

木構造に基づく特徴構成法による視覚的学習

Visual Learning by Feature Construction Based on Tree Structure

野宮 浩揮 上原 邦昭
Hiroki Nomiya Kuniaki Uehara

神戸大学大学院自然科学研究科
Graduate School of Science and Technology, Kobe University

In developing a feature based visual learning method, the selection of features highly affects the performance of the method. For example, although simple features such as pixel intensity can be used widely, they are inadequate for complex learning tasks. In this paper, we propose two types of new visual learning methods which can solve complex learning tasks and do not depend on learning tasks. In both methods, low-level features are combined into a higher-level feature that is represented by a tree structure. The one is the appearance based method that combines some small parts of the contours of an object into a key shape to recognize the object. The other is the region based method that combines some local features of an object into a global feature which can fully represent the whole object.

1. はじめに

近年、画像に内在する特徴を発見することによって、画像中の物体を認識する様々な視覚的学習法が提案されている [Nomiya 04][Krawiec 03]. 特徴に基づく視覚的学習法を構成する際、どのような特徴を用いるかが学習手法の性能に大きく影響する。例えば、最も基本的な特徴の一つとして、画素の諧調値が挙げられる。このような低レベルの特徴は、学習課題に依存することなく利用できるため、有用な特徴である。しかし、低レベルの特徴だけでは、複雑な認識課題に対して十分な認識精度を達成するのは困難である。

これに対し、予め十分な記述能力を持つ特徴を定義しておく、複雑な認識課題に対応する方法が考えられる。しかし、このような特徴は学習課題に依存するため、非常に限定された範囲の認識課題においては良い性能を示すが、汎用性に乏しい。また、十分な記述能力を持つ特徴は概して複雑であるため、人手による定義が困難である。

本稿では、木構造を用いて多種の低レベル特徴(これを基本特徴と呼ぶ)を組み合わせて、複雑な認識課題にも適用可能な高レベルの特徴を構成する手法を提案する。画像内において、物体は形状と領域成分によって表されるため、外観(輪郭など物体の形状を表す要素)と領域(物体中の特に識別に有用な部分の特徴)に基づいたアプローチが、視覚的学習に有効である。そこで、このようなアプローチを構成するための基礎として、外観ベースと領域ベースそれぞれの視覚的学習法を提案する。

外観ベースの手法では、まず訓練事例の集合から、主要な輪郭片(一本の直線または曲線で表される、輪郭の一部)の部分集合を抽出する。そして、多くの正事例に共通する輪郭片を基本特徴とする。最後に、輪郭片同士の位置関係に基づいて複数の輪郭片を組み合わせ、木構造で表される高レベルの特徴を構成する。この木構造を特徴木と呼ぶ。

領域ベースの手法では、色や明るさの分布に関する単純な特徴を、予め基本特徴として定義する。そして、論理/比較演算子を用いて基本特徴を組み合わせ、物体の特徴が現れている領域成分を表す特徴木を構成する。有用な特徴木を得るため、様々な基本特徴の組み合わせを用いて特徴木空間の探索を行う

必要がある。通常の特徴空間内で探索を行うと、効率的な探索が困難であるため、提案手法では反転空間と呼ばれる空間を導入し、効率的な探索を実現している。

2. 外観ベースの視覚的学習法

はじめに、外観ベースの視覚的学習法を提案する。外観ベースの視覚的学習法とは、主に物体の輪郭を用いて物体の識別に有効な特徴を抽出し、識別規則を構成する学習法である。提案手法は、(1) 訓練画像からの輪郭片の抽出、(2) 類似輪郭片のクラスタリングによる基本特徴の生成、(3) 木構造に基づく基本特徴の構造化のプロセスから構成される。

第1のプロセスにおいて、各訓練事例から主要な輪郭の一部を多数抽出する。ここで、主要な輪郭の一部とは、輪郭の中でも比較的長い一本の直線あるいは曲線を指し、これを輪郭片と呼ぶ。次に、第2のプロセスでは、訓練集合から抽出された輪郭片の集合のクラスタリングを行い、正事例にはよく現れるが、負事例にはほとんど現れないような輪郭片を発見し、これを基本特徴とする。そして、第3のプロセスで、基本特徴間の位置関係に基づき、複数の基本特徴を組み合わせ、高レベルの特徴を表す木構造、すなわち特徴木を構成する。

