

二次元図形の概念学習システム

松井良男 辰巳昭治
豊橋技術科学大学 情報工学系

本報告では、視覚情報からの例題学習において、訓練例の選択や学習終了の判断に対する教師の負担を軽減するという立場から、データ駆動型学習アルゴリズムに質疑応答機能を取り入れたシステムについて述べる。システムの対象領域は簡単な二次元図形世界であり、プリミティブと呼ばれる基本図形の視覚的情報(形、大きさ、重心の位置、隣接関係等)を表わす単項述語や二項述語を用いて、図形構造についての概念記述を学習していく。形成概念として、プリミティブ図形により構成される形、大小関係、位置関係等の静止画として示される訓練例から学習可能な概念の他に、移動、回転等の動画として示される訓練例から学習可能な概念を取り上げた。

A CONCEPT LEARNING SYSTEM IN THE WORLD OF TWO-DIMENSIONAL FIGURES

Yoshio MATSUI and Shoji TATSUMI

Department of Information and Computer Sciences, Toyohashi University of Technology
Tempaku-cho, Toyohashi, 440 Japan

This report describes a system which learns the conceptual descriptions of two-dimensional figures by means of visual information. To describe the learning concepts fourteen primitive figures are prescribed. Then the concept is made from a set of attributes of each primitive figures (adjacent relation, shape, size, color and location of centroid), as well as divided into the static-concept and the dynamic-concept. A data-driven method has been used as learning algorithm, but to cope with the drawbacks of conventional one (i.e. the selection of training instances and the terminational condition of learning), the function of a question and answer(Q&A-FUNCTION) is adopted in our system.

1. はじめに

コンピュータに視覚を持たせようという目的で、パターン認識や画像処理等の分野で得られた成果を取り入れながら、対象シーンの記述やその理解に焦点を当てた研究が人工知能の一分野として近年盛んになってきている。コンピュータビジョンと呼ばれるこの分野では、主に線画の抽出、線画の解釈等の技法に関する研究が進められてきたが、より一般的な人工知能の立場から、これまでに得られた技法や人工知能の他の分野で提案されている理論を基にして、対象シーンの意味する構造的な概念を記述する研究も行なわれてきつつある。

本研究では、対象領域を二次元図形世界に限定し、この領域に特有の特徴量(形、面積、重心等)を基本的なカテゴリとして、図形間の位置関係、大小関係、拡大縮小、回転、反転、プリミティブ図形によって構成される形等の図形構造の概念記述を例題からのデータ駆動型アプローチを用いて学習するシステムについて述べる。

例題からのデータ駆動型学習システムはWinston、Vere、Mitchellらによって提案されているが[1][2][3][4]、これらのシステムでは、教師すなわちユーザが逐次的に訓練例を選択してシステムに与える必要があり、実際に応用する場合、訓練例の選択や訓練終了の判断に対する教師の負担が大きいことが問題になる。

本システムでは、システムから教師への質問機能を取り入れることにより、これらの問題点に対処した。ここで行なわれる質問とは、システム自身が訓練例を生成し、それが学習しようとしている概念の訓練例になっているか否かを教師に問いかけることである。システムは、教師との質疑応答を基にして規則空間の探索を進めていき、正の訓練例を構成する述語の削除や一般化操作によって概念を形成した後、自律的に訓練を終了させる。探索空間は最初に与えられた正負の訓練例のスケルトンと正の訓練例を比較し、一般化する必要のある述語と必要のない述語を区別することにより狭められている。また、形成概念として、静止画として示された正負の訓練例から学習される概念(静的概念)の他に、動画として示された正負の訓練例から学習される概念(動的概念)を取り上げた。

2. 対象領域

二次元図形世界を構成する基本的な平面図形をプリミティブ(primitive)と呼ぶ。また、一枚の画像を一つのフレーム(frame)と呼び、各処理をフレーム単位で行なう。一つのフレームはそのなかに含まれているプリミティブの特徴、およびプリミティブ間の二項関係で図1のように特徴づけられる。表1の述語を用いたS式でフレームを表現する。

