectively, in case where large values of a performance metric represent good performance (expert level) and Gumbel distribution (block minimum) is used, the probabilistic index is defined as

$$I = P(X \le x) = F(x) = 1 - \exp\left(-e^{\frac{x-\mu}{\sigma}}\right). \tag{5.21}$$

The value of the expert performance level is obtained using inverse cumulative distribution F^{-1} for block maximum

$$x(p) = F^{-1}(1-p) = \mu - \sigma \ln\left(-\ln(1-p)\right), \tag{5.22}$$

and for block minimum

$$x(p) = F^{-1}(p) = \mu + \sigma \ln\left(-\ln(1-p)\right). \tag{5.23}$$

For example, using the data of Berlin marathon 2008 (Figure 5.3 bottom right) as the reference and defining the expert level to p = 99 %, an individual marathon runner whose result is 3:20.00, would receive probabilistic skill index value of 87.5%, and would have theoretical improvement potential of about 25 minutes to the pre-defined expert level performance (2:55.01 for the value of p = 99 %).

In the performance assessment of a man-machine system, there are often multiple performance metrics. Often the values of the metrics depend not only on the operator skills, but also on operating conditions and machine types. The use of reference data, parameterized distributions and pre-defined expert performance levels enables the statistical evaluation of the measured performance metric values. It also provides a systematic and objective tool for the assessment of the improvement potential.

Sometimes the influence of the operating conditions to the measured operator performance metric values is very significant. For example, in favorable conditions the operators generally get better values of the metrics than in difficult conditions. In this case the effect of the conditions can be considered for example by training different distribution parameters to different conditions. However, a disadvantage of this is that it requires much more reference data.

As a conclusion of the statistics of human performance, the use of practical statistical reference data is very important, to be able to evaluate the human performance measures. However, sometimes the measurement arrangements and the variation of human performance between consecutive repetitions of a task may result in a rather large variation to the single measurements. This random variation among the measurements of single repetitions of a task may well be normally distributed. However, the distribution of the single repetitions of a task should not be mixed with the distribution of the human performance levels. If the random variation of the single repetitions is larger or in within the same magnitude as the variation of the performance measure levels, human performance may easily be misinterpreted to be also normally distributed. Therefore, in order to get in touch with the actual underlying phenomenon, it is suggested that the human performance level is calculated using several repetitions of a task, and for example averaging the results. The reference group should also be selected such that the performances of the individuals are comparable. Due to this, the number of comparable samples may sometimes be rather small. It has been found in practice that the human performance levels of specialized individuals performing tasks that require expertise, are often skewed distributions, which can be modeled using GEV distributions [Pal09], [Rob95], [Ein09].

Chapter 6

Operator assistance by decision support and coaching

This chapter discusses systems that are purposed to assist the users in performing work or procedures related to the particular system or application. Common components to the assistance systems are the database (or knowledge base), the model (i.e., the user model within the application context), and the user interface. They can assist the operator for example by:

- Assisting the operator of a machine to perform certain difficult tasks (*active assistance, assist control systems*)
- Giving warnings if necessary (*warning systems*)
- Supporting the operator on decision making (*decision support systems*)
- Providing instructional feedback to help the operator to reach the desired performance (*coaching systems*)

Recently in automotive industry, a lot of examples can be found of systems that are designed to assist car drivers. The systems may for example guide the drivers to lower fuel consumption by gear shift indicators that indicate most fuel efficient gear. Alternatively, the feedback can be in the form of linguistic advice on how to drive more fuel efficiently, from the display of the on-board computer. Instead of passive advice or indicators, the form of feedback can also be active. Assist control systems actively assist the user to perform certain difficult control tasks. There are many applications of active assistance systems in automotive industry, for example parking assistance systems or electronic stability control systems. If the driving situation is regarded dangerous, the system can apply haptic feedback to warn the driver, for example by applying force or vibration through the accelerator pedal, seat or the steering wheel. The kind of systems that warn the user are called warning systems (WS).

Decision support systems (DSS) are computer-based information systems that support decision-making activities [Pow02]. DSS are most often used in business, finance, healthcare, transportation and logistics, usually at the organizational level but can also be applied on the user level. Decision support systems speed up the progress of problems solving by generating evidence to support decision-making process. DSS help the decision maker by compiling useful information from a combination of raw data, expert knowledge, or business models to identify and solve problems related to decision making.

In machine work, where the work procedures and techniques play an important role, it is important to promote the skills and knowledge of the operators through training and learning-by-doing procedures. Traditionally, the training of machine operators is performed with the presence of a human instructor. This kind of personal training can be both time consuming and expensive. This often leads to the situation where the human instructor is present at the basic training, and after that, the learners are on their own through most of the learning process. However, learning can be enhanced in learning environments that exploit the use an automatic instructor. The individual needs of learners and the most important areas of improvement can be focused more efficiently, objectively and without the presence of a human instructor. Learners are able to make sense of their experiences to identify poor decisions and actions, missing knowledge, and weak skills that deserve attention. Specific, computer-based learning systems that use an automatic instructor are called intelligent coaching (ICS) and tutoring systems (ITS). [Pal09]

The form of operator assistance in this thesis is focused to instructional advice (coaching) and decision support. On the contrary, active assistance systems are not discussed, because such systems should inherently be deterministic, which is quite opposite to the stochastic modeling approach of this thesis. The method for intelligent coaching system for mobile working machine operators, presented in this chapter, can be used to help the machine operator for example by providing information of the work performance, supporting decision making, helping to tune control parameters of the machine or the HMI to achieve better performance or as in [Pal09] coaching the user for more efficient work techniques and procedures.

6.1 Intelligent coaching and tutoring systems

Intelligent coaching and tutoring systems are widely researched in well-defined domains such as mathematics and sciences [Bre98]. There are a lot of examples in education, but also in military and flight training. Virtual reality is often used as the learning environment. Simulation based learning enables students to apply their knowledge and practice their skills in a variety of hypothetical situations. It is also possible to adjust the level of difficulty in virtual environments, which makes it easier for beginners to start with. For example, in a flight simulator the dynamics and controlling of the aircraft can be made much easier than in reality. Software simulations can present various scenarios to students and challenge them to apply their knowledge and skills to analyze situations, make appropriate decisions, and see the effects of their actions.

However, many computer-based training systems tend to be electronic textbooks that present facts and concepts using text and multimedia [Ong07]. They test the learner's understanding by asking multiple-choice or fill in the blank -questions. Because these systems focus on factual knowledge rather than on performance, the methods often produce trained novices who are familiar with the subject area, but do not have in-depth expertise needed for high performance. Also the presented scenarios in most virtual learning environments are predefined, and do not challenge the learners to generate own as they would have to in real life. Unlike for example teaching mathematics, the coaching or tutoring problems are ill-defined in many real world training systems [Ale08]. There may be several, possibly infinite number of equally or almost equally good solutions to the problem. There are also no means to directly measure the intent of a human performing a task, but only the resultant actions are observable.

Although there are already research results in the field of automatic assistance control and skill acquisition and the development of the tools are very promising for increasing the performance and stability of human operated machines in certain control tasks, the methods are mainly based on linear control theory. Consequently, the real-time utilization of the methods may not be expedient in machines, which perform generally nonlinear work with multiple ill-defined objectives that consist of not only continuous control tasks, but several separate and simultaneous tasks. Therefore, the focus in this thesis is not on finding a good predictive model of human operator's responses, and using it to provide feedback. Instead, the focus is rather to interpret the operators' actions and to incorporate the knowledge of domain experts and measurement database as the model of work performance. The results are used as instructions to the operator how to achieve better performance.

6.2 Intelligent coaching system for mobile working machine operators

The learning process of work machine operators consist a lot of learning-by-doing actions and involves a lot of tacit knowledge [Pol66], [Vaa05]. Training of mobile working machine operators in virtual learning environments and the transfer of skills to real world has been studied in [Ros00], [Ova05], [Ran03] and [Fre98]. It has been found that virtual learning environments are very effective in training of basic skills, and that the use of virtual environments at the very early stage of learning reduces the repair costs of the real machines. A drawback of virtual learning environments is that the level of reality is not close enough to the real environment. For example, in virtual reality, the vision is often limited, and the students have difficulties to discern the stereoscopic effect of the real environment. Ultimately, the students reach their final skill levels in the real environment.

The principle of the proposed method of intelligent coaching of mobile working machine operators is shown in Figure 6.1. The method consists of four stages. At first, the intents behind the actions of the learner, a human operator, are detected based on the available measurements (*plan recognition*). This is done by recognizing the work cycles and work tasks using Hidden Markov models (HMMs).

The performance and skill of the learner is evaluated based on task level performance and skill measures (*performance interpreter*). The performance of a human operator in a man-machine system is regarded to consist of both cognitive and motor-sensory skills. The performance of a human operator 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. The performance of the learner in different performance measures is interpreted as described in Chapter 5, the average performance from several repetitions.

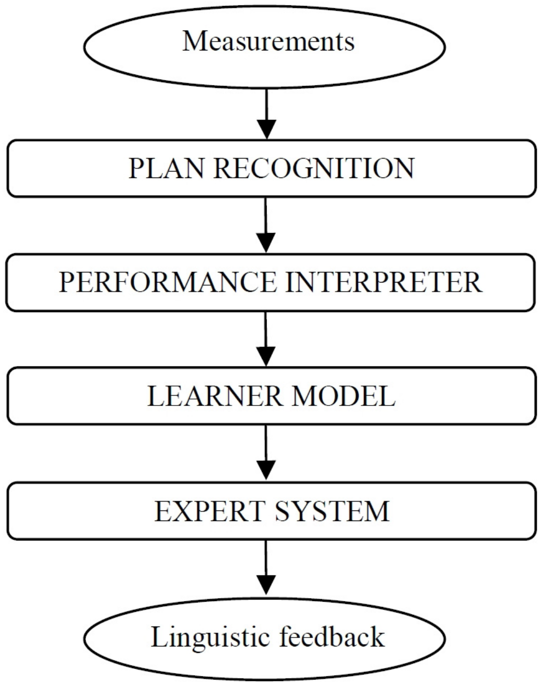

Figure 6.1: Structure of ICS for mobile working machine operators. [Pal09]

In normal working situation, the technical performance of a machine can be thought to remain at a relatively constant level. Then the overall performance of the man-machine system, relative to the expected performance for the machine type/model can be thought to depend mainly on the skills of the operator. The overall performance is also affected by the variation in the working conditions, which need to be considered [Hol05]. The randomness associated with performance levels of human operators is dealt using a statistical *learner model*, which compares the average measured performance levels of an individual learner (operator) to the average values of a statistical reference group. In the ICS, we may be interested in both the skill and work strategy or technique. However, often the work strategy cannot be directly observed from the measurements but can be obtained by an inference. Finally, a linguistic feedback system gives advice and suggestions to the learner. The advice and suggestions are formed using an *expert system*. The expert knowledge is formulated as rules of a fuzzy inference system (FIS).

6.2.1 Statistical learner model

A man-machine system is a time-varying stochastic process which has a lot of randomness in the performance. Most of the variation is due to the imperfect and noisy nature of the human motor-sensory processes and the skill associated with the task. The performance may also vary for example, due to the human emotions, fatigue and motivation. However, the human emotional factors cannot be measured in real-life industrial systems. Moreover, the performance of the machine varies also depending on the operating conditions and machine's technical condition.

To be able to evaluate the performance and assist an operator, the ICS needs to have a model of the performance. In [Pal09], a learner model that exploits the context of statistical reference was proposed. The statistical learner model is based on the following assumptions [Pal09]:

I. The most likely performance level of a human is interpreted as the average performance from several repetitions of a task.

II. In the long run, the variation caused by the human factors is assumed to be distributed so that the long term average describes well the most likely performance level of the operator.

III. The machine condition is assumed perfect and performance constant during the observation window.

IV. The greatest variations in the operating conditions can be compensated from the measurements.

V. The variation of the performance metrics (after compensation of operating conditions) measured from one operator is smaller than the differences between performance levels of operators with significantly different levels of skill.

VI. The learning rate is so slow that the performance can be approximated constant during the observation window.

VII. There exists a level of high performance that is very rarely exceeded.

An average performance level of the performance measures is calculated from an observation window. The length of the window has to be long enough in order to get enough samples so that the averages of the performance measures correspond to the most likely performance levels. The most likely performance and skill levels of the group of operators are modeled using a statistical distribution. With the assumptions above, the distribution of the individual performance levels can be described using a GEV distribution.

An advantage of using a statistical learner model is that same expert system can be used in different domains. If the expert knowledge is valid and usable in different domains, only the parameters of the statistical learner model have to be trained separately for each domain. Different domains could be for example different machine models that have significantly different performance. This is a significant advantage, because gathering the expert knowledge is often much more complicated task than updating the parameters of a statistical model.

6.3 Fuzzy expert system

Expert systems are computer programs that emulate the reasoning process of a human expert [Kan91]. Typically there is some amount of imprecision and uncertainty involved with the reasoning. The imprecision is due to the granulation of the linguistic formed expert knowledge. Other reasons for uncertainty are for example ill-defined structure of the problem formulation or imprecise measurements.

Fuzzy reasoning techniques [Zad65] can provide the basis for representing the imprecision in a way inherent to the expert's knowledge. Fuzzy techniques have been successfully used in several control applications. The theoretical basis behind fuzzy technique allows dealing with uncertainty in a manner that is well supported and the use of linguistic variables helps capturing the expert's knowledge and meaning in natural fashion. Fuzzy expert system can also be used in conjunction with both probabilistic and non-probabilistic variables. Fuzzy logic based controllers are designed to capture the key factors for controlling a process without requiring many detailed mathematical formulas. Due to this fact, they have many advantages in real time applications.

It is natural for humans to perform a wide variety of physical and mental tasks without any accurate measurements and extensive computations. From this point of view, in real-world problems there is much to be gained by exploiting the tolerance for imprecision, uncertainty and partial truth. [Zad02] Moreover, decision-making strategies can be generated using fuzzy logic based on linguistic instructions, priori knowledge and experience of domain experts. Indeed, fuzzy logic has proven to be successful in problems where exact mathematical formulation of the problem is hard or impossible, but a human operator has developed a solution through experience.

In fuzzy reasoning, the input variables are defined as fuzzy sets, and the values mapped into grades of membership, the membership functions. The process of converting a crisp input value to a fuzzy value is called fuzzification [Zad65]. There are multiple ways to implement the fuzzification, i.e. to choose how to partition the input value space and choose the shapes of the membership functions. All this is up to the expert's choice. In complex fuzzy systems, this may lead to tedious amount of work for the expert and it may become difficult to generate and to maintain the rule base of the expert system.

