Eye Directed View

Author Not Available

Follow this and additional works at: https://commons.erau.edu/space-congress-proceedings

Scholarly Commons Citation
https://commons.erau.edu/space-congress-proceedings/proceedings-1984-21st/session-2/4
manipulation of levers and switches. The ability to perform both tasks simultaneously with the same acquired ease that the radio operator accomplished his dual task is desirable. This paper investigates the feasibility of substituting an automaton for a human process in a man-machine control system. It also describes modifications to, and further development of, a commercially available automaton that transforms the natural inclination of an operator to look beyond the boundaries of a work scene, as presented on a video monitor, into the pan and tilt of television cameras. The automaton is unobtrusive, allows complete operator freedom, and approaches full capability of correct response with minimal operator orientation.

SYSTEM ANALYSIS

The primary purpose of substituting an automaton for a human process is to reduce the operational load of the operator. In order to evaluate a priori the effectiveness of the substitution, it is necessary to establish a comparison metric and an appropriate man-machine control model.

A good measure of the performance efficiency of a process is its duration while accurately completing a task. For a human, this process duration can be measured as reaction time (RT) which is the time between presentation of a stimulus and a detectable response. A good approximation of RT is Hick's law which is the relation:

\[ RT = a + bT \]  

where 'a' and 'b' are constants dependent on the task context and T is the average information transmitted per response. Note that T represents transmitted information and not input information (N). The constant 'a' represents constant time unrelated to T, and 'b' is the additional time needed to select a response correctly for additional units of information. The reciprocal of 'b' is thus considered to be the rate of information transmission. Hick's law represents the human operator as a fixed delay in series with a rate limited information channel (ref. 3). The goal of reducing RT to a minimum value can be met by optimizing components of 'a' and 'b'. There are some physical components of 'a' that are irreducible (e.g., nerve conduction and muscle contraction) but the information processing components can be affected by optimizing techniques. Two techniques that are relevant to increasing the efficiency of concurrent processes are generating compatible input and output codes, and practice (ref. 2). Congruent input and output symbology reduces the value of 'b' by simplifying the response selection process and practice minimizes the value of 'a' by shortening the perception and interpretation time requirement. Practice also affects 'b', especially if the input and output codes are greatly incompatible. Both of these techniques also have an optimizing effect on information transmission, causing T to approach N (ref. 3). The RT metric provides a simple and expressive measure for determining the effectiveness of automaton substitution in the work scene adjustment process. It allows the comparison of optimal human processing with optimal machine processing and also provokes a comparison of the relative difficulty of achieving optimization in either process.

Man-machine control systems can be classified according to the visual input to the human operator (ref. 3, 5 and 10). Usually, four classifications are identified:

1) Compensatory - input = the error signal (i.e. the difference between actual response and ideal response)

2) Pursuit - input = the ideal response and actual response with both inputs in motion

3) Preview - input = the same as pursuit with the addition of future ideal response displayed

4) Precognitive - input = knowledge of attributes of the ideal response rather than a direct view

A typical teleoperator task involves tracking a static target with the path to the target known prior to the commencement of the task (ref. 10). The a priori path knowledge dictates a preview classification. The static target indicates a compensatory system in that the varying distance between the effector and the target can be considered an error input. The static target restraint eliminates the pursuit classification. There is inevitably a precognitive influence, but its effect is negligible relative to the effect of a direct view of the input. The best approximation of a teleoperator control task is as a combination preview/compensatory system that degenerates to a compensatory tracking task in the vicinity of the target.

A man-machine control system in which the human operator is responsible for a scene adjustment task in addition to a teleoperator function is diagramed in figure 1. The scene adjustment process is activated when the target is not in view on the television monitor.
The operator's eye point of regard is positioned some place along the monitor's boundary while camera pan and tilt is commanded such that the scene is adjusted in the direction of the lookpoint. This task is similar to the teleoperator task with the exception that the a priori path knowledge is the operator's precognition of the work environment outside the perimeter of the monitor's screen. The error input is the difference between the image of the target and the center, or some other appropriate area, of the screen. The scene adjustment operation can be classified as a precognitive/compensatory tracking task.

Since both the teleoperator and scene adjustment processes degenerate to compensatory tracking tasks, they can be represented by the same model. The simplest representation of a human operator that gives realistic results is a quasilinear model (ref. 2, 3 and 5). The most common quasilinear model used to represent human operator dynamic characteristics in a compensatory tracking system is shown in figure 2. If the equalization factor is always optimal relative to Yc regardless of operator identification, then only the fixed liabilities factor is relevant in operator efficiency comparison. For a human performing a compensatory tracking task:

\[ T = RT = 0.12 \text{ sec. to } 0.20 \text{ sec.} \]
\[ T_N = \text{neuromuscular lag} = 0.1 \text{ sec. to } 0.2 \text{ sec.} \]

Furthermore, it is assumed that RT is optimized through practice and input/output congruency. Thus, the values given above for RT and \( T_N \) are considered to be optimal for humans.

