ランダム妨害図形を用いた画像ベースCAPTCHAの検討

田村 拓己[1,a)]　池田 匡視[2]　岡崎 直宣[1,b)]

概要：近年，WEB サービスが急激に普及する中で，それらの WEB サービスに対してボットと呼ばれる自動プログラムを使用し，不正にサービスを利用するという悪質な行為が問題となっている．このような問題を防止するために，CAPTCHA と呼ばれる反転チューリングテストが広く利用されている．しかし，近年，CAPTCHA を自動的に突破する技術が発達し，その脆弱性が多くの研究者に指摘されている．そこで本論文では，画像ベースの新たな CAPTCHA 方式を提案し評価する．本提案手法では，機械では実現することが難しい人特有の画像認識能力を利用し，高いユーザビリティと同時に，提示画像の中に答えとなる文字を全く表示しないことで OCR 機能を利用するボットと人間の高い判別率を実現する．

キーワード：CAPTCHA，ボット，反転チューリングテスト

An Examination of the Image-based CAPTCHA Using Random Obstruction Figures

Tamura Takumi[1,a)]　Ikeda Masami[2]　Okazaki Naonobu[1,b)]

Abstract: A reversal Turing test called CAPTCHA is used in many web-sites in order to prevent from bot-attack. However, the automatica lly breaking technology against CAPTCHA has developed, and many researchers have pointed out the vulnerability of CAPTCHA. In this paper, we propose a new image-based CAPTCHA using random obstruction figures, to achieve high usability and high res istance against OCR-based bot-attack. We paid attention to the image recognition ability unique to humans. In the proposed m ethod, the resistance against bot-attack is achieved by not allowing be displayed the characters of answer in the image.

Keywords: CAPTCHA, Bot, Inverted Turing test

1. はじめに

近年，WEB サービスの普及により，誰でも様々なサービスを利用することが可能となっている．しかし，それらの WEB サービスに対してボットと呼ばれる自動プログラムを使用し，不正にサービスを利用するという悪質な行為が最近問題視されている．このような問題を防止するためには，人間とボットを識別する反転チューリングテストが必要となり，現在，CMU の研究者によって開発された CAPTCHA と呼ばれる方式が広く利用されている [1]. CAPTCHA とはチャレンジ/レスポンス型テストの一種であり，対象者が人間であるか機械であるかを判別する．一般的に利用されている手法としては，歪曲やノイズが付加された文字列画像を WEB ページに提示し，閲覧者がその文字を判読できるか否かを試すものがある（図 1）.

しかし，近年，CAPTCHA を自動的に突破する技術が発達し，その脆弱性が多くの研究者に指摘されている．例えば，文字列の判読能力を試す CAPTCHA においては，すでに高機能な OCR（自動文字読取）機能を備えるボットが出回るようになっている [2][3]．その対策として，文字列に加える変形やノイズを大きくすることによってボットを排除

[1] 宮崎大学 工学部
Faculty of Engineering, University of Miyazaki
[2] 宮崎大学大学院 工学研究科
University of Miyazaki
[a)] tse926u@student.miyazaki-u.ac.jp
[b)] oka@cs.miyazaki-u.ac.jp

© 2012 Information Processing Society of Japan

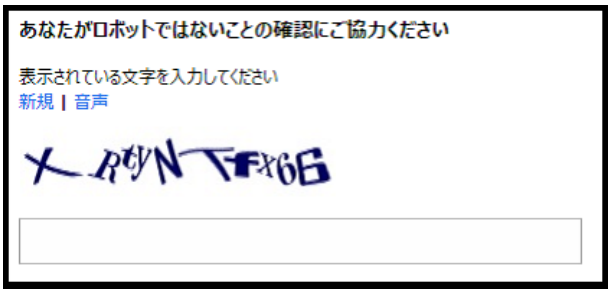

図 1 Microsoft で利用されている CAPTCHA（文字列 CAPTCHA）
Fig. 1 CAPTCHA used by Microsoft

する確率を向上させることはできるが，そのような文字は人間にとっても認識が困難になるため，人間の正答率まで低下させてしまう．この問題に対し，画像や音声をベースにした，人間のより高度な知識処理を利用する CAPTCHA [5] も提案されているものの，一部の手法ではボットによる突破が可能であるという指摘もされている [4]．さらに，ボットの能力（CAPCTHA 解読アルゴリズム，および PC の CPU パワー）の向上は留まるところを知らない．したがって，ボットがいかに高度になろうとも，ボットによる解答が依然として困難な，新たな CAPTCHA の導入が強く望まれる．ただし，CAPTCHA は，安全性とユーザビリティがトレードオフの関係になっていることに留意しなければならない．

