

平成 21 年 5 月 28 日現在

研究種目：基盤研究(C)
 研究期間：2006～2008
 課題番号：18500121
 研究課題名（和文） 視覚的注意への概念・空間関係導入に基づく不変・多態認識のための確率競合連想モデル
 研究課題名（英文） Probabilistic Model of Visual Attention and Perceptual Organization for Generic Object Recognition
 研究代表者
 渥美 雅保 (ATSUMI MASAYASU)
 創価大学・工学部・准教授
 研究者番号：00192980

研究成果の概要：視覚的注意とセグメンテーション，及びセグメントの知覚体制化は，人間の物体認識に不可欠な知覚過程であり，コンピュータビジョンシステムの実現においても，特に多くの物体で込み入った実世界画像にどのような物体が含まれるかを不変・多態的に認識する一般物体認識において，重要な基盤処理技術となる．本研究では，顕著性マップ上での並列前注意を，動的に形成されるマルコフ確率場での図地セグメンテーションを介して物体に基づく逐次選択的注意に統合させることを通じて，各概念カテゴリの物体に関する体制化された知覚を生成する一般物体認識のための新しい計算モデルを構築した．そして，本モデルをコンピュータビジョンシステムに実装し，いくつかの実験を通じてその有用性を確認した．

交付額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	1,700,000	0	1,700,000
2007年度	900,000	270,000	1,170,000
2008年度	900,000	270,000	1,170,000
年度			
年度			
総計	3,500,000	540,000	4,040,000

研究分野：知能情報処理

科研費の分科・細目：情報学・知能情報学

キーワード：視覚的注意，顕著性マップ，選択的注意，マルコフ確率場，セグメンテーション，知覚体制化，視覚作業記憶，物体認識

1. 研究開始当初の背景

視覚的注意とセグメンテーションは物体認識に不可欠な知覚過程である．視覚的注意は空間に基づく注意と物体に基づく注意の2つの様式で捉えられるが，空間に基づく注意がセグメンテーションを促進し，セグメンテーションされた物体に基づき注意が喚起されるという点において，注意とセグメンテ

ーションは相互作用する過程である．また，注意過程は前注意過程と焦点注意過程の2つの段階に区別できる．前注意過程では，視野全体から局所的な特徴が並列的に検出され，焦点注意過程では，これら局所の特徴が注意を向けた場所ごとに逐次・継時的に統合されてセグメントが知覚体制化され物体認識処理に渡される．一方，セグメンテーシ

ン、特に図地分離には境界所有割当とそれに続く面充填過程が大きく関与しそれらは注意に先立って遂行され注意を誘導することがわかってきている。また、セグメントの知覚体制化も物体認識に不可欠な過程であるが、知覚体制化が物体の部分から全体への注意を誘導するという知見も示されている。ところで、物体認識の重要な課題の1つである不変・多態認識に関しては多くの取り組みがなされているが十分なレベルに達してはなく、特に一般物体認識におけるそれら課題に関しては、研究ははじめられたばかりである。これら課題に対するアプローチの1つとして、各概念カテゴリに含まれる物体を構成するセグメント間の関係、及び物体とそのコンテキストとの空間的關係を統計的に獲得し認識に用いることが考えられるが、日常的な込み入った情景画像に対してそれら関係を自動的に抽出するためには、物体認識の基盤過程にある注意とセグメンテーション、及び知覚体制化に関して、人間のそれら知覚過程を模倣したモデルの構築が有効で重要であると考えられる。

2. 研究の目的

本研究では、一般物体認識のための視覚的注意とセグメンテーション、及びセグメントの知覚体制化の新しい計算モデルを構築する。本モデルでは、顕著性マップ上での空間的な並列前注意を物体に基づく逐次選択的注意に図地セグメンテーションを介して統合させる。ここで、図地セグメンテーションは前注意点の周りに動的に形成されるマルコフ確率場で遂行される。そして、注意を向けられた物体を構成するセグメント間、及び物体とそのコンテキストを与えるセグメントとの間に知覚体制化を生じさせる。具体的には、本モデルは次のような特徴を持つ。第1に、顕著性マップ上での並列前注意に基づく境界所有割当と動的に形成されるマルコフ確率場での合併を伴う面充填による多重の図地セグメンテーション、第2に、セグメントの注意度のその顕著性、閉領域性、及び注意バイアスに基づく計算と、その注意度に基づくセグメントへの選択的注意、そして、第3に、能動的なセグメント記憶と受動的なセグメント記憶からなる2レベルの視覚作業記憶の維持と、注意が向けられている能動的なセグメントの集まり内での知覚体制化作用である。物体を構成するセグメント間の知覚体制化は、物体の全体部分関係、及びその不変性・多態性に係わる物体の構成的知覚に役立つと考えられる。物体セグメントとその目立った周囲のセグメントとの間の知覚体制化は、ある物体の視覚的な共起コンテキストを与えるものであり、視空間の構成的知

