

A Variety of Automated Turing Tests for Network Security: Using AI-Hard Problems in Perception and Cognition to Ensure Secure Collaborations

John P. McIntire

*Air Force Research Laboratory
711th Human Performance Wing / RHCV
john.mcintire@wpafb.af.mil*

Lindsey K. McIntire

*Henry M. Jackson Foundation
for the Advancement of Military Medicine
lindsey.mcintire@wpafb.af.mil*

Paul R. Havig

*Air Force Research Laboratory
711th Human Performance Wing / RHCV
paul.havig@wpafb.af.mil*

ABSTRACT

There are a multitude of collaborative and network applications that are vulnerable to interference, infiltration, or attack by automated computer programs. Malicious programs can spam or otherwise disrupt email systems, blogs, and file sharing networks. They can cheat at online gaming, skew the results of online polls, or conduct denial-of-service attacks. And sophisticated AI “chat-bots” can pose as humans in order to gather intelligence from unsuspecting targets. Thus, a recurring problem in collaborative systems is how to verify that a user is a human and not a computer. Following the work of Coates et al. [1], von Ahn et al. [2], and others, we propose several AI-hard problems in perception and cognition that can serve as “CAPTCHAs,” or tests capable of distinguishing between human-level intelligence and artificial intelligence, ensuring that all collaborators interfacing a particular system are humans and not nefarious computer programs.

KEYWORDS: coordination and cooperation mechanisms, human factors in collaboration, human-machine collaborative interaction, awareness in collaborative systems, cognitive and psychological issues in collaboration

1. INTRODUCTION

Malicious automated programs constantly threaten, disrupt, and infiltrate collaborative systems enacted over the Internet or other distributed networks [2]. These

threats can include worms, denial-of-service attacks, or spamming (email, instant messaging, blogs, file sharing networks, etc.). Automated computer voting threatens the integrity of online polls, and similar cheating concerns plague online collaborative gaming. And automated programs can hack into websites or other password-protected systems by using repeated dictionary attacks. In all of these situations, automated computer programs threaten the safety, security, and usability of collaborative systems and distributed networks.

Additionally, automated programs equipped with artificial intelligence (AI) can pose as humans in collaborative environments for nefarious purposes. One can imagine a social or work situation in which users believe they are communicating with a trusted friend or colleague (or even a stranger), but are instead interacting with a computer program pretending to be human. This is not as far-fetched as it may seem; in the 1960's, an AI program designed to model a psychotherapist caused quite a stir when users became convinced that they were conversing with a caring human instead of a machine [3]. More modern “chat-bots” are considerably more sophisticated and more convincing, and pose correspondingly greater threats since a single malevolent user can dispatch thousands of chat-bots simultaneously for intelligence gathering purposes (e.g., see [4]).

The overall problem is that malicious automated computer programs can straightforwardly and with little resistance gain access to, interact with, or interfere with a variety of collaborative systems. Thus, a significant range of potential security threats can be easily avoided if access to these systems is restricted to only users that are verifiably human.

2. AUTOMATED TURING TESTS

So how can one tell if a user accessing a system is a person, and not a computer? A reliable method of distinguishing between human-level intelligence and artificial intelligence is needed. Such a method is referred to as a “Reverse Turing Test” [1] or as an “Automated Turing Test” [2], in reference to the well-known Turing Test for artificial intelligence. The original Turing Test was a challenge to AI designers to develop computer intelligence so advanced that human users could not tell whether they were conversing (typing back-and-forth) with a person or a machine [5]. In contrast, a Reverse or Automated Turing Test is meant to be a set of problems so advanced that a computer would be incapable of solving them while a human could easily do so, with the caveat that a machine should still be able to judge the success or failure of the test taker. Any such test can also be called a “HIP,” a Human Interaction Proof [6], or a “CAPTCHA,” a Completely Automated Public Turing Test to Tell Computers and Humans Apart [2].