そこで本論文では，画像ベースの新たな CAPTCHA 方式を提案する．本提案手法では，人間の視覚補完を利用することと，画像を使用することでユーザビリティを確保しつつ，提示画像の中に答えとなる文字を全く用いないことで，OCR 機能を備えるボットの突破率を低下させる．また，ランダムで多数の種類の妨害図形を用いることと，使用画像を毎回インターネット上で検索し収集することでデータベースを用いた攻撃に対して耐性を持たせる．

2. 関連研究

2.1 CAPTCHA について

CAPTCHA（Completely Automated Public Turing test to tell Computers and Humans Apart：コンピュータと人間を区別する完全に自動化された公開チューリングテスト）は 2000 年にカーネギーメロン大学の Luis von Ahn, Manuel Blum, Nicholas Hopper, John Langford によって作られた．人間には容易に解くことが可能であるが，コンピュータには解くことが難しいものを出題し，正しい解答をした者を人間と判断する，チャレンジ/レスポンス型テストの一種である．

CAPTCHA では，人間と機械を区別するために，画像や音声，文字列などを用いる方式がある．次節から既存の CAPTCHA について紹介する．

2.2 文字列 CAPTCHA

現在，最も広く利用されている CAPTCHA は文字列 CAPTCHA である．文字列 CAPTCHA には Gimpy [6], EZ-Gimpy [6]，r-Gimpy [7] などの種類がある．

Gimpy は，2 つの単語が重複して印刷されているものを 1 セットとしたとき，ある画像の中にそれを 5 セット表示し，その 10 個の単語の中から 3 つを答えさせそれが正しければ解答者を人間と判別するものである．EZ-Gimpy 及び r-Gimpy は Gimpy を単純化したもので，1 つの単語，あるいはアルファベットと数字をランダムに並べた文字列の画像を歪ませて表示し，その答えをテキストボックスに入力させ，解答が正しければ解答者を人間と判別する．

文字列 CAPTCHA のメリットとしては，システムとして単純であり，Web サイトに簡単に取り入れることが可能である点と，総当たり攻撃に高い耐性を持つ点が挙げられる．

デメリットとしては，近年の OCR（自動文字読取）の性能向上により，ボットでも簡単に文字を認識できるようになっていることが挙げられる．Gimpy，EZ-Gimpy では，まず，画像の中の文字をアルファベット 26 字と比較し，最も似ていると考えられる候補を 5 つほど選び出し，それらの組み合わせが単語になるかどうかを調べる．単語になったものを取り出し，その画像を歪め，最もよく元の画像と一致するスコアの単語を取り出すことによってこの CAPTCHA を破る．Mori らは，191 個の EZ-Gimpy に対して攻撃テストを行い，結果は 83 %の突破率であった [6]．これは，非常に高い数値であり，早急に対策が必要である．

2.3 画像 CAPTCHA

文字列 CAPTCHA と違い，画像 CAPTCHA は文字を使用しない．

ここでは例として Asirra [8] を挙げる．Asirra は，人間とボットのイヌとネコを見分ける能力の違いに基づいている．まず利用者に 12 枚のイヌまたはネコの画像を提示する．そして，利用者はネコ画像を全て選択する．ネコの画像を正しく選択できれば，利用者を人間と判別する（図 **??**）．

Asirra の長所として，イヌとネコの分類には専門的・文化的知識は必要とせず，人間は素早く正確に解くことができることが挙げられる．Elson らの 4717 枚の画像と 147 人の利用者を用いた実験では，人間は 98.5% の正確さでイヌとネコを見分けることができた．この場合，人間の利用者は 83.4% の正確さで，12 枚の画像を用いた Asirra を解くことができる．また，飼い主のいない動物の画像を用いることでそれらの動物の飼い主探しにつながるといった，他の利点も同時にある．

短所としては，テキストベースの CAPTCHA に比べて大きなスペースを使うこと，視覚障害のある人はアクセス

© 2012 Information Processing Society of Japan

できないこと，が挙げられる．また，近年，サポートベクターマシン（SVM）を用いた機械学習によって Asirra が破られた [4]．SVM はデータを 2 つのクラスに分離する超平面を作る教師つき学習方法である．8000 枚の画像で学習し 2000 枚の画像で実験したとき 82.7% の正確さで分類することができた．Asirra のテストを解くためには 12 枚の画像を分類する必要があるため，3 つの分類のうち一番精度の高い手法で Asirra の突破を試みた場合，突破できる可能性は 10.3% である．

