社団法人 情報処理学会 研究報告
IPSJ SIG Technical Reports

2007－ＣＳＥＣ－38（29）
2007／7／19

不鮮明化画像を利用した画像認証方式が有する特長
－囮画像生成の容易性－

山本 匠[1] 原田 篤史[2] 漁田 武雄[3] 西垣 正勝[4]

[1]静岡大学大学院情報学研究科 〒432-8011 浜松市中区城北 3-5-1
[2]三菱電機株式会社 〒247-0056 鎌倉市大船 5-1-1
[3]静岡大学情報学部 〒432-8011 浜松市中区城北 3-5-1
[4]静岡大学創造科学技術大学院 〒432-8011 浜松市中区城北 3-5-1
E-mail: [1] gs06059@s.inf.shizuoka.ac.jp [2] Harada.Atsushi@dw.MitsubishiElectric.co.jp
[3] isarida@inf.shizuoka.ac.jp [4] nisigaki@inf.shizuoka.ac.jp

あらまし 不鮮明な画像を利用することで覗き見攻撃に耐性のある「画像記憶のスキーマを利用した認証方式」が提案されている．本方式は，不鮮明な画像の特徴を有効に活用し囮画像を自動的に生成することが可能であり，従来の写真等を利用した画像認証方式において大きな問題であった囮画像の用意及びその更新が容易であるという特長を有する．本稿では，囮画像の自動生成法を検討するとともに，自動生成された囮画像を用いた場合の画像認証システムの安全性について基礎実験を行い評価する．

キーワード 画像認証，覗き見攻撃，不鮮明化画像，スキーマ，囮画像

Advantage of User Authentication Using Unclear Images
－Automatic Generation of Decoy Images－

Takumi YAMAMOTO[1] Atsushi HARADA[2] Takeo ISARIDA[3] and Masakatsu NISHIGAKI[4]

[1] Graduate School of Informatics, Shizuoka University, 3-5-1 Johoku, Naka, Hamamatsu, 432-8011, JAPAN
[2] Mitsubishi Electric Corporation 5-1-1 Ofuna Kamakura, 247-0056, JAPAN
[3] Faculty of Informatics, Shizuoka University, 3-5-1 Johoku, Naka, Hamamatsu, 432-8011, JAPAN
[4] Graduate School of Science and Technology, Shizuoka University, 3-5-1 Johoku, Naka, Hamamatsu, 432-8011, JAPAN
E-mail: [1] gs06059@s.inf.shizuoka.ac.jp [2] Harada.Atsushi@dw.MitsubishiElectric.co.jp
[3] isarida@inf.shizuoka.ac.jp [4] nisigaki@inf.shizuoka.ac.jp

Abstract We have recently proposed an user authentication system using "unclear images" as pass-images, in which only the legitimate users can understand their meanings by viewing the original images corresponding to the unclear pass-images. These unclear images are meaningless for illegal users. Hence it is difficult for illegal users to remember the unclear pass-images, even though they observe the legitimate users' authentication trial. This paper reports another advantage of the user authentication system using unclear images; the adaptation of unclear images enables the automatic generation of the decoy images which are displayed along with pass-images in the authentication window. Here, we explore the way of the automatic generation of decoy images.

Keyword image-based authentication, observing attack, unclear image, schema, decoy image

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．しかしながら，画像認証方式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であった．また，パスワード認証ではパスワードの漏洩[3]が問題となるが，画像認証においても依然として，画像の意味や内容を言葉で教えることによって認証を通過するのに十分な情報を他人に伝達することができてしまう．

上記で挙げた問題の解決を目標に，我々は，画像記憶を利用することで正規ユーザの記憶負荷を軽減しつつも，第三者が他人のパス画像を覗き見ても記憶負荷軽減の効果を得ることができず，かつ，正規ユーザが認証情報を他者に伝達することが難しい「スキーマを

利用したユーザ認証方式（以下，基本方式と呼ぶ）」を既に提案している[4]．詳細は2章で述べるが，基本方式は従来の画像認証方式に比べて覗き見攻撃やパス画像の漏洩への耐性向上を果たしており，本人認証に関しても高い認証成功率を維持している．

一方，従来の画像認証方式は覗き見攻撃に脆弱であることに加え，囮画像の用意（その頻繁な更新を含む）が難しいとされていた．適切な囮画像を潤沢に用意することができなければ，認証システムの安全性の低下やユーザの認証時における認識負荷の増大に繋がる可能性がある．