Due to the unique capabilities of human vs. machine intelligence, it turns out that the best-designed CAPTCHAs are those that pose AI-hard problems for the test takers. AI-hard problems are informally the most difficult artificial intelligence problems; they seem to require computational capabilities equivalent to or exceeding human-level intelligence [7]. The security of an AI-hard test is guaranteed as long as the problems used remain generally unsolvable by computer. And since the visual modality is by far the most utilized for human-computer interfaces, most CAPTCHAs to date have involved the testing of AI-hard problems in visual perception. For instance, a common CAPTCHA in current use involves text readability. Also known as “pessimal print,” these tests are the blurry, jumbled, and distorted letter, number, and/or word prompts that must be read and correctly typed as authentication for a variety of web services (Figure 1). These tests are able to routinely defeat most Optical Character Recognition machines or software, while being able to be read by most humans most of the time [1].



Figure 1. Pessimal Print of “HELLO”

Another common CAPTCHA invokes the problem of object recognition. In these tests, pictures of an object or objects are randomly picked from a database of imagery and presented to the user, who must then identify the particular object shown in the images (e.g., horse, dog, cup, etc.; see Figure 2). As long as the imagery is randomly distorted a bit each time it is presented (to avoid a simple template-matching hack), a computer should be generally unable to identify the objects [2].

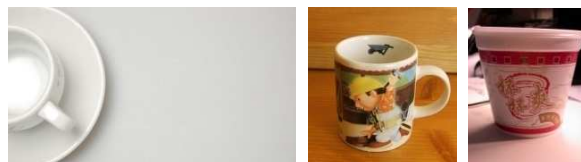


Figure 2. Multiple Images of a Target Object (Cup)

As ingenious as these current implementations of CAPTCHAs are, computer technology is progressing to the point that these types of tests are becoming vulnerable to AI attack [8], particularly tests of text readability (e.g., see [6,9]). To combat this potential issue, and to add to the current network defense toolset, we propose several further AI-hard perception and cognition problems that should remain impenetrable to computer attack for a relatively longer time, as they are probably not as “AI-easy” as text-readability or single object recognition in static imagery. In the visual domain, these tests include face recognition, motion perception, depth perception, mental rotation, and visuo-spatial memory; and in the auditory domain, spatial hearing. We hope that some of these ideas may prove useful in the growing battle to distinguish people from computers, or at the very least encourage further research on this important topic.

3. FACIAL FEATURE RECOGNITION

Despite its potential utility in security and surveillance applications, visual facial feature recognition by computers remains an extremely complex task given the huge possibility space of facial features that vary along a number of dimensions, including the rough geometrical relationships between the eyes, nose, mouth, chin, etc., their specific shapes, skin color, the presence/absence of facial hair, etc. [10]. The difficulty is exacerbated by the fact that recognition algorithms must make use of primarily two-dimensional imagery, which may have been taken under a variety of lighting conditions at various angles and distances, to identify features of a complex three-dimensional object, the face. But humans appear to be equipped with a dedicated region of brain tissue for processing facial feature information [11], and they are extremely good at it. And while humans rarely have difficulty identifying a familiar person who is

missing their usual mustache, wearing a hat, or donning sunglasses, such manipulations are notoriously problematic for computer recognition algorithms. We propose that two different areas within this field pose particularly difficult computational problems that seem to be easily solved by humans, and may thus serve as useful CAPTCHAs: identifying the faces of individual persons, and interpreting the emotional expressions of faces.

3.1. Individual Identification

Individual face identification by computer is generally accomplished using either 2D feature matching or template matching, or a mix of both, although there is a growing push for extracting depth information from the imagery to create 3D templates, with the goal of enhancing identification under a variety of conditions [10]. Despite the many advances in this field, individual face identification remains exceedingly difficult for computers to accomplish, and due to the advanced algorithms needed, it is an impractical option for someone attempting to break a facial recognition CAPTCHA.

