主観的輪郭を応用したCAPTCHAの強度向上手法

小宮山 哲俊† 梅澤 猛† 大澤 範高†

†千葉大学大学院融合科学研究科

あらまし 実在しない輪郭線が知覚される錯視である主観的輪郭を応用したCAPTCHAは，従来の文字型CAPTCHAと比較してOCRプログラムによる解読に対する耐性が高い．しかし，この手法はSVM等の機械学習手法による解読に対して脆弱であることが指摘されている．本研究では，主観的輪郭錯視を利用して表現する文字形状に回転等の歪み処理を加えることで，SVMによる解読に対する耐性を高める手法を提案する．SVMを用いて認識精度を調査した結果，既存手法が82%に対して提案手法は20%であり，認識精度が低減することを確認した．更に，被験者による読み取り実験結果から，ユーザに大きな負担を与えないことが示唆された．

Improving the strength of CAPTCHA based on subjective contour

Tetsutoshi Komiyama† Takeshi Umezawa† Noritaka Osawa†

†Graduate School of Advanced Integration Science, Chiba University

Abstract CAPTCHA based on subjective contour is more resistant to OCR than text-based CAPTCHA. However, previous work suggests that CAPTCHA based on subjective contour can be solved by machine learning models such as SVM. In this paper, we propose a new method to improve the strength of CAPTCHA based on subjective contour. The proposed method adds rotation to the character shape and delineates it. We have evaluated our CAPTCHA and showed that a success rate of our CAPTCHA is 20%, which is lower than that of an existing method. Furthermore, we carried out a user study and it suggested that human users did not feel burdened by our CAPTCHA.

1 はじめに

CAPTCHA[1]は，応答者が人間かプログラムかを判別するテスト手法であり，メールアカウントの不正取得や，掲示板・ブログへの不正書き込みを行う悪意あるプログラムを排除する目的で広く利用されている．現在一般的な文字判読型のCAPTCHAについては，光学文字認識（OCR）の技術発展に伴い，OCRプログラムによって自動認識されてしまう事例が増加している[2]．錯視を利用した読み取りテストによってOCRに対する強度を高めた手法も提案されているが，機械学習手法を利用したOCRプログラムによって，自動認識される事例がある．主観的輪郭と呼ばれる現象を応用したCAPTCHA[3]は，従来の文字型CAPTCHAと比較してOCRプログラムによる解読に対する耐性が高い．しかし，この手法も機械学習手法を利用したOCRプログラムによる自動認識に対して脆弱であることが指摘されている[4].

本研究では，主観的輪郭を利用して提示する文字の形状に回転等の歪みを加えることで，サポートベクターマシン（SVM）による解読に対する耐性を高める手法を提案する．また，文字形状に加える歪みの種類を変化させた 6 つのデータセットを用意し，SVM への耐性が高い歪みの種類を調査した．更に，被験者による読み取り実験を行い，提案手法がユーザに与える負担を調査した．

2　関連研究

2.1　文字型 CAPTCHA の解読

文字型 CAPTCHA の解読は，主に前処理，領域分割，文字の認識の 3 つの手順で行われる．前処理では，領域分割と文字認識を容易にするためにノイズの除去を行う．領域分割では，文字列を 1 文字ごとの領域に分割する．文字の認識では，SVM やニューラルネットワーク（NN）を利用して，各文字を認識する．歪みやノイズを加えた 1 文字分の画像における文字認識においては，SVM や NN を利用した手法は，人間より優れた認識精度を示すため，機械による解読への耐性を高める方法としては領域分割を困難にする手法が広く利用されてきた[2][5]．領域分割を困難にする手法として，文字色と背景色を似た色にする，文字の上に線などのノイズを描画する，文字と文字との間隔を狭める，といった手法が主に利用される．しかし，これらの手法の多くは Anti-pattern[2]，Gibbs de-noising[6]，SDNN（Space displacements neural network[7]といった手法に対して脆弱であることが指摘されている．

