# Complex Image Recognition and Web Security

Henry S. Baird[1]

Computer Science & Engineering Department
Lehigh University, Bethlehem, Pennsylvania USA
baird@cse.lehigh.edu
www.cse.lehigh.edu/~baird

**Summary.** Web services offered for human use are being abused by programs. Efforts to defend against these have, over the last five years, stimulated the development of a new family of security protocols able to distinguish between human and machine users automatically over GUIs and networks. AltaVista pioneered this technology in 1997; by 2000, Yahoo! and PayPal were using similar methods. Researchers at Carnegie-Mellon University [BAL00] and, then, a collaboration between the University of California at Berkeley and the Palo Alto Research Center [CBF01] developed such tests. By January 2002 the subject was called 'human interactive proofs' (HIPs), defined broadly as challenge/response protocols which allow a human to authenticate herself as a member of a given group: e.g. human (vs. machine), herself (vs. anyone else), etc. All commercial uses of HIPs exploit the gap in reading ability between humans and machines. Thus, many technical issues studied by the image recognition research community are relevant to HIPs. This chapter describes the evolution of HIP R&D, applications of HIPs now and on the horizon, relevant legal issues, highlights of the first two HIP workshops, and proposals for an image recognition research agenda to advance the state of the art of HIPs.

## 1 Introduction

In 1997 Andrei Broder and his colleagues [LBBB01], then at the DEC Systems Research Center, developed a scheme to block the abusive automatic submission of URLs [Bro01] to the AltaVista web-site. Their approach was to present a potential user with an image of printed text formed specially so that machine vision (OCR) systems could not read it but humans still could. In September 2000, Udi Manber, Chief Scientist at Yahoo!, challenged Prof. Manuel Blum and his students [BAL00] at The School for Computer Science at Carnegie Mellon University (CMU) to design an "easy to use reverse Turing test" that would block 'bots' (computer programs) from registering for services including chat rooms, mail, briefcases, etc. In October of that year, Prof. Blum asked the first author, of the Palo Alto Research Center (PARC), and Prof. Richard Fateman, of the Computer Science Division of the Univer-

sity of California at Berkeley (UCB), whether systematically applied image degradations could form the basis of such a filter, stimulating the development of PessimalPrint [CBF01].

In January 2002, Prof. Blum and the present authors ran a workshop at PARC on 'human interactive proofs' (HIPs), defined broadly as *a class of challenge/response protocols which allow a human to be authenticated as a member of a given group — an adult (vs. a child), a human (vs. machine), a particular individual (vs. everyone else), etc.* All commercial uses of HIPs known to us exploit the large gap in ability between human and machine vision systems in reading images of text.

The number of applications of vision-based HIPs to Web security is large and growing. HIPs have been used to block access to many services by machine users, but they could also, in principle, be used as 'anti-scraping' technologies to prevent the large-scale copying of databases, prices, auction bids, etc. If HIPs — possibly not based on vision — could be devised to discriminate reliably between adults and children, the commercial value of the resulting applications would be large.

Many technical issues that have been systematically studied by the image recognition community are relevant to the HIP research program. In an effort to stimulate interest in HIPs within the document image analysis research community, this chapter details the evolution of the HIP research field, the range of applications of HIPs appearing on the horizon, highlights of the first HIP workshop, and proposals for an image recognition research agenda to advance the state of the art of HIPs.

This paper is an expanded and updated version of [BK02].

### 1.1 An Influential Precursor: Turing Tests

Alan Turing proposed [Tur50] a methodology for testing whether or not a machine effectively exhibits intelligence, by means of an "imitation game" conducted over teletype connections in which a human judge asks questions of two interlocutors — one human and the other a machine — and eventually decides which of them is human. If judges fail sufficiently often to decide correctly, then that fact would be, Turing proposed, strong evidence that the machine possessed artificial intelligence. His proposal has been widely influential in the computer science, cognitive science, and philosophical communities [SCA00] for over fifty years.

However, no machine has "passed the Turing test" in its original sense in spite of perenniel serious attempts. In fact it remains easy for human judges to distinguish machines from humans under Turing-test-like conditions. Graphical user interfaces (GUIs) invite the use of images as well as text in the dialogues.

