EuroHaptics 2010, Amsterdam, July 8-10, 2010

A Turing-like Handshake Test for Motor Intelligence

Amir Karniel, Ilana Nisky, Guy Avraham, Bat-Chen Peles,
and Shelly Levy-Tzedek

The Computational Motor control Laboratory, Department of Biomedical Engineering,
Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel
{akarniel, nisky, guyavr, pelesb, shelly} @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. This test is limited to the linguistic aspects of machine intelligence. A salient function of the brain is the control of movement, with the human hand movement being 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, artificial, or a linear combination of the two). Instead of asking the interrogator whether the other party is a person or a computer program, we employ a forced-choice method and ask which of two systems is more human-like. By comparing a given model with a weighted sum of human and artificial systems, we fit a psychometric curve to the answers of the interrogator and extract a quantitative measure for the computer model in terms of similarity to the human handshake.

Keywords: Turing test, Human Machine Interface, Haptics, Teleoperation, Motor Control, Motor Behavior, Diagnostics, Perception, Rhythmic, Discrete

1 Introduction

As long ago as 1950, Turing proposed that the inability of a human interrogator to distinguish between the answers provided by a person and those provided by a computer would indicate that the computer can think intelligently [1]. The so-called "Turing test" has inspired many studies in the artificial intelligence community; however, it is limited to linguistic capabilities. We argue that the ultimate test must also involve motor intelligence – that is, the ability to physically interact with the environment in a human-like fashion - encouraging the design and construction of a humanoid robot with abilities indistinguishable from those of a human being. However, this ultimate Turing-like test for motor intelligence involves an enormous repertoire of movements. In this paper we present the first step towards the ultimate test by using the handshake test proposed by G.E. Loeb as the basis for our test [2]. The handshake is of interest not merely as a reduced version of the ultimate humanoid test but also due to its bidirectional nature, in which both sides actively shake hands and explore each other. Moreover, motor control research has concentrated on hand movements [3], generating a variety of hypotheses which could be applied to generate a humanoid handshake. Last but not least, the greatest progress in telerobotic technologies involves arm movements. The telerobotic interface is necessary to grant the human-computer discrimination significance, much as the teletype was necessary to hide the computer from the questioning human in the original Turing test.

Handshaking has been discussed in the social context [4, 5] but the development of artificial handshake systems is still in its infancy [6-10]. Nevertheless, research of human motor control has generated many theories as to the nature of hand movements, and the proposed Turing-like handshake test can be useful in identifying the aspects of these theories that are essential for producing a human-like handshake movement. In general terms, we assert that a true understanding of the motor control system could be demonstrated by building a humanoid robot that is indistinguishable from a human. Therefore, a measure of our distance from such demonstration could be most useful in evaluating current scientific hypotheses and guiding future neuroscience research.

We first describe the proposed one dimensional forced-choice Turing-like handshake test in section 2; we describe our experimental protocol in section 3; we analyze natural handshake movements in section 4; and demonstrate our experimental results for five simple handshake models in section 5. Finally, we introduce an international handshake test challenge in section 6.

2 The Turing-like Handshake Test

In this study, a Turing-like handshake test is used in order to compare handshake models and extract a quantitative measure of their similarity to a human handshake.

In the classical Turing test, the human interrogator compares the answers of a human and a computer. In the proposed handshake test we followed the original concept of three entities (***a human***, ***a computer***, and ***an interrogator***) and defined a forced-choice protocol. We administered the test with two PHANToM® Desktop haptic devices by SensAble Technology Inc, operated using SenseGraphics H3D API. The ***human*** as well as the ***interrogator*** are each asked to hold the stylus of a haptic device and to generate handshake movements through the telerobotic interface (Fig. 1f). The force feedback to the interrogator is a linear combination of the forces generated by both the ***human*** and the ***computer*** (Fig. 1a-c). The ***computer*** is a simulated handshake model which generates force signal as a function of time and one dimensional position of the hand and its derivatives,

