情報処理学会論文誌　Vol. 50　No. 9　2062–2076 (Sep. 2009)

不鮮明化画像を利用した暗示・応答型画像認証方式の提案

山本　匠[†1,†2]　漁田武雄[†3]　西垣正勝[†1,†4]

本論文では，一見すると無意味に見える不鮮明な画像を活用することで，暗示・応答型画像認証方式（Cue and Response 型画像認証方式）という新しいコンセプトの Challenge & Response 型画像認証方式を提案する．本方式では，不鮮明化画像を用いてユーザにチャレンジ（キュー）の意味を非明示的に伝える．登録時にオリジナル画像を見ている正規ユーザだけは，キューの意味を理解することができる．このように，システムから提示されるチャレンジの意味を攻撃者から隠すことで，レスポンス生成の方法が単純であっても，覗き見に対するある程度の安全性を確保することが可能である．これにより，正規ユーザであれば直感的な処理によって毎回のキューに対するレスポンスを生成することができ，かつ，攻撃者による有限回の覗き見に対する耐性を持つ，という 2 つの特長を有する画像認証方式が実現可能である．基礎実験を通じ，本方式の実現可能性を評価する．

Proposal of a Cue & Response Image-based User Authentication System Using Unclear Image

Takumi Yamamoto,[†1,†2] Takeo Isarida[†3] and Masakatsu Nishigaki[†1,†4]

We have recently proposed a user authentication system using “unclear images” as pass-images, in which only a legitimate user can understand their meanings by viewing the original images corresponding to the unclear pass-images. These unclear images are meaningless for unauthorized users. Hence it is difficult for unauthorized users to memorize the unclear pass-images, even though they observe the legitimate user’s authentication trial. In this paper we propose a new type of Challenge & Response authentication by using a feature of unclear image, which we call as “Cue & Response (Q&R) image-based user authentication”. In this authentication, unclear image are used to convey challenges (cues) to only the legitimate user. Thus, it is expected that even simple calculation of a response can achieve a certain level of robustness against observing attackers since the meaning of a challenge is hidden from unauthorized users. By doing this, only a legitimate user can respond to a randomly changing cue properly, while unauthorized users are prevented from impersonating with limited amount of observing attacks. We conduct fundamental experiments to study the availability of the proposed Q&R image-based user authentication system.

1. はじめに

現行のユーザ認証方式は，汎用性と利便性の高さからパスワード方式が主流となっているが，人間にとって長くランダムな文字列を記憶することは容易ではない．そのため，人間の画像認識能力の高さを利用して記憶負荷を軽減させる画像認証方式[1),2)]が注目されている．しかし，再認型の認証となる画像認証においては，毎回の認証時にパス画像がディスプレイ上に表示されるため，認証時の覗き見攻撃に対して脆弱となる．この問題への対策としては，画像認証方式をワンタイム化する方法[3)–7)]と認知心理学的に攻撃者の画像認識を妨害する方法[8)]に大別される．

画像認証をワンタイム化する方法においては，ネットワーク認証プロトコルで利用されている Challenge & Response 型認証方式（以降，C&R と略記する）にならい，画像認証を C&R 型に改良するための研究が行われている．しかし，人間は複雑な計算は不得手であるため，パスワードと乱数（チャレンジ）をハッシュ化してレスポンスを返すというようなことは不可能である．Sobrado らの方式[3),4)]は，pass-object を頂点とした凸包内部を選択させることで，比較的良好な C&R 型画像認証方式を実現している例といえるが，やはり，ユーザにとって，チャレンジに対するレスポンスを生成することは容易なことではない．Roth らの方式[5)]は，認証情報に付与されているグループ情報を回答させるというアイデアによって，レスポンスの生成に対するユーザの負荷を軽減させることに成功しているが，その代わりに，認証情報の入力回数が激増してしまう．

†1 静岡大学創造科学技術大学院
Graduate School of Science and Technology, Shizuoka University
†2 日本学術振興会特別研究員（DC）
Research Fellow of the Japan Society for the Promotion of Science (DC)
†3 静岡大学情報学部
Faculty of Informatics, Shizuoka University
†4 科学技術振興機構，CREST
Japan Science Technology and Agency, CREST

また，セキュア ID[9] のように認証情報を毎回更新するタイプのワンタイム画像認証方式も提案されている[6),7)]．ただし，画像認証においてパス画像を毎回覚え直すことはユーザにとって大きな負荷になってしまう．そこで，徐らの方式[6)] ではニーモニックを導入することによって，fakepointer[7)] では画像の短期記憶を活用することによって，それぞれ記憶負荷を抑える工夫をしているが，その記憶負荷は依然として大きい．

一方，画像認識を妨害する方法では，覗き見をする攻撃者にとってパス画像の記憶が困難となるように，モザイク化などの不鮮明化処理を施した一見無意味な画像をパス画像として使用する方法が提案されている[8)]．正規ユーザにのみオリジナル画像を見せ，スキーマ（オリジナル画像と不鮮明化画像の間の認知構造的なリンク）[10)] を学習させることにより，正規ユーザは不鮮明化画像を有意味な画像として認識できるようになり，パス画像を容易に記憶することができる．人間は画像の記憶に優れているという特性を有するものの，それは有意味な画像を記憶する場合に限ってのことであり，無意味に見える（意味を言語化できない）画像を記憶することはやはり難しい[11),12)]．ゆえに，他人のパス画像（不鮮明化画像）を覗き見て記憶することは，攻撃者にとって困難な作業となる．しかし，不鮮明化画像の認識は（困難ではあるが）不可能ではないため，認証の C&R 化，またはパス画像の定期的な変更が望まれる．

以上のように，前者のワンタイム画像認証における C&R 型画像認証方式には，安全性を維持したままレスポンス生成処理を簡素にしたいという要望があり，後者の不鮮明化画像認証方式には，認証を C&R 化したいという要望がある．そこで本論文では，両者を融合することによって，より効果的な画像認証方式を実現することを目指す．

具体的には，不鮮明化画像の特徴[8)] を活用し，認証のたびに異なる質問（チャレンジ）を正規ユーザのみが理解することができる形で提示する．チャレンジの意味を攻撃者から隠すことができれば，レスポンス生成の方法が単純であっても，覗き見に対する安全性をある程度確保することが可能であると考えられる．本論文ではこのような暗示的なチャレンジを「キュー（Cue）」と呼び，本方式を暗示・応答（Cue & Response：Q&R）型画像認証方式と呼ぶことにする．

2. 覗き見攻撃に対する既存の画像認証方式

覗き見攻撃に耐性を持たせることを目的とした画像認証方式を以下の 4 つに分けて，それぞれを簡単に説明し，課題を示す．

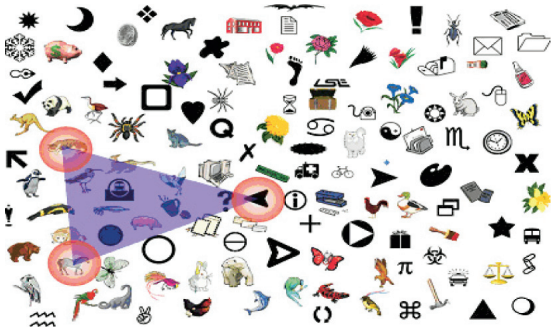

図 1　Sobrado らの認証システム[3)] の認証画面
Fig. 1　Authentication window in Sobrado, et al.'s scheme[3)].

