

Biometrics&CAPTCHA

11/29/2010

EJ Jung

Announcements

- Guest lecture on Wed. December 1st.
 - readings in the [Schedule page](#)
 - [report after the lecture - extra credit for quiz](#)

- Service lab presentation on Mon. December 6th
 - [be timely \(10 minutes\)](#)
 - [spend more time on slides 4 and 5](#)

- HR 535 for Wed. December 8th.

Biometric Authentication

UNIVERSITY of SAN FRANCISCO
department of computer science

- Nothing to remember
- Passive
 - Nothing to type, no devices to carry around
- Can't share (usually)
- Can be fairly unique
 - ... if measurements are sufficiently accurate

Problems with Biometrics

UNIVERSITY of SAN FRANCISCO
department of computer science

➤ Identification vs. authentication

- Identification = associating an identity with an event or a piece of data
 - Example: fingerprint at a crime scene
- Authentication = verifying a claimed identity
 - Example: fingerprint scanner to enter a building

➤ How hard are biometric readings to forge?

- Difficulty of forgery is routinely overestimated
- Analysis often doesn't take into account the possibility of computer-generated forgery

➤ Revocation is difficult or impossible

Biometric Error Rates (Benign)

UNIVERSITY of SAN FRANCISCO
department of computer science

- “Fraud rate” vs. “insult rate”
 - Fraud = system accepts a forgery (false accept)
 - Insult = system rejects valid user (false reject)
- Increasing acceptance threshold increases fraud rate, decreases insult rate
 - Pick a threshold so that fraud rate = insult rate
- For biometrics, U.K. banks set target fraud rate of 1%, insult rate of 0.01% [Ross Anderson]
 - Common signature recognition systems achieve equal error rates around 1% - not good enough!

- Face recognition (by a computer algorithm)
 - Error rates up to 20%, given reasonable variations in lighting, viewpoint and expression
- Fingerprints
 - Traditional method for identification
 - 1911: first US conviction on fingerprint evidence
 - U.K. traditionally requires 16-point match
 - Probability of false match is 1 in 10 billion
 - No successful challenges until 2000
 - Fingerprint damage impairs recognition
 - Ross Anderson's scar crashes FBI scanner

➤ Iris scanning

- Irises are very random, but stable through life
 - Different between the two eyes of the same individual
- 256-byte iris code based on concentric rings between the pupil and the outside of the iris
- Equal error rate better than 1 in a million
- Best biometric mechanism currently known

