

State of the Art in Anti- Automation Technologies in Today's Web

Markus Haudum
Stefan Panhuber

RockPort



Übersicht

- Einführung (HIP)
 - Yahoo
 - Altavista
- Variationen von CAPTCHAs
 - Gimpy, EZ-Gimpy
 - PIX
 - Bongo
 - Sounds
 - Formeln
 - Passfaces
- CAPTCHAs erstellen
 - Textuell
 - Akustisch





Einführung – HIP (CAPTCHA)

- HIP – Human Interaction Proof
 - Unterscheidung zwischen Mensch und Maschine
- CAPTCHA
 - Completely
 - Automatic
 - Public
 - Touring test
 - to tell Computers and Humans Apart.
 - Teile von HIPs (Web-Application)





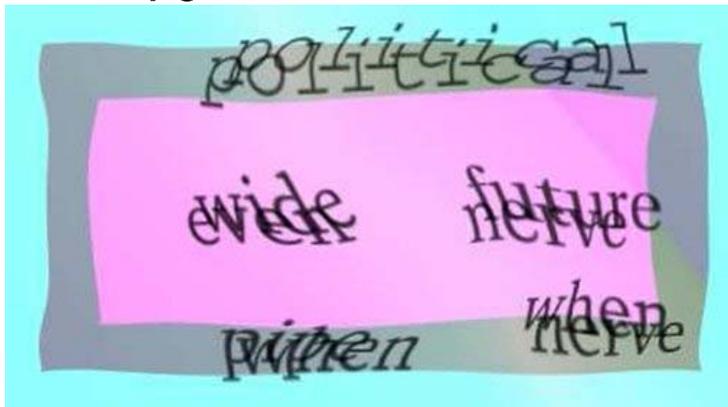
Einführung –Yahoo/Altavista

- 1997: Anfänge des Spam Problems.
- Altavista
 - Bots manipulierten Ranking von Webseiten.
- Yahoo
 - Automatische E-Mail Registrierungen zwecks Spam.
 - Werbe-Bots in Chatrooms.

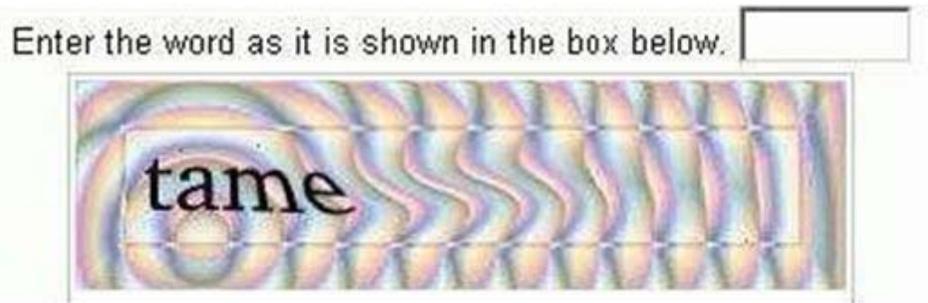


CAPTCHAs - Gimpy, EZ-Gimpy

- Konzept: Erkennung von Wörtern in einem manipulierten Bild.
 - Gimpy: Erkenne Drei Wörter.

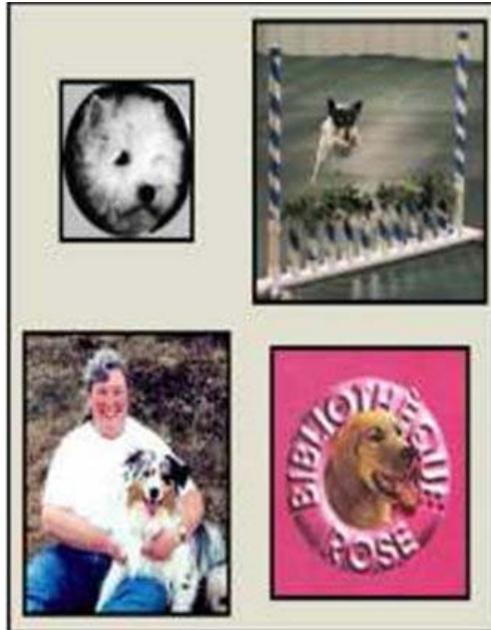


- EZ-Gimpy (Easy Gimpy): Erkenne ein Wort.



