

能動センシングと「センシング情報学」の事始め

出口 光一郎*

* 東北大学大学院情報科学研究科 宮城県仙台市青葉区荒巻字青葉 6-6-01
 * Graduate School of Information sciences, Tohoku University, 6-6-01
 Aoba-campus, Sendai, Miyagi, Japan
 * E-mail: kodedg@fractal.is.tohoku.ac.jp

キーワード：センシング情報学 (sensing informatics), 能動センシング (active sensing), 情報量 (information amount), 相互エントロピー (mutual information amount), 画像認識 (image understanding).
 JL 0009/12/5109-0808 © 2012 SICE

1. はじめに

本稿では、(1) センシング (計測) を「対象の状態に関する情報を獲得する行為」ととらえ、(2) この情報の獲得の場面での不確かさとあいまいさを能動的に減少させるためのセンシング戦略を情報理論の立場から見直し、(3) 一般化した「能動センシングの方法論」の構築を試みる。

ただし、「センシング情報学」と名付けたこのような試みはまだ事始めの段階であり、ここでは考え方のあらましを述べるにとどめている。

センシングでは、物理世界/外界から得られる計測結果としてのデータを、物理法則、すなわち、なにが測定できているかの物理モデルにあてはめることで、測定対象の状態を定量的に理解する。この物理モデルの共通性を異なる対象分野にも見出すことによって、分野を横断した統一的な視点によるセンシングを構成できる可能性をもつ。たとえば、電気分野での抵抗、インピーダンスなどの概念を機械や熱エネルギーの分野の物理量へと一般化することで、モデルの共通性を見出し、統一的なセンシング理論を構築してきた。このような試みは、センサの設計の場面で成功を納めてきている。しかし、測定対象に共通性を見出そうという努力に限定されてもいる。

センシングには、センサの設計と構成に加えて、そのセンサの配置やデータの読み取りとその処理などの能動的なセンシング行為が付随する。この能動センシングの行為は、図 1 のようにモデル化される。能動センシングでは、センシング行為によって獲得されるデータに応じてつぎの時点での行為の能動的な選択がなされる。どのようなセンサを

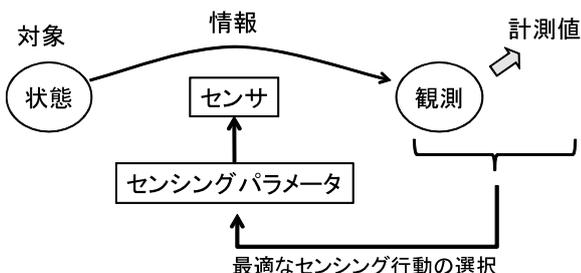


図 1 能動センシングのモデル。センシングを「状態」から「観測」への情報の流れと考える。

選び、それをどのように配置して何を計測するかに加え、ある計測データをつぎの計測や他のセンサによる計測にどう活かすか、たとえばどのようなセンサを組み合わせでつぎにどこを計測するか、といった選択である。このような枠組みは、広域で多様なセンシングを必要とする環境計測などのこれから重要となる大規模センシングへの体系的アプローチにとって、重要なものとなって来ている。

本稿では、「情報を効率よく獲得する」という立場から、情報理論に基づいて、この能動センシングのための統一的な考え方について考察する。特に、センシング行為における情報の流れに着目する。具体的には、センシング行為の良さを対象の状態に関する情報エントロピーの減少量で評価する。すなわち、あるセンシング行為をとると、対象の状態のあいまいさ、不確かさがどれだけ減少するかを評価する。この基準をもとに、「効率良く情報を獲得する行為としての能動センシング」という統一的なセンシング理論を構築する。なお、この着想に基づく物体認識システムの構築を文献 1) に発表しており、文献 2) で解説をしている。

2. センシング情報学の概要と課題

2.1 センシングにおける情報伝達のモデル

まず、センシング情報学の構想を述べるために一番簡略化したセンシングのモデルを示す。図 2 に示すように、計測対象 X は x_1, \dots, x_M の離散的などれかの状態をとるもの

