

An execution time prediction model for crew information processing in new special vehicles

Sinan Guo^{1,a}, Shuang Liu^{*2,b}, Chaoran Liang^{3,c}, Fang Xie^{4,d}, Sijuan Zheng^{5,e}

¹ School of Aeronautic Science and Engineering Beihang University, Beijing, China

² School of Aeronautic Science and Engineering Beihang University, Beijing, China

³ School of Aeronautic Science and Engineering Beihang University, Beijing, China

⁴ China North Industries Group Corporation Limited China North Vehicle Research Institute, Beijing, China

⁵ China North Industries Group Corporation Limited China North Vehicle Research Institute, Beijing, China

Abstract. Task execution time prediction modeling is of great significance to the safety of special vehicle crew and maintenance of overall combat effectiveness. In our study, a task execution time prediction model for crew information processing task is proposed considering the characteristics of information processing task, as well as the interaction mode of man-machine system of special vehicles. In addition, a model validation experiment was conducted adopting 20 subjects facing to two kinds of representative tasks, and both of them have two levels of complexities. The result shows that there is a highly positive correlation ($r = 0.999$, $p = 0.001$) between the model prediction and the experimental results of four tasks, which indicates that the model has a good applicability. Based on the rationality of model validation, the application research of the model was conducted. The result shows that for the sub-tasks with more objects and visual searching difficulty, the speech interaction mode can reduce the operation duration to a great extent; for simple sub-tasks with less operation steps and strong coherence, the touch interaction mode has certain advantages.

1 Introduction

Recently, with the development and application of information and computer technology, the informatization level of armored vehicle cabin has been continuously improved. Based on this new man-machine relationship, building the human performance model is of great importance for evaluating a human-machine system.

The goals, operators, methods, and selection rules (GOMS) model is an approach used to analyze the human-machine system interaction during the human-machine interface prototype period, which was firstly introduced by Card et al [1]. It is essentially a systematic description of how to estimate the mental and motor execution time to complete a task. There are several variants in the GOMS family [2], including Keystroke-Level Model (KLM) [3-6], Card, Moran, and Newell GOMS (CMN-GOMS),

Natural GOMS Language (NGOMSL) [7-9] as well as Critical-Path Method GOMS (CPM-GOMS) [10-12]. As a basic version of GOMS model, KLM is employed to predict the time requirements by expert users when completing a task through man-machine interaction system (error-free), and it has been widely used for the advantages of easy operation and fast analysis.

However, the original KLM model was introduced for PC interface interaction, and does not have good applicability for the human-machine interaction of new special vehicle interface. Therefore, based on the information processing task and interactive mode of human machine system in new special vehicles, the present study put forward a new modified KLM-GOMS model to predict the execution time of new special vehicle tasks. In addition, the modified model was validated through ergonomics experiment. Based on it, an application analysis of the model was carried out for comparing two different interactive modes.

The description and typical time of the operators in original KLM

Operator	Description	Typical time(s)
K	Keystroke (keyboard, mouse) or button press Average skilled user	0.12~1.2 0.2
P	use a mouse or trackball to point at a target	1.1
D	Draw n straight line segments with a total length of l	$T_D = 0.8 * n_D + 0.16 * l_D$
H	Home hands on the keyboard or other device	0.4
M	Mental preparation for a physical operator	1.35
R	System response	t

^azy1805305@buaa.edu.cn ^c13121252759@163.com ^dchristie_xie@163.com ^eclubjully@sina.cn

^b*Corresponding author: liushuangbh@163.com

*System response time depends on system itself

Table1. Mental preparation orientation rules

Rule0: Insert M in front of all Ks that are not part of an argument string. Place M in front of all Ps that select commands.
Rule1: When an operator following an M is fully anticipated in the operator just prior to M, remove the M.
Rule2: When a string of MKs belongs to a cognitive unit, delete all Ms except the first.
Rule3: When a K is a redundant terminator, delete the M prior to the K.
Rule4: When a K terminates a constant string, delete the M prior to the K; however, do not delete the M when the K terminates a variable string.