Another consideration when analyzing information processing system efficiency is system capacity. The capability of processing multiple tasks through parallel channels simultaneously is a desirable system feature. Theoretically, when \( T \) equals \( N \) through an information channel, the processing capability on that channel is not being utilized. In other words, the channel processor is available for use while simultaneously effecting throughput. In reality, \( T \) rarely if ever equals \( N \), but can approach \( N \) nearly enough to make parallel activity possible. The possibility of true parallelism is apparent in reflex actions and in abilities such as demonstrated by humans operating bicycles. However, for the most part, a human is single channel, limited capacity system that must multiplex in order to perform concurrent tasks (ref. 2, 3 and 5). The multiplex process is represented by \( Y_{HW} \) in figure 1 where \( T_{SW} \) is the switching time. The optimal switching time for a human is about 0.2 seconds (ref. 3).

The models and values presented so far have been gleaned from the literature and are validated through acceptance over time. They provide a benchmark for the design of machine based alternatives to human processing and for subsequent experiments to determine the effectiveness of any alternative.

SYSTEM DESIGN

The primary goal driving the design of an automaton capable of executing the scene adjustment task is to effect a marked increase in total system efficiency. Two criteria for successful achievement of this goal are reduction of the scene adjustment process time and reduction of switching time by implementing the process in parallel. The substitution of a processor with substantially higher transmission rates than a human can be easily accomplished given the current state of computing power. But to provide true parallel processing capability is more difficult. The automaton must be able to respond directly to human sensory input without any intervening human consciousness. For the scene adjustment application, the automaton should sense the operator's eye point of regard, determine whether or not to activate camera pan and tilt, calculate the resultant direction of camera movement, and execute the movement if warranted. It should be able to reduce the scene adjustment process time to a reflex response.

The automaton can be defined in terms of cerebral and motor responsibilities. The cerebral functions include eye position and eye rotation input from a sensor, processing of input data to determine lookpoint, and short term storage to provide continuity. The motor responsibilities are to maintain acquisition of the eye and to command camera motion to cause the desired scene adjustment. An instrument that is available off the shelf and possesses, or has the potential to possess, these functions is a remote oculometer. The most attractive asset of a remote oculometer is its ability to unobtrusively measure the angular deflection of the geometric axis of the eye and to determine the intercept of the axis with any plane of interest (ref. 7 and 8). This capability allows an oculometer based automaton to effectively read in data through the human's visual input subsystem without any cognitive effort required of the human. It subsequently processes the visual data and performs motor functions in parallel with the human.

A revised system diagram, with the automaton process \( Y_A \) substituted for the human...
process \( Y_{\text{na}} \), is shown in figure 3. Neither switching delay nor neuromuscular lag is a factor when \( Y_{\text{na}} \) is implemented. The parallel nature of the implementation eliminates the need for switching and the propagation delay \( 10^{-9} \) seconds through the microelectronics of the pan and tilt interface is negligible when compared with neuromuscular lag \( 10^{-1} \) seconds. The processing time \( (PT) \) of the oculometer based automaton ranges from 20 milliseconds to 37 milliseconds (see appendix A for the derivation of \( Y_A \) and determination of \( PT \)). This is an order of magnitude decrease in process lag relative to the human scene adjustment process.

The design synthesized in this section has been implemented and is operational. A technical description of the oculometer based scene adjustment system is presented in appendix B. It describes both the commercially available components and in-house modifications necessary to implement the design.

Verification of the machine's effectiveness has not yet been effected. The ultimate determination of the effectiveness of the scene adjustment automaton as an extension of the human is the length of time required for the human to adapt to its operation. Ideally, the adaptation period should have length zero. Practically it should be very much shorter than the amount of practice time necessary to optimize the human process. Experiments are planned to evaluate how well human and machine interact in systems of this type where the machine's action is a reflex response to a human sensed stimulus.

CONCLUSION

Continuing technological advances are allowing the practical application of machines that exhibit extended intelligence. Until recently, such machines have been relegated to the laboratory, because their requisite computing, sensory and motor components were dimensionally incompatible with a human. In order to perform not, so primitive tasks in response to a humanly perceived stimulus, a machine must do memory intensive computations and precisely measure the stimulus or the effects of the stimulus while interfacing unobtrusively with the human. This requires an extremely high density of active components on a physically small platform. Current VLSI techniques are making feasible the implementation of automats that can function intra-human as well as extra-human. The age of true symbiosis involving man and machine is quickly approaching if not already existent.