CAPTCHAs - PIX

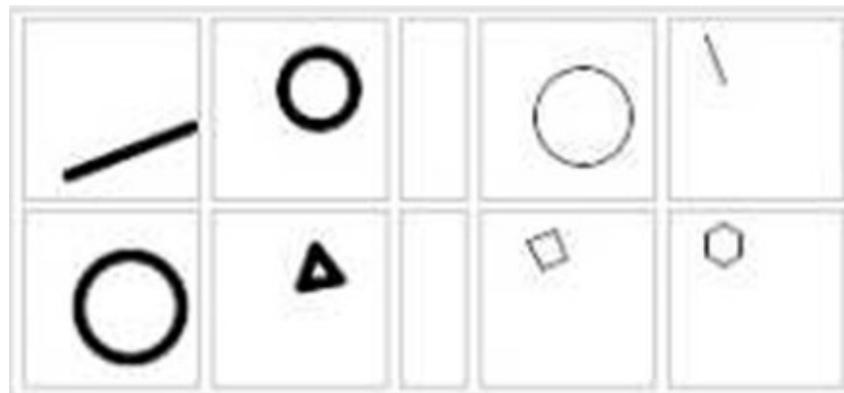
- Konzept: Erkenne aus einer Anzahl von Bildern die eindeutige Gemeinsamkeit.
- Umsetzung: In einer Datenbank sind Bildern Wörter zugeordnet.



- Lösung: Hund

CAPTCHAs - Bongo

- Konzept: Finde aus zwei getrennte Sets von Bildern die sich in genau einem Attribut unterscheiden jenes Attribut. Dann ordne den Sets basierend auf dem Attribut neue Bilder hinzu.



- Attribute
 - Liniendicke
 - Form der Objekte
 - Anzahl der Bilder
 - ...



CAPTCHAs - Sounds

- Konzept: Spiele eine modifizierte Sound-Datei ab, die der Benutzer erkennen muss.
- Einsatzgebiet: Kombination mit Gimpys.
- Barrierefreiheit: Blinden Menschen wird es durch die akustische Wahrnehmung ermöglicht CAPTCHAs zu lösen.



CAPTCHAs - Formeln

- Konzept: Präsentiere eine Gleichung, deren Lösung das CAPTCHA löst.
- Unterscheidung zwischen Erwachsenen und Kindern möglich

Please enter in the Calculation Result in the box below.

=

CAPTCHAs - Passfaces

■ Konzept

- Menschen erkennen Personen leicht wieder.
- Gib dem Benutzer Bilder von Menschen die er sich merken muss.
- Präsentiere dem Benutzer für jedes gemerkte Bild eine Matrix.
- Erkenne die gemerkte Person wieder.





Design - Anforderungen

- “Machines won't stay stupid forever. Range of problems they can solve is growing reasonably rapidly – it certainly isn't shrinking. ” - *Daniel Lopresti 2005*

- Kriterien für effektive CAPTCHAs:
 - Effektivität um Maschinen auszusperren
 - Tollerante, machbare Tests für Menschen

- “While today there are a large number of generative CAPTCHA's to choose from, someday we may run out of tests that meet both criteria (hard for machines, tolerable to humans). Should we be concerned? ”
-Daniel Lopresti 2005