この問題に対して，本稿では，基本方式[4]で用いられている不鮮明化画像の特長に着目し，従来の写真や絵を利用する画像認証方式では実現不可能な方法で，囮画像を用意・生成する方式を提案し，その実現可能性について検討を行う．

2. 画像記憶のスキーマを利用した認証方式

本章では，画像記憶のスキーマを利用した認証方式の基本方式について説明する．詳細については文献[4]を参照されたい．基本方式は，既存の画像認証方式における覗き見攻撃や，正規ユーザによる意図的な認証情報の（言葉による）漏洩の脅威を軽減することを目的としている．

2.1. コンセプト

画像認証方式にとって覗き見攻撃が脅威となるのは，正規ユーザのみならず覗き見攻撃者にとっても画像の記憶は容易であるからである．すなわち，認証画面にパス画像そのものが表示されるため，正規ユーザによる認証時の画像選択を覗き見られると，攻撃者にパス画像を容易に記憶されてしまう．そこで，基本方式では，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，モザイク化等の不鮮明化処理を施した一見無意味な画像（以下，不鮮明化画像）をパス画像として使用する．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[5]．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．

ただし，無意味な画像を記憶することは正規ユーザにとっても困難であるため，正規ユーザにのみ，パス画像の登録時に不鮮明化処理を施す以前の有意味なオリジナル画像を見せ，当該画像に不鮮明化処理を施したパス画像と合わせて記憶してもらうようにする．不鮮明化画像にはオリジナル画像の特徴がある程度残されているため，オリジナル画像を見ることによって，正規ユーザは不鮮明化画像の中にオリジナル画像の持つ意味を見出せるようになる．この結果，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[6]」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．人間は外界から得られる情報を，無意識のうちに，常時スキーマというフィルタを通して認識しており，ひとたび不鮮明化画像に対するスキーマを学習すれば，それ以降に当該不鮮明化画像を見た場合にも，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．

スキーマを認証に利用することで，不鮮明化処理を施したパス画像であっても正規ユーザは容易にこれを記憶でき，一方，スキーマを学習していない覗き見攻撃者には他人のパス画像を記憶することが困難であるという認証方式が実現できる．ここで重要なことは，正規ユーザ以外には，不鮮明化処理を施したパス画像からオリジナル画像の意味が類推できないようにすることである．つまり，他のユーザには当該パス画像に対するスキーマを学習させないようにする必要がある．

2.2. 不鮮明化画像の生成手順

認証に使用する画像について説明する．本方式では，多数の写真画像などの有意味なカラー画像 $I(x,y)$ (以下，オリジナル画像と記す)と，$I(x,y)$ に対してモザイク化などの不鮮明化処理を施した画像 $O(x,y)$ （以下，不鮮明化画像と記す）を使用する．以下に，本方式で採用した不鮮明化処理の手順を簡単に説明する．

【Step0】300×300 ピクセルの 256 色カラー画像 $I(x,y)$ を用意する．

【Step1】$I(x,y)$ をモノトーン化した後，ヒストグラム均一化処理をして，明るさ及びコントラストを調整した画像 $I'(x,y)$ を得る．

【Step2】$I'(x,y)$ に対し，6×6 ピクセルブロック単位でモザイク化処理を行い，画像 $I''(x,y)$ を得る．（各ブロックは，ブロック内の平均輝度で一色にぬりつぶされる．）

【Step3】$I''(x,y)$ のモザイク化処理された各ブロックを 1 画素とみなした画像 $M(k,l)$ (50×50 ピクセル)に対して，二次元 DCT 処理を行う．今回は簡単のため画像全体を 1 ブロックとして DCT を行った．

【Step4】Step3 で得られた DCT 係数の低周波成分及び

中～高周波成分の値にノイズとなるデータを与える．今回のシステムでは，図1におけるグレーの範囲に対応するDCT係数に，-100～100の値をランダムに代入し，DC成分は0とした．その後，IDCT処理によって画像 $M'(k,l)$ を得る．なお，乱数のシードには常に同じ値を設定し，同じ画像に対しては常に同じ不鮮明化画像が作成されるようになっている．

【Step5】$M'(k,l)$ の1画素を6×6サイズのブロックに伸長し，元画像の大きさに戻した後，再びヒストグラム均一化の処理を行って画像 $I'''(x,y)$ を得る．

