TOWARD THE CONSTRUCTION OF AN EFFICIENT SET OF ROBOT ARM OPERATOR PERFORMANCE METRICS

Tyler M. Akagi, Robert E. Schlegel, Randa L. Shehab, Kirby Gilliland, Tamy L. Fry and Quintin Hughes
University of Oklahoma
Norman, OK 73019

ABSTRACT

As part of a larger project to identify and validate relevant quantitative measures of robot arm operator proficiency, fifteen metrics of arm maneuvering and hand controller performance were defined and measured for 3-DOF translational movement tasks. Twelve freshly trained operators provided performance data for seven target-acquisition task scenarios involving a variety of distance combinations along the X, Y, and Z axes. Metrics included indicators of task component times, distance traveled, inefficient (inverse) motion, maximum velocities, amount of multi-axis control, and input control onset times along the three axes. Pairwise correlations of all measures and scatter plots of variables yielding strong intercorrelations were examined to determine the potential underlying causes of the significant relationships. By identifying subsets of metrics with explainable co-dependencies, the overall metric set can be reduced to a limited number of key metrics that serve as effective discriminators of operator performance.

INTRODUCTION

Remote manipulator systems (RMS) aboard the International Space Station and the Space Shuttle assist in space operations such as material management, transportation, unberthing, retrieval and berthing of payloads, and construction. Each RMS is remotely controlled by an astronaut using two hand controllers (one translational and one rotational) and an interface panel while monitoring robot arm movement on video displays. Astronauts undergo several phases of training to acquire the skill set necessary for successful operation of the arms. The initial Generic Robotics Training (GRT) program uses the Basic Operational Robotics Instructional System (BORIS), a computer graphics robot simulator. Although the training program offers instructors objective criteria for evaluating trainee progress, there are currently no quantitative operator performance data generated or recorded to aid instructors in their evaluations.

The analysis reported here is an element of a larger project to identify, validate, and present relevant quantitative measures of RMS operator proficiency in a format that enables instructors, trainees, and trained operators to appraise current skill levels and to determine candidate areas for additional training or practice. The project will likewise identify metrics that can serve as in-flight readiness-to-perform indicators prior to critical operational tasks. Category 1 skills judged most important by RMS operators and instructors are arm maneuvers, hand controller techniques, application of

frames, and situation awareness. The initial focus of the project has been on robot arm maneuvers and hand controller techniques. The ability to make smooth hand controller movements and provide multi-axis control was evaluated by developing theoretical models of control input ramping and multi-axis commanding.

For each skill, more than twenty-five potential metrics were identified for consideration. For each metric, an operational definition was constructed and mathematically formulated from the raw data extracted from the simulator. The current analysis was conducted to examine the interrelationships among fifteen of the metrics, including several metrics developed by Fry et al. (2003) and three new metrics constructed for this analysis. The fifteen metrics included indicators of task component durations, distance traveled, inefficient (inverse) motion, maximum velocities, amount of multi-axis control, and input control onset times (lag) along the three axes. The goal of the analysis was to identify subsets of metrics with explainable co-dependencies, and thus reduce the overall metric set to a limited number of measures serving as effective discriminators of operator performance.

METHOD

The methodology consisted of operator training, data collection, metric computation, correlation analysis, and graphical data analysis. See Bray et al. (2003) for more details on RMS/BORIS training.

Participants

Twelve University of Oklahoma students ranging in age from 18 to 25 years were paid to participate in the study. None of the participants had previous robot arm operator experience (actual or simulated).

Experimental Tasks

Participants performed seven 3-D target acquisition tasks on the BORIS simulator using the translational hand controller to align the robot end effector camera crosshairs within the target boundaries. Each task was presented to the participant as a set of X, Y, Z distances to move in either the positive or negative direction. Distance magnitudes were either 100, 200, or 300-cm within the simulation environment, yielding tasks with movement distance proportions of 1:1:3, 1:2:3, or 3:3:3 along the three axes. The participants were asked to make smooth inputs and to use multi-axis control. The order of tasks was counterbalanced across participants.

Equipment

The BORIS workstation consists of two video monitors and two hand controllers. The left monitor presents several soft control panels and digital displays of robot arm position and joint configuration. The right monitor provides views of the robot arm within the simulation room. The translational hand controller, operated by the left hand, directs motion along the X, Y, and Z axes. The rotational hand controller directs end effector orientation along the pitch, yaw, and roll axes. For data collection, the participant used only the translational hand controller and the right monitor with preset fixed camera views.

Experimental Design

A repeated-measures design was used with task characteristics, participants, and replications serving as the main factors. Task characteristics were examined in two ways: as seven unique tasks and grouped by movement distance proportion as shown in Table 1.