There are several ways in which individual facial feature recognition CAPTCHAs could be implemented. A large database of catalogued photos of faces could be drawn upon. In a single test, a user would view perhaps six different photographs at once, consisting of two different photos for three different people, and would have to match the three pairs of photos together that are of the same person (Figure 3). Or perhaps a large number of photos could be presented, of which only two are of the same person and the rest are of different people, and the user would have to identify the pair of photos that is of a single person. It must be kept in mind that a strict template matching program might be able to break such a CAPTCHA if a public database of photos is used, as human assistance could be provided to the malware by identifying all photos in the database before bots are sent to attack the system. But randomly distorting each image before presentation may help combat this potential problem.

There are a variety of other ways in which such a test could be implemented; another intriguing method is given by Rui & Liu [12]. The point is that identification of individuals through facial feature recognition might serve as a useful CAPTCHA test, due to the difficulty of implementing such capabilities in a computer. However, as the general field of object recognition in artificial intelligence continues to advance, comparable advances in facial feature recognition will probably advance likewise, highlighting the potential need for more advanced CAPTCHAs.



Figure 3. Individual Facial ID CAPTCHA: Which Pairs of Photos Go Together?

3.2. Emotional Expression Identification

Identifying the emotional expresses of faces involves a complex cognitive component that is not present in individual facial feature recognition tests. Thus, even if AI feature recognition capabilities become common and widely available, understanding the emotional content in faces adds a layer of complexity to the problem, decreasing the possibility of a computer accomplishing such a task. In contrast, most people can do this task effortlessly. Ekman [13] has shown that emotions conveyed through facial expressions in two-dimensional photographs of faces are easily identifiable and universally understood (culturally independent). Although there is still on-going debate regarding the exact number and type of emotions that form the basic emotional palette, there seems to be enough consensus for emotional-IDs to be useful as CAPTCHAs.

The most agreed upon set of universal emotions include anger, disgust, fear, joy, sadness, and surprise [13]. A CAPTCHA incorporating the identification of these emotions in facial expressions could be easily implemented, perhaps by requiring the correct identification of 80% on a series of 10 randomly selected faces displaying emotion. To accomplish this, a database of coded facial expression imagery would be required. Again, due to the possibility of a simple template-matching hack, slight random image distortions might be needed each time the test is given.

4. MOTION PERCEPTION TASKS

Motion perception is another complex human capability and AI-hard problem that may serve as a useful CAPTCHA. There are a variety of potential implementations that could be tested using this paradigm. One motion CAPTCHA proposal suggests using moving text characters against a background to cause perceptual “pop-out” of a pass-code, which must then be entered; or that the foreground can be interacted with through the mouse, and that by moving the foreground text, the pass-code becomes perceptible against the immobile background [14]. Our proposals differ from these methods in that the focus is on psychophysical perceptions of motion, such as involving direction discriminations or speed determinations, and do not at their core involve text/character or object recognition that happens to be dynamic, although those are certainly good ideas.

One rather simple implementation might ask a participant to view a quick “movie” or rapid succession of screen shots in which a single object is displaced on the screen, forming the perception of an object having moved in the x,y plane. The movement should start from a random position, and speed should vary across trials but always fall within the limits of human perceptibility. The optimal speeds required for stimulating the human visual motion system are between 4 and 64 degrees of visual angle per second, and the minimum integration time for motion perception is around 100 – 200 ms [15]. The participant must then select or otherwise indicate the trajectory of the object. The task is made difficult by using a multitude of visually-complex objects moving against a background of other objects or scenery, all in different directions. This would require the determination of movement direction of one particular target object, among several moving objects. This test would be particularly difficult, as object recognition *and* perception of movement direction would be necessary to pass the test. To add another layer of complexity, one might randomize the task instructions as to what target should be searched for; for example, on one trial the test might require indicating the direction of the object that is moving the fastest, is the largest, and is blue, and on the next trial, the target object might be slow and green.