3. 提案手法

本章では，3.1 節で提案手法の目的を述べ，3.2 節で満たすべき要件について説明し，3.3 節で，実際に提案手法を用いて CAPTCHA を作成する手順を紹介する．

以降，提案する CAPTCHA 方式を IC-CAPTCHA（Image-based Character input type CAPTCHA）と呼ぶ．

3.1 目的

既存手法では，OCR 機能を持つボットに対する耐性の低さとユーザビリティの低さが問題であった．従って提案手法では，画像 CAPTCHA におけるデータベース攻撃に対する耐性と総当たり攻撃に対する耐性に重点をおき，既存の文字列 CAPTCHA と比べて，OCR 機能を持つボットに対する耐性をもち，ユーザビリティに配慮した CAPTCHA を作成することを目的とする．

3.2 満たすべき要件

3.1 節より，満たすべき要件を 2 つ挙げる．

（1）データベース攻撃に対する耐性

画像 CAPTCHA における脆弱性にデータベース攻撃がある．データベース攻撃というのは，問題画像とその解を記録したデータベースを構築し，このデータベースを用いて問題を解く方法である．これは，画像 CAPTCHA のシステムに使用される画像枚数が有限であることが原因となる．従って，データベース攻撃に対する耐性を持つ CAPTCHA を生成するためには，画像 CAPTCHA システム内で使用する画像枚数に，なるべく制限が無いようなシステムであることが望ましいといえる．

（2）総当たり攻撃に対する耐性

総当たり攻撃というのは，暗号や暗証番号などで理論的にありうる全てのパターンを入力し解読する暗号解読法である．画像ベース CAPTCHA においては，方式自体が並べ替え方式，クリック方式，種類の分別方式，など解答の組み合わせの最大数が少ないものが多い．これは，総当たり攻撃において脆弱であるといえる．画像ベースの CAPTCHA の場合，たとえ総当たり攻撃に対して，一定時間内に規定回数以上の解答誤入力があった場合に規制をかける，解答時間に制限を設ける，などの対策を講じていたとしても，銀行 ATM に用いられている認証方式 PIN（1/10000）程度の強度を保つことが望ましいと考える．

3.3 IC-CAPTCHA システム

前提として，IC-CAPTCHA システムは画像から名詞を導き出せるであろう名詞群からなる名詞辞書と加工後画像のハッシュ値を登録したハッシュ辞典を持つ．以下に IC-CAPTCHA システムの画像生成手順を示す．

【IC-CAPTCHA システム画像生成手順】

© 2012 Information Processing Society of Japan

Step1. IC-CAPTCHA システムの持っている名詞辞書からランダムに 1 つの名詞を選ぶ.

Step2. その名詞，あるいは名詞に結びついている画像を検索エンジンを用いて検索し，その名詞に基づく画像を 1 枚取得する.

Step3. Step.2 の画像に，妨害図形の上書きをし，背景処理（回転，モザイク，ぼかし，色反転等）の画像処理を施す.

Step4. Step.3 の後の画像のハッシュ値をとり，そのハッシュ値でハッシュ辞典を検索し，まだ登録されていなければハッシュ辞典に登録する．もし登録されている場合は，Step.1 から画像を作り直す.

Step5. ユーザに画像を提示する.

Step6. ユーザは，画像から名詞を推測し，テキストボックスに名詞を入力する.

Step7. IC-CAPTCHA システムは，Step.1 で選んだ名詞とユーザの入力した名詞を比較し，マッチしたならば，ユーザを人間と認識し，認証する．マッチしなかった場合，2 回目までは Step.5 へ戻る．3 回目は，Step.1 へ戻り，名詞を変更する.

□

IC-CAPTCHA 作成手順のフローチャートを図 2 に示す.