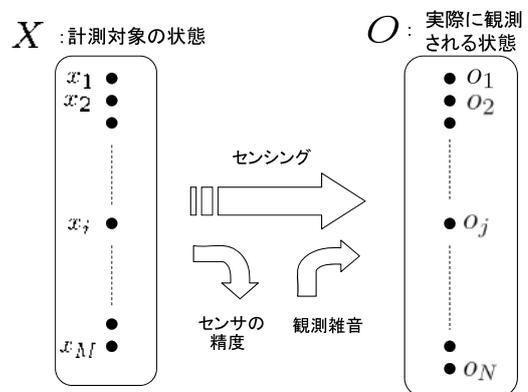


図 2 センシングにおける計測対象の状態から観測への情報の流れ

とする。たとえば上記の物体認識システムでは、対象物体が 1 から M のうちのどれかであることを、この対象の状態としている。一方、センシングでは、その画像から o_1, \dots, o_N の N 個のどれかの値として物体を特定する観測 O を得る。 M と N は同数とは限らない。この認識システムの例では、画像として何が観測されたかを表わす。対象の一部であったりある対象を異なる方向から見た画像であったりする。1 つの対象に対して観測は多数あったり、また一方で、異なる対象が見方によって同じに見えたりする。

温度などの物理的なセンシングでは、 X と O はこのような離散値の組ではないが、このモデルの連続値への拡張は容易であるので、離散的な場合で説明を続けよう。

今、実際に観測される O の値をもとに X の値を推定することが、ここでのセンシングである。ただし、必ずセンサの精度などによる情報の脱落と観測雑音その他による情報の付加がある（合わせて雑音と呼ぶ）。それらは、センシングにおける本質的な不確かさやあいまいさとして存在する。すなわち、確定的な値が得られるわけではない。

ここで注意をしなくてはならないことは、図 2 に示した対象の状態 X も、そもそもが、不確かさとあいまいさをもっていることである。これを克服することはここでの目的ではない。観測 O には、センシングの過程における不確かさとあいまいさが加わる。このセンシング過程での雑音に惑わされないようにすることが、ここでの目的である。

したがって、センシングでは、観測 O を得るごとに、 X のそれぞれの値 x_i について、それがどのような値であるかの確からしさを表わす生起確率分布 $p(x_i)$ (X に関する情報の信頼度と解釈される) を得るとというのが、ここでの基本である。

すなわち、観測ごとに状態に対する「不確かさ」や「あいまいさ」が付きまとう。いずれの場合も確率分布の分散が大きいかほどその観測は不確実であり、一方、分散が小さく、確率分布が単峰でピークが鋭いほどその観測は確実であるということを示している。

いま、同一の対象に対してセンシングを繰り返すとして、その「良いセンシング行動の系列」とは、状態 X に対して推定される確率分布 $p(x_i)$ をなるべく早く単峰の、しかも、なるべく小さな分散をもつ（すなわち、 X が本来的にもつであろう）分布とするものである。ある時刻 t でのセンシングのための行動を a_t で表わす。「望ましいセンシング行動」の系列 a_1, a_2, \dots, a_t とは、図 3 に示したように、その計測行動を順次繰り返すことで、 $p(x_i)$ が鋭い単峰となるように選ばれた系列であると言える。

2.2 情報の流れと情報エントロピー

確率分布の単峰性とその鋭さは、情報エントロピーによって評価できることが知られている。いま、 X の中で x_i が出現する確率を $p(x_i)$ とすると、「 x_i が出現した」という事象の情報量は $-\log_2 p(x_i)$ ビットであり、集合 X の持

