Study of Perceptual Similarity between Different Lexicons

Cinthia O. A. FREITAS[1], Flávio BORTOLOZZI[1] and Robert SABOURIN[2]
[1]Pontifícia Universidade Católica do Paraná (PUCPR), Graduate Program in Computer Science, Brazil.
cinthia@ppgia.pucpr.br, fborto@ppgia.pucpr.br
[2]École de Technologie Supérieure (ETS), Lab. d'Imagerie, de Vision et d'Intelligence Artificielle, Canada.
robert.sabourin@etsmtl.ca

Abstract. Our study investigates the perceptual feature similarity between different lexicons based on visual perception of the words and theirs representation through an observation sequence. We confirm that it is possible to apply databases, which are similar in terms of morphological/perceptual features for improving the recognition task. Having selected Portuguese and French words from legal amounts for this study, we improve the recognition rate of handwritten Portuguese words by adding samples of French words to the training database. The error rate obtained in the meta-class called "dezena"/"teen" is 36.10% lesser than that obtained in previous experiments reported in the work of Freitas et al. (2001).

1. Introduction

Reading has been described as a language code picked up through the visual or tactile system and then processed further, and a procedure involving an assembly of human activities. These activities include visual sensory input, accurate eye movements, and higher cognitive aspects of comprehension, reflecting the complexity of reading (Freitas et al., 2001, 2002). Perceptual features are visually important features of the word shape that have been cited in reading studies as being utilized in fluent reading (Madhavanath & Govindaraju, 1998, 2001). These features have been used for word recognition when the lexicon of possible words is small and static.

Normally, for recognition experiments, we do not use different databases collected in different situations, scanned through different machines, with writers of different nationalities. Our study investigates the perceptual feature similarity between different lexicons. These similarities were studied based on word representation through an observation sequence. We confirm that it is difficult, during these experiments, to apply some databases, which are different in terms of morphological or perceptual features.

Having selected Portuguese and French for this study, our idea is to demonstrate, based on experimental results, that it is possible to improve the recognition rate of handwritten Portuguese words by adding samples of French words to the training database based on their perceptual similarities. The scope of this study is limited to the off-line recognition of individual handwritten words from legal amounts. This paper presents the experiments carried out to validate the proposed hypothesis.

The paper is divided into seven sections. Section 2 presents the most important visual perception concepts. Section 3 summarizes the handwritten Portuguese word recognition problem. Section 4 explains the hypothesis to be analyzed and discusses the perceptual feature similarity of Portuguese and French words. Section 5 presents the databases used in the experiments and describes the word recognition method based on HMM. In Section 6, the experimental results are discussed, and, in the final section, our conclusion and future works are presented.

2. Visual Perception Concepts

We present in this paper only a summary about the visual perception concepts related to handwritten word recognition. The Gestalt Theory is beyond the scope of this paper. An excellent introduction to this subject can be found in Sekuler & Blake (1994). This theory describe some principles of organization, that is, these principles identified factors that tend to encourage the emergence of perceptual forms and promote the grouping of those forms, segregated form their surroundings.

Visual perception is a dynamic process and involves the observer and the object that is been observed. When the persons are looking at an object or a scene, they analyze the structure, solve ambiguities and do connections. In general way, the persons organize what they see and the human eye is the sense organ responsible by lecture, besides the fingers touching used by the visual incapable persons to "read" the Braille text. Summarizing the visual process, the eyes work as a sophisticated photographic machine carrying the visual information to the brain, where it will be processed.

When the persons perceived a visual field, the pattern that emerges as a figure, and not as the background, depends of the characteristics from the field and the relation among the objects inside the field. Therefore, the perception of the form is linked with the best relation between the form and our brain, as shown in Figure 1, where some letters "A" have a easer perception, that is, the letters are more interpretable or readable.

The human being has a tendency to interpret a visual stimulus as complete. This tendency is known in Gestalt theory as closure concept (Sekuler & Blake, 1994). Thus, you can read words and letters if these are poorly written, if not missing. For example, the first letter in the alphabet, letter "a", can appear in different writing styles, as shown in Figure 1. We can say, in terms of perceptual features, that the letter "A" is so different from "a", how "p" is different from "m". Moreover, "p", "b" and "d" are most similar between themselves than "b" and "B", "d" and "D". For human reader, the difference is not perceived each instant during the writing or reading process, for an adult person, an "A" is an "A", independently how it is written. Remember that the Gestalt principle of similarity referring to the perceptual tendency to group together objects that are similar to one another in texture, shape, and so on (Sekuler & Blake, 1994). For our study and according to Sekuler & Blake (1994) because only particular properties promote grouping, these properties may constitute the basic elements of perception. We call these visual elements *features*.

