A Syntactical Approach to Learn and Identify Bidimensional Image Models

Miguel Sainz Serra, Alberto Sanfeliu Cortés
Instituto de Robótica e infomática Industrial (CSIC-UPC)
c/Gran Capita 2-4 2a planta, 08034 Barcelona
[msainz-asanfeliu]@iri.upc.es

Abstract

In one hand, automatic generation of models from a set of positive and negative samples and a a-priori knowledge (if available) is a crucial issue for pattern recognition applications. In the other hand, a generic multipurpose 2D object model representation is very useful in object recognition in complex scenes. In this paper we present a new approach of 2D objects multi-purpose model representation based in context sensitive languages and automatic learning. To illustrate the model representation and the performances achieved two different applications have been developed: an outdoor traffic sign identifier and a human face identifier. Partial results of the recognition process of both applications are shown.

Keywords

Automatic learning, Generic multi-purpose model, Grammatical inference, Augmented regular expressions, Context sensitive grammars.

A Syntactical Approach to Learn and Identify Bidimensional Image Models

Abstract

In one hand, automatic generation of models from a set of positive and negative samples and a a-priori knowledge (if available) is a crucial issue for pattern recognition applications. In the other hand, a generic multipurpose 2D object model representation is very useful in object recognition in complex scenes. In this paper we present a new approach of 2D objects multi-purpose model representation based in context sensitive languages and automatic learning. To illustrate the model representation and the performances achieved two different applications have been developed: an outdoor traffic sign identifier and a human face identifier. Partial results of the recognition process of both applications are shown.

Keywords

Automatic learning, Generic multi-purpose model, Grammatical inference, Augmented regular expressions, Context sensitive grammars.

1.- Introduction

The recognition systems in the future must be capable of acquiring objects from real samples with limited human help and then use a generic model description to represent the learned objects. There exist few approaches to automatically acquire generic models. In this paper we deal with a new generic 2D model representation based on grammatical inference methods and a set of positive and negative samples. Two applications of this model representation proposed in [3] are presented. The first application is an identifier of traffic signs in outdoor road scenes and the second one is a human face identifier. Both applications use an automatic model learning method described in [3] and the same recognition process described in [2].

2.- Model representation

The chosen model representation is a two-level context sensitive language called PSeudoBidimensional Augmented Regular Expression (PSB-ARE), extracted from an image Im^i of the object to model, and it consists of a row-by-row description using ARE's[1] and a row-sequence ARE. This PSB-ARE is a data struct composed by the following fields:

1.- The set $\Sigma_{\text{PSB-ARE}} = \{a^0_{\text{PSB-ARE}}, \ldots, a^k_{\text{PSB-ARE}}\}$ corresponding to the model language vocabulary.
2.- A set of FSA_{row} containing all the allowed different FSA that can be used to construct the ARE expressions from the codified image rows.
3.- A set of linear systems matrices that joined with their proper FSA (from the FSA_{row} set), define the ARE_{row} of all the rows from the image Im^i.
4.- A FSA_{model} that describes the skeleton of the sequence of FSA_{row} from image Im^i, starting from the upper row until the lower row.
5.- A matrix to expand the FSA_{model} to an ARE_{model}.
6.- The height-width ratio of Im^i.

In addition to this PSB-ARE data struct, there are some features added to the model representation called *model pattern seeds* (MPS). These MPS are fully described in [3] and they are used in the recognition process to set hypothesis of the size and location of the objects candidates in the scenes.

3.- Recognition process

The recognition process is done in the following steps:

1.- Low level segmentation and Object Of Interest (OOI) location.
2.- Codification of the color pixels into language symbols.
3.- Size hypothesis extraction using the MPS.
4.- Model image generation by solving the ARE linear systems. The size hypothesis is used to solve

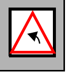
the ARE_{model} and ARE_{rows} and, as result of this, an image Im^{model} of the model is generated.
5.- Matching. This step is done by comparing row by row the object candidate with the model image using the Levenshtein distance.
In Fig. 1 some matching results of the two generated applications are shown.

Traffic Sign identifier

* Models:

M1 M2 M3 M4 M5 M6 M7 M8

* Statistical results:

Models	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
M1	8.8	11.6	11.7	37.7	57.8	70.2	18.4	19.6	16.4	3.5	67.7	17.5	--	--	19.6
M2	20.0	74.5	22.8	72.9	53.6	62.4	72.6	20.4	20.6	9.1	65.7	17.2	10.7	57.4	21.6
M3	68.2	61.5	33.3	66.6	38.4	34.6	16.0	13.2	20.0	15.0	64.1	15.0	7.7	17.1	10.4
M4	27.0	83.9	25.0	45.1	79.6	79.7	76.7	25.4	25.4	27.0	85.4	23.9	19.5	41.6	25.5
M5	14.6	15.5	16.2	13.0	4.8	16.8	15.9	9.1	12.5	54.7	1.2	20.6	17.3	--	2.5
M6	32.2	24.3	24.9	19.5	13.0	15.4	24.3	30.1	85.0	--	34.6	73.7	40.3	23.4	64.2
M7	27.1	31.3	24.0	20.5	27.2	12.0	37.0	21.4	89.1	25.9	26.2	83.5	80.6	32.3	73.5
M8	1.5	1.3	0	13.3	18.9	0.9	0.4	95.5	0.0	--	--	--	--	1.4	7.2
Best	M3	M4	M3	M2	M4	M4	M2/4	M8	M7	M5	M1/2/3	M7	M7	M2	M7
Correct	M4	M4	 80	M2	 70	 70	M2	M8		M5	M3	M7	M7	 80	

Human face identifier

* Models:

M1 M2 M3

* Matching results:

Models	N of source Images	% matching	# of pictures same face diff. shots	# of matches	# of test pictures	# of missmatches
M1	1	95.768	20	16 (80%)	15	0
M2	1	95.980	14	9 (65%)	15	3 (20%)
M3	1	92.647	11	11(100%)	16	3 (18.7%)

Fig. 1 .- Results

4.- References

[1] R. Alquezar and A. Sanfeliu, "Augmented regular expressions: a formalism to describe, recognize and learn a class of context -sensitive languages", *Pattern Recognition* (In press) (1997).

[2] Sanfeliu, A.; Sainz, M.: *Automatic recognition of bidimensional models learned by grammatical inference in outdoor scenes*, Proceedings of the IAPR Workshop on Structural and Syntactical Pattern Recognition, Leipzig 20 -23, August .(1996).

[3] Sainz, M.; Sanfeliu, A.: *Learning bidimensional context dependent models using a context sensitive language*, Proc. of the 13th International Conference on Pattern Recognition, Viena, 25-29 August 1996, Vol IV, Track D pp. 565-569.