【Step6】$I'''(x,y)$ に対して，$I''(x,y)$ を重み $w(0 \leq w \leq 1)$ の加重平均によって重ね合わせ処理を行い，画像 $O(x,y)$ を得る．

$$O(x,y) = wI''(x,y) + (1-w)I'''(x,y), \quad \forall(x,y)$$

今回のシステムでは，$w = 0.3$ とした．

Step4におけるDCT係数の操作による画像の劣化の程度には画像ごとに大きな差がでるため，Step4では比較的大きく画像を壊しておき，Step6の処理によってオリジナル画像の特徴を補完してバランスをとっている．Step4において各画像に応じて適切なDCT係数の調整が行えれば，Step6の処理は必要ない．

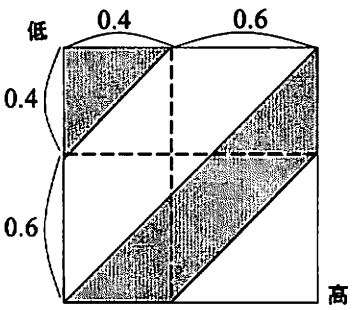


図1　不鮮明化処理におけるDCT係数の変更範囲

上記の手順に従ってオリジナル画像から得られる不鮮明化画像の例を図2に示す．図2左はオリジナルのカラー画像(予稿の印刷上はモノクロ)であり，図2右は不鮮明化処理後の画像である．不鮮明化画像は，オリジナル画像と比較して，モザイク化やDCT係数の操作によって大きく情報量が削減されているが，ある程度の特徴が残されていることが見てとれる．

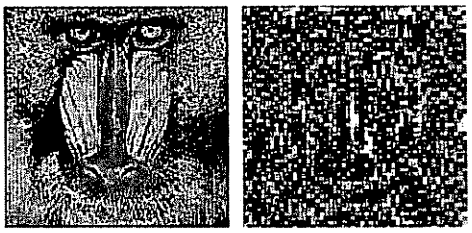


図2　画像の不鮮明化処理

2.3. 基本方式の手順

基本方式の登録・認証の手順を以下に示す．

● 登録フェーズ

【Step1】登録システムは，ランダムに選択した複数枚のオリジナル画像の一覧をユーザに提示する．

【Step2】ユーザは登録システムが提示したオリジナル画像の中から，パス画像にしたい画像を選択する．パス画像にしたい画像がなければ，Step1をやり直す．

【Step3】登録システムは，ユーザが選んだオリジナル画像と，それを不鮮明化した画像をユーザに提示する．ユーザはこれらの画像を納得がいくまで見ることが可能である．

【Step4】ユーザは，Step2の不鮮明化画像をパス画像として登録する．あるいは，当該画像に納得がいかなければ，Step1に戻ってやり直す．

【Step5】Step1～Step4を繰り返し行い，既定の枚数の画像をパス画像として登録する．

● 認証フェーズ

【Step1】認証システムは，ユーザのパス画像をランダムに1枚選び，かつ，ユーザのパス画像以外の不鮮明化画像をランダムに規定の枚数分選択する．そして，それらの画像を，ランダムな順番に並びかえて認証画面に表示する．

【Step2】ユーザは提示された不鮮明化画像の中から自身のパス画像を選ぶ．

【Step3】Step1～Step2を既定の回数だけ繰り返し，そのうちの一定回数以上，ユーザがパス画像の選択に成功すれば認証成功とする．

要求される認証強度に応じて，登録するパス画像の枚数，認証フェーズで提示される不鮮明化画像の枚数，認証フェーズにおける選択の繰り返し回数(ターン数)などが定められる．図3に，9択（囮画像8枚）認証システムの認証画面の例を示す．

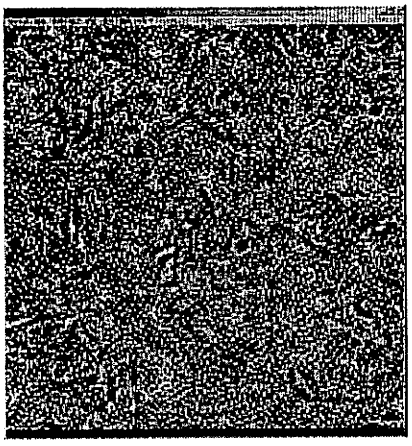


図 3　9 択認証システムにおける認証画面の例

3. 従来の画像認証における囮画像の問題

ユーザが記憶すべきパス画像だけでなく，パス画像を隠すために利用される囮画像（認証画面にパス画像と共に表示される複数の画像）を適切に用意することは画像認証方式において重要な手続きの 1 つである．