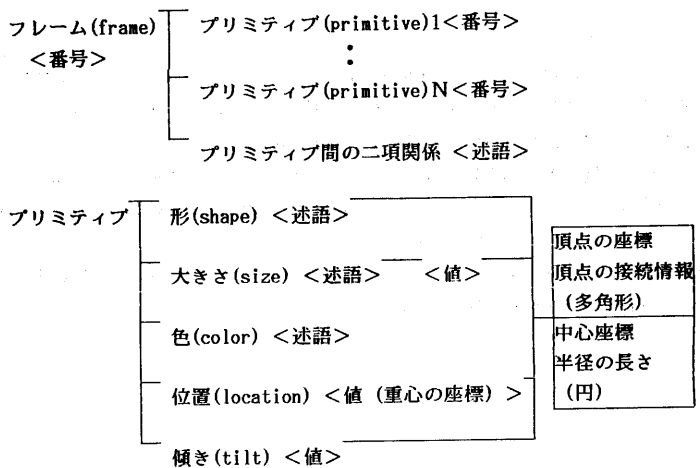


図1 フレームの構造

フレームには、その時間的順序に従って、番号1～Nがつけられる。各プリミティブは、前述の基本的特徴から構成されるが、(色情報以外の)これらの情報を構成する下位の情報として、頂点座標等の情報を持っている。頂点の接続情報とは、どの頂点とどの頂点が結ばれているかを示す情報であり、平面図形の辺を決定する。〈〉は記述形式を表し、記述されたものを属性値と呼ぶ。プリミティブ間の二項関係は隣接関係のみをとりあげ、交り、包含等は直接記述しない。本システムで取り上げた隣接関係は

- (1) P-P隣接：二つのプリミティブが一つの頂点を共有している。
- (2) P-L隣接：あるプリミティブの頂点が別のプリミティブの辺上にある。
- (3) L-L隣接：あるプリミティブの辺が別のプリミティブの辺上にあり、両者の長さが異なる。
- (4) S-L-L隣接：二つのプリミティブが一つの辺を共有している。

および、非隣接の五つである。大きさは、(下位情報によって計算されたプリミティブの面積) / (フレームの面積) の値により述語に変換される。色は、各プリミティブを構成するドットに付随している数値によって、述語表現される。一つのプリミティブは二つ以上の色情報を持たないと仮定する。

取り扱う概念は、静止画として示された正負の訓練例から学習される静的概念(プリミティブ図形によって構成される形、大小関係、位置関係に対する概念)と、動画として示された正負の訓練例から学習される動的概念(移動、拡大縮小、回転、反転に対する概念)に分類できる。静止画は一つのフレームで構成され、動画は時間的順序関係をもつ二つのフレームの系列から構成される。

表1. 使用する述語

形	cir(円) eIT(楕円) e-t(正三角形)	大きさ	v-s(微小な) s(小さい) m(中くらいの)
	i-t(二等辺三角形) r-t(直角三角形)		l(大きな) h(巨大な) s-値(数値表現)
色	tri(三角形) squ(正方形) rec(長方形)	隣接	bla(黒) blu(青) red(赤) pur(紫) gre(緑)
	rho(菱形) par(平行四辺形) tra(台形)		s-k(水色) yel(黄) whi(白)
重心	qua(四角形) r-p(正五角形) pen(五角形)	関係	not-tou(非隣接) p-p(隣接) p-l(隣接)
	l-値-値(位置) g-l-値(方向) g-d-値(距離)		l-l(隣接) s-l-l(隣接)

3. システムの構成

システムは、与えられた訓練例をもとにして概念形成を行なう学習ユニット、および形成された概念を利用し、簡単な照合手続きで判定処理や説明処理を行なう推論ユニットから構成される。図2にシステムの構成を示す。ここでは、紙面の都合上、中心的な役割を果たす学習ユニットのみについて説明する。