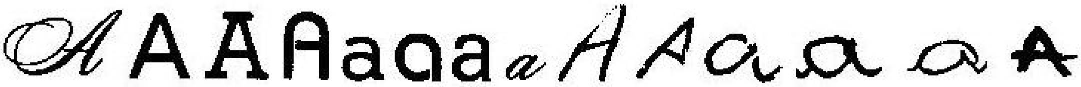

Figure 1: Letter "A" in different styles: printed and handwriting

3. Handwritten Portuguese Word Recognition

A "legal amount" corresponds to a numerical value which obeys a known grammar. The database comprises values between R$ 0,01 ("um centavo") and R$ 999.999,99 ("novecentos e noventa e nove mil, novecentos e noventa e nove reais e noventa e nove centavos"). From the numerical value, it is possible to define five meta-classes of words, such as: **1)"entos"/hundred:** words corresponding to numbers 100, 200, 300, 400, 500, 600, 700, 800, 900; **2)"enta"/"ty":** words corresponding to numbers 20, 30, 40, 50, 60 ,70, 80, 90; **3)"dezena"/"teen":** words corresponding to numbers 10, 11, 12, 13, 14, 15, 16, 17, 18, 19; **4)"unidade"/unity:** words corresponding to numbers 1, 2, 3, 4, 5, 6, 7, 8, 9, as shown in Figure 2; **5)key words (KW):** "mil", "reais" or "real" and "centavos" or "centavo".

We can also observe in Figure 2 the similarity of the suffixes and prefixes of the words in the lexicon, a fact which increases the complexity of the recognition problem. Observing, the Portuguese lexicon distribution of the training samples, the meta-class "dezena"/"teen" represents only 3.16% of the training database for the Portuguese lexicon. As a result, we have only a few samples for the HMM training models, and this is the principal explanation for the recognition rate observed in this meta-class by Freitas et al. (2001), see Table 2 (**I:** PFCCD).

Our word database, called PUCPR/LUCI, is a laboratory database which was generated based on a numerical value for the courtesy amount. Consequently, the words from meta-class "dezena"/""teen" are only written by the writer when the number at the middle position in the courtesy amount is the number "1". Otherwise, the words will belong to the "enta"/"ty" meta-class. This is an appropriate reason for collecting a few samples for this meta-class of words. Another important observation concerns the writing style distribution in PUCPR/LUCI, where the cursive is the writing style most frequently (72%) encountered in the training database. It is important to note that we specified our feature set for the cursive style knowing that another feature extraction approach is required for the block-print style.

4. Hypothesis: Perceptual Feature Similarity

The hypothesis analyzed in this paper is the following: Is it possible to improve handwritten Portuguese word recognition by using French words? Taking the considerations presented in Section 3 into account, our objective is to improve the word recognition rate in cases where the training database contains only a few examples for obtaining the HMM models. For us, this means improving the training of the models by inserting words from the French database. We insert French words because we are analyzing the perceptual feature similarity of the Portuguese and French lexicons. Even when objects do not appear to be identical, they may still resemble one another to varying degrees. For instance, you may be able to judge whether the baby in the carriage looks more like its mother or its father (Sekuler & Blake, 1994).

Now, considers one class of complex forms that you are constantly judging, letters of the alphabet. Reading would be very difficult indeed if all letters of the alphabet looked very much alike. So, the perceptual similarity of letters was defined in (Sekuler & Blake, 1994) based on their tendency to be confused. In the feature approach, it is necessary first to define a list of features, that is, to decide what features should make up the list.