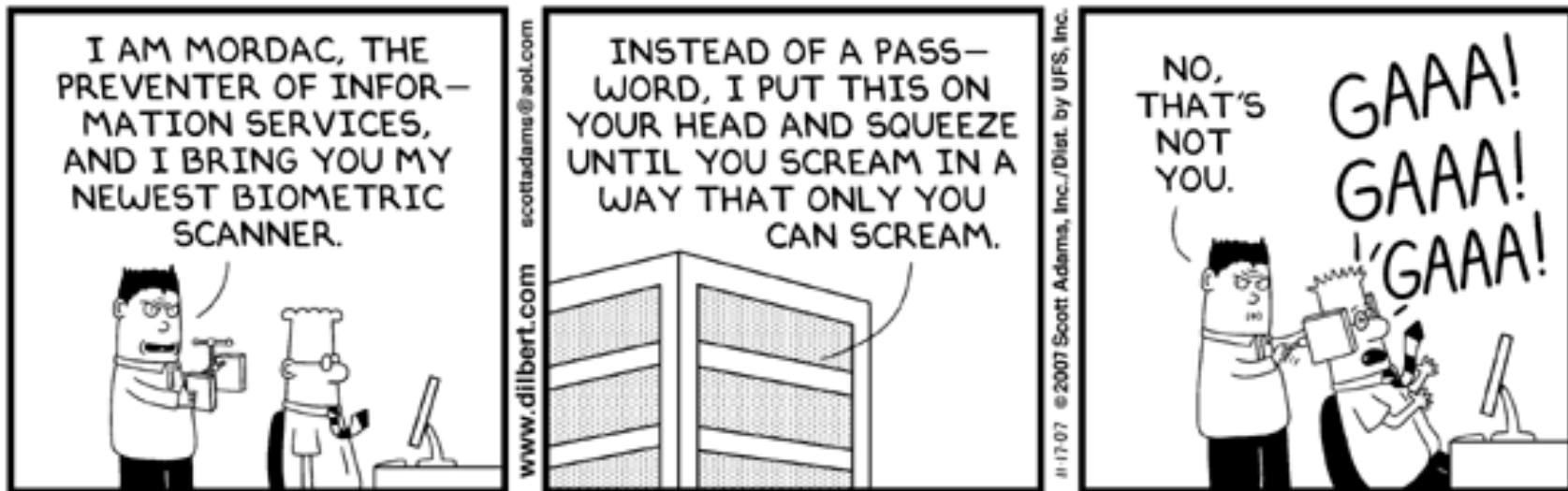
➤ Hand geometry

- Used in nuclear premises entry control, INSPASS (discontinued in 2002)

➤ Voice, ear shape, vein pattern, face temperature

Biometrics (3)

UNIVERSITY of SAN FRANCISCO
department of computer science



© Scott Adams, Inc./Dist. by UFS, Inc.

- Criminal gives an inexperienced policeman fingerprints in the wrong order
 - Record not found; gets off as a first-time offender
- Can be attacked using recordings
 - Ross Anderson: in countries where fingerprints are used to pay pensions, there are persistent tales of “Granny’s finger in the pickle jar” being the most valuable property she bequeathed to her family
- Birthday paradox
 - With false accept rate of 1 in a million, probability of false match is above 50% with only 1609 samples

Bypassing Biometrics

UNIVERSITY OF SAN FRANCISCO
department of computer science

The screenshot shows a web browser window displaying the BBC News website. The browser's address bar shows the URL `http://news.bbc.co.uk/2/hi/asia-pacific/4396831.stm`. The page header includes the BBC logo and navigation links for Home, News, Sport, Radio, TV, Weather, and Languages. Below the header, there are links for UK version, International version, and About the versions. The main content area features the BBC News logo, a 'WATCH One-Minute World News' button, and a globe graphic. The article title is 'Malaysia car thieves steal finger', written by Jonathan Kent, a BBC News reporter in Kuala Lumpur. The article text states: 'Police in Malaysia are hunting for members of a violent gang who chopped off a car owner's finger to get round the vehicle's hi-tech security system.' It further details that the car, a Mercedes S-class, was protected by a fingerprint recognition system and that the incident involved an accountant named K Kumaran.

BBC NEWS | Asia-Pacific | Malaysia car thieves steal finger - Windows

http://news.bbc.co.uk/2/hi/asia-pacific/4396831.stm

KeyGhost Keylogge... BBC NEWS | Asi...

BBC Home News Sport Radio TV Weather Languages

UK version International version | About the versions

BBC NEWS

WATCH One-Minute World News

News Front Page

Last Updated: Thursday, 31 March, 2005, 10:37 GMT 11:37 UK

E-mail this to a friend Printable version

Malaysia car thieves steal finger

By Jonathan Kent
BBC News, Kuala Lumpur

Police in Malaysia are hunting for members of a violent gang who chopped off a car owner's finger to get round the vehicle's hi-tech security system.

The car, a Mercedes S-class, was protected by a fingerprint recognition system.

Accountant K Kumaran's ordeal began when he was run down by four men in a small car as he was about to get into his Mercedes in a Kuala Lumpur suburb.

