Human Operational Errors in a Virtual Driver Simulation

Toru Nakata, Altion Simo, Koji Kitamura, and Takeo Kanade

Digital Human Research Cetner, National Institution of Advanced Industrial Science and Technology
Aomi 2-41-6 Koto-ku, Tokyo, Japan.
toru-nakata@aist.go.jp

Abstract

We propose a virtual user that simulates human machine interaction including errors in body actions and mistakes in cognitive decision. Occurrence of human error has deep relationships with control of user's body movement. For simulation a car driver, our virtual user model generates human-like body movements and error escalations in cognitive process. As an implementation example, we present a virtual user in a car driver seat that recognizes situation, plans actions, and interacts with human interfaces. The system simulates cognitive mistakes due to the gap between subjective understanding of the virtual user and objective state of interaction situation. Arm movements of the virtual user are generated from captured data of real human drivers to resemble in movement volatility.

1 Motivation for Virtual User Simulation

Human factor is the biggest problem. Today, more than 60% of accidents are caused by human errors (Perrow, 1984). Reliabilities of machines themselves are quite high, but human users perform badly in operating complex machines. For a human interface designer, reducing human error is still a main problem.

Real human subjects cost too much. In designing products that includes human-machine interaction, such as dash board panels, buttons, and switches for operational purposes at a car driver's seat, user tests are essential for increasing performance and comfort and decreasing human errors. Such ergonomics evaluation experiments that employ many human subjects cost time and money. Even the notion of the sufficient and practicable number of human subjects for such tests is often questioned (Bevan et al., 2003).

Virtual user models enable quick and bulk ergonomics experiments. Virtual user models in place of real human subjects alleviate the difficulties. Simulations in computers run quicker than real world experiments and allow a considerable number of experiments. User model simulation provides seamless designing process; the model is used throughout design and evaluation in Computer-Aided Design (CAD) systems.

Virtual user models detect weak points of human interface designs. Virtual users operate machines in a virtual simulation world and make human-like errors from time to time. Our simulation system calculates probabilities of errors by considering interactions between the interface's model (such as button arrangements, materials, and feedbacks) and the user's model (such as sensory inputs, poses, and body movement controls). Generation processes and rates of human errors are worth to be estimated, since severe accidents are often triggered by tiny human errors. Confusing interfaces will increase error probabilities. Virtual users will reveal confusing points of human interface design by making more mistakes there.

Conventional user models ignored generation of error mode and effects of body movement. While 'user modeling' is an established area of human machine interaction study, most of conventional user models do not consider possibilities of unexpected error modes or influences of body shapes and movements of users. The scope of the usability assessments were limited in expected error modes and tended to ignore geometric relationships between users and machines.

We present a human user model behavior at a car driver's seat as it is a representative case that includes both symbolic and physical aspects.

Table 1: History of User Modeling

Generation	Feature	Limitation	Works
Ancient-: Implicit	Designers image user behavior.	Design holes are left on unexpected user behavior.	--
1940's-: Transfer Function	User as linear system element.	Deal only numerical inputs and outputs. Cannot describe behaviour switching.	Craik (1947), Fitt (1954), Accot & Zhai (2003), Stassen (1987).
1950's-: Reflective behavior selection.	Explain switching of behavior by reflective decision algorithm.	Difficult to model non-reflective behaviors.	Tinbergen (1951), Walter (1951).
1970's-: Student Model	Estimate correct answer rate of student by degree of curriculum progress and difficulty of problems.	Difficult to generalize models over various educational curriculums.	Barr et al. (1975); Koffman & Blount (1975), VanLehn (1990).
1980's-: Cognitive Process Model	Block diagrams that trace cognitive processes of operators.	Conceptual. Not quantitative and not computational.	Broadbent (1958), Kahneman (1973), Wood et al. (1987); Wickens (1987), Rasmussen (1986).
1980's-: Error Rate and Efficiency Assessment	Estimate error rates and time consumptions by referring typical value of error rates.	Ignoring difference of situations.	GOMS model (Card, Moran & Newell, 1983); THERP (Swain & Guttman, 1983); Worledge (1985).
1990's-: Body Model. Digital Mannequins	User model including physical body and cognitive processors.	Body factors are not enough utilized to simulate human error generation.	Air MIDAS (Gore & Corker, 2000ab, 2001); Sakajo et al. (2002); Park et al. (2004).

