Quantification of Human Perception Using the Motor Human-Computer Interface Hand Tracker

DOI: 10.17691/stm2019.11.1.16
Received December 18, 2018

V.A. Antonets, DSc, Professor, Department of Psychophysiology, Faculty of Social Sciences;
Leading Researcher, Department of Radiophysical Methods in Medicine

1National Research Lobachevsky State University of Nizhny Novgorod, 23 Prospekt Gagarina, Nizhny Novgorod, 603950, Russia;
2Federal Research Center Institute of Applied Physics, Russian Academy of Sciences, 46 Ulyanova St.,
Nizhny Novgorod, 603950, Russia

Here, we propose an all-instrumental objective method (hand tracking) for the real time quantification of perception of stimuli presented in various modalities.

The method is based on measuring the dynamic error of performance by a subject of motor control over a real or virtual (computer generated) stimulus with the help of a contactless manipulator. The method got its name from its previous analogue — the eye tracking; in both cases, the measurement of psychophysiological characteristics of the subject is carried out by monitoring the parameters of his/her motor activities.

The above principle has been implemented in the form of a hardware-software complex Hand tracker, which has been tested for the ability to perceive the “segment orientation angle” and “figure color against the background” virtual stimuli.

With the first stimulus, individual diagrams of perception accuracy were obtained in healthy subjects who were asked to guess the segment inclination angle. The results allowed us to identify the parameters of the neural model simulating the respective primary recognizer.

With the second visual stimulus, we found that in patients with focal brain lesions, the thresholds of color differentiation were abnormally high and, unlike the normal ones, did not depend on the semantic content of the stimulus.

Since the Hand tracker provides for the simultaneous control over several (up to eight) independent parameters of audial and visual stimuli, the proposed method can be used in studying the intermodal connections and/or the collective behavior, when different team members control different parameters of the stimulus, trying to solve a common problem.

The article also addresses the prospect of implementing the hand tracking method in the form of special applications for smartphones with embedded movement sensors. Such a portable device could be used for epidemiological studies of perception (primary cognitive functions) to elucidate the age- and gender-related, regional, social and other stratification of populations.

Key words: virtual stimuli; perception; hand tracking; quantification of perception; cognitive diagnostics and rehabilitation; cognitive epidemiology.

Introduction

Knowledge of the quantitative characteristics of human perceptual systems is important for diagnosing primary cognitive disorders as well as for patient rehabilitation process. Specifically, we talk about controlling the technical, informational and exomotor interface components of simulators and systems that substitute the lost motor functions in patients.

The aim of the work was to substantiate, describe and demonstrate an all-instrumental objective method for quantifying the human perception and rationalize the prospects of its practical applications. The proposed method does not require any subject’s verbalization on his/her subjective feeling aroused by the stimulus.

The method is based on measuring the dynamic error of performance of motor control over a real or virtual (computer generated) stimulus [1]. Here, the dependence of this error on a physical parameter of the external stimulus is used for characterizing the sensory analyzer. The method is termed hand tracking by analogy with the device proposed Yarbus [2] for the study of spatial attention, which is currently called the eye tracker. Both devices are designed to measure the psychophysical characteristics of a person by monitoring his/her motor acts.

The rationalization of the method was carried out using real stimuli of two types: a) “load weight”, when voltage fluctuations of the shoulder biceps muscle were measured under loading with various weights; b) “tonal sound pitch” when frequency fluctuations of tonal sounds were compared for two cases: produced by a human voice and by an electronic voice generator controlled by a contactless manipulator.

The method is implemented in the form of a hardware-software complex (HSC) Hand tracker, which allows one to quantify the perception of audial and visual stimuli in real time.
The present article demonstrates the use of HSC for studying the perception of the virtual stimuli “segment orientation angle” and “color of a figure against a background”. With the first stimulus, individual diagrams of perception accuracy were obtained in healthy subjects who were asked to guess the angle of segment inclination. The results allowed us to identify the parameters of the neural model simulating the respective primary recognizer. With the second visual stimulus, we found that in patients with focal brain lesions, the thresholds of color differentiation were abnormally high and, unlike the normal ones, did not depend on the semantic content of the stimulus.