2.1 輪郭片の抽出

物体の輪郭は、複数の輪郭片、すなわち一本の直線または曲線の集合から構成されている。外観ベースの手法では、まず物体の輪郭片を抽出する必要がある。輪郭片を抽出するために、提案手法では3D物体の認識に用いられる、コンテキストパッチ(context patch)と呼ばれる輪郭片の表現を導入する [Selinger 99]. コンテキストパッチは、線分の集合として得られる輪郭から、認識に有用な輪郭片を得ることによって生成される。図1に、コンテキストパッチと輪郭片の関係の例を示す。

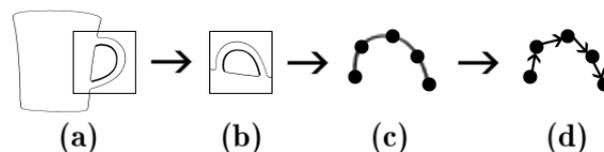


図1: コンテキストパッチと輪郭片

連絡先: 野宮 浩揮 (nomiya@ai.cs.scitec.kobe-u.ac.jp)

神戸大学 〒657-8501 兵庫県神戸市灘区六甲台町1-1

Tel: 078-803-6220, Fax: 078-803-6316

図1において、(a)は訓練画像を表しており、(b)はコンテキストパッチを表している。コンテキストパッチ内で太い線で示されている部分が輪郭片である。コンテキストパッチはサイズが固定されており、輪郭片の大きさと方向に合わせて、輪郭片が正方形領域の中心に位置するように正規化される。さらに、コンテキストパッチ内の輪郭片から、(c)に示すように輪郭上にほぼ等間隔に点を取り、(d)に示すような輪郭片のベクトル表現を得る。

2.2 基本特徴の生成

前節のプロセスにより、物体の輪郭から輪郭片の集合を得た。それぞれの輪郭片は非常に単純な形状であるため、基本特徴として用いることができる。しかし、抽出されたすべての輪郭片が物体の識別に有効であるとは限らない。例えば、正事例の中でもごく少数の事例に特有の形状であるとか、正事例だけでなく負事例にも多く含まれる形状などは、識別に有用であるとはいえない。したがって、輪郭片の集合中から、多くの正事例には含まれるが、負事例に含まれることはほとんどないような形状を示している輪郭片だけを基本特徴として用いる。

本稿では、識別に有効なパターンを発見するために用いることのできるパターンの類似尺度 AMSS (Angular Metrics for Shape Similarity) [Makio 05] を提案する。AMSS は複数の事例に共通して現れる輪郭片を抽出することができるので、これを用いて有用な輪郭片を発見する。AMSS では、輪郭片はベクトル列によって近似された形で表される。いま、2つの輪郭片 A, B がそれぞれ n 個のベクトルにより、 $A = \{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}$, $B = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ というベクトル列で表されるとすると、対応するベクトル間の類似度の総和により、 A と B の類似度を求められる。ベクトル間の対応は、類似度の総和が最大となるように動的計画法を用いて決定される。また、対応するベクトル \vec{v}_a, \vec{v}_b 間の類似度 $Sim(\vec{v}_a, \vec{v}_b)$ は、次のように表される。

$$Sim(\vec{v}_a, \vec{v}_b) = w(\vec{v}_a, \vec{v}_b)(1 - Dist(\vec{v}_a, \vec{v}_b)) \quad (1)$$

上式において、 $w(\vec{v}_a, \vec{v}_b)$ は類似度の総計に占める $Sim(\vec{v}_a, \vec{v}_b)$ の割合を表す。また、 $Dist(\vec{v}_a, \vec{v}_b)$ は \vec{v}_a, \vec{v}_b 間の距離を表し、次式で求められる。

$$Dist(\vec{v}_a, \vec{v}_b) = \frac{1}{2} \{Dist_a(\vec{v}_a, \vec{v}_b) + Dist_t(\vec{v}_a, \vec{v}_b)\} \quad (2)$$

ここで、 $Dist_a(\vec{v}_a, \vec{v}_b)$ および $Dist_t(\vec{v}_a, \vec{v}_b)$ はそれぞれ \vec{v}_a と \vec{v}_b のなす角 θ および長さの差異に基づいて次式で求められる。

$$Dist_a(\vec{v}_a, \vec{v}_b) = \frac{\theta}{\pi/2} \quad (3)$$

$$Dist_t(\vec{v}_a, \vec{v}_b) = \frac{||\vec{v}_a| - |\vec{v}_b||}{\max(|\vec{v}_a|, |\vec{v}_b|)} \quad (4)$$