覚に役立つと考えられる。

3. 研究の方法

(1) モデルの構築

本モデルは、画像の特徴空間上で計算される前注意のための顕著性マップ、図地セグメンテーションのための動的に形成されるマルコフ確率場の集まり、図地セグメントの集合を維持し知覚を体制化する視覚作業記憶、及びセグメントの選択的注意の機構から構成される。これら構成要素の確率的な数理モデルを定式化し、それらモデルに基づくコンピュータビジョンシステムを実装する。

(2) 実験による評価

本モデルに基づくコンピュータビジョンシステムの性能を実験により評価する。実験としては、視覚ポップアウト実験、図地反転実験、注意を引くべく設計された物体への注意とセグメンテーション実験、及び画像データベース(Caltech 画像データベース)の複数の概念カテゴリの画像に対する視覚的注意、セグメンテーション、及び知覚体制化の実験を行う。

4. 研究成果

(1) モデルの構築

画像特徴量：前注意、及び図地セグメンテーションのための画像特徴量として、明度・彩度・色相、それらのコントラストと方位成分を求める。明度コントラスト、彩度コントラスト、色相コントラストは、明度、彩度、色相をそれぞれ LoG(Laplacian of a Gaussian) フィルタを用いて畳み込むことにより計算される。ただし、色相の値 $[0, 2\pi)$ はそのカテゴリを表しているため、各点の周囲の点との相対色相値を求めて畳み込みを行う。明度方位成分、彩度方位成分、色相方位成分は、明度、彩度、色相をそれぞれ $0^\circ, 45^\circ, 90^\circ, 135^\circ$ のガボールフィルタを用いて畳み込むことにより、これら4つの方位に関して計算され、各点の方位成分はそれらの最大値として求められる。ただし、色相方位成分に関しては各点の周囲の点との相対色相値を求めて畳み込みを行う。

顕著性マップ：顕著性マップは、明度コントラスト、彩度コントラスト、色相コントラスト、及び方位成分から計算される。明度コントラストは、オン中心-オフ周辺型反応とオフ中心-オン周辺型反応に対応する2つの次元により構成される。彩度コントラストも同様に、オン中心-オフ周辺型反応とオフ中心-オン周辺型反応に対応する2次元で構成される。色相コントラストは、赤、黄、緑、シアン、青、マゼンタの6つの代表色相に対応する6次元で、方位成分は、4方位に対

応する4次元で構成される。各特徴に関する顕著性は、その特徴を構成する次元間で希少性基準と呼ぶ基準に従い競合的に計算される。希少性基準は、同一特徴の異なる次元の領域が画像内にあるとき、より少なく現れる次元の領域がより目立つことを保証する基準である。顕著性 $S(p)$ は、各点 p の明度コントラスト顕著性 $S_{BC}(p)$ 、彩度コントラスト顕著性 $S_{SC}(p)$ 、色相コントラスト顕著性 $S_{HC}(p)$ 、方位顕著性 $S_{OR}(p)$ から、重み w_i ($i=BC, SC, HC, OR$, ここで $w_{BC}+w_{SC}+w_{HC}+w_{OR}=1$)を用いて次のように計算される。

$$S(p) = \sum_{i=BC, SC, HC, OR} (w_i \times S_i(p)) \quad (1)$$

図地セグメンテーション：前注意は顕著性マップ上で顕著性の大きさに基づいて確率的に複数の点を選択する。そして、それら前注意点の周りに原始的な知覚対象を生成する。本研究では、これをプロトセグメントと呼ぶ。プロトセグメントとしては、ある大きさを上限に前注意点との明度と色相に関する類似点の集まりが選択される。前注意過程は境界所有割当を引き起こす。LoG フィルタによるオン中心-オフ周辺型とオフ中心-オン周辺型のコントラストは境界の両側に顕著な点を生成する。一般に、この顕著性は境界が凸領域を形成するとき、凸領域の内側、即ち図領域において大きくなる。これにより、境界の両側のいずれかに前注意が向けられ、かつより高い確率で図領域に対して境界所有割当が起こる。図地セグメンテーションは、プロトセグメントの周りに明度と色相のマルコフ確率場を動的に形成して図地ラベリングにより面充填を行うことにより遂行される。図地ラベリングはEM手続きにより求められ、EステップでのQ関数の計算には平均場近似を用いる。複数のセグメントが合併条件を満たす場合はそれらセグメントは合併される。

