Task Performance Evaluation for Industrial Robots

Billel Bouchemail*1, Abdelouahab Zaatri*2

*Welding and NDT Research Centre (CSC). BP 64 CHERAGA – ALGERIA 21000, Algeria
Bouchemail.billel@gmail.com

*Department of Mechanical Engineering LATA University of Constantine 25000, Algeria
azaatri@yahoo.com

Abstract——Integrated multi-modal Supervisory Control Systems (ISCS) are a new generation of complex and synergistic Human-Machine Interaction Systems (HMIS). This paper deals with multi-modal interaction and control applied to HRS. A task performance evaluation technique dedicated for multi-modal interaction and control is proposed. It enables comparison of task performance carried out by using different selection of control modes or by different operators. Objective and subjective performance measures are defined. Based on the analytical hierarchy process method (AHP) which takes into account qualitative and quantitative attributes and criteria.

Experimental results have been carried out and some preliminary results will be presented concerning industrial manipulators.

Keywords—Task performance evaluation, Robotics, human factors, Analytical Hierarchy Process, Adaptive supervisory control, Multi-modal interaction, Graphical-user interface.

I. INTRODUCTION

Under the pressure of various applications, advanced research in human-machine interaction systems is promoting a new generation of systems that can be named “Integrated Supervisory Control Systems or ISCS”. ISCS are complex, synergistic HMIS. They attempt to integrate multi-media techniques, modern simulation tools and high-level programming techniques. They provide various control modes such as tele-manipulation, vision, gesture, speech, etc. Such systems are meant to carry out complex tasks and missions in various environments. ISCS are present in many important industrial domains such as nuclear power plants, space missions, military operations, aircraft industry, undersea, tele-surgery and assistance to disable people, etc. An initial evaluation about the state of the art and the trends in this field can be inferred by surveying some ISCS that have been described in the literature, e.g. [1,2,3,4,5].

In HMIS, one main automation goal is the reduction of the human involvement in the task while increasing the system performance. In complex systems, task performance evaluation depends on numerous criteria and factors which may be quantitative, qualitative, or both. Some approaches [6,7,8] have attempted to provide means for task evaluation.

In ISCS, new problems emerge due to the possibility to carry out the same task by using various combinations of control modes. For instance, when the operator performs tasks by means of the ISCS under consideration, he/she has to select some control modes (vision, speech, gesture, etc.). Within each selected control mode, the operator may have also to allocate subtasks to the autonomous module, to the operator, or to both of them. This real-time process of action selection and dynamic task allocation is, in general, system, context and operator dependent. For each experiment, the composition, the order, the switching instants, the duration of the selected control modes, and the task allocation are all factors that influence substantially the task performance. The relationships between task performance and those factors are not easy to establish. One relevant question thus is how to evaluate task performance for different realizations of the same task achieved by different combinations of control modes of different operators.

A method for task performance evaluation and comparison of tasks is presented in this paper.

This paper deals with some aspects of a developed ISCS, which is characterized by multi-modal interaction and multi-level control. ISCS generates new problems due to the possibility to carry out the same task by using various combinations of control modes (vision, speech, etc.). One relevant question thus is how to evaluate and compare different realizations of the same task performed by using different combinations of control modes and/or by different operators. A technique for task performance evaluation is presented in this paper. Performance measures are defined mainly to express the task effectiveness and the workload sustained by the human operator. It is based on the Analytical Hierarchy Process method that takes into account and fuses both qualitative and quantitative attributes and criteria.

This paper is organized as follows. Section 2 presents a brief overview of the ISCS. Section 3 presents the implemented multimodal operator interface. Section 4 presents the method of task performance evaluation. Section 5 concludes the paper.
II. RESEARCH BACKGROUND

In a previous work, we have used the AHP technique to evaluate HR task performance for integrated multi-modal supervisory control systems. It has been applied to industrial manipulator, taking into account qualitative and quantitative attributes and criteria [9].

The analysis of the state of the art concerning the application of AHP technique to HRI, we can divide the application of the AHP method as an evaluation procedure or as an evaluation and decision making procedure.

