IJDAR (2001) 3: 196–206

International Journal on Document Analysis and Recognition

Industrial bank check processing: the *A2iA CheckReader*™

Nikolai Gorski, Valery Anisimov, Emmanuel Augustin, Olivier Baret, Sergey Maximov

Artificial Intelligence and Image Analysis (A2iA), 40 bis, rue Fabert, 75007 Paris, France; e-mail: gorski@a2ia.com, http://www.a2ia.com, Tel.: +33-1-44420080, Fax: +33-1-44420081

Received October 13, 2000 / Revised December 4, 2000

Abstract. This paper presents the current state of the *A2iA CheckReader*™ – a commercial bank check recognition system. The system is designed to process the flow of payment documents associated with the check clearing process: checks themselves, deposit slips, money orders, cash tickets, etc. It processes document images and recognizes document amounts whatever their style and type – cursive, hand- or machine printed – expressed as numerals or as phrases. The system is adapted to read payment documents issued in different English- or French-speaking countries. It is currently in use at more than 100 large sites in five countries and processes daily over 10 million documents. The average read rate at the document level varies from 65 to 85% with a misread rate corresponding to that of a human operator (1%).

Key words: Bank check processing – Payment systems – Handwriting recognition – Automatic reading – Document analysis – Intelligent character recognition

1 Introduction

Bank checks are, probably, the most widespread documents. Nearly one hundred billion checks circulate yearly all over the world. The great bulk of them are still processed manually by human operators, the most common and labor-consuming operation being document amount reading and typing. Automation of bank check processing is an important and promising application of document recognition techniques [2, 3, 12–14, 22]. It uses both recent theoretical achievements of pattern recognition and document analysis [5, 6, 10, 15, 16, 24, 28, 30], and practical approaches developed in adjacent application areas, such as postal automation or form recognition [3, 4, 8, 25, 29]. A comprehensive overview of the state-of-the-art in the field of check recognition can be found in [2]. Recent publications devoted to various aspects of check processing are [1, 17, 18, 20, 27, 31]. Engineering bank check recognition systems are presented in [7, 9, 14, 19, 21, 26]. As far as we know, some of these systems are commercially available and are in real use [9, 12, 18].

Bank check processing is usually performed either in big clearing centers or at branch agencies. They are equipped with fast scanners/sorters, archiving systems, and videocoding terminals for operators who make primary data entry. Operators look at document images one by one and enter document amounts. While doing this, operators naturally make errors: about 1% of documents (on average) appear to be entered with incorrect amounts. There exist a number of schemes to reduce this error rate to a demanded very low level of one error per 10^4 documents – such as double keying, balancing with partial control sums, etc. However, primary data entry remains the kernel of traditional check processing and its most resource-consuming part.

A common way to automate this process is to replace a human operator with an *intelligent character recognizer* (ICR) that is able to do the operator's job with the same (or lower) error rate. Thus, check clearing technology is not disturbed, which is essential for integrators who develop, install, and support industrial banking systems. Usually, an ICR is allowed to reject some documents: operators will further process these items in the traditional way. Then, if an ICR processes, for example, 67% of documents with the same substitution rate as that of a human operator and rejects the remaining 33%, it saves about 2/3 of human work power. By "substitution" we mean the error rate measured only on those items that are accepted by an ICR. (For a human operator, substitution- and error rates appear to be identical.)

In this paper, we present a commercial ICR system called the *A2iA CheckReader*™. It is designed to replace a human operator in payment document clearing, i.e, to read amounts of bank checks and associated documents. Besides checks themselves (Fig. 1), a real document flow also contains other items, such as deposit slips, debit- and credit forms, money orders, cash tickets, etc. (Fig. 2). As checks, these items also contain amounts to be recognized. Normally, all payment documents are scanned in a common stream and should be processed