### 1.2 Robot Exclusion Conventions

The Robot Exclusion Standard, an informal consensus reached in 1994 by the robots mailing list (`robots@nexor.co.uk`), specifies the format of a file (the `http://.../robots.txt` file) which a web site or server may install to instruct all robots visiting the site which paths it should not traverse in search of documents. The Robots META tag allows HTML authors to indicate to visiting robots whether or not a document may be indexed or used to harvest more links (cf. `www.robotstxt.org/wc/meta-user.html`).

Many Web services (Yahoo!, Google, etc) respect these conventions. Some 'abuses' which HIPs address are caused by deliberate disregard of these conventions. The legality of disregarding the conventions has been vigorously litigated but remains unsettled [Bar01,Pli02]. Even if remedies under civil or criminal law are finally allowed, there will certainly be many instances where litigation is likely to be futile or not cost-effective. Thus there will probably remain strong commercial incentives to use technical means to enforce the exclusion conventions.

The financial value of any service to be protected against 'bots' can not be very great, since a human can be paid (or in some other way rewarded) to pass the CAPTCHA (an acronym for Completely Automated Public Turing Test to Tell Computers and Humans Apart, coined by Prof. Manuel Blum, Luis A. von Ahn, and John Langford of CMU). Of course, minimum human response times — of 5–10 seconds at least — may be almost always slower than a automated attack, and this speed gap may force reengineering of the 'bot' attack pattern. Nevertheless, this may be simpler—and more stable—than actively engaging in an escalating arms race with CAPTCHA designers. There are widespread, but so far unsubstantiated, reports of systematic "farming out" of CAPTCHAs, in which humans are encouraged and rewarded (by, for example, according to an often-repeated rumor, access to porn sites) to pass CAPTCHAs [Tho02].

### 1.3 Primitive Means

For several years now web-page designers have chosen to render some apparent text as image (e.g. GIF) rather than encoded text (e.g. ASCII), and sometimes in order to impede the legibility of the text to screen scrapers and spammers. A frequent use of this is to hide email addresses from automatic harvesting by potential spammers. To our knowledge the extent of this practice has not been documented.

One of the earliest published attempts to automate the reading of imaged-text on web pages was by Lopresti and Zhou [DZ00]. Kanungo et al [KLB01] reported that, in a sample of 862 sampled web pages, "42% of images contain text" and, of the images with text, "59% contain at least one word that does not appear in the ... HMTL file."

### 1.4 First Use: The Add-URL Problem

In 1997 AltaVista sought ways to block or discourage the automatic submission of URLs to their search engine. This free "add-URL" service is important to AltaVista since it broadens its search coverage and ensures that sites important to its most motivated customers are included. However, some users were abusing the service by automating the submission of large numbers of URLs, and certain URLs many times, in an effort to skew AltaVista's importance ranking algorithms.

Andrei Broder, Chief Scientist of AltaVista, and his colleagues developed a filter (now visible at [Bro01]). Their method is to generate an image of printed text randomly (in a "ransom note" style using mixed typefaces) so that machine vision (OCR) systems cannot read it but humans still can (Figure 1). In January 2002 Broder told the present authors that the system had been in use for "over a year" and had reduced the number of "spam add-URL" by "over 95%." (No details concerning the residual 5% are mentioned.) A U.S. patent [LABB01] was issued in April 2001.

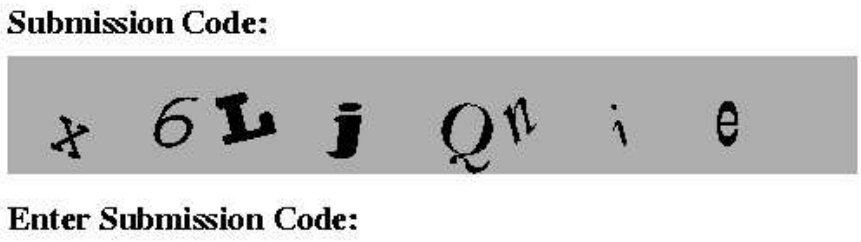


**Fig. 1.** Example of an AltaVista challenge: letters are chosen at random, then each is assigned to a typeface at random, then each letter is rotated and scaled, and finally (optionally, not shown here) background clutter is added.