Forging Handwriting

UNIVERSITY of SAN FRANCISCO
department of computer science

[Ballard, Monroe, Lopresti]

graphic language target	crisis management target	solo concert target
graphic language human forgery	crisis management human forgery	solo concert human forgery
graphic language generative forgery	crisis management generative forgery	solo concert generative forgery

Generated by computer algorithm trained
on handwriting samples

Making an Artificial Finger from a Residual Fingerprint

Materials

A photosensitive coated Printed Circuit Board (PCB)
“10K” by Sanhayato Co., Ltd.



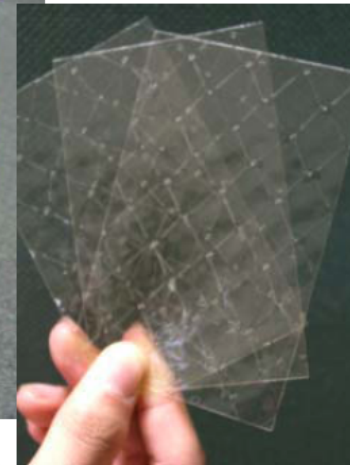
320JPY/sheet



Solid gelatin sheet
“GELATINE LEAF ”
by MARUHA CORP



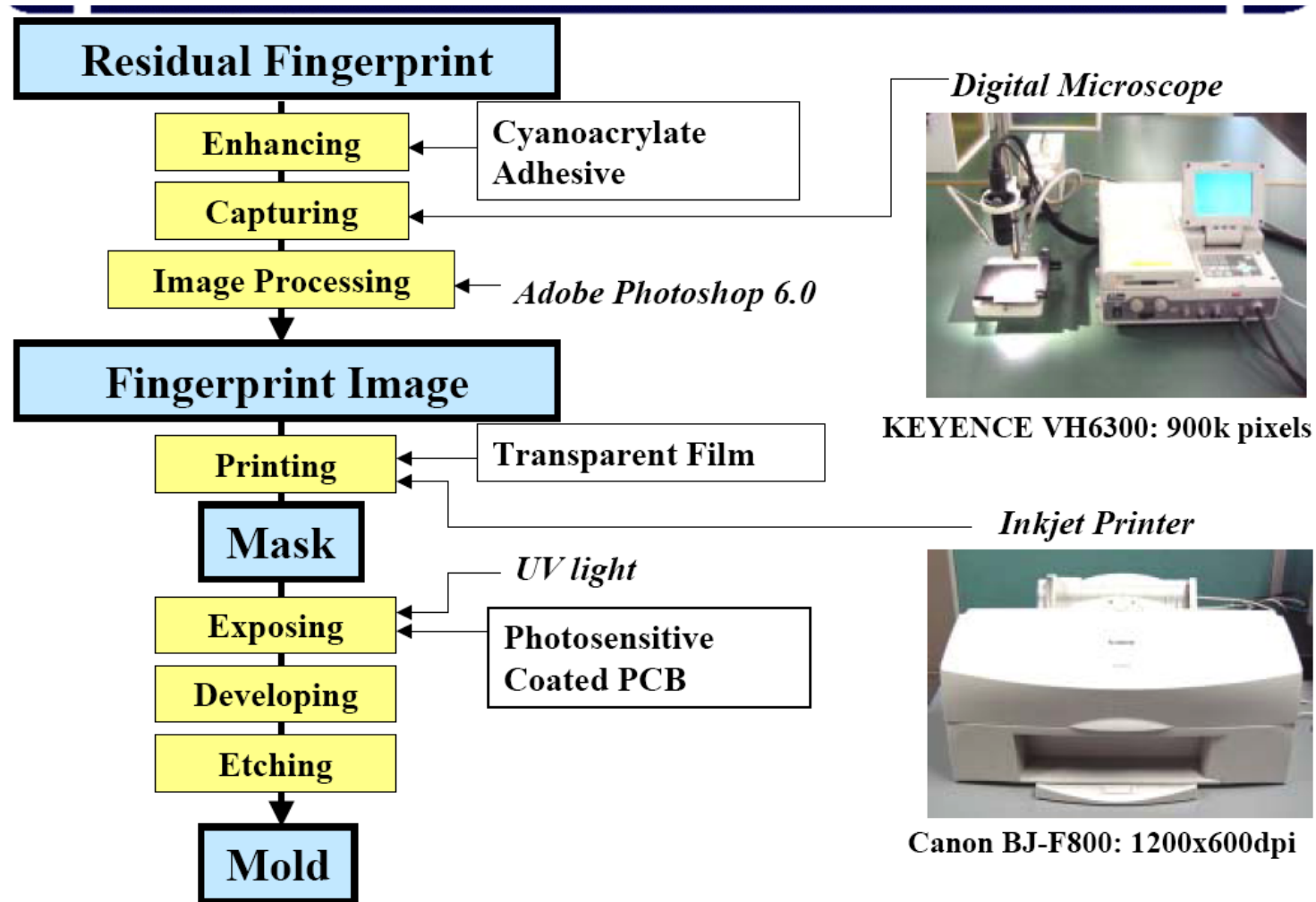
200JPY/30grams



Cloning Process

UNIVERSITY of SAN FRANCISCO
department of computer science

[Matsumoto]



Fingerprint Image

UNIVERSITY of SAN FRANCISCO
department of computer science

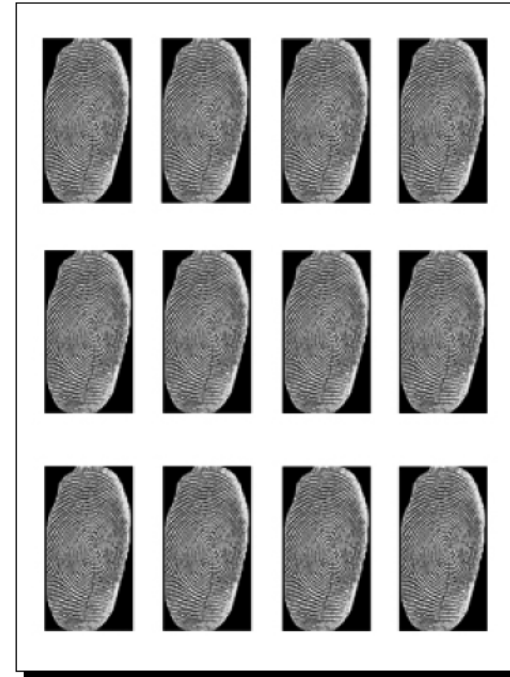
[Matsumoto]



