

ON THE LEARNING OF CONCEPTS FROM VISUAL EXAMPLES

メタデータ	言語: en 出版者: 静岡大学大学院電子科学研究科 公開日: 2008-03-31 キーワード (Ja): キーワード (En): 作成者: Held, Andreas メールアドレス: 所属:
URL	http://hdl.handle.net/10297/1210

氏名・(本籍)	HELD ANDREAS (スイス)
学位の種類	博 士 (工 学)
学位記番号	工博甲第 114 号
学位授与の日付	平成 7 年 3 月 24 日
学位授与の要件	学位規則第 4 条第 1 項該当
研究科・専攻の名称	電子科学研究科 電子応用工学専攻
学位論文題目	ON THE LEARNING OF CONCEPTS FROM VISUAL EX- AMPLES (視覚的な例に基づく概念の学習)

論文審査委員	(委員長)		
	教授	竹 内 康 博	
	教授	鈴 木 淳 之	教授 阿 部 圭 一
	助教授	下 平 美 文	助教授 中 谷 広 正
	助教授	山 口 高 平	

論 文 内 容 の 要 旨

The learning of concepts from visual input is a field that draws from the areas of both, computer vision and artificial intelligence. The problems encountered in designing such a system are, therefore, manifold. In the present work, some of those problems are addressed, and a system that learns simple model descriptions from sequences of binary images is proposed. The system consists of several independent modules that are grouped together. As the aim is to develop a system that does not rely on any supervision, the input is provided in the form of images, either acquired by a scanner or any other reading device. Four major modules can be distinguished; they are labelling, description, interpretation, and explanation.

Based on some visual input, which has been restricted to binary images, the labelling stage attempts to extract salient components of the input object. Salient components are assumed to be components that give a concise but meaningful description of some object, meaningful in an intuitive sense. Since human intuition cannot generally be captured by a machine, the concept of maximal approximately convex subparts (MACS) is proposed and it is shown that this notion can represent some of the semantics of an object, solely based on its geometrical structure. The decomposition into MACS is then used to derive a description of the relations between MACS. This description is in the form of a directed graph, called relational network, whose links quantitatively describe relations. To

obtain such quantitative relations, it is proposed to generalize commonly used predicate relations, as, for instance, used in semantic networks, into so-called generalized predicates. Generalized predicates can be viewed as the “fuzzification” of usual predicates, and they quantitatively describe relations among MACS. Each generalized predicate consists of a sum of scaled Gaussian functions; through generalization, this sum of Gaussians will be adjusted to mirror more general relations and thus provide more general class descriptions.

Once a description of an instance has been obtained, it can be matched with previously acquired concepts or object class descriptions. Since the basic structure of an instance or a concept is that of a graph, this matching corresponds to the subgraph isomorphism problem, which is known to be NP-complete. The feasibility of the matching can be restored, however, by employing a heuristically augmented state-space search. Provided a satisfactory match is found, the new instance and the concept are generalized to mirror the enlarged class. Generalization is carried out in two steps: generalization of the structure of the concept, and generalization of the links and the generalized predicates. Generalizing the structure of the concept is straightforward; however, the relational network is now transformed into a conceptual network. The difference between the relational network and the conceptual network is that the latter has a third dimension along which concept discrimination is carried out. Generalization of the predicates, on the other hand, has to ensure that the definition of the generalized predicates is closed under the used generalization operator. In the present system, an adopted version of MYCIN's generalization operator has been used. As the steps of matching and generalizing correspond to identifying an instance with some recorded descriptions, an interpretation of the input image becomes possible.

Finally, the results obtained from generalization have to be explained. As this explanation should be easily understood by a human operator, a description in near natural language is extracted. Such a description is obtained by first transforming the conceptual network into a simple semantic network. Based on the extracted semantic network, it is then relatively easy to produce near natural language, based on a simple set of production rules.

As stated above, the main task of the system is the acquisition of concepts or of model structures. Using the same mechanisms, however, it is as well possible to recognize new instances based on some concept database. Recognition is not complete, though, as it is only possible to recognize or to describe an instance in terms of the recorded concepts. Hence, the approach might be called recognition by experience. On the other hand, even if it is not possible to completely recognize a new instance, partial recognition based on substructures of the recorded concepts might be possible. Finally, employing the hierarchical abstraction of concepts that becomes possible in conceptual networks, certain clusterings of instances can be observed. Those clusterings tend to identify common substructures among several instances. It is those substructures or partial concepts that might give valuable hints during recognition.

論文審査結果の要旨

本論文は、複数の例示図形を与えて、それらの図形から形成される図形概念を学習するシステムの構築法について論じている。

対象とする図形は、シルエットのような2次元の2値画像である。第1章で本研究の位置づけがなされ、第2章において、構築したシステムの概要と、従来の研究との比較の上でそのシステムの特徴が述べられている。このシステムは、分解・記述・学習・概念データベース・自然言語による説明の5つのモジュールから成る。

第3章では、入力された図形を顕著な成分に分解する方法が説明される。まず、近似的な凸性を計る新しい尺度を提案し、その良さを論じている。ついで、これに基づいて、与えられた図形をMACS (Maximal Approximately Convex Subpart) と呼ぶ成分にオーバーラップを許して分解するアルゴリズムを示している。これは、モーフォロジー演算に組み合わせて作られているが、特に、Elongationと呼ばれる新しい演算が用いられている。

第4章では、記述とマッチングが説明される。上の分解結果を用いて、記述モジュールは構成要素間の関係を記述した関係ネットワークを作る。ネットワークの節点は各成分であり、辺はオーバーラップする成分間の関係、すなわち、相対的な長さ・幅、交差角度、重心間の距離を表す。これらの属性は、それぞれ正規分布の和の形の「一般化述語」によって記述される。新たに入力された図形とすでにある概念モデルとのマッチングは、関係ネットワーク・レベルのマッチングと一般化述語の相関との2段階で行われる。

第5章では、マッチングの結果を利用して、新たな入力図形をも含むように概念モデルを拡張する方法が述べられている。このとき、共通する部分概念の抽出により、概念構造の階層化が試みられる。第6章では、人間とのインタフェースを改善するために、概念モデルにたいし自然言語による説明機能を付加している。

第7章では、以上のシステムを実際に図形概念の学習に用いた例が示されており、おおむね予想通りに機能することが確かめられている。第8章には結論と今後の課題がまとめられている。

以上のように、本論文は、図形を例示することにより、それから直接に図形概念を学習するという先例の少ない研究を行っており、MACSへの分解、関係ネットワークと一般化述語による記述とそのマッチング、それに基づく概念モデルの拡張といった点で、有益な提案をしている。また、これに基づいて構築されたシステムは、図形集合の例示にたいし適切な概念モデルを作成しており、学習システムとして成功している。

よって、本論文は画像理解と学習とを結びつけた研究として高く評価することができ、博士(工学)の学位を授与するにふさわしい業績であると認める。