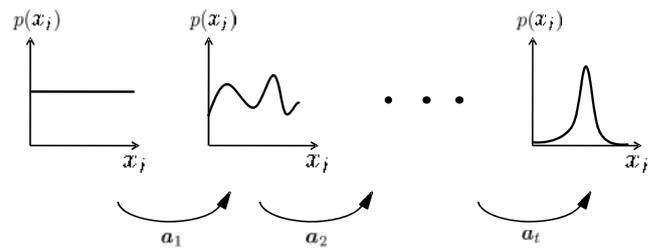


図 3 能動センシングの系列による「不確かさ」や「あいまいさ」の克服

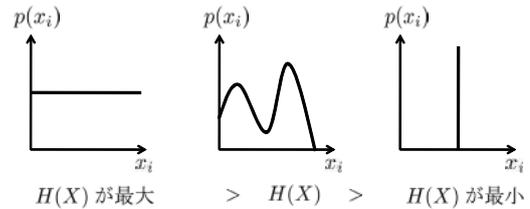


図 4 情報エントロピー $H(X)$ は、その情報源からどのような情報がもたらされるかがまったくわからないとき、つまり、どの生起確率も等しいときに最大、逆に、ある一つの事象しか生じないときに最小、生起確率に偏りがあるときは、その偏りの程度に応じてその中間の値をとる。

つ平均情報量（情報エントロピー）は、

$$H(X) = - \sum_{i=1}^M p(x_i) \log_2 p(x_i) \quad (1)$$

で与えられる。図 4 に示すように、 x_i があらかじめ特定できない状況であるほど、 $H(X)$ は大きくなる。

したがって、良い計測とは、なるべく大きくエントロピーを減少させる行為ということになる。ここでは、この情報エントロピーの減少量を最大化するセンシング行為を選択する指針を与える。

そこで、対象の状態 X とその観測 O の双方のもつ情報量と、その相互の情報の流れとの関係を定量化する必要がある。これは、下記に示すように、「相互エントロピー」と呼ばれる量で評価できる。したがって、各センシング行為における計測値からの状態の推定において、この相互エントロピーをセンシング対象のモデルからどのように定式化でき、それを最大化する計測行動 a_t をどのように導くかが、ここでの課題である。

まず、上記のセンシングのモデルにおける情報の流れについて見てみよう。図 5 に示すように、対象の状態が x_i であったとして、観測はそれぞれ条件付き確率 $p(o_j|x_i)$ に応じて o_j を得る。これは、「何が起きているのか」のモデルであって、この状況下でわれわれのなすことは、図 6 に示すように、 o_j を得て、 x_i を推定することである。ここでは、条件が入れ替わった条件付き確率 $p(x_i|o_j)$ が基本となる。

観測 O を得たときの X については、条件付き平均情報

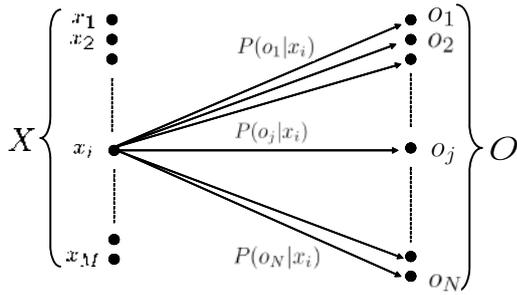


図5 情報の流れ. 状態 x_i に対して, 確率 $p(o_j|x_i)$ で, それぞれ観測 o_j を得る.

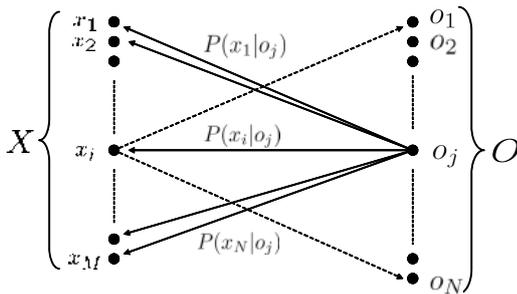


図6 観測の集合 O から状態の集合 X を推定する時のあいまいさ (図5を逆の立場から見たものにあたる)

量 (条件付きエントロピー) が,

$$H(X|O) = - \sum_i \sum_j p(x_i, o_j) \log_2 p(x_i|o_j) \quad (2)$$

と与えられる.