2 An overview of original KLM

2.1 Steps

The original KLM model consists of the following steps: (1) Select representative task scenarios; (2) Decompose task to the level that can list the keystroke level action units; (3) List all the action units, named **Operator** in the model, involved in the representative tasks. (4) Insert a mental operator when user must need to stop and think; (5) Find the standard execution time for each operator; (6) Add up the execution time for the listed operators (7) The total is the task execution time.

2.2 Description

The descriptions and typical time of operators are shown in Table 1.

2.3 Mental preparation placement rules

Mental preparation needs to be inserted into the corresponding operators at the time point when the user is thinking, and the insertion of mental activity is governed by certain cognitive behavioral assumptions, as shown in Table 2.

3 Modified KLM model

3.1 Limitation of original KLM model

The original model had the following limitations when applying it to the new special vehicles.

(a) The description of some operators and their typical time need to be updated for the interaction of new special vehicles.

(b) The original model lacks some operators which are contained in the new special vehicle tasks, such as voice interaction and verbal report.

(c) As a kind of sub-task, the information search operation cannot be covered by the operator M.

(d) The crew must complete a series of task operations

according to the task instruction system. The process of reading and understanding the instructions result in time expenditure. However, the original model does not introduce this type of behaviors.

3.2 Extension of operators

Considering the above problems, our study carried out an extension on the definitions of operators and their typical time.

(a) Redefinition of original operators

- **K_S**: Finger with touch screen to complete pressing and lifting sub-operation units instantaneously. The average time of this operator was 0.2s obtained by Sethawong et al [13].
- **K_H**: Press a button for starting voice interaction, including two units: a button press and waiting for system response.
- **P**: Use fingers to point at a target.
- **H**: Home hands from the screen to the initial position
- **R**: System response. In Table III, several requirements for the maximum response time of the system under normal working conditions of military equipment are listed, referring to GJB2873-97.

Table2. Maximum response time of military system in GJB2873-97

System response type	Response time definition	maximum response time (s)
Keystroke response	From key to reaction	0.1
Keystroke echo	From key to character appear	0.2
Change pages	From issuing the command to the first few rows appearing	1

(b) New operators

- **S**: Enter voice commands or report task status. Although there are individual differences in speaking speed, the normal speaking speed of ordinary Chinese is about 100~200 words per minute. In the

current study, the average speaking speed is set as 150 words/minute, and thus the operator time can be expressed as (1)

$$T_s = \text{num} * 60 / 150 \quad (1)$$

In (1), Where num is represents the number of spoken words.

- W: Search targets in the designated area with selective attention. It is assumed that the target exists, and search operation will be terminated when the target is determined. The typical tasks include: locating specific targets on the map, obtaining key information in the text box and other subtasks. As in (2), its typical time is identified by serial self-terminating search (SSTS) [14].

$$T = a_p + bN / 2 \quad (2)$$

In (2), b represents the time to detect a non-target and determine that it is not the target, a_p represents the reaction time when the target is found, N represents the number of objects in the interface, $bN/2$ represents the target which will be found in the average level of searching time. In order to simplify the calculation, a_p and b are drawn as 1s to present average reaction time.

- U: Read and understand text instructions, get specific operation targets and instructions. As the average reading comprehension speed for Chinese is 300 words/minute referring to the research conclusion obtained by Liao et al [15], the typical time for U can be expressed as (3).

$$T_U = \text{num} * 60 / 300 \quad (3)$$

In the actual tasks, the prompt text of each subtask is about 15 words. Therefore, the typical time can be simplified as 3s.

3.3 Typical task analysis

Taking the typical task of armored vehicle commander as an example, the task analysis is carried out in the two stages of “task decomposition” and “strike planning”. According to the field investigation and the interview with the armored vehicle crew, the operational steps of the actual subtasks are different in the number of steps required due to the different mission situations. Therefore, two levels of complexity are set to verify the applicability of the model. The specific number of subtasks for two different levels of complexity is shown in Table 4.

4 Model Validation

An experiment was conducted to verify the modified model.

