Video Article

One Dimensional Turing-Like Handshake Test for Motor Intelligence

Amir Karniel[1], Guy Avraham[1], Bat-Chen Peles[1], Shelly Levy-Tzedek[1], Ilana Nisky[1]

[1]Biomedical Engineering, Ben-Gurion University

Correspondence to: Amir Karniel at akarniel@bgu.ac.il

Abstract

In the Turing test, a computer model is deemed to "think intelligently" if it can generate answers that are not distinguishable from those of a human. However, this test is limited to the linguistic aspects of machine intelligence. A salient function of the brain is the control of movement, and the movement of the human hand is a sophisticated demonstration of this function. Therefore, we propose a Turing-like handshake test, for machine motor intelligence. We administer the test through a telerobotic system in which the interrogator is engaged in a task of holding a robotic stylus and interacting with another party (human or artificial). Instead of asking the interrogator whether the other party is a person or a computer program, we employ a two-alternative forced choice method and ask which of two systems is more human-like. We extract a quantitative grade for each model according to its resemblance to the human handshake motion and name it "Model Human-Likeness Grade" (MHLG). We present three methods to estimate the MHLG. (i) By calculating the proportion of subjects' answers that the model is more human-like than the human; (ii) By comparing two weighted sums of human and model handshakes we fit a psychometric curve and extract the point of subjective equality (PSE); (iii) By comparing a given model with a weighted sum of human and random signal, we fit a psychometric curve to the answers of the interrogator and extract the PSE for the weight of the human in the weighted sum. Altogether, we provide a protocol to test computational models of the human handshake. We believe that building a model is a necessary step in understanding any phenomenon and, in this case, in understanding the neural mechanisms responsible for the generation of the human handshake.

Video Link

The video component of this article can be found at https://www.jove.com/video/2492/

Protocol

1. Preparing the System

1. Hardware requirements:
    - Two PHANTOM desktop robots by SensAble Technologies, Inc.
    - 2 Parallel cards.
    - Minimum system requirements: Intel or AMD-based PCs; Windows 2000/XP, 250 MB of disc space.

2. Software requirements:
    - Driver-
      Download drivers from the SensAble technologies website http://www.sensable.com according to the computer operating system.
    - H3DAPI-
      Download the H3DAPI source code and install according to the instructions displayed in the installation walkthrough section in http://www.h3dapi.org/modules/mediawiki/index.php/H3DAPI_Installation
      We updated the code to fit our requirements that were not met in the existing code. The updated files can be downloaded from the handshake tournament website http://www.bgu.ac.il/~akarniel/HANDSHAKE//index.html
      Each file should be placed in the appropriate folder as described in the website above.
      After performing the changes, download CMake and compile the code. The compiling instructions can be found in http://www.h3dapi.org/modules/mediawiki/index.php/H3DAPI_Installation
    - X3D and python
      - Download python 2.5 from http://www.python.org/download/releases/2.5.5/
      - Download the python and x3d codes from the handshake website and place the code files in a dedicated directory, for instance: "C:\codeDirectory"
      - Each handshake force model should be written in a separate python file. See Figure 1 for an example of a spring force model.

    - psignifit toolbox version 2.5.6 for Matlab, available at http://www.bootstrap-software.org/psignifit/

```
#This is an example of a force function
#of a linear spring

#input:
# x - vector of position samples
# t- vector of time samples

#output:
# f - force

def function (x,t):

    springConst = 70

    f= (x[0]-x[-1])*springConst

    return f
```

Figure 1. Force function in python. An example of a spring force model for a handshake

2. Experimental Protocol

- Open the command window and change to the python and x3d code directory by typing: *cd C:\codeDirectory*.
- Create a folder with the subjects' names in *C:\codesDirectory*.
- In order to run the experiment type: h3dload *code_name.x3d*.
- Create a new random file by typing: random_file_name.txt. The random file defines the order in which the different handshakes appear.
- Enter the subjects' names exactly as in the previously created folder.
- Following the original concept of the classical Turing test, each experiment consists of 3 entities: a human, a computer, and an interrogator. Two subjects (human and interrogator) each hold the stylus of one Phantom haptic device and generate handshake movements. They are asked to follow the instructions that appear on the screen (e.g., :"press Page Up for the first handshake"), for handshake forces to be applied. In all of the following methods, each trial consists of 2 handshakes, and the subjects are required to compare between them. The computer is a simulated handshake model which generates a force signal as a function of time and the one dimensional hand position and its derivatives.
  (1)$F_{model}(t) = \Phi[x(t),t]$ $0 \le t \le T$
  $F[x,t]$ stands for any causal operator, e.g., non-linear time-varying mechanical model of the one dimensional stylus movement, and T is the duration of the handshake. In the current study T=5 seconds.