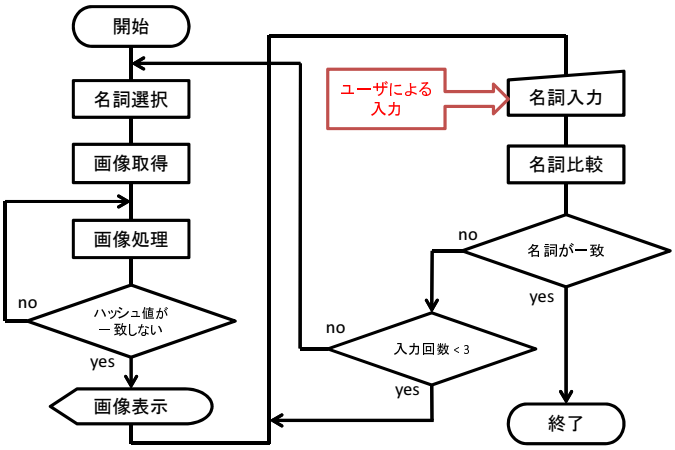

図 2 IC-CAPTCHA のフローチャート
Fig. 2 Flowchart of IC-CAPTCHA

まず，Step.1 で使用する名詞辞書の単語登録数がそのまま総当たり攻撃における耐性になる．名詞辞書の単語登録数が少ないと総当たり攻撃に脆弱となってしまうが，他の方式と違い，画像の選び方や並べ方と関連しないため，名詞辞書の登録数を増やすことは比較的容易である．ただし，実用レベルでの使用を考えるならば，名詞辞書の登録数は，10000 語程度まで増やす必要がある.

また，Step.3 では検索エンジンを用いて画像を毎回検索し，収集することでデータベース攻撃に耐性を持たせる．もし，同じ画像を加工することになったとしても，ランダムな妨害図形と背景処理を施すため，加工後に全く同じ画像になることはほぼないと思われる．万が一，全く同じ加工画像を生成したとしても，ハッシュ辞典を用いることで全く同一の加工画像は廃棄するため，全く同じ加工画像がユーザにさらされることはない.

IC-CAPTCHA システムでは，画像内に答えと結びつく文字列は全く表示されないため，OCR 機能を持つボットに対しての耐性は，考慮する必要はない.

4. 実装

4.1 開発環境

開発言語は C++，ライブラリは Open CV を用いて，仮想 PC 上の Ubuntu11.10 にて実装した.

4.2 実装プログラム

本研究では，IC-CAPTCHA システムの CAPTCHA 生成画像が安全性と利便性において重要であり，IC-CAPTCHA システムの根幹を成しているため，IC-CAPTCHA 作成手順の Step.1，Step.3，Step.5 を実装した.

Step.2 の Web 検索を用いた画像収集では，検討していた Google 画像検索において自動プログラムを用いた使用に制限があったため，画像収集については，予め，素材となる複数の画像をそれぞれの名詞ごとに収集し，名詞ごとにフォルダを分別した．その名詞自体と，その名詞フォルダに何枚の画像が格納されているかを示したテキストファイルより辞書の読み込みを行った.

実装プログラムの画像処理として，背景処理では，モザイク，ぼかし，色反転，画像回転を用い，上書きする妨害図形には，円，楕円・扇型，ポリゴン（多角形），文字を用いた．実際に実装したプログラムによる IC-CAPTCHA システムの生成画像は図 3 のようになる.

図 3 生成画像の例（りんご）
Fig. 3 exsample of IC-CAPTCHA image (apple)

5. 評価および考察

提案手法が既存手法のユーザビリティを改善し，また，

© 2012 Information Processing Society of Japan

画像 CAPTCHA の脆弱性であるデータベース攻撃に耐性を持つのか調査するため，実装した CAPTCHA 作成プログラムで生成した画像を用いて，データベース照合評価とユーザビリティ評価を行った．

5.1 データベース照合評価

データベース評価では，Google 画像検索を用いて，提案手法の生成画像を検索することで，元の名詞が推測されるか，類似画像として名詞の画像が検出されるかを調査した．

2 つのパラメータを変更し，5 パターンの調査をした結果，妨害図形数 11 のときの該当率が一番低く，値は 10.5% であった．パラメータとして，画像の出現範囲も変更したが，その数値に関係なく，妨害図形数の増加に比例して該当率は減少したことから，画像検索を用いたデータベース照合においては妨害図形数が重要であることがわかった．また，名詞によって該当率にばらつきがあった．これは，名詞によって，特徴点の数と大きさが違うためだと考えられ，車や飛行機など，特徴点の多い名詞は該当率が高かった．

5.2 ユーザビリティ評価

ユーザビリティ評価では，情報システム工学科の大学生 10 名（A〜J）に，文字列 CAPTCHA と画像 CAPTCHA（Asirra），提案手法の IC-CAPTCHA（3 つのパターン）を，各手法 10 回ずつ解いてもらい，その後，アンケート調査を実施した．アンケート項目を表 1 に示す．