Table 1. Task Characteristics.

Movement Proportion (MP)	X, Y, Z Movement
1:1:3	100cm, 100cm, 300cm
	100cm, -300cm, -100cm
	300cm, 100cm, -100cm
1:2:3	100cm, -200cm, -300cm
	300cm, -100cm, -200cm
	200cm, 300cm, 100cm
3:3:3	300cm, 300cm, 300cm

Table 2 provides definitions of the operator performance metrics used in the reported analysis.

Table 2. Operator Performance Metrics.

Measure	Calculation/Description
Completion time (T_{comp})	Total elapsed task time.
Correction time (T_{corr})	Time to task completion from the moment each of the three axis inputs has been reduced to zero commanded velocity at least once.
Distance traveled (*Dist*)	Length of the path followed by the end effector during the task.
Degree of inverse motion (D_n)	Sum of the distances traveled along all axes in a direction contrary to that called for by the specific task.
Lag (L_i) L_X, L_Y, L_Z L_2, L_3 ($L_1 = 0$) L_{total}	Time from initial movement to the initial input of a command along a given axis. Where i corresponds to either the axis (*X, Y, Z*), the order (*1, 2, 3*) in which motion along that axis was initiated, or the summation of lags.
Maximum component velocity (VC_{max})	The maximum single-axis commanded velocity observed across the *X, Y,* and *Z* axes.
Maximum velocity (V_{max})	The maximum vector-combined commanded velocity (*V*) during the duration of a given participant-task, where: $$V = \sqrt{VC_X{}^2 + VC_Y{}^2 + VC_Z{}^2}.$$
Percentage of biaxial input (*%Bi*)	Percentage of time during which two axes are commanded simultaneously.
Percentage of triaxial input (*%Tri*)	Percentage of time during which three axes are commanded simultaneously.
Percentage of multi-axis input (*%MA*)	Percentage of time during which more than one axis is commanded simultaneously.

Procedure

Participants were trained through the first three lessons (two hours each) of GRT. Lesson 1 was taught to introduce the participants to robotics terminology using figures and 3-D models. Lesson 2 introduced the participants to the simulation itself and allowed them to use the hand controllers to operate the robot arm. Lesson 3 provided instruction in the use of various display views. Participants also received two hours of hands-on

simulation practice time to become better practiced with operating BORIS. Participants completed two test sessions at different times. Each session included the seven 3-D tasks along with other 1-D and 2-D tasks.

RESULTS

The collected dataset was used to compute pairwise correlations of all measures. Of the 210 correlations of the fifteen measures, 25 correlations exceeded 0.4. Scatter plots of these variable pairs were examined to determine the potential underlying causes of the significant relationships. The goal of the scatter plot analysis was to determine which of the independent variables (movement proportion, task, participant, replication) might contribute to the strong correlations. To accomplish this goal, separate plots were constructed for each independent variable, with individual points on the plots coded to represent the variable level. Table 3 summarizes those factors for which a distinctive pattern was observed.

Table 3. Scatter Plot Analysis (x = distinctive pattern).

Var 1	Var 2	r	ID	MP	Task	Rep
T_{comp}	T_{corr}	0.835				
T_{comp}	$Dist$	0.430		x		
T_{comp}	D_n	0.459	x	x		x
T_{comp}	$\%Tri$	-0.461		x		
T_{comp}	$\%MA$	-0.506				
T_{comp}	VC_{max}	-0.403	x			x
T_{corr}	$Dist$	0.413		x		
T_{corr}	D_n	0.503	x	x		x
T_{corr}	$\%Tri$	-0.507		x		
T_{corr}	$\%MA$	-0.670				
$Dist$	D_n	0.559	x	x		x
$Dist$	V_{max}	0.426	x	x		
L_X	L_2	0.466	x	x	x	x
L_Y	L_2	0.808	x		x	
L_Y	L_{total}	0.442			x	
L_Z	L_3	0.937			x	
L_Z	L_{total}	0.902			x	
L_2	L_3	0.440	x			
L_2	L_{total}	0.577	x			
L_3	L_{total}	0.987				
VC_{max}	V_{max}	0.858	x	x		
V_{max}	$\%Bi$	-0.413	x	x		
$\%Bi$	$\%Tri$	-0.446		x		
$\%Bi$	$\%MA$	-0.409		x		
$\%Tri$	$\%MA$	0.635		x		

MP=Movement Proportion, ID=Participant, Rep=Replication

CONCLUSIONS

The question addressed by this analysis was the extent to which a significant relationship between two metrics is driven by one of the individual variables. Figures 1 and 2 illustrate the graphical analysis approach taken toward this end.