The principle of the proposed fuzzy expert system is that it should be able to automatically assess the operator's actions during normal operation. It is important to note that there is always some amount of uncertainty involved with the measuring and modeling a man-machine system that operates in varying conditions, which makes it impossible to find exactly correct solutions to the inference problems. Instead, the focus of the expert system is to make observations and to give useful suggestions that may potentially help the operator to achieve better overall performance [Pal09]. To be able to exploit and combine the knowledge of all domain experts in machine work (operators, instructors, researchers, etc.), the task of forming the expert system should also be made as simple as possible. To achieve this, the knowledge base is constructed in linguistic form, including the inputs and outputs, the numerical values of the parameters rely on data. The reasoning problem, as such, should also be constructed from irreducible basic elements as simple and as few as possible.

6.3.1 Fuzzification

In the approach of this thesis, the values of the input variables are based on statistical reference and cumulative distributions. Then, each input has value between 0 and 1, and the interpretation of the value is based on probability. However, due to the imprecision and uncertainty of the values of the performance measures, the crisp values are converted to fuzzy values. For each input and variable selected, two or more membership functions are defined. The qualitative categories for each membership functions are for example: *low*, *average* or *high*. The chosen type of membership function is Gaussian [Pal09]. The Gaussian membership function has the shape of normal distribution and two parameters, center c and scale σ. Gaussian membership function is formally defined as

$$\mu(x) = \exp\left(-\frac{(x-c)^2}{2\sigma^2}\right). \tag{6.1}$$

Using the property of cumulative distribution function, that values are always between 0 and 1, it is possible to define default values for membership function parameters. Consequently, the expert does not need to focus on the numerical values of variables, it is only required that the expert has the qualitative knowledge to be able to understand significance of the relative values, which have been converted to fuzzy variables. This may ease the workload of the expert, especially if the number of inputs variables is large or the values depend on operating conditions or machine types.

Choosing that the input space is equally partitioned to two membership functions (for example *low* and *high*) such that centers are c = {0, 1} and the membership functions intersect at point 0.5 with membership grades 0.5, the value of scale parameter σ can be calculated from (6.1) as

$$\sigma = \frac{1}{\sqrt{-\ln\left(\frac{1}{2}\right)2^3}} \tag{6.2}$$

Respectively, when input space is equally partitioned to three membership functions such that centers are c = {0, 0.5, 1} and the membership functions intersect at points 0.25 and 0.75 with membership grades 0.5, the scale parameter σ value is

$$\sigma = \frac{1}{\sqrt{-\ln\left(\frac{1}{2}\right)2^5}} \tag{6.3}$$

Gaussian membership functions have the good properties that they are continuous and value of the membership function is always nonzero. Due to this, the rule base does not need to be complete, and nonzero membership values always yield an inference result. However, the choice of the membership function is finally up to the expert. If appropriate, the expert may as well choose other parameterizations or shapes, such as for example sigmoidal or bell shaped membership functions.

6.3.2 Inference based on fuzzy similarity

Fuzzy similarity measures have been successfully used in many inference and pattern recognition systems that imitate the reasoning process of a human expert. It is natural for a human expert to make decisions and draw conclusions from partially true premises, and to choose the one which is the most similar to ideal case, according to the expert's own knowledge. In this thesis, Łukasiewicz valued [Luk70], [Pav79] fuzzy similarity measures are used to generate feedback and suggestions of the ICS by interpreting the relative values of machine operators' skill and performance measures according to the expert's knowledge.

Same way as notion of fuzzy subset generalizes that of the classical subset, the concept of similarity can be considered as a many-valued generalization of the classical notion of equivalence [Zad71]. In Łukasiewicz structure, mean or weighted mean of many fuzzy similarities is a fuzzy similarity, too [Tur99]. Moreover, instead of arithmetic mean, other means such as geometric and harmonic mean can be used as well [Luu03]. In some cases this is useful, since the use of Łukasiewicz structure and weighted mean as fuzzy similarity measure may express better the experts' knowledge in decision making and be more intuitive, compared to other frequently used algebraic structures in fuzzy logic. Particularly, if the decision is based on multiple, possibly vague criteria, it may be better to base the decision on the value of most of the criteria rather than requiring all the criteria to be satisfied [Yag88].

A Łukasiewicz many-valued logic similarity based fuzzy algorithm is used to construct fuzzy IF–THEN inference system. The algorithm utilizes Łukasiewicz many-valued equivalence. Many-valued similarity is regarded in the sense of fuzzy similarity relation in a crisp set. Starting from the Łukasiewicz well-defined many-valued logic, it is possible to construct a method performing fuzzy reasoning such that the inference relies only on expert's knowledge and on well-defined logical concepts. The output of the inference is a linguistic expression, and therefore technical defuzzification methods (like Center of Gravity) are not needed to determine the final output. [Nii03], [Tur99]

A basic observation is that any fuzzy set generates a fuzzy similarity, and that these similarities can be combined to a fuzzy relation which is also a fuzzy similarity. This induced fuzzy relation is called total fuzzy similarity. Fuzzy IF–THEN inference systems are actually problems of choice: each IF-part of the rule base is compared with an actual input value, and the most similar case fires the corresponding THEN-part. Thus, the main distinction between the approach based on similarity and other fuzzy inference systems is that only one IF–THEN rule determines the output. [Nii03], [Tur99]

The system designer's knowledge of the process is stored as rules in the knowledge base (rule base). Thus, the rules have a basic influence on behavior of the system and should therefore be acquired thoroughly. The development of rules is time consuming, and designers often have to translate their process knowledge into appropriate rules. In [Sug85] and [Zim96], there are mentioned ways to derive fuzzy control rules:

- Operators or control engineer's experience and knowledge
- Fuzzy modeling of the operator's control actions
- Fuzzy or crisp modeling of the process
- Heuristic design rules
- On-line adaptation of the rules

It is also possible to include extra demands in the algorithm. For example, in some cases the degree of total fuzzy similarity of the best alternative should be greater than some fixed value between 0…1, particularly, if the rule base is incomplete. Sometimes all the alternatives possessing a high fuzzy similarity should be indicated, or the difference between the best candidate and second one should be larger than a fixed value between 0…1. All this depends on an expert's choice.

Continuous *t-norms* are important in fuzzy logic. In Łukasiewicz algebra, in the real unit interval, *t-norm* is defined as

$$x \otimes y = \max\{0, x + y - 1\}, \tag{6.4}$$

and *equivalence* relation as

$$x \leftrightarrow y = 1 - |x - y| . \tag{6.5}$$

Moreover, Łukasiewicz valued similarity on set X is defined as

$$\forall x: S\langle x,x\rangle = 1 \tag{6.6}$$

$$\forall x,y: S\langle x,y\rangle = S\langle y,x\rangle \tag{6.7}$$

$$\forall x,y,z: S\langle x,y\rangle \otimes S\langle y,z\rangle \le S\langle x,z\rangle \tag{6.8}$$

Any fuzzy set (A, μ_X) on a reference set A generates a fuzzy similarity S on A, defined by

$$S(x,y) = \mu_X(x) \leftrightarrow \mu_X(y)\,, \tag{6.9}$$

where x, y are elements of A. Moreover, if $\mu_X(y) = 1$, then

$$S(x,y) = \mu_X(x)\,. \tag{6.10}$$

Consider n Łukasiewicz valued similarities S_i, $i = 1, \ldots, n$ on a set X. It has been shown in [Tur99] that

$$S\langle x,y\rangle = \frac{1}{n}\sum_{i=1}^{n} S_i\langle x,y\rangle \tag{6.11}$$

is a Łukasiewicz valued total similarity on X. Generally, the weighted mean

$$S\langle x,y\rangle = \frac{1}{M}\sum_{i=1}^{n} m_i \cdot S_i\langle x,y\rangle, \quad M = \sum_{i=1}^{n} m_i\,, \quad m_i \in \{0,1,2,\ldots\} \tag{6.12}$$

is also a Łukasiewicz valued total similarity on X [Tur99].

The rule base, constructed by the expert knowledge, is a collection of IF-THEN rules. For example, let the given inputs $x_1, \ldots, x_n$ be objects that are described using [0, 1] - valued fuzzy sets $A_1, \ldots, C_k$. Then, the rule base is:

Rule 1. IF x_1 is A_1 AND x_2 is B_1 AND x_3 is C_1 THEN z is D_1

Rule 2. IF x_1 is A_2 AND x_2 is B_2 AND x_3 is C_2 THEN z is D_2

…

Rule k. IF x_1 is A_k AND x_2 is B_k AND x_3 is C_k THEN z is D_k

The rules can be seen as an idealization, which has membership value 1 for all $i = 1, \ldots, n$. By (6.10) and (6.11) the total fuzzy similarity is then

$$S\langle x,1\rangle = \frac{1}{n}\sum_{i=1}^{n}\mu_X(x). \tag{6.13}$$

Each IF-part of the rule base is considered as a crisp case and the actual input values are compared separately with each IF-part. In other words, the combination of multiple similarities is calculated between actual inputs and each IF-part of the rule base. By (6.12) and (6.13) this is equivalent to counting weighted means, for example:

$$\big(m_{1,1}\,\mu_{A1}(x_1) + m_{1,2}\,\mu_{B1}(x_2) + m_{1,3}\,\mu_{C1}(x_3)\big)\,/\,M_1 = \text{Similarity(actual,Rule 1)}$$
$$\big(m_{2,1}\,\mu_{A2}(x_1) + m_{2,2}\,\mu_{B2}(x_2) + m_{2,3}\,\mu_{C2}(x_3)\big)\,/\,M_2 = \text{Similarity(actual,Rule 2)}$$
$$\ldots$$
$$\big(m_{k,1}\,\mu_{Ak}(x_1) + m_{k,2}\,\mu_{Bk}(x_2) + m_{k,3}\,\mu_{Ck}(x_3)\big)\,/\,M_k = \text{Similarity(actual,Rule } k\text{)}$$

where $m_{k,n}$ are the weights of the input fuzzy sets, which can be used to emphasize the mutual importance of the corresponding input variables in each rule. M_k is sum of the weights for each rule k.

The output D, which is a linguistic expression, is chosen according to the rule (ideal object), which has the highest total similarity with the actual input combination. Therefore, defuzzification procedures are not required. Note that it is allowed that some D_i and D_j with $i \neq j$, are the same. If the highest total fuzzy similarity is not unique, the expert can define criteria how to distinguish the mutual importance of the rules. If appropriate, the expert may choose multiple rules as well. If the expert is unable to define rules to each combination, the rule base can be incomplete. This is a useful feature, if there are a lot of combinations of inputs.

Fuzzy inference systems can certainly be constructed in many ways depending on the application. However, when the output of the inference system is a decision or a conclusion based on the knowledge (rule) base, the benefit of using Łukasiewicz fuzzy similarity measures as the basis of the inference is that the similarity between multiple objects can be constructed as the weighted mean of many fuzzy similarities. Furthermore, defuzzification procedures are not required. These properties may considerably reduce the effort of constructing a fuzzy expert system.

Chapter 7

Case: Analysis of work and performance in cut-to-length forest harvesting

In this chapter, work and performance of a group of operators in CTL forest harvesting is analyzed by applying the methods proposed in the thesis. The aim of this chapter is to summarize in a concise manner, what kind of useful information can be discovered about the machine work by a sophisticated manipulation and analysis of the existing measurement data. All the steps of the analysis proposed in this thesis are performed to one data set. The results are also briefly compared to the results of earlier publications related to the subject, [Pal04], [Hol05], [Rep06], [Pal06], [Pal09], [Ter09] and [Ter10a]. These publications have slightly different contexts, measurement data and machines, which would make it rather difficult to compare their results against each other and evaluate their significance for the man-machine application.

At the beginning, the reader is first introduced to the mechanized timber harvesting and the machines that are used. The measurement data of the test case example is collected from eight operators from two machines. All the operators worked in similar operating conditions and with similar machines in order to be able to compare their results against each other more easily. In the first step of the analysis, the performance of the two machines is evaluated using the index-based machine performance evaluation framework presented in Chapter 3. The results of the machine performance evaluation provide an overview on the performance and condition of the machines and verify that the machines are in acceptable condition. It is necessary to evaluate the machine performance and condition before it is possible to evaluate the human performance. The results of machine performance evaluation also show, that among the operators, there is noticeable variance in the productivity and fuel economy.

In the next step, the separate work tasks and work cycles are recognized from the measurement data using the HMM based recognition method proposed in Chapter 4. Recognition of the separate work tasks enables a more detailed analysis of the work and it is the basis of the task level operator skill evaluation and work technique analysis. The operator skill evaluation is performed using the methods presented in Chapter 5. A simplified analysis of harvester operator's work techniques and how the techniques affect to productivity is also carried out.

Finally, an intelligent coaching system for harvester operators, with the structure proposed in Chapter 6, is formed. The coaching system is able to make observations on the work and if necessary, it can provide useful hints and suggestions to the operators how they potentially could reach better performance.

7.1 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 (FT) is common whereas in Europe the cut-to-length (CTL) method is dominant. Full-tree harvesting machines are more conventional, having one or two processing functions, whereas CTL harvesting machines are technologically more sophisticated and have more versatile processing operations. Generally CTL is more environmental friendly, because the machines are lighter. CTL is more suitable to small scale harvesting and forest thinning. It is better in varying conditions and it reduces the cost of transportation. [Uus10]

Two separate machines are used in the CTL method, forest harvester and forwarder. A harvester is used for felling, delimbing and bucking (i.e. cutting the trees into logs) of timber. 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. A CTL harvester is shown in Figure 7.1.

The most complex and important part of the harvester is the harvester head. Main functions of the harvester head are sawing, feeding, delimbing of branches, as well as measuring log length and diameter profile. Trees are felled and the stems are cut to logs with a hydraulically actuated chain saw. Once sawing is complete, the stem of the tree is fed to the next cutting point with feeding rollers rotated by hydraulic motors. The grip of the feeding is ensured by pressing the feeding rollers against the stem with a hydraulic cylinder. The stem is grasped into the harvester head with two pairs of delimbing knives, which are closed circumferentially around the tree trunk. Tree branches are cut by means of the feeding force and inertia of the moving stem when the stem is fed between the delimbing knives.

In a modern forest machine, the different parts of the distributed control system (DCS) are interconnected via a Controller Area Network (CAN) bus. The bus enables two-way digital communication between the control modules. The cabin is equipped with the controls for operating the machine functions and with a computer and a display, the operator is provided with a lot of information about the worksite, the currently processed tree and the operating state of the machine. The machines are also equipped with GPS and navigation systems.