さて, O を知った後では X のわからなさは減少したはずで,

$$H(X) \geq H(X|O) \quad (3)$$

である. この差, すなわち, O を知ることで X のどれだけがわかったことになるのが相互エントロピーで,

$$I(X; O) = H(X) - H(X|O) \quad (4)$$

で与えられる. この値は, 平均情報量の定義式に戻れば,

$$\begin{aligned} I(X; O) &= H(X) + H(O) - H(XO) \\ &= H(O) - H(O|X) \\ &= I(O; X) \end{aligned} \quad (5)$$

でもある. ここで, $H(XO)$ は, (x_i, o_j) というペアの同時生起確率をもとにした情報エントロピー,

$$H(XO) = - \sum_i \sum_j p(x_i, o_j) \log_2 p(x_i, o_j) \quad (6)$$

である ((2)式とのわずかな違いに注意).

2.3 相互エントロピーの応用

本稿で提案している手法とその枠組みは, 上記の相互エントロピーを基盤としている. 上記のように, 相互エントロピーによって2つの集合 X, Y の要素の間でなんらかの対応がつくかを評価することができる. この手法による異なる種類の画像の重ね合わせ (位置合わせ) は, 特に医用画像の分野で多くの応用があり, CT 画像と NMR 画像など異なる「属性」を観察した画像間での重ね合わせに用いられている³⁾. 基本となるのは, 見ているものは違っても, 同じ点の画素同士には何らかの対応関係はあるはずだということである.

たとえば, X の要素 x と Y の要素 y にある特別な1対1関係があると, (x, y) の組について高い同時生起確率をもつ領域が, 局所的な偏りをもつ. 図4の類推から, この偏り具合は $H(XY)$ で評価できる. ただし, x や y の分布がもともと局所的であれば, x, y のペアも局所的になってしまう. そこで, X, Y それぞれの局所性を表わす $H(X), H(Y)$ でこれを正規化しておく必要がある. これが, 相互エントロピーを表わす (5) 式の解釈である.

すなわち, $I(X; Y)$ は, X と Y の間に1対1の関係が成り立つ度合いを表わすことになる. 同一対象の画像でも, それを画像中で追跡したい場合, 実用上は照明の変化による濃淡の変化 (それも, 局所ごとに異なる場合が多くある) や, 姿勢変化による形状の変形, 対象への距離の変化によるスケールや解像度の変化などに対応する必要がある. 対象の変形, 拡大・縮小にも対処した画像内の対象のトラッキングへの応用, ビジュアルサーボへの応用の例が多く報告されている^{4), 5)}.

また, 信号や画像をごく少数の基底の重ね合わせで表現するスパースコーディング, その展開形である独立成分分析などにおいても, 相互エントロピーは威力を発揮している⁶⁾.

3. 能動センシング

3.1 センシング行動のモデル

2章で提起した能動センシングのモデルを具体化したチャンネルモデルを, 図7に示す. ここでも, センシングにおいては対象の状態 x_t がこの時点 t でのセンシング行動 a_t によって「観測チャンネル」を通して観測者へ伝達され, 観測 o_t を得ることを示している^(注1). そこで, 図に示すように, センシング行動 a_t という条件付きの確率分布 $p(o_t|a_t)$ を得たことにあたる.

能動センシングでは, まず, 直前の時刻 $t-1$ までの観測データ o_0, o_1, \dots, o_{t-1} をもとに, 対象がどのような状態をとるであろうかを, 条件付確率分布 $p(x_t|o_{t-1}, \dots, o_0)$ という形で推定する. われわれは, 時刻 t では, この

(注1) 以下, x や o の添え字 t は時刻を表わし, 状態や観測の番号 i, j は省いてある. すなわち, x_{it}, o_{jt} を x_t, o_t と表わす

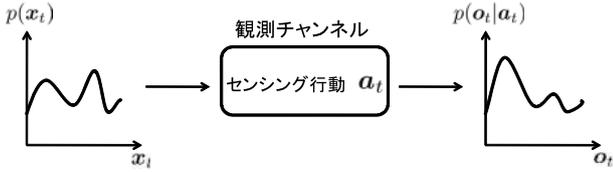


図7 対象の状態 x_t と観測 o_t を入出力として結ぶチャンネルモデルと、センシングにおける情報の流れ。 $p(x_t)$ は実際には過去の観測に依存している（依存するように a_t を選ぶ）が、その関係は省いてある。