To the present authors, these do not seem to present a difficult challenge to modern machine vision methods. The black characters are widely separated against a background of a uniform grey, so they can be easily isolated. Recognizing an isolated bilevel pattern (here, a single character) which has undergone arbitrary affine spatial transformations is a well-studied problem in pattern recognition, and several effective methods have been published [SWI99,LBP98]. The variety of typefaces used can be attacked by a brute-force enumeration.

### 1.5 The ChatRoom Problem

In September 2000, Udi Manber of Yahoo! described this "chat room problem" to researchers at CMU: 'bots' were joining on-line chat rooms and irritating

the people there, e.g. by pointing them to advertising sites. How could all 'bots' be refused entry to chat rooms?

CMU's Prof. Manuel Blum, Luis A. von Ahn, and John Langford articulated [BAL00] some desirable properties of a test, including:

- the test's challenges can be automatically generated and graded (i.e. the judge is a machine);
- the test can be taken quickly and easily by human users (i.e. the dialogue should not go on long);
- the test will accept virtually all human users (even young or naive users) with high reliability while rejecting very few;
- the test will reject virtually all machine users; and
- the test will resist automatic attack for many years even as technology advances and even if the test's algorithms are known (e.g. published and/or released as open source).

Theoretical security issues underlying the design of CAPTCHAs have been addressed by Nick Hopper and Manuel Blum in [HB01].

The CMU team developed a "hard" 'GIMPY' CAPTCHA which picked English words at random and rendered them as images of printed text under a wide variety of shape deformations and image occlusions, the word images often overlapping. The user was asked to transcribe some number of the words correctly. An example is shown in Figure 2.


**Fig. 2.** Example of a "hard" GIMPY image produced by the Carnegie-Mellon Univ. CAPTCHA.

The non-linear deformations of the words and the extensive overlapping of images are, in our opinion, likely to pose serious challenges to existing

machine-reading technology. However, it turned out to place too heavy a burden on human users, also: in trials on the Yahoo! website, users complained so much that this CAPTCHA was withdrawn.

As a result, a simplified version of GIMPY ("easy" or "EZ" GIMPY) , using only one word-image at a time (Figure 3), was installed by Yahoo!, and is in use at the time of writing (visible at `chat.yahoo.com` after clicking on 'Sign Up For Yahoo! Chat!'). It is used to restrict access to chat rooms and other services to human users. According to Udi Manber, Chief Scientist of Yahoo!, it serves up as many as a million challenges each day.

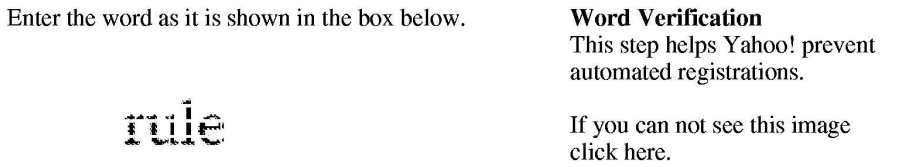


**Fig. 3.** Example of a simplified Yahoo! challenge (CMU's "EZ GIMPY"): an English word is selected at random, then the word (as a whole) is typeset using a typeface chosen at random, and finally the the word image is altered randomly by a variety of means including image degradations, scoring with white lines (shown here), and non-linear deformations.

The variety of deformations and confusing backgrounds (the full range of these is not exhibited in the Figure) poses a serious challenge to present machine-vision systems, which typically lack versatility and are fragile outside of a narrow range of expected inputs. However, the use of one English word may be a significant weakness, since even a small number of partial recognition results can rapidly prune the number of word-choices.

### 1.6 Screening Financial Accounts

PayPal (`www.paypal.com`) is screening applications for its financial payments accounts using a text-image challenge (Figure 4). We do not know any details about its motivation or its technical basis.

This CAPTCHA appears to use a single typeface, which strikes us a serious weakness that the use of occluding grids does little to strengthen.

A similar CAPTCHA has recently appeared on the Overture website (click on 'Advertiser Login' at `www.overture.com`).