3.1 学習ユニット

学習ユニットは例題からのデータ駆動型アルゴリズムを用いて、概念形成を行うが、従来提案されてきたアルゴリズムは教師負担に関して、

* システムにより形成される概念の妥当性を判断しながら、逐次的に訓練例を選択する必要がある。

* 訓練の終了条件が明確でなく、教師が判断しなければならない。

が欠点としてあげられる。本ユニットでは質疑応答機能を取り入れ、次の枠組で処理を行うことにより、これらの欠点に対処する。概念形成は提供される情報の水準(level)と質(quality)により決定される[1]。ここで取り上げている視覚情

報における例題学習では訓練例として低水準の情報が与えられるため、重要でない詳細情報を無視して情報の一般化を行ない、概念を形成する。尚、情報の質として、他のデータ駆動型学習システムと同様、雑音を含まない高品質の訓練例を仮定している。

処理の枠組

- (1) 教師から正負二つの訓練例を受け取り、学習概念を正の訓練例に初期化した後、一般化する必要のある述語と必要のない述語を区別する。
- (2) 一般化が必要な述語だけに注視点移して訓練例を生成する。
- (3) 生成されたモデルに対する賛否をユーザに与えてもらいながら学習概念を修正する。
- (4) 一般化すべき述語が無くなるまで、(2)~(3)を繰り返す。

通常、一般化は概念を構成する述語の削除だけで行われるが、本ユニットではより一般的な意味を表わす述語へ置換していく一般化方法も取り入れている。例えば隣接関係や形、大きさを表現する述語は、次のように順序付ける。隣接では、その隣接状態における位置関係の自由度を基準にし、大きな自由度を持つ述語を一般的であるとする。形を表わす述語は、辺の長さの関係や角の大きさを比較して、順序付けを行う。

(一般的な隣接関係) P-P ← P-L ← L-L ← S-L-L (特殊な隣接関係)

(一般的な三角形) TRI ← R-T ← I-T ← E-T (特殊な三角形)

(特殊な大きさ) V-S → S → (一般的な大きさ) M ← L ← H (特殊な大きさ)

システムは一般化を行おうとしている述語が、どの程度一般的であるのかを判断し、述語を置換する。順序付けられない色情報の述語に対しては削除だけで一般化を行う。重心の方向関係のように反対の状態を意味する述語を持つ二項述語に対しては変項の順序替えが行われる。

一般化する必要のある述語と必要でない述語は、次の分類規則で区別される。

一般化が必要な述語..... $\{\forall x \mid x \in S1 \cap S2\}$

一般化が必要でない述語..... $\{\forall y \mid y \in S1 \cap \neg S2\}$

但し、S1.....正の訓練例に含まれている述語を要素とする集合

S2.....正負の訓練例のスケルトンに含まれている述語を要

素とする集合

即ち、正の訓練例を構成する述語と負の訓練例を構成する述語の違いをみつけ、その違いのために正の訓練例が負の訓練例になっていると仮定し、負の訓練例とマッチングのとれなかった正の訓練例の述語は学習概念に対し必要不可欠なものであると解釈する。不十分な一般化をさけるため、分類規則は構造的特徴を表現する二項述語だけに適用する。

訓練例の生成は、他の特徴に与える影響度が大きい特徴を基準にして行い、先に決定された制約条件をその他の特徴に伝搬させ、訓練生成における探索空間を狭めていく方向で進められる。影響度が最も大きい特徴は隣接関係であると考えられる。例えば、この特徴がL-Lに決定されていれば、形に対する一般化の可能性としてCIRなどを考える必要はなくなる。この訓練例の生成原理を図3に示す。学習は一般化すべきプリミティブがなくなった時、終了する。