$$\boldsymbol{F}_{\text{model}}(\boldsymbol{t}) = \Phi\left[\boldsymbol{x}\left(\boldsymbol{t}\right),\boldsymbol{t}\right] \qquad \boldsymbol{0} \le \boldsymbol{t} \le \boldsymbol{T}\,, \tag{1}$$

where $\Phi[x,t]$ stands for any causal operator, e.g., non-linear time-varying mechanical model of the one dimensional stylus movement generated by artificial handshake, and T is the duration of the handshake. The force feedback to the ***human*** is generated purely by the ***interrogator***, in order to preserve the mutual and adaptive characteristics of the human handshake movement.

A ***trial*** consists of two handshakes, each lasting 5 seconds. In one of the handshakes in each trial – the ***stimulus*** – the interrogator interacts with a combination of forces that comes from the human and a ***computer*** handshake model:

$$F = \alpha_{stimulus} \cdot F_{human} + (1-\alpha_{stimulus}) \cdot F_{stimulusModel} \quad (2)$$

$\alpha_{stimulus}$= {0, 0.11, 0.22, 0.33, 0.44, 0.55, 0.66, 0.77, 0.88, 1}

The other handshake – the ***reference*** – is a fixed combination of forces generated from the human and a reference model:

$$F = \alpha_{reference} \cdot F_{human} + (1-\alpha_{reference}) \cdot F_{referenceModel} \quad ; \quad \alpha_{reference}=0.5 \quad (3)$$

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

We fitted a psychometric curve to the answers of the interrogator (Fig. 1g) [16]. Such curve describes 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), where the probability to answer that the stimulus is more human-like is at chance level is extracted from the curve and used to calculate a quantitative grade for each model's closeness to a human handshake – Model Human-Likeness Grade (MHLG) - according to:

$$MHLG = PSE + 0.5 \quad (4)$$

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

3 The Handshake Experimental Protocol

In each of two experiments we performed, three ***computer*** handshake models were compared - two tested models and one base model. The comparison was performed by the following four tests: (1) Base vs. Base; (2) Base vs. Test #01; (3) Base vs. Test #02; and (4) Test #01 vs. Test #02.

Each experimental block consisted of 40 trials: 10 linear combinations of the stimulus and the human (eq. 2) for each of the four model combinations mentioned above. As described in section 2 each trial consisted of two handshakes: a stimulus and a reference handshake. The order of the trials within each block was random and predetermined. Each experiment consisted of 11 blocks, the first block was not analyzed and used for general acquaintance with the system and the task. For each test we fitted a psychometric curve and calculated the MHLG from the extracted PSE.