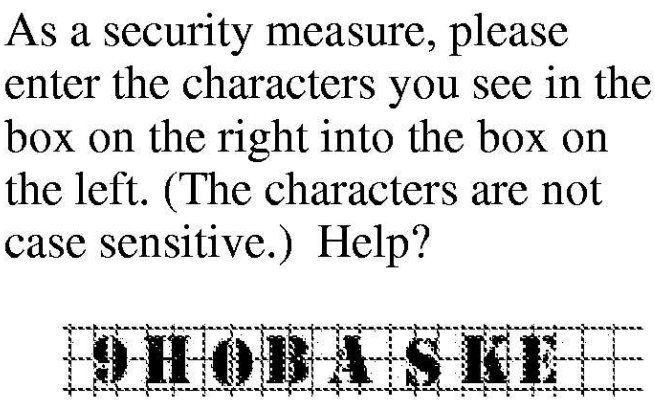


**Fig. 4.** Example of a PayPal challenge: letters and numerals are chosen at random and then typeset, spaced widely apart, and finally a grid of dashed lines is overprinted.

### 1.7 PessimalPrint

A model of document image degradations [Bai92]—approximating the physics of machine-printing and imaging of text—was used to generate the "PessimalPrint" challenges illustrated in Figure 5.

**Fig. 5.** Example of a PessimalPrint challenge: an English word is chosen at random, then the word (as a whole) is typeset using a randomly chosen typeface, and finally the word-image is degraded according to randomly selected parameters (with certain ranges) of the image degradation model.

An experiment assisted by ten UC Berkeley graduate-student subjects and three commercial OCR machines located a range of model parameters

in which images could be generated pseudorandomly that were always legible to the human subjects and never correctly recognized by the OCR systems. In the current version of PessimalPrint, for each challenge a single English word is chosen randomly from a set of 70 words commonly found on the Web; then the word is rendered using one of a small set of typefaces and that ideal image is degraded using the parameters selected randomly from the useful range. These images, being simpler and less mentally challenging than the original GIMPY, would in our view almost certainly be more readily accepted by human subjects.

### 1.8 BaffleText

Chew and Baird [CB03] noticed vulnerabilities of reading-based CAPTCHAs to dictionary and computer-vision attacks, and also surveyed the literature on the psychophysics of human reading, which suggested fresh defenses available to CAPTCHAs. Motivated by these considerations, they designed "BaffleText," a CAPTCHA which uses non-English 'pronounceable words' to defend against dictionary attacks, and Gestalt-motivated image-masking degradations to defend against image restoration attacks. An example is shown in Figure 6.

**Fig. 6.** Example of a BaffleText challenge: a nonsense (but English-like) word was generated pseudorandomly, the word (as a whole) was typeset using a randomly chosen typeface, an mask was generated, and the word image was damaged using the mask.

Experiments on human subjects confirmed high human legibility and user acceptance of BaffleText images. They also found an image-complexity measure that correlated well with user acceptance and assisted the generation of challenges lying within the ability gap.

### 1.9 ScatterType

In response to reports (*e.g.* [MM03,CLSC05]) that several CAPTCHAs in wide use could be broken by segment-then-recognize attacks, Baird *et al* developed ScatterType [BR05,BMW05], whose challenges are images of machine-print text whose characters have been pseudorandomly cut into pieces which have

**Fig. 7.** Example of a ScatterType challenge: a nonsense (but English-like) word was generated pseudorandomly, the characters (separately) typeset using a randomly chosen typeface, an mask was generated, the character images cut into pieces, and the pieces scattered pseudorandomly.

then been forced to drift apart. An example is shown in Figure 7. This scattering is designed to repel automatic segment-then-recognize computer vision attacks. Results from an analysis of data from a human legibility trial with 57 volunteers that yielded 4275 CAPTCHA challenges and responses show that it is possible to locate an operating regime—ranges of the parameters that control cutting and scattering—within which human legibility is high (better than 95% correct) even though the degradations due to scattering remain severe.