例えば，毎回の認証で常に同じ囮画像のセットを利用してしまうと，攻撃者が認証画面中の画像一枚一枚に当たりをつけ，その画像を選択して認証に失敗したならば，その画像はパス画像ではないというように，パス画像の候補が徐々に絞られていく問題（exhaustive-attack）がある．

また，複数の画像の中から自分が記憶したパス画像を選択する再認型の画像認証方式において，過去にパス画像もしくは囮画像として用いた画像を囮画像にしたり，自分が撮影した写真のように自分にとって馴染みの深い画像を囮画像としたりすると，結果的に正規ユーザは認証画面中の全ての画像に対して再認を起こすことになり，パス画像の認識に混乱をきたす可能性があると考えられる．そのため，囮画像にはなるべく正規ユーザが再認を起こさない，つまり，正規ユーザが見たことのない画像を使うことが望ましい．

これらの問題を解決するために，毎回の認証で全ての囮画像を一新する，つまり，一度囮画像として利用した画像は以降の認証では囮画像として利用しないという方法が考えられる．しかしながら，認証の度に必ず表示される画像がパス画像であると容易に推測することができるため問題（intersection-attack）が残る．

以上より，ある一定枚数の囮画像は前回の認証から引き継ぎ，残りの囮画像は正規ユーザが見たことの無い全く新しい画像を用いるという折衷案が適切と考えられるが，残念ながらこの方法にも依然として，どのようにして全く新しい囮画像を追加すればよいのかという問題が残る．あらかじめ大量の囮画像を端末に保存しておいたり，ネットワークを介して自動的にダウンロードしたりという方法も考えられるが，特にストレージや通信量に厳しい制約のある携帯端末等ではこのような対応を採ることは難しい．

このように，従来の画像認証方式においては，囮画像の用意及びその更新に大きな問題を抱えている．そこで本稿では，この問題に対して，基本方式[4]で用いられている不鮮明化画像の特長に着目し，従来の写真や絵を利用する画像認証方式では実現不可能な方法で，囮画像を生成する方式を提案する．

4. 不鮮明化画像の特長を利用した囮画像の生成

4.1. 不鮮明化画像の特長

本節では，文献[4]で用いられている不鮮明化画像の特長を紹介し，その特長を活用した囮画像の用意及びその更新について検討する．

はじめに，図 4 に示されている 3 枚の不鮮明化画像を見てもらいたい．

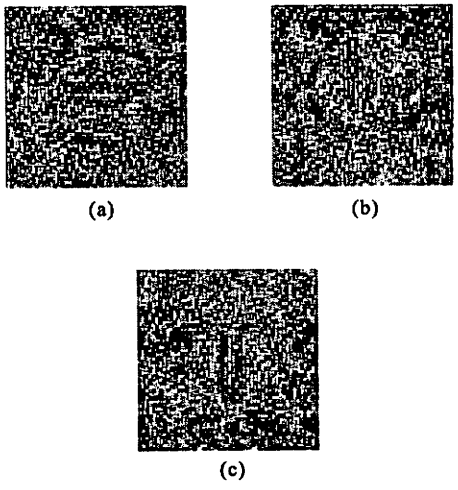


図 4　不鮮明化画像の特徴を利用した不鮮明化画像の加工例

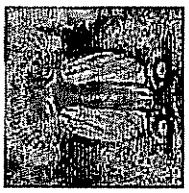
図 4(a)の不鮮明化画像は図 1 の不鮮明化画像を時計回りに 90 度回転（今後特に断りが無い限り，時計回りを回転方向の基準とする）させた画像である．また図 4(b)の不鮮明化画像は図 1 のオリジナル画像と図 5 のオリジナル画像を一定の割合で重ね合わせ，基本方式と同様の不鮮明化処理を施して生成した画像（わかりやすくするために，不鮮明化の度合いを弱めている）である．最後に図 4(c)は，図 1 のオリジナル画像を 16 個のセグメントに分割し，ある規則に従ってセグメン

