Pattern Recognition Letters 33 (2012) 331–341

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Symbol recognition using spatial relations

K.C. Santosh [a,*], Bart Lamiroy [b], Laurent Wendling [c]

[a] LORIA – INRIA, 615 rue Jardin Botanique, 54600 Villers-lès-Nancy, France
[b] LORIA – Nancy Université, BP 239 – 54506 Vandoeuvre-lès-Nancy Cedex, France
[c] LIPADE, Université Paris Descartes, 75270 Paris Cedex 06, France

ARTICLE INFO

Article history:
Received 16 March 2011
Available online 25 October 2011
Communicated by A. Koleshnikov

Keywords:
Vocabulary
Spatial relations
Attributed relational graph
Symbol recognition

ABSTRACT

In this paper, we present a method for symbol recognition based on the spatio-structural description of a 'vocabulary' of extracted visual elementary parts. It is applied to symbols in electrical wiring diagrams. The method consists of first identifying vocabulary elements into different groups based on their types (e.g., *circle*, *corner*). We then compute spatial relations between the possible pairs of labelled vocabulary types which are further used as a basis for building an attributed relational graph that fully describes the symbol. These spatial relations integrate both topology and directional information.

The experiments reported in this paper show that this approach, used for recognition, significantly outperforms both structural and signal-based state-of-the-art methods.

1. Introduction

1.1. Motivation

Symbol recognition – the core part of graphical document image analysis and recognition systems – plays an important role in a variety of applications such as automatic recognition and interpretation of circuit diagrams (Okazaki et al., 1988), engineering drawings (Yang et al., 2007) and architectural drawings (Lladós et al., 2001; Valveny and Martí, 2003), maps (Samet and Soffer, 1996), musical notations (Rebelo et al., 2010), mathematical expressions (Chaudhuri and Garain, 2000), as well as optical characters (Yuen et al., 1998). Therefore, a symbol can be defined as a graphical entity with a particular meaning in the context of a specific domain.

Research on graphics recognition has an extremely rich state-of-the-art literature, aimed to localise/recognise symbols depending on the applications. (Cordella and Vento, 2000; Lladós et al., 2002) show that these methods are particularly suited for isolated line symbols, not for composed symbols connected to a complex environment. In order to exploit the information embedded in those documents, one needs to be able to extract visual parts and formalise the possible links that exist between them. This combination of symbol localisation based on extracted visual parts is going to be the core of this paper and is very much inspired by a real world industrial problem (Tombre and Lamiroy, 2008; Santosh et al., 2009). It consists in identifying a set of known symbols in aircraft electrical wiring diagrams, in order to bootstrap simulation algorithms. The main challenges come from the fact that the test symbols come in a wide variety of different forms. Symbols may either be very similar in shape, and only differ by slight details or either be completely different from a visual point of view. Symbols may also be composed of other known and significant symbols and need not necessary be connected.

The rest of the paper is organised as follows. An overview of pertinent literature is given in Section 1.2, followed by a brief explanation of our proposed method in Section 2. We explain the way we describe symbols in Section 3, which mainly includes the concept of using spatial relations. We derive a symbol matching method from it in Section 4. Full experiments are reported in Section 5 and confront our method with current state-of-the-art algorithms. It includes a comprehensive experimental result analysis. We conclude in Section 6.

1.2. State-of-the-Art

1.2.1. Symbol representations

Symbol recognition is a particular application of pattern recognition. Existing approaches, specifically those based on feature based matching, can be sorted into three classes: statistical, structural and hybrid. As respective examples , among others, one can cite (Yang, 2005; Zhang et al., 2006; Lladós et al., 2001; Yang, 2005).

Under statistical approaches, global signal based descriptors (Yuen et al., 1998; Kim and Kim, 2000; Tabbone et al., 2006; Belongie et al., 2002; Zhang and Lu, 2002, 2004) are usually quite fault tolerant to image distortions, since they tend to filter out small

* Corresponding author.
E-mail addresses: Santosh.KC@inria.fr (K.C. Santosh), Bart.Lamiroy@loria.fr (B. Lamiroy), Laurent.Wendling@parisdescartes.fr (L. Wendling).

0167-8655/$ - see front matter
doi:10.1016/j.patrec.2011.09.040

detail changes. This is unfortunately an inconvenient in our context. Moreover, they difficultly accommodate with connected or composite symbols. For instance, when symbols are combined, approaches that rely on centroid detection like (Yuen et al., 1998) tend to fail. Others, like shape context (Belongie et al., 2002) are sensible to occlusions on the symbol boundaries. Overall, they are generally not well adapted for capturing small detail changes, since they are specifically conceived to filter those out. In these statistical approaches, signatures are simple with low computational cost. However, discrimination power and robustness strongly depend on the selection of optimal set of features for each specific application.

