Defining Writer’s Invariants To Adapt the Recognition Task

Ali Nosary  Laurent Heutte  Thierry Paquet  Yves Lecourtier
Laboratoire Perception, Systemes, Information (PSI)
Universitè de Rouen
F76821 Mont-Saint-Aignan cedex, France
E-mail: Ali.Nosary@univ-rouen.fr

Abstract

This work investigates the automatic reading of unconstrained omni-writer handwritten texts. It shows how to endow the reading system with adaptation faculties to each writer’s handwriting. The adaptation principles are of major importance to make robust decisions when neither simple lexical nor syntactical rules can be used e.g. for free lexicon and full text recognition. The first part of this communication defines the concept of writer’s invariants. In the second part we explain how the recognition system can be adapted to the current handwriting by exploiting the graphical context defined by the writer’s invariants. This adaptation is guaranteed thanks to the writer’s invariants, by activating interaction links over the whole text between the recognition procedures of word entities and those of letter entities.

Keywords: Omni-writer handwritten text recognition, writer’s invariants, variability measure, adaptation, interaction.

1. Introduction

Making machines learn to read omni-writer handwritten texts requires, on the one hand, sophisticated and highly adapted algorithms of pattern recognition and, on the other hand, the management of the various interpretation levels, from the graphical to lexical levels, by trying to exploit the human reader expertise. Therefore, as a human reader may do, an automatic reading system must be able to meet two different requirements. It should have omni-writer capabilities in order to recognise any handwriting. It should also have mono-writer capabilities in order to take into account the potential fantasy of each writer. Conventional systems which consider the recognition of omni-writer handwritten texts as a problem of character/word recognition have been pushed to their limits, as they try to recognize handwritten words one independently from the others in a sequential manner [8] [7]. In general, a segmentation of words into pseudo-characters or graphemes is first performed; even if word segmentation still remains a hard problem, it has been extensively studied and several efficient algorithms have been proposed [2]. Note that sophisticated algorithms for cursive handwritten word recognition have also been proposed in the last few years to cope with the problem of word segmentation [1]. Then pseudo-characters are recognized using either well-adapted features or multiple classifier combination techniques [6]. Finally, the independent word hypothesis are analysed in a post-treatment stage using syntactical constraints [10]. Let us recall that conventional omni-writer systems must be trained on large databases in order to learn the inter-writer properties even if recent studies have tried to counteract the problem of handwriting variability by classifying handwritings into families of handwriting styles to further select specific recognizers [4]. However, all of these systems still only use the inter-writer properties (either globally or within a particular group of writers).

In this communication we introduce the concept of adaptation which allows to make robust decisions when neither simple lexical nor syntactical rules can be used as, for example, in the case of free lexicon unconstrained handwritten text recognition. We first define the concept of writer’s invariants that allow to derive new contextual graphical knowledge. We show that the writer’s invariants allow to define a variability measure of the handwriting to be recognized. Furthermore, this variability measure allows to evaluate the level of legibility of handwritten texts. In the second part of this communication we show how the adaptation of the recognition system to the writing can be carried out thanks to the writer’s invariants.
2. Writer’s Properties

The fundamental property of handwriting, which makes written communication possible, is that there exist inter-writer invariants since the morphological differences between patterns representing distinct letters are more important than those between different alphabets of a same letter (inter-writer variability). We postulate now that each writer draws the same letters using the same patterns (his own handwriting references): we call these references "the writer’s invariants".

2.1. Writer’s Invariants

The writer’s invariants, reflecting the morphological redundancy of his handwriting, can be defined as the set of similar patterns or graphemes extracted from the segmentation of his handwriting using a particular segmentation technique.

The detection of the morphological invariants is performed using an automatic classification of graphemes issued from the segmentation of the handwritten text. The graphemes are first normalized into a fixed size. Among the different clustering methods, a simple and fast sequential clustering algorithm that does not need to know a priori the number of clusters has been retained [5]. One characteristic of this sequential clustering method is that the choice of the templates depends on the order of presentation of the elements in the grapheme set. To cope with this problem, multiple sequential clusterings are iterated with a random selection order of the elements. Finally only the elements that are always grouped together, when iterating the clustering procedure, are retained to constitute a cluster, while elements that are not always assigned to the same group constitute singular elements e.g. clusters with a single element. This algorithm requires, as any clustering method, an index of proximity and an adequate choice of the regrouping threshold. The number of generated invariant clusters is a function of the selected threshold. Among the several types of proximity measures tested, the correlation similarity measure has been retained. Figure 1 gives an example of the clusters obtained after four iterations of a sequential clustering.

As one can see, the proposed method succeeds in determining regular patterns that occur in the text. Particularly, one can remark on the example that the writer uses two different alphabets to draw letter a. Of course, this primary result does not allow to evaluate the robustness of the method over a large set of handwritten texts. We show in the following paragraph that the concept of writer’s invariants allows to order a sample of handwritten texts according to their respective level of stability by introducing a new variability measure.