Figure 7.1: Cut-to-length harvester.

Despite of the high degree of automation in a harvester, the task field of the operator is broad and quite demanding. The work of the operator requires both motor-sensory and cognitive skills and knowledge and involves a lot of planning and decision making tasks. The human operator has the main responsibility of the efficiency and environmental impact of timber harvesting. The operator drives the base machine in sometimes very difficult terrain. Usually the harvester operator also plans the strip route for the logs to be carried to the roadside.

The harvester head is attached to the tip if the boom crane and the operator controls the hydraulic actuators of the crane using joysticks. Boom operation plays a very significant role in terms of the overall performance of the machine and the most important requirements are speed and maneuverability. The operator has to learn the different dynamics of the flexible manipulator and for example the dynamics of falling trees.

Processing of the tree stems involves many tasks that require a lot of quick decision making skills. The control computer of the harvester has software that calculates the optimal lengths of the harvested logs in order to maximize the timber value. However, it is the responsibility of the operator to decide the final log lengths and timber assortments, depending on the quality the harvested timber. Especially at thinnings, the operator may also have to choose which trees remain and which are cut down. Overall work preplanning and pre-trained work techniques suitable for different situations are the requirements for efficient harvester work [Ova04]. These are skills, which are developed through experience.

7.1.1 Performance monitoring

The modern forestry machines are equipped with performance and condition monitoring software. The software helps contractors, operators and maintenance staff in the improvement and utilization of the machines. With the use of the performance monitoring software it is possible to check that the machines keep operating at maximum productivity and optimum efficiency. Figure 7.2 shows the overview page of performance monitoring software for John Deere harvesters, TimberLink H. The software provides detailed information on individual work cycles and stages, highlighting the changes in technical performance of the machine. [JDF10]

The software is focused on performance monitoring, showing relevant measures that describe overall performance, for example productivity and fuel consumption. These measures are affected by the technical performance of the machine, as well as the skills of the operator and the operating conditions. However, a motivated operator can use these measures to improve work procedures and performance. The measures serve as reference points, which the operator can follow up, experiment and see the effects of changes in for example working technique. TimberLink software has also been used studying fuel efficiency in harvesting work [Tik08].

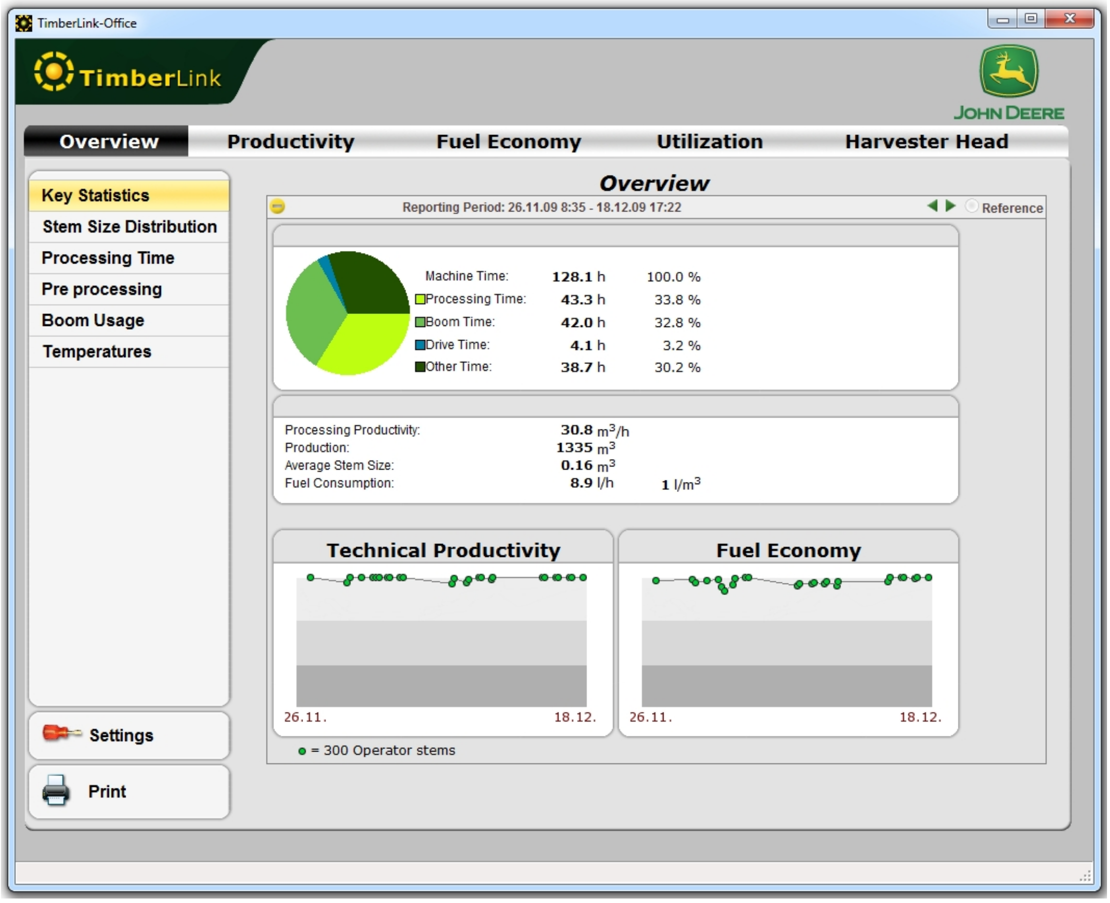

Figure 7.2: Harvester performance monitoring software, TimberLink.

7.1.2 Test case setup

The test case consists of normal working data measurements, collected from eight harvester operators. The operators worked on two separate harvesters of the same model (John Deere 1070D), four operators on both machines. The test involved professional harvester operators, each having their individual backgrounds, abilities and characteristics. To emphasize this human factor and individuality in the results, the operators are referred with names instead of letters or numbers. However, in order to preserve the anonymity of the test operators, their real names are replaced with fictious names.

The test case harvesters and operators are referred as:

Harvester 1 (*H1*): *Harry*, *Andy*, *Robert*, and *Mike*

Harvester 2 (*H2*): *Pete*, *Larry*, *Oscar* and *Tommy*

The measurement data was collected using TimberLink software without any additional measuring equipment. The measured work periods were also recorded on video to verify the results of the tests.

The worksite consisted of relatively small thinning harvesting. Thinning operation is regarded more demanding than regeneration harvesting, since the thinning work involves more decision making tasks than in regeneration harvesting, where all trees are cut down. In thinnigs, the operators have to for example choose the most suitable strip routes, the harvested trees and they have to avoid damage to the remaining trees. Due to the fact that it is extremely difficult to keep the working conditions as similar and comparable as possible for all operators, each operator worked only about an hour for the recorded measurements. Some characteristics of the worksite are:

- Total number of harvested trees: 952
- Mean breast height diameter: 10.1 cm
- Mean stem length (commercial part): 6.2 m
- Mean stem volume (commercial part): 46 dm^3

Since the recorded work time was not particularly long, it is possible that randomness in environmental factors may have affected the results. It is also possible that some of the operators may have speeded up their work pace for the measurements into a level they would not be able to keep up in the long run. However, the measured work time is enough to reveal the characteristics of the operators' skills and work techniques. In [Ter10a] it has been found that the standard deviations of the skill metrics calculated from individual operators are generally smaller than the differences between the metric values of different operators. Thus, the operators have individual skill levels and a work styles that remain relatively constant.

Confidential details of the data acquisition and preprocessing systems and also the work task detection system are not shown. The function of these systems is explained on the level that is needed for comprehension of the methods presented in this thesis. Also, only normalized units (ratios, percentages) are used throughout the case study. In addition, to be able to illustrate the core principles of the developed methods also to readers, who are not familiar with the mechanized timber harvesting technology, a lot of simplifications and generalizations about the harvesting work analysis are made.

7.2 Machine performance evaluation

The overall performance of forest harvester (man-machine system) depends mainly on

- Skills and knowledge of the human operator
- Technical performance and limitations of the machine
- Operating conditions

The task dependent skill evaluation and operator assistance procedure, proposed in Chapters 5 and 6, assumes that the machine condition is perfect and technical performance is constant during the observation window. Although the machine has automated functions that are comparable from one work cycle to another, there are still many work phases, which depend totally on the actions of the operator. Thus, the operator performance cannot be deterministically separated from the machine technical performance. Therefore the technical performance must be checked first, before one can draw conclusions about the skills of the operator.

The machine performance is evaluated applying the methods introduced in Chapter 3. In this illustrative example, four low level indices related to different machine functions are calculated. These low level indices are referred as *MI1, MI2, MI3* and *MI4*. Also two high level overall performance indices are calculated, which are productivity and fuel consumption per harvested timber. These high level indices are referred as *ProductivityIndex* and *FuelEconomyIndex*. Productivity is defined as in Chapter 2, the relative OEE productivity metric, which is the ratio of theoretical ideal work cycle time and the average measured work cycle time. Usually productivity in forestry is defined by the ratio of the produced timber volume and the elapsed work time. However, the use of OEE definition in this case reduces the effect of slightly different stem volumes to the productivity and therefore emphasizes the variation caused by other factors such as the operators' skills.

The calculated performance indices are shown in Figure 7.3. The four operators of harvester *H1* are on left side of the figure and the operators of harvester *H2* are on right side. The most evident observation from the index data is that the low level indices are roughly on the same level for all the operators of the same machine. This is the expected result, since the low level indices were designed to indicate the performance level of the machine functions, which should depend on the condition of the machine, and not on the operators' skills. Index *MI4*, especially on harvester *H2*, is somewhat lower than the other low level indices, indicating that the condition of the machine functions related to index *MI4* should be checked. However, the level is not critical, and according to the machine index data, the condition of the two machines is acceptable and the assumption, that the machine condition is perfect and technical performance is constant during the observation window, is valid.

However, it can be seen from the high level indices, that there are some noticeable differences between the operators and harvesters. One can notice that fuel economy of harvester *H2* is lower than on *H1*. The machine index data suggests that the difference between the harvesters may be caused by lower performance in machine functions related to *MI4*. Another significant observation is that the operators with lower productivity also have lower fuel economy. Reasons for good productivity have been studied in [Ova04]. It has turned out, that often the harvester operators with high productivity, also use fuel efficient working techniques. Efficient work technique shows for example in avoiding unnecessary dragging of the stems on the ground and minimizing reverse driving.

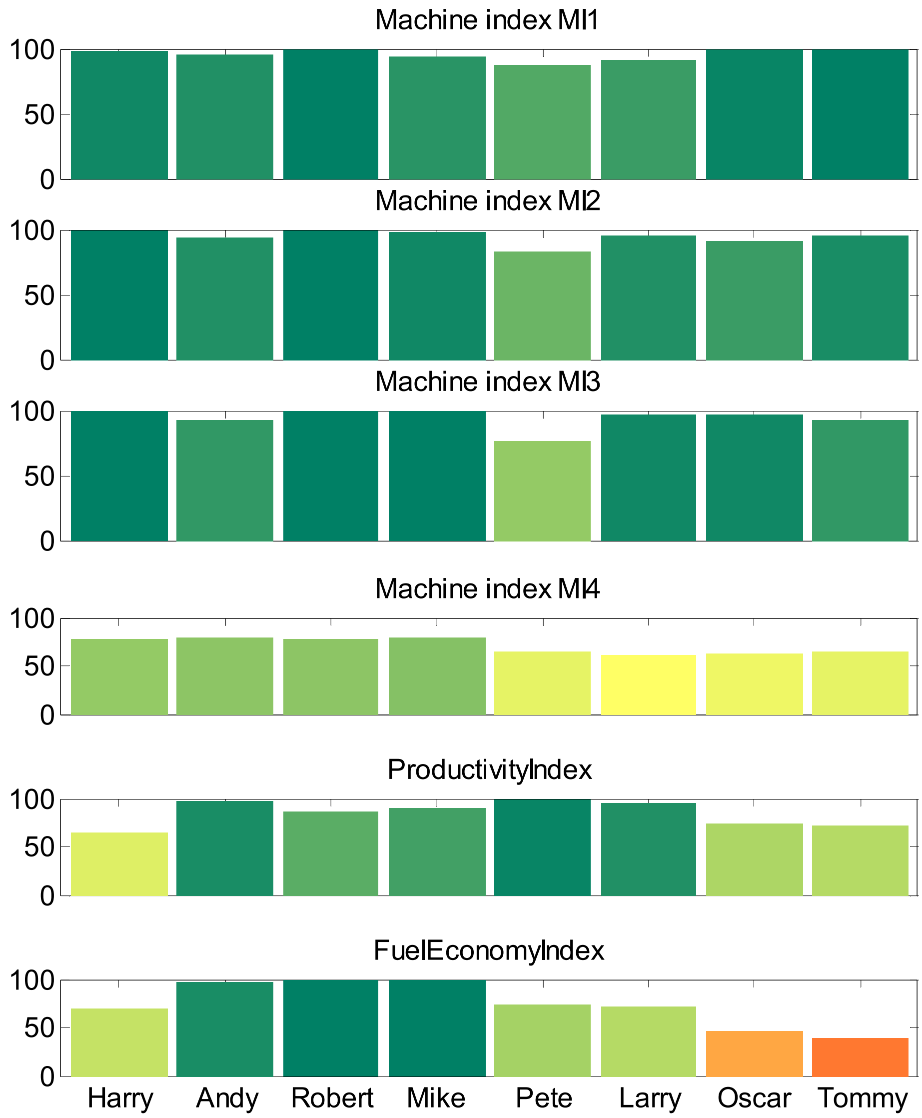

Figure 7.3: Machine performance indices for harvester. Four machine functions related indices (top), productivity index and fuel economy index.

It seems evident that both productivity and fuel consumption depend also on the skills and working technique of the operator and not only the machine condition. Therefore these high level indices cannot be used alone to evaluate the performance of the machine. The evaluation of the machine technical performance and condition should be based primarily on the low level machine indices, which are calculated from performance measures of separate machine functions.

7.3 Recognition of work tasks

The machine performance evaluation revealed two important issues:

1) The technical condition of both of the harvesters is acceptable
2) There exists noticeable differences in productivity and fuel consumption

However, the high level performance indices were unable to tell the reasons for the differences in productivity and fuel consumption from one operator to another. Also, based on only the index data, it is not possible to suggest the operators, what they could do in a different way to improve their productivity and fuel economy.