2.1 C&R 型画像認証方式

C&R 型画像認証の代表的な方式として Sobrado らが提案する方式[3),4)] がある．この方式では，チャレンジとして，システムから多数のアイコンがランダムに配置された画面が提示される．ユーザは，あらかじめ登録しておいた複数の pass-object（3 つ以上）を画面の中から探し出し，pass-object を頂点とした凸包内部を選択することでレスポンスを返す（**図 1**）．

この作業を複数回繰り返すことで認証の可否を判断する．ユーザからのレスポンスを「凸包の内部」という曖昧な形で返すため，覗き見攻撃者に凸包を構成する pass-object が一意に漏洩しない．しかし，正規ユーザにとって多数の object（アイコン）の中から特定の pass-object を探し出す作業は容易なことではなく，認識負荷の点で問題を残している．

一方，Roth らが提案する方式[5)] では，認証情報に付与されているグループ情報を回答させるというアイデアによって，レスポンスの生成に対するユーザの負荷を軽減させることに成功している．この方式では，認証時には 0〜9 までの数字が並べられた画面が表示される．各数字の背景は白もしくは黒のどちらかの色でランダムに塗られており，それによって各数字がどちらの色のグループに含まれるのかが示されている．ユーザは自分の暗証番号の数字の背景色が白なのか黒なのかを答える（**図 2**）．これを各桁につき複数回行う．ユーザが色を答えるたびに画面上の数字の背景色はランダムに塗り直される．

暗証番号を記憶している正規ユーザはレスポンス（背景色）を容易に返答することが可能である．しかし，レスポンスの選択肢が限られる（白と黒の 2 択）ので，十分な総当たり数

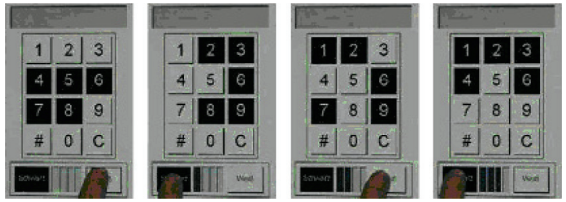

図 2　Roth らの認証システム[5] の認証画面の例
Fig. 2　Authentication window in Roth, et al.'s scheme [5].

図 3　徐らの認証システム[6] の認証画面の例
Fig. 3　Authentication window in Jo, et al.'s scheme [6].

を確保するためには問答を繰り返す必要があり，入力回数が激増してしまう．

2.2　パス画像更新型認証方式

パス画像更新型認証方式は，セキュア ID[9] のように認証情報を毎回更新するタイプのワンタイム画像認証方式である．しかしながら，画像認証においてパス画像を毎回覚え直すことはユーザにとって大きな負荷になってしまう．そこで，徐らはストーリづけによる記憶補完（ニーモニック）を導入することによって，パス画像更新時の記憶負荷を軽減しようと試みている[6]（**図 3**）．しかし，ストーリによる記憶負荷軽減の効果が十分でないこと，および，パス画像更新のたびにストーリを考えること自体がユーザの負荷となることなどの問題が残る．

一方，fakepointer[7] ではパス画像の短期記憶を活用することで記憶負荷を抑える工夫をしている．fakepointer は暗証番号をパス画像（背景画像）に合わせることによって認証が

図 4　fakepointer[7] の認証画面の例
Fig. 4　Authentication window in fakepointer [7].

行われる（**図 4**）が，認証のたびにパス画像が変更される．ここでパス画像は，時間的・空間的に異なる通信路を介して，認証操作の直前にユーザに送られてくる．すなわち，ユーザがパス画像を記憶していなければならない時間は，パス画像が届いてから認証操作を行うまでの短期間のみとなる．しかし，パス画像の具体的な取得方法に疑問が残るうえに，パス画像の短期記憶の負荷が本当に低いかどうかに関する実験や評価もなされていない．

2.3　画像認識妨害型認証方式

画像認識妨害型認証方式では，覗き見をする攻撃者にとってパス画像の認識が困難となるように，モザイク化などの不鮮明化処理を施した一見無意味な画像（**図 5** 右）をパス画像として使用する[8]．無意味に見える画像は，それ単体だけでは正規ユーザにも認識・記憶が困難である．しかし正規ユーザにのみオリジナル画像（図 5 左）を見せることにより，正規ユーザは不鮮明化画像を有意味な画像として認識できるようになり，パス画像を容易に記憶することができる．これは，不鮮明なパス画像に対する「スキーマ[10]」を正規ユーザに学習させていることに相当する．ここでスキーマとは，人間が外界からの情報を知覚した際に無意識のうちに蓄積している「その情報をどのように認識・記憶したかという知識構造」を意味する認知心理学用語である．人間は，ひとたび不鮮明化画像に対するスキーマを学習すれば，それ以降にその不鮮明化画像を見た場合に，スキーマを活用することによって簡単にその意味を再認識することが可能になる．

しかし，不鮮明化画像の認識は困難ではあるが不可能ではないため，認証の C&R 化，ま

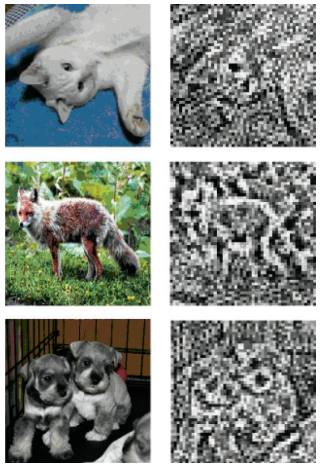

図 5　原田らの認証システム[8] で用いられる不鮮明化画像（右）とそのオリジナル画像（左）の例
Fig. 5　The original images and the corresponding unclear images used in Harada, et al.'s scheme [8].

たは，（記憶負荷を増加させることなく）パス画像を定期的に変更できると望ましい．

2.4　チャレンジ隠蔽型 C&R 画像認証方式

Sasamoto らの方式[13] は，チャレンジそのものを秘密の通信路を介してユーザに提示することで，ビデオカメラによる複数回の盗撮に対しても高い耐性を実現している．認証時にはディスプレイから与えられる明示チャレンジ（visible challenge）と触覚デバイスから与えられる隠蔽チャレンジ（hidden challenge）とからレスポンスを生成する．隠蔽チャレンジは触覚デバイスの掌によって知覚される．触覚デバイスはユーザの掌によって隠されており，隠蔽チャレンジは外部からは見えない．隠蔽チャレンジを外部から取得することは難しく，覗き見（ビデオカメラによる撮影）によって取得可能な情報（明示チャレンジとレスポンス）だけでは，正規ユーザになりすますことは困難である．

しかし Sasamoto らの方式では，チャレンジの隠蔽を実現するために特殊な装置（知覚デバイス）を必要とする（**図 6**）．

3.　Q&R 型画像認証方式

3.1　コンセプト

本論文では，2 章で概説した既存方式のうち，2.1 節で示した C&R 型画像認証方式と 2.3 節で示した画像認識妨害型認証方式（不鮮明化画像認証方式）に焦点をあてる．C&R 型画像認証方式には，その安全性を維持したまま，画像認識妨害型認証方式のようにレスポンス生成処理を簡素にしたいという要望がある．一方，画像認識妨害型認証方式には，その認識容易性を維持したまま，C&R 型画像認証方式のように何らかのワンタイム性を追加したいという要望がある．そこで本論文では，C&R 型画像認証方式と画像認識妨害型認証方式の両者を融合することによって，より効果的な画像認証方式を実現することを目指す．

図 6　Sasamoto らの認証システム[13]
Fig. 6　Authentication system proposed by Sasamoto, et al. [13].

具体的には，「不鮮明なパス画像に関する情報を言葉で与えただけでは，スキーマを持たない攻撃者はパス画像を特定するに足る情報量を得ることができない」という不鮮明化画像の特徴[8] を活用し，認証のたびに異なる質問（チャレンジ）を正規ユーザのみが理解することができる形で提示する．

通常の C&R 型認証においては，攻撃者もチャレンジとレスポンスの両者を観測することができるため，1 方向性を有する複雑な計算によってチャレンジからレスポンスを生成しなければ，覗き見攻撃に耐えられない．これに対し，チャレンジの意味を攻撃者から隠すことができれば，レスポンス生成の方法が単純であっても，覗き見に対する安全性をある程度確保することが可能であると考えられる．

ここで重要なことは，提案方式においては，2.4 節のチャレンジ隠蔽型 C&R 画像認証方式のようにチャレンジそのものを物理的に隠蔽するのではなく，チャレンジの意味を認知心理学的に攻撃者から隠している点である．これによって，特殊な装置を使うことなく，チャレンジ隠蔽型 C&R 画像認証方式と同等の効果が期待される．