2 Requirements of User Model Architecture for Human Error Simulation

We consider requirements for user models to realize cognitive and physical simulations for usability assessments. We regard importance especially on ability to generate error modes.

2.1 Conventional User Models

Classic user models lack mechanisms for physical and cognitive simulation (Table 1). The history of user modeling started with transfer function model in the cybernetic age. Consideration on user body shapes was included since 1990's. Cognitive factors were taken account since 1980's. Human error has been included in the scope of human modeling after student models in computer-aided instruction of 1970's. Based on those precedent studies, we can consider unification of physicality, cognitive process, and error mode generations.

2.2 Generate Human Errors by Escalating Cognitive Gaps

Severe human error is cognitive gap that may grow to severe misunderstanding. Human error can be defined as disagreement between operator's expectations and actual results (Norman, 1986). Even though minor cognitive gaps are common and not preventable, severe misunderstandings of users must be prevented. Assuring human error safety is preventing cognitive gaps from growing to severe accidents.

Mission of safety design is to prevent escalation of cognitive gaps. Cognitive gaps that do not attract users' attentions may grow to severe misunderstandings without being fixed by users' error recovery efforts (Figure 1). On the other hand, escalating but detectable errors are just inefficient and not severely dangerous. For safety, designers' mission is to anticipate and prevent errors that may escalate to unrecoverable without being detected.

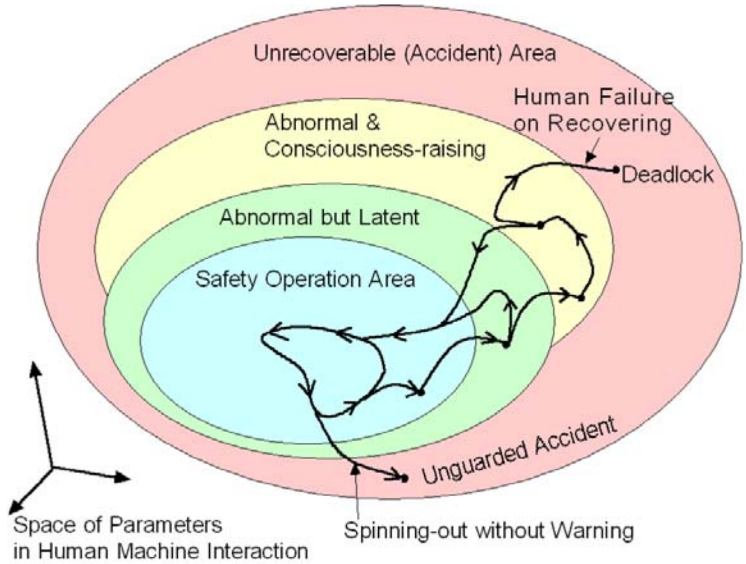


Figure 1: Error escalations. Errors that are detectable in recoverable period are rather safe. Errors that escalate under cover may cause severe accidents.

Cognitive gaps are necessary to explain spontaneous recovery behavior of user. Cognitive discord between subjective understandings and objective state drives users to check situation. When a mismatch between the subjective understanding and objective condition of circumstance is detected, users begin to check the objective condition and recover errors.

Many conventional user models were built as models of flawless users. Some of the conventional user models have ability to generate the mismatches, but most of them assume that the operator models recognize error immediately. Therefore, the simulations ignore possibility of error escalations. When single action errors happen in the simulations, the systems stop simulation processes and do not observe aftermaths of single errors.

Simulator should manage separately virtual user's memories for subjectivity and objectivity to produce the cognitive gaps. The mechanism to simulate cognitive gap escalations consist in separated architecture of virtual user memory: memory for virtual user's subjective understanding of the situation, and memory for objective description of the situation (Figure 2). The architecture imitates latent growths of cognitive gaps in real human users; virtual users do not recognize cognitive gaps until disagreements occur between their subjective understandings and sensory inputs.

2.3 Generate Unexpected Error Modes

Severe error is unexpected error. Predicting erroneous behavior is more difficult than estimating normal behavior. While normal operations can be described as one-way sequences, erroneous behaviors have mutual and complex dependencies. Causes and consequences of errors will stochastic and often unexpected.