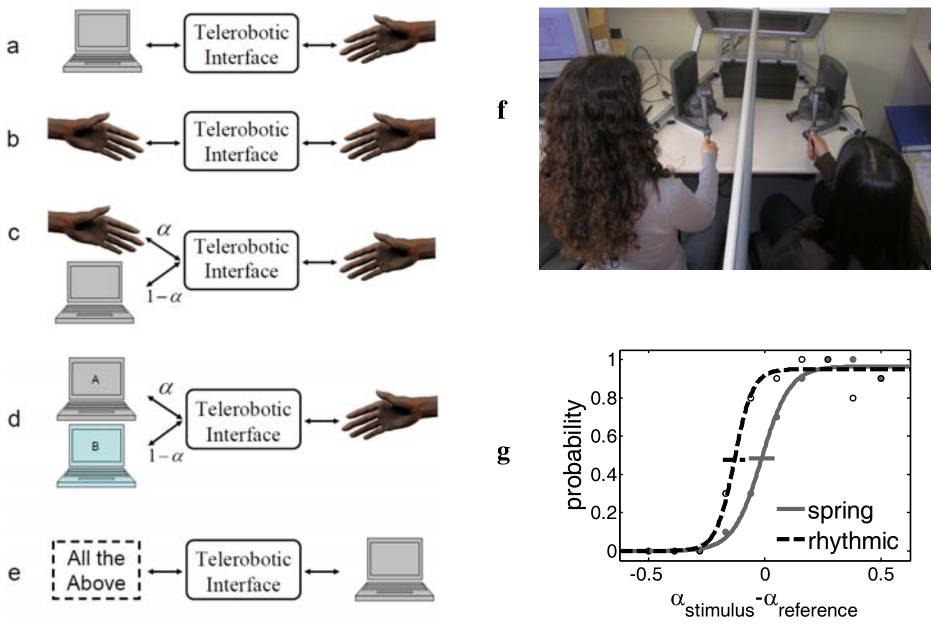


Fig. 1. Illustration of the handshake test. In the basic Turing-like handshake test the human interrogator has to distinguish between a simulated handshake (a) and a natural handshake (b). To extract a quantitative measure for the human-like behavior of simulated handshake we propose a forced-choice method in which the interrogator has to probe combinations of natural and simulated handshakes (c) and answer which one is more human-like. Using the same method one can compare between two simulated handshakes (d). Eventually, one can also consider an artificial interrogator (e) and test the reverse handshake hypothesis asserting that a simulated handshake indistinguishable from a natural handshake would be useful as an interrogator handshake to distinguish between other handshakes. The actual tests were performed using two haptic devices (f). The data was analyzed by fitting psychometric curves to the answers of subjects (g), where the vertical axis represents the probability of the subject to answer that the stimulus handshake felt more human-like. In the base curve (solid gray), both stimulus and reference models where passive springs, whereas in the test curve (dashed black), stimulus is based on a rhythmic handshake model and reference on a passive spring.

4 Smoothness and Harmonicity of a natural handshake

In approaching the challenge of creating a robotic model of the human handshake that will be indistinguishable from an actual human handshake we first assessed two of the quintessential properties of the human handshake: whether it is rhythmic or discrete in nature, and its relative smoothness. We hypothesized that a model which captures these properties will likely be more similar to the human handshake than one that did not account for the rhythmicity and smoothness of the handshake.

We asked five participants to hold the stylus of the haptic device and to generate handshake movements with the experimenter via a telerobotic system and recorded their position over time.

We evaluated the rhythmicity and the smoothness of the human handshake using two metrics: mean squared jerk ratio (MSJR), and harmonicity. The relative smoothness of a movement is obtained by dividing its mean squared jerk by the mean squared jerk of the corresponding maximally smooth movement [11]. This ratio would approach a value of one for a highly smooth movement, whereas a value much greater than one would imply the movement is highly fragmented. Harmonicity of the movement is determined by the features of the acceleration trace around movement reversals, and provides a measure of the harmonic or inharmonic nature of the movement [12]. A harmonicity value approaching one indicates a highly harmonic movement, whereas a value approaching zero implies mechanical energy is dissipated in the vicinity of movement reversal [13]. Detailed descriptions of the methods of calculating these metrics have been provided elsewhere [14, 15].

The average amplitude of the handshake movements was 10.0±1.0cm (mean±SD), and average frequency was 2.5±0.1Hz. MSJR values were close to unity (1.5±0.4), indicating a highly smooth movement. Harmonicity values from all analyzed trials almost uniformly equaled one (0.98±0.02), indicating a highly harmonic nature of the movement.

These results suggest that a model of the human handshake that produces a smooth, harmonic, movement is more likely to be perceived more human-like than one that produces a jerky, discrete movement.

5 Demonstrating the Turing-like Handshake Test

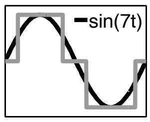
One interrogator participated in two experiments. In each experiment we compared two test models and one base model (see Table 1). The following models were used as computer-generated handshake models:

$$F_1 = Kx \quad F_2 = Kx+Bv \quad F_3 = \sin(7t) \quad F_4 = \mathrm{ceiling}(\sin(7t) - \sin(\pi/2-1)) \qquad \textbf{(5)}$$

Where F_1, a simple passive model, was used as the Base model; F_2, a viscoelastic model, was used in experiment I with different values for the stiffness K and the viscosity B for Test Models #01 and #02; and F_3, F_4 were used in experiment **II** as the rhythmic and the discrete force generators.