提案方式においては，認証のたびに異なる質問が正規ユーザのみが知覚できる形で提示さ

れる．本論文では，このような暗示的なチャレンジを「キュー」と呼び，本方式を暗示・応答（Cue & Response：Q&R）型画像認証方式と呼ぶ．Q&R 型画像認証方式は，正規ユーザであれば直感的な処理によってキューに対するレスポンスを生成することができ，かつ，攻撃者による有限回の覗き見に対する耐性を持つ，という 2 つの特長を有する C&R 型画像認証方式となっている．

3.2 認識方式

先行研究[14] では，「不鮮明なパス画像に関する情報を言葉で与えただけでは，スキーマを持たない攻撃者はパス画像を特定するに足る情報量を得ることができない」という不鮮明化画像の特徴[8] を活用し，認証時に言葉によってパス画像に関する手がかりを提示することで，覗き見攻撃者によるなりすまし成功率を増大させることなく，正規ユーザの認証時の認識負荷を軽減する方式を提案している（**図 7**）．

本論文ではこのアイデアを拡張し，先行研究[14] で用いられた手がかり情報の提示を，正規ユーザのみにチャレンジ（キュー）を認識させる手段として利用する（**図 8**）．キューは，パス画像に対する部位情報を言語手がかりで示したもの（例：「左目」，「右耳」，「尻尾」，「左前足」など）であり，認証のたびに変化する．ユーザは，囮画像に紛れているパス画像を見つけたうえで，キューによって指示された部位に対応する場所をクリックすることによってレスポンスを返す．

スキーマを持たない攻撃者には，不鮮明化画像の意味を認識することは困難であるため，指示された部位に対応する場所を正しくクリックすることは容易なことではない．一方，不鮮明化画像の意味（スキーマ）を知っている正規ユーザであれば，指示された部位をクリックすることは容易である．キューにより指定される部位は認証のたびに変化する（ある認証フェーズで「左耳」をクリックしている瞬間を覗き見られたとしても，次回の認証においてはたとえば「左前足」という指示に変わる）ため，覗き見攻撃に対する耐性も増加する．また，この方法は，パス画像選択の総当たり数を増やすことを可能にするというメリットもある．

しかし，この方法においては，たとえば「左目」というキューに対する正規ユーザのレスポンスを覗き見ていた攻撃者は，「正規ユーザがクリックした場所の付近が，正規ユーザが認識している左目である」という情報を得ることができてしまう．攻撃者は，複数回これを繰り返すことにより，正規ユーザが認識する不鮮明化画像の全体像（たとえば動物の写真画像の場合，顔，手，足，胴体の位置関係など）を認識することが可能かもしれない．

そのため本論文では，キュー（部位情報）を言語手がかりとして直接的に示すのではなく，別の不鮮明化画像中の部位としてユーザに暗示的に示す方法を採用する．以降，キューを示すために用いられる不鮮明化画像を参照画像と呼ぶことにする．

図 7　先行研究[14] の概観
Fig. 7　Authentication system with verbal cue [14].

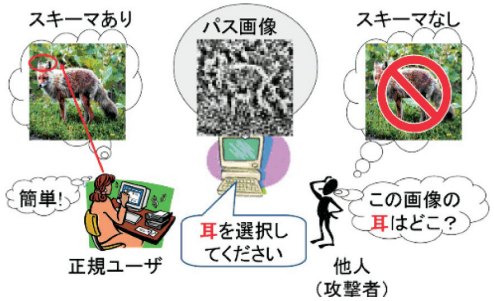

図 8　言語てがかりによって選択する部位を指示する認証方式の概観
Fig. 8　Authentication system where user responds by clicking a position designated by verbal cue.

正規ユーザはパス画像登録時に，参照画像とそれに対応するオリジナル画像も一緒に記憶し，参照画像のスキーマを学んでおく．すなわち，パス画像に対するスキーマと参照画像に対するスキーマの両方を学習しておく．認証時には，システムはパス画像と複数の囮画像を認証ウインドウに提示する（**図 9** 左）．同時に，参照画像中の任意の部位（以降，パス部位と呼ぶ）を選び，その位置に目印をつけた形でこれを参照ウインドウに表示する（図 9 右）．参照画像のスキーマを有する正規ユーザは，参照画像中の目印からパス部位を認識することができる．また，正規ユーザはパス画像のスキーマも学習しているので，認証ウイン

図 9　認証ウインドウ（左）と参照ウインドウ（右）の例[⋆1]
Fig. 9　Example of the authentication window (left-side) and the reference window (right-side).

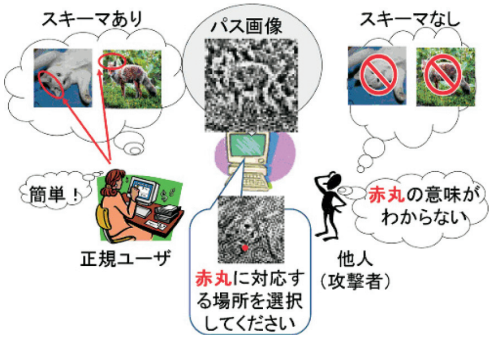

図 10　提案方式（Q&R 型画像認証システム）の概観
Fig. 10　Proposed system (Q&R image-based user authentication system).

ドウの中からパス画像を見つけたうえで，パス画像におけるパス部位をクリックすることが可能である．すなわち，参照画像上の目印によって提示されたパス部位が「右足」であったとすると，正規ユーザはパス画像の「右足」付近をクリックすることになる．

攻撃者は参照画像に対するスキーマを有していないため，参照画像上の目印を覗き見たとしても，その位置に何が映っているのかを類推することは容易ではない．さらに，攻撃者はパス画像に対するスキーマも持っていないため，正規ユーザのレスポンスを覗き見たとしても，クリックの位置に何が映っているのかを類推することも難しい．これにより，パス部位を攻撃者に理解できない形で正規ユーザにのみ提示することが可能となる．

パス部位の位置は，毎回の認証でランダムに決定される．参照画像上に目印として提示されるパス部位が毎回の認証のキューであり，ユーザによるパス画像上のパス部位のクリックがキューに対するレスポンスである．**図 10** に提案方式（Q&R 型画像認証方式）の概観を示す．

言語手がかりを用いる方式においては，覗き見によって得られたパス画像のパス部位と言語手がかりとの対応情報（たとえば，「パス画像中の座標位置 (X1, Y1) が眼」）から，攻撃者は「鼻は眼の下のあたりにある」，「顔の下に体がある」などというように，パス画像の構造を予想していくことが可能である．しかし参照画像を用いる方式では，攻撃者が覗き見によってパス画像と参照画像のパス部位の対応情報を得たとしても，それは「無意味に見えるもの」と「無意味に見えるもの」との対応付けでしかないため，そこから画像の構造を推測していくことは困難になる[⋆2]．

3.3　不鮮明化画像における部位情報の登録

スキーマは，正規ユーザが不鮮明化画像の意味を再認識するにあたっての大きな手がかりとなる．しかしそれは，「不鮮明化画像を見た際に，正規ユーザの頭の中に鮮明なオリジナル画像が再び蘇る」という現象が起こっているわけではないということに注意しなければならない．本来人間は不鮮明な画像やランダムドットのような画像を見た場合にも，無意味な点と点，線と線の間に何らかのまとまった関連を見つけ，そこに何らかの意味を見い出そうとする性質を持っている[15),⋆3]．スキーマは，この「不鮮明化画像から意味を見い出す作業」を後押しするものであり，「オリジナル画像を不鮮明化した際に失われた情報を修復し，正規ユーザの頭の中に鮮明なオリジナル画像を復元する」までの効力はない．

このため，画像によっては，オリジナル画像と不鮮明化画像とで正規ユーザが認識する部位の有無・位置・大きさが異なってくるようなケースが発生する．そこで本方式では，オリジナル画像上で認識される部位を用いるのではなく，不鮮明化画像をユーザに見せ，その中で「ユーザにとって部位として認識できる場所」を登録してもらうという方法をとる．

⋆1 図の大きさの都合上，赤丸は実際のサイズより相対的に大きく記してある．

⋆2 攻撃者が覗き見を繰り返すうちに，過去に表示されたチャレンジ（キュー）と同じチャレンジが提示された場合は，攻撃者がその時点で取得しておいたレスポンスをリプレイすることによって，なりすましが可能である．この攻撃に対しては，言語手掛かりを用いる場合も参照画像を利用する場合も耐性はない．