3. 2 静的概念の形成アルゴリズム

1. 教師から正負二つの訓練例をS式表現で受け取る。学習概念Hを正の

訓練例に初期化する。

2. Hを構成するプリミティブの位置情報を二項述語に変換する。

- (1) 隣接関係、あるいは重心間の距離関係をもとにして関連性の強い二つのプリミティブを見つける。
- (2) 重心間を結んでできる直線の角度、および重心間の距離から、次のような二項述語を作り出す。

G-L-〈値〉(A B): 重心間の方向関係を表わす二項述語

G-D-〈値〉(A B): 重心間の距離関係を表わす二項述語

3. 3.1 二つの訓練例のマッチングをとり、スケルトンを発見する。

- (1) 二つの訓練例で正の訓練例になっているものをP-T-I、負の訓練例になっているものをN-T-Iとする。P-T-IとN-T-Iのどちらか一方にしか含まれていない述語名を持つ述語があれば、それらを取り除き、P-T-I、N-T-Iを整理する。
- (2) 変項の種類で探索空間の刈り込みを行ないながら、整理したP-T-I、N-T-Iの変項をユニフィケーションしていき、二つの訓練例に共通な最大論理積をすべて求める。最大論理積が一つだけ求められた場合、これをスケルトンSKEとする。その他の場合、求められた論理積どうしの論理積をとり、SKEとする。

3.2 スケルトンと正の訓練例を比較して、一般化が必要な特徴と必要でない特徴を区別する。

- (1) SKEの変項にあわせてP-T-Iの変項をユニフィケーションする。
- (2) ユニフィケーションしたP-T-Iを構成する述語のうちで、SKEには含まれていない二項述語に重み1、その他の述語に重み0.5を割当てて新にHとする。

4. 4.1 Hをもとにして新しい訓練例を生成し、それが学習しようとしている概念の訓練例になっているか否かをユーザから知らせてもらう。

4.2 4.1で得られた情報をもとにしてHを一般化していく。

5. Hを構成するすべてのプリミティブの重みが1になるまで4を繰り返す。

3. 3 動的な概念の形成アルゴリズム

1. 教師から正負二つの訓練例をS式表現で受け取る。

(各々の訓練例はそれぞれ二つのフレームの系列から構成されているため、システムは四つのフレームのS式表現を受け取る。)

2. 2.1 正の訓練例を構成する二つのフレーム間に共通な最大論理積を求め、変化した特徴を発見する。

- (1) 正の訓練例を構成する二つのS式表現を、それぞれP-T-1、P-T-2とする。(但し、P-T-1は、P-T-2よりも過去の事柄を表現しているS式表現である。)
- (2) P-T-1とP-T-2のマッチングをとり、共通な最大論理積を求めP-GLPとする。
- (3) P-GLPの変項にあわせてP-T-1、P-T-2の変項をそれぞれユニ

フィケーションする。

- (4) ユニフィケーションしたP-T-1を構成する述語のうちで、P-GLPにも含まれている述語を取り除き、新たにP-T-1とする。
- (5) (4)と同様にして、新たなP-T-2を作り出す。

2.2 負の訓練例を構成する二つのフレーム間に共通な最大論理積を求め、変化した特徴を発見する。

- (1) 負の訓練例を構成する二つのS式表現を、それぞれN-T-1、N-T-2とする。(但し、N-T-1は、N-T-2よりも過去の事柄を表現しているS式表現である。)
- (2) 2.1の(2)、(3)、(4)、(5)と同様にして、新たなN-T-1、N-T-2を作り出す。