Instead of using singular objects as the targets, entire moving fields of dots could be used, as is common in psychophysical experiments of motion perception. If two overlapping fields of dots are both moving in different directions relative to each other, the perception is of a single field moving in a direction that is an average of the two if the movement directions fall within about 10 degrees of each other [16]. If they move in directions that differ by more than this, the perception is of two “sheets” of dots moving in different directions over top of each other. A CAPTCHA based on this principle might ask

how many moving fields are present, and what direction are they moving? If the field direction difference is less than 10 degrees, we would expect humans to see a unified percept of a single moving field, while a computer, even one equipped with a simplistic motion-perception system, would likely see two moving fields instead.

A similar possibility is using a random dot kinematogram (RDK), a commonly used tool in motion psychophysics, where each dot within a field of dots all move on different trajectories, and they change direction each frame, but within a limited range of possible directions. A range of up to 30 degrees allows direction discriminations to be as good as if all dots were moving in the same direction [15]. The result of an RDK display is a single unified percept of coherent motion that is an average of all vectors present, and appears somewhat like a flowing river of dots; indeed, it is also referred to as a *flow field*. In this instance, a participant could easily indicate the overall direction of motion of the flow field, but a computer lacking a complex visual system equipped with human-like motion perception mechanisms should fare poorly.

4.1. Form-from-Motion Perception

Besides direction discriminations, other aspects of visual motion perception may serve as useful CAPTCHAs. A fascinating property of the visual system is that it can easily derive the three-dimensional shape of an object even if given extremely sparse visual information via a two-dimensional display, especially if the object is rotating or otherwise moving relative to the background. This phenomenon is somewhat analogous to the *motion parallax* cue to depth (see Section 4) but is probably more based upon 3D object recognition, as opposed to a focus on primarily depth extraction. It is called *form-from-motion* perception.

To elicit this perception, a display similar to RDKs is needed, in that a moving field of dots is presented. In contrast to RDKs, though, the dots move as though they are spots of light upon three-dimensional objects and surfaces, and that the viewpoint is rotating around the fixed object(s) of interest. Any given static image of the scene shows nothing but a disorganized field of dots. But when dynamic, the result is the strong and undeniable percept of three-dimensional object structure. This type of *form-from-motion* CAPTCHA would require users to view the dynamic display and identify the object in the scene, perhaps choosing among several randomly selected possibilities. As shown in Figure 4a, no object is visible in any particular static scene, but the object “pops out” perceptually when the static scenes are temporally combined in a form-from-motion test (Figure 4b). The complexity of this type of test would serve as an excellent

foil against computational attack, as it requires 3D object recognition from extremely sparse dynamic visual information.

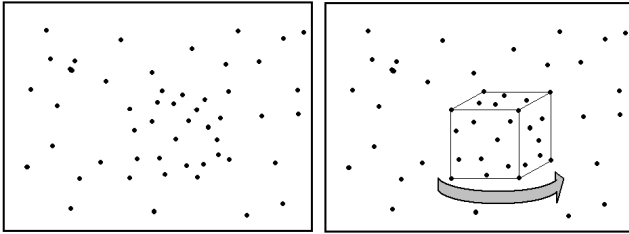


Figure 4. (a) Static Scene – No Form is Perceptible (b) Relative Motion of Dots Creates the Perception of a 3D Shape (Rotating Cube)

5. DEPTH PERCEPTION

Depth perception is another particularly vexing computational problem that should prove difficult for machines to perform for the foreseeable future. Because it is a cognitively sophisticated task likely requiring extensive computational resources and advanced architectures to model accurately, tests involving depth judgments might serve as useful CAPTCHAs. For present purposes, *binocular* depth perception can be ignored due to the lack of commonly available 3D displays and the potentially large numbers of the general population that possess stereoscopic vision anomalies (perhaps as high as 30%) [17], as these limit its usefulness as a CAPTCHA. We will instead focus on *monocular* depth perception (also called pictorial, perspective, or 2.5D depth) tasks that do not require the use of two eyes.