Designing safety is finding unexpected error modes. Detected safety holes are no longer major dangers and become safe after countermeasures are taken.

Simulation of human error should be generative to indicate possibility of unexpected error modes. Mere randomness in simulation algorithm is not enough to make usability simulations generative; types of error modes will not increase. Simulation systems should not stop the process of simulation when an error happens. By continuing the process, usability simulation can include aftermaths of the error that may escalate to severe accidents or be recovered. Reactions of virtual users make simulation processes complicated and error mode generative.

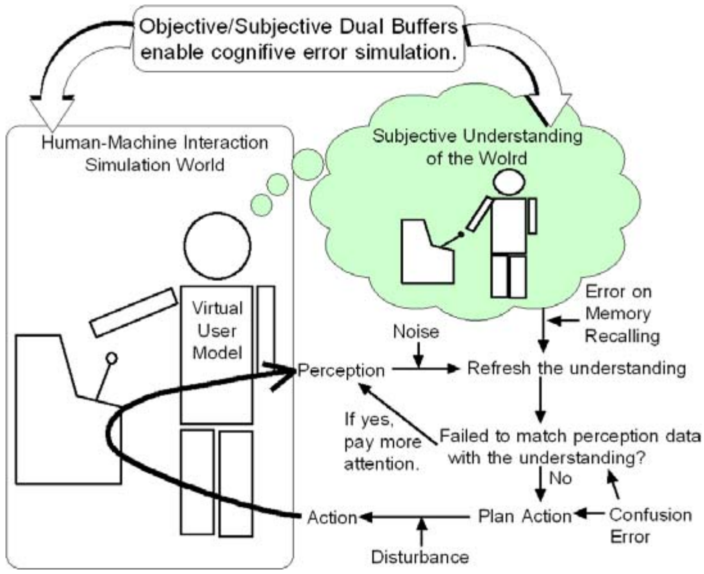


Figure 2: Mechanism of cognitive gap escalation.

Conventional user error estimations tried to forecast human error probability under limited contexts. Most of the traditional simulation systems are likely to estimate human error probabilities (HEP). Although HEP implies an aspect of usability, they do not represent processes of unexpected error occurrences. Accuracy of results for usability experiments is often discussed. Estimations made by traditional user models, such as error rates, are repeatedly questioned. 'Theoretical values' of usability estimations are unlikely to exist. Even if some formulas would be correct for particular users and situation, they are not useful for unexpected error modes. Usability simulation should not be considered as a weather forecast.

We utilize usability simulations as accelerated tests. Features of virtual users, such as body size and cognitive ability, should be adjustable even beyond the ranges of real humans. Simulations of interactions with virtual users with extremely low ability will be accelerated tests of input devices, since human errors will occur more frequent than interactions with real humans. The results will not be accurate in values but they will make useful suggestions about defects of the input devices.

2.4 Include Human Body

Spatial interaction between users and input devices deeply affect usability performance. Spatial conditions have strong influence on error generation on perception and actuation of users. Scales of input devices for daily use are usually in the same order as human body. A certain portion of human errors, such as misidentification of buttons, are caused by relationships between user bodies and input devices. Designs of human interfaces require considerations on human body shape and body movement.

Conventional error assessments have been ignoring or underestimating the role of human body shape. For one reason, human body structure is too complicated. Another reason, human interfaces to be assessed in conventional research are often much larger than human body size. Some of the conventional works employed operator models to evaluate designs of consoles or control rooms for nuclear power plants (Furuta et al., 1996). Details of human body shape were not important.

Today we can use digital mannequins, which are human models including anthropometric factor. Conventional digital mannequins do not include mechanisms of perception and decision making, which will enable to utilize digital mannequins as virtual users for usability tests.

3 Implementation a Virtual Car Driver: Simulation of Behaviors toward Peripheral Input Devices

We designed a prototype virtual user system that simulates behaviors and errors of a driver's hand on the passenger side. The behaviors in a car are important for safety and utilities of cars, even so we often make slips. These behaviors are relatively limited and constrained by spatial condition of car cockpits, so that they are easy to be modeled.

As an evaluation of our simulation, we try forecasting driver's strategy of hand movement routing. Even though it is difficult to estimate exact rates of positioning errors, the system can estimate the order of error rates. Based on the order, the virtual driver of the system plans a strategy of hand movement routing.