Since the HSC Hand tracker allows for controlling several (up to eight) independent parameters of audial and visual stimuli, the proposed method can be used in studying the intermodal relations and/or the collective behavior. In the latter, different team members control different parameters of the stimulus, trying to solve a common problem: for example, to achieve a certain size and color of a circle. In both cases, the measurement procedure is also the training one.

A part of the Hand tracker potential can be implemented in the widely used smartphones equipped with movement sensors. Such a portable device can be used for epidemiological studies of perception (primary cognitive functions) in various age, gender, regional, social and other groups of the population.

Instrumental study of perception: methods and problems

To date, two methods for quantifying human perception of sensory stimuli are best known.

The first method was proposed at the beginning of the XVIII century by Bouguer [3], who studied the human perception of brightness of luminous objects for the purpose of improving marine navigation. He established the constancy of the ratio of the barely perceptible difference in brightness \( \Delta R \) between two luminous objects to the background brightness \( R \); this ratio is now called the relative differential threshold of perception. More than 100 years after that, Weber and others showed that if the stimulus magnitude is far from the thresholds of perception and pain, the Bouguer rule holds for stimuli of other modalities [4, 5]. This rule is usually expressed:

\[
\frac{\Delta R}{R} \cong k,
\]

where the value of \( k \) depends on the modality of the stimulus; in the experiments of Bouguer, \( k \approx 1/64 \).

Assuming equality of the relative change in the magnitude of sensation and the relative change in the magnitude of the stimulus, which is subjectively evaluated by a person, i.e.,

\[
\frac{\Delta S}{S} = \frac{\Delta R}{R},
\]

Fechner [6] concluded about a logarithmic relation between the magnitude of sensation \( S \) and the magnitude of the stimulus \( R \).

The second method, created by Stevens [7], is based on the assumption that the relative change in the sensation magnitude \( \Delta S/S \) depends on the relative change in stimulus \( \Delta R/R \) by a linear relationship, where the constant \( a \) depending on the modality of the stimulus can be either less or greater than a unit:

\[
\frac{\Delta S}{S} = a \cdot \frac{\Delta R}{R}.
\]

From relation (3), for all values of \( a \neq 1 \), i.e. other than a unit, there is a power relationship between the sensation \( S \) and the stimulus \( R \):

\[
\frac{S}{S_0} = (\frac{R}{R_0})^a.
\]

When \( a=1 \), which corresponds to the Fechner hypothesis, there is a logarithmic relation between the sensation magnitude \( S \) and that of the stimulus \( R \) (3).

For the stimuli that are far from the thresholds of perception and pain, the formula (4) was confirmed experimentally by Stevens [7]. For many modalities of sensations, the values of \( k \) and \( a \) have been measured and introduced into the reference manuals (for example, [8]).

A common shortcoming of both methods regarding the diagnosis and rehabilitation is that the subjects are aware of the emerging sensation and the verbal expression of the subjective assessment of its magnitude.

This technique is cumbersome, takes time and does not guarantee an objective verbal evaluation of the subjective feeling.

In [9], we demonstrated a fundamentally different all-instrumental method for quantifying the perception in real time; the proposed approach is free from subjectivism and verbalization. It is based on determining the dynamic error occurred in controlling a real or virtual (computer-generated) stimulus. In this case, the quantitative dependence of the dynamic error on a physical parameter of the external stimulus is used to characterize the sensory analyzer. This approach is based on a simple fact noted by Sechenov [10] that a human makes use of sensations as feedback for controlling his/her functions, e.g., movements.

Description and rationalization of the using method

The rationale for quantifying the perception was obtained from experiments with real stimuli: “load weight” [11, 12] and “tonal sound pitch” [13–15]. Thus, we measured micro-fluctuations of the shoulder biceps muscle force (with its antagonist — triceps muscle relaxed) needed to balance a load held on the distal part of the forearm bent at the elbow joint. We found that the average value of force fluctuations \( f_{\text{fluct}} \) changed with changes in the mass \( m \) of the load [11, 12]. In addition, these changes in the average fluctuation \( f_{\text{fluct}} \) in
response to a change in the load mass $\Delta m$ obey relation (3) with an inaccuracy of 10%:

$$\frac{\Delta f_{\text{fluct}}}{f_{\text{fluct}}} = \beta \frac{\Delta m}{m}.$$  

(5)

The coefficient $\beta$ had a value of 2.0–2.5 in different subjects (10 people).