In the field of graphical user interface design, reference [10] has proposed a model for the design method of the GUI for AV remote controller based on AHP technique. It has been used to reduce the human error. Seven evaluation criteria are used as alternatives and decision on four design strategies for the user interface; vision assistance, cognition assistance, operation assistance, and memorizing to obtain the most suitable interface for every user (Multi-interface strategies).

In terms of evaluation techniques, reference [11] presents a methodology based on the AHP which can be used to evaluate and rank different approaches to automating the functions and tasks planned for the NASA space station. The primary goal will be to optimize the human-machine mix of functions and skills at both the module and subsystem levels. Because of the large number of factors involved in the model, the overall problem is decomposed into four sub-problems individually focusing on human productivity, economics, design, and operations, respectively. Every sub-problem is divided in multiple levels of criteria’s. Special attention is given to the impact of advanced automation on human productivity, we note that the human productivity criterion is used as goal in the first sub-problem and is based on the human-machine interface as an important factor. And we note that in the operations problem, the acceptability has been proposed as criterion and based on the ease of use factor as a qualitative attribute to evaluate the operation.

Another example in reference [12] shows that a quantitative evaluation is performed by using the AHP approach on a case study system to make analytic comparison between human only and human-robot collaboration in cellular manufacturing system. Indeed, several researches show the possibility to integrate the AHP method with other techniques to evaluate or make a decision. In reference [13], a fusion of analytic hierarchy process and brain limbic system control strategy is suggested to control the mobile robot navigation in unknown environments. In this example of application, the optimal gap is used as goal based on three criteria’s to decide the optimal gap; the distance from the gap to the target, the gap width, and the amount of robot rotation to get the gap.

III. OVERVIEW OF THE ISCS

The physical system (see Figure 1) has already been described in [5]. It consists of a Kuka-361 industrial robot controlled by an interactive version of COMRADE [14]. COMRADE is a software package designed at K.U.Leuven to facilitate the development and tuning of compliant motion control. A force sensor and a CCD camera are mounted at the end-effector of the controlled system (the remote robot). The CCD camera is linked to a video-monitor to display the captured scenes to the operator and to a DSP frame grabber. The grabbed pictures are transmitted to a Silicon Graphics workstation, which hosts the image processing routines [15]. A human speech input system with a limited vocabulary is also implemented [16].

At the control site, the operator uses a 6 DOF hand controller, a vision module and a speech module to monitor the remote robot. Feedback information is ensured by direct vision if applicable, selected images sensed by the CCD camera and other sensor measurements such as force and position.

Figure 1 Overview of the experimental system

IV. INTEGRATED MULTI-MODAL OPERATOR INTERFACE

To access and to manage the different ISCS functionalities, an integrated multi-modal operator interface has been implemented. The operator can select the control modes he/she prefers or considers as suitable one for achieving a specific task. The available control modes range from tele-manipulation to practically autonomous control including traded, shared, superimposed, some interactive functions, Human Demonstration Programming (HDP) [15, 16], vision-based [5, 17] and speech-based controls. The operator interface serves for selection of and switching between control modes for simulation, tele-programming and monitoring. To keep, as much as possible, similar condition for the operator to supervise a task in the real world as well as in the virtual world, the same multi-modal operator interface is employed.

In the designed ISCS, flexible mode management and change have been considered. For instance, the described ISCS can manage sequentially the selected control modes. But, it can also be set to mixing modes for any controlled variables (joint or Cartesian) at any time. This possibility enables the operator to intervene on-line during the wait-state or during autonomous operations.

This ISCS is enhanced by some other functions that can assist the operator such as the possibility to automatically center the view of the camera on a relevant point designated by the operator anywhere in an image, focusing on the object of interest.
V. TASK PERFORMANCE EVALUATION WITH ISCS

A. Performance measures in complex dynamic systems

To evaluate and compare different experiments of a given task performed by the same or by different operators, attributes and criteria of performance are required. However, numerous criteria and measures can be defined. Examples of such criteria can be minimizing the completion time, minimizing the used energy, increasing the safety, maximizing the accuracy, minimizing the rate of errors and accidents, maximizing the ease of the operation, maximizing the human confidence on the system, etc.