セグメントの注意度：セグメントの注意度は、セグメントの顕著性、閉領域性、及び注意バイアスに基づき計算される。セグメントの顕著性は、セグメントが面としてその周囲から目立っている程度とセグメント内に目立ったスポットが存在する程度により定められる。前者を面注意度、後者をスポット注意度と呼ぶ。面注意度は、セグメントの明度と色相の平均特徴量とその周りの地領域の平均特徴量、即ち図地セグメンテーションによる図セグメントと地セグメントの平均特徴量間の距離により求められる。スポット注意度は、セグメント内の各点の顕著性の最大値により求められる。閉領域性は、画像内でセグメントが閉じているかどうか、即ちセグメントが画像の縁にかかっていないかどうかで判定される。画像の縁にかかる点数がある閾値未満のときセグメントは閉じているとみなす。注意バイアスとは、ある特徴を持

った領域への先験または経験的な注意性向を与える制約で、本モデルでは顔らしい領域への注意バイアスを扱う。顔らしさは、本モデルでは、簡単に、セグメントの色相と縦横比で判定している。このとき、セグメント f の注意度 $A(f)$ は次のように求められる。

$$A(f) = (f, \cdot) \times (\alpha \times G(f) + \beta \times P(f) + \gamma \times B(f)) \quad (2)$$

ここで、 $G(f)$ は面注意度、 $P(f)$ はスポット注意度、 $B(f)$ は注意バイアスで、 α, β, γ はそれぞれそれらの重み係数である。また、 (f, \cdot) は f が閉領域のとき1、そうでないときを返す関数で、 γ はセグメントが閉じていないときの注意減少率である。

視覚作業記憶：セグメントは視覚作業記憶に維持される。視覚作業記憶のセグメントのうちある数のセグメントがそれらの注意度に基づき確率的に焦点注意の対象となる。そして、その中で注意度の大きいセグメントに選択的注意が向けられる。焦点注意の対象となるセグメントはリハーサルされる能動的な記憶であり、それ以外は受動的な記憶で受動的記憶は視覚キャッシュに相当するものである。受動的な記憶はある一定時間経過すると視覚作業記憶から消失する。

アルゴリズム：注意とセグメンテーションは概ね次のステップを繰り返すことにより遂行される。

[Step 1] 顕著性マップまたは視覚作業記憶から注意度に基づき確率的に前注意点またはセグメントを選択する。前注意点に対してはプロトセグメントを求める。

[Step 2] プロトセグメントまたはセグメントの周囲に一定のマージンでマルコフ確率場を拡大し図地ラベリングを行うことをある一定回数、またはセグメントが収束するまで繰り返す。また、視覚作業記憶内の受動的記憶セグメントに関しても同様に図地ラベリングを拡大する。複数のセグメントが合併条件を満たす場合は、それらセグメントを合併する。

[Step 3] セグメントの注意度を計算し、視覚作業記憶を更新する。ここで、Step 1で選択されたプロトセグメントまたはセグメントをStep 2で合併を伴って展開したセグメントの集まりが能動的記憶セグメント集合を構成する。

[Step 4] 能動的記憶セグメント集合から注意度の大きいある数のセグメントを選択的注意セグメントとして選択する。

[Step 5] 視覚作業記憶において、選択的注意セグメントとそれに隣接する能動的記憶セグメントのペアを記憶する。

知覚体制化：知覚体制化は、選択的注意が向けられたセグメントとそれと同時に知覚される能動的記憶セグメントとの間の群化

により生ずる．本モデルでは，近接，特に隣接という最も簡単な要因のみを扱うが，他の要因による群化に拡張することも可能である．これにより，1つの物体を構成する複数のセグメントの知覚体制化やある物体セグメントのそのコンテキストとの知覚体制化が達成される．