## 2 The First International HIP Workshop

The first NSF-sponsored workshop on Human Interactive Proofs was held January 9-11, 2002, at the Palo Alto Research Center. There were thirty-eight invited participants, with large representations from CMU, U.C. Berkeley, and PARC. The Chief Scientists of Yahoo! and Altavista were present, along with researchers from IBM Research, Lucent Bell Labs, Intertrust STAR Labs, RSA Security, and Document Recognition Technologies, Inc. Prof. John McCarthy of Stanford University presented an invited plenary talk on "Frontiers of AI".

As a starting point for discussion, HIPs were defined tentatively as

> *automatic protocols allowing a person to authenticate him/herself — as, e.g., human (not a machine), an adult (not a child), himself (no one else) — over a network without the burden of passwords, biometrics, special mechanical aids, or special training.*

Topics presented and discussed included:

- Completely Automatic Public Turing tests to tell Computers and Humans Apart (CAPTCHAs): criteria, proofs, and design;
- secure authentication of individuals without using identifying or other devices;
- catalogues of actual exploits and attacks by machines to commercial services intended for human use;
- audio-based CAPTCHAs;
- CAPTCHA design considerations specific to East-Asian languages;

- authentication and forensics of video footage;
- feasibility of text-only CAPTCHAs;
- images, human visual ability, and computer vision in CAPTCHA technology;
- human-fault tolerant approaches to cryptography and authentication;
- robustly non-transferable authentication; and
- protocols based on human ability to memorize through association and perform simple mental calculations.

Some details of the HIP2002 workshop are available on-line at

`www.parc.com/istl/groups/did/HIP2002`

including the Program and Participants' list.

## 3 The Second International HIP Workshop

The 2nd International Workshop on Human Interactive Proofs (HIP2005, May 19-20, Bethlehem, PA) brought together twenty-six researchers, engineers, and business people interested in technologies to protect networked services from abuse by programs (bots, spiders, phishers, etc.) masquerading as legitimate human users.

Attendees participated in an intensive day and a half of plenary talks, panels, and group discussions sharing the state of the art and identifying urgent open problems. Nine regular papers, published in the refereed, on-site, 141-page hardcopy proceedings [BL2005], established the framework of discussion which embraced three broad topics:

- Performance Analysis of HIPs and CAPTCHAs
- HIP Architectures
- HIPs within Security Systems

Three working groups delved into the topics of "Evaluation Methodologies for HIPs," "Assuring High Performance in HIPs," and "Present and Future HIP Technologies."

Dr. Patrice Simard of Microsoft Research presented an invited talk on "HIP Design: Synthesis, Analysis, and Usability." At the workshop banquet, Dr. Andrei Broder of IBM Research gave the Keynote Address on "The Story Behind Patent No. 6,195,698 (the First CAPTCHA)."

Complete lists of the participants and the regular papers, details of the program, and slides of some of the talks are available at the workshop website `http://www.cse.lehigh.edu/prr/hip2005`. Summaries of the working group discussions will be posted there.

The workshop was organized by Professors Henry Baird and Daniel Lopresti of the Computer Science and Engineering Department at Lehigh University.

## 4 Implications for Image Recognition Research

The emergence of 'human interactive proofs' as a research field offers a rare opportunity (perhaps unprecedented since Turing's day) for a substantive alliance between the image recognition and the theoretical computer science research communities, especially theorists interested in cryptography and security.

At the heart of CAPTCHAs based on reading-ability gaps is the choice of the family of challenges: that is, defining the technical conditions under which text-images can be generated that are reliably human-legible but machine-illegible. This triggers many image recognition research questions, including:

- Historically, what do the fields of Computer Vision and Pattern Recognition suggest are the most intractable obstacles to machine reading, *e.g.*: segmentation problems (clutter, etc); gestalt-completion challenges (parts missing or obscured); severe image degradation?
- What are the conditions under which human reading is peculiarly (or even better, inexplicably) robust? What does the literature in cognitive science and the psychophysics of human reading suggest, e.g.: ideal size and image contrast; known linguistic context; style consistency?
- Where, quantitatively as well as qualitatively, are the margins of good performance located, for machines and for humans?
- Having chosen one or more of these 'ability gaps', how can we reliably generate an inexhaustible supply of distinct challenges that lie strictly 'inside' the gap?