The relation (5) obtained by solely physical measurements and relation (3), obtained by a subjective psychophysical assessment, differ from each other only by the denoting symbols and, possibly, by the values of $\alpha$ and $\beta$. Therefore, the physical measurement of fluctuations (inaccuracies) of the muscle force is the objective measurement of perception in real time. Thus, the error of control is what creates the sensation and perception of gravity and, apparently, in this context, they can be simply identified with each other.

This interpretation of errors in controlling a stimulus was confirmed by studies on fluctuations of the frequency of a tonal sound generated by a subject’s voice [13, 15] or a sound generator [14, 15]. In the latter case, the frequency was controlled by changing the distance between a fixed coordinate plane and a contactless manipulator freely movable by the subject’s hand (Figure 1 (a)).

This distance was measured using an ultrasonic locator and then transformed into a signal (the stimulus) that controlled the sound generator frequency. Since the subject’s hand did not lean on any support, he/she was only able to change the sound frequency based solely on hearing, i.e., the same way as a natural voice production. Under both conditions (the voice and the generator), the measured frequency fluctuation $\Delta f$ was found proportional to the reproduced frequency $f$. This relation corroborates with formulas (1) and (3) and supports the idea that the relative error in controlling a stimulus (which necessarily implies the existence of feedback) precisely characterizes the quality of perception. In singing a tonal sound, the $\Delta f/f=0.15$. When a tonal sound is produced by a controlled generator, the $\Delta f/f$ depends on the coefficients $K$ (Hz/mm), which links the manipulator movement with a change in the frequency of the sound generator. With $K=0.17f$, the value of $\Delta f/f$ from the generator becomes the same as that produced by voice.

Thus, there were no fundamental differences in the quality of voice and motor control of the sound pitch. This could be expected because the production of a tonal sound using a motor act is the usual work of musicians playing string and wind instruments or thereminoxx, which can all be considered as controlled sound generators with feedback.

Later, we used the contactless manipulation technique to objectively measure the perception of not only sound but also visual stimuli of various types.

The hardware-software complex Hand tracker

The above rationale was used for the development of a manual tracking system [16, 17] that allowed us to measure in real time the quantitative parameters of perception of audial and visual stimuli.

This device makes it possible to form a virtually unlimited variety of visual stimuli and modes of their presentation. It implements the following types of visual stimuli: homothetic compression–extension (change in size) of flat figures; linear compression–extension of flat figures along any chosen axis; rotation of flat figures; movement of flat figures; change in color of the background figures; compression–extension of a straight line or segment of a curve; rotation of a line segment or curve segment; movement of a line segment or curve segment.

The proposed HSC Hand tracker also provides for formation of a wide variety of audial stimuli and modes of their presentation: a tonal or noise sound varying in its volume; a frequency-varying tonal sound; a sequence of paired sound pulses with a variable interpulse interval and a variable intensity; a binaural sequence of sound

---

**Figure 1. A contactless manipulator used to control the parameters of audial (a) and visual (b) stimuli (schematic representation) [1]**
pulses with a variable repetition frequency and a variable
interaural delay (laterometry); a binaural sequence of
sound pulses with a variable repetition period and a
content that can be home-made or quoted from external
sources.

At the choice of the operator, the subject can control
any stimuli using a contactless manipulator (Hand
tracker itself), a joystick, arrows, assigned keyboard
keys, a mouse, or a touchpad panel. In the HSC, the
operator sets the factors for converting the magnitude of
the control action (for example, moving the manipulator
or the joystick) into the size of the change in the stimulus
parameter. The background color and the outline of
the figures (color, thickness) are also assigned by the
operator. The basic stimulus can be a figure, a line or
a sound from the library of stimuli or a quote from an
external source.

The operator creates a test scenario by assigning a
rule for changing the parameters of images and sounds.
The rule can be specified by a formula or a discrete
numerical series.

Each of the created scenarios, i.e. a set of parameters
and modes of stimulus presentation, can be saved in
the database as a configuration and called up for later
experiments. The results of each experiment, as well as
the subjects’ personal data (in unlimited numbers), can
also be stored in the database.