An Enhanced Fingerprint

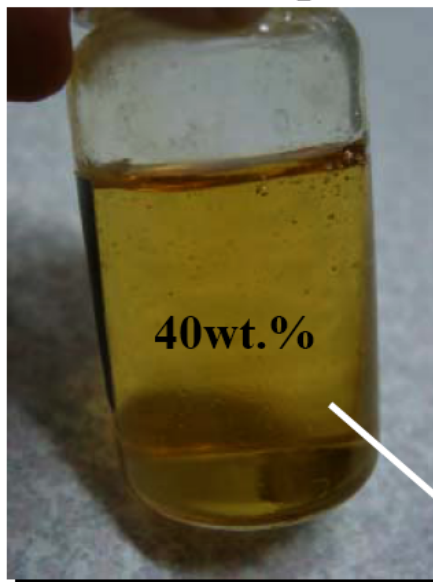


A Fingerprint Image

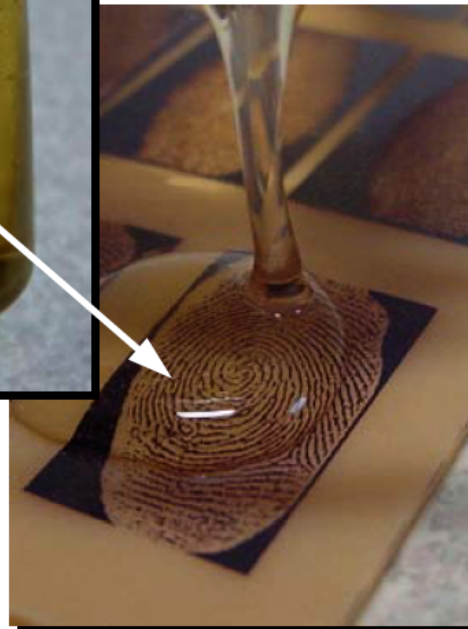


A Mask with Fingerprint Images

Gelatin Liquid



**Drip the liquid
onto the mold.**



**Put this mold into
a refrigerator to cool,
and then peel carefully.**



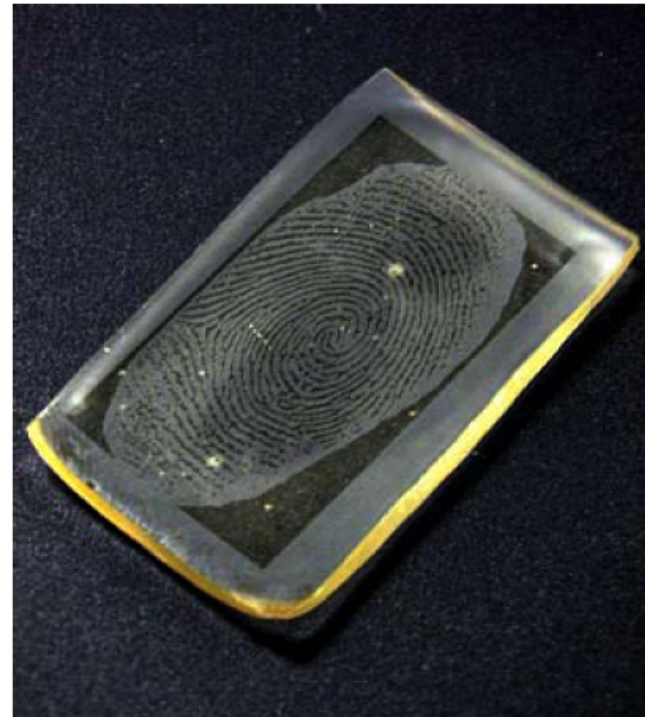
The Mold and the Gummy Finger

UNIVERSITY of SAN FRANCISCO
department of computer science

[Matsumoto]



Mold: 70JPY/piece
(Ten molds can be obtained
in the PCB.)



Gummy Finger: 50JPY/piece

Side By Side

UNIVERSITY of SAN FRANCISCO
department of computer science

[Matsumoto]

Pores can be observed.



Enhanced Fingerprint



Captured Fingerprint Image of
the Gummy Finger
with the device H (a capacitive sensor)

Play-Doh Fingers

UNIVERSITY of SAN FRANCISCO
department of computer science

[Schuckers]

- Alternative to gelatin
- Play-Doh fingers fool 90% of fingerprint scanners
 - [Clarkson University study](#)
- Suggested perspiration measurement to test “liveness” of the finger



CAPTCHA

- stands for Completely Automated Public Turing test to tell Computers and Humans Apart
- Reverse Turing test
 - Turing test: how to tell an intelligent computer
 - from Wikipedia
 - it proceeds as follows: a human judge engages in a natural language conversation with one human and one machine, each of which try to appear human; if the judge cannot reliably tell which is which, then the machine is said to pass the test.
 - remember Blade Runner?
- Human Interactive Proof

Robots can do more and faster

- Botnets can do even more
- Crawlers may ignore robot.txt
- Bots leave malicious contents as comments, postings, emails and collect informations
- Web spam is legal (spam is not)
 - btw, <http://www.ncsl.org/programs/lis/CIP/hacklaw.htm>

Motivation for attack

➤ Search engine

- more links, higher ranking
- e.g. Google's page rank

➤ Advertisement

- mimic "word of mouth"