Table3. Design of task complexity level

Task stage	Level	Number of subtasks
Task decomposition	High	11
	Low	8
Strike planning	High	10
	Low	6

4.1 Subject

There are 20 subjects participating in the experiment, all of them are adult males, aged between 21 and 43 years ($M=30.55$ years, $SD=7.37$ years), all engaged in armored equipment related fields. All subjects were right-handed, no color blindness and color weakness, normal vision or corrected vision, and normal hearing.

4.2 Apparatus

This experiment is based on a new type of armored vehicle simulation platform, which consists of touch screen displays, control components, simulator hosts and servers etc. It can simulate a variety of armored vehicle combat tasks with high fidelity and record accurate time point of specific operations such as a single touch input or voice input.

4.3 Procedure

The subjects were required to complete four types of trails, including "task decomposition" with the high and low complexities, as well as "strike planning" with the high and low complexities. The specific task sequence was carried out in the order of the Latin square design to eliminate the learning and fatigue effects. Before a formal experiment, all the subjects were required to use the system proficiently through the task manual and experimenters' guidance.

4.4 Results

Firstly, the experimental data was filtered and integrated by “Pandas” and “Xlwings” modules of python3.0, and then SPSS Statistics 23.0 was used to conduct further statistical analysis. The descriptive statistics results of the four types of tasks are shown in Table V.

Since all the Coefficient of Variation are higher than 15% for the four types of tasks, there probably exists several discrete or extreme values in each group of data. Results of discrete or extreme data location are shown in Figure 1 and those data was deleted in the follow-up analysis.

4.5 Result analysis

After preprocessing, the average experimental data and the model prediction data are shown in Table 6. The results of correlation analysis show that there is a significant positive correlation between predicted time of the model and actual completion time for the four types of tasks ($r = 0.999$, $p = 0.001$). In addition, in Fig. 2, the theoretical prediction results are consistent with the experimental results in the changing trends, which validates the availability of the model proposed in our study. However, when the task complexity increases, the deviation between the model prediction and the experiment result tends to increase.

Table4. Descriptive statistics results of the four types of tasks

	Numbers	Mean(s)	Std. Dev.(s)	C.V	Min(s)	Max(s)
Low complexity-Task decomposition	20	65.45	41.565	63.51%	31	180
Low complexity-strike planning	20	51.30	45.624	88.94%	24	236
High complexity-Task decomposition	20	109.35	29.918	27.36%	72	172
High complexity-strike planning	20	144.55	70.657	48.88%	46	327

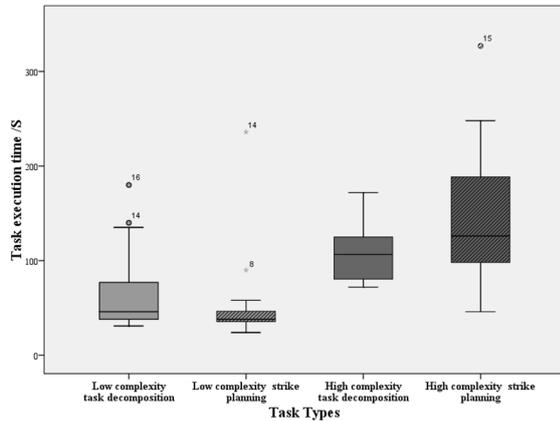


Fig1. Results of discrete or extreme data with subjects' number

Table5. Comparison of experimental average and model prediction time under different conditions

Task type	Model prediction (s)	Experiential results(s)
Low complexity-Task decomposition	51.05	54.94
Low complexity-strike planning	39.8	38.89
High complexity-Task decomposition	80.2	109.35
High complexity-strike planning	95.1	134.95