APPENDIX A

The scene adjustment automaton's transfer function was determined by simulating a typical oculometer input. A 2.5 hertz square wave, alternating between 0 and -1.5 volts D.C. that represented a lookpoint change of approximately 7.7 inches on a typical fixation plane was input to the automaton. The resulting output was displayed, along with the input, on a dual trace oscilloscope so that the transform pairs could be identified. As shown in figure 4, the input was a step function and the output a delayed step function. The transfer function is derived as follows:

\[
\begin{align*}
F(s) &= 1/s \\
G(s) &= 1/s \cdot e^{-PTs} \\
H(s) &= G(s)/F(s) = e^{-PTs}
\end{align*}
\]

\( PT \) ranged from 20 milliseconds to 37 milliseconds, averaging 28.47 milliseconds.

APPENDIX B

The scene adjustment automaton is comprised of four subsystems; the eye movement sensor, the processor, the sensor controller, and the visual scene controller.

Eye Movement Sensor. This is an electro-optical instrument that illuminates the eye with IR radiation in the 0.8 to 0.9 μm band. The pupil and corneal reflections are imaged by collection optics onto a vidicon tube. The illumination and collection optics are bore-sighted so that the corneal reflection always appears in line with the center of corneal curvature. The corneal reflection and the center of the pupil move differentially only with rotary motion of the eye relative to the sensor, thereby allowing measurement of eye direction. The electro-optical sensor is a part of the commercially available oculometer package but, depending upon the application, needs modification.

There are two modifications required to increase the utility of the oculometer in eye control applications.

First, the sensor must be physically small enough to allow flexibility in adapting to varying installation requirements. The electro-optical sensor for the scene adjustment automaton is approximately 1/3 of its original size at 5.75" x 3.5" x 14.5". Its depth can be further reduced by substituting a solid state image sensor for the vidicon tube. A sensor that uses a CCD imager has been developed at Langley but is not presently in use. Further flexibility has been incorporated by providing two entry
ports and by using tubes of varying length coupled with 90 degree folding mirrors to allow variation of the optical path.

Second, the human operator must be allowed freedom of head movement. The oculometer allows approximately one cubic foot of head movement. It accommodates one square foot of movement in any plane normal to the axis of the electro-optical sensor by automatically tracking the operator with a closed loop moving mirror system. Change in focus due to head movement along the sensor's axis had been automatically adjusted initially, but due to instability in the automatic focusing sub-system, it has since been adjusted manually. Because a stable automatic system is needed, an improved automatic focusing feature has been developed based on the effect that a lens system, uncorrected for astigmatism, has on the image of a point source off the axis of the objective lens. When point source \( Q \) in figure 5 is moved away from the lens, its image appears as an ellipse whose horizontal axis is the major axis \( E \) in figure 5. When \( Q \) is moved closer, the major axis of the ellipse is the vertical axis \( F \) in figure 5. This effect can be used to drive a focusing lens in the proper direction until the image is refocused at the circle of least confusion (figure 5). Astigmatism can be introduced in the electro-optical sensor by yawing the objective lens with respect to the optical path. The fact that at focus there is now a circle and not a point does not degrade the eye direction measuring function of the oculometer. The measurement depends on the relative positions of the centers of the pupil and corneal reflections and not on their respective sizes.

Processor. The processor is a 16 bit microcomputer with 24K words of memory and an 800 nanosecond cycle time. Its primary function is to determine look angle based on the digitized pupil and corneal image, and derive the gaze vector intercept on selected objects from their geometrical relation to the electro-optical sensor. Secondary functions include head tracking, auto-focus and external device control.

Most of the software modifications have been done in support of the implementation of new features. Modifications that have been done to increase the efficiency of the processor function involve efforts to distribute processor responsibility. With the advent of the computer on a chip, the delegation of headtracking, auto-focus, and external device control functions has become feasible.

Sensor Controller. This controller performs the eye tracking, eye illumination and auto-focusing tasks. The eye is continuously illuminated and continuously positioned in the center of the imaging surface regardless of head position by a moving mirror sub-system. The pitch and yaw mirrors, which track vertical and horizontal eye translations respectively, are mounted in a box that is positioned at the sensor's input port. The mirrors are driven by galvanometers that respond to a signal that is proportional to the difference between the center of the pupil image and the center of the imager. The eye tracking task has recently been delegated to a single chip microprocessor resulting in a fourfold increase in tracking rate. The mirror position is currently updated every video field (16.7 milliseconds) with a lag of 1 field in mirror response. Previously, the position update took place every other field with a 1/4 field lag in mirror response.