In addition to studying the perception of individual
visual and audial stimuli, the HSC allows one to study
intermodal interaction, since audial and visual stimuli
can be presented simultaneously so to let the subject
control both of them within the same manipulation. This
set up provides for the option to compare errors in the
perception of sounds and images, both individually and
together.

The HSC Hand tracker provides for the possibility of
simultaneous independent control of several (up to
eight) parameters of audial and visual stimuli. Therefore,
control over several stimulus parameters or several
combinations of these parameters can be handled by
two or more subjects; this would make it possible to
compare the quality of their perceptions and study their
collective behavior in controlling the stimuli. This means
that the HSC Hand tracker, developed as a human-
machine perception interface, can also help creating a
human-human perception interface and harmonizing the
subjective perception images of interacting subjects.

The proposed method includes a number of classic
psychological tests.

Example of use: studying the perception
of the “angle of segment orientation”
visual stimulus

The subject was tasked with using a manipulator
(Figure 1 (b)) to control the orientation of a segment on
the computer monitor so that the segment always stayed
parallel to the model segment presented by the operator
on the same screen. Then, deviations of the model
segment were measured.

The operator presented three series of stimuli:
by increasing from –90 to +90 angular degrees in
increments of 22.5 angular degrees, then by decreasing
and also by a random order. The subject maintained
control over each angle for 5 s. Every 0.1 s, a digitized
measurement of the segment orientation angle was
made. Seventy people who considered themselves
healthy took the tests once; 10 of them passed the test
cycle 4 times, with one test per day for a week.

Figure 2 (a) shows the results of a typical test.
Along the circumference, there are miniature graphs
showing the time variation of the angular orientation
of the test segment for each orientation angle of the
model segment. In the center, there is a diagram (in
the polar coordinates) showing the dependence of the
root-mean-square dynamic deviation (the error) of the

Figure 2. A typical result obtained in one subject (a) and the averaged orientation diagram for ten subjects (b)
Along the circumference, the angles of orientation of the model segment are shown; the radial coordinate indicates the mean
square dynamic error of the angular orientation of the controlled segment with respect to the orientation of the model (in angular
degrees) [1]
controlled segment inclination angle on the model segment inclination angle. According to the results, individual orientation charts may vary on different days. Therefore, to assess the present state of the analyzer, the measurements must be made in real time. The diagram shows that with the least error was made when the subject worked with the vertical and horizontal segments. This result was true for the orientation diagram averaged for the 10 subjects who underwent four test cycles (Figure 2 (b)).

Measuring the individual orientation diagrams (the errors) opens up the possibility for evaluating the parameters of the neural network that functions as a primary recognizer. In Figure 3, the experimental and model diagrams are compared: those are diagrams of dynamic errors of control over the segment inclination angle.

Diagram in Figure 3 (b) is shown not in angular degrees but in arbitrary units, since there was no sufficient experimental data for the absolute calibration of the radial axis. However, we successfully used this data to identify the parameters of inter-neuronal connections and their excitation thresholds [18–20]. Interestingly, the discrepancy between the experimental and theoretical diagrams as a function of the magnitude of the receptive field (the number of neurons comprising the receptive field) has a minimum. In this model, the optimum performance of the primary recognizer is achieved for a network of approximately $10^2$ neurons, which work well with the number of neurons in the real recognition column [21].

Example of use: studying the color discrimination in the “figure on the background” visual stimulus in patients with focal brain lesions

To study the visual perception in patients with focal brain lesions, the HCS was used in the “campimetry” mode, i.e. for the measurement of color discrimination thresholds. As patients with brain lesions (and motor disorders) had difficulties in controlling the presented stimuli using a contactless manipulator, we could not measure the thresholds of color discrimination in dynamics. Therefore, we opted to measure the static thresholds of color distinction. In this setting, we presented colored shapes against a colored background; the patient was asked to control the shape color using buttons on the computer keyboard. We used three types of colored figures with different semantic content: a) geometric figures — a circle, a square, and a rectangle; b) images of animals and fruits; c) letters of the Russian alphabet.

The background color and the color of the stimuli were specified with the help of the HLS (hue, lightness, saturation) model [22]. The color tone $H$ of the background was set by a scale from 0 to 240 arb. units and increments of 60 units with fixed saturation levels $S=220$ units and lightness $L=100$ units.