2.2　錯視を応用した CAPTCHA

文字型 CAPTCHA が OCR プログラムにより自動認識される事例の増加に伴い，人間の高度な視覚処理機能により引き起こされる，錯視を利用した読み取りテストを利用した手法の研究が進められている．その一つとして，主観的輪郭という錯視現象を利用した手法が挙げられる[4]．主観的輪郭は存在しない輪郭線が知覚される錯視現象である．主観的輪郭を利用した CAPTCHA の画像例を図 1 に示す．輪郭線の知覚を誘導するための図形（誘導図形）を配置することで，文字の輪郭線を直接描画することなく文字‘A’を提示している．人間は錯視効果により文字を読み取ることができるが，文字の輪郭線を特徴量とした従来のパターンマッチング手法による OCR プログラムにとっては輪郭線が存在しないため認識が困難となる．しかし，この手法は SVM を利用した攻撃に対して脆弱であることが指摘されている[4]．

図 1　主観的輪郭を利用した CAPTCHA の一例

図 2　提案手法の一例

3　提案手法

本研究では，従来型の OCR だけではなく，SVM に対しても高い耐性を持った CAPTCHA を実現するため，主観的輪郭を利用して提示する文字形状に，歪みや回転を加える手法を提案する．提案手法の一例として，文字列‘AAR’の各文字を回転処理により変形し，主観的輪郭を応用して表現した画像を図 2 に示す．主観的輪郭で表現する文字の形状が増えることで，SVM よる認識率の大幅な低減が期待できる．

3.1　歪み処理

本研究では，波変換，回転，透視変換の 3 種類の画像処理を利用し，主観的輪郭で提示する文字の形状に変形を加えた．また，各歪み処理の一例として，文字‘A’に各歪みを加えた画像を図 3 に示す．

波変換

X 軸または Y 軸のどちらかの方向を選択し，正弦波による座標変換を行う．

回転

画像の中心を軸に文字画像を回転する．

透視変換

文字画像を任意の四辺形に変形する．

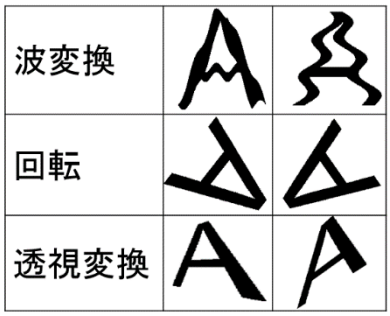

図 3 歪み処理の種類

3.2 文字画像生成アルゴリズム

本研究では，提案手法による文字画像を人間が認識できるように，以下に示す手順で錯視文字画像を生成した．

1 文字の表現に用いる誘導図形の個数：m
誘導図形候補リストの上限：n
誘導図形の直径： R

① 任意の文字画像を生成する．
② 文字画像に歪みを加える．
③ 空の誘導図形座標リスト，候補座標リストをそれぞれ用意する．
④ 文字上の点を一点選択する．
⑤ 誘導図形座標リストが空の場合，誘導図形座標リストに④で選択した点を追加し，④に戻る．
⑥ ④で選択した点と誘導図形座標リスト内の誘導図形座標とのユークリッド距離が，R 以下の場合（他の誘導図形と重なる場合），選択した点を候補リストに追加し，⑦に進む．
④で選択した点と他の誘導図形座標とのユークリッド距離が R より大きい場合（他の誘導図形と重ならない場合），選択した点を誘導図形座標リストに追加し，⑧に進む．
⑦ 候補リストが n を超えた場合，候補リストから誘導図形座標リストの全要素の座標との距離が最遠の点を選択し，誘導図形座標リストに追加した後に候補リストを空にする．
⑧ 誘導図形座標リストのサイズが m に満たない場合，④に戻る．
⑨ 文字画像と同サイズの画像を用意し，誘導図形座標リスト内のすべての座標に，誘導図形を描画する．この際，誘導図形の重心を誘導図形座標とする．

4 強度評価

本研究では，SVM による提案手法の認識精度を調査した．文字列画像に加える歪み処理を変化させた 6 種類のデータセットを用意し，SVM への耐性が高い歪みの種類を調査した．