It is well known in the image recognition field that low-quality images of printed-text documents pose serious challenges to current image pattern recognition technologies [RJN96,RNN99]. In an attempt to understand the nature and severity of the challenge, models of document image degradations [Bai92,Kan96] have been developed and used to explore the limitations [HB97] of image pattern recognition algorithms. These methods should be extended theoretically and be better characterized in an engineering sense, in order to make progress on the questions above.

The choice of image degradations for PessimalPrint was crucially guided by the thoughtful discussion in [RNN99] of cases that defeat modern OCR machines, especially:

- thickened images, so that characters merge together;
- thinned images, so that characters fragment into unconnected components;
- noisy images, causing rough edges and salt-and-pepper noise;
- condensed fonts, with narrower aspect ratios than usual; and
- Italic fonts, whose rectilinear bounding boxes overlap their neighbors'.

Does the rich collection of examples in this book suggest other effective means that should be exploited?

To our knowledge, all image recognition research so far has been focused at applications in *non-adversarial environments*. We should look closely at new security-sensitive questions such as:

- how easily can image degradations be normalized away?
- can machines exploit lexicons (and other linguistic context) more or less effectively than people?

Our familiarity with the state of the art of machine vision leads us to hypothesize that no modern OCR machine will be able to cope with the image degradations of PessimalPrint. But how can this informed intuition be supported with sufficient experimental data?

CMU's Blum et al. [BAL00] have experimented, on their website `www.captcha.net`, with degradations that are not only due to imperfect printing and imaging, but include color, overlapping of words, non-linear distortions, and complex or random backgrounds. The relative ease with which we have been able to generate PessimalPrint, and the diversity of other means of bafflement at hand, suggest to us that the range of effective text-image challenges at our disposal is usefully broad.

There are many results reported in the literature on the psychophysics of human reading which appear to provide useful guidance in the engineering of PessimalPrint and similar reading-based CAPTCHAs. [LPSS85] reports on studies of the optimal reading rate and reading conditions for people with normal vision. In [LKT97] an ideal observer model is compared quantitatively to human performance , shedding light on the advantage provided by lexical context. Human reading ability is calibrated with respect to estimates of the intrinsic difficulty of reading tasks in [PBFM02], under a wide range of experimental conditions including varying image size, white noise, and contrast, simple and complex alphabets, and subjects of different ages and degrees of reading experience. These and other results may suggest which image degradation parameters, linguistic contexts, style (in)consistencies, and so forth provide the greatest advantage to human readers.

How long can a CAPTCHA such as PessimalPrint resist attack, given a serious effort to advance machine-vision technology, and assuming that the principles — perhaps even the source code — defining the test are known to attackers?

It may be easy to enumerate potential attacks on vision-based CAPTCHAs, but a close reading of the history of image pattern recognition technology [Pav00] and of OCR technology [NS96] in particular support the view that the gap in ability between human and machine vision remains wide and is only slowly narrowing. We notice that few, if any, machine vision technologies have simultaneously achieved all three of these desirable characteristics: high accuracy, full automation, and versatility. Versatility — the ability to cope with a great variety of types of images — is perhaps the most intractable of these, and so it may be the best long-term basis for designing CAPTCHAs.

Ability gaps exist for other varieties of machine vision, of course, and in the recognition of non-text images, such as line-drawings, faces, and various objects in natural scenes. One might reasonably intuit that these would be harder and so decide to use them rather than images of text. This intuition is not supported by the cognitive science literature on human reading of words. There is no consensus on whether recognition occurs letter-by-letter or by a word-template model [Cro82,KWB80]; some theories stress the importance of contextual clues [GKB83] from natural language and pragmatic knowledge. Furthermore, many theories of human reading assume *perfectly formed* images of text. However, we have not found in the literature a theory of human reading which accounts for the robust human ability to read despite extreme segmentation (merging, fragmentation) of images of characters.

The resistance of these problems to technical attack for four decades and the incompleteness of our understanding of human reading abilities suggests that it is premature to decide that the recognition of text under conditions of low quality, occlusion, fragmentation, and clutter, is intrinsically much easier — that is, a significantly weaker challenge to the machine vision state-of-the-art — than recognition of objects in natural scenes. There is another reason to use images of text: the correct answer to the challenge is unambiguously clear and, even more helpful, it maps into a unique sequence of keystrokes. Can we put these arguments more convincingly?