3.1 Features of Car Driver's Blind Operations and Errors

As target of simulation, we chose driver's operations toward peripheral equipments such as levers, radio buttons and navigation system.

Drivers perform the peripheral operations often in blind. Drivers are usually watching forward outside. The peripheral operations require arm movements through certain distances. Arm movement varies among trials even in same driver. Feedback information that drivers get steadily is only perceptions of self-arm posture and contacts between hand and machines. The limitation of feedback makes the operations much erroneous, inefficient and complex.

Most common errors are missing and mistaking objective buttons of instruments on dashboard. In order to generate the error we employ disturbance on body control as the only cause of error. Yet other causes of human errors exists in every steps of human cognitive process as we showed in Figure 2, we develop a prototype simulation system with limited cause of error.

3.2 Algorithm of Virtual User

The virtual driver makes its decisions along the following ten steps (Figure 3 and 4). We designed the algorithm to allow escalation of mismatch for subjective understanding and objectivity of the simulation state.

A) Select a target to move randomly.
B) Consider the possible route to reach the target by learning from history of successes and failures.
C) Calculate spatial difference between the position of the target and subjective-understood position of the hand.
D) Move the left hand without watching. The virtual driver intends to move the hand along the line straightly directed towards the target. It may contain positioning errors.
E) Check contact between the hand and buttons or surfaces of the dashboard.
F) In case of touching a surface, the virtual driver restricts the hand movement along the surface but not poking into it. Subjective understanding of the hand position is modified to agree with the contact of the hand and the surface.
G) When the hand touched a button, the virtual user compares the material type of the button and that of the goal target.
H) If the types of material match, the virtual user thinks that he reached the correct target. Then back to step A to continue the simulation.
I) In case of mismatch, the virtual user watches the hand to correct the subjective understanding of the hand position with respect to the objective state. Then back to step B to retry.

The system simulates error occurrences and recoveries in driving scene. The process produces a difference between subjective understanding and objective situation of the hand position. The virtual driver has to take actions and fix it.

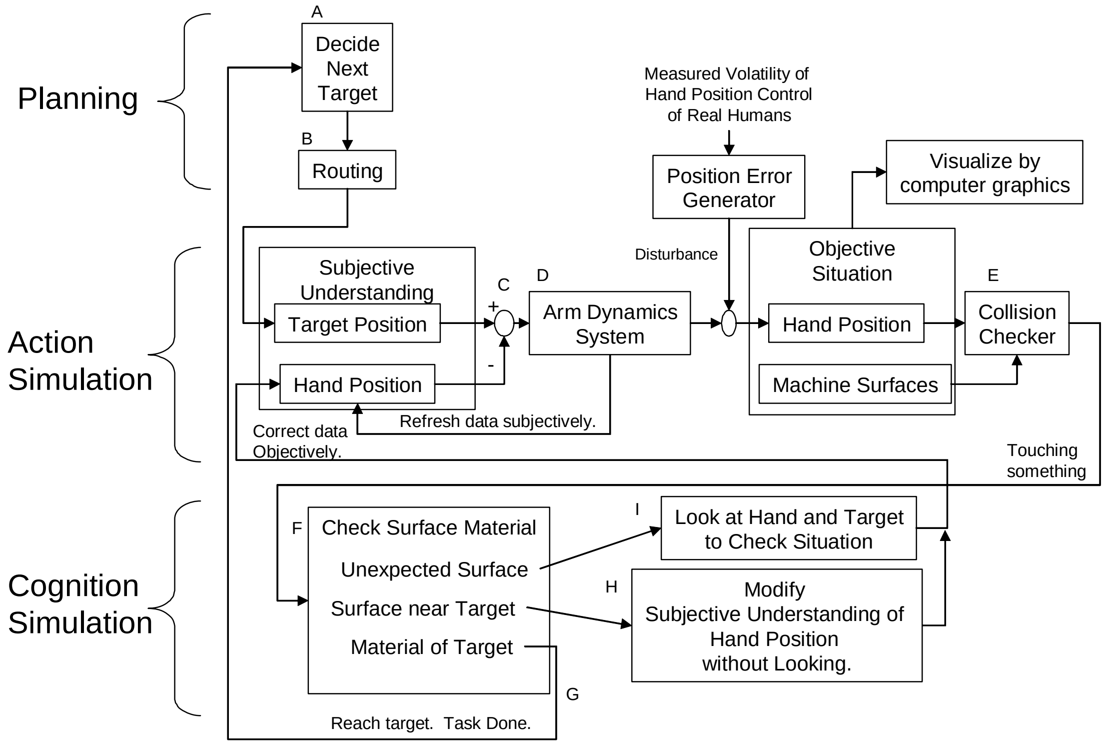
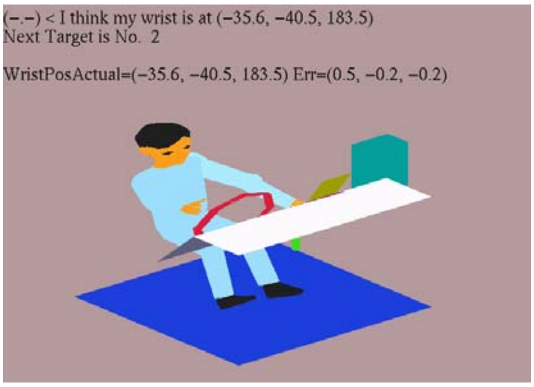
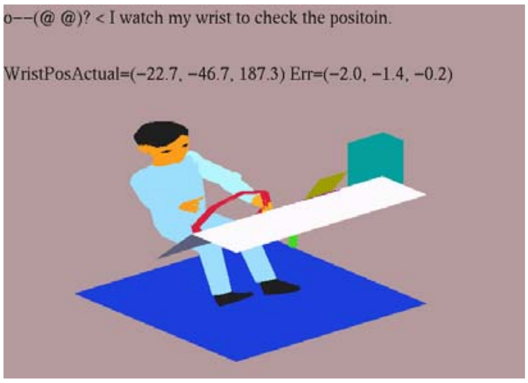