1. Conducting the "pure" test and calculating the model human likeness grade $MHLG_p$
   The experiment starts with 12 practice trials in which all the handshakes (n=24) are human, such that the subjects simply shake hands with each other through the telerobotic system. The purpose of these practice trials is to enable the participants to be acquainted with a human handshake in the system.
   In the experiment, we compare four computer models. Each experimental block consists of 4 trials in which we compare the 4 tested models to a human handshake. One of the handshakes in each trial is an interaction with a force generated from one of the four models (a computer), and the other is with a human (the second subject). Therefore, the subjects function as both humans and interrogators. The order of the trials within each block is random and predetermined. Each experiment consists of 10 blocks, such that each computer handshake is repeated 10 times. An initial unanalyzed block is added for general acquaintance with the system and the task.
   For each model, the proportion of handshakes in which the subject chooses the model handshake over the human handshake as more human-like is calculated, to provide a value which is 0.5 when the model is indistinguishable from a human. We multiply this value by two in order to obtain the $MHLG_p$, such that MHLG=0 is clearly non-human like and MHLG=1 means that the tested handshake is indistinguishable from the human handshake.
2. Conducting the "weighted human-model" test and calculating the model human likeness grade $MHLG_w$
   In this protocol, there is only one interrogator subject. The other subject functions as the human entity in the handshakes.
   The experiment starts with 30 practice trials in which the interrogator experiences one human handshake and one computer handshake in each trial. In the end of the trial they are asked to choose which of the 2 handshakes was the human handshake. If they succeed, the screen displays "Correct!", and if they didn't choose the right handshake, a "Wrong!" message appears.
   After the practice block the experiment is conducted as follows: A trial consists of two handshakes. In one of the handshakes- the stimulus - the interrogator interacts with a combination of forces that comes from the human and a computer handshake model.
   (2) $F = \alpha_{stimulus} \cdot F_{human} + (1-\alpha_{stimulus}) \cdot F_{stimulusModel}$
   $\alpha_{stimulus}$ is equally distributed from 0 to 1, e.g:
   $\alpha_{stimulus}= \{0, 0.142, 0.284, 0.426, 0.568, 0.710, 0.852, 1\}$
   The other handshake - the reference - is a fixed combination of forces generated from the human and a reference model:
   (3) $F = \alpha_{reference} \cdot F_{human} + (1-\alpha_{reference}) \cdot F_{referenceModel}$; $\alpha_{reference}=0.5$
   At the end of each trial the interrogator is requested to choose the handshake that felt more human-like.
   In each experiment we compare two test models and one base model.
   Each experimental block consists of 24 trials comprising each of the linear combinations of the stimulus and the human (eq. 2) for each of the 3 model combinations:
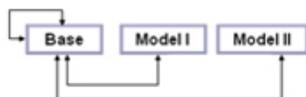
The order of the trials within each block is random and predetermined. Each experiment consists of 10 blocks, such that each combination is repeated 10 times. An initial unanalyzed block is added for general acquaintance with the system and the task.

We fit a logistic psychometric function[1] to the answers of the interrogator using the psignifit toolbox version 2.5.6 for Matlab, available at http://www.bootstrap-software.org/psignifit/ , with a constrained maximum likelihood method for estimation of the parameters, and find confidence intervals by the bias-corrected and accelerated (BCa) bootstrap method. The curve displays the probability of the interrogator to answer that a stimulus handshake is more human-like, as a function of $\alpha_{stimulus} - \alpha_{reference}$. The point of subjective equality (PSE) is extracted from the 0.5 threshold level of the psychometric curve, indicating the difference between the $\alpha_{stimulus}$ and $\alpha_{reference}$ for which the handshakes are perceived to be equally human like. The PSE is used for calculating the $MHLG_w$ according to:

(4) $MHLG_w = 0.5 - PSE$

A model which is perceived to be as human-like as the reference model yields the $MHLG_w$ value 0.5. The models that are perceived as the least or the most human-like possible, yield $MHLG_w$ values of 0 or 1, respectively.