トを並び替えることにより生成した画像に対し，基本方式の不鮮明化処理と平滑化処理（ぼかし処理）を組み合わせて生成した画像である．

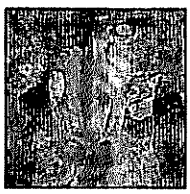
既に2章で図1のオリジナル画像とそれに対応する不鮮明化画像を見ているにも関わらず，図4の3枚の不鮮明化画像の意味（何が映っていて，どのような状態になっているか等）を認識することは難しかったのではないだろうか．

一方，図6に示すように，オリジナル画像で同じことを行った場合，一目見るだけで，その画像の意味（何が映っていて，どのような状態になっているか等）を容易に認識することができるだろう．

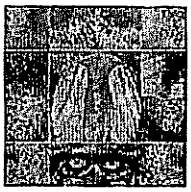
図5　図4(b)の生成にあたって図2のオリジナル画像に重畳されているもう1枚のオリジナル画像



(a)



(b)



(c)

図6　図4(a)～(c)に対するオリジナル画像

このように，不鮮明化画像の場合は，画像を加工（回転させたり，他の不鮮明化画像と重ね合わせたり，セグメントに分割した上で並び替えたり）しても，それが加工された画像なのかどうかを判断することが困難な画像といえる．この特長を活用すれば，加工した不鮮明化画像を囮画像として利用することができると考えられる．一方，図6のように，オリジナル画像に対して同様の加工を行った場合，明らかに加工された画像だと認識できてしまい，これを囮画像として利用することは理に適っていない．

また，図2の画像を知っている者であっても4.1節の説明を聞くまでは図4の画像の意味の認識が難しかったという事実から，本方式を用いれば，正規ユーザが記憶しているパス画像や正規ユーザにとって馴染みの深い画像からでも，正規ユーザが混乱をきたすことの無い囮画像を生成することが可能だと考えられる．また，加工を組み合わせることにより，多種多様な囮画像を容易に生成することが可能であり，従来の画像認証方式における囮画像の用意に関する問題を解決することができると期待される．

4.2. 囮画像の生成手順

本節では，図4に例示した3種類の囮画像の生成手順について説明する．

a)　囮画像1（画像の回転）

【Step0】300×300 ピクセルの 256 色カラー画像 $DI(x,y)$ を用意する．

【Step1】$DI(x,y)$を回転し得た画像を $RDI(x,y)$とする．回転の角度は，90度，180度，270度の3種類である．

【Step2】$RDI(x,y)$をオリジナル画像 $I(x,y)$として2.2節に示した不鮮明化処理を（Step1～Step6）を行い，得られた不鮮明化画像を囮画像1とする．

b)　囮画像2（2枚の画像の重畳化）

【Step0】300×300 ピクセルの 256 色カラー画像 2 枚 $DI_1(x,y)$，$DI_2(x,y)$を用意する．

【Step1】$DI_1(x,y)$に対して$DI_2(x,y)$を重み$w(0 \leq w \leq 1)$の加重平均によって重ね合わせ処理を行い，画像$ODI_{1,2}(x,y)$ を得る．

$$ODI_{1,2}(x,y) = (1-w)\, DI_1(x,y) + w\, DI_2(x,y),\ \forall(x,y)$$

今回のシステムでは，$w = 0.5$ とした．

【Step2】$ODI_{1,2}(x,y)$を回転し得た画像を$RODI_{1,2}(x,y)$とする．回転の角度は，0度（回転無），90度，180度，270度の4種類である．

【Step3】$RODI_{1,2}(x,y)$をオリジナル画像$I(x,y)$として2.2節に示した不鮮明化処理を（Step1～Step6）を行い，得られた不鮮明化画像を囮画像2とする．

c)　囮画像3（画像のセグメントの並び替え）

囮画像3の生成に関する不鮮明化処理は，囮画像1,2における不鮮明化処理と若干異なる．そのため，不鮮明化処理部分についても説明する．

【Step0】 300×300 ピクセルの 256 色カラー画像 $DI(x,y)$を用意する.

【Step1】 $DI(x,y)$を図 7 左のように, 75×75 ピクセル単位のセグメントに分割し, さらに図 7 右のようにセグメントを並び替える.セグメントの並び替えにより, 画像$SDI_1(x,y)$を得る.

【Step2】 $SDI_1(x,y)$をモノトーン化した後, ヒストグラム均一化処理をして,明るさ及びコントラストを調整した画像$SDI_2(x,y)$を得る.