For us, perceptual similarity presupposes that the ascenders, descenders and loops occur in the same grapheme in the words, and that the perceptual feature extraction procedure is stable, independent of the database. These similarities are presented in Table 1, which shows the positions of the perceptual features. Moreover, in support of our hypothesis, we observed the following: **a)** The languages (Portuguese and French) have the same origin: Latin; **b)** The lexicons contain some identical words: “onze”, “quatorze” and “quinze”, **c)** The lexicons contain some identical morphemes: “*qu*at*r*o” and “*qu*at*r*e”, “*qu*arenta” and “*qu*arente”, “cin*qu*enta” and “cin*qu*ante”. A morpheme is a linguistic element, such as “qu” and “tr”, **d)** The lexicons contain some words which are similar in terms of perceptual features (ascenders, decenders and loops): “três” and “trois”, “doze” and “douze”, “treze” and “treize”, “trinta” and “trente”, “sessenta” and “soixante”.

So, the idea is to work at a high level of representation: the observation sequence extracted from the word images based on the feature set. It is important note that we are merging the representations of two different databases based on features extraction. Our feature extraction procedure is based on perceptual features (ascenders, descenders and loops) and on concavity and convexity deficiencies (PFCCD). Concavity and convexity deficiencies in the word body are extracted and labeled as presented in Freitas et al.(2001). These deficiencies are obtained by labeling the background pixels of the input image. The entire alphabet is composed of 29 different symbols.

Based on feature extraction and visual perception, we intend to explore the shape information contained in this feature set and their consonant representation (Table 1: *t*, *q*, *z*, *d*). Another aspect is that, given that the first letter of the word is very important in the recognition process (Schomaker & Segers, 1998), in our case the similarity of the lexicons is increased since the selected words (Portuguese and French) have the same first letter (Table 1). We also intend to explore the fact that the vowels exhibit the same behavior for the reader, as observed by Schomaker & Segers (1998) and as expected for our system.

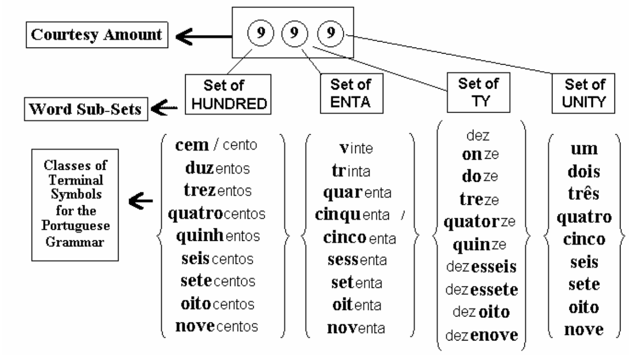

Figure 2: Brazilian legal amounts: meta-classes of words

Table 1: Perceptual Feature Similarity

Portuguese	French	Portuguese	French
três	trois	quinze	quinze
quatro	quatre	trinta	trente
onze	onze	quarenta	quarante
doze	douze	cinquenta	cinquante
treze	treize	sessenta	soixante
quatorze	quatorze		

5. Word Recognition Method and Databases

The experiments were carried out using the following databases: **a) Portuguese lexicon:** handwritten word database called PUCPR/LUCI, consisting of 39 classes of words, 300 dpi, 256 gray levels (Freitas et al., 2001) and **b) French lexicon:** handwritten word database from IRESTE/University of Nantes (France), called IRONOFF (IReste ON/OFF Dual Database), consisting of 29 classes of words from Form B: B64-B93 fields (Viard-Gaugin, 1999) and (Viard-Gaugin et al., 1999), 300dpi with 8 bits per pixel.

The Hidden Markov Model (HMM) theory has been successfully used to model writing variability. The theoretical formulation of HMM is beyond the scope of this paper. An excellent introduction to this subject can be found in Rabiner & Juang (1993). It correctly integrates different modeling levels (morphological, lexical, syntactical) and also provides efficient algorithms to determine an optimum value for the model parameters. Our HMM word models are based on a Global Approach and a *Bakis Topology*. The lexicon size allows us to consider one model for each word class. Model training is based on the Baum-Welch Algorithm and the Cross-Validation process (Rabiner & Juang, 1993). The objective of the Cross-Validation process is to monitor the general outcome during the training process. We also evaluate the use of the *a priori* probability of each word class $p(\lambda_i)$ in the training database during the recognition process (Freitas et al., 2002).

6. Experimental Results

Our experiments take into account the following conditions: **a) HMM model training:** the Portuguese database is merged with the French database considering the words presented in Table 1. Thus, we are modifying 11 HMM models and **b) HMM model validation and recognition experiments:** Portuguese validation and test database.