3. Conducting the "added noise" test and calculating the model human likeness grade $MHLG_n$

    Similarly to the weighted model-human test, there is only one interrogator subject. The other subject functions as the human entity in the handshakes. The practice block is also the same as in the previous method.

    After the practice, in one of the 2 handshakes - the stimulus - the interrogator interacts with a computer handshake model.

    The other handshake - the reference - is a force generated from a combination of the human and white noise with a frequency range filtered according to the frequencies that appear in the human handshake.

    (5) $F = \alpha \cdot F_{human} + (1-\alpha) \cdot F_{noise}$;

    α is equally distributed from 0 to 1, e.g:

    $\alpha = \{0, 0.142, 0.284, 0.426, 0.568, 0.710, 0.852, 1\}$

    At the end of each trial the interrogator is requested to choose the handshake that felt more human-like.

    The experiment is built in the same way the weighted model-human test mentioned above is built. However, while the calibration in the weighted human-model test is performed by comparing different combinations of a base model to a fix combination of itself, in this experiment the base model is substitute by the noise.

    The PSE is extracted from the psychometric curve and defines the $MHLG_n$

    (6) $MHLG_n = 1 - PSE$

    The models that are perceived as the least or the most human-like possible, yield $MHLG_n$ values of 0 or 1, respectively.

3. Representative Results:

Figure 2 demonstrates the results of one subject for each of the 3 methods. The tested models in all three experiments are 2 viscoelastic models- KB1: spring K=50 N/m, damper B=2 Ns/m; KB2: spring K=20 N/m, damper B=1.3 Ns/m. In the weighted model-human test, the $MHLG_w$ is evaluated by comparing each of the tested models to the elastic base model K=50 N/m.
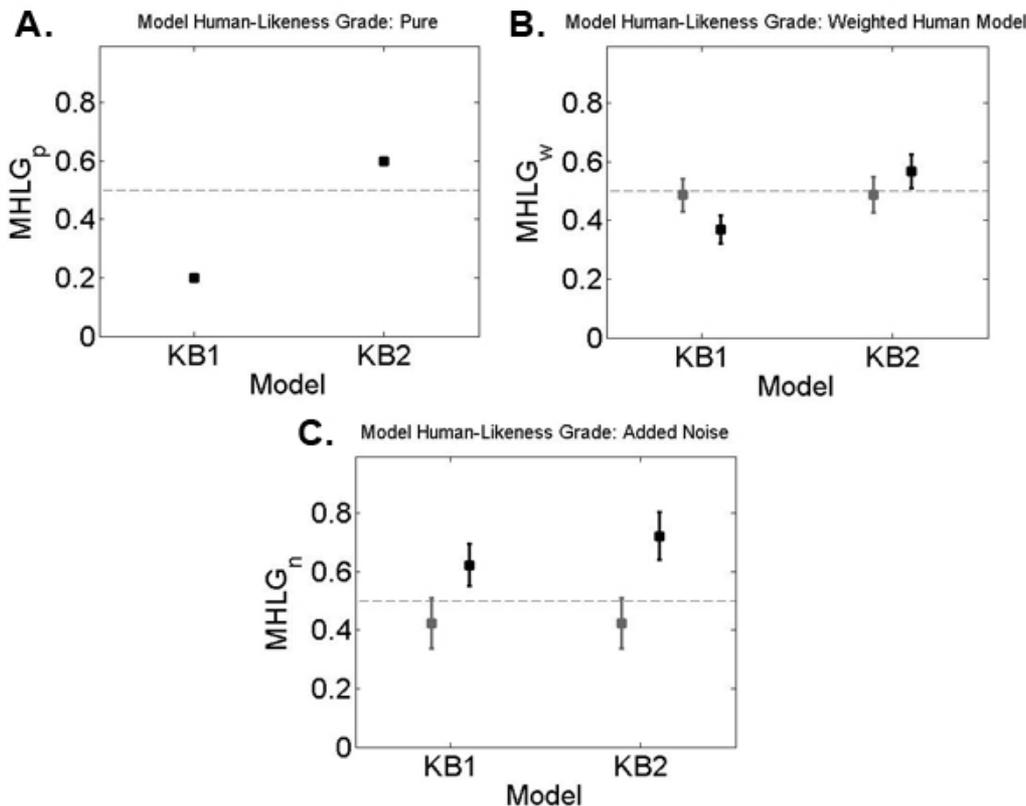

