メンタルローテーションを利用した画像CAPTCHAの提案
An image-based CAPTCHA using mental rotation

池谷勇樹† 可児潤也† 米山裕太† 西垣正勝†
Yuki IKEYA† Junya KANI† YutaYONEYAMA† Masakatsu NISHIGAKI†

† 静岡大学大学院 情報学研究科
† Graduate school of Informatics, Shizuoka University

要旨

メールアカウントの不正取得やブログへのスパムコメントの書き込みといった Web サービス提供サイトに対する自動プログラム（マルウェア）による DoS（Denial of Service，サービス不能）攻撃が定常的に行われている．そのため，マルウェアによる Web サービスの不正利用と，人間による正規利用を識別するために CAPTCHA が広く利用されている．しかし，近年のマルウェアの高度化に伴い，広く使われている文字判読型の CAPTCHA や，動物画像の判別を用いた CAPTCHA が突破されてしまっており，人間のより高度な認知能力に基づいた CAPTCHA が必要とされている．既にそのような CAPTCHA がいくつか提案されているが，出題画像の自動生成やマルウェアに対する解読耐性に関する課題が残っていた．本稿では，人間の高度な認知能力の 1 つであるメンタルローテーションに着目した画像 CAPTCHA を提案する．メンタルローテーションとは，1 つの視点から写された 2 次元物体や 3 次元物体を頭の中で回転させ，異なる視点から写された形姿を認識する能力のことである．提案方式のプロトタイプを実装し，被験者 20 名に対して基礎実験を行い，提案方式の正答率，所要時間，利便性について評価した．また，提案方式と既存の CAPTCHA を比較し，考察を行った．

キーワード

CAPTCHA，画像認識，メンタルローテーション，WEB セキュリティ

1. はじめに

メールアカウントの不正取得やブログへのスパムコメントの書き込みといった Web サービス提供サイトに対する自動プログラム（マルウェア）による DoS（Denial of Service，サービス不能）攻撃が定常的に行われている．

CAPTCHA は，このようなマルウェアによる Web サービスの不正利用と，人間による正規のサービス利用とを識別するために必須となる技術である．人間には容易に解答できるがコンピュータには判別が困難である問題をユーザに出題し，正解できたユーザを人間だと判定する．

現在広く使われている文字判読型の CAPTCHA（図 1）[1]や動物画像の判別を用いた Asirra（図 2）[3]など，画像の利用が典型的な手法となっている．近年は，OCR 機能によって文字判読型 CAPTCHA を破るマルウェア[2]や機械学習によって Asirra を破る自動プログラム[4]が報告されており，人間のより高度な認知能力を利用した画像 CAPTCHA が研究されている[5][7]．しかし，これらの既存技術には，出題画像の自動生成や出題画像の解読耐性に関する課題が残っている．

これらの課題の克服のために，本稿では，人間の高度な認知能力の 1 つであるメンタルローテーションに着目し，これを利用した画像 CAPTCHA を提案する．提案方式は，出題画像の自動生成が可能であり，解読耐性の向上についても期待される．

以下，2 章で関連研究として既存の画像 CAPTCHA を紹介する．3 章で提案方式について説明し，4 章にて基礎実験の結果を報告する．5 章で提案方式についての考察を行い，6 章で本稿をまとめ，今後の課題を述べる．

図 1 文字判読型 CAPTCHA の認証画面例

2. 関連研究

2.1. Asirra

OCR を用いて文字判読型 CAPTCHA を解読するマルウェアに対抗するために提案された画像 CAPTCHA として Asirra がある[3]．Asirra

では, 合計 12 枚の犬と猫の出題画像を表示し，それらの画像の中から猫の画像だけを全て選択できたユーザを人間として判定する（図 2）．画像の意味を理解することは人間の高度な認知メカニズムの 1 つであり，マルウェアによる不正解答は不可能であると考えられていた．

しかし，猫の画像の特徴や犬の画像の特徴を抽出し，機械学習技術を利用することによって，Asirra は破られ得るという研究報告がなされた[4]．Asirra が破られた原因として，Asirra が画像の表面的な意味を問うものであったためと考えられる．人間のより高度な認知処理に基づく CAPTCHA が求められている．