ただし、 $\theta > \pi/2$ または $Dist_t(\vec{v}_a, \vec{v}_b) > LengthThreshold$ ($LengthThreshold$ は長さの差異の閾値を表す) ならば、 \vec{v}_a と \vec{v}_b はまったく似ていないと判断し、 $Sim(\vec{v}_a, \vec{v}_b)$ を 0 とする。

類似度をすべての輪郭片の組み合わせに対して求め、類似している輪郭片をクラスタリングしていく。そして、多数の正事例を含み、かつ負事例をほとんど含まないクラスを識別に有効な輪郭片とし、これを基本特徴として用いる。

2.3 特徴木の構成

前節のプロセスにより、基本特徴を定めることができた。しかし、これらの基本特徴は、単に正事例に多く見られ、負事例にほとんど見られない形状を表しているだけなので、物体中の位置、角度に関する情報を持たない。したがって、基本特徴の

みでは高い認識精度は得られない。そこで、基本特徴間の位置と角度の関係に基づき、複数の基本特徴を組み合わせ、より高レベルな特徴を生成する。提案手法では、基本特徴を組み合わせる際に木構造を用いる。特徴を木構造で表現すると、節点の数により特徴の複雑さを制御できるため、学習課題の複雑さによらず適切な特徴を構成できるという利点がある。

図1 (a) の画像に対する、基本特徴を組み合わせ構成される木構造 (特徴木) の例を図2に示す。ただし、この図は簡潔な説明のため、特にカップの取っ手部分のみについて作成したものであり、実際の特徴木はカップ全体に関する基本特徴の組み合わせから構成されることに注意されたい。

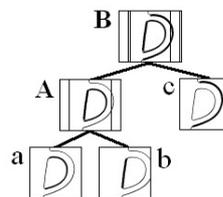


図2: 外観ベースの特徴木の例

特徴木は基本特徴に相当する葉 (図中の a, b, c) と、基本特徴間の位置関係を記述する節点 (図中の A, B) により構成される。各節点について、太い線で示されている線分が輪郭片を表している。節点 A の図は、基本特徴 a と b を適切な位置に配置して合成したものを示しており、A が a と b の組み合わせにより、取っ手の内側のループを表現していることがわかる。実際には、A は a と b の位置関係を記述するパラメータ (相対的な距離、角度) のみを保持する。さらに、A の記述する形状は、基本特徴 c と組み合わせることによって、より上位の特徴を表す節点 B を構成する。つまり、B は実質的には3つの基本特徴 a, b, c から取っ手全体の形状を表現している。

このような特徴木は、基本特徴の組み合わせについて、AMSS に基づいて位置関係の類似度を求めることにより構成される。まず、すべての基本特徴の対について、類似度の高い対をグループ化し、その位置関係を記述する節点を生成する。例えば、図2の例では、類似度の高い基本特徴である a, b をグループ化し A を生成する操作に相当する。次に、すべての基本特徴および基本特徴のグループの対について、類似度の高い対をさらにグループ化し、位置関係を記述する節点を生成する。図2の例では、A, c をグループ化し B を生成する操作に相当する。そして、これ以上類似している対がなくなるまでグループ化を繰り返し、最終的に構成された特徴木を出力する。

このように特徴木を構成し、単純な線分で表される輪郭片の集合から、物体を特徴づける形状を構成する。物体の形状には、識別に有効な情報が多く含まれているので、外観ベースの視覚的学習法により、高い認識精度を実現できる。しかし、画像中の物体を構成するもう一つの重要な要素として、領域成分がある。物体の外観に関する情報とともに、識別対象の物体に特有な領域内の特徴を用いれば、より高い識別能力が得られる。次章では、領域ベースの視覚的学習法について述べる。

3. 領域ベースの視覚的学習

本章では、領域ベースの視覚的学習手法を提案する。領域ベースの手法では、物体内の特定の領域における特徴を用いて、識別規則を構成する。これにより、物体に内在する特徴的な部分に注目でき、効率的な識別が可能になる。提案手法では、予めいくつかの基本特徴と、その関係を記述するための論