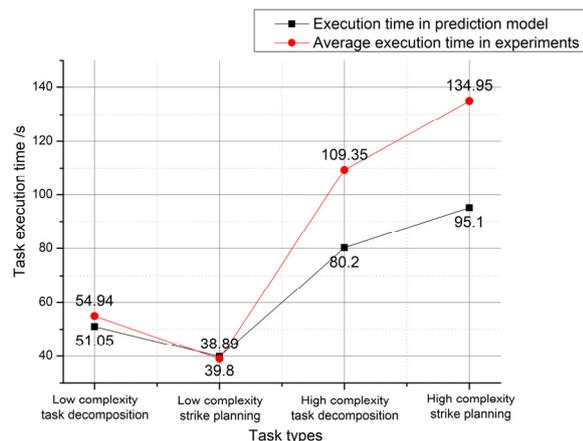


Fig2. Comparison of model prediction results and experimental

Table6. Results of task analysis under different interaction modes

NO.	Touch interaction	Voice interaction
1	Move the hand to the initial position -H	Read task input and understand its requirements-U
2	Read task input and understand its requirements -U	Mental preparation, determine the speech input paradigm-M
3	Mental preparation, search the target location indicated by the task -M	Press the voice input button to enter the voice input state-K_H

results

5 Application analysis

This section applies the model proposed above to predict the time required for different human-machine interaction modes to complete the same task. According to the predicted results, the two human-computer interaction modes, touch interaction and voice interaction, are compared and analyzed from the performance perspective of task completion time, thus it can provide the theoretical analysis basis for the further optimization of the task flow of follow-up special vehicles.

5.1 Task Analysis of touch and voice interaction modes

Through the field research and the armored vehicle crew interview, a task analysis of the "path planning" stage in touch and voice interaction mode was carried out. Taking the subgoal of "setting the starting point of the path" as an example, results of task analysis under different interaction modes are shown in Table 7.

5.2 Comparative analysis of touch and voice interaction model

The prediction results of subtask and total time for "path planning" task with two interaction modes are shown in Table 8, the total time gain percentage of using voice interaction is 25.2%.

The time gain percentage of each subtask is shown in Figure 3, the two subtasks of "Set the path start point" and "Set the path end point" both show a higher time gain percentage which have the characteristics of more goals and greater difficulty in visual search. On the other hand, the subtask of "Generate routing plan" shows a negative percentage using voice, which has simple and fewer steps, as well as strong operational coherence.

4	Point your finger to the "start point" button -P	According to the speech input paradigm, input the starting point target (about 4 words)-S
5	Touch and click the "starting point" menu (touch button) -K_S	System response time, time from key to response (maximum acceptable time)-R
6	Touch response time, time from keystroke to response (maximum acceptable time)-R	
7	Search the location of starting point target in the map (about 10 Suspicious target)-W	
8	Point to the target location -P	
9	Touch target (touch button)-K_S	
10	Touch response time, time from keystroke to response (maximum acceptable time)-R	

Table7. Results of each subtask and total task execution time prediction

Name of subtasks	Touch interaction (s)	Voice interaction (s)	Time gain percentage (%)
Choose an equipment	9.15	8.35	8.74
Set the path start point	13.55	6.35	53.14
Set the path end point	13.55	7.15	47.23
Generate routing plan	6.25	6.45	-3.20
Save the plan	6.15	5.55	9.76
Issue plan	6.25	5.65	9.60
Total task	61.05	45.65	25.2

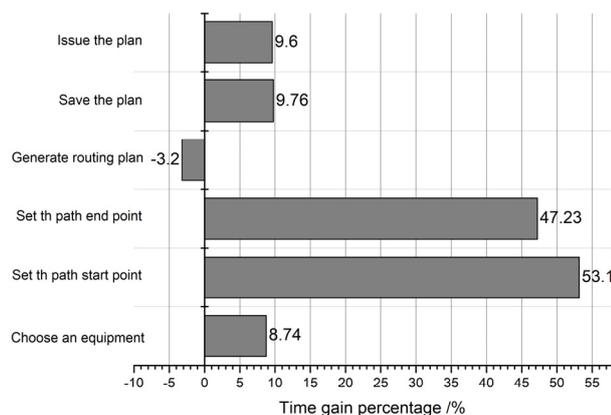


Fig3. Time gain percentage of each subtask

6 Conclusions

Based on KLM-GOMS performance modeling method, the present study proposed a time prediction model of information processing performance for armored vehicle occupants. Subsequently, an ergonomic experiment based on a high-fidelity armored vehicle simulation platform was conducted, and the model was verified using the experimental performance data. Finally, according to the application of the proposed model, the execution time required for different human-machine interaction modes when completing the same task is analyzed and the

conclusions are obtained.