➤ Phishing

- disguise as suggestions and recommendations

Motivation Beyond the Web

- Prevent dictionary attacks in any password system (Pinkas & Sander)
 - after failures, as for CAPTCHA and the password
- Deter massive attacks
 - botnets may not pass CAPTCHA
 - humans are much slower
 - ask for CAPTCHA for any suspicious activity

Precursors

- Unpublished manuscript by Moni Naor first mentions automated Turing test in 1997, but not proposed or formalized.
- Alta Vista patent in 1998 first practical example of using slightly distorted images of text to deter bots.
 - broken later by OCR

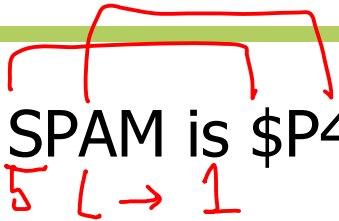
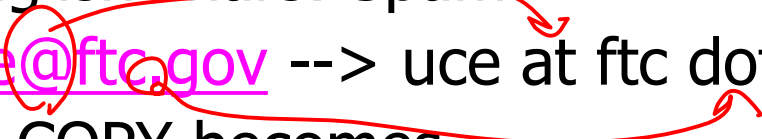
Definition

- In 2000, formalized by Luis von Ahn, Manuel Blum & Nicholas J. Hopper of Carnegie Mellon; John Langford of IBM
- “A CAPTCHA is a cryptographic protocol whose underlying hardness assumption is based on an AI problem.”
- www.captcha.net
- Advancing AI and security together
 - battle of breaking and improving

General Approaches

- Text (ASCII/Unicode)
- Image
- Speech
- Animation
- 3-D
- Combinations of all above

ASCII/Unicode ©4Ptçh4

- Change text to look-alike: SPAM is \$P4M. Fools simplest text matching.

- Accented or non-English chars: Spám
- Chars to words: uce@ftc.gov --> uce at ftc dot gov

- URL/HTML entities: COPY becomes
¢0Ρ¥ or %430P%59
- Better than nothing, but easy to crack
- This is not technically CAPTCHA

Text Based CAPTCHAs

- Gimpy, ez-gimpy
 - Pick a word or words from a small dictionary
 - Distort them and add noise and background
- Gimpy-r
 - Pick random letters
 - Distort them, add noise and background
- Simard's HIP
 - Pick random letters and numbers
 - Distort them and add arcs

Text Based CAPTCHAs



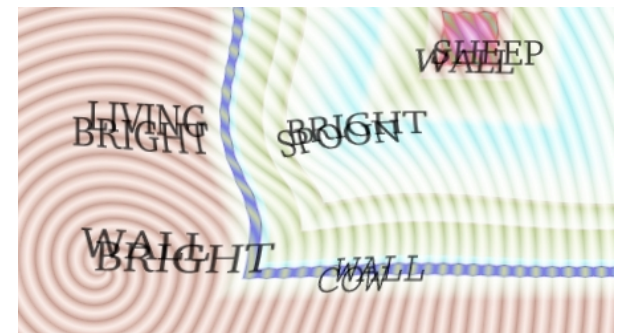
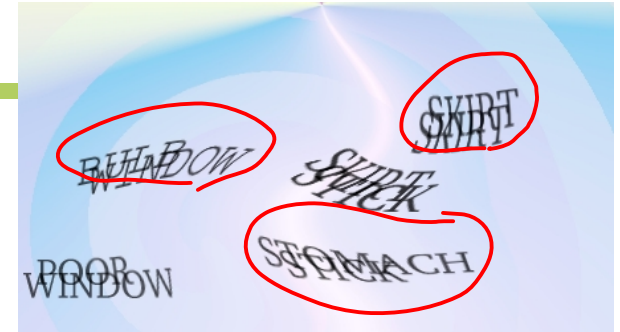
sciarna