【Step3】 $SDI_2(x,y)$に対し, 6×6 ピクセルブロック単位でモザイク化処理を行い, 画像$SDI_3(x,y)$を得る. (各ブロックは,ブロック内の平均輝度で一色にぬりつぶされる.)

【Step4】 $SDI_3(x,y)$ においてモザイク化処理されている各ブロックを 1 画素とみなした画像$M_1(k,l)$ (50×50 ピクセル)に対して, 図 8 に示すグレーの領域を 5×5 の平滑化フィルタでぼかし, 画像$M_2(k,l)$を得る.

【Step5】 $M_2(k,l)$ の 1 画素を 6×6 サイズのブロックに伸長し, 元画像の大きさに戻し, 画像$SDI_4(x,y)$を得る.

【Step6】 $SDI_4(x,y)$を回転し得た画像を$RSDI_1(x,y)$とする. 回転の角度は, 0 度 (回転無), 90 度, 180 度, 270 度の 4 種類である.

【Step7】 $RSDI_1(x,y)$ のモザイク化処理された各ブロックを 1 画素とみなした画像$M_3(k,l)$ (50×50 ピクセル)に対して, 二次元DCT処理を行う. 今回は簡単のため画像全体を 1 ブロックとしてDCTを行った.

【Step8】 Step3 で得られたDCT係数の低周波成分及び中～高周波成分の値にノイズとなるデータを与える. 今回のシステムでは, 図 1 (2.2 節の不鮮明化処理と同じ領域)におけるグレーの範囲に対応するDCT係数に, -100～100 の値をランダムに代入し, DC成分は 0 とした. その後, IDCT処理によって画像$M_4(k,l)$ を得る. なお, 乱数のシードには常に同じ値を設定し,同じ画像に対しては常に同じ不鮮明化画像が作成されるようになっている.

【Step9】 $M_4(k,l)$ の 1 画素を 6×6 サイズのブロックに伸長し, 元画像の大きさに戻した後, 再びヒストグラム均一化の処理を行って画像$RSDI_2(x,y)$ を得る.

【Step10】 $RSDI_2(x,y)$ に対して, $RSDI_1(x,y)$ を重み$w(0 \leq w \leq 1)$ の加重平均によって重ね合わせ処理を行い, 画像$O(x,y)$ を得る.

$$O(x,y) = w\,RSDI_1(x,y) + (1-w)RSDI_2(x,y), \quad \forall(x,y)$$

今回のシステムでは, $w = 0.3$ とした. 得られた不鮮明化画像を囮画像 3 とする.

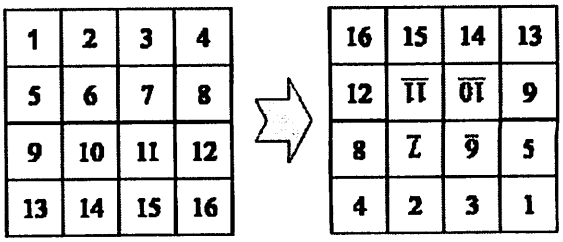


図 7 セグメント化 (左)
及びセグメントの並び替え (右)

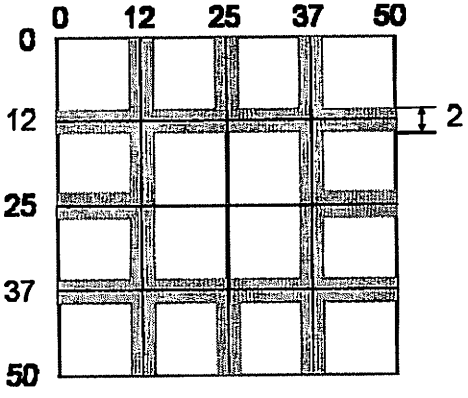


図 8 平滑化処理 (ぼかし処理) を行う領域

次章では, 以上の手順で作成された 3 種類の囮画像を用いた実験を行い, 本囮画像生成方式の実用性を検証する.

5. 評価実験

5.1. 実験1(識別可能性評価)

● **実験の目的**

オリジナル画像を加工 (回転させたり, 他の画像と重ね合わせたり, セグメントに分割した上で並び替えたり)した上で不鮮明化処理を行い生成した囮画像と, 加工を行わずに生成した通常の不鮮明化画像とを比較して, 両者の見分けがつかないかどうかを基礎実験により検証する. 本節では, 囮画像の生成手法ごとに識別可能性評価実験を行った.