理/算術演算子を定義する。そして、木構造に基づいて基本特徴と演算子を構造化し、識別規則を表す特徴木を構成する。

一方、有用な特徴木を得るため、様々な基本特徴と演算子の組み合わせを用いて、特徴木空間における探索を行う必要がある。通常の特徴木空間における探索は、特徴木の評価基準を定めるのが難しく、探索に要する計算量が非常に大きくなるため、提案手法では反転空間と呼ばれる空間を導入し、適切な評価基準の設定と効率的な探索を実現している。

3.1 特徴木の構造

領域ベースの提案手法も、外観ベースの手法と同じく、基本特徴を構造化して特徴木を構成する。しかし、領域ベースの手法では、予め定義された基本特徴と、その関係を記述する演算子により特徴木を構成するため、特徴木の構造が異なる。本節では、領域ベースの特徴木の構造について述べる。

領域ベースの手法で効果的な特徴構成を行うためには、多様な基本特徴が必要となる。提案手法では、画素の階調値の和、ストローク密度関数、閉じ率、ペリフェラルパターン、ランレングスの5種類の特徴を基本特徴としている。さらに、これらの基本特徴をより有効に利用するために、画像に対してフィルタリングを行う。フィルタリングには、水平/垂直 Sobel フィルタ, Laplacian フィルタ, Gaussian フィルタを用いている。フィルタリングは基本特徴値を得る前に行われ、主に物体の垂直/水平方向のエッジや輪郭を抽出したり、ノイズを低減する。

ここで、領域ベースの特徴木の例を、図3に示す。

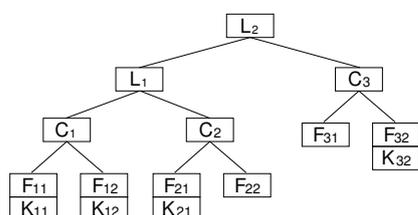


図3: 領域ベースの特徴木の例

領域ベースの特徴木は、基本特徴に相当する葉と、論理/算術演算子に相当する節点により構成される。多様な演算子を用いて基本特徴を階層的に組み合わせるため、最終的に構成される特徴は、物体の識別に有用な多くの局所的特徴を捉えることができる。図3において、L, C, F, K はそれぞれ論理演算子、算術比較演算子、基本特徴、フィルタリングオペレータを表す。フィルタリングオペレータを持つ葉に対しては、基本特徴値を得る前にフィルタリングが行われるが、持たない葉に対してはフィルタリングは行われず、訓練画像から直接基本特徴値を得る。論理演算子として AND, OR, XOR を用いており、算術比較演算子として以下の8種類の演算子を用いている。

- = : 等しい
- \approx : ほぼ等しい ($|F_1 - F_2| < \frac{F_1 + F_2}{20}$)
- \geq [\leq] : 等しいかより大きい [小さい]
- $>$ [$<$] : より大きい [小さい]
- \gg [\ll] : 十分に大きい [小さい] ($F_1 > 2F_2$ [$2F_1 < F_2$])

また、左から1, 2, 3番目および一番右の葉はフィルタリングオペレータを持っているので、特徴値 $F_{11}, F_{12}, F_{21}, F_{32}$ を計算する前に、それぞれフィルタ $K_{11}, K_{12}, K_{21}, K_{32}$ によりフィルタリングが行われる。構成された特徴 $f(x)$ は、事例 x に対して真または偽を出力する。事例 x は、 $f(x)$ の出力が真のとき正事例に分類され、偽のとき負事例に分類される。

3.2 効率的特徴構成のための手法

特徴構成を行う際には、有用な特徴木を得るために、特徴空間内における探索が必要となる。しかし、探索には二つの大きな問題が伴う。一つは、探索空間すなわち特徴空間が極めて広いこと、効率的な特徴空間の探索が困難であり、探索のための計算量が非常に大きくなるという問題である。もう一つは、探索の際に特徴木を評価するための基準の問題である。提案手法では、正事例と負事例の分離度を評価基準としている。これは認識精度と直接関係しており、有効な評価基準であると考えられる。しかし、特徴空間において効率的に求められる分離度を定式化するのは困難である。そこで、これらの問題を解決するため、特徴空間の代わりに反転空間と呼ばれる空間を導入する。反転空間内では、特徴の組み合わせを幾何学的に表現できるため、特徴の有用性の評価や探索を効率的に行える。