The next step of the analysis is the recognition of work phases and tasks. The work tasks of each work cycle (each processed stem) are recognized using the methods proposed in Chapter 4. The HMM models of the work task recognition system are not trained using the same particular data set that in this test case example.

To simplify the illustration of the work task analysis in this example, the actual work tasks in the cycles that are obtained from the work task recognition system, are combined into four very general work task categories. Each of the four presented task categories consists of multiple separate work tasks. Although the work tasks are presented this way, recognition of all the separate work tasks inside these categories is necessary to be able to calculate for example the metrics for task level skill evaluation and work technique analysis.

The four general work task categories are:

I. *Driving* during the work cycle

II. Work tasks related to *crane work* (moving the processing head to the tree, moving the felled tree stem, etc.)

III. Work tasks related to *decision making* (Choosing the tree to be felled, bucking decisions)

IV. Work tasks related to execution of *machine functions* (cutting, feeding of stem, etc.)

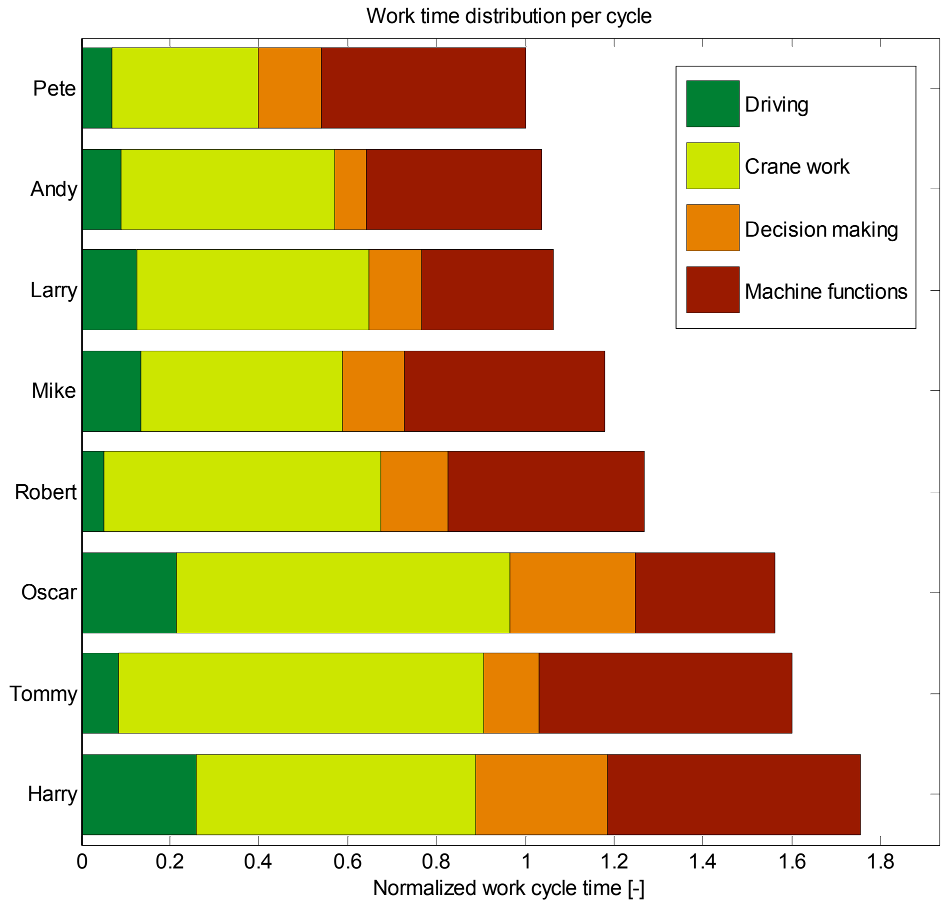

Figure 7.4: Comparison of the test operators in average normalized work cycle times and the distribution of work cycle time into general work categories.

The average normalized work cycle times of the eight operators are shown in Figure 7.4. The work cycle times are normalized to the fastest operator work cycle time. The operators are sorted according to their average work cycle time. The average time distribution inside work cycles can also be seen in Figure 7.4. The longest average work cycle time is approximately 1.7 times longer than the fastest, which leads to significant difference in productivity. Assuming the average stem volumes are the same, 1.7 times longer work cycle time results in 41% lower productivity in timber volume. Similar differences in productivity are reported also in [Ova04].

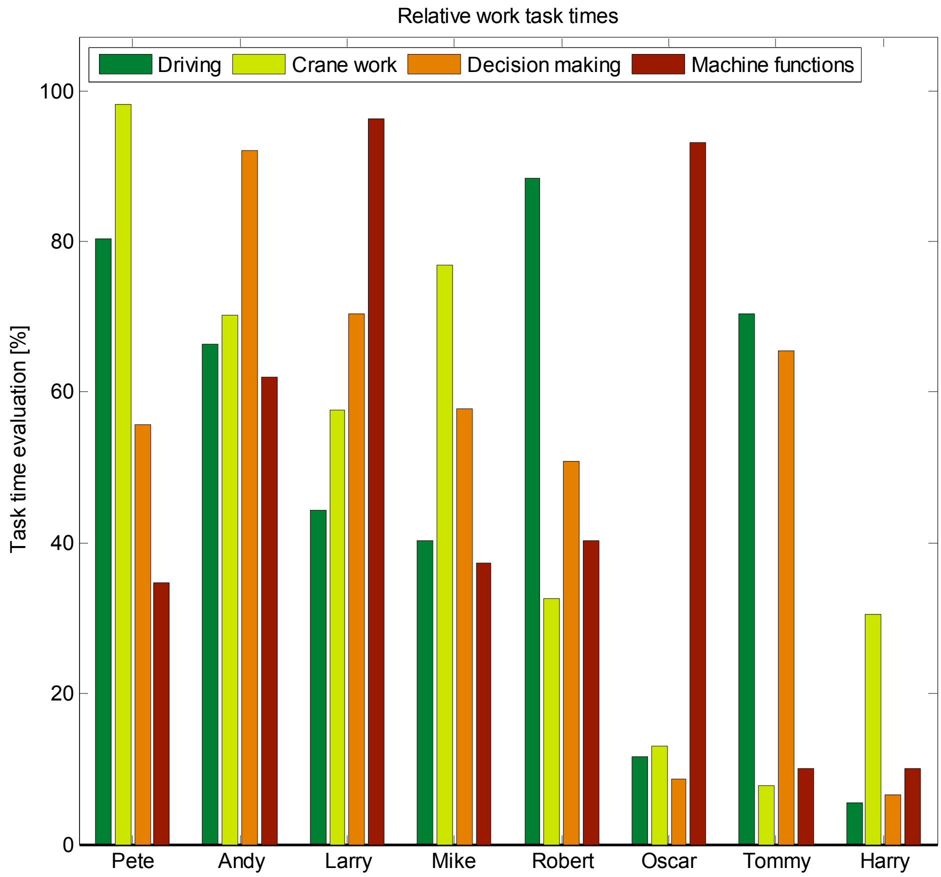

Figure 7.5: Relative work task times of the test operators.

However, since the measured work time for each operator was only about an hour, it must be noted that very likely the operating conditions were not equally favorable for all operators. For example, the average distances between the harvested trees, terrain and the quality of the timber were not the same, which naturally leads to slightly different driving times. Therefore, for a more reliable comparison, the productivities of the operators should actually be compared after much longer periods of working time from variety conditions.

For the illustration of the areas of improvement for each operator, the separate work task times of the operators are compared in Figure 7.5. The operators are sorted according to their average work cycle time as in Figure 7.4. The values shown in Figure 7.5 are relative values of the average work task times, which are obtained using similar statistical reference procedure presented in Chapter 5 that is also used in skill evaluation. The distribution of average work task times is modeled using Gumbel distribution (block maximum form). The relative values of each operator are obtained using the cumulative distribution function (5.20). The parameters of the distribution are obtained using (5.17) and (5.18), by using the average work task times of the eight test operators. By using only the data from this kind of small group of test operators, the obtained distributions do not necessarily generalize and present well the distributions of the work task times of harvester operators in general, but even so, it enables the comparison of the operators inside this test group.

The results of the work task analysis show that *crane work* tasks are very important for achieving good productivity. There is also a lot of variance in the total length of *crane work* tasks between the operators. In order to be efficient in *crane work* tasks, a good motor-sensory skill and an efficient work technique are required [Ter10a]. The five most productive operators are relatively fast in all the four work task categories, but for the three least productive operators, there are clearly some areas that require attention. This information alone could be enough for an experienced and motivated professional operator to be able improve productivity. However, it would be more valuable and interesting to examine in more detail the factors that lead to the differences.

The recognition rates of work were not computed for this data set. However, in studies of similar work cycle and task recognition systems [Pal06] and [Aul09], the reported recognition rates have been good. The errors were mostly related to the timing of the transition from one state to another, and completely missed states were rare. It is noted that the state transitions are actually more or less gradual, which makes timing of the transitions an ambiguous task. This means that the state transitions in the manually selected training data may also be inaccurate. In [Pal06], the recognition rates (percentages of correctly recognized HMM states) for forwarder work were 90.3% for loading and 95.5% for unloading. In [Aul09], the work of harbor crane was modeled using HMMs, also with successful recognition rates.

7.4 Operator skill evaluation

Task level skill evaluation of harvester and forwarder operators was studied in [Ter10a]. The results showed that skill metrics derived using the presented definition of machine operator skill (Definition 5.1) make a clear distinction between the operators in both of the CTL forest machine applications. In [Ter09], a hierarchical method for evaluation of the skill components of the Definition 5.1 was introduced. The results of these studies show, that it is important to evaluate skill on the task-level, since the operators' skills may differ from one task to another. Crane control skill was found to be very important in both harvester and forwarder applications, but in addition harvester work requires a lot of cognitive, decision making skills. It was also shown in [Ter09] that the possession of only good machine controlling skills does not necessarily imply productive and efficient work. The operator must also possess a good working technique and strategy.

In this test case example, the recognition of work tasks revealed the most important areas of improvement and showed that there are very large differences between operators are in the work tasks related to *crane work*. However, the work task recognition did not answer the question: what are the main reasons for the differences between the operators. Hence, skill evaluation and an analysis of work technique are required.

The skills of these eight operators are evaluated by applying the methods introduced in Chapter 5. The evaluated skill metrics are shown in Figure 7.6. Seven different skill metrics are used, which are related to

- Average simultaneous boom movements in three separate work tasks (metrics *C1, C2, C3*)
- Decision making times in two separate work tasks (metrics *D1, D2*)
- Task sequence smoothness in two separate operational phases (metrics *S1, S2*)

The work techniques of the operators do not have much effect on these seven skill metrics, so these two aspects can be analyzed separately. According to [Gel02] and [Ter10a], high average simultaneous crane control movements correlate with motor-sensory skill, while task sequence smoothness and decision making times are more related to cognitive skills. [Ter10a]

Similarly as in previous figures, the operators in the Figure 7.6 are sorted according to their average work cycle time (productivity). The distribution of the average skill metric values among the operators is modeled using Gumbel distribution. Metrics *C1, C2, C3* use block minimum form, others the block maximum form. Relative values, obtained using cumulative distributions of the each skill metric, are shown. The parameters of the distributions are obtained using (5.17) – (5.19), by using the average skill metric values of the eight test operators. As previously, the distribution approximations based on no more than the data of the test operators enables only a comparison between the test operators.

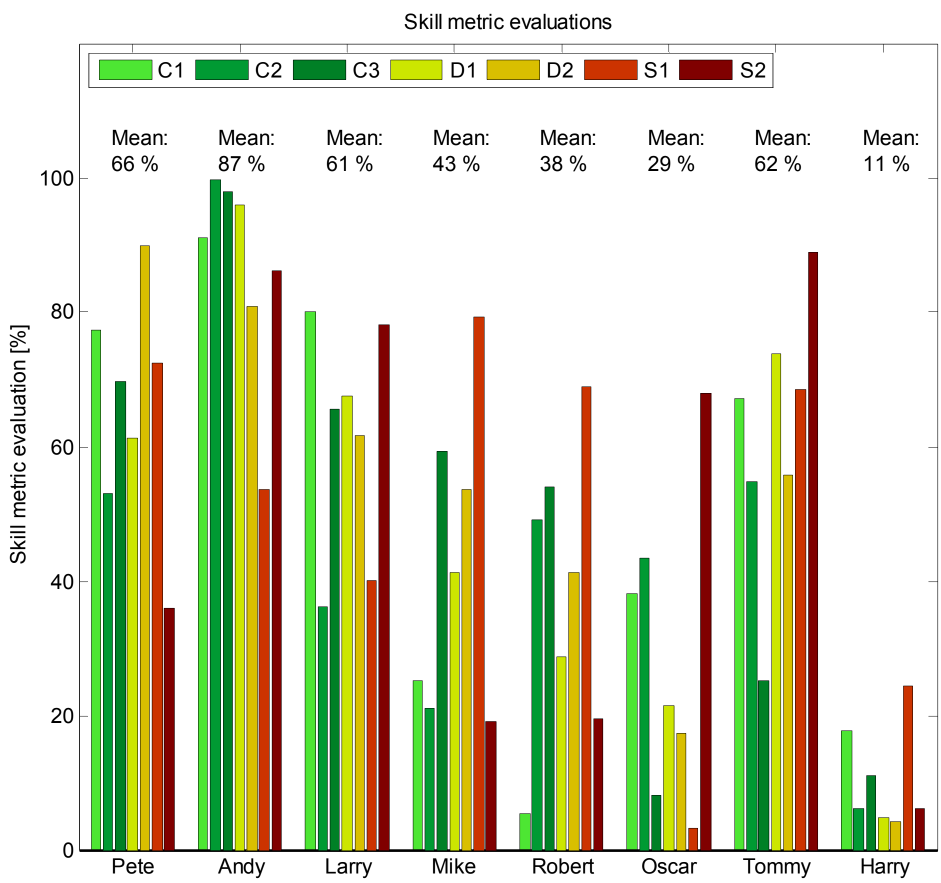

Figure 7.6: Skill metric evaluations of the test operators.

It seems natural to think that the operators, who reach good productivity, are skilled and also reach good values of skill metrics. The Figure 7.6 shows that good skill is clearly related to high productivity. The three most productive operators reach better than average values of skill metrics. Based on these skill metrics, it can be seen that operator *Andy* is clearly the most skilled operator of this group. *Andy*'s good skills are especially related to machine control skills (metrics *C1*, *C2*, *C3*).