In our study, we seek to compare two different achievements of the same task driven by the ISCS. From human factors viewpoint, three characteristics are important to evaluate a human-machine interaction system: the reaction time, expressing the speed of execution, the errors, expressing the quality of the expected result, and the human workload that can be objective and/or subjective and expressing the human involvement. Obviously, the aimed goal is to select the case that presents the shortest execution time, gives the best quality results and exhibits the least workload.

In this application, we will focus on two main measures: the speed and some aspects of the workload. The speed measure will be related to the task completion time (CT). Concerning the workload, which is considered in the general case as the demand placed upon the operator [18, 19], distinction can be made between physical and mental workload. While the physical workload can be measured, the mental workload is not directly measurable. The physical involvement of the operator will be related to the tele-operation time (TI) and to the number of interactions of the operator with the GUI (NI). The mental workload will be estimated according to the human judgment about the ease or the difficulty to perform the task. For each experiment, CT, TI and NI are automatically computed once the task is achieved while the mental workload EU is estimated according to the subjective appreciation of the operator.

B. Illustrative experiments and performance evaluation

To illustrate the issue of performance evaluation in ISCS, an application example is presented in relationship to the following set-up (Fig 2). Fig 2 shows the image of the working environment as seen by the camera. In this image, one can identify a circular object (named A) and square object (named B).

Example 1

To illustrate the issue of performance evaluation in ISCS, two application examples are presented in relationship to the following set-up (Figure 2). Figure 2 shows the image of the working environment as seen by the camera held by the Kuka-361 robot end-effector. In this image, one can identify a cylindrical object posed at the middle of the left side of the black box. This cylindrical object contains a hole in its centre, recognised as a black inner circle and called A. In the top of the right side of the big black box, one can recognise a cube on top of which there is a white paper. This white area is called B. The completion time is recorded with respect to the experiment number (Figure 3).

Example 2

This example concerns two performance of the same task which has been carried out by following different combinations of control modes. The first task performance consists of using vision-based control to go from an initial point to A, then to use tele-manipulation to move from A to B. The second task performance consists of moving from the initial position to A by using vision-based control then again using vision-based control to move from A to B.
This example is presented for illustrating the application of the technique used for comparing various achievements of the same given task. Within this technique, quantitative as well as qualitative factors and criteria may be merged together for estimating the task performance. The quantitative factors considered are the completion time CT, the tele-manipulation time TT, and the number of interactions NI. CT, TT and NI are automatically recorded at the end of each experiment. The qualitative and subjective factor of preference expressing the degree of Ease of Use (EU) felt by the operator when carrying out the two task achievements is estimated. Many other factors qualitative as well as quantitative can be added. For qualitative and subjective factors and criteria, a fuzzy quantification using a certain methodology that will be shown latter (see Table 2), is required. The results obtained for two task achievements are presented in Table 1. In this table, the EU has been quantified as can be seen in the last column.

The AHP model presented in fig. 4 is a hierarchy with a main goal (appropriate assistance) at the top of the model, followed by the alternatives as the solution candidates in the middle layer, and ended with criteria required to achieve the goal on the bottom of the model.

<table>
<thead>
<tr>
<th>TABLE 1.</th>
<th>VALUES OF THE DIFFERENT ATTRIBUTES (CRITERIA) FOR DIFFERENT ALTERNATIVES</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT(sec)</td>
<td>TT(sec)</td>
</tr>
<tr>
<td>Experiment 1</td>
<td>94</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>146</td>
</tr>
</tbody>
</table>

How to decide which alternative is best with regard to the results shown in table 2 in case of multi-criteria problems where qualitative and quantitative factors influence the system performance? The so-called Analytical Hierarchy Process can be a suitable tool for solving such problems. It has been adapted here for comparing different achievements of a given task and extracting the best alternative with respect to given criteria. This application exemplifies the use of the AHP method.