3. P-T-1とP-T-2を合成し、P-T-Iとする。さらにN-T-1とN-T-2を合成してN-T-Iとする。静的概念の形成アルゴリズムを適用する。

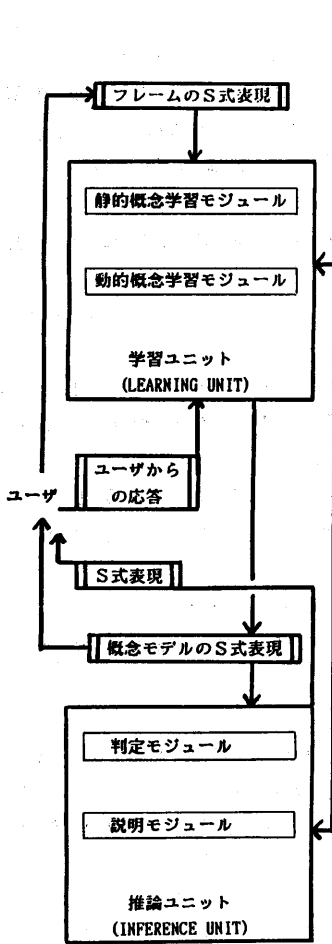


図2 システムの構成

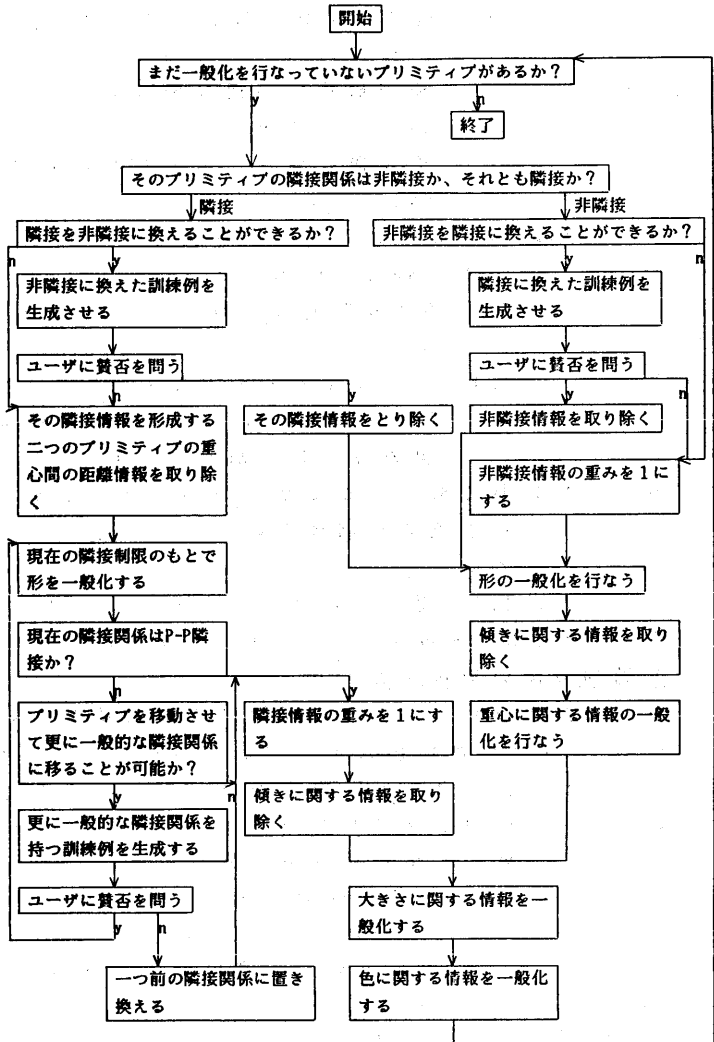


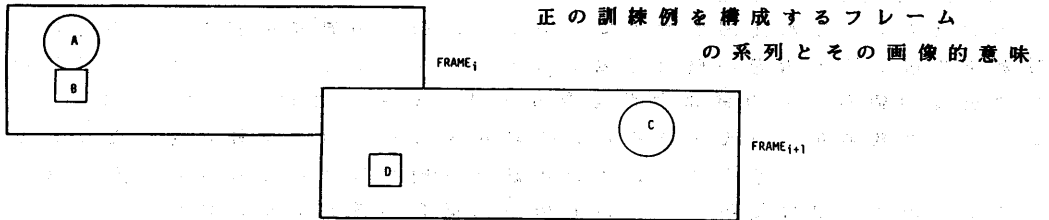
図3 訓練例の生成原理

4. 処理結果

"右に移動する"という動的概念の処理に使用した訓練例とその画像的意味、および概念形成のプロセスを図4に示す。図5は"人が走る"という動的概念の処理結果であるが、ここでは使用した訓練例の画像的意味と形成された概念を示すだけにとどめる。各フレームは表1の述語を用いて表現されている。システムが生成した訓練例の数は"右に移動する"の概念形成に対して7個であり、"人が走る"の概念形成に対しては33個であった。