Monocular cues to depth include relative size, texture gradient, occlusion, linear perspective, atmospheric haze, shadows, and if including dynamic cues, motion parallax. For instance, the occlusion depth cue suggests that if one object is occluding (overlapping or blocking) another, the occluded object appears further in depth (Figure 5). Thanks to the ubiquity of 2D and 3D graphics programs and engines, all of these cues can be easily created or otherwise elicited for testing on a standard two-dimensional computer display.

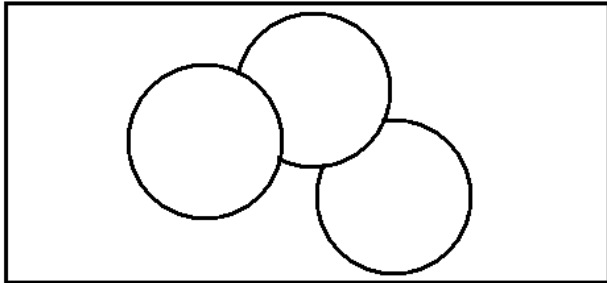


Figure 5. Occlusion Depth Cues

What this means for CAPTCHA development is that imagery can be captured and/or generated containing multiple objects at various depths. A test taker would have to correctly order these objects (from closest to farthest in depth) or perhaps would judge which object is closest in depth to a reference object that is randomly selected, while viewing the imagery via a two-dimensional standard computer display. This idea is very similar to the “3-D CAPTCHA” proposed by Kaplan [18], except the focus here is more upon depth ordering than on strictly object recognition of three dimensional objects, although combining the two might make the task even more difficult for a computer to solve.

To add further complexity to this task, wire-frame type objects might be over or under-sized relative to each other, and overlapping or overlaying (but not completely occluding) other objects in the scene, to increase the difficulty of performing figure/ground segmentation of objects. An analogous technique for two-dimensional imagery has shown some success, utilizing this principle of figure/ground confusions by resizing and overlapping a variety of images and requiring test takers to identify the center-points of multiple images within a single combined image [19]. To solve our proposed problems, a computer program would have to complete two AI-hard tasks simultaneously: object recognition, to identify the appropriate reference objects in the scene, and depth perception, to identify the spatial position of objects along the z-axis and to order them correctly. This highlights the computational complexity of what appears to be a simple, everyday task for us humans. But a computer lacking the ability to process depth cues and to conceptualize those objects in space (particularly when shown only a 2D image of a 3D scene, i.e., 2.5D) should reliably fail a depth perception CAPTCHA test.

6. MENTAL ROTATION TASKS

Following the ground-breaking research of Shepard and his colleagues [20] into mental imagery, mental rotation tasks have been used in cognitive psychology for decades. These tasks require the mental rotation of an object through two or three dimensions. Several implementations of these tests are possible, including matching a rotated object to one of several reference objects; or indicating the amount of rotation of a single standard object. Two dimensional test objects can be used, although the use of three dimensional objects (utilizing projection via two-dimensional displays) is recommended as this adds the complexity of depth extraction of 3D objects to the task (see an example in Figure 6).

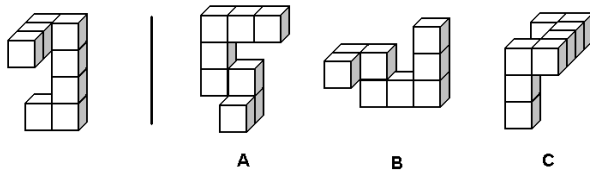


Figure 6. A Shepard & Metzler-type 3D Mental Rotation Task

Another common measure of mental rotation or spatial visualization abilities is the Paper Folding task seen on common intelligence tests (e.g., Stanford-Binet scale) and even on some standardized tests. In these tasks, a folded up piece of paper is shown, and a hole-punch is indicated in the folded paper. The test-taker must then answer, from several possibilities shown, what the piece of paper will look like when completely unfolded (see Figure 7). Computers would presumably be unable to do this task since it seems to require human-like spatial visualization capabilities.