C. The Analytic Hierarchy Process (AHP)

The AHP method was originally developed by Saaty in 1980 [20]. It is a powerful and flexible decision making aid for complex, multi-criteria problems where both qualitative and quantitative factors have to be incorporated. It is useful in situations where the decision maker has to choose between different alternatives on the basis of how well they meet various attributes. It is presented and at the same time illustrated with the example at hand. This method involves the following main steps:

- Develop the hierarchical representation of the problem. At the top of the hierarchy is the overall objective. The alternatives are at the bottom. Between the top and the bottom levels are the relevant attributes of the decision problem.
- Establish the overall objective function. The overall objective function for each alternative k is estimated by the following expression:

\[ F_k = \sum_{i=1}^n w_i c'_i \]

where: \( w_i \) is the relative weight of the attribute \( i \), \( c'_i \) is the relative score of alternative \( k \) for an attribute \( i \), and \( n \) is the total number of attributes. The evaluation of an alternative requires thus the estimation of the relative weights \( w_i \) and the scores \( c'_i \).
- Estimate the weights \( w_i \) of the attributes. To this end, the comparison of the importance of each attribute with respect to the others is required. This importance is to be measured in an integer-valued 1-9 scale, with each number \( a_{ij} \) having the interpretation shown in Table 2 [21].

<table>
<thead>
<tr>
<th>TABLE 2.</th>
<th>TABLE OF PAIRWISE COMPARISON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>Interpretation</td>
</tr>
<tr>
<td>( a_{ij} )</td>
<td>1 objectives i and j are of equal importance</td>
</tr>
<tr>
<td></td>
<td>3 objective i is weakly more important than objective j</td>
</tr>
<tr>
<td></td>
<td>5 objective i is strongly more important than objective j</td>
</tr>
<tr>
<td></td>
<td>7 objective i is very strongly more important than objective j</td>
</tr>
<tr>
<td></td>
<td>9 objective i is absolutely more important than objective j</td>
</tr>
<tr>
<td></td>
<td>2,4,6,8 are intermediate values</td>
</tr>
</tbody>
</table>

This comparison is represented in a matrix called the pairwise comparison matrix, which is, for the example at hand, the matrix A. An element of A expresses the estimation of the importance of the attribute i with respect to the attribute j accordingly to the scale presented in Table 2.

<table>
<thead>
<tr>
<th>TABLE 3.</th>
<th>EVALUATION VALUES FOR THE FOUR ATTRIBUTE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CT</td>
</tr>
<tr>
<td>CT</td>
<td>1</td>
</tr>
<tr>
<td>TT</td>
<td>1/4</td>
</tr>
<tr>
<td>NI</td>
<td>1/5</td>
</tr>
<tr>
<td>EU</td>
<td>1/6</td>
</tr>
</tbody>
</table>

For instance, in the present case, \( a_{12} = 4 \), means simply that the first attribute CT is midway between weakly and strongly more important with respect to the second attribute TT. \( a_{14} =
3 means that TT is weakly more important than DD. Of course, reciprocity implies, \( a_{ii} = 1 \) and \( a_{ii} = 1 a_{i} \).

The relative weights of the attributes are computed according to the following formulas:

\[
w_j = \frac{\sum_{i=1}^{n} a_{ij}^{\text{norm}}}{n} \quad \text{where} \quad a_{ij}^{\text{norm}} = \frac{a_{ij}}{\sum_{j=1}^{n} a_{ij}} \quad \text{and} \quad \sum_{i=1}^{n} w_j = 1
\]

The numerical values of \( w_j \) are \( (w_1 = 0.60, w_2 = 0.20, w_3 = 0.12, w_4 = 0.08) \).