Figure 2. The MHLG values of two viscoelastic models according to the "pure" test protocol (a), the "weighted model-human protocol" (b) and the "added noise" protocol (c). The error bars in (b) and (c) represent the psychometric curves' confidence intervals. The black bars represent the MHLG grades for the models, and the gray bars represent those of the base model in (b) and the noise in (c).

The results demonstrate that the viscoelastic model KB2 is perceived as more human like than the other viscoelastic model KB1 using all three evaluation methods.

Discussion

We have presented a new protocol for a forced-choice Turing-like handshake test administered via a simple telerobotic system. This protocol is a platform for comparing artificial handshake models, rather than a platform for determining absolute human likeness. This protocol was presented in a few conferences[2-5]

We have shown here that this test is helpful in finding the parameters of the passive characteristics of motion that provide the most human-like feeling. It can be used in further studies in order to develop a model for a handshake which will be as human-like as possible. We will employ this platform in the first Turing-like handshake tournament which will take place in summer 2011 [[http://www.bgu.ac.il/~akarniel/HANDSHAKE/index.html]], where competing models will be graded for their human likeness. The ultimate model should probably consider the nonlinearities and time-varying nature of human impedance[21], mutual adaptation with the interrogator and many other aspects of a natural human handshake which should be tested and ranked using this forced-choice Turing-like handshake test.

The proposed test is one dimensional and performed via a telerobotic interface, and therefore is limited: it hides many aspects of the handshake such as tactile information, temperature, moisture, and grasping forces. Nevertheless, in several studies a telerobotic interface was used for exploring handshakes[6-11]and other forms of human-human interaction[12]. In addition, in this version of the test, we didn't consider the duration of the handshake, the initiation and release times, its multi-dimensional nature and the hand trajectories before and after the physical contact. There are also many types of handshakes depending on gender and culture of the person[13-14] and therefore, one cannot expect to generate a single optimal human-like handshake model. Nevertheless, we believe that the simplicity of the proposed test is an advantage, at least at this preliminary stage of the study. Once the key features of such one dimensional handshake are properly characterized we can move on to consider these limitations and extend the test accordingly.

It should be noted that a Turing-like handshake test could be reversed, with the computer instead of the person being asked about the identity of the other party. In this framework, we consider the following reverse handshake hypothesis: the purpose of a handshake is to probe the shaken hand; according to the reverse handshake hypothesis, the optimal handshake algorithm - in the sense that it will be indistinguishable from a human handshake - will best facilitate the discrimination between people and machines. In other words, the model will yield the best handshake such that an appropriately tuned classifier can distinguish between human and machine handshakes.

If the reverse handshake hypothesis is indeed correct it yields a clinical application for our test: identifying motor impairments in people suffering from various neurological motor-related disorders, such as Cerebral Palsy (CP). Previous studies have shown differences in the kinematic parameters between CP patients and healthy subjects when performing reaching movements[15-16]. We recently showed that the characteristics of movements differ between healthy individuals and individuals with CP when shaking hands through a telerobotic system[4]. These findings strengthen our claim that people with motor impairments can be distinguished from healthy people by examining and exploring the handshake movement of each individual. One should also note that the test discussed herein is a perceptual test and recent studies distinguish between perception and action[17-20]. Future studies should explore three versions of the test in order to accurately assess the nature of the human-like handshake: (1) a psychometric test of the perceived similarity; (2) a motor behavior test (motormetric test) that will explore the motor reaction of the interrogator which may differ from his/her cognitively perceived similarity; (3) an ultimate optimal discriminator which attempts to distinguish between human and machine handshakes based on the force and position trajectories.