(2)実験による評価

視覚ポップアウト実験：視覚的注意の基本性能を評価するために行った明度・色相・方位ポップアウト実験の様子を図1(a)に示す．図1(b)はそれらの顕著性マップである．希少性基準により，それぞれ「白丸」「赤丸」「斜め線」に対する顕著性が高くなっていることが観測される．明度ポップアウトに関して，注意選択の連続する25回の繰り返しにおいて，各回の選択的注意数を4とした場合，100注意のうち47の注意が「白丸」に向けられ，25回の繰り返しのうち22回で少なくとも1つの注意が「白丸」に向けられた．これより，1つの「白丸」にほぼ常に注意を集中させながらその周りの9個の「黒丸」にも注意を分散させていることがわかり強いポップアウトが確かめられた．色相，方位に関しても同様にポップアウトが観測された．

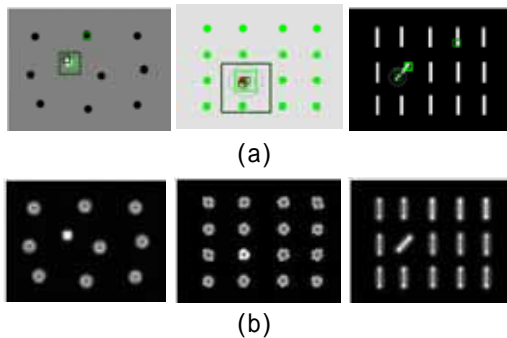


図1: 明度・色相・方位ポップアウト実験

図地反転実験：図2(a)に図地反転実験で用いたルビンの杯の画像を示す．図2(b)はその顕著性マップである．図2(c)，(d)に示すように，それぞれ「杯」に相当する中央の白い図形，「向き合った顔」に相当する周辺の黒い図形が図セグメントとして抽出され，異なる注意で選択された．セグメンテーションでは，中央の白い図形と周辺の黒い図形の境界上の顕著性が高い前注意点からその点と同じ特徴量の面に図セグメントが広がる．周辺の黒い図形が抽出される過程においては，一時的に，図2(e)，(f)に示すような「右向きの顔」と「左向きの顔」に相当するセグメントが展開された．これは，「杯の柄」の辺りの境界の複雑さゆえにその辺りの顕著性が他の境界部より高く，セグメンテーションが「杯の柄」の辺りの前注意点から起こるた

めである．

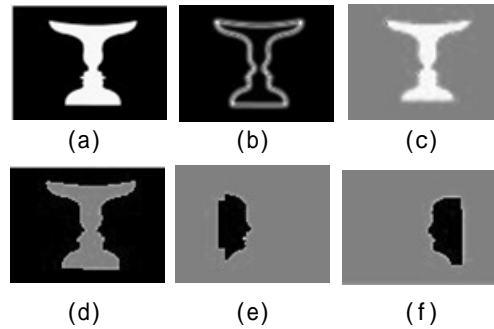


図2: 図地反転実験

注意を引くべく設計された物体への注意とセグメンテーション実験：注意を引くべく設計された物体として非常口サイン，交通標識を含む画像を図3(a)，(c)にそれぞれ示す．非常口サインを含む画像では，図3(b)に示すように非常口サインセグメントに第一に注意が向けられ，また，背景の壁セグメントにも時折注意が向けられた．交通標識を含む画像では，図3(d)に示すように2つの交通標識セグメントに高い頻度で注意が向けられ，また，図3(e)に示すような背景にある樹木，空，建物の壁，石壁，道路のセグメントにも時折注意が向けられた．これらより，注意を引くべく設計された物体のセグメントに高い頻度で注意が向けられることが確かめられた．

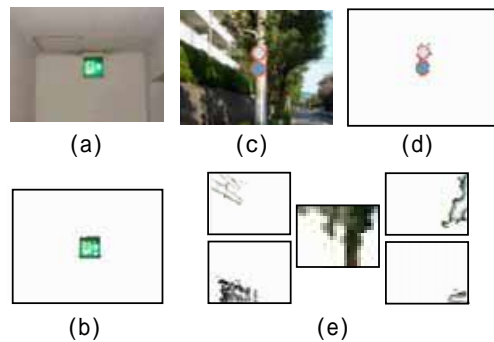


図3: 注意を引くべく設計された物体の実験

Caltech 画像データベースの複数の概念カテゴリの画像に対する視覚的注意，セグメンテーション，及び知覚体制化の実験：Caltech 画像データベースは，指定カテゴリの物体のみを含む物体画像が多いが，日常の背景の中に指定カテゴリの物体が存在するシーン画像を特に選択して，それら概念カテゴリの物体に対する選択的注意とセグメンテーション，及び知覚体制化を評価した．図4(a)に，実験で用いたいくつかの概念カテゴリの物体を含む画像を示す．実験では，10個の概念カテゴリの各々について2枚の画像を用い