3.3 反転空間の導入

一つ目の問題を解決する上で有効な要素として、反転空間 (inverted space) と呼ばれる空間が提案されている [Mehra 89]。反転空間は、事例と特徴の関係を記述するフレームワークの一つであり、特徴の組み合わせを幾何学的に表現できるため、効率的に特徴空間を探索することができる。反転空間内では、特徴空間とは逆に、事例が次元 (軸) に相当し、特徴が空間内の点 (またはベクトル) に相当する。したがって、特徴は単純なベクトル演算によって組み合わせることができ、特徴が変わると空間が変わってしまう特徴空間に比べて、特徴を組み合わせる際の計算量を大幅に低減することができる。提案手法では、特徴空間の代わりに反転空間内で特徴を表し、探索を行う。

図4に反転空間の例を示す。(a) は特徴空間を示しており、(b), (c) はそれぞれ対応する正, 負の反転空間を示している。正, 負の反転空間の軸はそれぞれ正事例, 負事例のみで構成されている。反転空間での特徴の有用性を評価するために、等値線 (equiparameter line) と呼ばれる直線が定義されている。等値線は次の条件をみたす特徴の集合で構成される。

$$e_1 = e_2 = \dots = e_N \quad (5)$$

ここで、 e_i は i 番目の事例の特徴値、 N は正事例または負事例の数を表している。

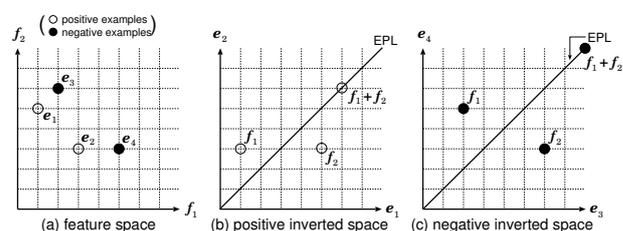


図4: 反転空間の例

図4 (b), (c) において、“EPL” と表記されている直線が等値線である。例えば、既存の特徴 f_1, f_2 から、新しい特徴 $f' = f_1 + f_2$ を生成することを考える。(b), (c) より、 f' は正負双方の反転空間内において等値線上にあることがわかる。したがって、特徴 f' は最適な特徴であるといえる。実際、 f' の値はすべての正事例について6、すべての負事例について8となるため、特徴 f' により正事例と負事例を完全に分離できる。

3.4 特徴木の評価基準

反転空間のフレームワークが適切な特徴木の評価基準を提供するため、二つ目の問題は反転空間に基づいて評価基準を定めれば解決できる。提案手法では、二つの基準を用いて特徴木

の評価基準を定義する。特徴と等値線の間の距離が短いほど、特徴の識別能力が高いため、一つ目の基準としてこの距離を用いる。また、特徴を正と負それぞれの反転空間内の点として表したとき、両者間の距離が大きいかほど識別に有利であるため、この距離を二つ目の基準として用いる。これら二つの基準値は、判別分析で用いられている級内分散と級間分散によって近似できる [Mehra 89]。したがって、級間分散の級内分散に対する比を正事例と負事例の分離度とみなすことができる。この基準を視覚化した図を図5に示す。

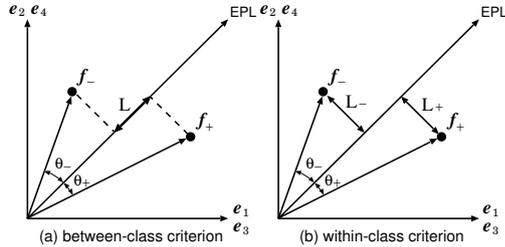


図5: 特徴木の評価基準

図5において、 f_+ および f_- はそれぞれ正と負の反転空間における特徴ベクトルを表す。横軸と縦軸は、図4 (b) [(c)] における e_1 と e_2 [e_3 と e_4] に対応する。級間分散は図5 (a) 中の線分 L の長さに相当し、級内分散は図5 (b) 中の線分 L_+ と L_- の長さの和に相当する。以上より、特徴の評価基準 E を以下のように定義する。

$$E = \frac{L}{L_+ + L_-} = \left| \frac{F_+ \cos \theta_+ - F_- \cos \theta_-}{F_+ \sin \theta_+ + F_- \sin \theta_-} \right| \quad (6)$$

ここで、

$$F_+ = \frac{|f_+|}{\sqrt{n_+}}, \quad F_- = \frac{|f_-|}{\sqrt{n_-}}$$

$$\begin{pmatrix} |f_+| \quad |f_-| & : & \text{特徴ベクトル } f_+ \text{ [} f_- \text{] の長さ} \\ n_+ \quad n_- & : & \text{正事例 [負事例] 数} \\ \theta_+ \quad \theta_- & : & f_+ \text{ [} f_- \text{] と等値線のなす角} \end{pmatrix}$$