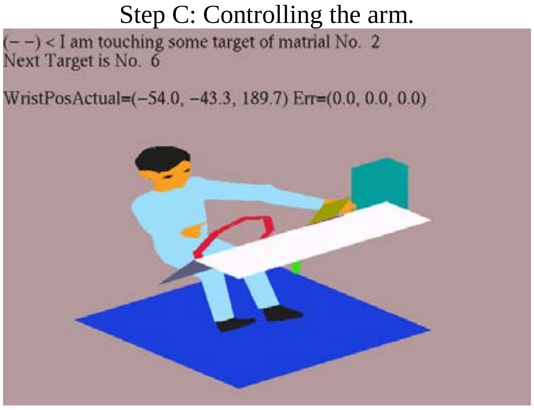
Figure 3: Algorithm of Virtual User



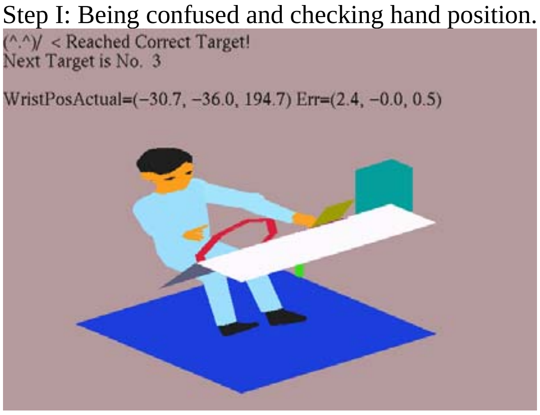
Step C: Controlling the arm.



Step I: Being confused and checking hand position.



Step F: Checking touching material.



Step G: Recognize the target.

Figure 4: Screenshots of Virtual Driver Simulation

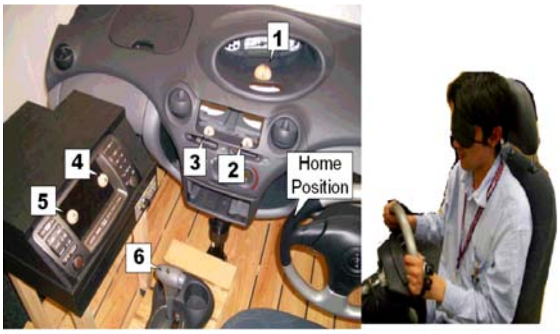


Figure 5: Experimental Driver's Cockpit. Left: Targets and Home Position of Driver's Hand. Right: Subject Driver wearing an eye mask and the 3D-position tracker on the left hand.