In general terms, we assert that understanding the motor control system is a necessary condition for understanding the brain function, and that such understanding could be demonstrated by building a humanoid robot indistinguishable from a human. The current study focuses on handshakes via a telerobotic system. We assert that by ranking the prevailing scientific hypotheses about the nature of human hand movement control using the proposed Turing-like handshake test, we should be able to extract salient properties of human motor control or at least the salient properties required to build an artificial appendage that is indistinguishable from a human arm.

Disclosures

No conflicts of interest declared.

Acknowledgements

References

1. Wichmann, F. & Hill, N. The psychometric function: I. Fitting, sampling, and goodness of fit. Perception & Psychophysics 63, 1293 (2001).
2. Avraham, G., Levy-Tzedek, S. & Karniel, A. in Computational Motor Control Workshop (Be'er-Sheva, Israel, 2009).
3. Avraham, G., Levy-Tzedek, S. & Karniel, A. in Neural Control of Movement (Florida, USA, 2010).
4. Avraham, G., Levy-Tzedek, S., Peles, B. C., Bar-Haim, S. & Karniel, A. in Computational Motor Control Workshop (Be'er-Sheva, Israel, 2010).
5. Karniel, A., Nisky, I., Avraham, G., Peles, B. C. & Levy-Tzedek, S. in EuroHaptics (Amsterdam, Netherlands, 2010).

6. Jindai, M., Watanabe, T., Shibata, S. & Yamamoto, T. in The 15th IEEE International Symposium on Robot and Human Interactive Communication (Hatfield, UK 2006).
7. Kasuga, T. & Hashimoto, M. in Internatioanl Conference on Robotics and Automation. IEEE (Barcelona, Spain 2005).
8. Ouchi, K., Hashimoto, S. Handshake Telephone System to Communicate with Voice and Force. Proceedings of the 6th IEEE International Workshop on Robot and Human Communication; 1997 Sept 29-Oct 1, 466-471 IEEE (Sendai, Japan 1998).
9. Bailenson, J. & Yee, N. Virtual interpersonal touch and digital chameleons. Journal of Nonverbal Behavior 31, 225-242 (2007).
10. Miyashita, T. & Ishiguro, H. Human-like natural behavior generation based on involuntary motions for humanoid robots. Robotics and Autonomous Systems 48, 203-212 (2004).
11. Wang, Z., Lu, J., Peer, A., & Buss, M. in EuroHaptics (Amsterdam, The Netherlands 2010).
12. IiZuka, H., Ando, H. & Maeda, T. in proceedings of AsiaGraph (2010).
13. Chaplin, W., Phillips, J., Brown, J., Clanton, N. & Stein, J. Handshaking, gender, personality, and first impressions. Journal of Personality and Social Psychology 79, 110-117 (2000).
14. Stewart, G., Dustin, S., Barrick, M. & Darnold, T. Exploring the handshake in employment interviews. Journal of Applied Psychology 93, 1139-1146 (2008).
15. Van Der Heide, J., Fock, J., Otten, B., Stremmelaar, E. & Hadders-Algra, M. Kinematic characteristics of reaching movements in preterm children with cerebral palsy. Pediatric research 57, 883 (2005).
16. Ronnqvist, L. & Rosblad, B. Kinematic analysis of unimanual reaching and grasping movements in children with hemiplegic cerebral palsy. Clinical Biomechanics-Kidlington 22, 165-175 (2007).
17. Goodale, M. & Milner, A. Separate visual pathways for perception and action. Trends in neurosciences 15, 20-25 (1992).
18. Ganel, T. & Goodale, M. Visual control of action but not perception requires analytical processing of object shape. Nature 426, 664-667 (2003).
19. Pressman, A., Nisky, I., Karniel, A. & Mussa-Ivaldi, F. Probing Virtual Boundaries and the Perception of Delayed Stiffness. Advanced Robotics 22, 119-140 (2008).
20. Nisky, I., Pressman, A., Pugh, C. M., Mussa-Ivaldi, F. A. & Karniel, A. in EuroHaptics 213-218 (Amsterdam, Netherlands, 2010).
21. Karniel. Computational Motor Control, in Encyclopedia of Neuroscience, M.D. Binder, N. Hirokawa, and U. Windhorst, Editors., Springer-Verlag: Berlin 832-837 (2009).