## 5 Acknowledgments

Our interest in HIPs was triggered by a question — could character images form the basis of a Turing test? — raised by Manuel Blum of Carnegie-Mellon Univ., which in turn was stimulated by Udi Manber's posing the "chat room problem" at CMU in September 2000.

## References

[Bai92] H. S. Baird, "Document Image Defect Models," in H. S. Baird, H. Bunke, and K. Yamamoto (Eds.), *Structured Document Image Analysis*, Springer-Verlag: New York, 1992, pp. 546-556.

[BAL00] M. Blum, L. A. von Ahn, and J. Langford, *The CAPTCHA Project*, "Completely Automatic Public Turing Test to tell Computers and Humans Apart," `www.captcha.net`, Dept. of Computer Science, Carnegie-Mellon Univ., and personal communications, November, 2000.

[Bar01] D. P. Baron, "eBay and Database Protection," Case No. P-33, Case Writing Office, Stanford Graduate School of Business, 518 Memorial Way, Stanford Univ., Stanford, CA 94305-5015, 2001.

[BL05] H. S. Baird and D. P. Lopresti (Eds.), *Proceedings, 2nd Int'l Workshop on Human Interactive Proofs (HIP2005)*, Bethlehem, PA, Springer-Verlag, Lecture Notes on Computer Science, LNCS Vol. No. 3517, Berlin, 2005. [ISBN-10 3-540-26001-3]

[BMW05] H. S. Baird, M. A. Moll,and S-Y Wang, "A Highly Legible CAPTCHA that Resists Segmentation Attacks," in H. S. Baird and D. P. Lopresti (Eds.), *Proc., 2nd Int'l Workshop on Human Interactive Proofs (HIP2005), May 19-20, Bethlehem, PA*, Springer-Verlag, Lecture Notes on Computer Science, LNCS Vol. No. 3517, Berlin, 2005.

[BK02] H. S. Baird and K. Popat, "Human Interactive Proofs and Document Image Analysis," *Proc., 5th IAPR Int'l Workshop on Document Analysis Systems*, Princeton, NJ, Springer-Verlag (Berlin) LNCS 2423, pp. 507–518, August 2002.

[BR05] H. S. Baird and T. Riopka, "ScatterType: a Reading CAPTCHA Resistant to Segmentation Attack," *Proc., IS&T/SPIE Document Recognition & Retrieval XII Conf,*, San Jose, CA, January 16–20, 2005.

[Bro01] AltaVista's "Add-URL" site: `altavista.com/sites/addurl/newurl`, protected by the earliest known CAPTCHA.

[CB03] M. Chew and H. S. Baird, "BaffleText: a Human Interactive Proof," Proc., 10th SPIE/IS&T Document Recognition and Retrieval Conf. (DRR2003), Santa Clara, CA, January 23–24, 2003.

[CBF01] A. L. Coates, H. S. Baird, and R. Fateman, "Pessimal Print: a Reverse Turing Test," Proc., IAPR 6th Intl. Conf. on Document Analysis and Recognition, Seattle, WA, September 10-13, 2001, pp. 1154-1158.

[CLSC05] K. Chellapilla, K. Larson, P. Y. Simard, & M. Czerwinski, "Building Segmentation Based Human-Friendly Human Interactive Proofs (HIPs)," in H. S. Baird & D. P. Lopresti (Eds), *Proc., 2nd Int'l Workshop on Human Interactive Proofs (HIP2005)*, LNCS Vol. No. 3517, Springer (Berlin), pp. 1–26, May 2005.

[Cro82] R. G. Crowder, *The Psychology of Reading*, Oxford University Press, 1982.

[DZ00] D. Lopresti and J. Zhou, "Locating and Recognizing Text in WWW Images," *Information Retrieval*, May, 2000, Vol. 2, No. 2/3, pp. 177–206.

[GKB83] L. M. Gentile, M. L. Kamil, and J. S. Blanchard *'Reading Research Revisited* , Charles E. Merrill Publishing, 1983.