⋆3 このような人間の認知の仕組みは，ゲシュタルト心理学派の研究者が提唱してきた，人間が個別の情報をより簡潔な意味を持つ全体的な情報として認識するという，いわゆるゲシュタルトの法則によって説明される場合もある．

不鮮明化画像がまったく無意味な画像であった場合には，不鮮明化画像の中から部位として認識できる場所を見つけ出す作業は人間にとって容易なことではない．しかし本方式においては，正規ユーザは，まずオリジナル画像と不鮮明化画像の両者を見てスキーマを獲得することにより，不鮮明化画像の中に意味を見い出すことができている．よって，登録フェーズにて不鮮明化画像の中の部位情報をシステムに回答することは正規ユーザにとっては大きな負荷にはならないと考えられる．

4. 基 礎 実 験

提案方式の有効性を確かめるために基礎実験を行い検証する．

4.1 本人認証実験

a) 実験の目的

正規ユーザにとって，参照画像から与えられるキュー（パス部位）を正しく認識し，パス画像中のパス部位を的確に選択することが可能かどうかを確認する．

b) 実験方法

先行研究[8] のシステムと比較ができるように，被験者に記憶してもらうパス画像の数は 4 枚とし，9 択 × 4 ターンの試行を 1 回の認証とする．実験で使用する画像も，先行研究[8] のシステムにあわせ，様々な種類の動物が写っている背景つきの写真画像 100 枚とした．今回の実験では，被験者に記憶してもらう参照画像は 1 枚とした．

本実験の被験者は本学情報学部学生 10 名である．全被験者には，それぞれ 4 枚のパス画像と 1 枚の参照画像を記憶してもらう．パス画像 4 枚と参照画像 1 枚はすべて異なる画像である．また，実験で使用したすべての画像に対して，被験者自身に部位情報を登録してもらう．今回の実験では動物の写真画像を使用しているため，目，鼻，耳，右前足，左後足，胴体，尻尾などが部位として登録されることになる．ただし，被験者の手間を軽減させるため，まずは実験実施者が手動で部位を登録しておき，被験者にその位置・大きさを修正[⋆1] してもらうようにした．図 5 のきつねと犬の画像に対し，実験実施者が登録した部位の例を**図 11** および**図 12** に示す．ただし，3.3 節で述べたように，被験者による部位情報の登録は（オリジナル画像ではなく）不鮮明化画像を見ながら行ってもらうことに注意されたい．このため，登録される部位情報は，オリジナル画像上で認識されるものとは異なりうる．

⋆1 実験実施者が前もって登録しておいた部位の中で，被験者が認識できないものがあった場合にはこれを削除したり，実験実施者が登録したもの以外に被験者が認識できる部位があった場合にはこれを追加登録したりすることも自由に許した．

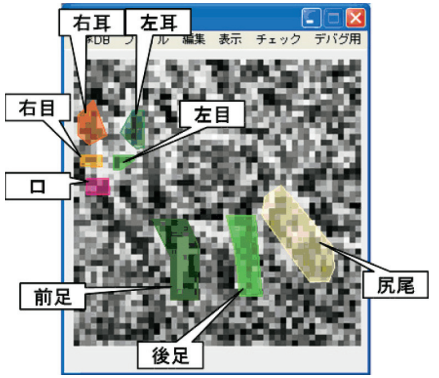

図 11　部位登録画面の例 1
Fig. 11　1st example of position registration.

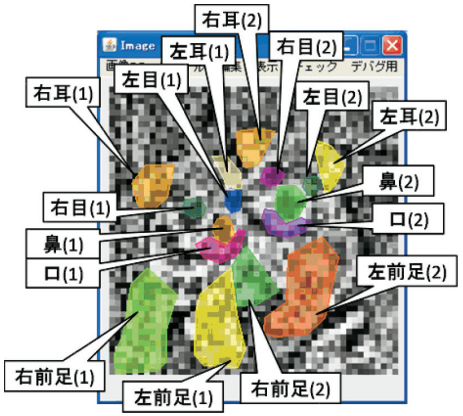

図 12　部位登録画面の例 2
Fig. 12　2nd example of position registration.

なお，図 11 および図 12 に付されている各部位の名称は，説明を分かりやすくするために記したものであり，実験の際に被験者には提示されない．また，図 12 の部位の名称には「(1)」，「(2)」というラベルが記されているが，図 5 の犬の画像において左の犬を犬 (1)，右の犬を犬 (2) としている．

認証時には，パス画像 1 枚と囮画像 8 枚が表示されている認証ウインドウ（図 9 左）と，参照画像 1 枚が表示されている参照ウインドウ（図 9 右）が被験者に提示される．パス画像，囮画像，参照画像のそれぞれの画像はすべて，300 × 300 pixel の大きさで画面に表示される．参照画像上にはパス部位が赤い丸でプロットされる．参照画像上に表示されるパス部位は，部位の重心の位置に半径 3 pixel でプロットすることとした．

被験者は，参照画像中のパス部位（赤い丸）を認識し，認証ウインドウ中の 9 枚の画像の中から自分のパス画像を探し出したうえで，パス画像上におけるパス部位を選択する．なお，画像においては左右の概念（動物自身の右なのか，画像を見ている被験者から見て右側なのか）が曖昧になるため，今回の実験では左右の区別はしないこととした．また，画像の中に複数の動物が存在する画像においては，何番目の動物であるかを指示することは煩雑であると考え，今回の実験ではどの動物であるかは区別しないこととした．

これを図 11，図 12 の例を用いて簡単に説明する．たとえば，参照画像中のキュー（パス部位）が「左前足」であり，その際のパス画像が図 11 であった場合，被験者はパス画像（図 11）の中の「左前足」をクリックしても「右前足」をクリックしても認証される．また，参照画像中のキュー（パス部位）が同様に「左前足」であり，その際のパス画像が図 12 であった場合は，被験者はパス画像（図 12）の中の犬 (1) の「左前足 (1)」と「右前足 (1)」，犬 (2) の「左前足 (2)」のどれをクリックしたとしても認証されることになる．

今回の実験では，登録される部位情報の数は 1 枚の画像あたり平均 8.2 カ所（標準偏差 = 2.13，最小値 = 5，最大値 = 16）であった．ただし，上記のように左右および動物の順序を区別せずにカウントした場合（たとえば，「右前足」と「左前足」は合わせて 1 つと数えた場合）は，部位情報の数は 1 枚の画像あたり平均 5.6 カ所（標準偏差 = 0.88，最小値 = 4，最大値 = 7）であった．

パス画像を変えながらこの操作を 4 ターン行って，認証可否の判定を行う．キューとなるパス部位は，登録されている複数の部位情報の中から，ターンごとにランダムに選択される．

パス画像（および参照画像）登録日から 1 日後と 8 日後に，各被験者につき 10 回ずつ認証を行ってもらった．登録後，被験者は認証実験以外の場でパス画像，参照画像，および，それらのオリジナル画像を確認することはできない．

c) 実験結果

実験結果を**表 1** に示した．表中，「認証成功率」は，各認証試行において認証に成功した（1 回の認証において，4 ターンのパス部位選択すべてに成功した）割合である．一方，「ターンごとの成功率」は，各認証試行時に行う 4 ターンのパス部位選択（9 択の不鮮明化画像の中からパス画像 1 枚を探し，その中のパス部位を選択するタスク）を独立にとらえ，1 ターンごとの成功率を表したものである．なお，今回はユーザの画像認識における曖昧性を吸収するために，登録されている各部位の領域境界から $\theta = 10$ pixel まで（領域境界から 10 pixel 未満）を選択成功範囲とした．表 1 は比較のために，$\theta = \{0, 10, 15, 20\}$ に対する認証成功率についても示した．「画像選択の成功率」は，パス部位の選択については無視し，パス画像の選択のみを考慮した認証試行としてとらえた場合（先行研究の画像認識妨害型認証方式[8] のシステムに相当する）の認証成功率とターンごとの成功率を示してある．また，1 回の認証に要した時間の平均，最短時間，最長時間をそれぞれ「認証時間の平均」，「認証時間の最短値」，「認証時間の最長値」として示した．

表 1　本人認証実験の結果
Table 1　Result of authentication by authorized users.