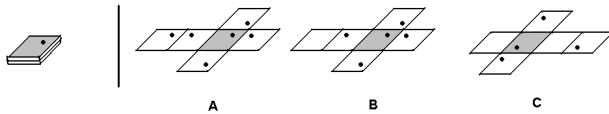


Figure 7. A Paper Folding Task Shown in Perspective (2.5D) Depth

What makes these tests potentially useful as CAPTCHAs is that they typically require both object matching and some form of mental rotation ability of two and/or three dimensional objects. In humans, this ability is made possible through the use of a cognitively sophisticated visuo-spatial “scratchpad” or “sketchpad” [21], upon which the tested object is imagined in the “mind’s eye” and the representation of the object is mentally rotated, in a manner analogous to actually having to rotate a physical object through space; larger rotational transformations require more mental computation and thus take more time to complete [20, 21]. A computer program lacking a human-like visuo-spatial sketchpad, upon which spatial objects can be rotated and subsequently compared accurately against a stored 2D or 3D representation, should fare poorly in comparison to a human taking these types of tests.

7. VISUO-SPATIAL MEMORY TASKS

The game “Memory” is a visuo-spatial short-term memory task that is relatively easy for humans to do, but might be very difficult for a computer to perform, making it a potentially good CAPTCHA. In the game “Memory,” cards are turned faced-down and arranged in a grid. On

any given turn, the player must select any two cards, which are then flipped over for a brief time before being returned face down in their original positions. The object of the game is to match each card with its pair by remembering the spatial locations of each card face. On the face of each card may be an object, a simple shape, a word, a number, or even complex imagery. Part of the difficulty of the task lies in the fact that card parity may be ambiguous until several cards are overturned and viewed; card pairs are not necessarily identical images. They may be defined as pairs based upon membership in a high-level conceptual category (women, sporting equipment, prime numbers, etc.), or as lower-level categories (numbers, letters, shapes, animals, etc.). Of course, the simplest and most straightforward implementation would define pairs as exact replicas of each other; but again, random image distortions might be necessary to avoid template-matching hacks in this instance.

For the purposes of developing a CAPTCHA “Memory” game, we would want a fairly small grid to reduce time on task and to ensure that the vast majority of human test takers would pass. A 4 x 4 grid would be only mildly challenging for a human but probably still too complex for a computer lacking visual cognition. The imagery on the card could very plainly be letters, numbers, or simple shapes so that implementation via computer programming system would be easy, without requiring extensive databases of catalogued imagery. To add difficulty to the game (if needed), one could add imagery, graphics, complex shapes, etc. Difficulty might also be raised by cutting single images in half and using the half-images as the game pairs, or defining pairs on high level conceptual categories that should not be immediately obvious to a silicon brain.

8. SPATIAL (3D) AUDITORY TASKS

Although the previous CAPTCHAs have all focused on only *visual* perception and/or cognition issues, there is growing interest in using auditory CAPTCHAs, particularly to supplement visual CAPTCHAs in the case of visually-impaired human test takers. However, current auditory CAPTCHA implementations involve spoken word/number recognition tasks, and are at serious risk of being broken with current pattern recognition technologies. To make the task much more difficult for a computer to perform, we propose using spatial (3D) auditory listening tasks, which requires that test-takers possess only stereo headphones and a standard sound card.

Spatial (3D) audio creates the compelling illusion of sound coming from a specific point or direction in space, through the use of normal headphones [22]. Test-takers

could very easily listen to perhaps 10 different sequentially-presented sound items, of which they would be required to denote on-screen (or otherwise signal) the spatial direction of the perceived sounds. A computer lacking complex binaural sound processing would have severe difficulty signaling the correct spatial direction. Or perhaps, current audio CAPTCHA identification tasks might be made more difficult by requiring the listener to identify the sound/word coming from a particular direction in the presence of a spatially noisy environment, an amazing human capability called the *cocktail party phenomenon* [23]; again a computer would surely have great difficulty with this task.