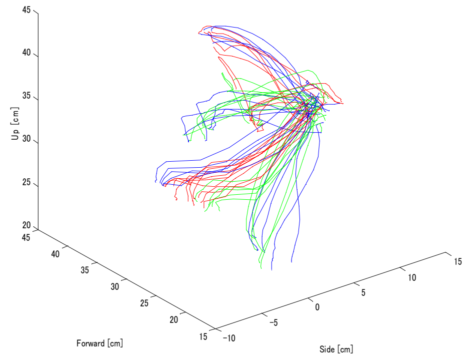


Figure 6: Trajectories of a real human hand's movements from the home position to the targets and returns.

3.3 Generation of Plausible Arm Movement

Movements of human arms draw complicated curves. We invent a method to imitate arm movements that generates imitations of real human movements with plausible precision and volatility of movements.

3.3.1 Capturing Real Human Arm Movement

In order to obtain referential data, we measured movements of the left hand of real human subjects in a mockup of a car driver's seat.

The mockup consists of a steering wheel, a selector lever, surface of dashboard, and a box of navigation machine (Figure 5). We used a dashboard of Toyota Vitz/Iris of right-hand drive.

We prepared six target positions (e.g. radio buttons and the top of the selector lever) to be touched on the mockup. For other target than the selector lever, wooden hemispheres are attached to the target positions, so that subjects can distinguish the targets from other part of the mockup by touching.

An ultrasonic tracker attached on the back of left hand of each subject enables to detect the position. The positioning system used here is IS-600 of Intersence Inc. Its resolution is about 5 mm. Sampling frequency is set to 127 Hz.

Each subject wears an eye mask to prevent from obtaining visual information. We expected that subject behaviors would resemble driver's blind operation during driving.

Each subject is, at first, told to hold a certain point of the steering wheel. The subject moves the arm to a target after the identification number of the target is called by computer voice. Target numbers are called 30 times with 4 seconds intermissions. The order of calling is randomized, and same number is not called in succession. After touching the target, the subject returns the left hand to the home position of the steering wheel immediately.
Figure 6 shows examples of trajectories of a male subject.

3.3.2 Modeling Arm Control

Let $\mathbf{r}[n]$ denote a positional vector included by trajectory of real-human motion.

For purpose of generating plausible movements for virtual users, we have to generate imitational trajectories $\tilde{\mathbf{r}}[n]$ that are not identical to original $\mathbf{r}[n]$ but resemble on dynamic characteristics.

Let us consider structure of each $\mathbf{r}[n]$ by separating into foundations and additives. Trajectories of human motion have bases of certain lines or smooth curves governed by intention of humans. We call the base of movement a 'trend'. Small and unstable movements that are often accompanied with trends can be regarded as noises on movement control. We name the noise-like component 'volatility'. Volatilities are produced by disturbance or uncertain control of human body.

Trend with random walk model (we call it T+RW model in this paper), which has been developed in finance engineering for modeling stock price movements (Black & Scholes, 1976), is easy to implement.

Let us calculate a step vector $\mathbf{tr}$ that directs to the end of the trajectory from the start.

$$\mathbf{tr} := \frac{\mathbf{r}[n_{end}] - \mathbf{r}[n_{start}]}{n_{end} - n_{start}} . \qquad (1)$$

We define volatility in T+RW model as un-linearity of the trajectory. The volatility can be evaluated as average of the distance between points on the trajectory and the line that connects the start and the end points of the trajectory:

$$vol := \underset{n:\text{all}}{mean}(|\,\mathbf{r}[n] - n \cdot \mathbf{tr}\,|) . \qquad (2)$$

An imitated trajectory can be generated by T+RW model as the following:

$$\tilde{\mathbf{r}}[n] = \tilde{\mathbf{r}}[n-1] + \mathbf{tr} + \mathbf{G}(vol) \qquad (3)$$

$\mathbf{G}(vol)$ is Gaussian probabilistic vector which standard deviation is the original volatility. The sequence of $\tilde{\mathbf{r}}[n]$'s will be an imitation of the trajectory. T+RW model keeps the characteristics of original data with respect to the direction of the trend and overall average of the volatility.