Figure 2 displays the MHLG values of the four models that were tested. One can see that the interrogator perceived the rhythmic model and the second viscoelastic model as more human-like than the base model. On the other hand he perceived the discrete model as well as the first viscoelastic model as less human-like than the base model.

Table 1 Experimental protocol parameters

Experiment	Base Model	Test Model #01	Test Model #02
I	K=50 N/m	K=50 N/m; B=2 N·sec/m	K=20 N/m; B=1.3 N·sec/m
II	K=50 N/m	Rhythmic 	Discrete

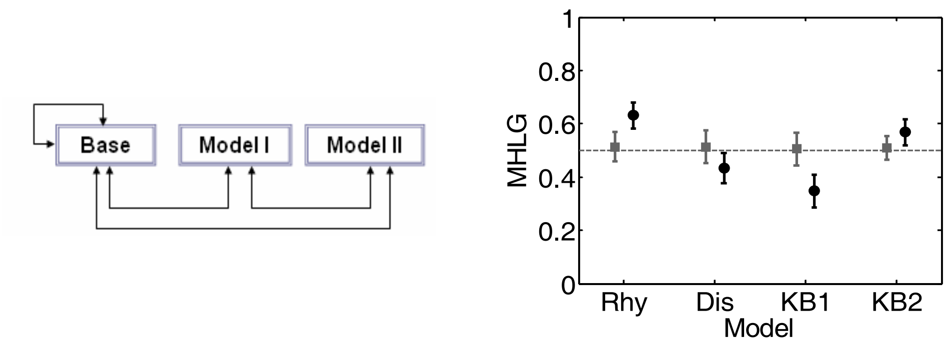


Fig. 2. Demonstrating the Turing-like Handshake test with human, interrogator and three models. In each experiment we compared three models by performing four model comparisons as illustrated in the left panel. The right panel depicts the results of the test using the Model Human-Likeness Grade (MHLG) Plot. The tested models, as appear in Table 1, are: Rhy- rhythmic; Dis- Discrete; KB1- spring K=50, damper B=2; KB2- spring K=20, damper B=1.3. Black circles and gray squares are the MHLG of the test and base models, respectively, each compared to the base model. Error bars are 95% confidence intervals of the MHLG.

6 Discussion

In this study we propose a forced-choice Turing-like handshake test administered via a simple telerobotic system. We demonstrated the proposed handshake test by comparing four type of models: (i) a simple passive spring model (ii) passive viscoelastic models (iii) a rhythmic power source, and (iv) a discrete power source of the same frequency and amplitude (see equation 5, and Table 1). Using the Turing-like handshake test, and the proposed measure of Model Human-Likeness Grade (MHLG), we found that a rhythmic handshake model is perceived as more human-like than a discrete one. This preliminary finding is consistent with our analysis of rhythmicity: the human handshake is highly rhythmic. Moreover, this test is helpful in finding the parameters of the passive characteristics of motion that provide the most human-like feeling. Further studies are required to validate these results on a large group of subjects and, more importantly, to develop a model for a handshake which will be as human-like as possible. Here we provide a platform for comparing handshake models. The four models we propose are far from being indistinguishable from a natural human handshake as we did not consider the nonlinearities and time-varying nature of human impedance, 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.

There are many limitations to the proposed test which is one dimensional and consists of a telerobotic system and therefore hides many aspects of the handshake such as tactile information, temperature, moisture, and grasping forces. Moreover in this version of the test we didn't consider the duration of the handshake, the initiation and release times 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 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 (see Fig. 1e). 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, once a model that best mimics the human handshake is constructed, it itself can then be used to probe handshakes generated by humans and/or other machines.