反転空間内では、特徴が幾何学的に表されるため、特徴の有用性が容易に評価でき、探索をより効率的に行える。

3.5 有用な特徴木の探索

特徴木を構成するための基本特徴の組み合わせの数は膨大であるため、特徴空間全体を探索することは非常に困難であり、多くの計算量を必要とする。この問題に対して、提案手法ではランダム探索法を導入して、探索の効率化を図っている。単純なランダム探索は、計算量を大幅に削減できるが、それほど性能の良くない特徴木しか見つけられないことも多い。そこで、提案手法では単にランダムに探索を行うだけでなく、有用性の高い既存の特徴木の周辺を重点的に探索し、この問題に対処している。以下に、提案手法における特徴木の探索手順を示す。

はじめに、いくつかの特徴木をランダムに生成する。次に、生成された特徴木の有用性を評価し、有用性の高い特徴木を選択する。そして、選択された特徴木をランダムに組み替えて、新しい特徴木を生成する。このように、評価と組み替えを繰り返すため、生成される特徴木の性能は徐々に良くなる。また、探索が局所解に陥ることを避けるため、特徴木の組み替え過程で、ランダムに生成された特徴木をいくつか加える。

図6に示すように、2つの特徴木の部分木を組み替えて、新しい特徴木を生成する。図6 (a) および (b) は組み替え対象の

特徴木を表しており、(c) は組み替えの結果生成された新しい特徴木を表している。組み替え対象となる部分木は特徴木 A と B からランダムに選択され (図6の場合、 S_A と S_B がこれに相当する)、組み替えの結果 (c) が生成される。

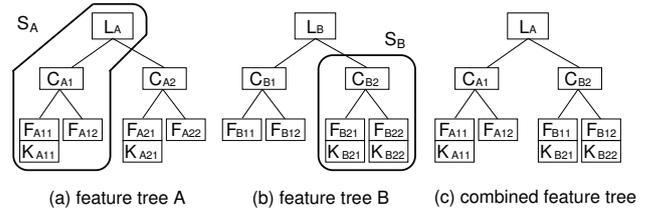


図6: 特徴木の組み替え

評価と組み替えの過程において、各特徴木は式 (6) の基準によって評価される。最終的に、探索の過程で得られた評価値の高い特徴木が学習結果として出力される。

4. まとめ

本稿では、画像内の物体を表現するために必要な要素である物体の形状と領域成分に基づいて、外観ベースおよび領域ベースの視覚的学習手法を提案した。いずれの手法も、画像を構成する基本的な要素に基づき、高レベルな特徴を発見するため、学習課題に依存せず、複雑な認識課題に適用が可能である。

今後の課題としては、実世界の画像データを用いて実験を行い、各提案手法の性能を検証すること、既存の視覚的学習手法との比較を行い、提案手法の有用性を検証し、改善すべき点についての検討を行うことが挙げられる。また、本稿では外観ベースと領域ベースそれぞれの手法を提案したが、物体の形状と領域成分双方を考慮した視覚的学習手法を構成できれば、より高い認識能力を実現できると考えられるので、両手法を統合した視覚的学習手法を構成する必要がある。

参考文献

- [Nomiya 04] H. Nomiya and K. Uehara: Visual Learning by Set Covering Machine with Efficient Feature Selection, *Proc. of the 16th European Conference on Artificial Intelligence*, pp. 525–529, (2004).
- [Krawiec 03] K. Krawiec and B. Bhanu: Visual Learning by Evolutionary Feature Synthesis, *Proc. of the 20th International Conference on Machine Learning*, pp. 376–383, (2003).
- [Selinger 99] A. Selinger and R. C. Nelson: A Perceptual Grouping Hierarchy for Appearance-Based 3D Object Recognition, *Computer Vision and Image Understanding*, Vol. 76, No. 1, pp. 83–92, (1999).
- [Makio 05] K. Makio, Y. Tanaka and K. Uehara: Discovery of Skills from Motion Data, *New Frontiers in Artificial Intelligence: Joint Proc. of the 17th and 18th Annual Conferences of JSAI*, (2005).
- [Mehra 89] P. Mehra, L. A. Rendell and B. W. Wah: Principled Constructive Induction, *Proc. of International Joint Conference on Artificial Intelligence*, pp. 651–656, (1989).