Evaluation of all metrics must be conducted with the knowledge that many metrics have inherent relationships based on their definitions: one metric may form a substantial component of another metric. In the case of the metrics examined in this study, T_{corr} and D_n are direct components of T_{comp} and $Dist$, respectively. Neither T_{corr} nor D_n can be non-zero until the velocity along one or more of the input axes is reduced to zero or reversed. Furthermore, because much of the input recorded before corrections (T_{corr}) are initiated is multi-axial, and because D_n typically results from uniaxial input, the existence of either tends to decrease the magnitude of $\%MA$.

With knowledge of these inherent relationships, several interesting conclusions can be drawn. Most importantly, 11 of the 25 significant correlations were impacted in some fashion by differences in participant skill or performance. In these cases, either performance measure of the pair could be used to discriminate participant performance (see Figure 2). Other examples besides correction time and degree of inverse motion are the numerous correlations involving lag times and maximum velocities.
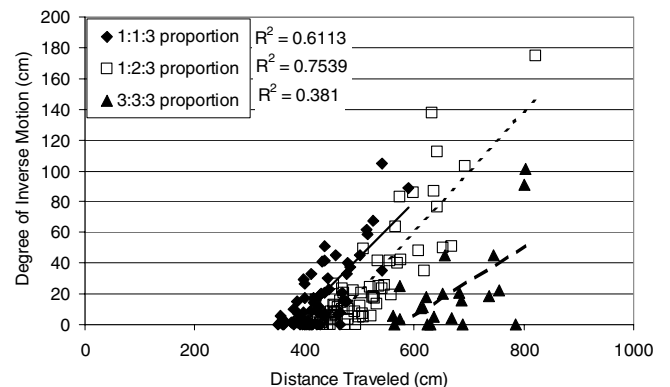

Figure 1. Distance Traveled vs. Degree of Inverse Motion Grouped by Movement Proportion.
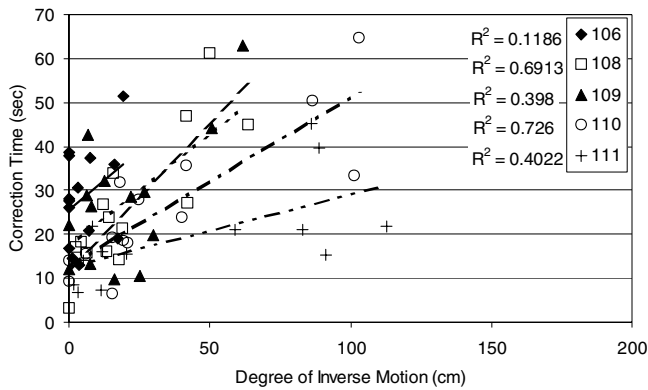
Figure 2. Correction Time vs. Degree of Inverse Motion Grouped by Participant.

Another relationship observed through scatter plot analysis was the effect of the minimum value of *Dist* associated with each of the movement proportions on the relationship among other metrics. Observable patterns as a function of movement proportion were found for 14 of the 25 significant correlations. The plots of T_{comp} *vs. Dist* and *Dist vs. %MA* grouped by movement proportion reveal that the minimum *Dist* has a marked effect on the strength of these relationships. Figure 3 illustrates the necessity of isolating data from tasks with different movement proportions in order to observe the true relationship between *Dist* and *%MA*. Within each task ratio the linearity and associated correlation of a trend line is fairly high, and between task ratios these trend lines exhibit fairly consistent patterns. The magnitude of these relationships is greatly diminished, however, when data are combined across all movement proportions. The R^2 coefficient drops to an insignificant 0.0908.
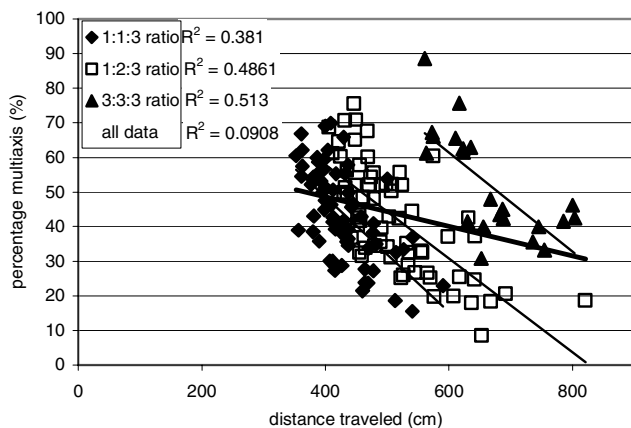

Figure 3. Distance Traveled vs. Percentage Multi-Axis Grouped by Movement Proportion.