図 2　Asirra の認証画面例

2.2.　3D 画像 CAPTCHA

YUNiTi.com [5]は，人間が有する「3 次元物体の認識能力」を利用した 3D 画像 CAPTCHA を実装し，運用している． 出題画像には 3 次元オブジェクトが 3 個表示されており，それぞれのオブジェクトが何であるかを 18 個の候補画像の中から正しく選択できたユーザを人間として判定する（図 3）．候補画像は，18 個のそれぞれのオブジェクトの正立画像である．出題画像は，候補画像のオブジェクトのいずれかを別の視点から写した画像となっており，その視点は出題の度に異なる．

しかし，この方式は，出題画像とともに候補画像の一覧が全て表示されている上に，候補画像が 18 種類しかない．過去の出題画像から，様々な視点から描画されたオブジェクトの画像を，オブジェクトの種類ごとに数百枚以上収集しておけば，テンプレートマッチングによってマルウェアが正解オブジェクトを特定でき得るという問題が指摘されている[6].

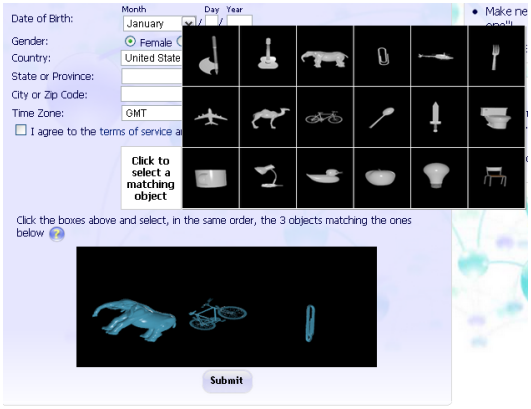

図 3　3D 画像 CAPTCHA の認証画面例

2.3.　4 コマ漫画 CAPTCHA

4 コマ漫画 CAPTCHA [7][8]は，人間の「ユーモアを解する能力」を利用した CAPTCHA である（図 4）．4 コマ漫画の各コマの順番を入れ替えて 4 枚の出題画像として表示し，正しい順番に並べることができたユーザを人間として判定する．人間は起承転結の崩れを認識し，ユーモアを理解して 4 コマ漫画を正しい起承転結の順番に再構築することができる．また，4 コマ漫画を利用しているため，ユーザの利便性の向上も期待される（エンターテイメント性を有するため，ユーザが楽しんで問題を解くことができる）方式となっている．しかし，起承転結を備えた 4 コマ漫画の自動生成が難しいという問題がある．

図 4　4 コマ漫画 CAPTCHA の認証画面例
（出展：左から 1 番目の図：文献[9]の p.25 の 4 コマ漫画の 1 コマ目，2 番目の図：同，p.25 の 4 コマ目，3 番目の図：同，p.25 の 3 コマ目，4 番目の図：同，p.25 の 2 コマ目）

3. 提案方式

3.1. コンセプト

人間は，空間認識能力に長けている．このため，3 次元オブジェクトが写っている 2 次元画像から，そのオブジェクトの 3 次元形状を理解することができる．この「2 次元画像から 3 次元物体を認識する能力」は，人間が有する高度な認知メカニズムの 1 つであると考えられる[10]．また，人間であれば，ある 1 つの視点から写された 2 次元オブジェクトや 3 次元オブジェクトを頭の中で回転させ，異なる視点から写された形姿を認識することができる．このような，人間の「物体を頭の中で回転させ，比較する能力」は心的回転（メンタルローテーション）と呼ばれている[11][12]. 上記の 2 つの認知能力を使用することによって，人間は，ある 3 次元オブジェクトを異なる視点から写した 2 枚の 2 次元画像を見た際に，そこに写されている 3 次元オブジェクトの形状を推測し，一方の 2 次元画像からもう一方の 2 次元画像にどう視点を変えたのか理解することができる．

本稿では，メンタルローテーションを利用した画像 CAPTCHA（以下，Doko-Soko CAPTCHA と呼ぶ）を提案する．Doko-Soko CAPTCHA では，1 つの 3 次元オブジェクトを 2 つの異なる視点から写した 2 枚の 2 次元画像を自動生成し，一方の 2 次元画像を出題画像として，もう一方の 2 次元画像を回答画像として使用する．出題画像には 3 次元オブジェクトの任意の 1 部位にマーカが付加されている．回答画像にはマーカがない．Doko-Soko CAPTCHA の出題画像および回答画像の例を図 5, 図 6 に示す. ユーザは，出題画像のマーカ（赤い丸）の部位が回答画像ではどこに当たるのかを回答する．人間であれば，出題画像の 3 次元オブジェクトを頭の中で回転させ，回答画像の 3 次元オブジェクトと比較することによって，回答画像における正解部位（出題画像のマーカ部位に対応する部位）を認識可能である．