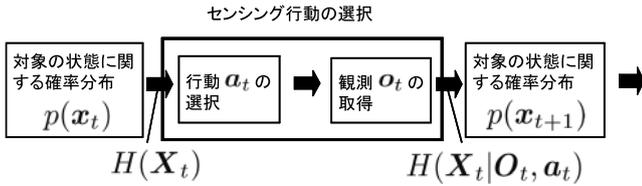


図8 能動センシングのステップ。より良いセンシング行動は、エントロピーをどれだけ改善できるか、すなわち、 $H(X_t)$ と $H(X_t|O_t, a_t)$ の差をどれだけ大きくできるかで評価できる。

$p(x_t|o_{t-1}, \dots, o_0)$ に基づいて観測計画を立て、つぎの時点での最適なセンシング行動 a_t を選択する。その結果、時刻 t での具体的な観測値 o_t を得る。

このときに、どのような観測値を得るかは、そのときの対象の状態 x_t とセンシング行動 a_t とによって、やはり確率的に表現される。すなわち、状態 x_t という条件の下での観測 o_t の起こりうる条件付確率は、 $p(o_t|x_t, a_t)$ と表現される。ここで、 $p(o_t|x_t, a_t)$ は、また、チャンネルにおける雑音の混入をモデル化するものでもある。一般には、平均が状態 x_t によって決まり、分散が選択された行動 a_t 、または、状態 x_t と行動 a_t の両方によって決まるガウス分布としてモデル化できる。

3.2 センシングの評価としての情報エントロピー

さて、あるセンシング行動 a_t によって観測 o_t を得るとすると、推定 $p(x_t)$ が $p(x_t|o_t, a_t)$ に基づくものとして改良できるものとしよう。ここで、この改良の程度は、上記のエントロピーで評価できる。すなわち、エントロピー $H(X)$ をどれだけ減少できるかで、そのセンシング行動の良さを評価する。

X_t, O_t を、時刻 t での状態と観測の取りうる値の集合であるとする、観測後の条件付きエントロピーは、

$$H(X_t|O_t, a_t) = - \sum_{x_t \in X_t} \sum_{o_t \in O_t} p(x_t, o_t, a_t) \log p(x_t|o_t, a_t) dx_t do_t \quad (7)$$

で与えられる。この間のセンシングのステップとエントロピーの変化を図8に示す。

すると、より良いセンシング行動は、 $H(X_t)$ と

$H(X_t|O_t, a_t)$ の差をどれだけ大きくできるかで評価できる。この差、

$$I(X_t; O_t|a_t) = H(X_t) - H(X_t|O_t, a_t) \quad (8)$$

は、まさに、前述の相互エントロピーであり、状態 x_t から観測 o_t へ情報が伝搬するとき、チャンネルで付加された正味の情報量ともみなすことができ、情報の不確かさをどれだけ減少させることができたかを示す。

したがって、最適なセンシング行動は、

$$a_t^* = \arg \max_{a_t} I(X_t; O_t|a_t) \quad (9)$$

を選ぶことで得られる。

ただし、実際に(8)式によって相互エントロピーを得るには無理がある。この式の後半の $H(X_t|O_t, a_t)$ を(7)式によって得るためには、確率分布 $p(x_t|o_t, a_t)$ が必要であるが、ここでのセンシングのモデルでは、実際には得ることができない。というか、この確率分布が得られるのであれば、それはもっと直接的にセンシングができていたことを意味するからである。

しかし、相互情報量エントロピーには、(5)式に示した都合の良い性質があり、上記(8)式は、

$$I(X_t; O_t|a_t) = H(O_t) - H(O_t|X_t, a_t) \quad (10)$$

という置き換えができ、さらに、 $H(O_t|X_t, a_t)$ は、まさに、図7に示したチャンネル雑音のエントロピーそのものにあたる。すなわち、ここで必要なものは、センシングにおけるチャンネル雑音のモデルであるが、これも、前節で述べたように、多くの場合はガウス分布を用いて妥当なモデル化ができる。