た．図 4(b)に，各概念カテゴリの物体に関して抽出されたセグメントを示す．「People」，「School-bus」，「Butterfly」画像の例に見るように物体を構成するセグメントがほぼ正確に抽出されることが確認された．しかし，「Horse」画像の胴体部分のように類似特徴量のセグメントが連結したり，「Iris」画像の花のように複数の同一特徴量セグメントがまとまったりするケースも観測された．注意選択の連続する 15 回の繰り返しにおいて，これら概念カテゴリ物体のセグメントのいずれかが選択的注意セグメントとして選択された回数は，「People」画像で 12 回，「School-bus」画像で 11 回，「Horse」画像で 14 回，「Iris」画像で 11 回，「Butterfly」画像で 9 回であった．実験で用いた全画像に関して，概念カテゴリ物体のセグメントのいずれかに選択的注意が向けられた回数は注意選択の 15 回の繰り返し中に平均で 11.9 回であった．一方，それら以外のセグメントに選択的注意が向けられた回数は平均で 7.2 回であった．これらより，ある概念カテゴリの物体を含む画像という観点で人により注意された物体に対して，本モデルが選択的注意を向けることを確認することができた．また，「People」画像での顔への選択的注意の回数は 15 回の繰り返し中に平均 8 回と人物を構成する他のセグメントと比較しても高く，注意バイアスとして導入した顔らしい物体への注意性向の有効性が確認された．

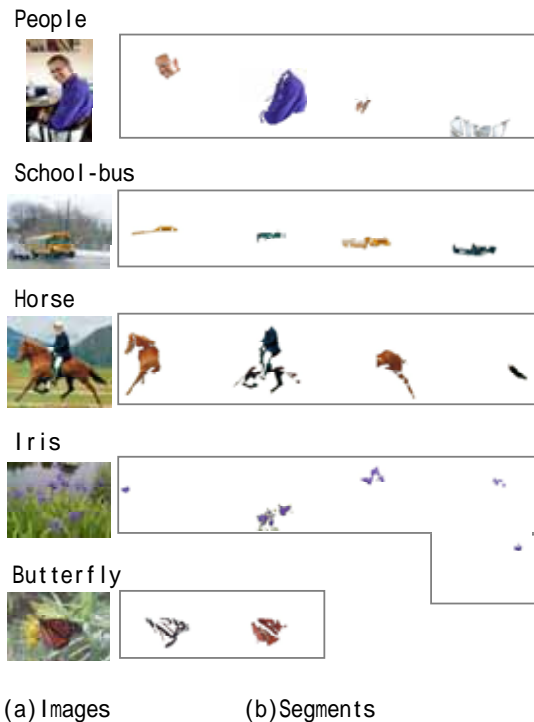


図 4：概念カテゴリ物体セグメントへの注意

図 5 に，図 4 の各概念カテゴリの物体セグメントに関する知覚体制化の例を示す．本モデルにおいて，知覚体制化は選択的注意セグメントとそれに隣接する能動的記憶セグメントの間で生ずるが，「People」カテゴリ，「School-bus」カテゴリ，「Horse」カテゴリ，「Butterfly」カテゴリの例に見るように，概念カテゴリ物体を構成するセグメント間に知覚体制化が多く発生した．これらは，物体の全体部分関係，及びその不変性・多態性に係わる物体の構成的知覚に役立つと考えられる．また，「Horse」カテゴリにおける「馬」の部分と「芝」セグメント，「Iris」カテゴリにおける「花」と「茎・葉」セグメント，「Butterfly」カテゴリにおける「蝶」と「花」や「葉」セグメントといった物体セグメントとその目立った周囲のセグメントとの間にも知覚体制化が観測された．これらは，ある物体の視覚的な共起コンテキストを与えるものであり，視空間の構成的知覚に役立つと考えられる．セグメントの注意過程にこれらセグメントの構成的知覚に関する知識を組み入れてセグメントの注意度を制御することにより体制化された知覚の生成の促進が可能となる．