Having multiple originals of trajectory data, T+RW model can improve accuracy of trajectory imitation.

Let $\mathbf{tr}_m$ denote the trend vector calculated from motion data of m-the measurement. We get the average of the trends vectors.

$$\mathbf{tr}_{average} = \underset{m}{average}(\mathbf{tr}_m) . \qquad (4)$$

Likewise, denotes the volatility of m-th original data. Let us calculate the average of the volatilities.

$$vol_{average} = \underset{m}{average}(vol_m) \qquad (5)$$

Using the averaged data of trend and volatility, we will get more statistically precise data:

$$\tilde{\mathbf{r}}[n] = \tilde{\mathbf{r}}[n-1] + \mathbf{tr}_{average} + \mathbf{G}(vol_{average}) \qquad (6)$$

We use T+WR model with averaging over 5 samples.

3.3.3 An application: Assessing the Safest Hand Path That the Interface Can Afford.

Step B in the algorithm of the virtual user includes a learning process to plan a route on which the movement is expected to be the most successful.

Define $SR(A \to B)$ is the rate of successes among trips from point A to B. As shown Figure 7, the virtual user considers alternative routes that detour via another point: the virtual user decides to detour via C to A, if some point C exists, such that

$$SR(A \to B) < SR(A \to C) \cdot SR(C \to B) . \qquad (16)$$

Otherwise the virtual driver decides to move to B directly.

In the simulation, short movements are estimated as successful. Long movements tend to induce noise accumulations on movements which result in gaps between subjective recognition and objective condition of the hand's position. Short movements provide more feedback information for the hand position: the virtual driver recognizes and fix the hand position by relaying these positions.

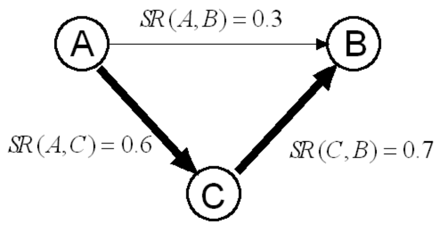


Figure 7: Diagram of highway. When direct movement from A to B is unlikely to be successful according to operation history, the virtual user searches successful detours.

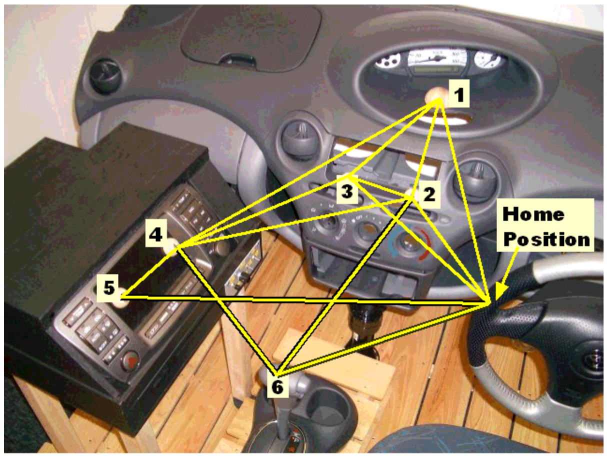


Figure 8: Highways for hand movement acquired by learning.

We name the detour 'highway', since the detours are likely to offer successful and easy movement control. Figure 8 shows a result of acquired highways after learning with 10,000 trips. As an advantage of simulation, the system executed the huge number of trips in 10 seconds.

4 Conclusion

We proposed a virtual user model that simulates modes of human errors by discussing requirement for use simulators to generate unexpected error modes. In simulation, the user model generates human-like body movement by analyzing captured motion data of real human subjects, and assessed interface's performance on providing safety path of hand movements. We will evaluate our system on correctness of estimating difficult-to-use points. We plan to expand the system by adding other error sources to generate different error levels, such as memory error and procedural error.

References

Accot, J, & Zhai, S. (2003). Refining Fitt law models for bivariate pointing. *Proc. CHI2003*, ACM Press, 193-200.

Barr, A. Beard, M. & Atikinson, R. C. (1975). The computer as a tutorial laboratory: the Stanford BIP project. *Int. J. Man-Machine Studies 8*, 567-596.

