物体のサイズ感を利用した 3DCG 画像 CAPTCHA の評価

Evaluation of 3D CG CAPTCHA Using Object Size Sense

○西原 大貴[1)]，新井 イスマイル[2)]，梶村 好宏[1)]

Daiki Nishihara[1)], Ismail Arai[2)], Yoshihiro Kajimura[1)]

[1)]明石工業高等専門学校電気情報工学科，[2)]奈良先端科学技術大学院大学 総合情報基盤センター

1. はじめに

Web サービスに対する，自動プログラムを用いた機械攻撃を防ぐ技術の一つとして，CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart)と呼ばれる人間か機械かを識別するテストが利用されている．それらのうち文字判別型 CAPTCHA は現在広く利用されているが，近年では OCR(Optical Character Recognition)技術の発展などにより，機械攻撃によって破られる可能性が高まってきた．すなわち，CAPTCHA は，人間にとって解読しやすいこと(利便性)の他に，機械攻撃耐性(安全性)が確保されている必要がある．一方で，CAPTHCA には，出題が自動生成可能である(自動生成性)という要求も存在する[1]．これを満たさない場合，出題の総数は有限となり，データベースを参照する機械攻撃が可能となる．従って，CAPTCHA にはこれらの 3 要件が要求される．

これを満たす既存研究として藤田らは，常識的な形状をした異なる 2 つの 3 次元オブジェクトをマージしてめり込ませることで生成した非現実オブジェクトをユーザに選択させる 3DCG 画像 CAPTCHA 手法(以下，非現実画像 CAPTCHA)を提案した[1]．しかしながら，これは，輪郭抽出技術を応用した機械攻撃により破られる可能性が考えられる．

これに対し，我々は「サイズ感」に着目することで，輪郭抽出技術などにより各オブジェクトの形状や名称が限定されたとしても容易には解読されないと期待できる手法を提案する．

2. 関連研究

非現実画像 CAPTCHA は，「常識からの逸脱を認識する能力」が人間特有の高度な認知能力であることに着目し，複数の 3D オブジェクトの中に配置された 1 体の非現実オブジェクト(異なる 2 つの 3D オブジェクトをマージし，めり込ませたもの)をクリックさせる．これは，3DCG を用いることで無数の出題を自動生成でき，常識を持つ人間は容易に回答できる一方で，機械は人間の常識を備えることが困難なため通常と非現実のオブジェクトを見分け難く，CAPTCHA の 3 要件を満たすとした．

特に，安全性について，論文中で機械学習を用いてめり込み部分の境界を検出する手法や総当たり攻撃にも耐えうるとされた．しかしながら，輪郭抽出技術の応用など，その他の攻撃手法により，めり込みを検出できる可能性が考えられる．

図 1 提案手法のイメージ

3. 物体のサイズ感を利用した手法の提案

提案手法では図 1 に示すように，「背景」3D オブジェクトを基準として，複数の「物体」3D オブジェクトを配置した画像を出題し，その中から全オブジェクトに対して 1 つだけ非常識な大きさの「正解」オブジェクト(この例ではテーブル上の白いコップ)を選択できたユーザを人間とみなす．

従って，ユーザは解読のためにサイズ感をとらえる，すなわち背景の場所や，物体の 3D 空間内の位置関係を解読する必要があり，常識を持つ人間は容易に回答できると期待できる．一方，機械は輪郭抽出などにより，配置された物体の正体をおおむね解明し物体の実サイズを得る可能性があるが，出題画像内のサイズ感の把握ができない限り機械による突破は不可能である．また，3D オブジェクトを任意の位置に配置するため，無数の出題を生成できる．以上により提案手法は CAPTCHA の 3 要件を満たす．

提案手法で用いる背景及び物体の 3D オブジェクトのモデルは，実世界でのサイズ情報と共に，予めデータベースに大量に登録されているとする．その中から背景を 1 つ選択したのちに，その背景に対して大きすぎず，かつ小さすぎないオブジェクトを無作為に選択する．これは，例えば閉ざされた室内空間に，車など本来は屋外にあるような大きなオブジェクトが配置されると通常より大きく見え，また大きな空間に小さいオブジェクトが配置されると見えにくくなると筆者が感じたためである．