一方，マルウェアは，立体認識の技術を利用することができる．立体認識の分野では，1 つの 3 次元オブジェクトを異なる複数の視点から撮影した画像からそのオブジェクトの 3 次元形状情報を再構築する技術が研究されている[12]. 現在の技術においては，1 つの 3 次元オブジェクトを異なる 2 つの視点から撮影した 2 枚の画像から，その 3 次元オブジェクトの立体形状をほぼ正確に同定することが可能となっている．しかし，2 枚の画像に対してある程度以上の歪が加わっている場合には，3 次元オブジェクトの立体形状の同定は格段に難しくなる．

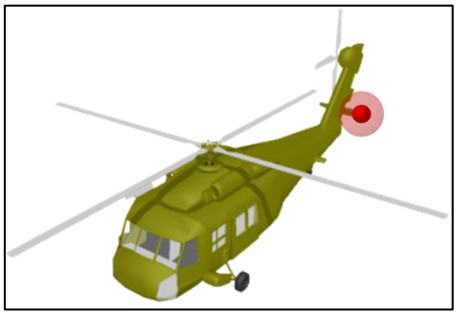

図 5(a)　方式 α の出題画像の例

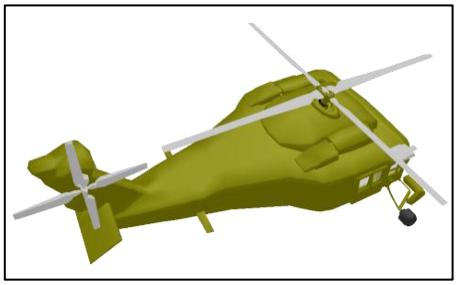

図 5(b)　方式 α の回答画像の例

3.2. 手順

Doko-Soko CAPTCHA の手順を以下に示す．なお，システムには大量の 3 次元オブジェクトのモデルが登録されていることを前提とする．

① システムは，出題画像に利用する 3 次元オブジェクト（以下，出題用オブジェクト）のモデルをランダムに選ぶ.
② システムは，①で選んだ出題用オブジェクトに対して，マーカ部位をランダムに選ぶ.
③ システムは，出題画像の視点をランダムに選ぶ.
④ システムは，出題画像を生成する（出題画像にはマーカも描画されている）.
⑤ システムは，①で選んだ出題用オブジェクトをランダムに加工することによって，回答画像に利用する 3 次元オブジェクト（以下，回答用オブジェクト）を生成する．マーカの部位も加工に応じた位置に移動する.
⑥ システムは，回答画像の視点をランダムに選ぶ.
⑦ システムは，回答画像を生成する（回答画像にはマーカは描画されていない）.
⑧ システムは，出題画像と回答画像を表示する.
⑨ ユーザは，回答画像において「出題画像内のマーカが付加された部位（②で選ばれた部位であり，⑤の加工に応じた位置に移動している）に対応する部位」を回答する.
⑩ システムは，正答できたユーザを人間，

正答できなかったユーザをマルウェアとして判別する．

立体認識の技術を利用した解読に対する対策のために，手順⑤にて出題用オブジェクトを「加工」していることに注意されたい．ここでは，加工の例として，出題用オブジェクトにアフィン変換を施して変形することによって回答用オブジェクトを生成する方式（方式α）と，出題用ブジェクトを「出題用ブジェクトに類似したオブジェクト」に置換することによって回答用オブジェクトを生成する方式（方式β）を示す．

方式αの画像例が図5である．図5の例では，回答画像（図5(b)）の生成の際に，出題用オブジェクト（ヘリコプタ）の頂点データの x 軸，y 軸，z 軸がそれぞれに独立に任意の倍率で拡大縮小され，ここでは「丸く太ったヘリコプタ」に変形させられている．なお，回答用オブジェクトだけでなく，出題用オブジェクトに対しても加工を施すことも可能である．

方式βの画像例が図6である．図6の例では，出題用オブジェクトが猫であるのに対し（図6左），回答用オブジェクトが馬に置換されている（図6右）．なお，ここで，登録されている3次元オブジェクトモデルのそれぞれに対して，オブジェクト間の部位どうしの対応関係を記したデータベースがシステム内に管理されている必要がある．