4. 画像データベース照合による物体認識

4.1 物体認識のためのカメラパラメータ選択

以上の議論の具体的な適用例を示す。ここでの例では、物体をカメラでとらえ、画像データベースと照合することで物体認識を行う。そのときのカメラパラメータの最適な選択を上記の枠組みの中で決定しようというものである。

認識対象の物体を、図9に示す。対象は、ほんの一部分を除いて全く同じ外見(特徴)をもつ物体であり、そのわずかな物体間の違いをうまく見分けることができる視点を選んで、効率的に認識することができるかを検証するための実験である。

図10に示すように、カメラの正面にターンテーブルに載った物体がある。センシング対象の状態 x は、物体の番号(物体は何か)とその姿勢(カメラからの距離と回転角)である。センシング行動 a が表わすものは、カメラのズームと物体の乗るターンテーブルの回転角の選択である。観測 o は画像であり、その各画素における濃淡値からなるベクトルとして表わす。

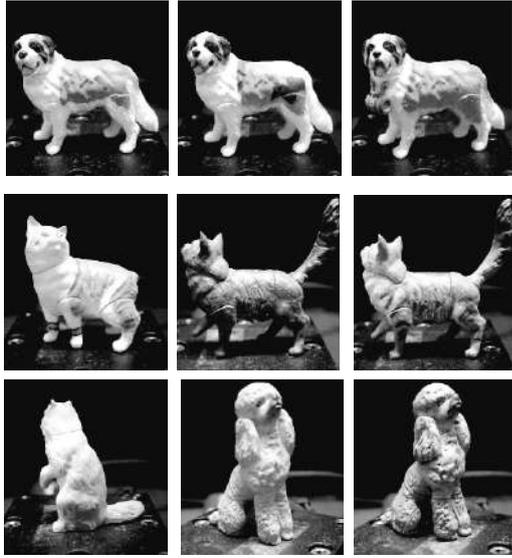


図9 実験で用いた9個の対象物体 x_1, \dots, x_9 . 背中
の模様, 首にさげた樽の有無などの細かい部分のみが異なる
物体が含まれている.

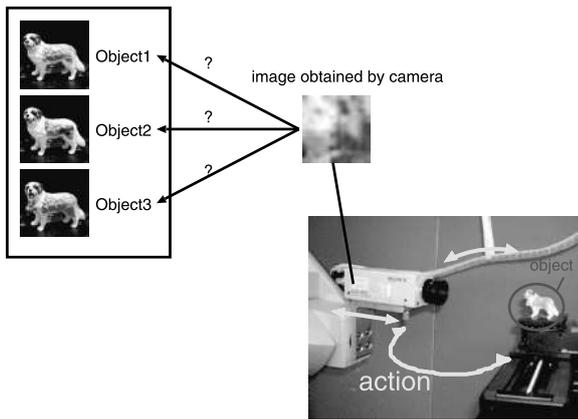


図10 能動的な物体認識のためのセンシング行動. 最適な
カメラのズームと回転テーブルの回転角を選択する.

この設定で, テーブルに乗っている物体が何であるか
とその姿勢を, カメラのズームとテーブルの回転角を選びつ
つなるべく早く (少ない観測で) 認識する. 具体的には, カ
メラの視点選択 (行動) を行った上で得る画像をデータベ
ースより求めたモデルと照合することで, 認識対象がどの物
体であるかを認識する. 最終的に, 特定の物体であるとい
う確率が十分に高くなるまで, 行動して認識することを繰
り返す.