		実験実施日			
		1 日後		8 日後	
		ターンごとの成功率	認証成功率	ターンごとの成功率	認証成功率
許容する誤差（領域境界からの pixel 数：θ）	0	93.75% (375/400)	78.00% (78/100)	92.75% (371/400)	76.00% (76/100)
	10	98.25% (393/400)	93.00% (93/100)	97.00% (388/400)	90.00% (90/100)
	15	98.25% (393/400)	93.00% (93/100)	97.25% (389/400)	91.00% (91/100)
	20	99.00% (396/400)	96.00% (96/100)	98.00% (392/400)	94.00% (94/100)
画像の選択成功率		100% (400/400)	100% (100/100)	99.75% (399/400)	99% (99/100)
認証時間の平均[sec]		39.29		37.51	
認証時間の最短値[sec]		17.41		16.14	
認証時間の最長値[sec]		129.00		240.64	

θ が大きくなるほど，本人拒否率は低下する（正規ユーザにとっては，おおよその場所をクリックすれば認証に成功する）一方で，他人受入率（どれがパス画像かを推定できた攻撃者が，パス画像をランダムにクリックした場合であっても認証に成功する可能性が高まる）が増加する．よって，パス部位を含む登録部位の面積は重要なセキュリティパラメータとなる．よって，**表 2** に θ の値に対する登録部位 1 つあたりの平均面積を示した．

d)　考察

キューによって暗示されたパス部位を正しく認識し，パス画像上のパス部位を正確に選択できた割合は，1 日後，8 日後とも 90%以上であり，提案システムの有効性が確認された．

表 2　本人認証実験における登録部位 1 つあたりの面積
Table 2　Mean area of each position registered.

		許容する誤差（領域境界からの pixel 数：θ）							
		θ =0		θ =10		θ =15		θ =20	
		パス画像	参照画像	パス画像	参照画像	パス画像	参照画像	パス画像	参照画像
部位の左右，ならびに，何番目の動物の部位であるかを区別した場合	面積の平均 [pixel^2]	2341.49	2089.61	4682.18	4239.15	6177.93	5618.82	7875.82	7190.45
	面積の標準偏差 [pixel^2]	2359.59	1989.67	3455.03	2947.70	4035.46	3450.68	4632.20	3950.12
	画像面積に対する割合[%]	2.60	2.32	5.20	4.71	6.86	6.24	8.75	7.99
部位の左右の区別や，何番目の動物の部位であるかといった区別は行わない場合	面積の平均 [pixel^2]	3459.50	2855.80	6917.82	5793.50	9127.76	7679.06	11636.34	9826.94
	面積の標準偏差 [pixel^2]	3757.49	2799.88	5832.79	4415.29	7023.27	5335.14	8308.71	6299.90
	画像面積に対する割合[%]	3.84	3.17	7.69	6.44	10.14	8.53	12.93	10.92

ただし，パス画像の選択だけであれば，両日ともほぼ 100%であるため，既存の画像認識妨害型認証方式と比べ，正規ユーザの認証時の負荷が若干増大したことが分かる．また，表 1 と表 2 より，確かに，θ が大きくなるほど（パス部位として許容される領域の面積が大きくなるほど）本人拒否率が低下していることが確認できる．

認証負荷増大の問題を今後解決していくために，本実験における失敗の傾向について分析を行った．その結果を**表 3** に示す．表 3 は，本人認証実験のターンごとにおける選択失敗のすべてを，以下に示す失敗のケースごと（(A)〜(E)）に分類し，各々の発生頻度を示したものである．

(A) パス部位の領域境界から 10 pixel 以上の離れた位置を選択した．
(B) 参照画像上のパス部位の認識を間違えた．
(C) 登録されていない部位を選択した．

表 3　失敗の傾向
Table 3　Varieties of authentication failure and their frequency.

		1 日後の失敗の傾向	8 日後の失敗の傾向
失敗の傾向	A	3	10
	B	1	0
	C	2	0
	D	0	1
	E	1	1

(D) パス画像そのものを忘れた．
(E) うっかりミス．

(A) に分類された失敗は，パス画像の大まかな構図はスキーマを使って認識できるものの，時間が経過するにつれて当該部位から少し離れた別の場所が対応部位として認識されるようになってしまったことが原因である．本実験では認証可否についてのフィードバックを被験者に与えなかったが，実運用では，認証に成功した際に必要に応じて登録部位の位置を再度確認させるなどの対策が考えられる．

(B) に分類された失敗は，参照画像の記憶が曖昧であったために起こったと考えられる．特に，「目」と「耳」など互いに近くに配置されやすい部位がキューとして提示された場合に，パス部位が「目」なのか「耳」なのかを被験者が混乱している傾向があった．画像によっては顔の部位が非常に密集している可能性がある．部位の密集具合に応じて密集している部位を 1 つの大きな部位に置き換えるなどの工夫が必要である（たとえば，目，鼻，口の 3 つの部位を顔という 1 つの部位に置き換える）．

(C) に分類された失敗は，部位情報を登録する際に（実験実施者が前もって登録しておいた部位の中で）被験者自身が認識しにくいと感じた部位を登録から除外したにもかかわらず，認証時には除外した部位を認識することができてしまい，部位の選択に混乱してしまったために起きた．この問題に対しても，被験者への認証可否のフィードバックや登録部位の再度確認などの対策が有効であろう．

(D) に分類された失敗は，従来の画像認識妨害型認証方式においても共通の問題ではあるが，いかに不鮮明化画像を効率良く記憶してもらうかといった工夫が必要である．

なお，表 1 において，8 日目の「認証時間の最長値」が 240.64 秒と非常に長くなっているのは．この被験者が認証時にパス画像を見つけることができず，考え込んでしまったため

である．

(E) に分類された失敗は，実験後に当該被験者から，キューの認識は正確にできていたのだが，パス画像中の部位を選択する際に，うっかり間違えて別の部位を選択してしまったと報告を受けたものである．

4.2　覗き見攻撃実験

a)　実験の目的

攻撃者が過去に覗き見したキュー（参照画像上のパス部位）に対するレスポンス（正規ユーザがクリックした画像とその位置）の情報を用い，現在表示されているキュー（パス部位）に対するレスポンスを推測することが難しいかどうかを確認するために，認証試行を攻撃者に複数回覗き見されたことを想定した覗き見攻撃実験を行った．

なお，提案方式においては，攻撃者がなりすましを行うにあたっては，(i) 9 択の認証ウインドウの中からパス画像を発見したうえで，(ii) その中のパス部位を回答する必要がある．このうち，(i) に関する攻撃実験については，正規ユーザの肩越しからの覗き見により，パス画像を特定する攻撃を仮定した実験が先行研究[8] で行われている．そこで，本論文では (ii) に関する攻撃実験のみを行う．すなわち，パス画像についてはすでに不正者によって特定されてしまっていることを仮定し，そのうえで，パス部位の推測に関する攻撃成功率を測定する．

b)　実験方法

本実験では，パス画像と参照画像のペア（2 枚 1 組）を被験者に提示する．攻撃者が過去に正規ユーザの認証試行を n 回（$n = 1$，2，3）覗き見たことを想定し，パス画像および参照画像上の対応する部位を任意に n カ所選び，それぞれの画像上に丸印でプロットする．

パス画像と参照画像とで対応している部位どうしは同じ色でプロットされ，n カ所の部位は互いに異なる色（赤色以外）でプロットされる[⋆1]．また，参照画像には，現在の認証に対するキュー（パス部位）が赤い丸印でプロットされる．攻撃者は，パス画像の中のパス部位（参照画像の赤丸の部位に相当する部位）を推測する．ここで，すべての丸印は，今回の認証システムに合わせ，部位の重心の位置に半径 3 pixel でプロットした．攻撃実験の画面の例を**図 13** に示す．なお，図 13 に付されている各部位の名称は，説明を分かりやすくするために記したものであり，実験の際に被験者には提示されない．