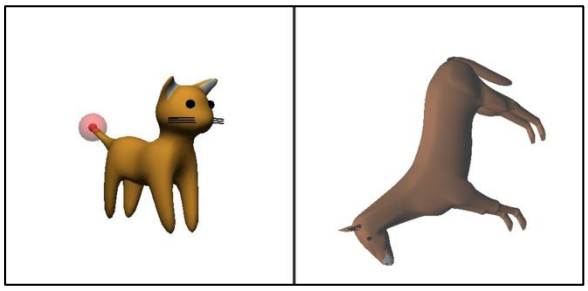

図6　方式βの画像例
（左：出題画像，右：回答画像）

Doko-Soko CAPTCHA においては，回答画像にはマーカが描画されておらず，出題画像と回答画像の情報だけを用いて，マルウェアが回答画像におけるマーカ部位を同定することは困難であると考えられる．これに対し，システムは回答画像におけるマーカの位置を知っている．これが「落とし戸」となり，システム（機械）が「マルウェア（機械）には認識できない問題」を自動生成し，かつ，システム（機械）自身がユーザの解答に対する正解判定を行うことが可能となっている．使用するオブジェクト，マーカの位置，視点の位置が認証の度にランダムに選ばれるため，システムが大量の3次元オブジェクトのモデルを有していれば，ほぼ無数の問題を自動生成することができる．よって，テンプレートマッチングを利用した解読に対する耐性も高いのではないかと期待される．

3.3.　実装

Doko-Soko CAPTCHA のプロトタイプの実装を行った．今回は方式αを実装した．Doko-Soko CAPTCHA の認証画面例を図7に示す．図7左が出題画像であり，図 7 右が回答画像である．出題画像中に描画されている赤い球がマーカである．ユーザは，出題画像中のマーカ部位が回答画像上のどの位置にあるかを同定し，マウスクリックによって回答する．ユーザがクリックした箇所（ディスプレイ上の座標）と正解部位の位置（ディスプレイ上の座標）の距離が閾値以下であれば認証成功とした．図7の例では，出題画像においてマーカが猫の右耳を示していることが分かるため，回答画像における猫の右耳をクリックすれば正解となる．

実装において留意した点を以下に述べる．

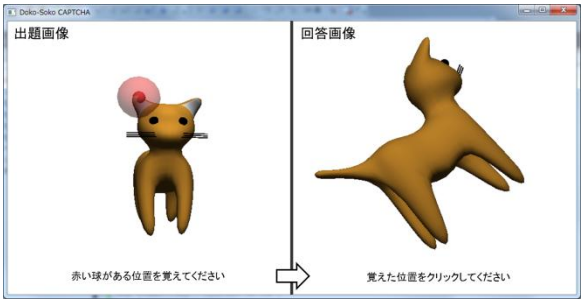

図7　Doko-Soko CAPTCHA の認証画面例
（左：出題画像，右：回答画像）

3.3.1.　マーカ

今回実装したプロトタイプでは，出題用オブジェクトの3次元モデルにおける頂点情報の中からランダムに1点を選択し，その点をマーカの中心の座標とした．なお，オブジェクトによっては，頂点が密集している部位がある．そのような場合は，密集部位の頂点がマーカとして選ばれる可能性が高くなってしまい，マルウェアにマーカ位置を推測する手掛かりを与えてしまう．このため，マーカ位置をランダムに選択するにあたっては，頂点座標の分布の偏りを考慮する必要がある．（今回の実験では，頂点座標の分布にそれほど偏りのない3次元オブジェクトを使用することで対処した．）

回答オブジェクトにおける「出題用オブジェクトのマーカとして選択された頂点」に対応する頂点の座標が正解座標となる．ここで，頂点座標は3次元データであるのに対し，ユーザに

よるマウスクリックは（ディスプレイ上の座標情報として得られるため）2 次元データである．このため，Doko-Soko CAPTCHA では，3 次元オブジェクト上の正解部位がディスプレイ上ではどの座標にあたるかを計算した上で，その 2 次元正解座標とクリックされた座標との距離によって正解判定を行っている．正解範囲は，正解座標を中心とした円の内部であり，今回の実装では 30 ピクセルを円の半径とした．