● **実験方法**

本実験では, 4.2 節の a)～c)にて説明した方法で生成される囮画像 1, 2, 3 のそれぞれの画像セットに加え, 比較のために, 以下の 2 つ種類の囮画像セットも実験

の対象とした．

a) 囮画像 2'（回転無）

囮画像 2 の生成手順における step2 で，常に 0 度回転を用いて（すなわち回転処理を行わずに）生成した画像

b) 囮画像 3'（回転無）

囮画像 3 の生成手順における step6 で，常に 0 度回転を用いて（すなわち回転処理を行わずに）生成した画像

本実験システムは，被験者に，図 9 に示すような 2 択の認証画面を提示する．2 枚の内，どうちか 1 枚がパス画像（無加工の不鮮明化画像）であり，もう 1 枚が囮画像（囮画像 1，2，3，2'，3'の内のいずれか）である．被験者はパス画像を知らされておらず，2 枚の画像の中で自分が直感的にパス画像だと思うものを選択する．この作業をパス画像と囮画像のペアをランダムに選び直しながら各被験者につき囮画像の種類ごと 5 回ずつ行ってもらった．

2 択システムとした理由は，攻撃者役の被験者にとって非常に有利な条件であっても，本方式で生成した囮画像（加工された不鮮明化画像）とパス画像（加工されていない不鮮明化画像）とを区別することが困難であるかを確認するためである．また，被験者には，各識別実験を始める前に，各実験で用いられる囮画像（囮の不鮮明化画像とその不鮮明化処理をする前の画像）の例をいくつか見せ，各囮画像生成手法に対する知識を深めてもらった．なお，事前説明の際に被験者に例示した囮画像は識別実験には用いられない．

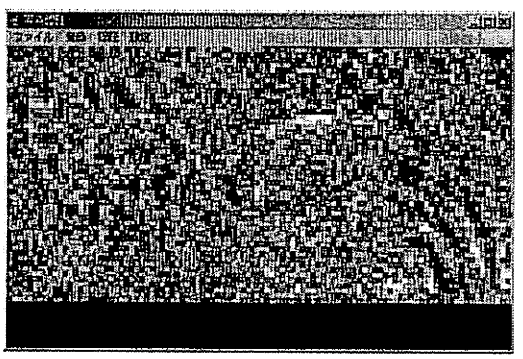


図 9　2 択システムの画面例

被験者は本学情報学部学生 5 名である．各識別試行において，制限時間は設けていない．本稿の実験で使用した画像は，様々な種類の動物が写っている背景つきの写真画像 90 枚である．実験に用いた動物の写真画像はインターネット上で公開されている画像などから収集した．本稿の図中に示した写真画像は，著作者により自由な使用が認められている画像である．

● 実験結果

識別評価実験の実験結果を表 1 に示す．表中，「成功率」は 5 人の各被験者につき 5 回ずつ行った識別試行の全体の成功率（パス画像となる無加工の不鮮明化画像と囮画像となる加工済みの不鮮明化画像とを区別できた割合）を表す．

表 1　各囮画像に対する識別評価実験の結果

	囮画像 1	囮画像 2	囮画像 2' 回転無	囮画像 3	囮画像 3' 回転無
成功率	16/25 (64%)	13/25 (52%)	49/25 (80%)	20/25 (80%)	15/25 (60%)

識別評価実験より，囮画像 2 及び囮画像 3'を使ったときの識別成功率が低くなっていることが見てとれる．しかしながら，2 択の選択画面において，無作為に（つまりどちらがパス画像か判断できない状況で）正しくパス画像を選択できる確率は 50%であることを鑑みると，今回利用した囮画像の中で，最もパス画像と見分けがつきにくかった囮画像 2（識別成功率 52%）及び囮画像 3'（識別成功率 60%）であっても，人間は囮画像とパス画像との間に少なからず何らかの違いを認識しているものと推測できる．また，実験規模も十分なものとは言えず，今後は，今回の識別実験の結果を踏まえ，より規模の大きな実験を通じて，パス画像との見分けが難しい囮画像の検討を行っていく必要がある．