⋆1 たとえば参照画像には左目しか写っていない場合などには，パス画像中の 2 つの部位（左目と右目）が参照画像中のパス部位（左目）と対応する．その場合，同じ部位はすべて同じ色で表示される．

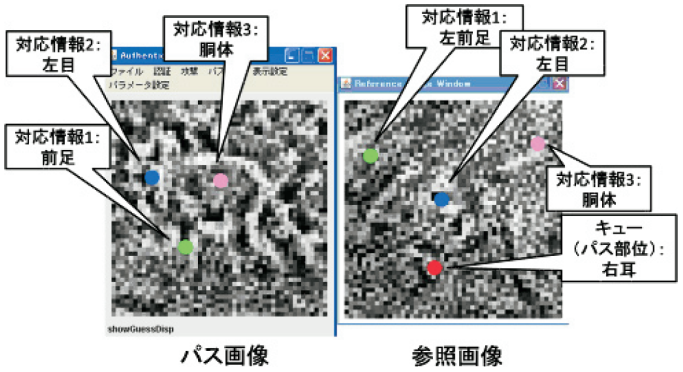

図 13　攻撃実験（過去の覗き見回数が 3 回）の例[⋆2]
Fig. 13　Example of the window for observing attack experiment.

本実験では，4.1 節の本人認証実験のシステムにあわせ，4 枚のパス画像と 1 枚の参照画像を 1 組のパス画像セットとして，攻撃を行った．すなわち，1 組のパス画像セットには，パス画像と参照画像のペアが 4 組（ただし，4 組のペアの参照画像は同一）含まれることになる．今回は，計 5 組のパス画像セットを用意した．5 組のパス画像セットの中には同じ画像は含まれていない．4.1 節の被験者とは別の被験者 1 名に正規ユーザ役を引き受けてもらい，本実験で用いるすべてのパス画像と参照画像（計 25 枚）に対して，4.1 節の認証実験のときと同じ方法で部位を登録してもらった．本実験における攻撃者役の被験者は，4.1 節の被験者 10 名と同じである．

以下に詳細な実験手順を示す．ここで，パス画像セット i に含まれる j 番目のパス画像を P(i, j) (i = 1〜5，j = 1〜4) とし，パス画像セット i の参照画像を R(i) と記す．

1)　実験システムは，画像セット 1〜5 の中から 1 つの画像セット x をランダムに選ぶ．
2)　実験システムは，画像セット x に含まれる 4 枚のパス画像の中から 1 枚のパス画像 P(x, y) をランダムに選ぶ．
3)　実験システムは，パス画像 P(x, y) に含まれる部位の中から 1 つのパス部位 z1 をランダムに選ぶ．また，参照画像 R(x) に含まれる部位の中から z1 に対応するパス部位を探し，その部位の重心に赤い丸を記す．

⋆2 図の大きさの都合上，丸印は実際のサイズより相対的に大きく記してある．

4) 実験システムは，攻撃者役の被験者に P(x, y) と R(x) の組を提示する．参照画像 R(x) のパス部位の上には赤丸がプロットされている．
5) 被験者は，参照画像 R(x) 上の赤丸（パス部位）をキューとし，パス画像 P(x, y) の中から対応するパス部位を推測し，マウスのクリックによりその位置を回答する．パス画像 P(x, y) 上のパス部位 z1 をクリックできた場合には，覗き見成功と判定される．
6) 実験システムは，パス画像 P(x, y) の中に含まれる部位の中から z1 以外のパス部位 z2 をランダムに選び，その部位の重心に緑の丸を記す．また，参照画像 R(x) に含まれる部位の中から z2 に対応するパス部位を探し，その部位の重心に緑の丸を記す．
7) 実験システムは，攻撃者役の被験者に P(x, y) と R(x) の組を提示する．参照画像 R(x) のパス部位 z1 の上には赤丸がプロットされている．パス画像 P(x, y) と参照画像 R(x) の部位 z2 の上には緑丸がプロットされている．
8) 被験者は，参照画像 R(x) 上の赤丸（パス部位）をキューとし，パス画像 P(x, y) の中から対応するパス部位を推測し，マウスのクリックによりその位置を回答する．被験者はパス画像 P(x, y) と参照画像 R(x) の緑丸の対応を，パス部位の推測に利用できる．パス画像 P(x, y) 上のパス部位 z1 をクリックできた場合には，覗き見成功と判定される．
9) z1 および z2 以外のパス部位 z3 を追加し，6)–9) と同様の手順で攻撃実験を行う．部位 z3 の上には青丸が記される．
10) さらに，z1，z2，z3 以外のパス部位 z4 を追加し，同様の攻撃実験を行う．部位 z4 の上にはピンクの丸が記される．この時点で被験者に提示される画像例が図 13 である．
11) y を変え，3)〜10) の攻撃実験を繰り返す．ただし，1 度使用した y は選ばれない．3)–10) を 4 度繰り返した時点でパス画像セット x のパス画像 P(x, y) が使い尽くされる．
12) x を変え，2)〜11) の攻撃実験を繰り返す．ただし，1 度使用した x は選ばれない．2)–11) を 5 度繰り返した時点で 5 種類のパス画像セットが使い尽され，すべての攻撃実験が終了する．

c) 実験結果

実験結果を**表 4** に示す．表中，「成功率」は，各被験者がパス画像セット 5 組分の攻撃を行った攻撃試行全体の成功率（参照画像中のキューからパス画像中のパス部位を正しく選択できた割合）を表示された覗き見情報の数（n）ごとに示したものである．なお今回は，4.1 節の実験と同様，参照画像のキューによって指示されたパス部位が「右前足」であった場合，パス画像における「右前足」と「左前足」のどちらを選択しても正答とした．また，1 枚の画像に複数の動物が写っている場合には，何番目の動物の「右前足」と「左前足」を選択しても正答とした．表 4 においても，領域の境界から θ pixel（$\theta = \{0, 10, 15, 20\}$）広げたときの成功率をそれぞれ示している．

表 4 覗き見攻撃実験の結果
Table 4 Result of experiment for observing attack.

		覗き見情報の数			
		0 個	1 個	2 個	3 個
成功率（領域境界から θ pixel までを境界範囲としたときの成功率）	θ =0	24.0%	29%	33.0%	33.0%
	θ =10	31.0%	40%	44.0%	43.0%
	θ =15	38.0%	45.0%	52.0%	51.0%
	θ =20	44.0%	51.0%	57.0%	57.0%

d) 考察

覗き見回数（n）が増加するにつれて攻撃成功率も増加していることが見てとれる．本人認証の実験にて設定した $\theta = 10$ [pixel] の許容範囲に対して，覗き見情報なし（$n = 0$）では約 3 割，覗き見情報 3 個（$n = 3$）では約 4 割程度の攻撃成功率であった．

先行研究[8] で実施された 9 択 × 4 ターンの認証システムに対する攻撃実験において，攻撃者が覗き見によってパス画像を特定することに成功する確率は約 60%であることが報告されている．よって，「先行研究[8] の攻撃成功率」×「表 4 の攻撃成功率」により求められる提案方式の攻撃成功率は，およそ 20〜25%であると結論付けられる．パス部位という秘密情報が追加されている分，提案方式の覗き見攻撃耐性は既存の画像認識妨害型認証方式よりも当然高くなっている．

しかし，「覗き見情報なし（$n = 0$）」であっても，攻撃者はキュー（パス部位）に対応する場所を 30〜40%の割合で選択することができてしまっている．これは，今回用いた不鮮明化画像の意味が攻撃者にある程度類推可能であったことを意味している．また，実験後のヒアリングの結果，多くの被験者が今回の実験で用いた画像の特徴を活用してパス部位を推定していたことが分かった．すなわち，今回は四足哺乳動物の写真画像を本実験に用いたため，被験者は画像の上部には動物の「頭」，画像の下部には動物の「足」がある可能性が高いといったことや，「目」，「耳」，「鼻」，「口」は比較的お互い近い位置関係にある可能性が高いといったことを仮定することで，パス画像と参照画像の部位を効率的に推測していた．