出題画像内に描画するマーカは，大きな半透明の球の中に小さな不透明の球が入っている．マーカを一つの球で表した場合，マーカや視点の位置によってはマーカ（の一部）が隠れてしまい，マーカの中心を認識することが難しくなるケースがある．マーカを二重の球で表すことにより，内側の小さな球が隠れてしまっても，「内側の小さな球が見えない」という事実を補助情報として用いることでマーカの中心の位置の認識が容易となる．

3.3.2. 視点

3 次元オブジェクトの後ろ側にマーカが隠れてしまうと，人間であっても解答が困難になる．このため，出題画像においては，マーカを正面近くから見ることができる位置の中から視点が選ばれるような制約を追加している．

一方で，回答画像の視点に対しては，出題画像の視点から Y 軸（垂直軸）を中心として 45 度以上離れた位置の中からランダムに選ばれるような制約を追加した．回答画像において「出題画像の視点」に近い視点が選ばれた場合，生成される出題画像と回答画像がほぼ同じになる可能性が高まり，両方画像を比較することでマルウェアが正解箇所を解読する危険性が生じるためである．

4. 基礎実験

4.1. 目的

ユーザ（人間）は Doko-Soko CAPTCHA に正答することが可能であることを確認する．また，実験後に被験者に対してアンケート調査を行い，Doko-Soko CAPTCHA の利便性について調査する．

4.2. 実験方法

本実験の被験者は情報系の大学生 20 名である．各被験者に，5 問の Doko-Soko CAPTCHA を連続して解いてもらった．すべての被験者にとって Doko-Soko CAPTCHA は初見であるため，5 問の内の始めの 2 回は練習という位置付けで，3～5 問目の 3 問を実験の本番として扱うことにした．

今回の実験に使用した 3 次元オブジェクトは 5 種類（A～E）である．練習では，必ずオブジェクト A，B の順番で出題用オブジェクトが選ばれる．本番では，オブジェクト C～E の中からランダムな順番で 1 回ずついずれかのオブジェクトが選ばれる．マーカ位置および視点については，3 章で説明した制約の下，毎回ランダムに選ばれる．

今回は，マーカの選ばれ方や，出題画像と回答画像の視点の選ばれ方については，被験者には知らせていない．被験者には，マーカは（外側の大きな半透明の球ではなく）内側の小さな不透明の球の中心であるので，それを意識して回答を行うよう指示した．

今回の実験では，各問題に対して，正解したかどうか，解答にかかった所要時間，クリックされた位置を測定した．また，実験終了後に，被験者にアンケートに回答してもらった．アンケートの質問項目を以下に示す．質問①，③，⑤は 1～5 点の点数付けで回答してもらった．

① 簡単に解けたか（簡単なら 5）
② 質問①で 2 や 1 を選択した場合，その理由は何か
③ 面倒だと感じたか（面倒ではないなら 5）
④ 質問③で 2 や 1 を選択した場合，その理由は何か
⑤ 面白いと感じたか（面白いなら 5）
⑥ 質問⑤で 5 や 4 を選択した場合，その理由は何か
⑦ 何問までなら続けて解いても良いと思うか．また，その理由は何か
⑧ 実際の Web サービスの利用の場面で CAPTCHA を解くことが要求された場合，文字判読型 CAPTCHA と Doko-Soko CAPTCHA のいずれかを選ぶことができたらどちらを選ぶか．また，その理由は何か

4.3. 実験結果

4.3.1. 正答率と所要時間

ユーザごとに正答率と平均所要時間をまとめた結果を表 1 に示す．今回の実験では使用する 3 次元オブジェクトの順番をランダムで決定したため，表 2 に実験順でまとめた結果を，表 3 にオブジェクト別にまとめた結果を示した．

表 1 より，Doko-Soko CAPTCHA の正答率は全ユーザの平均で 77.3%（全ユーザで 20 名×3 回＝60 回の試行を行った内，成功が 44 回，失敗が 16 回）である．実用に供するには正答率が低いため，今後改良を行なっていく必要がある．表 2 より，1 回目の正答率が最も低くなっ

ているため，ユーザの慣れによって正答率が上がることが予想される．表 3 より，CAPTCHA に使用する 3 次元オブジェクトによって正答率が大きく変わることが分かる．