Bevan, N., Barnum, C., Cockton, G., Nielsen, J., Spool, J., & Wixon, D. (2003). The "magic number 5": is it enough for web testing? *Proc. CHI 2003.* ACM Press, 698 – 699.

Black, F. & Scholes, M. (1976). The pricing options and corporate liabilities. *J. Financial Economics 3*, 167-179.

Broadbent, D. E. (1958). *Perception and Communication.* Pergamon Press.

Card, S. K., Moran, T. P. & Newell, A. (1983). *The Psychology of Human-Computer Interaction.* Lawrence Erlbaum.

Craik, K. W. J. (1947). Theory of the human operation in control systems 1: the operator as an engineering system. *British J. of Psychology 38*, 56-61.

Feyen, R. G. & Liu Y. (2001). The queuing network-model human processor (QN-MHP): an engineering approach for modeling cognitive performance. *Proc. Human Factors and Ergonomics Soc.*

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Experimental Psychology 47*, 318-391.

Furuta, K. Shimada, T. & Kondo, S. (1996). Simulating interaction between operator and environment. *Proc. Int. Conf. on Probabilistic Safety Assessment and Management,* 15-20.

Gore, B. F. & Corker K. M. (2000a). Human performance modeling: identification of critical variable for national airspace safety. *Proc. IEA/HFES 2000.*

Gore, B. F. & Corker, K. M. (2000b). A systems engineering approach to behavioral predictions of an advanced air traffic management concept. *Proc. IEEE 19th Digital Avionics Systems Conf.*

Gore, B. F. & Corker K. M. (2001). Human error modeling predictions: increasing occupational safety using human performance modeling tools. *Proc. Computer-Aided Ergonomics and Safety Conf.*

Hollnagel, E. (1993). The phenotype of erroneous actions. *Int. J. of Man-Machine studies 39*, 1-32.

Hollnagel, E. Cacciabue, P. C. & Mancini, S. (1993). The system response generator: an interactive analytical tool to evaluate simplicity and effectiveness of procedures. *Proc. 4th European Conf. Cognitive Science Approaches in Process Control*, 92-101.

Horonf, A. J., & Halverson, T. (2003). Cognitive strategies and eye movements for searching hierarchical computer displays. *Proc. CHI 2003.* ACM Press, 249-256.

Kahneman, D. (1973). *Attention and Effort*. Prentice Hall.

Koffman, E. B., & Blount, S. E. (1975). Artificial Intelligence and automatic programming in CAI. *Artificial Intelligence 6*, 215-234.

Newcombe, N. S., & Huttenlocher, J. (2003). *Making Space.* MIT Press.

Norman. D. A. (1986). Cognitive engineering. In Norman, D. A. & Draper, S. W. (Eds.). *User Centered System Design*, Erlbaum.

Park, W., Chaffin, D. B. & Martin, B. J. (2004). Toward memory-base human motion simulation: development and validation of a motion modification algorithm. *IEEE Trans. SMC 34*, 3, 376-386.

Perrow, C. (1984). *Normal Accidents*, Basic Books.

Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction: An approach to cognitive engineering.* Elsevier.

Sakajo, S., Nakagawa, T., & Terashita, N. (2002). Human error evaluation methods for fieldworks. *Proc. SICE Sensing Forum*, 65-70. (In Japanese).

Stassen, H. G. (1987). Human supervisor modeling: some new developments. *Int. J. Man-Machine Studies 27*, 613-618.

Swain, A.D. & Guttmann, H. E. (1983). *Handbook of human reliability analysis with emphasis on nuclear power plant applications. NUREG/CR-1278.*

Tinbergen, N. (1951). *The study of instinct.* Clarendon Press.

Vanlehn, K. (1990). *Mind bugs*. MIT Press.

Walter, W. (1951). A machine that learns. *Scientific American, 185 (2)*, 60-63.

Wickens, C. D. (1987). Information processing, decision-making, and cognition. In Salvendy, G. (Ed.). *Handbook of Human Factors.* John Wiley and Sons, 72-107.

Woods, D. D., Roth, E. M. & Pople, H. (1987). *An artificial intelligence based cognitive environment simulation for human performance assessment. NUREG/CR-4862.*

Worledge, D. H. (1985). Some useful characteristics of performance models. *Proc. IEEE Conf. Human Reliability*, 43-52.