However, it can also be seen from the figure that it is possible achieve very good productivity with around the average skill levels and that good skills do not necessarily lead to high productivity. This same conclusion was also made in [Ter09]. For example, the work cycle times of operator *Pete*, *Andy* and *Larry* are quite the same, but according to the skill metrics, *Andy* is clearly the most skilled of these three operators. On the contrary, operator *Tommy* has better than average values in skill metrics, but his productivity is only the second lowest of the group. His relatively good skills are mainly related to cognitive skills, planning and decision making abilities (metrics *D1*, *D2*, *S1*, *S2*). Another interesting observation about operator *Tommy* is that although being quite skilled, according to Figure 7.3, his fuel economy is the lowest in the group. Intuitively, the most obvious reason for the differences of productivities in these cases is not the skill, but it could be found in the efficiencies of the operators' working techniques.

As a conclusion, the analysis of operators' average work task times and skill evaluation reveals that good skills are clearly related to high productivity. However, good productivity and fuel economy can be also reached with around the average skills. The possession of good skills does not guarantee high productivity and high fuel economy, also a good working technique is required. It must also be noted that skill metrics defined this way are indirect measures of different components of skill. Their values certainly include some degree of inherent uncertainty and imprecision and they should be considered as indicative and suggestive, rather than treating them as precise answers.

7.5 Work technique analysis

The working technique analysis in this case example is concentrated only on work techniques related to *crane work* tasks. A lot of simplifications are made and the analysis is based on some of the most common crane work techniques and does not cover the entire *crane work* time. In thinning harvesting, the operators have to choose the strip route and which trees are cut down. They must consider the remaining trees and avoid tree damage, which considerably affects to the crane work technique. The remaining "edge trees" along the side of the strip route restrict the crane movements, and therefore the operators need to master a lot of different crane work techniques and to be able to quickly choose the most efficient technique for each position and situation. [Ova04, Ova06, Ova09]

Based on the available measurements, five general crane work technique categories were defined for the analysis. The idea is to categorize the boom crane movements and movement combinations that the operator uses while positioning the harvester head to the removable trees, moving the trees after felling cut and positioning the tops of the trees. The techniques are referred to *technique1, technique2 ... technique5*, and they are classified according to the combinations of crane motions during consecutive *crane work* tasks. The techniques are labeled in ascending order according to the average work task time consumption, such that *technique1* is on the average the quickest and *technique5* the slowest. The crane movement distances are also estimated to be longer than average when the operator uses *technique4* and *technique5*, which additionally increases the fuel consumption. Moreover, it has been found in many studies that the unnecessary work movements and distances should be minimized in order to increase productivity of harvesting [Har91], [Ova09]. The whole *crane work* time is classified to some of these five techniques or to category *other*. The category *other* includes the *crane work* time when none of these five techniques is applied. It also includes of the work time of other, less frequent *crane work* tasks, which cannot be categorized to these working techniques, for example clearing of small trees from the worksite.

The use of different crane work techniques is presented in Figure 7.7. The figure shows crane work time per work cycle, normalized by the fastest total crane work time per work cycle. As in the previous figures, the operators are sorted according to their average work cycle time. The figure shows that operator *Pete* uses the shortest time in crane work, while *Tommy* works with the crane 2.5 times longer per work cycle. Due to the fact that the remaining trees restrict the crane movements, different work techniques are needed for different situations. The operators use all the five techniques, but there are differences how the operators master the techniques and how successfully they choose the most efficient one in each situation.

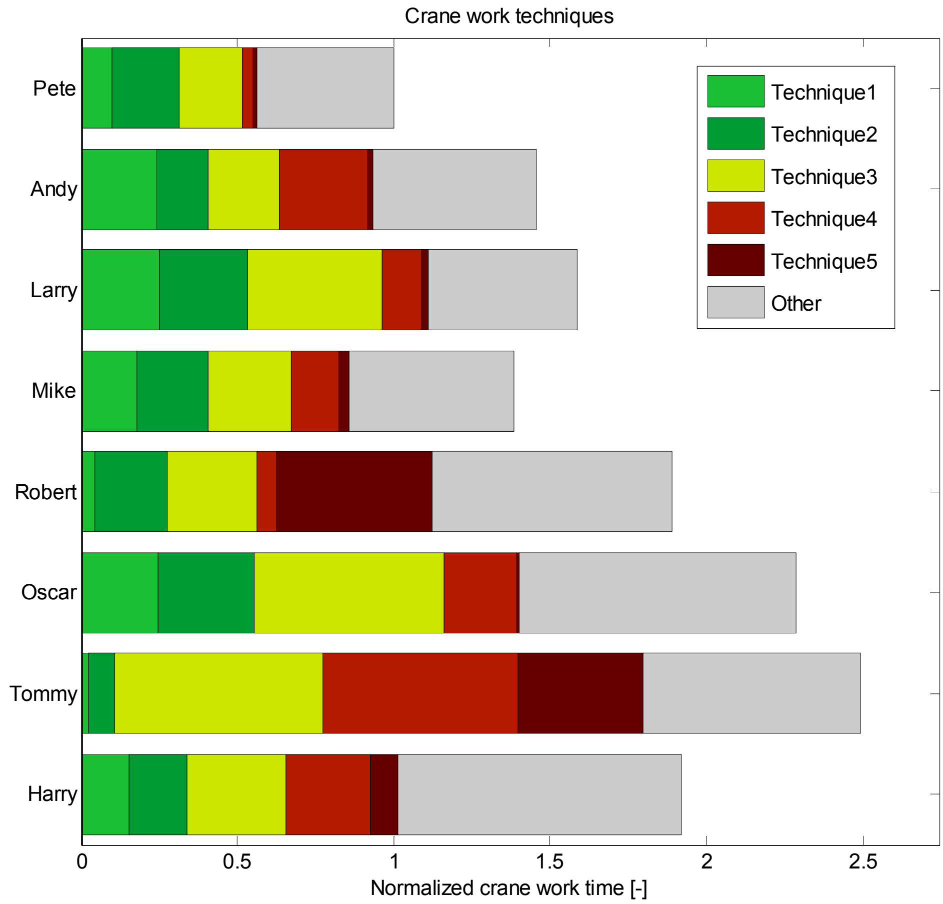

Figure 7.7: Crane work techniques.

The differences between the work techniques of operators *Pete* and *Tommy* are compared in Figure 7.8. The figure shows the normalized frequencies, how often each of the techniques is used. It can be seen clearly that *Pete* cuts down the most of the trees using *technique1* and *technique2*, which are the two fastest crane work techniques. Meanwhile, *Tommy* uses *technique3*, *technique4* and *technique5* more commonly. Since *technique4* and *technique5* are the two slowest crane work techniques on the average, there is a huge difference between the operators in crane work time, although both operators are about as skilled based on the skill metrics shown in Figure 7.6. The *Tommy's* low value of *FuelEconomyIndex* in Figure 7.3 is also explained with the frequent use of *technique4* and *technique5*, which results in longer than average stem moving distances.

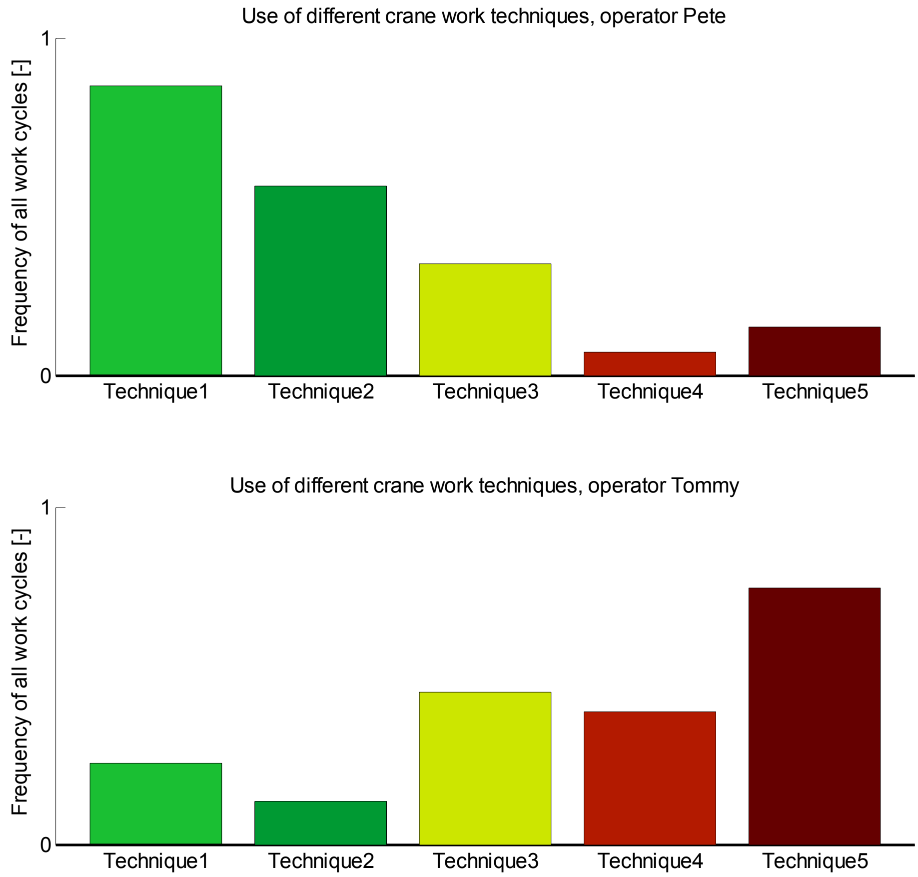

Figure 7.8: Comparison of crane work techniques of operators *Pete* and *Tommy*.

7.6 Intelligent coaching system for harvester operators in crane work tasks

Finally, an intelligent coaching system for harvester operators is formed. The system of this case example is relatively simple and it concerns only the work tasks related to *crane work* at rather general level of detail. In practice, the system should provide the operators more detailed information and cover also the other work tasks. However, the main purpose this example is to illustrate how the methods proposed in Chapter 6 can be used. The ICS observes if there is potential for improvement in crane work and if necessary, it suggests where the operator should focus on to improve work performance.

The membership functions of the fuzzy input variables are calculated using statistical reference procedure proposed in Chapter 5 and the fuzzification and membership function parameterization proposed in Chapter 6. Thus, the inference behind the suggestions for the operator is based on linguistic (qualitative) expert knowledge, input data measured from the operator and corresponding statistical reference.

The following linguistic expert knowledge is used:

I. If crane work is efficient and productivity is high, there is *no need to give suggestions on work technique*.

II. If there is potential for improvement in crane work, and the most efficient work techniques are not used, the operator should *focus on using more efficient work techniques*.

III. If there is potential for improvement in crane work, and mainly efficient work techniques are used, and crane control skill is low, the operator should *focus on controlling the crane and check that the crane control parameters are suitable*.

This knowledge is then formulated into a rule base of a fuzzy inference system:

Rule1. IF *CraneWorkImprovementPotential* IS *Low*
AND *Productivity* IS *High*
THEN *Suggestion* IS *NoSuggestions*

Rule2. IF *CraneWorkImprovementPotential* IS *High*
AND *CraneWorkTechnique* IS NOT *FocusedOnEfficientTechniques*
THEN *Suggestion* IS *FocusOnCraneWorkTechnique*

Rule3. IF *CraneWorkImprovementPotential* IS *High*
AND *CraneWorkTechnique* IS *FocusedOnEfficientTechniques*
AND *ControlSkillLevel* IS *Low*
THEN *Suggestion* IS *FocusOnCraneControl*

The input variables are calculated from the measurements, using the procedure presented in Chapter 6. The input variables are defined as

- *CraneWorkImprovementPotential* is the calculated from average *crane work* task time. It has two membership functions, *Low* and *High*. *High CraneWorkImprovementPotential* means that the average measured crane work time is relatively long compared to the statistical reference.
- *Productivity* is calculated from the average measured work cycle time. It has two membership functions, *Low* and *High*. *High Productivity* means that the measured average work cycle time is short compared to the statistical reference.
- *CraneWorkTechnique* is calculated from the measured *crane work* times, assigned to different crane work techniques. It has one membership function, *FocusedOnEfficientTechniques*. It means that the ratio of the times

$$\frac{technique1 + technique2}{technique1 + technique2 + technique4 + technique5} \tag{7.1}$$

  is high compared to the statistical reference.
- *ControlSkillLevel* is calculated from the average of the skill metrics *C1, C2* and *C3*. It has one membership function, *Low. Low ControlSkillLevel* means that the average of the measured skill metrics *C1, C2* and *C3* is low compared to the statistical reference.

The output of the fuzzy inference system is the suggestion of the rule corresponding to highest total fuzzy similarity with the actual inputs. Each of the inputs in each rule was considered equally important, so weighting of the inputs was not used. The calculated total fuzzy similarities between the rules and the inputs are shown in Figure 7.9. The figure also shows the output of the inference system, the corresponding suggestion. Only one rule (suggestion) was chosen for each operator, but for some operators it would also have been reasonable to give multiple rules.

As a result of inference, the fuzzy inference system decides that it is not necessary to give any suggestions related to crane work technique to operators *Pete, Andy, Larry* and *Mike* (highest similarity with *Rule1*). These four operators also reached the highest productivity of this group. Though, *Larry* and *Mike* have also relatively high value of total fuzzy similarity also with *Rule3*, which would suggest focusing on crane control.