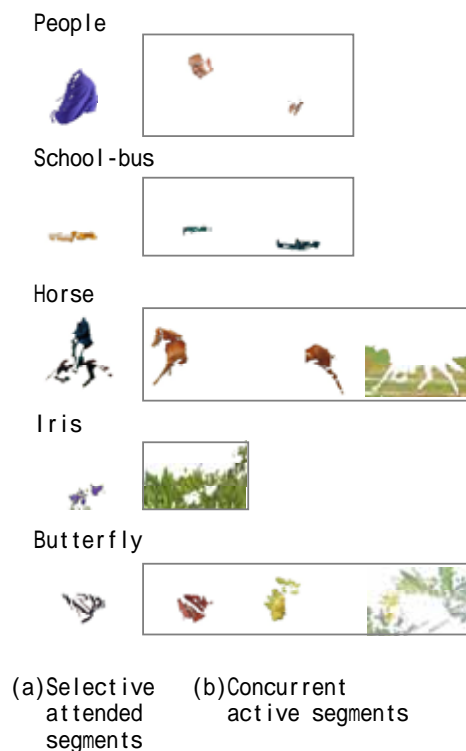


図 5：知覚体制化

このように，Caltech 画像データベースの複数の概念カテゴリの画像を用いた実験を通じて，指定カテゴリの物体またはその部分セグメントに選択的注意が向けられること，注意を向けられた物体を構成するセグメン

ト間またはその周囲のコンテキストとの間に知覚体制化が生ずることを示して、本モデルの有効性を確認した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計4件)

Masayasu Atsumi, Attention-based Segmentation on an Image Pyramid Sequence, (J.Blanc-Talon; S.Bourenane; W.Philips; D.Popescu; P.Scheunders (Eds.)), Advanced Concepts for Intelligent Vision Systems-10th International Conference, ACIVS2008 Proceedings-, Lecture Notes in Computer Science, Vol.5259, pp.625-636, 2008, 査読有.

Masayasu Atsumi, A Dynamical Model of Attention and Segmentation on an Image Pyramid Sequence, Proceedings of Joint 4th International Conference on Soft Computing and Intelligent Systems and 9th International Symposium on Advanced Intelligent Systems (CD-ROM), pp.1255-1260, 2008, 査読有.

Masayasu Atsumi, Stochastic Selective Attention and its Shift on the Visual Attention Pyramid, Proceedings of 14th Korea-Japan Joint Workshop on Frontiers of Computer Vision, pp.463-468, 2008, 査読有.

Masayasu Atsumi, Stochastic Attentional Selection and Shift on the Visual Attention Pyramid, Proceedings of the 5th International Conference on Computer Vision Systems, (CD-ROM), 10p., 2007, 査読有.

[学会発表](計7件)

渥美雅保, 注意に基づき動的に形成されるマルコフ確率場での知覚体制化, 第23回人工知能学会全国大会論文集(CD-ROM), 4p., 2009年6月17-19日, サポートホール高松(香川県高松市).

Masayasu Atsumi, Attention-based Segmentation on an Image Pyramid Sequence, 10th International Conference on Advanced Concepts for Intelligent Vision Systems, October 20-24, 2008, Juan-les-Pins, France.

Masayasu Atsumi, A Dynamical Model of Attention and Segmentation on an Image Pyramid Sequence, Joint 4th International Conference on Soft

Computing and Intelligent Systems and 9th International Symposium on Advanced Intelligent Systems, September 17-21, 2008, Nagoya, Japan.
渥美雅保, 画像ピラミッド系列上での選択的注意に基づくセグメンテーションとグルーピング, 第22回人工知能学会全国大会論文集(CD-ROM), 2J1-3, 4p., 2008年6月11-13日, 旭川市ときわ市民ホール.
Masayasu Atsumi, Stochastic Selective Attention and its Shift on the Visual Attention Pyramid, 14th Korea-Japan Joint Workshop on Frontiers of Computer Vision, January 23-26, 2008, Oita, Japan.

渥美雅保, 視覚注意ピラミッド上での確率的な注意選択とシフト, 第21回人工知能学会全国大会論文集(CD-ROM), 3C0-3, 4p., 2007年6月18-22日, 宮崎市ワールドコンベンションセンターサミット.

Masayasu Atsumi, Stochastic Attentional Selection and Shift on the Visual Attention Pyramid, 5th International Conference on Computer Vision Systems, March 21-24, 2007, Bielefeld, Germany.

6. 研究組織

(1)研究代表者

渥美 雅保 (ATSUMI MASAYASU)

創価大学・工学部・准教授

研究者番号: 00192980

(2)研究分担者

なし

(3)連携研究者

なし