画像データベースは, 前掲の9種類の物体について, そ
れぞれ, ズームが6段階, 回転が12段階の組み合わせで
全72種類よりなり, 図11に示すようにそのときの物体
と回転やズーム等の姿勢とでラベル付けしてある. その際
に, この72種類のそれぞれの視点につき照明などの条件
を変えて画像を得て, それらをガウス分布にあてはめるこ
とで $p(o|x, a)$ や, $p(o)$ をあらかじめ得て, (10) 式の計算

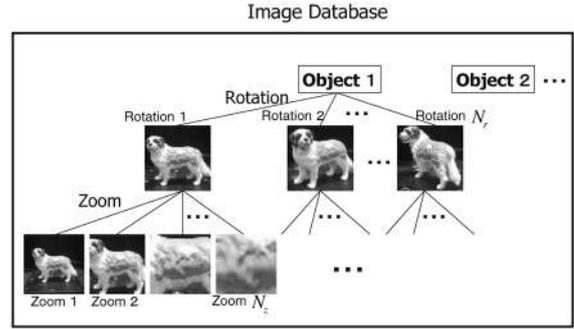


図11 データベースの例. 各画像について, どの物体に対
してどのような姿勢 (回転とズーム) で得たものかをラベ
ル付けしておく.

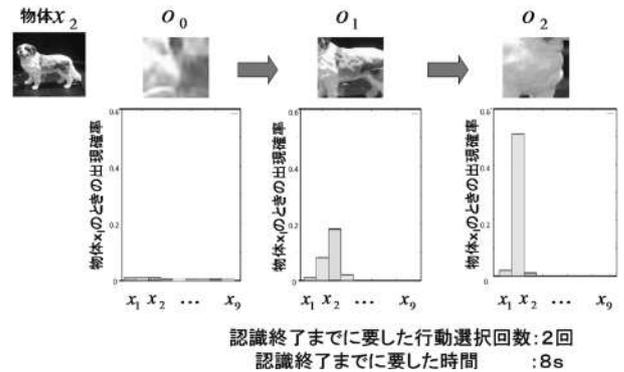


図12 行動選択を伴う物体認識. つぎつぎに行動選択をし
て画像を得ることで認識結果もつぎつぎと更新される.

の準備をしておく.

そして, 画像から, データベースのうちどの物体であ
るか, もしくはデータベースには存在しない物体であるか
を認識する.

4.2 認識結果の逐次的な更新

相互エントロピーによる評価である (9) 式に基づいて能
動的にカメラを動かすことによりつぎつぎに画像を得る. 行
動 a_t により観測 $p(o_t|a_t)$ を得ることになる. すると, 図
12に一例を示すように認識結果もつぎつぎに更新される.
ここで, 画像がどの物体のものであるかの確率 $p(x_t|o_t, a_t)$
は次式で求まる.

$$p(x_t|o_t, a_t) = \frac{p(o_t|x_t, a_t)p(x_t|a_t)}{p(o_t|a_t)} \quad (11)$$

式中の $p(x_t|a_t)$ は, 行動 a_t をとったときに物体が x_t だ
ると判定される確率である. これはこの段階では求まら
ないが, $t-1$ ステップ目までの観測で物体が x_t であると
判定された確率を利用して,

$$p(x_t|a_t) = p(x_t|o_{t-1}, a_{t-1}) \quad (12)$$

とする.