4.1 実験条件

4.1.1 歪み処理

文字画像に強い歪みを加えた場合，人間の認識率が低下することが予測される．そこで，本研究では，文字画像に加える歪みの強さと人間による文字認識の可否を事前に確認し，人間による認識が容易になるように，以下の条件に従い，各データセットに歪み処理を施した．

波変換

波の振幅は，0 以上 12 未満の範囲内から無作為に決定した．また、周波数は 0 以上 0.03 以下の範囲内から無作為に決定した．

回転

回転角度は 0 度以上 360 度未満の範囲内から無作為に決定した．

透視変換

透視変換では，長方形の文字画像を任意の四辺形に変形することで歪みを加える．変換後の四辺形の各頂点の座標を以下の範囲内から無作為に抽出し，画像の変換を行った．
文字画像の横幅と高さをそれぞれ w，h

（pixel）とする．

四辺形の左上の座標：
（-w/5≦x≦w/5　，-h/5≦y≦h/5）
四辺形の右上の座標：
（4w/5≦x≦6w/5　，-h/5≦y≦h/5）
四辺形の右下の座標：
（4w/5≦x≦6w/5　，4h/5≦y≦6h/5）
四辺形の左下の座標：
（-w/5≦x≦w/5　，4h/5≦y≦6h/5）

4.1.2　主観的輪郭文字画像

本研究では，文字画像のフォントサイズを180pixel，誘導図形の形状は正円，円の直径は45pixelとして，3.2項の生成アルゴリズムに従い，主観的輪郭文字画像を生成した．また，実験用データセットとして，上記条件で生成した画像を32×32pixelに縮小した画像を用いた．

また，実験に使用した6種類のデータセットに加えた歪みの組み合わせを表1に示す．

4.2　実験手順

まず，主観的輪郭を利用して‘A’から‘Z’の26種類の文字を表す画像（32×32pixel）をそれぞれ2,000件ずつ，計52,000件用意した．つぎに，用意した画像をテストデータと学習データとして26,000件ずつに等分した．得られた学習データ26,000件を使ってSVMによる学習を行い，one-against-all法に従って画像を分類した．この際，SVMのカーネルとして，学習時間が短く，特別なパラメータ調整を必要としない，線形カーネルを利用した．

‘O’，‘C’，‘Q’等の文字は人間による誤認識が頻発することが予測される．そのため，実際に提案手法を運用する際には，‘O’，‘C’，‘Q’等の文字の誤認識については，許容する必要がある．そこで，以下の文字の組み合わせについてSVMが誤認識しても，識別が成功したとみなした場合の認識精度を調査した．

① 特別な処理をしない
② ‘C’と‘O’，‘O’と‘Q’
③ ‘F’と‘P’
④ ‘M’と‘W’
⑤ ‘N’と‘Z’
⑥ ‘C’と‘O’，‘O’と‘Q’，‘F’と‘P’
⑦ ‘C’と‘O’，‘F’と‘P’，‘M’と‘W’，‘N’と‘Z’，‘O’と‘Q’

実験結果を表2に示す．歪みを加えない場合，回転処理を加えた場合，透視変換を加えた場合の認識精度はそれぞれ81.8%，19.5%，41.2%となり，歪みを加えることで認識精度が大幅に低減することがわかる．また，条件①と条件②の認識精度の違いを比較すると，すべてのデータセットにおいて，差

表1　データセットの種類

データセット	波	回転	透視変換
1	×	×	×
2	×	○	×
3	○	×	×
4	×	×	○
5	×	○	○
6	○	○	×

表2　SVMによる提案手法の認識精度(%)