adviser



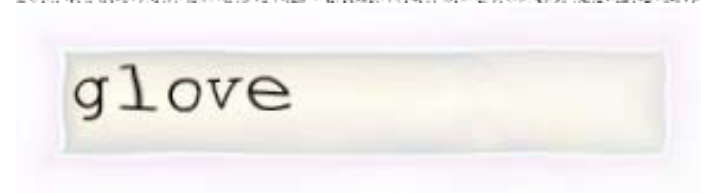
Gimpy

- First generation
 - Pick a word from dictionary
 - Random placement, font, distortion, background pattern
 - Overlapping words serve as noise.
- Frequently cracked and improved.
 - <http://www.cs.sfu.ca/~mori/research/gimpy/>
- In current version, 5 pairs of overlapped words. User identifies 3 words.



EZ-Gimpy

- Pick a word or words from a small dictionary
- Distort them and add noise and background
- 99% success in breaking
 - Distortion Estimation Techniques in Solving Visual CAPTCHAs, CVRP 2004



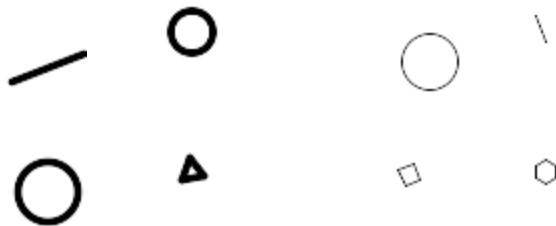
Gimpy-r

- Pick random letters
 - Distort them, add noise and background
- 78% success in breaking Gimpy-r
- Distortion Estimation
Techniques in Solving Visual CAPTCHAs, CVRP 2004

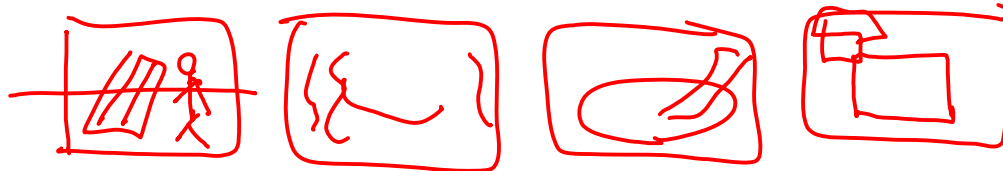


Bongo

- Visual pattern recognition puzzle
- Example: thick vs. thin
- User is presented with a new block and needs to pick left or right



- Image recognition with keywords
- Procedure
 - display four images with the same keyword
 - provide a random set of keywords to choose from
 - user needs to pick the common keyword



ESP-Pix



LIVESTOCK

Choose a word that relates
to all the images.



TIP: You can type the first letter
of a word and then use the down
arrow to find it.

Submit

Beating CAPTCHA

➤ OCR-base attacks

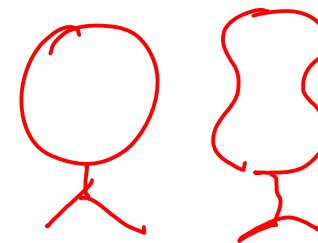
- <http://sam.zoy.org/pwntcha/>
- *Pretend We're Not a Turing Computer but a Human Antagonist*

➤ Heuristics

- vary position, warp, noise, background, colors, overlap, randomness, font, angles, language,