Results indicate that beyond the effects of minimum *Dist*, task design plays a role in five of the 25 significant

correlations, all five involving the lag metrics. Figure 4 illustrates the effect of task on the relationship between L_2 and L_x for tasks of MP 1:2:3. As the magnitude of the *x*-component decreases, the correlation between L_2 and L_x increases. For these particular lag metrics to be useful proficiency measures, it may be necessary to restrict their use to intra-task performance evaluations.
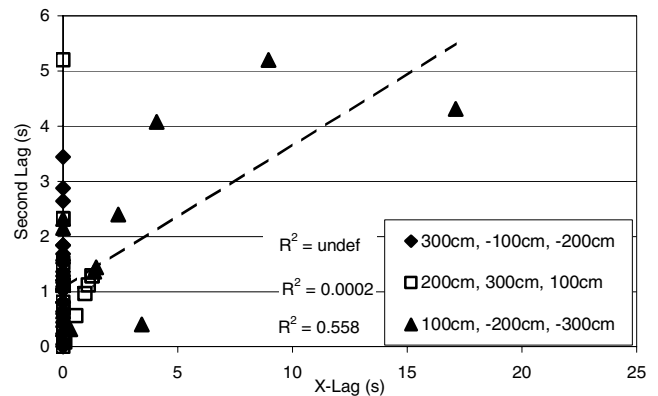

Figure 4. Second Lag vs. X-Lag Grouped by Task.

The final relationship observed through scatter plot analysis was the effect of task replication on the relationships between metrics. For five of the 25 significant correlations, a distinct pattern emerged distinguishing performance between the first and second test sessions. Figure 5 illustrates that both D_n and T_{corr} generally decreased in overall magnitude and variability from replication 1 to replication 2. This result indicates that further testing may be required to account for learning effects that may significantly impact the data; however, the result also provides evidence toward validation of the metrics as measures of operator performance.
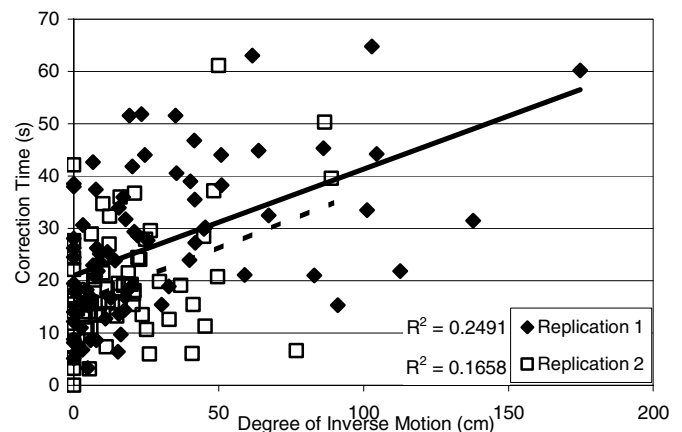

Figure 5. Correction Time vs. Degree of Inverse Motion Grouped by Replication.

In summary, the directed plotting of results permitted the identification of critical aspects of the task scenarios used to evaluate performance. In particular, the significance of task and movement proportion in determining the true correlations involving several metrics was highlighted. These results have significant impact on the design of the task scenarios used to assess differences between operators or throughout training. In addition, the large number of correlations driven by operator differences, as well as the smaller number of correlations impacted by replication, provide encouragement for the successful validation of these metrics as reliable quantitative measures of operator performance. Taken as a whole, the results of scatter plot analysis indicate a particular subset of metrics whose relationships are decidedly co-dependent and robust to changes in experimental design. After the elimination of lag metrics most likely affected by task design, this subset includes: T_{comp}, T_{corr}, *Dist*, D_n, VC_{max}, and *%MA*. Metrics within this subset exhibited consistent correlations that were most descriptive of the observed variable-metric relationships. Such consistency indicates the possibility that the overall metric set may be reduced and still form the basis for effective objective performance evaluation. Further analyses of the relationships within this subset are thus warranted.

Future research should focus on continued validation of the reduced metric subset. To accomplish this goal, this protocol has been administered to a set of seven expert RMS operators consisting of NASA astronauts and trainers. The data collected from these operator performances will be analyzed to allow determination of the efficacy of the metric subset in quantifying skill levels.

ACKNOWLEDGMENTS

REFERENCES

Bray, T.M., Shehab, R.L, Schlegel, R.E., Civan, A., & Walker, D.E. (2003). Effectiveness of 2-D views for 6-D robotics simulation maneuvers. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting (CD-ROM).* Santa Monica, CA: Human Factors and Ergonomics Society.

Fry, T.L., Schlegel, R.E., Shehab, R.L., & Gilliland, K. (2003). Developing metrics to evaluate NASA RMS training performance. In *Proceedings of the IIE Annual Conference*, Portland, OR.