[HB97] T. K. Ho and H. S. Baird, "Large-Scale Simulation Studies in Image Pattern Recognition," *IEEE Trans. on PAMI*, Vol. 19, No. 10, pp. 1067–1079, October 1997.

[HB01] N. J. Hopper and M. Blum, "Secure Human Identification Protocols," In: C. Boyd (Ed.) Advances in Crypotology, Proceedings of Asiacrypt 2001, LNCS 2248, pp.52 -66, Springer-Verlag Berlin, 2001

[Kan96] T. Kanungo, *Document Degradation Models and Methodology for Degradation Model Validation*, Ph.D. Dissertation, Dept. EE, Univ. Washington, March 1996.

[KLB01] T. Kanungo, C. H. Lee, and R. Bradford,"What Fraction of Images on the Web Contain Text?", *Proc., 1st Int'l Workshop on Web Document Analysis*, Seattle, WA, September 8, 2001 (ISBN 0-9541148-0-9) and also at `www.csc.liv.ac.uk/~wda2001`.

[KWB80] P. A. Kolers, M. E. Wrolstad, and H. Bouma, *Processing of Visible Language 2*, Plenum Press, 1980.

[LABB01] M. D. Lillibridge, M. Abadi, K. Bharat, and A. Z. Broder, "Method for Selectively Restricting Access to Computer Systems," U.S. Patent No. 6,195,698, Issued February 27, 2001.

[LBP98] T. Leung, M. Burl, and P. Perona, "Probabilistic affine invariants for recognition," *Proc., IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recogn.*. pp. 678–684, 1998.

[LKT97] G. E. Legge, T. S. Klitz, and B. S. Tjan. "Mr. chips: An ideal-observer model of reading," *Psychological Review* 104(3):524-553, 1997.

[LPSS85] G. E. Legge, D. G. Pelli, G. S. Rubin, and M. M. Schleske, "Psychophysics of reading: I. normal vision," *Vision Research*, 25(2):239-252, 1985.

[MM03] G. Mori and J. Malik, "Recognizing Objects in Adversarial Clutter: Breaking a Visual CAPTCHA," Proc., IEEE CS Society Conf. on Computer Vision and Pattern Recognition (CVPR'03), Madison, WI, June 16-22, 2003.

[NS96] G. Nagy and S. Seth, "Modern optical character recognition." in *The Froehlich / Kent Encyclopaedia of Telecommunications*, Vol. 11, pp. 473-531, Marcel Dekker, NY, 1996.

[Pav00] T. Pavlidis, "Thirty Years at the Pattern Recognition Front," King-Sun Fu Prize Lecture, 11th ICPR, Barcelona, September, 2000.

[PBFM02] D. G. Pelli, C. W. Burns, B. Farell, and D. C. Moore, "Identifying letters," *Vision Research*, [accepted with minor revisions; to appear], 2002.

[Pli02] P. Plitch, "Are Bots Legal?," The Wall Street Journal, Dow Jones Newswires: Jersey City, NJ, `online.wsj.com`, September 16, 2002.

[RNN99] S. V. Rice, G. Nagy, and T. A. Nartker, *OCR: An Illustrated Guide to the Frontier*, Kluwer Academic Publishers, 1999.

[RJN96] S. V. Rice, F. R. Jenkins, and T. A. Nartker, "The Fifth Annual Test of OCR Accuracy," ISRI TR-96-01, Univ. of Nevada, Las Vegas, 1996.

[SCA00] A. P. Saygin, I. Cicekli, and V. Akman, "Turing Test: 50 Years Later," *Minds and Machines*, 10(4), Kluwer, 2000.

[SWI99] D. Shen, W. H. Wong, and H. H. S. Ip, "Affine-invariant Image Retrieval by Correspondance Matching of Shapes," *Image and Vision Computing*, No. 17, pp. 489–499, 1999.

[Tho02] C. Thompson, "Slaves to Our Machines," *Wired* magazine, pp. 35–36, October, 2002

[Tur50] A. Turing, "Computing Machinery and Intelligence," *Mind*, Vol. 59(236), pp. 433–460, 1950.