One should also note that the test discussed herein is a perceptual test and recent studies distinguish between perception and action [17,18]. Future studies should explore three versions of the test in order to accurately assess the nature of human-like handshake: (1) a psychometric test of the perceived similarity; (2) a motor behavior 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.

Acknowledgments. The first author wishes to thank Gerry Loeb for useful discussions about the proposed Turing-like handshake test. This research was supported by the Israel Science Foundation (grant No. 1018/08). SL was supported by a Kreitman Foundation postdoctoral fellowship. IN was supported by a Kreitman Foundation fellowship, and a Clore scholarship.

References

1. Turing, A.M.: Computing Machinery and Intelligence. Mind, A Quarterly Review of Psychology and Philosophy LIX (1950)
2. Loeb, G.E., Otten, B.: T-shirt logo (human machine handshake(Computational Neuroscience: Motor Control, Cold Spring Harbor, NY (1986)
3. Shadmehr, R., Wise, S.P.: The Computational Neurobiology of Reaching and Pointing: A Foundation for Motor Learning. MIT Press (2005)
4. Chaplin, W.F., Phillips, J.B., Brown, J.D., Clanton, N.R., Stein, J.L.: Handshaking, gender, personality, and first impressions. Journal of Personality and Social Psychology 79 (2000) 110-117
5. Stewart, G.L., Dustin, S.L., Barrick, M.R., Darnold, T.C.: Exploring the handshake in employment interviews. Journal of Applied Psychology 93 (2008) 1139-1146
6. Jindai, M., Watanabe, T., Shibata, S., Yamamoto, T.: Development of Handshake Robot System for Embodied Interaction with Humans. The 15th IEEE International Symposium on Robot and Human Interactive Communication, Hatfield, UK (2006)
7. Kasuga, T., Hashimoto, M.: Human-Robot Handshaking using Neural Oscillators. Internatioanl Conference on Robotics and Automation. IEEE, Barcelona, Spain (2005)
8. Ouchi, K., Hashimoto, S.: Handshake Telephone System to Communicate with Voice and Force. IEEE International Workshop on Robot and Human Communication (1997) 466-471
9. Bailenson, J.N., Yee, N.: Virtual interpersonal touch and digital chameleons. Journal of Nonverbal Behavior 31 (2007) 225-242
10. Miyashita, T., Ishiguro, H.: Human-like natural behavior generation based on involuntary motions for humanoid robots. Robotics and Autonomous Systems 48 (2004) 203-212
11. Hogan, N., Sternad, D.: On rhythmic and discrete movements: reflections, definitions and implications for motor control. Exp Brain Res 181 (2007) 13-30
12. Guiard, Y.: On Fitts's and Hooke's laws: simple harmonic movement in upper-limb cyclical aiming. Acta psychologica 82 (1993) 139-159
13. Buchanan, J., Park, J., Shea, C.: Target width scaling in a repetitive aiming task: switching between cyclical and discrete units of action. Experimental Brain Research 175 (2006) 710-725
14. Levy-Tzedek, S., Ben Tov, M., Karniel, A.: Early switching between movement types: indication of predictive control? (submitted)
15. Levy-Tzedek, S., Krebs, H., Song, D., Hogan, N., Poizner, H.: Non-monotonicity on a spatio-temporally defined cyclic task: evidence of two movement types? . Experimental Brain Research (in press 2010)
16. Wichmann, F.A., Hill, N.J.: The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys 63 (2001) 1293-1313
17. Goodale, M.A, Milner, A.D. Seperate visoual pathways for perception and action, Trends Neurosci. 15 (1992) 20–25
18. Pressman, A., Nisky, I., Karniel, A., Mussa-Ivaldi, F.A.: Probing Virtual Boundaries and the Perception of Delayed Stiffness. Advanced Robotics 22 (2008) 119-140