本実験でユーザが解答に失敗した主な理由を 3 つ挙げる．1 つ目は，オブジェクトの左右を混同による間違いであり，全 16 回の失敗の内，この間違いによるものは 5 回であった．2 つ目は，クリック位置のわずかなずれであり，全 16 回の失敗の内，この間違いによるものは 4 回であった．もし，正解範囲の半径を 30 ピクセルから 35 ピクセルに広げたとすると，正答率は 80.0%にまで上がる．適切な正解範囲については今後検討していく．3 つ目は，画像の奥行きの分かりにくさによる間違いである．特に，3 次元オブジェクトを真正面，真後ろ，真上，真横，真下から見る視点が選ばれてしまった場合，画像の奥行きが分かりにくくなり，間違いやすくなることが分かった．奥行きの誤認識は，左右の間違いやクリック位置のずれにも関連するため，視点の選択についての検討が必要である．

表 1 より，Doko-Soko CAPTCHA の 1 問あたりの平均所要時間は 5.4 秒である．最短所要時間は 3.2 秒，最長所要時間は 8.9 秒である．一般的な文字判読型 CAPTCHA の所要時間が 10 秒程度であるため[8]，Doko-Soko CAPTCHA は短い時間で解くことができると言えるだろう．表 2 より，実験順による所要時間の大きな変化は見られなかった．表 3 より，3 次元オブジェクトによって所要時間の差が見られた．短い時間で解くことができる 3 次元オブジェクトの特徴についての検討を進めて行きたい．

表 2　実験結果（実験順）

	平均正答率	平均所要時間[秒]
1 回目	60.0% (12/20)	5.9
2 回目	80.0% (16/20)	5.5
3 回目	80.0% (16/20)	5.2

表 3　実験結果（オブジェクト別）

	平均正答率	平均所要時間[秒]
オブジェクト C	80.0% (16/20)	6.1
オブジェクト D	55.0% (11/20)	4.6
オブジェクト E	85.0% (17/20)	5.4

表 1　実験結果（ユーザ別）

	正答率	平均所要時間[秒]
ユーザ 01	2/3	6.7
ユーザ 02	2/3	5.3
ユーザ 03	3/3	3.9
ユーザ 04	3/3	8.7
ユーザ 05	3/3	3.9
ユーザ 06	2/3	8.9
ユーザ 07	3/3	6.5
ユーザ 08	2/3	4.6
ユーザ 09	1/3	4.3
ユーザ 10	3/3	5.4
ユーザ 11	2/3	7.1
ユーザ 12	2/3	3.9
ユーザ 13	2/3	3.2
ユーザ 14	3/3	3.7
ユーザ 15	2/3	4.8
ユーザ 16	3/3	6.4
ユーザ 17	2/3	3.4
ユーザ 18	1/3	5.9
ユーザ 19	1/3	6.1
ユーザ 20	2/3	4.5
平均	73.3% (44/60)	5.4

表 4　アンケート結果

	①簡単さ	③面倒のなさ	⑤面白さ	⑦回数	⑧どちらを選ぶか
ユーザ 01	2	4	4	3	文字判読
ユーザ 02	2	4	4	3	文字判読
ユーザ 03	4	4	4	2	Doko-Soko
ユーザ 04	4	2	4	1	Doko-Soko
ユーザ 05	4	5	4	3	Doko-Soko
ユーザ 06	5	5	3	2	Doko-Soko
ユーザ 07	4	4	5	4	Doko-Soko
ユーザ 08	4	5	3	2	Doko-Soko
ユーザ 09	2	5	5	2	Doko-Soko
ユーザ 10	2	1	4	2	Doko-Soko
ユーザ 11	4	5	5	3	Doko-Soko
ユーザ 12	3	5	4	3	Doko-Soko
ユーザ 13	3	5	4	3	文字判読
ユーザ 14	4	3	4	3	文字判読
ユーザ 15	2	5	5	3	Doko-Soko
ユーザ 16	4	5	5	3	Doko-Soko
ユーザ 17	4	4	5	3	Doko-Soko
ユーザ 18	3	5	5	3	Doko-Soko
ユーザ 19	2	4	4	3	文字判読
ユーザ 20	3	5	5	3	Doko-Soko
平均	3.3	4.3	4.3	2.7	

4.3.2. 利便性

アンケートの結果を表 4 に示す.

質問①「簡単に解けたか(簡単であれば 5 点)」については，4 と回答したユーザが最も多く，平均値は 3.3 点となった．難しい（2 や 1）と回答したユーザには，質問②でその理由を書いてもらった．理由としては，「立体的に見ることが難しい」，「左右の把握が難しい」，「(回答画像において）クリックする箇所が見えなくなると難しい」が挙げられていた.