Besides global signal based descriptors, another idea is to decompose the symbols into either vector based primitives like points, lines, arcs etc. or into meaningful parts like *circles*, *triangles*, *rectangles* etc. These methods fall under structural approaches. They are then represented as attributed relational graphs (ARG) (Bunke and Messmer, 1995; Conte et al., 2004), region adjacency graphs (RAG) (Lladós et al., 2001), constraint networks (Ah-Soon and Tombre, 2001) as well as deformable templates (Valveny and Martí, 2003). Their common drawback comes from error-prone raster-to-vector conversion. Those errors can increase confusions among different symbols. Furthermore, variability of the size of graphs leads to computational complexity in matching. However, structural approaches provide a powerful representation, conveying how parts are connected to each other, while also preserving generality and extensibility.

To describe the symbols, it is necessary to handle relations between the decomposed parts. The following paragraph gives an overview of existing work on spatial relations and their proper usages.

1.2.2. Spatial relations

Effects of spatial relations on recognition performance have been examined comprehensively for scene understanding, document analysis and recognition tasks (Biederman, 1972; Bar and Ullman, 1993; Xiaogang et al., 2004; Pham and Smeulders, 2006). Spatial relations can be either topological (Egenhofer and Franzosa, 1991; Egenhofer and Herring, 1991; Papadias et al., 1995) directional (Bloch, 1999; Matsakis and Wendling, 1999; Wang and Keller, 1999) and metric in nature. For example, topological configurations are handled in (Xiaogang et al., 2004) with a few predicates like *intersection*, *interconnection*, *tangency*, *parallelism* and *concentricity* expressed with standard topological relations as described in (Egenhofer and Herring, 1991).

In a similar way, various directional relation models have been developed for a wide range of different situations.

- If the objects are far enough from each other, their relations can be approximated by their centres based on the discretised angle (Miyajima and Ralescu, 1994). This approach is robust to small variations of shape and size.
- If they are neither too far nor too close, relations can be approximated by their *minimum bounding rectangle* (MBR) (Lee and Hsu, 1992; Jungert, 1993; Papadias et al., 1995; Papadias and Theodoridis, 1997) as long as they are regular.
- Approaches like *angle histograms* (Wang and Keller, 1999) tend to be more capable of dealing with overlapping, something the previous ones have difficulties with. However, since they consider all pixels of a shape, their computational cost increases dramatically.
- Other methods, like *F-Histograms* (Matsakis and Wendling, 1999) use pairs of longitudinal sections instead of pairs of points, also at the cost of high time complexity.
- Another well-known approach uses fuzzy landscapes (Bloch, 1999), and is based on fuzzy morphological operators.

Previously mentioned approaches address only either topological or directional relations. Managing both comes at high computational costs. Even then, no existing model fully integrates topology. They rather have various degrees of sensitivity to or awareness of topological relations. While methods like (Xiaogang et al., 2004) focus on topological information only, our approach unifies both topological and directional information into one descriptor (Santosh et al., 2010) without any additional running time cost.

Placing spatial relations in the context of recognition and symbol description, one should note that spatial relations also have a language-based component (related to human understanding e.g., to the *right* of) that can be formalised in a mathematical way (e.g., the 512 relations of the 9-intersection model Egenhofer and Herring, 1991). Therefore, qualitative and quantitative relations are another way to do categorisation of spatial relations. Consider an example, an object $\mathbb{A}$ extending from *Right* (98%)–*Top* (2%) with respect to $\mathbb{B}$ is expressed as *Right–Top* $(\mathbb{A}, \mathbb{B})$. This spatial predicate remains unchanged upto a reasonable change of the objects' shape and position. Taking this into account, our work uses more natural relations than the all-or-none nature of standard relations (Freeman, 1975).

In the following section, we explain our proposed method by focusing on using spatial relations for describing and matching symbols.

2. Proposed recognition method

Our recognition method is based on a spatio-structural description of extracted visual parts that compose a symbol. This means that, to describe a symbol, we compute spatial relations between previously extracted visual parts. Without any other consideration, it is obvious that the size of the resulting relational graph is potentially very large and variable from one symbol to another. However, when grouping visual parts together according to their types (e.g., *circle*, *corner*) and by labelling them accordingly (see Section 3.1), we can eliminate all the combinatorial problems inherent to graph matching, without sacrificing recognition quality or expressive power.

We compute the spatial relations (see Section 3.2) between the distinct labelled attributes for building an attributed relational graph (ARG – see Section 3.3), achieving at the same time integration of both topological and directional information.

Since each vertex represents a different class of visual parts, the graph has a uniquely and distinctly labelled vertex set. Vertex and edge matching thus becomes trivial and can be done in near-constant time.

3. Symbol description

As mentioned in Section 2, we first define our visual vocabulary in Section 3.1. In Section 3.2 we explain the way we compute pairwise spatial relations and finally use both in Section 3.3 to build an ARG and completely describe the symbol.

3.1. Visual vocabulary

We define a set of well controlled visual elementary parts as a *vocabulary* (Santosh et al., 2009). While, in the general case, this vocabulary can be of any kind from any type of bags-of-features, related to what is visually pertinent in the application context under consideration, our current vocabulary is related to electrical symbols. It can be easily extended to adapt to other domains. Such visual elementary parts are extracted with the help of image treatment analysis operations as described in (Rendek et al., 2004). Shortly, we discuss on how we accomplished it.