続いて, この確率分布をもとに, 次ステップでの最適な
センシング行動, すなわち, 対象の回転角とカメラのズー

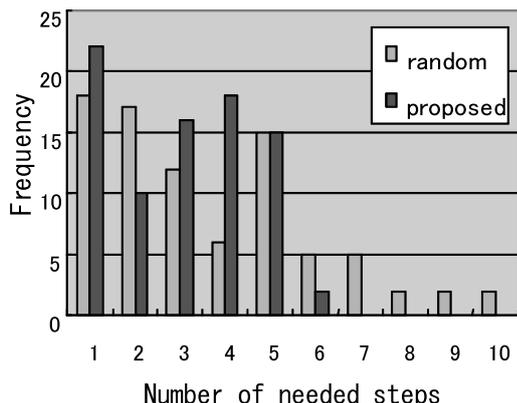


図 13 認識終了までに要した行動選択の回数

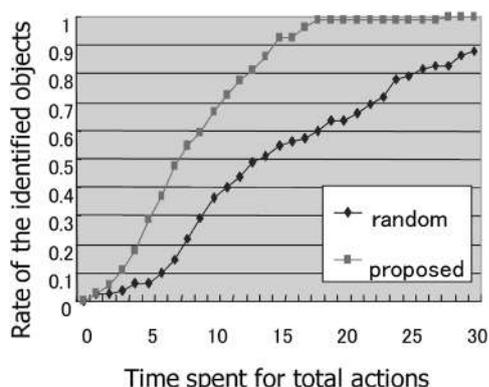


図 14 認識終了までに要した時間

ムを選択し、つぎの画像を得る。そして、(11)式で得られる確率がある特定の1つの物体に対してのみ高くなるまで、このステップを繰り返す。

4.3 物体認識の実験の設定と結果

この実験の目的は、似通った外見をもつ物体がデータベースに存在する場合に、そのわずかな物体間の違いをうまく見分ける効率的な視点の選択を検証することである。

各ステップで得られた画像を用いて、どの物体であるかの確率を求め、それを元にセンシング行動の選択をし、観測を繰り返していく。ある特定の物体に対する画像の出現確率が0.4以上、かつその物体である確率が0.9以上である場合は認識を終了する。そうでない場合はつぎの行動を選択し、画像を取得し、判定を繰り返した。

図13は行動選択をランダムに行ったときと提案手法を用いたときにすべての試行について認識終了までに要した行動選択回数を示している。提案手法の場合はランダムの場合と比べると、要した行動選択回数が全体的に少ないことがわかる。また、提案手法は多くても6回で認識を終えている。つぎに、図14は行動選択をランダムに行った場合と提案手法を用いた場合に認識終了までに行動選択に要

した時間を示している。提案手法がほぼすべての試行で認識が終了したときに要した時間内では、ランダムの場合では全試行の6割程度しか認識を終えていなかった。

5. まとめ

本稿では、「センシング = 情報を得る行為」とし、計測行為の良さを対象の状態に関する情報エントロピーの減少量(相互情報量)で評価する手法を提案し、センシング情報学の理論体系を構築した。ここでの考え方は、センシングを対象から観測への情報の流れと捉えるところにある。その上で、この情報の流れをスムーズにするセンシング行為という観点で、能動センシングの情報学の確立を試みた。

もう一つの課題として、センシングのコストについても考慮する必要がある。たとえ本稿の意味で最適な行動であっても、その行動へ移行するために多大な時間を要したり、装置上の困難さがあつたりして、戦略として成り立たない場合はいくらでもありうる。実際に応用するためには、能動センシングの最適性について、より広い検討が必要であろう。

(2012年6月27日受付)

参考文献

- 1) K. Deguchi: An Information Theoretic Approach for Active and Effective Object Recognitions, SICE Journal of Control, Measurement, and System Integration, 1-1, 33/39 (2008)
- 2) 出口, 本多: センシングのための情報と数理, コロナ社 (2009)
- 3) J.P.W. Pluim *et al.*: Mutual-information-based registration of medical images: a survey, IEEE Trans. Medical Imaging, 22-8, 986/1004 (2003)
- 4) A. Dame and E. Marchand: Accurate Real-time Tracking Using Mutual Information, Proc. ISMR, 47/56 (2010)
- 5) N. Dowson and R. Bowden: Mutual Information for Lucas-Kanade Tracking(MILK): An Inverse Compositional Formulation. IEEE Trans. PAMI, 30-1, 180/185 (2008)
- 6) 村田 昇: スパースコーディングの基礎理論と画像処理への応用, 情報処理学会研究報告, 2006-CVIM-153 (2006)
- 7) K. Deguchi and H. Ohtsu: An Information Theoretic Approach for Active and Effective Object Recognitions, ICPR2006, Tue-P-II-2 (2006)

[著者紹介]

出口 光一郎 君 (正会員, フェロー)



1976年東京大学工学系研究科修士課程修了(計数工学)。同年より、東京大学工学部助手、講師、山形大学工学部助教授、東京大学助教授を経て、98年東北大学大学院情報科学研究科教授、現在に至る。コンピュータビジョン、画像計測の研究に従事。2011年より横断型基幹科学技術研究団体連合会長を務めている。