質問③「面倒だと感じたか（面倒でなければ 5 点)」については，5 と回答したユーザが最も多く，平均値は 4.3 点となった．面倒（2 や 1）と回答したユーザには，質問④でその理由を書いてもらった．理由としては，「立体物の構造を考えなければならないので面倒」，「毎回必ず解くことになるなら面倒」が挙げられていた.

質問⑤「面白いと感じたか(面白ければ 5 点)」については，4 と回答したユーザが最も多く，平均値は 4.3 点となった．面白い（5 や 4）と回答したユーザには，質問⑥でその理由を書いてもらった．主な理由としては，「画像を使っていて面白い」，「ゲーム感覚でできて面白い」，「空間認識能力を試されるので面白い」が挙げられていた.

質問⑦「何問までなら続けて解いても良いか」については，3 回という回答したユーザが最も多く，13 人であった．次に多かったのは 2 回と回答したユーザであり，5 人であった．4 回および 1 回と回答したユーザがそれぞれ 1 人いた．多くのユーザが 4 問以上連続では解きたくないと感じている．その主な理由としては，「多すぎると失敗してしまうから」，「時間が掛かり過ぎると面倒だから」が挙げられていた．また，「Web サービスの重要性よって連続で解いても良いと思える問題数は変わる」という意見や，「歳をとったら 1 問でも苦痛だと思う」という意見もあった.

質問⑧「文字 CAPTCHA と Doko-Soko CAPTCHA のどちらを選ぶか」については，文字判別型 CAPTCHA を選んだユーザが 5 人，Doko-Soko CAPTCHA を選んだユーザが 15 人であった．文字判別型 CAPTCHA に不便を感じていたユーザは，Doko-Soko CAPTCHA を選んだようであった．文字判別型 CAPTCHA を選んだ主な理由としては，「文字判別型 CAPTCHA の方が分かりやすいから」，「キーボードのみで操作を行えるから」，「Doko-Soko CAOTCHA は正確に答えようとすると時間が掛かるから」が挙げられていた．Doko-Soko CAPTCHA を選んだ主な理由としては，「文字判別型 CAPTCHA は難しいから」，「マウスのみで操作を行えるから」，「画像の方が楽しいから」が挙げられていた．難しさや入力デバイスについては真逆の理由が挙げられていたため，ユーザによって適している CAPTCHA が異なることが分かった.

5. 考察

2 章に示したように，人間の高度な認知処理を利用した画像 CAPTCHA においては，出題画像の自動生成および解読耐性に課題が残っていた．本章では，Doko-Soko CAPTCHA がこれらの問題を改善していることを示す.

5.1. 出題画像の自動生成

4 コマ漫画 CAPTCHA [7]は人間の高度な認知処理に基づく画像 CAPTCHA であるが，4 コマ漫画（出題画像）の自動生成が難しいという問題があった．Doko-Soko CAPTCHA は，3 次元コンピュータグラフィックスを利用して毎回新しい出題画像を生成することが可能であり，出題画像の自動生成を達成している．多数の 3 次元オブジェクトのモデルをシステムに登録しておき，使用するオブジェクト，オブジェクトの大きさ，マーカを付加する位置，視点の位置といったパラメータを変更することによって，出題画像と回答画像のペアを無数に生成することができる.

5.2. 解読耐性

YUNiTi.com [5]の 3D 画像 CAPTCHA は，人間が有する「3 次元物体の認識能力」を利用している点，出題画像の自動生成を達成している点で秀逸な画像 CAPTCHA である．しかし，「複数の候補画像の中から一番近い画像を選ぶ」という形の質問形態となっているため，テンプレートマッチングに対する脆弱性を残していた.

YUNiTi.com の 3D 画像 CAPTCHA においては，解答の候補となる 3 次元オブジェクトの数が 18 種類しかないため，この問題が非常に顕著となる．解答の候補となる 3 次元オブジェクトの数が少ない場合，過去の出題画像をアーカイブすることによって，オブジェクトごとに「数百の異なる視点から描画されたオブジェクトの画像」を収集することは，非現実的ではない．3D 画像 CAPTCHA の出題画像はいずれかのオブジェクトを任意の視点から描画した画像であるため，前もってあらゆる視点からの画像をアーカイブしておけば，アーカイブ画像群の中に出題画像に近い画像が必ず存在することにため，パターンマッチングによって正解オブジェクトを特定でき得る[6].