データセット	条件①	条件②	条件③	条件④	条件⑤	条件⑥	条件⑦
1(歪みなし)	81.8	83.7	82.5	82.1	81.8	84.4	84.6
2(回転)	19.5	22.4	20.1	20.4	20.0	22.3	24.2
3(波)	75.8	77.8	76.7	76.1	75.8	78.6	78.9
4(透視変換)	41.2	44.3	43.0	41.9	41.3	46.1	47.0
5(透視変換+回転)	9.9	12.8	10.3	10.4	10.2	13.2	14.0
6(波変換+回転)	17.9	20.8	18.5	18.9	18.3	21.4	22.8

は 4%以内に収まっており，ユーザによる‘C’と’O’，‘O’と’Q’の誤認識を許容したとしても，セキュリティ強度は大きく低減しないと考えられる．

5　ユーザの読み取り負担評価

主観的輪郭を利用した CAPTCHA に回答する際に，ユーザが感じる負担を調べるために被験者による文字読み取り実験を行った．提示手法による違いを調べるため，1）主観的輪郭を利用した英字 3 字の画像，2）主観的輪郭を利用した上で文字形状に回転を加えた英字 3 字の画像，3）主観的輪郭を利用した上で文字形状に透視変換と回転の両方を加えた英字 3 字の画像の 3 通りについて読み取り実験を行った．被験者が各画像の読み取りに掛かる時間を計測した後，質問紙調査によって負担の程度を調査した．質問は Q1 から Q3 までの 3 つで，1（強く不同意）～5（強く同意）の 5 段階による回答を得た．

Q1．主観的輪郭を利用した CAPTCHA の負担は軽い

Q2．主観的輪郭を利用した上で回転を加えた CAPTCHA の負担は軽い

Q3．主観的輪郭を利用した上で透視変換と回転を加えた CAPTCHA の負担は軽い

実験は被験者 9 人に対して行い，3 種の CAPTCHA 手法についてそれぞれ用意した 100 枚の画像から 各種 1 枚ずつ計 3 枚の画像を被験者毎にランダムに提示した．また，被験者が回答を誤った場合には，同じ CAPTCHA 手法による別画像を再提示し，正解するまで解答を繰り返すこととした．

実験の結果，3 つの CAPTCHA 提示手法それぞれの読み取りに要した平均時間は，それぞれ 1）5.8 秒，2）6.9 秒，3）12.2 秒であった． また，正解率はそれぞれ 1）90%，2）90%，3）70%であった． 質問紙による負担調査の結果は図 4 に示す．質問紙による負担調査の結果，歪みを加えない手法と回転を加えた手法については負担を重く感じたユーザはほとんどいなかった．これに対して，透視変換と回転を加えた手法については多くのユーザが負担を重く感じていた．また，全ての提示画像について読み取りの負担を「非常に重い（強く不同意）」と 回答した被験者はいなかった．

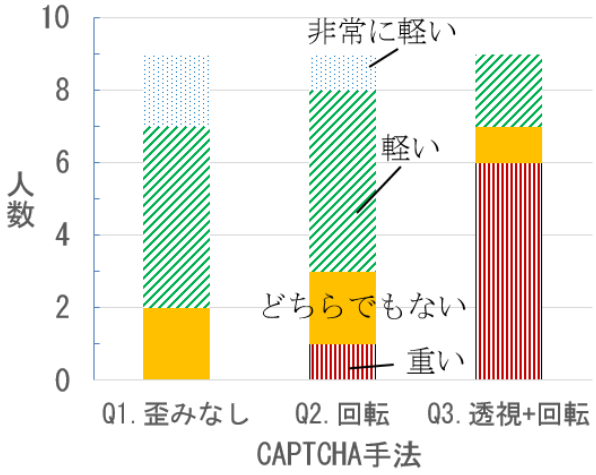

図 4　ユーザの読み取り負担

表 3　文字数と SVM による認識精度(%)

データセット	1 文字	2 文字	3 文字	4 文字
歪みなし	83.7	70.1	58.6	49.1
回転	24.2	5.9	1.4	0.3
透視+回転	14.0	2.0	0.3	0.03