Design – Text CAPTCHA

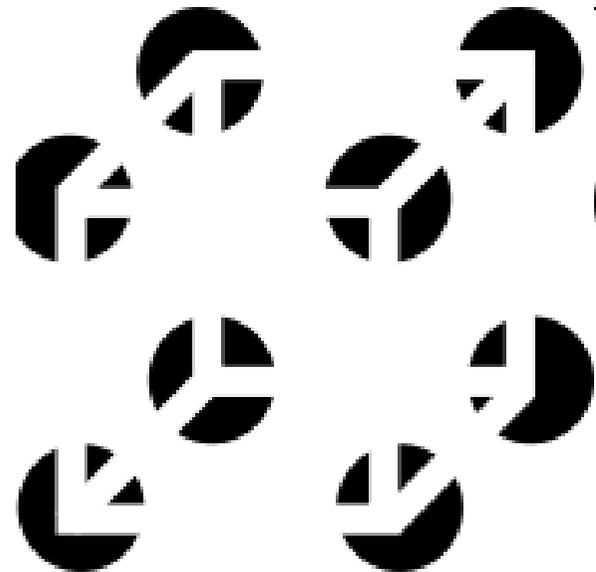
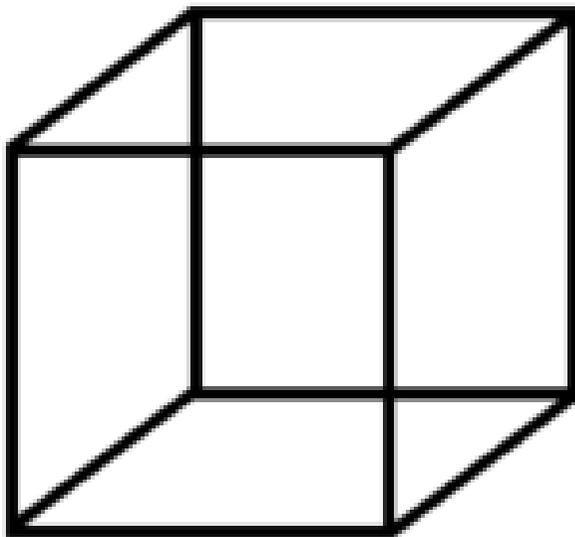
- Wahrnehmung: Mensch gegen Maschine
 - Bildqualität
 - Hintergrundrauschen
 - Belichtung
 - Bildeigenschaften
 - Strecken
 - Zerren
 - Rotation
 - Bildfehler
 - Farbmischung
 - Zerlegung
 - Doppelbuchstaben
 - Kritzeleien
 - Absichtliche Fehler



Design – Text CAPTCHA

■ Gestalt Laws

- Gestaltpsychologie
- Anfang 20. Jahrhunderts



Design – Text CAPTCHA

■ Gestalt Laws I

- Law of closure



- Law of similarity

OXX
XOX
XXO





Design – Text CAPTCHA

■ Gestalt Laws II

● Law of proximity: * * * * *

* * * * *

* * * * *

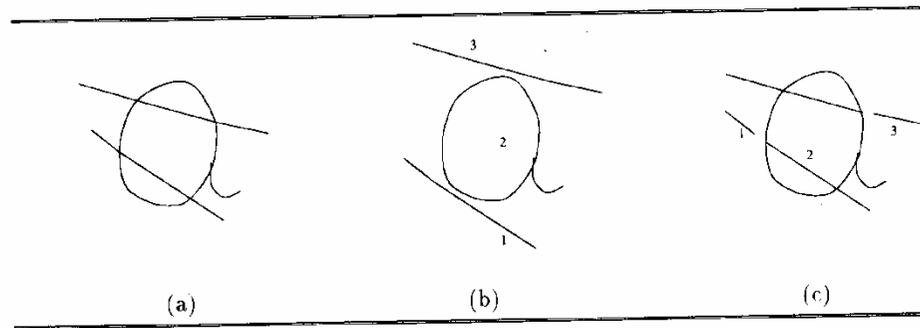
● Law of symmetry: [] [] []