9. IMPLEMENTATION

How difficult would it be to implement these types of tests in computers? We focused our efforts on these particular cognitive and perceptual problems precisely because (1) they should be relatively easy to implement in a computer and (2) they can be easily graded by computer. In contrast, CAPTCHAs such as answering trivia questions or comprehending the material in a paragraph of text are very effective at distinguishing people from computers, but generating a huge variety of these tests and grading the answers are problematic for a computer to do properly and would limit their widespread applicability. If generating, presenting, and grading such tests cannot be automated (performed by computer), then it is difficult to use them as “Automated Turing Tests.”

Many of the tests proposed in this paper are psychophysical in nature, and the discovery of and research into these phenomenon were in fact made possible by computer technologies that automated their scientific testing in laboratories. In other words, these topics have been so successfully and extensively researched because their “tests” are easy to implement and grade with computers, implying their usefulness as CAPTCHAs. More cognitively complex tasks, such as face recognition, emotion identification, or object recognition, might pose more difficulty for a computer attempting to *generate* random tests, but this may be worked around using large stored databases of pre-defined test stimuli, with which a computer can draw upon to present randomized tests and then grade them.

10. CONCLUSIONS

The ability to reliably distinguish people from computers is an extremely important issue for the security of many modern network platforms, particularly those platforms involving distributed communications and collaborations. Social networks, chat-rooms, email systems, webpages, online voting and gaming, blogs, and file sharing

networks are some of the collaborative applications that are especially vulnerable to automated computer infiltration, exploitation, and/or attack.

We have reviewed several outstanding AI-hard problems in both visual and auditory perception and cognition that we believe hold particular promise as CAPTCHAs, or tests capable of easily distinguishing between computer and human-level intelligence. Facial feature recognition, motion and depth perception, mental rotation, visual memory, and spatial hearing are all incredibly complex human behavioral phenomena layered with mixes of lower-level, bottom-up processes and high-level, top-down influences, creating largely unpredictable and chaotic outcomes that for the time being lay outside the realm of computational replication.

The true beauty in these “Automated Turing Tests” is that they pose the same fundamental challenge posed by Alan Turing in his original Turing Test; that of humankind developing computer intelligences so advanced that they can convincingly pose as humans. Turing’s challenge has, for the most part, gone seemingly unfulfilled, but this will not always be the case. Hopefully, this review has given security developers some ideas as to how to verify humanness in collaborative systems, and also clarified for AI developers some of the outstanding hurdles their field will face as their work proceeds. In any case, as long as the problems outlined above remain impenetrable to full understanding, which may be for a long while yet, they should easily provide solid security for human authentication in collaborative computer workspaces and systems.

ACKNOWLEDGEMENTS

We would like to thank Dr. Scott Watamaniuk at Wright State University for his expertise on visual motion perception and Dr. Regina Schmidt at AFRL for her assistance regarding facial recognition perception.

REFERENCES

- [1] A.L. Coates, H.S. Baird, and R.J. Fateman, “Pessimist Print: A Reverse Turing Test,” *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR '01)*, Seattle, WA, 2001, pp. 1154-1159.
- [2] L. von Ahn, M. Blum, and J. Langford, “Telling Humans and Computers Apart Automatically,” *Communications of the ACM*, Vol. 47, No. 2, 2004, pp. 57-60.
- [3] J. Weizenbaum, “ELIZA – A Computer Program For the Study of Natural Language Communication Between Man and Machine,” *Communications of the ACM*, Vol. 9, No. 1, 1966, pp. 36-45.