The previous results reported in Freitas et al. (2001) (**I** and **II**) and the current results obtained (**III** and **IV**) are presented in Table 2. The improvement achieved in the recognition and error rate for each meta-class of words is shown in Table 3. The global improvement (1.43%) is not very much higher, but we did obtain an improvement for all meta-classes. A significant improvement in the recognition rate is observed for the "dezena"/"teen" meta-class (+ 17.33%), showing, as expected, a better HMM training. The error rate obtained in this meta-class is 36.10% lesser than that obtained in previous experiments reported in Freitas et al. (2001) and the error rate in average is 3.96% lesser than before (see Table 3).

In carrying out the recognition procedure over the validation database, we performed an error analysis. We observed improvements with 22 words and a loss with 7 words, while 10 words maintained their rates. It is important to remember that only 11 HMM models were modified, demonstrating that we did, in fact, improve the cursive handwritten recognition rate. By visual analysis, it was possible to verify that 80% of the easily recognized images are cursive.

Based on these results, we can conclude that the French database selected (IRONOFF) was appropriate for our study. We can also conclude that the major improvement for the "dezena"/"teen" meta-class derives from the enlargement of the training database. We can now compare this meta-class on the same basis as the others. With the results obtained, we can say that maximum performance can be achieved based on the perceptual feature set from the Portuguese database for cursive writing.

Table 2: Recognition Experiment (%)

Experiments	TOP 1	TOP 3	TOP 5
I: PFCCD	67.66	86.65	92.21
II: PFCCD + pwc	**70.61**	**88.08**	**92.84**
III: Portuguese + French	69.09	88.00	93.52
IV: Portuguese + French + pwc	**71.96**	**89.45**	**94.17**

pwc = *a priori* probability of each word class

Table 3: Improvement by Meta-Class

Meta-Class	I (%)	III (%)	Recognition Rate(%)	Error Rate(%)
Unity	70.53	71.40	+ 0.87	- 2.95
Teen	52.00	69.33	+ **17.33**	**- 36.10**
Ty	56.12	58.22	+ 2.10	- 4.79
Hundred	64.05	64.27	+ 0.22	- 0.61
KW	78.26	78.86	+ 0.60	- 2.76
Average	**67.66**	**69.09**	+ 1.43	- 3.96

7. Conclusion

In this paper, we explored the perceptual feature similarity of two different lexicons: Portuguese and French. Normally, for recognition experiments, we do not use different databases collected in different situations, scanned through different machines, with writers of different nationalities. However, the experiments have shown the viability of our approach, which focuses on perceptual feature similarity. Moreover, these similarities were studied based on word representation through an observation sequence. We confirm that it is difficult, during these experiments, to apply some databases which are different in terms of morphological or perceptual features, e.g. English and French databases. Future work will provide a complementary feature set study.

References

Freitas, C. O. A., Bortolozzi, F., Sabourin, R. Handwritten Word Recognition: An Approach based on Mutual Information for Feature Set Validation, IEEE Sixth International Conference on Document Analysis and Recognition, 2001, 665-669.

Freitas, C.O. A. Visual Perception and Handwritten Word Recognition (Percepção Visual e Reconhecimento de Palavras Manuscritas). Titular Professor Monograph, PUCPR, Brazil, 2002.

Madhavanath, S. & Govindaraju, V. Preceptual features for off-line handwritten word recognition: a framework for prediction, representation and matching. *Advances in Pattern Recognition*, august, 1998, 524-531.

Madhavanath, S. & Govindaraju, V. The role of holistic paradigms in handwritten word recognition. *IEEE Transactions on Pattern Analysis and machine Intelligence*, Vol. 23, No. 2, February, 2001. 149-164.

Sekuler, R. & Blake, R. Perception. 3rd ed. McGraw-Hill, Inc. 1994.

Schomaker, L. and Segers, E. A method for the determination of features used in human reading of cursive handwriting. In *6th International Workshop on Frontiers in Handwriting Recognition*, 1998, 157-168.

Viard-Gaudin, C. The ironoff user manual. IRESTE, University of Nantes, France, 1999.

Viard-Gaudin, C.; Lallican, P.M.; Knerr, S.; Binter, P. The ireste on/off (ironoff) dual handwriting database. *IEEE International Conference on Document Analysis and Recognition*, 1999. 455-458.

Rabiner, L., Juang, B.H. Fundamentals of Speech Recognition. Prentice Hall Inc., 1993.