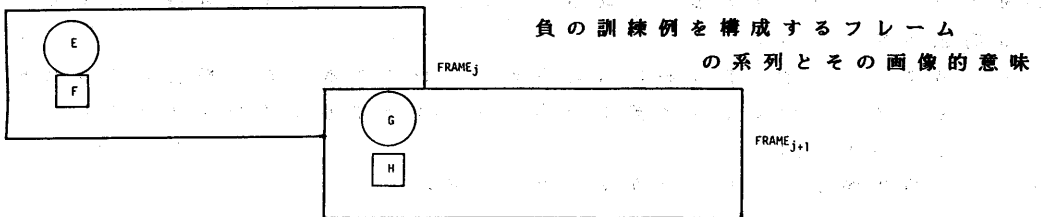
(CIR (A) WHI (A) S-0.039 (A) L-100-60 (A) SQU (B) BLA (B) S-0.020 (B) L-100-125 (B) P-L (A B))

(CIR (C) WHI (C) S-0.039 (C) L-500-60 (C) SQU (D) BLA (D) S-0.020 (D) L-100-125 (D) NOT-TOU (C D))



(CIR (E) WHI (E) S-0.039 (E) L-100-60 (E) SQU (F) BLA (F) S-0.020 (F) L-100-125 (F) P-L (E F))

(CIR (G) WHI (G) S-0.039 (G) L-100-40 (G) SQU (H) BLA (H) S-0.020 (H) L-100-125 (H) NOT-TOU (G H))



***FINDING-OF-GREATEST-LOGICAL-PRODUCT

(CIR-(A)-WHI-(A)-M-(A) SQU-(B)-BLA-(B)-M-(B))

P-T-I

((G-L-U (A B) G-D-65 (A B) P-L (A B)) (G-L-L (B A) G-D-40C (B A) NOT-TOU (A B))

)

正の訓練例を構成するフレーム
の系列のスケルトン

合成された正の訓練例

***FINDING-OF-GREATEST-LOGICAL-PRODUCT

(CIR (E) WHI (E) M (E) G-L-U (E F) SQU (F) BLA (F) M (F))

N-T-I

((G-D-65 (A B) P-L (A B)) (G-D-85 (A B) NOT-TOU (A B))

*****T-I*****

((G-L-U (A B) G-D-65 (A B) P-L (A B)) (G-L-L (A B) G-D-400 (B A) NOT-TOU (A B))

)

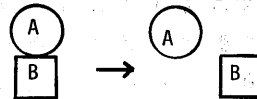
負の訓練例を構成するフレーム
の系列のスケルトン

合成された負の訓練例

*****IS-THIS-RIGHT-MOVE?-Y/N*****

N

N



概念の形成プロセスと
その画像的意味

*****T-I*****

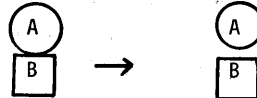
((G-L-U (A B) G-D-65 (A B) P-L (A B)) (G-L-U (A B) G-D-400 (B A) NOT-TOU (A B))

)