これらの問題に対応する手段として，画像における一般的な知識（画像の構造：画像の上部に頭があり，下部には足があるなど）を崩したうえで不鮮明化する方法が考えられる．たとえば，歪めた画像を不鮮明化してパス画像や参照画像に用いてやれば，一般的な知識と歪められた画像の構造がマッチせず，攻撃者がパス画像や参照画像の内容を推測することを困難にすることができると考えられる．一方，正規ユーザは登録時に歪められた状態のオリジナル画像を見ることで，歪められたパス画像のスキーマを学習することができ，たとえ歪められていても正しくパス画像を認識することができると考えられる．

5. 検　　討

5.1 利便性について

先行研究の画像認識妨害型認証方式[8]と比べ，提案方式は若干認証時の負荷が増大していた．しかし，実験結果から得られた知見をもとに大半の問題は解決可能であると考えられる．また，提案方式の認証成功率は，登録日から 1 日後では 93%，8 日後では 90%程度であるが，カラー人工画像を使った従来の再認型画像認証方式[1]では 1 週間後の認証成功率が 90%程度（失敗ログインの割合が 10%）であることが報告されていることから，カラーの人工画像を使った方式と比べても提案方式の記憶負荷はそれほど高いものではないと考えられる．

また，Sobrado らの C&R 型画像認証方式[3),4)]では，初心者向けの非常に単純化したシステム[⋆1]を用いた場合に，認証に平均 70 秒強（1 日目：最短時間 = 24.08 [sec]，最長時間 = 150.42 [sec] [⋆2]）の時間を要したことが報告されている．一方，提案方式の本人認証実験で要した認証時間の平均は 40 秒弱（1 日目：最短時間 = 17.41 [sec]，最長時間 = 129.00 [sec]，8 日目：最短時間 = 16.14 [sec]，最長時間 = 240.64 [sec] [⋆3]）である．Sobrado らの方式においてユーザビリティ実験が 1 日目しか行われていない点，および，平均認証時間などに鑑みると，提案方式の認識負荷および作業負荷は Sobrado らの方式に比べ低いと考えることができる．

⋆1 記憶している 5 個の pass-object の中から 3〜5 個の pass-object がランダムに選択され，囮のアイコンと一緒に認証画面に表示される．認証画面には pass-object を含めて 43〜112 体のアイコンが表示される．1 回の認証における問答の繰返しは 5 回である．

⋆2 ただし認証に非常に時間を要した 1 名の被験者をアウトライヤとして実験結果から除いている．アウトライヤになった被験者は，2 番目に認証時間が長かった被験者の，2 倍以上認証に時間を要したと報告されている．

⋆3 8 日目の認証時間の最大値が長くなったのは，4.1 節 d) に示したように，1 名の被験者が認証時にパス画像を忘れてしまったためである．8 日目において 2 番目に長い認証時間は 112 秒であった．

5.2 安全性について

a) 覗き見攻撃

提案方式においては，攻撃者がなりすましを行うにあたっては，(i) 9 択の認証ウインドウの中からパス画像を発見したうえで，(ii) その中のパス部位を回答する必要がある．よって，提案方式の攻撃成功率は，「先行研究[8]の攻撃成功率」×「4.2 節の表 4 の攻撃成功率」より求められ，その結果は約 20〜25%であった．パス部位という秘密情報が追加されている分，提案方式の覗き見攻撃耐性は先行研究の画像認識妨害型認証方式よりも当然高くなっている．

Roth らが提案する C&R 型画像認証方式[5]では，0〜9 までの数字を白か黒かの 2 グループに分け，グループ情報を答えることで，攻撃者が一意に暗証番号を推測することを困難にしようとしているが，1 回の認証における一連の作業すべてのスナップショットを撮られると暗証番号が一意に特定されてしまうという問題がある．入力をさらに曖昧化することによって，1 回の認証を覗き見られただけでは暗証番号を一意に特定することを不可能にした改良方式も提案されているが，その場合は総当たり攻撃に対する安全性が低下してしまう．以上より，提案方式は Roth らの C&R 型画像認証方式と比べ，十分な安全性を有しているといえる．

Sobrado らの C&R 型画像認証方式[3),4)]では，3 つ以上の pass-object が構成する凸包内をクリックするという曖昧入力により，人間の目視に対しては高い覗き見攻撃耐性を確保している．しかし，小島らが行ったビデオ撮影を想定した覗き見に対する安全性の評価においては，Sobrado らの認証方式の簡易版プロトタイプシステム（3 体の pass-object と 197 体の囮アイコンを 10×20 に整然に配置し，その中から探し出した pass-object を頂点とした凸包内部を選択するという作業が 1 回の認証行為となる）に対して，ビデオ撮影を用いた 1 回の認証行為の覗き見によって pass-object の候補が約 1/8 に絞られること，および，約 8 回程度の覗き見により pass-object が特定されるという結果が報告されている[16]．よって，提案方式の安全性は Sobrado らの方式と比べても低くはないと考える．

Sasamoto らの方式[13]は，チャレンジそのものを秘密の通信路を介してユーザに渡すことで，ビデオカメラによる複数回の盗撮に対しても高い耐性を実現している．しかし，特殊な装置（触覚デバイス）が必要となることから，提案手法との比較対象からは外す．

b) ランダムクリック攻撃

従来の再認型画像認証方式[1),2)]も提案方式も，複数の画像の中からパス画像を探し出すという点で同じである．簡単のために N 枚の画像の中から 1 枚のパス画像を選ぶ形の認証方

式を想定すると，攻撃者が当て推量（ランダムクリック）によりパス画像を選択する確率は $1/N$ である．提案方式では，$1/N$ の確率でパス画像を見つけた後に，さらにパス部位を選択する必要があるため，その分，ランダムクリック攻撃の耐性が向上しているといえる．ただし，4 章で行った実験においては，ユーザの部位選択における曖昧性を吸収するために，パス部位の境界から θ pixel 離れた領域内をクリックできれば認証成功と判定している．この場合，θ が大きくなるほどランダムクリックによる他人受入れが増加することになる．同様に，本実験においては，パス部位における左右の区別や何番目の動物であるかといった区別は行っておらず，これに関してもランダムクリックによる他人受入れを増加させる原因となっている．

Man らの報告[17] によると，Sobrado らの方式は画面の中心をクリックすれば非常に高い確率で正規ユーザになりすますことが可能であることが示されている．著者らもこれを確かめるために，Sobrado らの認証方式を単純なモデルに置き換え[⋆1]，プログラムによりシミュレーションを行った結果，認証画面の中心をクリックすることで 3 割強の確率[⋆2]で pass-object が構成する凸包内部が選択される結果となることが分かった．これはすなわち，Sobrado らの C&R 型画像認証方式における凸包内の選択という認証行為 1 回あたりのなりすまし成功率が約 30%であることを意味している．

以上より，ランダムクリック攻撃に対しては，提案手法は従来の再認型画像認証方式や Sobrado らの C&R 型画像認証方式と同程度の安全性を有していると考えられる．他人受入れの問題に関しては，今後の課題として検討していく予定である．

c)　Intersection 攻撃および Exhaustive 攻撃

提案方式においても Intersection 攻撃および Exhaustive 攻撃の脅威は大きい．これらの攻撃は，画像認証全般に共通する問題である．ただし，提案方式では不鮮明化画像を利用しているため，オリジナル画像を用いている画像認証と比べ，目視による Intersection 攻撃および Exhaustive 攻撃についてはその耐性が向上していると考えられる．これらの攻撃に対しては，今後も対応を検討していく予定である．

⋆1 pass-object の数は 3 体に固定し，認証画面には pass-object を含め 66〜100 体のアイコンがランダムな位置に整然と配置される．利便性と安全性を考慮すると凸包が小さすぎても大きすぎても問題となるため，認証画面の面積を S としたときに凸包の面積 T が $S/27 < T < S/3$ となるようランダムな位置に pass-object を配置している．

⋆2 pass-object を認証画面にランダムに配置する試行を 10,000 回繰り返し，そのうち，pass-object により構成される凸包が認証画面の中心点を内包する割合を求めた．

d)　Educated Guess 攻撃

提案方式では不鮮明化画像を利用しているため，攻撃者が画像の意味を推測することは難しいと考えられる．それゆえ，提案方式における Educated Guess 攻撃の脅威は，オリジナル画像をそのまま用いるタイプの画像認証方式に比べ低いと考えられる．