- [4] P. Naughton, "Flirty Chat-Room 'Bot' Out to Steal Your Identity," FOXNews [web article]. (Dec 2007). Available: <http://www.foxnews.com/story/0,2933,316473,00.html>.
- [5] A.M. Turing, "Computing Machinery and Intelligence," in A. Collins & E.E. Smith (Eds.) READINGS IN COGNITIVE SCIENCE: A PERSPECTIVE FROM PSYCHOLOGY AND ARTIFICIAL INTELLIGENCE, Morgan Kaufmann Publishers, Inc., San Mateo, CA, 1988.
- [6] K. Chellapilla and P.Y. Simard, "Using Machine Learning to Break Visual Human Interaction Proofs (HIPs)," Advances in Neural Information Processing (NIPS '04), MIT Press, 2004.
- [7] "AI-complete," [webpage]. (September 2008). Available: <http://en.wikipedia.org/wiki/AI-complete>.
- [8] K. Chellapilla, K. Larson, P. Simard, and M. Czerwinski, "Designing Human Friendly Human Interaction Proofs (HIPs)," Conference on Human Factors in Computing Systems (CHI), 2005.
- [9] G. Mori and J. Malik, "Recognizing Objects in Adversarial Clutter: Breaking a Visual CAPTCHA," Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR '03), 2003.
- [10] R. Brunelli and R. Poggio, "Face Recognition: Feature versus Templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, October 1993, pp. 1042-1052.
- [11] N. Kanwisher and G. Yovel, "The Fusiform Face Area: A Cortical Region Specialized for the Perception of Faces," *Philosophical Transactions of the Royal Society B*, Vol. 361, 2006, pp. 2109-2128.
- [12] Y. Rui and Z. Liu, "ARTiFACIAL: Automated Reverse Turing Test Using FACIAL Features," Microsoft Technical Report, April, 2003, <http://research.microsoft.com>.
- [13] P. Ekman, "Universals and Cultural Differences in Facial Expressions of Emotion," Nebraska Symposium on Motivation, University of Nebraska Press, Lincoln, NE, 1972, Vol. 19, pp. 207-283.
- [14] P. Qvarfordt, E.G. Rieffel, and D.M. Hilbert, "Motion and Interaction Based CAPTCHAs," United States Patent 20080127302, published May 29, 2008.
- [15] S.P. McKee and S.N.J. Watamaniuk, "The Psychophysics of Motion Perception," in A. Smith and R. Snowden (Eds.) VISUAL DETECTION OF MOTION, Academic Press, New York, 1994.
- [16] S.N.J. Watamaniuk, Wright State University, Department of Psychology, Dayton, OH. *Personal communication*.
- [17] I.P. Howard, SEEING IN DEPTH: BASIC MECHANISMS, Vol. 1, University of Toronto, Toronto, 2002, p. 487.
- [18] M.G. Kaplan, (September, 2008). The 3-D CAPTCHA [website]. Available: <http://spamfizzle.com/CAPTCHA.aspx>.
- [19] R. Datta, J. Li, and J.Z. Wang, "IMAGINATION: Image-based CAPTCHA Generation System," ACM International Conference on Multimedia, November, 2005, Available: <http://wang.ist.psu.edu/imagination/imagination.ppt>.
- [20] R. Shepard and J. Metzler, "Mental Rotation of Three Dimensional Objects," *Science*, Vol. 171, Issue 972, 1971, pp. 701-703.
- [21] R.L. Solso, M.K. MacLin, and O.H. MacLin, COGNITIVE PSYCHOLOGY, 7th ed., Pearson Education, Inc., Boston, MA, 2005.
- [22] F.L. Wightman and D.J. Kistler, "Headphone Simulation of Free-Field Listening. II: Psychophysical Validation," *Journal of the Acoustical Society of America*, Vol. 85, No. 2, 1989, pp. 868-878.
- [23] A.W. Bronkhorst, "The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions," *Acta Acustica*, Vol. 86, 2000, pp. 117-128.