The pupil and corneal image is focused onto the imager by a servo driven lens. The image of the corneal reflection is used to determine the condition of focus and, if defocused, the direction of the focal plane relative to the current position of the focusing lens. All of the information necessary to make these determinations exists in the corneal image as a result of the astigmatism introduced into the lens system by yawing the objective lens. The focusing software currently resides in the automaton's processor, but will be relocated to the microprocessor responsible for eye tracking.

A constant illumination of the eye is maintained by a control loop that compares the brightness of the pupil return with a preset value.

Visual Scene Controller. An external device controller has been developed to add motor capability to the oculometer's existing sensing and processing capabilities. The controller consists of three functional parts, an area of regard identifier which has general utility, a proportional rate x-y step generator which is specific to motor applications in an x-y plane, and a camera pan and tilt unit, which is specific to the scene adjustment task.

The area of regard identifier uses the lookpoint coordinates, as normally output by the oculometer, to determine which of a maximum of 64 areas is being observed. The area boundaries are preset to any desired configuration. For the eye directed view application, an approximately 0.5 inch wide border around the edge of a television monitor is divided into rectangular areas. A lookpoint intrusion into any area produces both an analog and a digital code unique to that area. This code is used by the step generator to command pan and tilt.
The step generator develops x and y pulse trains whose relative rates indicate the direction in which the eye is moving when the operator's point of regard enters the border. The pulse trains are converted to analog ramp signals that drive the pan and tilt actuators. A camera mounted on the pan and tilt mechanism will continue to scan in the desired direction until either the point of regard enters another border area or returns to some preselected area of the monitor. The camera will then either change direction of scan or cease scanning.

The present configuration of the scan adjustment automaton allows 30 degrees of pan and 30 degrees of tilt. The pan and tilt are currently accomplished by the deflection of a galvanometer driven mirror assembly that reflects the scene onto the camera lens. Future implementation will use standard camera pan and tilt mechanisms. An advantage of using galvanometers as actuators is their almost instantaneous response to change in input. This is ideal for laboratory use in that it allows an experimenter to observe the effect of a wide variety of scene scan rates on the human operator. For practical applications, the durability of the standard actuators, which include stepper and servo motors, makes them the better implementation choice.

Figures 6 and 7 are schematics of the oculometer as originally configured and the oculometer as modified for the scene adjustment function, respectively. Figure 8 shows the reconfigurable control station used at Langley as a testbed for teleoperator studies. In this photograph, it is set up for eye directed view activity.

**SYMBOLS**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>time lag related to perception</td>
</tr>
<tr>
<td>b</td>
<td>time lag related to selection</td>
</tr>
<tr>
<td>N</td>
<td>input information</td>
</tr>
<tr>
<td>PT</td>
<td>automaton's processing time</td>
</tr>
<tr>
<td>Q</td>
<td>point source of light</td>
</tr>
<tr>
<td>Q'1</td>
<td>image of point source in meridional plane</td>
</tr>
<tr>
<td>Q'2</td>
<td>image of point source in sagittal plane</td>
</tr>
<tr>
<td>RT</td>
<td>human reaction time</td>
</tr>
<tr>
<td>T</td>
<td>average information transmitted</td>
</tr>
<tr>
<td>TL</td>
<td>lead equalization</td>
</tr>
<tr>
<td>TN</td>
<td>neuromuscular lag</td>
</tr>
<tr>
<td>TSW</td>
<td>multiplex lag</td>
</tr>
<tr>
<td>T</td>
<td>time constant</td>
</tr>
<tr>
<td>X_A</td>
<td>automaton scene adjustment process</td>
</tr>
<tr>
<td>X_C</td>
<td>controlled process</td>
</tr>
<tr>
<td>X_CA</td>
<td>controlled process, arm</td>
</tr>
<tr>
<td>X_CP</td>
<td>controlled process, pan and tilt</td>
</tr>
<tr>
<td>X_HSA</td>
<td>human scene adjustment process</td>
</tr>
<tr>
<td>X_HSW</td>
<td>human multiplex process</td>
</tr>
<tr>
<td>X_H_T</td>
<td>human teleoperator process</td>
</tr>
</tbody>
</table>

**REFERENCES**


\[ Y_{HT} = Y_{HSA} \text{ EQUALIZATION FACTOR} \]  

WHERE

\[
\text{EQUALIZATION} = K \frac{T_{LS} + 1}{T_{IS} + 1} 
\]

\[
\text{FIXED LIABILITIES} = \frac{e^{-T_s}}{T_{NS} + 1} 
\]

\[ Y_{HSW} = e^{-T_{SW}s} \]

\[ Y_A = e^{-P_T s} \]

\[ Y_{CA} = \text{CONTROL DYNAMICS OF ARM} \]

\[ Y_{CP} = \text{CONTROL DYNAMICS OF PAN AND TILT MECHANISM} \]

FIGURE 2
FIGURE 6