Doko-Soko CAPTCHA は，「複数の候補画像の中から一番近い画像を選ぶ」という質問形態をとっておらず，テンプレートマッチングに対す

る耐性も向上していると期待される．5.1 節で示したように，出題画像と回答画像のペアを無数に生成することによって，過去の画像をアーカイブすることも困難になっている．

また，3.1 節で説明したように，立体認識技術を利用した攻撃については，出題画像と回答画像とで 3 次元オブジェクトの変形（方式α）や置換（方式β）するという対処を行っている．

しかし，マルウェアによる攻撃手法は多様であり，Doko-Soko CAPTCHA の解読耐性が理論的に証明されているわけではない．特に機械学習を利用した解読などについては，今後の早急に分析を行っていきたい．

6. まとめと今後の課題

本稿では，人間の高度な認知処理の 1 つであるメンタルローテーションを利用した画像 CAPTCHA である「Doko-Soko CAPTCHA」を提案し，プロトタイプの実装と基礎実験による評価を行った．基礎実験では被験者 20 人に Doko-Soko CAPTCHA を解いてもらい，平均正答率 77.3%，平均所要時間 5.4 秒という結果が得られた．1 問あたりの解答所要時間は短いものの，正答率については今後改良を行なっていく必要がある．アンケートによる利便性の調査については良好な結果が得られた．

今後の課題としては，Doko-Soko CAPTCHA に適した 3 次元オブジェクトの検討（メンタルローテーションを行いやすいオブジェクトを利用することによって，正規ユーザの正答率が向上する），Doko-Soko CAPTCHA の視点の選択範囲の検討（マーカ位置の認識が容易となるような視点を選択することによって，正規ユーザの正答率が向上する），正解判定の範囲の検討（範囲が大きいほど正規ユーザの正答率は向上するが，ブルートフォース攻撃に対して脆弱となる），および，方式βの実装と評価，等が挙げられる．今回の基礎実験やアンケート調査の結果を参考にしながら，これらの検討を進めていきたい．

参考文献

[1] Unlocking Google's Gmail CAPTCHA <http://www.gmailhelp.com/2009/10/unlocking-googles-gmail-captcha/>, 2013 年 3 月 12 日アクセス

[2] J.Yan, A.S.E.Ahmad, 'Breaking Visual CAPTCHAs with Naïve Pattern Recognition Algorithms', 2007 Computer Security Applications Conference, 2007, pp.279-291

[3] ASIRRA - Microsoft Research, <http://research.microsoft.com/en-us/um/redmond/projects/asirra/>, 2013 年 3 月 12 日アクセス

[4] P.Golle, 'Machine Learning Attacks Against the ASIRRA CAPTCHA', 2008 ACM CSS, 2008, pp.535-542

[5] YUNiTi.com - Social Networking At Its Best, <http://www.yuniti.com/>, 2013 年 3 月 12 日アクセス

[6] TechnoBabble Pro: How they'll break the 3D CAPTCHA , <http://technobabblepro.blogspot.jp/2009/04/how-theyll-break-3d-captcha.html>, 2013 年 3 月 12 日アクセス

[7] 鈴木徳一郎，山本匠，西垣正勝，「４コマ漫画 CAPTCHA の提案」，2009 年暗号と情報セキュリティシンポジウム予稿集，2009，3D3-3（CD-ROM）

[8] 上原章敬，鈴木徳一郎，山本匠，西垣正勝，「４コマ漫画 CAPTCHA の検討」，第 52 回コンピュータセキュリティ合同研究発表会予稿集，2011，2-B(13)（CD-ROM）

[9] 植田まさし，「新コボちゃん 8」，芳文社，2006

[10] 吉村浩一，「知覚は問題解決過程―アーヴィン・ロックの認知心理学」，ナカニシヤ出版，2001

[11] Shepard, R and Cooper, L, 'Mental images and their transformations.', MIT Press, Cambridge, MA, 1982

[12] Shepard, R and Metzler. J, 'Mental rotation of three dimensional objects.', Science, New Series, Vol.171, No.3972, 1971, pp.701-703

[13] 安居院猛，長尾智晴，「画像の処理と認識」，昭晃堂，1992