At first, the subject was presented with an image where the background color and the figure color were identical. By pressing the keyboard arrow, the subject increased the value of $H$ by increments of 1 arb. unit until he/she detected the figure. In parallel, when the subject pressed the Enter key, the detection threshold of the color tone $H$ was recorded by the system; if the figure was recognized incorrectly, the system recorded this result as an error. Either of the two results was reported by an arising icon. With a new keystroke, the subject was presented with a new image with a different color background value $H$.

The primary data were collected from a group of neurological patients with confirmed diagnoses of “acute ischemic cerebrovascular accident”, “acute cerebral hemorrhagic disorder”, and “status after endoscopic removal of a left lateral ventricle cyst” (5 men, 4 women, 23–74 years old). The control group included conditionally healthy students from the Department of Psychophysiology of the National Research Lobachevsky State University of Nizhny Novgorod (5 males, 25 females, age 18–41 years).

The inclusion criteria for the study group were: a stable condition (satisfactory or moderate-severe), the absence of inflammatory diseases, the absence of severe somatic diseases in the off-remission stage at the time of the study, the proven ability of verbal contact, the understanding of relevant conversations, and the absence of severe motor deficit.

The study was conducted in accordance with the Helsinki Declaration (2013). An informed consent was obtained from each patient.

To study a relation between the semantic content of the stimulus and the thresholds of color discrimination, the Student’s t-criterion was used to compare the data from the study and control groups (see the Table) presented with colored stimuli that were uniform in color tone but different in semantics.
According to the Student’s criterion, there were statistically significant ($p<0.05$) differences in the color discrimination thresholds between the study and control groups in respect to each type of color stimuli. At the same time, the color discrimination thresholds did not significantly differ between the subjects of the study group even if the semantics of the stimuli varied. In the control group, for color tone #60 (green, [22]), there was a significant ($p<0.05$) decrease in the color discrimination threshold for Russian letters images as compared to object images.

Thus, the proposed approach made it possible to reveal a qualitative abnormality of visual perception in patients with local brain lesions in two ways:

- the color discrimination thresholds in patients are higher than normal for all types of stimuli that differ in the semantic content;
- the semantics of the visual stimulus does not change the color discrimination threshold for any of the tested colors.

**Conclusion**

The examples of practical use of the all-instrumental quantification of perception, implemented in the form of the hardware-software complex Hand tracker, allow us to conclude that this method can be effectively used for laboratory and clinical studies on the perception of stimuli of various modalities. In this case, the measurement procedure itself is also a training one.

Unlike the traditional methods, the hand tracking has a much higher response rate and makes it possible to obtain the quantitative data on perception that have been unavailable previously.

The presence of several parallel channels for presenting the stimuli and controlling their parameters opens up the possibilities for objectively studying the perception of multimodal stimuli and the interaction between team members that collectively control the parameters of a virtual stimulus (for example, the size and color of an object). Thus, the hardware-software complex Hand tracker can be interpreted as a human-human sensory interface, which provides the interacting subjects with the instrument to coordinate their subjective sensory images of the presented objects.

The explosive growth in the number of smartphone users provides the opportunity to develop a hand tracking method for epidemiological studies of perception (primary cognitive functions), taking into account regional, age, gender, social and other stratification of populations. The presence of embedded sensors in smartphones allows one to create special software applications [17, 23] that implement visual and audial perceptual tests. Moreover, some widely used applications and smartphone options can be modified into a perceptual test involuntary performed by the user (at his/her discretion). This development can produce surprising results, just as the creation of a network of meteorological stations made it possible to discover the existence of cyclones and anticyclones.

**Acknowledgments.** The author expresses his deep appreciation of S.A. Polevaya and M.M. Cirkova for their contribution to the research described in the article.

**Research funding.** This work was supported by grants from the Russian Foundation for Basic Research No.13-04-12063, No.05-08-33526-a, No.97-06-80286-a, No.00-06-80141-a and project No.0035-2014-0008 as part of the state-supported research program in the Institute of Applied Physics of the Russian Academy of Sciences.

**Conflict of interest** does not exist.

**References**