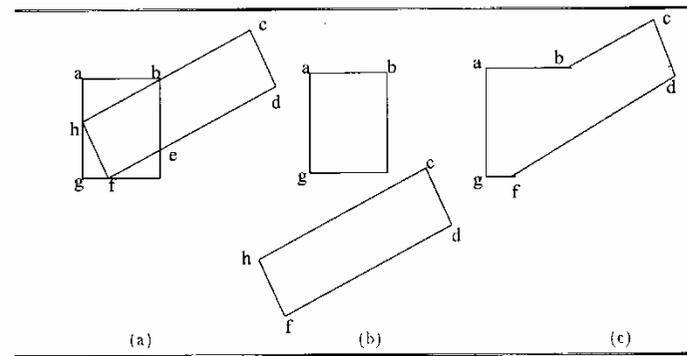
Design – Text CAPTCHA

■ Gestalt Laws III

- Law of continuity



- Law of familiarity



Design – Text CAPTCHA

■ Gestalt Laws IV

- Law of Figure and Ground



- Memory and internal metrics

NWOD-EDISPU GNIDEB
READING UPSIDE-DOWN





Design – Akustische CAPTCHA

■ Design Regeln

- **Schwerer Dialog**
 - Komplizierte Satzbau
 - mehrdeutige Wörter
- **Kompliziertes Sprechsignal**
 - Rauschen
 - Schlechte Qualität
- **Akustische Illusionen**
 - Für den Menschen leicht erkennbar
 - Maschinen filtert etwas anderes heraus





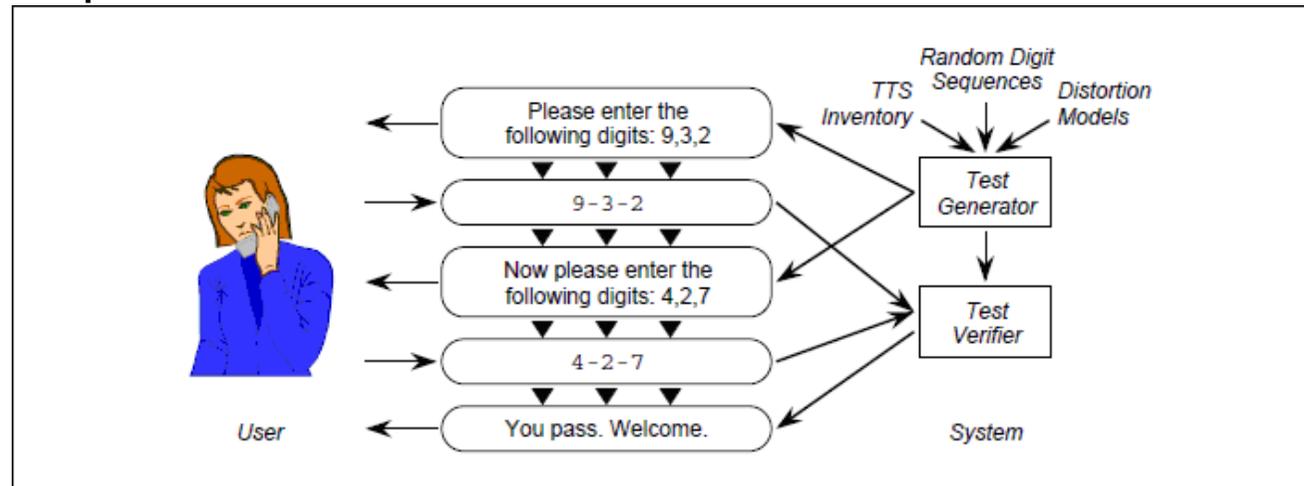
Design – Akustische CAPTCHA

- Probleme mit akustischen CAPTCHAs
 - Keine Tastatur vorhanden (öffentliche Internetstationen, usw.)
 - Wörter die sehr ähnlich Klingen, bzw. für das selbe Wort gehalten werden.
 - Viele Menschen können nicht richtig Buchstabieren



Design – Akustische CAPTCHA

■ Beispielprotokoll



■ Intelligente Ansätze (Beispiele)

- Welche Nummer kommt vor/nach einer bestimmten Nummer m ($0 < m < 1000$)
- Welcher Wochentag kommt vor/nach dem Wochentag x



Ende...

- Vielen Dank für eure Aufmerksamkeit!!!