表 1 アンケート項目
Table 1 Questionnaire item

質問事項	印象語と得点
解いていて楽しかったか？	楽しくない 1 点← →5 点 楽しい
解くことは面倒だったか？	面倒だ 1 点← →5 点 面倒ではない
解くことは簡単だったか？	難しい 1 点← →5 点 簡単だ
CAPTCHA が使いやすかったか？	使いにくい 1 点← →5 点 使いやすい
Web サービス上で使いたいか？	使いたくない 1 点← →5 点 使いたい

また，CAPTCHA を解いてもらう際に，CAPTCHA の解答までに要する時間とその正否を調査した．

結果は表 2，3 のようになった．

表 2 アンケート結果
Table 2 questionnaire results

質問事項	IC-CAPTCHA（提案手法）	文字列 CAPTCHA	Asirra（画像 CAPTCHA）
解いていて楽しかったか？	4	1.8	3.9
解くことは面倒だったか？	4.9	1.5	3.3
解くことは簡単だったか？	4.9	2.5	4.4
CAPTCHA が使いやすかったか？	4.8	2.1	3.7
Web サービス上で使いたいか？	4.4	2.4	3.4

表 2 より，5 つの質問事項全てで，IC-CAPTCHA > 画像 CAPTCHA > 文字列 CAPTCHA となった．また，表 3 より，IC-CAPTCHA は 2 つの既存手法より正答率が高く，平均所要時間が短いことがわかる．したがって，IC-CAPTCHA は，実際に使用する際のユーザビリティにおいて非常に優れているといえる．IC-CAPTCHA において，妨害図形数が多いパラメータの方が平均所要時間が短い理由としては，IC-CAPTCHA を解いてもらう順番を妨害図形数 7，9，11，としたため，被験者が徐々に IC-CAPTCHA の解答手順に慣れたためだと考えられる．

表 3 所要時間と正答率
Table 3 The required time and correct answer rate

	正答率（%）	平均所要時間（*sec*）
IC-CAPTCHA 提案手法:妨害図形数 7	99%	4.7709
IC-CAPTCHA 提案手法:妨害図形数 9	100%	4.2343
IC-CAPTCHA 提案手法:妨害図形数 11	100%	3.7683
文字列 CAPTCHA	75%	15.195
Asirra	95%	14.04

6. まとめ

本論文では，既存の CAPTCHA 手法の問題点を改善する新たな CAPTCHA 手法である IC-CAPTCHA システムの提案を行った．また，データベース照合とユーザビリティアンケートの 2 つの面から評価と考察を行い，IC-CAPTCHA システムの有効性について示した．

今後は，画像を加工する際の妨害処理の精度を向上させるとともに，ハッシュ辞典検索にブルームフィルターを用いて，検索時間の短縮等を行いたい．

参考文献

[1] L. von Ahn, M. Blum, N. Hopper, and J. Langford, “CAPTCHA: Telling humans and computers apart,” Advances in Cryptology, Eurocrypt’03, vol.2656 of Lect. Notes Comput. Sci.,pp.294-311, 2003.

[2] J. Yan and A.S.E. Ahmad, “Breaking visual CAPTCHAs with native pattern recognition algotithms,” 2007 Computer Security Applications Conference, pp.279-291, 2007.

[3] K. Chellapilla and P.Y. Simard, “Using machine learning to break visual human interaction proofs（HIPs）,” Advances in Neural Infomation Processing Systems, vol.17, pp.265-272, 2005.

[4] P.Golle, “Machine learning attacks against the asirra CAPTCHA,” Proc. 15th ACM conference on Computer and Communications Security, pp.535-542, 2008.

[5] L. von Ahn, B. Maurer, C.McMillen, D. Abraham, and M. Blum, “reCAPTCHA:Humen-based character recognition via Web security measures,” Sience, vol.321, no.5895, pp.1465-1468, 2008.

[6] Greg Mori, Jitendra Malik, “Recognizing Objects in Adversarial Clutter:Breaking a Visual CAPTCHA,” cvpr,pp.134, 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition （CVPR ’03） - Volume 1, 2003.

[7] Gabriel Moy, Nathan Jones, Curt Harkless, and Randall Potter, “Distorition Estimation Techniques in Solv-

© 2012 Information Processing Society of Japan

ing Visual CAPTCHAs," proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), 2004.

[8] Jeremy Elson, John Douceur, Jon Howell and Jared Saul,"Asirra: a CAPTCHA that exploits interest-aligned manual image categorization," Proceedings of the 14th ACM conference on Computer and Communications Security, pp. 366-374, October 2007.

© 2012 Information Processing Society of Japan