6　考察

6.1 CAPTCHA の強度

1 回の CAPTCHA チャレンジで提示する文字数に関する検討を行う．画像中に含まれる文字数が 1~4 文字の場合の SVM による認識精度を試算した結果を表 3 に示す．この際，回転処理を加える手法については，一文字あたりの認識精度として条件⑦の結果を利用した．歪みを加えない手法については，一文字あたりの認識精度として条件②の結果を利用した．表 3 を見ると，透視変換と回転処理を加えた主観的輪郭 CAPTCHA を 4 文字で運用することで，SVM による認識精度を

0.03%まで低減できることが試算される． また，比較的ユーザへの負担が少ないと考えられる回転処理を施した主観的輪郭 CAPTCHA を 4 文字で運用することで，SVM による認識精度を 0.3%まで低減できることが試算される．

6.2 ユーザの負担

ユーザが各 CAPTCHA の読み取りに要した時間を比較すると，回転を加える手法は，歪みを加えない手法と比べて，その差は 1.1 秒と大きな違いは見られない．これに対して，透視変換と回転を加える手法は，歪みを加えない手法と比べて，その差は 6.4 秒と大きな違いが見られる．また，正解率について比較すると，回転を加える手法は，歪みを加えない手法の正解率はどちらも 90%であるのに対して，透視変換と回転を加える手法の認識精度は 70%と低いことがわかる．従って，読み取りの所要時間および正解率の点で，歪みを加えない既存の手法と比較すると，回転を加える手法がユーザに与える負担増は限定的であるものの，透視変換と回転を加える手法がユーザに与える負担は大きいと考えられる．

つぎに，ユーザが感じる負担感について，図 5 の Q1 と Q2 を比較すると，歪みなしの手法については負担が重いと感じた被験者が 1 人もいないのに対して，回転を加える手法については 1 人の被験者が負担を重く感じていた．しかし，その他の被験者については，どちらも同じ評価をしており，2 つの提示手法がユーザに与える負担の差は大きくないと考えられる．また， Q2 と Q3 を比較すると，透視変換と回転を加える手法の負担については重いと答える回答が多いことがわかる．これらのことから，透視変換と回転を加える手法と比べて，回転のみを加える手法がユーザへ与える負担は小さいことがわかる．

7 結論

本研究では，主観的輪郭錯視を利用して表現する文字の形状に歪みや回転を加えることで，SVM による攻撃に対して高い耐性をもつ手法を提案した．強度評価実験の結果，提案手法を用いることで，SVM による攻撃に対する耐性が大幅に高まることが分かった．更に，被験者による読み取り実験の結果，主観的輪郭錯視を利用して表現する文字の形状に回転を加える文字の提示手法は，人間に大きな負担を与えないことが示唆された．

今後は，被験者数を増やした実験を行うことで統計的な有意性について検証していきたい．また，SIFT 特徴量などの回転に強い特徴量を利用した文字認識手法に対する耐性も評価していく必要がある．

参考文献

[1] M. Blum, L. A. von Ahn, and J. Langford. “Telling humans and computers apart automatically”, Communications of the ACM, 2004. pp. 56-60

[2] E. Bursztein, M. Martin, and J. Mitchell. “Text-based CAPTCHA strengths and weaknesses”, Proc. 18th ACM CCS, 2011. pp. 125-138

[3] 橋本弥弦. “主観的輪郭の CAPTCHA への応用“, 早稲田大学大学院 理工学研究科 修士論文 (2008)

[4] 小宮山哲俊, 梅澤猛,大澤範高. “主観的輪郭を応用した CAPTCHA の強度評価" FIT2015, 2015

[5] J. Yan and A. S. El Ahmad. “Low-cost Attack on a Microsoft CAPTCHA”, Proc. 15th ACM CCS, 2008. pp. 543-554

[6] S. Gema and D. Geman. “Stochastic relaxation, Gibbs distributions and the Bayesin restoration of images*”, J. Appl. Statistics 20, 1993. pp. 25-62

[7] O. Matan, C. J. C. Burges, Y. Lecun and J. S. Denker. “Multi-digit recognition using a space displacement neural network”, NIPS, 1993, pp. 488-495