- Estimate the score \( c_{ik} \) of the alternative \( k \) for an attribute \( i \). To determine these scores, one constructs for each attribute a pairwise comparison matrix \( B' \) in which the rows and the columns are the possible alternatives. In case where there is a combination of quantiative and qualitative attributes such as with the current example, the quantitative measures as well as the qualitative judgement are first rescaled according to the scale given in Table 2. The values and judgements given in Table 1 have first to be fuzzified according to the scale provided in Table 2. In the current example, there are two alternatives (denoted alt1 and alt2), the comparison matrix noted \( B' \) is given by:

\[
B' = \begin{bmatrix}
\text{alt1} & \text{alt2} \\
\text{alt1} & 1/b_{12} \\
\text{alt2} & b_{21} \\
1 & 1
\end{bmatrix}
\]

An element \( b_{ki} \) compares the alternative \( k \) with the alternative \( i \) for the attribute \( j \). For each attribute \( i \), a matrix \( B' \) has to be built based on the data given in Table 2. The dimension of the \( B' \) matrix is of \( mxm \) where \( m \) is the total number of alternatives.

The scores \( c_{ik} \) are then estimated according to the following formula:

\[
c_{ik} = \frac{\sum_{j=1}^{m} b_{ki,\text{norm}}}{m} \quad \text{where} \quad b_{ki,\text{norm}} = \frac{b_{ki}}{\sum_{j=1}^{m} b_{kj}} \quad \text{and} \quad \sum_{k=1}^{m} c_{ik} = 1
\]

One simple evaluation of \( B' \) is:

- Consistency tests: before applying the AHP method, the matrixes \( A \) and which are obtained by human judgement are first submitted to a test of consistency. This test checks whether the judgements made to compare attributes with each other and objectives are consistent or not. A Consistency Index (CI) is defined by:

\[
CI = \frac{\lambda_{\text{max}} - n}{n-1}
\]

Where \( n \) is the number of attributes, and \( \lambda_{\text{max}} \) is the principal eigen value of the matrix \( A \). \( \lambda_{\text{max}} \) is given by:

\[
\lambda_{\text{max}} = \frac{1}{n} \sum_{i=1}^{n} (A w_i^T)
\]

To check the consistency, compute the ratio: \( CI \) in which \( RI \), called the Random Index, is given in Table 3 with respect to the dimension \( n \) of the matrix under consideration [20].
consistency for the $B$ matrixes since their dimension is $2$, so referring to the Table 3, $RI = 0$ for this case.

- Compare the alternatives and extract the best one

The evaluation of the overall objective $F_k$ for each alternative is estimated. The best alternative $F_{opt}$ is selected based on the optimisation criteria:

$$F_{opt} = \max_k F_k$$

Given the matrices $A$, $B^i$ and the optimisation criteria, the application of this technique is straightforward. The optimality criteria consist of minimising CT, TT, NI and maximising EU. Thus, $B^i$ should be transposed in order to use minimisation criteria for all the measures in the overall objective. The overall objective function adapted to this problem for the two alternatives are respectively $F_1 \equiv 0.4$ and $F_2 \equiv 0.6$.

The first experiment corresponding to the first task achievement has been selected by the AHP method as a best solution with respect to the defined measures and criteria. By personalising the coefficients of the comparison matrices $A$ and $B^i$, this technique becomes user-adaptive. Since, the ISCS uses the same GUI for simulation as well as well for monitoring in real world, this technique is thus applicable in virtual environments as the same manner as in real applications.

VI. CONCLUSION

This paper has presented a technique for improving the performance of an ISCS, which is characterised by a multi-modal interaction and control.

Because of their multi-modal structure, task performance evaluation of ISCS is a difficult issue. In effect, a same task can be performed by using different sequences of control modes. Furthermore, the task performance evaluation also depends on human factors, which are inherent to the human involvement into task execution. Despite the complexity of the ISCS and of the tasks to be performed, objective and subjective performance measures has been defined in relationship to task effectiveness and to the workload sustained by the human operator. A technique exploiting the AHP method, which incorporates both quantitative and qualitative attributes and criteria, has been used for the evaluation and comparison of task performance. Our experiments have proven its capability to evaluate and compare different task achievements.