(1) Firstly, the model built in our study shows a high degree of consistency between the prediction results and the experimental results for different types of tasks as well as the same type tasks with different complexities, therefore the model has good applicability.

(2) Secondly, in the follow-up model application research, voice interaction can significantly reduce the operation time for subtasks with more objects and complicated visual searching; while for subtasks with less operation steps, simple and strong coherence, touch interaction has certain advantages.

There are still some limitations in the current study. The model doesn't consider the situation of mis-operation during the operation process, while in the actual operation mis-operation is inevitable. Furthermore, the typical time of some operators are set as constant, however, due to individual differences, the distribution tends to be asymmetric in reality. In view of the above problems, further work will be performed for optimizing the model.

References

1. S. K. Card, T. P. Moran, and A. Newell, *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, USA, 1983.
2. B. E. John and D. E. Kieras, "The GOMS family of userinterface analysis techniques: Comparison and contrast," *ACM Transactions on Computer-Human Interaction*, vol. 3, no. 4, pp.320–351, 1996.
3. W. Jorritsma, P.-J. Haga, F. Cnossen, R. A. Dierckx, M. Oudkerk, and P. M. van Ooijen, "Predicting human performance differences on multiple interface alternatives: KLM, GOMS and CogTool are unreliable," *Procedia Manufacturing*, vol. 3, pp. 3725-3731, 2015.
4. A.D. Rice and J. W. Lartigue, "Touch-level model (TLM) evolving KLM-GOMS for touchscreen and mobile devices," in *Proceedings of the 2014 ACM Southeast regional conference*, 2014, pp. 1-6.R. Nicole, "Title of paper with only first word capitalized," *J. Name Stand. Abbrev.*, in press.
5. K. El Batran and M. D. Dunlop, "Enhancing KLM

- (keystroke-level model) to fit touch screen mobile devices," in Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services, 2014, pp. 283-286.
6. S. K. Card, T. P. Moran, and A. Newell, "The keystroke-level model for user performance time with interactive systems," *Communications of the ACM*, vol. 23, no. 7, pp. 396-410, 1980.
 7. N. Parameshwara, J. H. Kim, W. Guo, and K. S. Pasupathy, "NGOMSL simulation model in an emergency department," in 2016 Winter Simulation Conference (WSC), 2016: IEEE, pp. 1938-1949.
 8. S. Lee and R. J. Koubek, "Communicative NGOMSL: Development of an evaluation method for a text-based communication system," *Int J Hum-Comput Int*, vol. 27, no. 11, pp. 1013-1036, 2011.
 9. D. Kieras, "A guide to GOMS model usability evaluation using NGOMSL," in Handbook of human-computer interaction: Elsevier, 1997, pp. 733-766.
 10. W. D. Gray, B. E. John, and M. E. Atwood, "The precis of Project Ernestine or an overview of a validation of GOMS," in Proceedings of the SIGCHI conference on Human factors in computing systems, 1992, pp. 307-312.
 11. C. J. Lin, T.-L. Hsieh, and S.-F. Lin, "Development of staffing evaluation principle for advanced main control room and the effect on situation awareness and mental workload," *Nucl Eng Des*, vol. 265, pp. 137-144, 2013.
 12. C. J. Lin and T.-L. Hsieh, "Development of task analysis method for operator tasks in main control room of an advanced nuclear power plant," *J Nucl Sci Technol*, vol. 53, no. 1, pp. 102-111, 2016.
 13. P. Setthawong and R. Setthawong, "Updated Goals Operators Methods and Selection Rules (GOMS) with Touch Screen Operations for Quantitative Analysis of User Interfaces," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 9, no. 1, p. 258, 2019.
 14. C. D. Wickens, J. G. Hollands, S. Banbury, and R. Parasuraman, *Engineering psychology and human performance*. Psychology Press, 2015.
 15. J. Q. Liao and W. S. Zhang, "The Reading Velocity of Chinese Characters." *Chinese Journal of Ergonomics*, vol. 2, no. 1, pp. 38-41, 19