2.2. A New Variability Measure of Handwriting

In order to evaluate the quality of the obtained writer’s invariants for each handwritten text, it is necessary to be able to assess the quality of grapheme clusters. But, in our unsupervised context, this evaluation depends on the handwriting quality as well as on the clustering parameters. Therefore, for fixed parameters, we have decided to evaluate the correlation between the handwriting quality of a handwritten text and the obtained clusters. We thus define a variability measure for a handwritten text $T$ as follows: assume that the handwritten text has been segmented into $N$ graphemes; assume also that these graphemes have been grouped into $n$ different clusters (the writer’s invariants) and that a cluster $C_i$ contains $n_i$ graphemes; then we define the variability measure as:

$$E(T) = \frac{\sum_{i=1}^{n} P(C_i) \log_2 \frac{P(C_i)}{\log_2 N}}{N}$$

where $P(C_i) = \frac{n_i}{N}$

This measure is derived from Shannon’s entropy measure where the denominator ensures that: $0 \leq E(T) \leq 1$.

Therefore, $E(T) = 1$ means that there are as many clusters as graphemes (in the case of absolute disorder). On the contrary, $E(T) = 0$ means that all the graphemes are grouped in a single big cluster. In practice, we show that any writing has a different amount of redundant patterns distributed in different clusters: the variability measure will thus range from a minimum value for very stable writings up to a maximum value for very distorted writings.

Experiments have been conducted on a database of 75 handwritten pages corresponding to the same text and written by 75 different writers. The segmentation of each handwritten page into graphemes generates around 500 and 800 graphemes. For each handwritten
3.1. Specific Knowledge Modeling

Considering that the whole text of the writer has been segmented into graphemes using well-known techniques encountered in the literature [2], each grapheme (corresponding to a letter or not) is characterized by:

**Intrinsic Morphological Knowledge (IMK):** any knowledge that can be extracted from the grapheme pattern alone, such as a set of features detected on the grapheme image for example.

**Contextual Morphological Knowledge (CMK):** any knowledge about the grapheme pattern that can be extracted from its environment, such as the invariant cluster the grapheme belongs to.

Now the following symbolic knowledge about each grapheme can be provided by different treatments:

**Intrinsic Symbolic Knowledge (ISK):** any knowledge about the possible letter (label) that can be associated to the grapheme considered alone (e.g. obtained from IMK) using classical recognition schemes that exploit inter-writer invariants.

**Contextual Symbolic Knowledge (CSK):** any knowledge about the possible letter that can be associated to the grapheme by referring to its context. For example symbolic knowledge about a grapheme can be derived from the invariant cluster it belongs to using the hypothesis made about its neighbors. Symbolic knowledge can also be derived thanks to the use of the lexical constraints applied to the word the grapheme belongs to.

3.2. Exploiting Contextual Knowledge

An attempt to illustrate how a recognition system can exploit this knowledge at word and grapheme levels is shown in figure 3. Assume that handwritten words have already been localized and that the segmentation into graphemes has been performed for each of them. Assume also that IMK and CMK have been extracted.
for each grapheme. Then the following links can be activated at the grapheme level. a) A character recognition procedure can provide ISK to each grapheme. b) The ISK of each grapheme can activate word level procedure. c) CSK for each grapheme can be derived from lexical constraints applied at word level. d) CSK of each grapheme can also be derived from its morphological neighbors (the invariant cluster (Ci) it belongs to) e) Global CSK of each grapheme can provide symbolic hypothesis for a writer invariant. f) A coherent analysis of each invariant cluster can reinA the similar letter hypothesis for the similar patterns.

**Figure 3.** Illustration of the role of the writer’s invariants in the interactive recognition system

Assume for example that a lexical analysis cannot disambiguate the letter hypothesis e and l for the grapheme_6. Then thanks to the writer’s invariants it is possible to refer to the letter hypothesis made on grapheme_6 that belongs to the same cluster but occurs in a different lexical context (words). Then since there is no ambiguity in letter hypothesis of grapheme_6, due to its lexical context, letter hypothesis on grapheme_6 can be disambiguated by means of the writer’s invariants. The activation links described above provide a general framework that can be used to implement various strategies in the reading system. Depending on the strategy used, a global coherence of the recognition hypothesis can be reached at each of the two interpretation levels (Word, Grapheme). Note that the same principles of interaction could be applied between text level and word level thanks to the use of syntactical constraints.

4. Conclusion

We have introduced in this paper the concept of writer’s invariants which can be defined as the set of similar patterns extracted from the segmentation of a handwriting. We have also shown the pertinence of this concept over a sample of handwritten texts using a variability measure which allows to evaluate the quality of the handwriting to be recognized. We have also discussed how a reading system could adapt itself during the recognition to the current handwriting by introducing interaction between treatments at different contextual levels (graphical, symbolic, lexical). Note that the concept of interactivity has already been investigated in [9] [3]. In a sense, we can view our model as an extension of this concept. Future works on our model will focus on its implementation using the multi-agent paradigm on a distributed system architecture.

References