5.2. 実験 2（本人認証評価）

● 実験目的

本方式を用いれば，端末に登録されているパス画像を 4.2 節の手順で加工することで，囮画像として利用することができると考えられる．つまり，認証を行うにあたり，あらかじめ用意しておくものはパス画像だけでよいということになる．

そこで，本実験では，4.2 節の手順によって，本人がスキーマを持っているパス画像から生成された画像を囮画像として用いることによって，正規ユーザが混乱をきたす等の悪影響があるかどうかについて検証する．

● 実験方法

本実験では，図 3 に示すような 9 択の認証フェーズ（認証画面中にパス画像 1 枚と囮画像 8 枚が提示される）を 4 ターン行って 1 回の認証とするシステムを構築した．正規ユーザが記憶すべきパス画像の枚数は 4 枚であり，ターン毎に 4 枚のパス画像の中から 1 枚が

ランダムに重複無く選ばれ認証画面に表示される．本実験システムで利用する囮画像は，4 章で説明した囮画像 1〜3 とした．囮画像 1〜3 の画像セットの中からターン毎に 8 枚がランダムに選ばれ，1 枚のパス画像とともに認証画面に表示される．

被験者は本学情報学部学生 5 名である．パス画像登録後，1 日後と 8 日後に，各被験者につき 5 回ずつ認証を行ってもらう．なおパス画像登録後，被験者はパス画像を確認することはできない．

● **実験結果**

実験結果を表 2 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 ターンのパス画像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 ターンのパス画像選択（9 択の不鮮明化画像の中からパス画像 1 枚を選択するタスク）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．また，ターンごとのパス画像選択にかかった回答時間の平均を「ターンごとの平均回答時間」として記した．

表 2　本人認証の実験結果

認証実施日	1 日後	8 日後
認証成功率	22/25 (88 %)	22/25 (88%)
ターンごとの成功率	97/100 (97%)	97/100 (97%)
ターンごとの平均回答時間(秒)	14.28	13.36

本人認証の実験結果から，たとえ，本人がスキーマを持っているパス画像から囮画像を生成したとしても，正規ユーザ（被験者）は自分のパス画像を高い確率で認識できていることが確認できる．

1 日後と 8 日後における計 6 回の選択ミスについては，実験データを解析してみたところ，全ての選択ミスが，囮画像 2 における 0 度回転（回転無）の画像，もしくは囮画像 3 における 180 度回転の画像を誤って選択したことによるものであった．実験後の聞き取り調査でも，「囮画像 2 における 0 度回転の画像，もしくは囮画像 3 における 180 度回転の画像においては，パス画像を加工することによってこれらを生成した場合には，加工後もパス画像の特徴を残しており，本物のパス画像との区別が難しい」という意見があった．このように，パス画像の特徴が残るような囮画像の生成法を用いた場合には，正規ユーザ（被験者）に混乱を与えてしまう．しかし，囮画像 2 における 0 度回転の画像，および囮画像 3 における 180 度回転の画像以外の囮画像に対しては，このような弊害は確認されていないことから，これら 2 種類の囮画像生成法に関しては使用を控えてやればよいと考えられる．

今後，スキーマを持つ本人に対してはパス画像の認識がより容易になり，スキーマを持たない他人に対してはパス画像の認識がより困難となるような囮画像の生成手法を検討していきたい．

6. おわりに

本稿では，従来の画像認証方式における「囮画像の用意及びその更新が難しい」という問題に対し，不鮮明化画像の特長を活用した囮画像の生成手法によって，その解決を図った．画像認証方式の実現において，あらかじめ十分大量の囮画像を端末に保存しておいたり，ネットワークを介して自動的にダウンロードしたりすることなく，囮画像を用意できることは，非常に画期的である．本稿の実験規模は十分なものではないが，基礎実験により本方式に対する基礎的な知見を得ることができたと考える．

参　考　文　献

[1] Rachna Dhamija, Adrian Perrig, "Deja Vu: A User Study Using Images for Authentication", 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] Richard E. Smith 著，稲村雄 監訳：認証技術 パスワードから公開鍵まで，オーム社，2003 年．

[4] 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.

[5] 太田信夫，多鹿秀継 編著，"記憶研究の最前線"，北大路書房，2001.

[6] W. F. Brewer: Schemata, In R. A. Wilson & F. C.Keil (Eds.), MIT Encyclopedia of the Cognitive Sciences, pp.729-730, 1999.

謝辞　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た．ここに謝意を表する．