表 1　検証結果の平均

	前回	今回	既存
正答率	0.65	0.74	0.65
回答時間 [s]	4.64	4.79	3.71

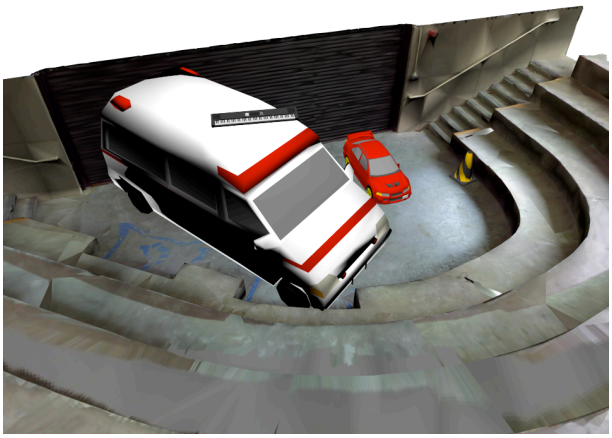

図 2　提案手法で正答率が低かった出題

なお，データベースに登録する物体は，人間が常識的なサイズ感を捉えにくいもの以外に限定する．具体的には，前回の検証[2]により，鉢植えなどのように，物体の特性としてサイズが一意に定まらないものや，ライフル銃などのようにサイズが一意に定められるとしても一般的に馴染みがないものは，サイズ感を捉えにくいという結果を得ている．

また，出題画像では，物体は背景に対して宙に浮かせず，互いに重ならないよう任意の位置に配置する．出題の 3DCG 画像生成時のカメラ位置は，配置した物体が全て映る範囲内で，無作為に定める．

4.　利便性の検証と考察

機械あるいは電気系の学科に属する 14 人の被験者に下記 3 つの出題画像群に回答してもらい，各出題画像の解答時間と正答率を記録した．全ての検証において，描画するオブジェクト数は 4 体とした．

1. 前回の検証[2]と同じ 26 枚の画像(正解オブジェクトの倍率は 0.5～0.75, 1.5～2 の中で乱数値).
2. 改善後の今回の手法において，各背景(3 種類)，正解オブジェクトの倍率 0.5，1.5 倍(2 種類)の全通りを組み合わせた 6 枚の画像．3D モデルには Web 上から収集した 3 種類の背景およびサイズ感が固定できる 22 種類の物体を用いた[1].
3. 既存手法の非現実画像 CAPTCHA[1]を再現し，生成した 6 枚の画像．2 つ目の検証と同じ 3D モデルを使用した．

本検証で得られた各手法の検証結果について，被験者の平均正答率および平均回答時間を表 1 に示す．また，今回の提案手法において，被験者の正答率が低かった出題例を図 2 に示す．

表 1 によると，前回手法に比べ今回手法は，正答率に改善が見られ，回答時間は同程度であると言える．また，正答率は今回の提案手法が既存手法を上回った．しかしながら，前回の検証[2]で理想とした 90%程度には達していない．下記に各不正解画像の原因とその対策を挙げる．

1 つ目の原因は，サイズ感が掴みにくいオブジェクトの使用である．ロードコーンやソファは筆者の先入観により，サイズ感が固定されると考えたために選択したが，被験者にとっては曖昧だったと思われる．対策として，将来的には，物体の実世界でのサイズ情報を Web 上で検索し，ショッピングサイトや百科事典などから自動収集する予定であるが，その際に複数の文献でサイズが一致するような物体に限定すればよい．ところが，キーボードは，サイズが明確であり，一般的な認知度も低くないと思われるが，サイズ感を捉えて正解できた例が少なかった．音楽を趣味としない人には分かりにくいなど，様々な原因を推測できるが，個々の物体の特性が要因であるため，根本的な対策は困難である．システムを運用する中で，学習によって正解率の低いオブジェクトを排除するなどの対策を講ずるしかない．

2 つ目の原因は，特定の 2 物体のみを比較してしまったことである．正解オブジェクトの近くに配置された物体のみと大きさを比較し，誤った方を選択してしまった例がいくつか見られた．図 2 はその一例であり，赤い車(正解オブジェクト)が通常より小さいが，それを基準に手前の救急車が大きく見え，誤って選択した被験者が多かった．大きさが大きく異なる物体同士が近くにあると比較しにくいことなどが原因だと思われる．今後の課題として，物体の大きさと配置による結果の違いを検証したい．

5.　おわりに

本研究では，物体のサイズ感を利用した 3DCG 画像 CAPTCHA 手法を提案した．前回の検証結果より改良し再検証した結果，正答率が改善され，既存手法に対して優位性が見られた．また，検証結果から，被験者が正解しにくいオブジェクトについて議論し対策を述べた．今後は，新たに見出した課題を解決し，機械攻撃手法を実装して，耐性の確認を行う．

参考文献

[1] 藤田真浩，池谷勇樹，可児潤也，西垣正勝，情処学論，Vol. 56, No. 12, pp. 2324‐2336 (2015).

[2] 西原大貴，新井イスマイル，物体のサイズ感を利用した 3DCG 画像 CAPTCHA の検討，情処研報，Vol.2016-CSEC-75，No.5 (2016).

[1] メタセコ素材！(http://sakura.hippy.jp/meta/)，Vynatic:Musical Keyboard(https://sketchfab.com/models/ae3dee84c9e54f8e88e90df0b45eda5c，CC-BY)にて公開されている素材を用いた．