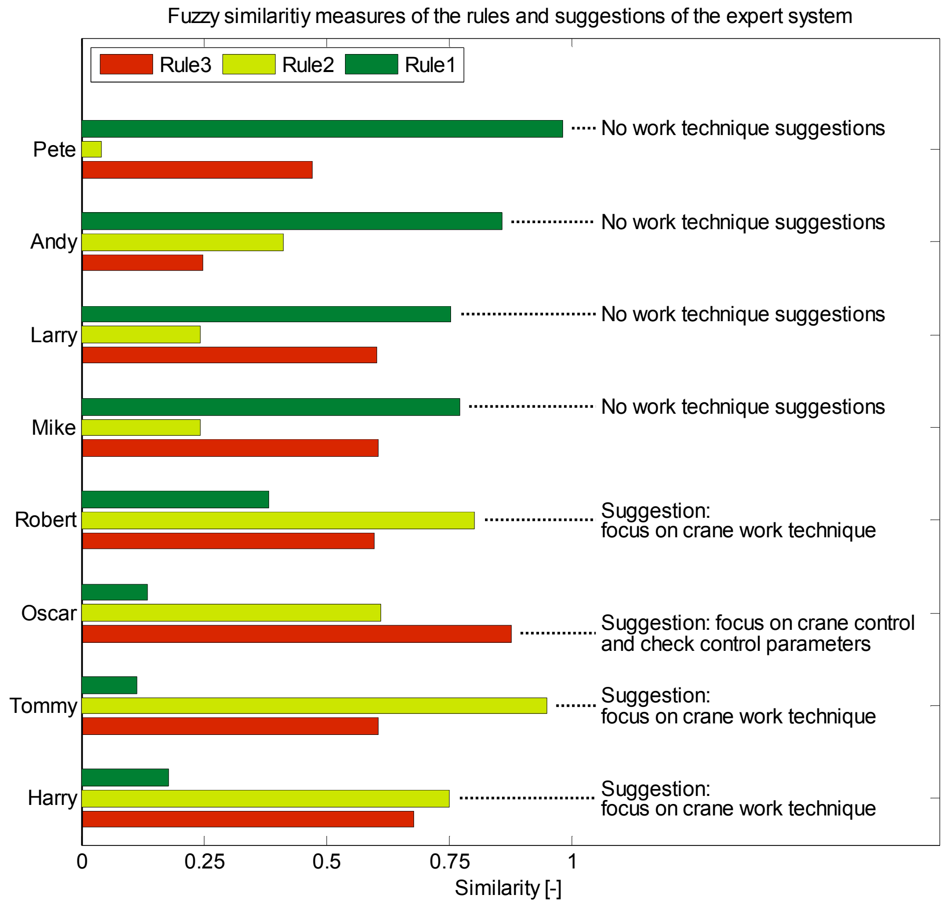

Figure 7.9: Total fuzzy similarity between the rules and the actual inputs and the corresponding suggestions of the expert system.

According to the fuzzy inference system, operators *Robert, Tommy* and *Harry* are advised to focus on using the more efficient crane work techniques (highest similarity with *Rule2*). This suggestion is quite obvious, especially for *Tommy*, after examining Figures 7.7 and 7.8. Operator *Harry* also has relatively high value of total fuzzy similarity with both *Rule2* and *Rule3*, which would suggest focusing on both crane work technique and crane control. The suggestion for operator *Oscar* is to focus on crane control (highest similarity with *Rule3*).

An important observation is that the fuzzy inference system worked reasonably well and without the need of tuning the parameters of the FIS inputs manually. This is very useful feature, since in practice, the FIS could possibly be much more complex and require more extensive expert knowledge, and it could turn out to be rather tedious task to tune all the parameters of each of the inputs separately. As the fuzzification of the inputs was based on the statistical distributions of the average levels of individual operators' performance metrics, basically, all that was needed here was to formulate the expert knowledge into a FIS rule base.

This kind of statistical approach reveals the most obvious deviations of the measurements from the median of the corresponding reference data. The aim of the ICS is to provide interesting and useful information from the process, which can possibly help the operator to reach better performance. It also allows some amount of uncertainty and imprecision in the measurements. Obviously, a real system should incorporate a larger variety of situations and a lot more extensive rule base. However, this example shows that the use of statistical distributions eases the design and parameterization of a fuzzy expert system.

7.7 Summary of results and discussion

This chapter summarized what kind of details and useful information of machine work can be revealed using the methods proposed in this thesis. The proposed methods concern machine performance monitoring, work task recognition and the analysis of operator skill and work technique. The results of the operator analysis in the test case example are scored up in a summary shown in Figures 7.10 and 7.11. The results of the test case showed, that the condition of the two machines can be regarded acceptable. Although the operators worked in similar operating conditions and used similar machines, the performance evaluation results show significant differences between the productivities of the operators. The fuel efficiency was also found to vary among the operators, such that productive, efficient work is generally also fuel efficient.

Performance assesment and condition monitoring in forest harvester application using performance indices has been studied in [Hol05]. The performance evaluation was based on normal working data, without additional sensors. In [Rep06], these index-based performance measures were used as indirect measures for unsupervised fault detection. The results in these publications show, that performance indices point out changes in machine performance, which would have been difficult to detect otherwise due to the variation of operating conditions and the complexity of the system.

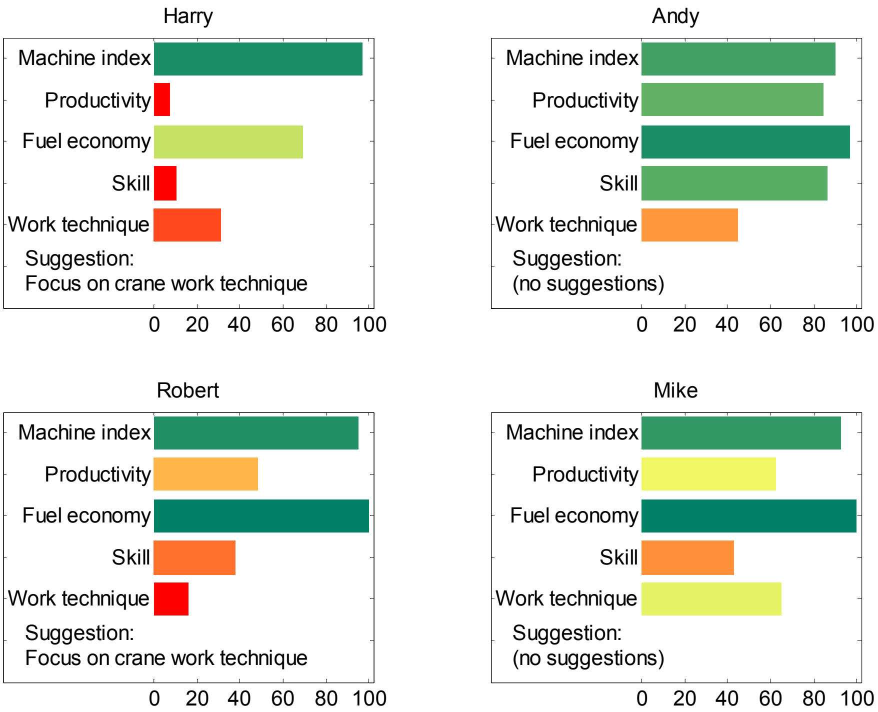

Figure 7.10: Summary of harvester H1

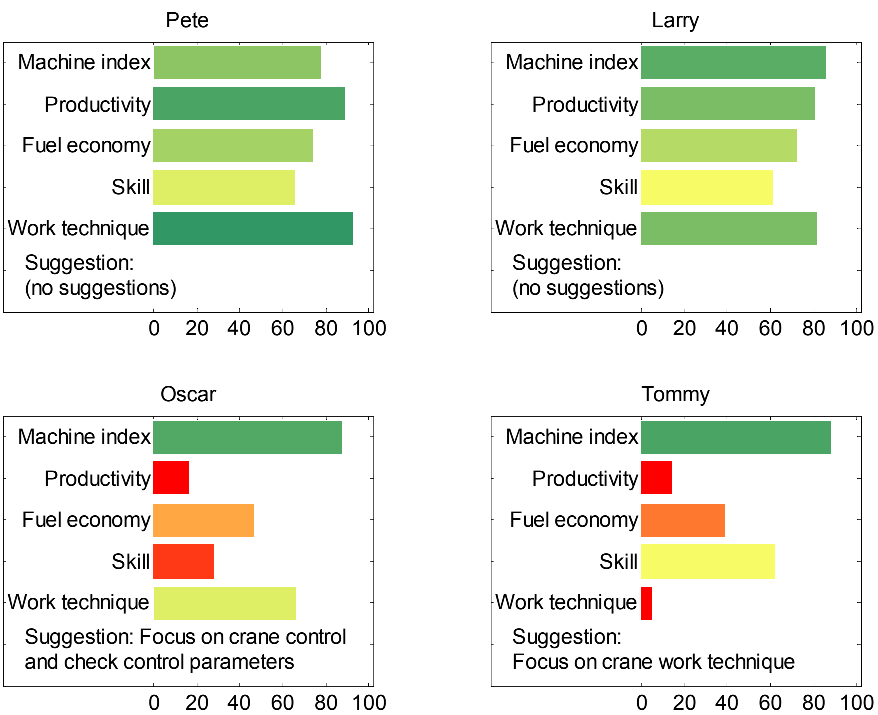

Figure 7.11: Summary of harvester H2

The performance indices have also been shown to be useful for indirect indicators of such component faults, which cause slightly deteriorated performance of the system. However, the fault detection accuracy compromise between false alarms and detection performance. Also, identification of the failed component real-time is generally not possible using indirect measures, fault identification requires direct measurements and possibly also specialized test sequences [Pal04].

The distribution of the operators' work time was calculated using the work task recognition procedure using HMMs proposed in this thesis. The procedure was first proposed in [Pal06] for forwarder work. A similar approach was also successfully applied in the recognition of work cycles in harbor cranes [Aul09]. The reported recognition rates in these studies have been good, and most of the recognition errors relate to timing of state transitions or abnormal, exceptional situations that might appear during the work. Recognition of the separate work tasks has many benefits. It reveals the most important areas of improvement of the individual operators and also enables work task dependent skill evaluation of the operators and more detailed analysis of work technique.

The skill evaluation of the operators showed that there are noticeable differences in skills between the operators. The operators possess different kinds of skills in different work tasks. This shows the importance of task level skill evaluation and the assessment of different components of skill. The case example results show that the possession of good skills commonly implies good productivity. However, the results also show that in the work of a harvester operator, extremely good skills are not necessarily required to reach good productivity; it can also be reached by using efficient work techniques.

The operators' work techniques related to crane work tasks were analyzed. The work time in crane works tasks was classified to different work techniques, which were ordered according to the average work time per work cycle using that particular technique. The results showed that harvester operators in thinning work use multiple work techniques. However, the operators who choose more frequently the most efficient techniques spend less time performing crane work. This is one of the most obvious causes in the differences between operators' productivities and fuel efficiencies.

The ICS uses the expert knowledge related to crane work and aims to help the machine operator to reach better skills and performance by making observations and giving suitable suggestions to the operator. The expert knowledge of the ICS is formulated as rules of a fuzzy inference system. The ICS is based on linguistic expert knowledge, input data and corresponding statistical reference. The fuzzification of inputs and its parameterization is based on the statistics of the variables, and therefore only qualitative expert knowledge is required.

The simple expert system of the presented case example consisted of only three rather general rules and was focused only on work tasks related to crane work. In practical implementations, the number of rules would likely be larger, and other work tasks should be considered as well. The system was formulated in a way that it analyzes the long term performance of the operators, which can be used in operator training. In addition to this, the system can also be implemented in a way that it can give real time feedback to assist the operators during normal work.

The potential improvement that can be gained using operator assistance systems is very large. The over 40% gap in productivity between professional operators, found in the literature and also in the measurements of the case example, is economically very significant. It is also ecologically significant, since the results show that an efficient and productive machine work is fuel efficient as well. Since there are differences in skill and talent between human individuals by nature, it is unrealistic to expect that any kind of feedback and training would raise the skill and performance of each and every operator to the highest levels. However, even smaller improvements would make a considerable difference in the long run.

In the future research related to coaching systems, it would be interesting to conduct a long term follow up of the operators who receive the feedback about their work, and to see how the feedback and the suggestions of the ICS affect to their work performance.

Chapter 8
Conclusion

In the work of mobile machines, the overall performance is affected by the technical limitations of the machine, the operating conditions, as well as the skills and work techniques of the machine operator. In the recent studies of the machine work, the role of the human operator has been found very significant. Nevertheless, the studies and development related machine work have been based on manual classification of the work tasks by an external observer. Any practical and wide-spread means to automatically detect the work tasks and work cycles performed by a human operator have not been available before. This limits the coverage and availability of research material, makes the classification of the work more or less subjective, and makes it more difficult to utilize the material real time for example in the training of the operators.

One of the main contributions of the thesis, a method for work task and work cycle recognition using Hidden Markov models (HMM) was presented in Chapter 4. The approach of this thesis in modeling of man-machine process is essentially the modeling of human actions, decisions and performance while performing machine work. The actions are modeled as a network of tasks, which are consequently modeled as the states of the HMM. The main advantage of the proposed approach is that separate operational phases, work tasks and work cycles are recognized practically real-time during normal work. Additionally, the work task recognition by HMMs is based on the control commands given by the operator, and additional sensors are not required. As a result, the recognition of work tasks enables a more detailed analysis of the work process and offers an opportunity to develop the work of the operators.

Based on the recognition of work tasks, a method for skill evaluation of the machine operators was presented in Chapter 5. The principle of the skill evaluation method is that the operators' skills are evaluated at the task level. The evaluation method of the operator's skill levels is data-driven and uses the statistical values measured from a reference group. Since the metrics used in skill evaluation may be affected by the machine technical condition, the machine performance and condition must also be verified. An index-based method for machine performance evaluation was presented in Chapter 3.

A method of using an intelligent coaching system to assist the work of mobile work machine operators was presented in Chapter 6. The ICS is based on qualitative expert knowledge related to the man-machine work process, work task dependent skill and performance measures and corresponding statistical reference. The intelligent coaching system makes observations and gives suitable feedback to the operator in the form of linguistic suggestions. The expert knowledge is formulated as rules of a fuzzy inference system. An intelligent coaching system can enhance the training of the operators by focusing on the individual needs of learners and the most important areas of improvement and without the need of the presence of a human instructor. As well as providing detailed information for operator training based on long term performance, the system can also be implemented in a way that it can give real time feedback to assist the operators during normal work. Although versatility and situation awareness of a human instructor is much better compared to an automatic instructor, obvious benefits of an automatic feedback system are the availability of long term evaluation at low cost and the objectivity of the feedback.

The potential of improvement in productivity of man-machine work that could be gained using performance monitoring and operator assistance systems is substantial. A case example analysis of cut-to-length forest harvester work, which used the methods proposed in this thesis, was presented in Chapter 7. The case example shows that work task level analysis of the work process can reveal substantial amount of useful information and details about the machine work, which have not been available before. Moreover, the methods presented in this thesis for work task and work cycle recognition, task level skill evaluation of machine operator, work technique analysis and the ICS are based on the measurements and performance measures that are already available from the process. Thus, implementation of the methods does not increase the cost and complexity of the system, since it is not necessary to mount additional measuring equipment.

Recently, the demands of reducing emissions and increasing fuel efficiency of machine work are becoming ever more essential. It has been shown, that a skilled machine operator, who has an efficient working technique, often performs machine work productively and fuel efficiently at the same time. Therefore, the efforts in enhancing operators’ skills and knowledge will not only increase the productivity of the work, but makes it also more fuel efficient. Furthermore, as the actions and work tasks performed by the machine operator become recognizable, the advantages are not limited only to the development of machine operators’ skills; the methods presented in this thesis will also offer new opportunities and valuable tools to evaluate and further develop the machine technology and work procedures to meet new demands of the future.