*****IS-THIS-RIGHT-MOVE?-Y/N*****

N

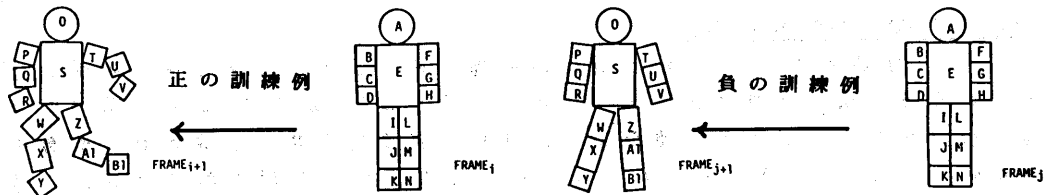
N



形成された概念「(AからCへラベルの変わるプリミティブがある) ∩ (Aの重心はBの重心の左にある) ∩ (AとCの重心間の距離が増大している) ならば、「右へ移動する」という動的概念を意味する。」

((A C) (G-L-L (A C)) (G-D-* (A C)))

図4 "右に移動する"の処理結果



((G-L-U (F G) G-L-U (M N) L-L (B E) S-L-L (B C) S-L-L (G D) L-L (E F) S-L-L (F G) S-L-L (G H) L-L (E I) L-L (E L) S-L-L (I J) S-L-L (J K) S-L-L (I L) S-L-L (J M) S-L-L (K N) S-L-L (L M) S-L-L (M N)) (G-L-L (F G) G-L-L (M N) P-L (B E) P-P (B C) P-L (C D) P-L (E F) P-L (F G) P-L (G H) P-L (E I) P-L (E L) P-L (I J) P-L (J K) P-L (L M) P-L (M N) NOT-TOU (I L) NOT-TOU (J M) NOT-TOU (K N)))

形成された概念

図5 "人が走る"の処理結果

5. あとがき

本論文では、二次元図形世界において、教師との質疑応答を基に自分自身で訓練例を生成し図形構造の概念形成を行うシステムについて述べた。学習状況としては視覚的情報からの例題学習を取り上げている。このような学習状況のとき、システムは教師から与えられる正の訓練例あるいは負の訓練例をもとにして、学習しようとしている概念についての仮説を一般化していかなければならない。ここでは、現在までに提案されている構造的な概念学習システムにおける欠点として教師の負担が大きいことを問題にし、その負担を軽減させるためシステムからユーザへの質問という機能を取り入れた。これにより、教師が訓練選択あるいは訓練の終了条件について考えることは不必要になり、その負担は軽減されるが次のような問題が残っている。

1. 教師によって一番初めに与えられる正負の訓練例が一般化に大きな影響を及ぼす。(教師に対する質問が多くなりすぎて逆に負担を大きくしてしまうことがある。)
2. 一般化が不十分なことがあったりする。
3. 雑音に弱い。
4. S式で表現された訓練例に対する妥当性の判断が困難である。

教師に対する質問を制限することは、プリミティブを特徴づける特徴量の間の関連性に対するヒューリスティックを利用して、ある程度可能である。一般化の不十分さについても領域に特有なヒューリスティックを用いて改善することができるが、し過ぎについてとのトレード・オフが問題になる。今後の課題としては、雑音に関する検討、およびユーザ・インターフェースを高めるためS式を図形表現に変換する図形生成ユニット構築があげられる。さらに、図形として与えられた訓練例を取り扱えるシステムへ拡張する必要があるが、本システムで用いた特徴と同様な視覚的情報を画像から抽出したのち、言語を生成する研究も行われており[5]、拡張は十分可能であると考えられる。

参考文献

- [1] Feigenbaum et al. (eds.), "The handbook of artificial intelligence 1~3", William Kaufmann, 1981~1982.
- [2] Winston, P.H., and Horn, B.K.P., "LISP", Addison-Wesley, 1981.
- [3] Winston, P.H., "The psychology of computer vision", McGraw-Hill, 1975.
- [4] Winston, P.H., "Artificial Intelligence", Addison-Wesley, 1977.
- [5] 高木, 伊東, 六沢, 北岡, 清水, 小原, "二次元図形世界における視覚情報からの日本語文の生成", 信学論(D), J67-D, 2, pp.216-223, 1984年2月