➤ Accessibility problem for vision-impaired users

- audio as well as visual
- <http://www.w3.org/TR/turingtest/>



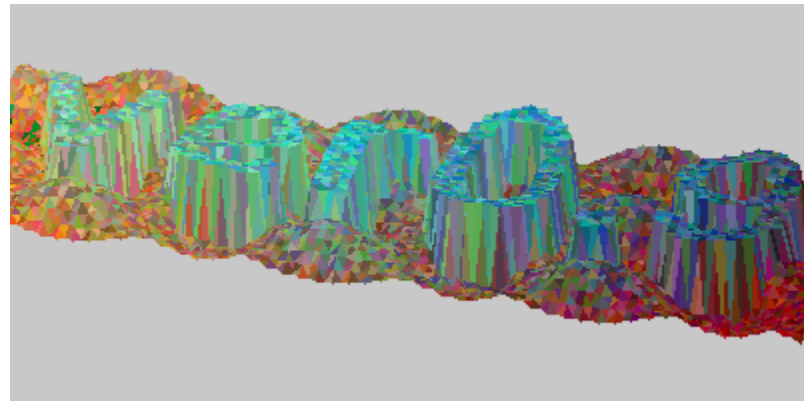
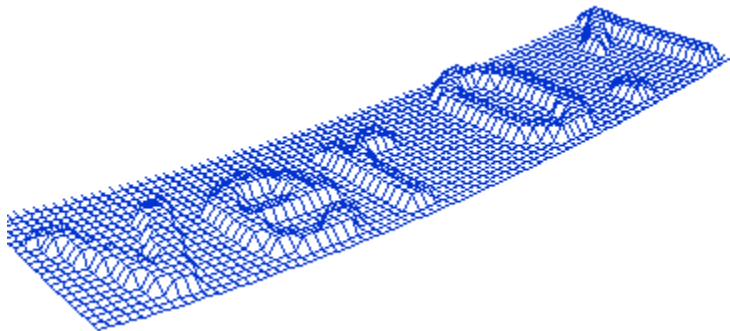
Speech CAPTCHA

- Spell in synthesized or recorded voices
- Voice recognition vs. user's miss rate
- Use with visual CAPTCHA for increased accessibility
 - may help attackers guess correctly

Animated CAPTCHA

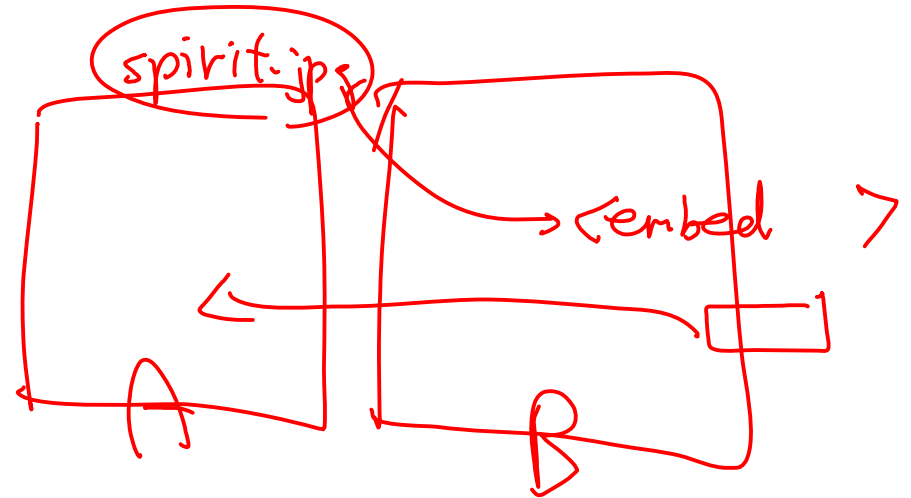
- Can use Flash, MPEG, animated GIF
- Often combined with speech
- Weaknesses of Image CAPTCHA apply
- Usually *easier* to crack due to extra data for pattern matching to analyze
- Much higher processor and traffic load
- Not practical in most cases

- tEABAG_3D
 - <http://www.ocr-research.org.ua/index.php?action=teabag>
- Renders the password in 3D image
- More difficult to crack then 2D images
- More resources on server
 - [high load graphic processing](#)
- Can be combined with other methods



Beating CAPTCHA by humans

- Man-in-the-middle
 - copy CAPTCHA from the target
 - post on the attacker's website
 - forward the answer to the target
- CAPTCHA factory
 - <http://taint.org/2008/03/05/122732a.html>
- Reuse the session id
 - http://www.puremango.co.uk/cm_breaking_captcha_115.php



Adopt CAPTCHA for yourself?

➤ Free software

- <http://captcha.net>