References

[Ale08] Aleven, V., Ashley, K., Lynch, C. and Pinkwart, N., Intelligent Tutoring Systems for Ill-defined domains: Assessment and Feedback in Ill-Defined Domains. In: *The 9th international Conference on Intelligent Tutoring Systems*, 2008.

[Aul09] Aulanko, S. and Tervo, K., Modeling and evaluation of harbor crane work. In: *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 869-874, 2009.

[Aul10] Aulanko, S. and Tervo, K., Modeling and analysis of harbor crane work efficiency using work cycle recognition. In: *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 2010.

[Bar70] Baron, S., Kleinman, D. L., and Levison, W. H., An Optimal Control Model of Human Response Part II: Prediction of Human Performance in a Complex Task. *Automatica*, **6**: 371-383, 1970.

[Bar90] Baron, S., Kruser, D. S., and Huey, B. M., Quantitative Modeling of Human Performance in Complex, Dynamic Systems. *Panel on Human Performance Modeling, Committee on Human Factors*, National Research Council, 1990.

[Bau70] Baum, L. E., Petrie, T., Soules, G. and Weiss, N., A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics* **41**(1): 164-171, 1970.

[Bjo91] Björheden, R., Basic time concepts for international comparisons of time study reports. *Journal of Forest Engineering* **2**(2): 33-39, 1991.

[Brd09] Brdiczka, O., Langet, M., Maisonnasse, J. and Crowley, J. L., Detecting human behavior models from multimodal observation in a smart home. *IEEE Transactions on Automation Science and Engineering* **6**(4), 2009.

[Bre98] Breuker, J., What are Intelligent Coaching Systems and Why are they (in)evitable? In: *IEEE Colloquium on Artificial Intelligence in Educational Software*, 1998.

[Bru92] Brummer, J. N. L., Characterization of digital channels using hidden Markov models. In: *Proceedings of South African Symposium on Communications and Signal Processing*, pp. 183-188, 1992.

[Bun00] Bunks, C., McCarthy, D. and Al-Ani, T., Condition-based maintenance of machines using hidden Markov models. *Mechanical Systems and Signal Processing* **14**(4): 597-612, 2000.

[Chi01] Chiang, L., Russel, E. and Braatz, R., *Fault Detection and Diagnosis in Industrial Systems*. Springer, 2001.

[Col01] Coles, S., *An Introduction to Statistical Modeling of Extreme Values*. Springer, 2001.

[Dav92] Davidson, J. B. and Schmidt, D. K., Modified Optimal Control Pilot Model for Computer-Aided Design and Analysis, *NASA Technical Memorandum* 4384, 1992.

[Dav94] Davidson, J. B. and Schmidt D. K., Extended Cooperative Control Synthesis, *NASA Technical Memorandum* 4561, 1994.

[Dem77] Dempster, A. P., Laird, N. M. and Rubin, D. B., Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B*, **39**(1): 1-38, 1977.

[Dud00] Duda, R.O., Hart, P.E. and Stork, D.G., *Pattern Classification*, 2nd ed. Wiley-Interscience, 2000.

[Ein09] Einmahl, J. H. J. and Smeets, S. G. W. R., Ultimate 100m world records through extreme-value theory. *CentER Discussion Paper Series*. 2009.

[Eli98] Eliasson, L., *Analyses of single-grip harvester productivity*. Doctoral thesis. Swedish University of Agricultural Sciences, Department of Operational Efficiency, Umeå, 1998.

[Fit52] Fitts, P. M. and Peterson, J. R., Information Capacity of Discrete Motor Responses. *Journal of Experimental Psychology* **67**(2): 103-112, 1964.

[For07] Fortuna, L., Graziani, S., Rizzo, A. and Xibilia, M.G., *Soft sensors for monitoring and control of industrial processes*. Springer-Verlag, 2007.

[Fre98] Freedman, P., Forestry machine simulators: looking for added value in training. In: *Canadien Woodlands Forum Annual Meeting*, 1998.

[Ge08] Ge, M., Xu, Y. and Du, R., An intelligent online monitoring and diagnostic system for manufacturing automation. *IEEE Transactions on Automation Science and Engineering* **5**(1): 127-139, 2008.

[Gel02] Gellerstedt S., Operation of the single-grip harvester: Motor-sensory and cognitive work. *International Journal of Forest Engineering*, **13**(2): 35-47, 2002.

[Goe81] Goertzel, T. and Fashing, J., The Myth of the Normal Curve: A Theoretical Critique and Examination of its Role in Teaching and Research. *Humanity and Society* **5**: 14-31, 1981.

[Gul95] Gullberg, T., Evaluating operator-machine interactions in comparative time studies. *Journal of Forest Engineering* **7**(1): 51-61, 1995.

[Gum41] Gumbel, E. J., The return period of flood flows. *The Annals of Mathematical Statistics* **12**(2): 163-190, 1941.

[Han00] Hånell, B., Nordfjell, T. and Eliasson, L., Productivity and costs in shelterwood harvesting. *Scandinavian Journal of Forest Research* **15**: 561-569, 2000.

[Han05] Hansen, R. C., *Overall Equipment Effectiveness (OEE)*. Industrial Press, 2005.

[Har06] Harashima, F. and Suzuki, S., Human adaptive mechatronics - interaction and intelligence. In: *IEEE International Workshop on Advanced Motion Control*, 2006.

[Har99] Harris, T. J., Seppala, C. T. and Desborough, L. D., A Review of Performance Monitoring and Assessment Techniques for Univariate and Multivariate Control Systems. *Journal of Process Control* **9**: 1-17, 1999.

[Har88] Harstela, P., Principle of comparative time studies in mechanized forest work. *Scandinavian Journal of Forest Research* **3**: 253-257, 1998.

[Har91] Harstela, P., *Work studies in forestry*. Silva Carelica 18, 1991.

[Har05] Harstela, P., The competence of forest-machine operators and tacit knowledge. In: *Proceedings of the International Seminar on Simulator-based Training of Forest Machine Operators*, 2005.

[Hic52] Hick, W. E., On the rate of gain of information. *Quarterly Journal of Experimental Psychology* **4**: 11-26, 1952.

[Hid06] Hidaka, K., Saida, K., and Suzuki, S., Relation between skill level and input-output time delay. In: *Proceedings of the IEEE International Conference on Control Applications*, pp. 557-561, 2006.

[Hym53] Hyman, R., Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology* **45**: 188-196, 1953.

[Hol05] Hölttä, V., Repo, M., Palmroth, L. and Putkonen, A., Index-based performance assessment and condition monitoring of a mobile working machine. In: *The 2005 ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2005.

[Hol09] Hölttä, V., *Plant performance evaluation in complex industrial applications*. Doctoral dissertation, Department of Automation and Systems Technology, Helsinki University of Technology, Espoo, Finland, 2009.

[Ita97] Itabashi, K., Hayakawa, K., Suzuki, T., Okuma, S., and Fujiwara, F., Modeling of the peg-in-hole task based on impedance parameters and HMM. In: *Proceedings of the IEEE International Conference on Intelligent Robots and System*s, pp. 451-457, 1997.

[Iwa06] Iwase, M., Hatakeyama, S., and Furuta, K., Analysis of Intermittent Control Systems by Human Operations. In: *Proceedings of the IEEE Conference on Industrial Electronics*, 2006.

[JDF10] John Deere Forestry – TimberLink. [web page]. Cited 7 Sep 2010. Available at: http://www.timberoffice.com/english/products/timberlink/.

[Jua91] Juang, B. H., and Rabiner, L. R., Hidden Markov Models for Speech Recognition, *Technometrics* **33**: 251-272, 1991.

[Kan91] Kandel, A., *Fuzzy expert systems*. CRC Press, 1991.

[Kar04] Kärhä, K., Rönkkö, E. and Gumse, S.-I., Productivity and cutting costs of thinning harvesters. *International Journal of Forest Engineering* **15**(2): 43-56, 2004.

[Kle98] Klein, G., *Sources of Power: How People Make Decisions*. Cambridge, MA: MIT Press, 1998.

[Koi97] Koivo, A. J. and Repperger, D. W., Skill evaluation of human operators. In: *IEEE International Conference on Computational Cybernetics and Simulation* **3**: 2103-2108, 1997.

[Kot00] Kotz, S. and Nadarajah, S., *Extreme Value Distributions: Theory and Applications*. World Scientific, 2000.

[Kur04] Kurihara, K., Suzuki, S., Harashima, F., and Furuta, K., Human adaptive mechatronics (HAM) for haptic system. In: *IEEE Conference on Industrial Electronics Soc*iety **1**: 647-652, 2004.

[Lev83] Levinson, S. E., Rabiner, L. R., and Sondhi, M. M., An Introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition. *Bell System Technical Journal* **62**(4): 1035-1073, 1983.

[Luk70] Łukasiewicz, J., *Selected Works*. Cambridge University Press, Cambridge, 1970.

[Luu03] Luukka, P., and Saastamoinen, K., Fuzzy similarity based classification in normal Łukasiewicz algebra with geometric and harmonic mean. In: *3rd EUSFLAT, an international conference in fuzzy logic and technology*, pp. 242-245, 2003.

[Mac91] MacKenzie, I. S., *Fitts' law as a performance model in human-computer interaction*. Doctoral dissertation. University of Toronto, Canada, 1991.

[Man01] Mäntylä, V.-M., *Discrete hidden Markov models with application to isolated user-dependent hand gesture recognition*. VTT Publications 449, 2001.

[Meg06] Megali, G., Sinigaglia, S., Tonet, O., and Dario, P., Modelling and evaluation of surgical performance using hidden Markov models. *IEEE Transactions on Biomedical Engineering* **53**(10): 1911-1919, 2006.

[Mur03] Murphy, T. E., Vignes, C. M., Yuh, D. D., and Okamura, A. M., Automatic Motion Recognition and Skill Evaluation for Dynamic Tasks. In: *Proceedings of Eurohaptics*, pp. 363-373, 2003.

[Nii03] Niittymäki, J. and Turunen, E., Traffic signal control on similarity logic reasoning. *Fuzzy Sets and Systems* **133**(1): 109-131, 2003.

[Nor88] Norman, D. A., *The Psychology of Everyday Things*. Basic Books, New York, 1988.

[Nur06] Nurminen, T., Korpunen, H. and Uusitalo, J. Time consumption analysis of the mechanized cut-to-length harvesting system. *Silva Fennica* **40**(2): 335-363, 2006.

[Ong07] Ong, J., Automated Performance Assessment and Feedback for Free-Play Simulation-Based Training. *Performance Improvement* **46**(10): 24-31, 2007.

[Ova04] Ovaskainen, H., Uusitalo, J. and Väätäinen, K., Characteristics and significance of a harvester operators' working technique in thinning. *International Journal of Forest Engineering* **15**(2): 67-77, 2004

[Ova05] Ovaskainen, H., Comparison of harvester work in forest and simulator environments. *Silva Fennica* **39**(1): 89-101, 2005.

[Ova06] Ovaskainen, H., Uusitalo, J. and Sassi, T., Effect of edge trees on harvester positioning in thinning. *Forest Science* **52**(6): 659-669, 2006.

[Ova07] Ovaskainen, H. and Heikkilä, M., Visuospatial cognitive abilities in cut-to-length single-grip timber harvester work. *International Journal of Industrial Ergonomics* **37**(9-10): 771-780, 2007.

[Ova09] Ovaskainen, H., *Timber harvester operators' working technique in first thinning and the importance of cognitive abilities on work productivity*. Dissertationes Forestales 79. University of Joensuu, Faculty of Forest Sciences, Finland, 2009.

[Pal04] Palmroth, L. and Hölttä, V., Rapid control prototyping in test sequence design for condition monitoring of a hydraulic valve. In: *IEEE Mechatronics & Robotics*, Part I, pp. 218-223, 2004.

[Pal06] Palmroth, L. and Putkonen, A., Work cycle detection in human operated machines using Hidden Markov Models. In: *The 8th International Conference on Motion and Vibration Control*, 2006.

[Pal09] Palmroth L., Tervo K. and Putkonen A., Intelligent coaching of mobile working machine operators. In: *Proceedings of the IEEE 13th International Conference on Intelligent Engineering Systems*, pp. 149-154, 2009.

[Pav79] Pavelka, J., On fuzzy logic I, II, III. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* **25**: 45-52, 119-134, 447-464, 1979.

[Pol66] Polanyi, M., *The Tacit Dimension*. Doubleday & Co, 1966.

[Pow02] Power, D. J., *Decision support systems: concepts and resources for managers*. Quorum Books, 2002.

[Qiu05] Qiu, H., Liao, H. and Lee, J. Degradation assessment for machinery prognostics using hidden Markov models. In: *Proceedings of ASME International 20th Biennial Conference on Mechanical Vibration and Noise*, pp. 531-637, 2005.

[Rab86] Rabiner, L. R. and Juang, B. H., An introduction to hidden Markov models. *IEEE ASSP Magazine* **3**(1): 4-16, 1986.

[Rab89] Rabiner, L. R., A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE* **77**(2): 257-286, 1989.

[Ran03] Ranta, P., Possibilities to Develop Forest machine Simulator Based Studying. In: *The Eleventh International PEG Conference: Powerful ICT for Teaching and Learning*, 2003.

[Ras83] Rasmussen, J., Skills, Rules and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models. *IEEE Transactions on Systems, Man, and Cybernetics SMC*-**13**(3), 1983.

[Rep06] Repo, M., Hölttä, V. and Palmroth, L., Unsupervised fault detection of forest harvester head functions. In: *Preprints of SAFEPROCESS*, 2006.

[Rob95] Robinson, M. E. and Tawn, J. A., Statistics for Exceptional Athletics Records. *Journal of the Royal Statistical Society*. Series C (Applied Statistics), **44**(4): 499-511, 1995.

[Ros00] Rose, F. D., Attree, E. A., Brooks, B. M., Parslow, D. M. and Penn, P. R., Training in virtual environments: transfer to real world tasks and equivalence to real task training. *Ergonomics* **43**(4): 494-511, 2000.

[Ros01] Rosen, J., Hannaford, B., Richards, C. G. and Sinanan, M. N., Markov Modeling of Minimally Invasive Surgery Based on Tool/Tissue Interaction and Force/Torque Signatures for Evaluating Surgical Skills. *IEEE Transactions on Biomedical Engineering* **48**(5): 579-591, 2001.