ただし，4.2 節の実験からは，不鮮明化画像の意味が攻撃者にある程度類推可能であるという結果が得られている．この問題に対しては，不鮮明化アルゴリズムを改良したり，画像を歪めたりすることによって画像の構造を崩したうえで不鮮明化するなどの方策を検討する必要がある．

e)　部位に関する攻撃

提案方式ではパス部位がキューとして暗示的に提示されるため，この情報が攻撃に利用される可能性がある．たとえば，攻撃者が参照画像におけるパス部位（キュー）の意味を類推することができた場合，対応する部位を含んでいないと推測される画像は囮画像であると判断することができる．また，攻撃者が何らかの方法でパス画像の特定には成功している場合には，参照画像におけるパス部位（キュー）の意味を類推することができれば，その分，パス画像の中からパス部位を推測する作業は容易となるだろう．

不鮮明化画像の特徴を活用して，部位の位置を類推するという攻撃も考えられる．たとえば，不鮮明化画像中に特徴的なエッジが認識できる場所は部位として登録されている可能性が高いかもしれない．

以上の問題は不鮮明化画像に起因するものであるため，不鮮明化処理の改善が対策の鍵となると考えている．

また，今回の実験で用いた認証システムでは，参照画像上に表示されるパス部位（キュー）は，部位の重心の位置に半径 3 pixel の赤い丸でプロットされるという方式となっている．このため，キューのバリエーションは部位情報の登録個数と等しい値となる．よって，攻撃者が認証行為の覗き見を繰り返せば，すべてのキューとそれに対するレスポンスを収集することができてしまう．

この問題に対しては，「認証が繰り返されるうちに同じ部位を再びパス部位として使用することになった際には，パス部位（キュー）を示す赤丸の位置を変更する」という対策が考えられる．パス部位（キュー）を示す赤丸は，正規ユーザが見た際にパス部位の領域を特定することができる位置であれば，パス部位の重心以外の位置に表示してもかまわない．このような簡易な改良によって，キューのバリエーションを部位情報の登録個数よりもある程度増やすことは可能であるだろう．

6. おわりに

本論文では一見すると無意味に見える不鮮明な画像の特徴を活用することで，Q&R という新しいコンセプトに基づく C&R 型の画像認証方式を提案した．提案方式は，正規ユーザであれば直感的な処理によってキューに対するレスポンスを生成することができ，かつ，攻撃者による有限回の覗き見に対する耐性を持つ，という 2 つの特長を有する画像認証方式となっている．基礎実験を行い，提案方式の有効性を評価した．

提案方式は，既存の C&R 型画像認証方式[3)–5)] および画像認識妨害型認証方式[8)] と比べ，利便性および安全性の面で有望な結果を示している．ただし，提案方式においても，多数の覗き見に対しては十分な耐性を有していない．多数回の覗き見に対しても十分な耐性を持つよう，今回の実験結果および考察から得られた知見をもとに改善を行っていく予定である．

謝辞　本研究は科研費（No.20-6290）の研究助成を受けている．また，本研究は一部，（財）セコム科学技術振興財団の研究助成を受けている．

参考文献

1) Dhamija, R. and Perrig, A.: Deja Vu: A User Study Using Images for Authentication, *Proc. 9th USENIX Security Symposium*, pp.45–58 (2002).
2) 高田哲司，小池英樹：あわせ絵：画像登録と利用通知を用いた正候補選択方式による画像認証方式の強化法，情報処理学会論文誌，Vol.44, No.8, pp.2002–2012 (2002).
3) Sobrado, L. and Birget, C.J.: Graphical passwords, The Rutgers Scholar, *An Electronic Bulletin for Undergraduate Research*, Vol.4 (2002). http://rutgersscholar.rutgers.edu/volume04/sobrbirg/sobrbirg.htm（2009 年 4 月確認）
4) Wiedenbeck, S., Waters, J., Sobrado, L. and Birget, C.J.: Design and evaluation of a shoulder-surfing resistant graphical password scheme, *Proc. Working Conference on Advanced Visual Interfaces* (*AVI'06*), pp.177–184 (2006).
5) Roth, V., Fischer, K. and Freidinger, R.: A PIN entry method resilient against shoulder surfing, *Proc. 11th ACM Conference on Computer and Communications Security* (*CCS'04*), pp.236–245 (2004).
6) 徐　強，西垣正勝：ニーモニックに基づくワンタイムパスワード型画像認証の実現可能性に関する検討，情報処理学会研究報告，2006-CSEC-32, pp.317–322 (2006).
7) 高田哲司：fakePointer：映像記録による覗き見攻撃にも安全な認証手法，情報処理学会論文誌，Vol.49, No.9, pp.3051–3061 (2008).
8) 原田篤史，漁田武雄，水野忠則，西垣正勝：画像記憶のスキーマを利用したユーザ認証システム，情報処理学会論文誌，Vol.46, No.8, pp.1997–2013 (2005).
9) RSA Security Inc.: RSA SecurID. http://www.rsa.com/node.aspx?id=1158（2009 年 4 月確認）
10) Brewer, F.W.: Schemata, MIT Encyclopedia of the Cognitive Sciences, Wilson, R.A. and Keil, F.C. (Eds.), pp.729–730 (1999).
11) 太田信夫，多鹿秀継（編著）：記憶研究の最前線，北大路書房 (2001).
12) 松川順子：ランダム図形の命名作用と再認，心理学研究，Vol.54, pp.62–65 (1983).
13) Sasamoto, H., Christin, N. and Hayashi, E.: Undercover: Authentication Usable in Front of Prying Eyes, *Proc. 26th Annual SIGCHI Conference on Human Factors in Computing Systems* (*CHI'08*), pp.183–192 (2008).
14) 山本　匠，原田篤史，漁田武雄，西垣正勝：画像記憶のスキーマを利用した認証方式の拡張—手掛かりつき再認方式，情報処理学会研究報告，2006-CSEC-34, pp.411–418 (2006).
15) K.T. スペアー，S.W. レムクール（著），苧阪直行（訳）：視覚の情報処理—〈見ること〉のソフトウエア，サイエンス社 (1986).
16) 小島悠子，山本　匠，西垣正勝：覗き見攻撃耐性と利便性を有する画像認証方式に関する一検討，情報処理学会研究報告，2009-CSEC-44, pp.91–96 (2009).
17) Man, S., Hong, D. and Matthews, M.: A shoulder-surfing resistant graphical password scheme — WIW, *Proc. International Conference on Security and Management* (*SAM'03*), pp.105–111 (2003).

(平成 20 年 12 月 1 日受付)
(平成 21 年　6 月 4 日採録)

山本　　匠（学生会員）

2006 年静岡大学情報学部情報科学科卒業．2007 年 9 月同大学大学院修士課程修了．現在，同創造科学技術大学院博士課程，日本学術振興会特別研究員（DC）．情報セキュリティに関する研究に従事．

漁田　武雄

1950 年生．1976 年広島大学大学院教育学研究科博士課程後期中退．同年広島大学教育学部助手．1988 年国立特殊教育総合研究所研究員．1982 年静岡大学教養部講師．現在，静岡大学情報学部情報社会学科教授．文学博士．人間の記憶の文脈依存機構の解明に関する研究に従事．著書等としては『目撃証言と文脈依存記憶』（現代のエスプリ 350，目撃者の証言：法律と心理学の架け橋，至文堂）等がある．日本心理学会，日本認知心理学会，日本基礎心理学会各会員，アメリカ心理学会国際会員．

西垣　正勝（正会員）

1990 年静岡大学工学部光電機械工学科卒業．1992 年同大学院修士課程修了．1995 年同博士課程修了．日本学術振興会特別研究員（PD）を経て，1996 年静岡大学情報学部助手．1999 年同講師，2001 年同助教授．2006 年より同創造科学技術大学院助教授．2007 年より准教授．博士（工学）．情報セキュリティ，ニューラルネットワーク，回路シミュレーション等に関する研究に従事．