[Ros03] Rosen, J., Chang, L., Brown, J. D., Hannaford, B., Sinanan, M. N. and Satava, R., Minimally Invasive Surgery Task Decomposition – Etymology of Endoscopic Suturing. In: *Studies in Health Technology and Informatics - Medicine Meets Virtual Reality* **94**: 295-301, 2003.

[Sha49] Shannon, C. E. and Weaver, W., *The mathematical theory of communication*. Urbana: University of Illinois Press, 1949.

[She02] Sheridan, T. B., *Humans and Automation: System Design and Research Issues*. Wiley-Interscience, 2002.

[Sol08] Solis, J. and Takanishi, A. Enabling autonomous systems to perceptually detect human performance improvements and their applications. In: *Proceedings of the IEEE Conference on Automation Science and Engineering*, pp. 259-264, 2008.

[Sug85] Sugeno, M., Nishida, M., Fuzzy control to model car. *Fuzzy Sets and Systems* **16**: 103-113, 1985.

[Suz04] Suzuki, S. Tomomatsu, N. Harashima, F. and Furuta, K. Skill evaluation based on state-transition model for human adaptive mechatronics (HAM). In: *Proceedings of the International Conference on Industrial Electronics Society* **1**: 641-646, 2004.

[Suz05a] Suzuki, S., Kurihara, K. Furuta, K. Harashima, F. and Pan, Y., Variable Dynamic Assist Control on Haptic System for Human Adaptive Mechatronics. In: *Proceedings of the IEEE Conference on Decision and Control and the European Control Conference*, 2005.

[Suz05b] Suzuki S., Suzuki, Y., Baba, Y. and Furuta, K., Human skill and brain activity in machine operation. In: *Proceedings of the IEEE Annual Conference on Industrial Electronics Society*. pp. 1967-1972, 2005.

[Tay11] Taylor, F. W., *The Principles of Scientific Management*. Harper Bros., 1911.

[Ter06] Tervo, K., *Evaluation of tree stem feeding process performance using hidden Markov models*. Master's thesis, Helsinki University of Technology, Espoo, Finland, 2006.

[Ter07] Tervo, K. and Hölttä, V., Hidden Markov Models in Step Response Evaluation. In: *Proceedings of The 61st Meeting of the Society for Machinery Failure Prevention Technology*, 2007.

[Ter09] Tervo, K., Palmroth, L. and Putkonen, A., A hierarchical fuzzy inference method for skill evaluation of machine operators. In: *The 2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 2009.

[Ter10a] Tervo, K., Palmroth, L. and Koivo, H., Skill Evaluation of Human Operators in Partly Automated Mobile Working Machines. *IEEE Transactions on Automation Science and Engineering* **7**(1): 133-142, 2010.

[Ter10b] Tervo, K., *Human adaptive mechatronics methods for mobile working machines*. Doctoral dissertation, Department of Automation and Systems Technology, Helsinki University of Technology, Espoo, Finland, 2010.

[Tik08] Tikkanen, L., TimberLink as a tool for measuring the fuel consumption of a harvester. In: *The Nordic-Baltic Conference on Forest Operations*, 2008.

[Tur99] Turunen, E., *Mathematics Behind Fuzzy Logic*. Physica-Verlag, Heidelberg, 1999.

[Uus10] Uusitalo, J., *Introduction to Forest Operations and Technology*. JVP Forest Systems Oy, 2010.

[Yag88] Yager, R. R., On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on Systems, Man and Cybernetics* **18**(1): 183-190, 1988.

[Yan97] Yang, J., Xu, Y., and Chen, C.S., Human Action Learning via Hidden Markov Model. *IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans*. **27**(1): 34-44, 1997.

[Yin00] Ying, J., Kirubarajan, T., Pattipati, K. R. and Patterson-Hine, A. A hidden Markov model-based algorithm for fault diagnosis with partial and imperfect tests. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* **30**(4): 463-473, 2000.

[Yu04] Yu, W., Dubey, R. and Pernalete, N., Robotic therapy for persons with disabilities using hidden Markov model based skill learning. In: *Proceedings of the IEEE International Conference on Robotics and Automation*. **2**: 2074-2079, 2004.

[Vit67] Viterbi, A., Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, *IEEE Transactions on Information Theory*, **13**(2): 260-269, 1967.

[Vaa04] Väätäinen, K. and Ovaskainen, H., The importance of the harvester operator in CTL-harvesting operations. In: *NSR Conference on Forest Operations*, 2004.

[Vaa05] Väätäinen, K., Ovaskainen, H. Ranta, P. and Ala-Fossi, A. *Hakkuukoneen kuljettajan hiljaisen tiedon merkitys hakkuutulokseen työpistetasolla.* Finnish Forest Research Institute, Research Papers 937, 2005. (In Finnish)

[Wan03] Wang, K., *Intelligent condition monitoring and diagnosis systems: a computational intelligence approach*, IOS Press, 2003.

[Zad65] Zadeh, L. A., Fuzzy sets. *Information and Control* **8**(3): 338-353, 1965.

[Zad71] Zadeh, L. A., Similarity relations and fuzzy orderings. *Information Sciences* **3**(2): 177-200, 1971.

[Zad02] Zadeh, L. A., From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions. *International Journal of Applied Mathematics and Computer Science* **12**(3): 307-324, 2002.

[Zhu08] Zhu, C., Cheng, Q. and Sheng, W., Human intention recognition in smart assisted living systems using a hierarchical hidden Markov model. In: *Proceedings of the IEEE International Conference on Automation Science and Engineering*, pp. 253-258, 2008.

[Zim96] Zimmermann, H.-J., *Fuzzy Set Theory*, Kluwer Academic Publishers, Dordrecht, 1996.

Appendix A

Matlab code for calculating the sequence of HMM states in the "Morning" model

```
function hmm_states = hmm_example_morning
% HMM_EXAMPLE_MORNING Generates test data for task network model
% "Morning", estimates parameters of HMM and calculates the most
% likely path of states using viterbi algorithm and the observations
% of the test data.
%
% Outputs:
% hmm_states - path of states calculated using Viterbi algorithm
%
% Author: Lauri Palmroth, TUT/ASE 2010

N = 9; % number of (hidden) states, N
M = 20; % number of discrete observation symbols, M

% Generate test data
hmm_test_data = create_test_data;
Nmornings = length(hmm_test_data);

% Estimation of HMM parameters
[A,B,p] = estimate_hmm_parameters(hmm_test_data,N,M);

% Calculate the most likely path of states using viterbi algorithm
hmm_states = [];
for d = 1:Nmornings % calculate states for each morning
    tmp_states = hmm_viterbi_algorithm(hmm_test_data{d}.O,A,B,p,N);
    hmm_states = [hmm_states, tmp_states];
end

end % end of function
```

```
function hmm_test_data = create_test_data
% CREATE_TEST_DATA Creates test data for "Morning" -model
%
% Outputs:
% hmm_test_data - struct with fields
%   day - day of week
%   O - sequence of observations
%   Q - sequence of states
%
% Author: Lauri Palmroth, TUT/ASE 2010

% 1st sequence: "Monday"
hmm_test_data{1}.day = 'Monday';
hmm_test_data{1}.O = [9,4,13,17,19,20,17,15,17,19,20,18,16,17,17, ...
    15,19,1,20,16,15,17,19,20,19,1,20,16,17,17,18,18,18,18,14,10, ...
    9,5,10,10,10,11,11,9,9,4,9,2,3,2,1,3,2,1,1,3,12,12,10,5,9,9, ...
    10,10,6,8,8,8,10,7];
hmm_test_data{1}.Q = [1,1,1,1,2,2,1,1,1,2,2,1,1,1,1,1,2,2,2,1,1, ...
    1,2,2,2,2,2,1,1,3,3,3,3,3,3,3,4,4,5,5,5,5,5,5,5,5,5,6,6,6,6, ...
    6,6,6,6,6,7,7,7,7,9,9,9,9,8,8,8,8,8,8];
% 2nd sequence: "Tuesday"
hmm_test_data{2}.day = 'Tuesday';
hmm_test_data{2}.O = [9,4,17,18,18,18,18,14,10,10,9,5,10,10,10, ...
    11,11,9,4,9,2,3,2,1,3,10,12,12,12,12,5,10,9,9,10,10,10,10,10, ...
    11,11];
hmm_test_data{2}.Q = [1,1,3,3,3,3,3,3,3,3,4,4,5,5,5,5,5,5,5,5,6, ...
    6,6,6,6,7,7,7,7,7,7,9,9,9,9,9,9,9,9,9,9];
% 3rd sequence: "Wednesday"
hmm_test_data{3}.day = 'Wednesday';
hmm_test_data{3}.O = [9,9,4,13,19,20,17,19,16,15,19,1,20,9,4,18, ...
    18,18,18,14,10,10,9,5,10,10,10,11,11,9,4,9,2,3,2,1,3,2,1,3,2, ...
    3,5,9,4,9,9,2,3,2,1,3,2,3,10,12,12,10,5,10,9,10,10,10,10];
hmm_test_data{3}.Q = [1,1,1,1,2,2,1,2,1,1,2,2,2,1,1,3,3,3,3,3,3, ...
    3,4,4,5,5,5,5,5,5,5,5,6,6,6,6,6,6,6,6,6,5,5,5,5,5,6,6,6,6, ...
    6,6,6,7,7,7,7,7,9,9,9,9,9,9];
% 4th sequence: "Thursday"
hmm_test_data{4}.day = 'Thursday';
hmm_test_data{4}.O = [9,4,13,19,20,19,1,20,16,17,19,20,9,4,18,18, ...
    18,14,10,9,5,10,10,11,11,11,9,9,4,9,2,3,1,1,1,5,3,12,9,10,12, ...
    10,10,10,11,4,11];
hmm_test_data{4}.Q = [1,1,1,2,2,2,2,2,1,1,2,2,1,1,3,3,3,3,3,4,4, ...
    5,5,5,5,5,5,5,5,5,6,6,6,6,6,6,7,7,7,7,9,9,9,9,9,9];
% 5th sequence: "Friday"
hmm_test_data{5}.day = 'Friday';
hmm_test_data{5}.O = [9,13,4,19,20,19,16,15,19,20,1,13,4,13,9,9, ...
    11,10,9,9,3,2,3,12,11,10,11,11,12];
hmm_test_data{5}.Q = [1,1,1,2,2,2,1,1,2,2,2,1,1,1,1,5,5,5,5,5,6, ...
    6,6,9,9,9,9,9,9];

end % end of function
```

```
function [q] = hmm_viterbi_algorithm(O,A,B,p,N)
% HMM_VITERBI_ALGORITHM Calculates the most likely path of HMM states
% using the Viterbi algorithm
%
% Inputs:
% O - sequence of observations
% A - state transition probabilities, size: NxN
% B - observation probabilities, size: NxM
% p - prior probabilities of states
% N - number of discrete states
%
% Outputs:
% q - path of most likely states
%
% Author: Lauri Palmroth, TUT/ASE 2010

% Avoid numerical problems by using logarithmic scale in computations
logA = log(A);
logB = log(B);
logp = log(p);
negInf = -1e+020; % approximate of negative infinity

% Allocate space
T = length(O); % length of state sequence
q = zeros(1,T); % path of states
V = zeros(N,1); % incremental quantity

% Initialization step
for i = 1:N
    V(i) = logp(i) + logB(i,O(1));
end
VOld = V; % V at time t-1
phi = zeros(N,T); % state sequence back tracking array

% Loops
for t = 2:T % loop through sequence
    for j = 1:N % loop through next states
        bestValue = negInf; % best transition probability
        bestTransition = 0; % best, the most probable transition
        for i = 1:N % loop through previous states
            val = VOld(i) + logA(i,j);
            if val > bestValue
                bestValue = val;
                bestTransition = i;
            end
        end
        % update incremental quantity V
        V(j) = logB(j,O(t)) + bestValue;
        % update state sequence back tracking array
        phi(j,t) = bestTransition;
    end
    VOld = V;

end

% Termination step, decide the final state
[~, finalState] = max(V);
```

```
% Finally, trace back the path of states
q(T) = finalState;
for t = 1:(T-1)
    q(T-t) = phi(q(T-t+1),T-t+1);
end

end % end of function
```

```
function [A,B,p] = estimate_hmm_parameters(hmm_test_data,N,M)
% ESTIMATE_HMM_PARAMETERS Estimates parameters of a HMM
%
% Inputs:
% hmm_test_data - test data struct with fields
%   Q - sequence of states
%   O - sequence of observations
% N - number of (hidden) states
% M - number of discrete observation symbols
%
% Outputs:
% A - state transition probabilities, size: NxN
% B - observation probabilities, size: NxM
%
% Author: Lauri Palmroth, TUT/ASE 2010

% set pseudotransitions and pseudo-observations to avoid
% zero probabilities
A = ones(N);
B = ones(N,M);
p = ones(N,1);

% sum up state transitions and observations for all mornings
% of the training data
Nmornings = length(hmm_test_data)-1;
for d = 1:Nmornings
    tmp_O = hmm_test_data{d}.O; % sequence of observations
    tmp_Q = hmm_test_data{d}.Q; % sequence of (known) states
    [tmpA,tmpB,tmpp] = hmm_counter(tmp_O,tmp_Q,N,M);
    A = A + tmpA;
    B = B + tmpB;
    p = p + tmpp;
end

% normalize to give frequency estimate.
RowSum = sum(A,2);
A = A./repmat(RowSum,1,N);
RowSum = sum(B,2);
B = B./repmat(RowSum,1,M);
RowSum = sum(p);
p = p/RowSum;

end % end of function
```

```
function [A,B,p] = hmm_counter(O,Q,N,M)
% HMM_COUNTER Sums up state transitions and observations of
% input sequences
%
% Inputs:
% Q - sequence of states
% O - sequence of observations
% N - number of (hidden) states
% M - number of discrete observation symbols
%
% Outputs:
% A - number of state transitions, size: NxN
% B - number of observations, size: NxM
% p - first state of the sequence
%
% Author: Lauri Palmroth, TUT/ASE 2010

T = length(O); % length of state sequence
A = zeros(N);
B = zeros(N,M);
p = zeros(N,1);

% count up the state transitions from the state path
for t = 1:T-1
    A(Q(t),Q(t+1)) = A(Q(t),Q(t+1)) + 1;
end
% count up the observations for each state
for t = 1:T
    B(Q(t),O(t)) = B(Q(t),O(t)) + 1;
end
% the first state
p(Q(1)) = p(Q(1)) + 1;

end % end of function
```