Correspondence to: N. Gorski

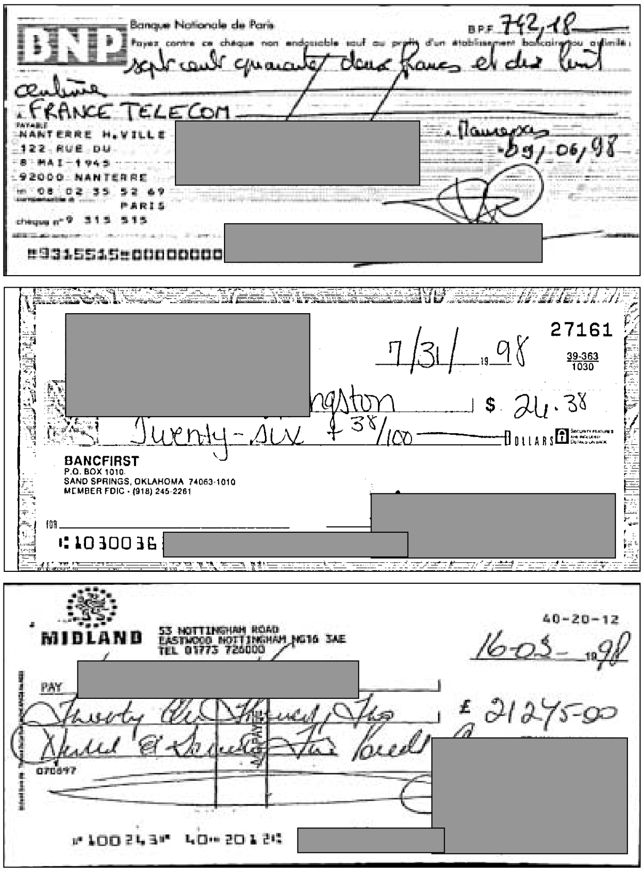


Fig. 1. Binary images of typical French-, US-, and British bank checks. Confidential information fields are shaded

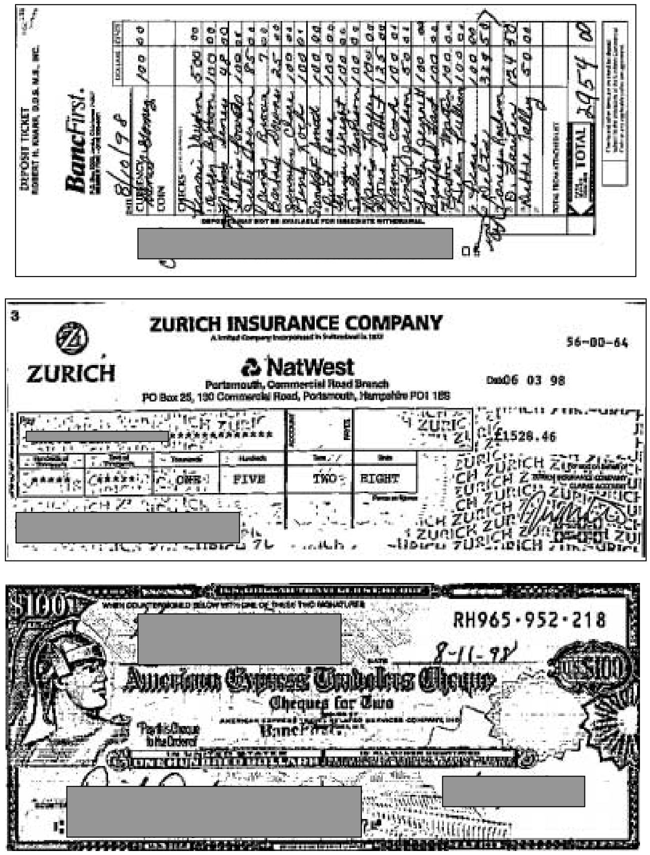


Fig. 2. Samples of non-check payment documents. Confidential information fields are shaded

in a similar way. Thus, the main task of the system is detection and recognition of amounts on all documents of the flow. Sometimes the system should also be able to read other information fields (e.g., account numbers, dates, etc.), but this area of the work is beyond the scope of the present paper.

The amount in different payment documents can be expressed in quite a different way. Sometimes it is machine printed (or pre-printed), in other cases it can be handwritten. The style of a handwritten amount can be purely cursive or hand-printed. In particular, check amounts are handwritten in 60–80% of cases. Often, the amount is indicated in a document twice: as a numeral expressed in digits (courtesy amount), and as a phrase expressed in words (legal amount) – see Fig. 1. The ICR system should locate amount fields of all types and styles in the document image and then recognize them to make a decision regarding what the true document amount is. All this is realized in the *A2iA CheckReader*™. To fulfil its task the system performs automatic localization, extraction, and cleaning of amount field(s); segmentation of the extracted amount image into characters and words; recognition of characters and words; interpretation of amount; and decision making regarding whether to accept or reject the document.

The paper is organized as follows: Sect. 2 presents the architecture and main features of the *A2iA CheckReader*™. Section 3 describes country- and language-specific aspects of system implementation for different countries. Section 4 gives an overview of implemented document processing techniques and approaches. Finally, Sect. 5 contains experimental results.

2 System architecture and design principles

The *A2iA CheckReader*™ is designed to process payment documents written in French or English. It is able to recognize both courtesy and legal amounts on various types of payment documents: handwritten personal or business checks, machine-printed checks, deposit tickets, credit/debit forms, money transfer orders, travelers checks, etc.

Figure 3 shows the current organization of the system which is currently different from what has been reported in our previous papers [12, 19]. As can be seen from Fig. 3, there are two main processing chains: one for documents with machine-printed amount(s), and the other for documents with handwritten amount(s). A special module distinguishes these two document types automatically, based on measurements of properties of potential amount zones. To minimize processing time, the more time-consuming legal amount recognition process is started only in those cases when the courtesy amount recognizer cannot make a decision by itself.

Conceptually, the *A2iA CheckReader*™ is based on the following key ideas:

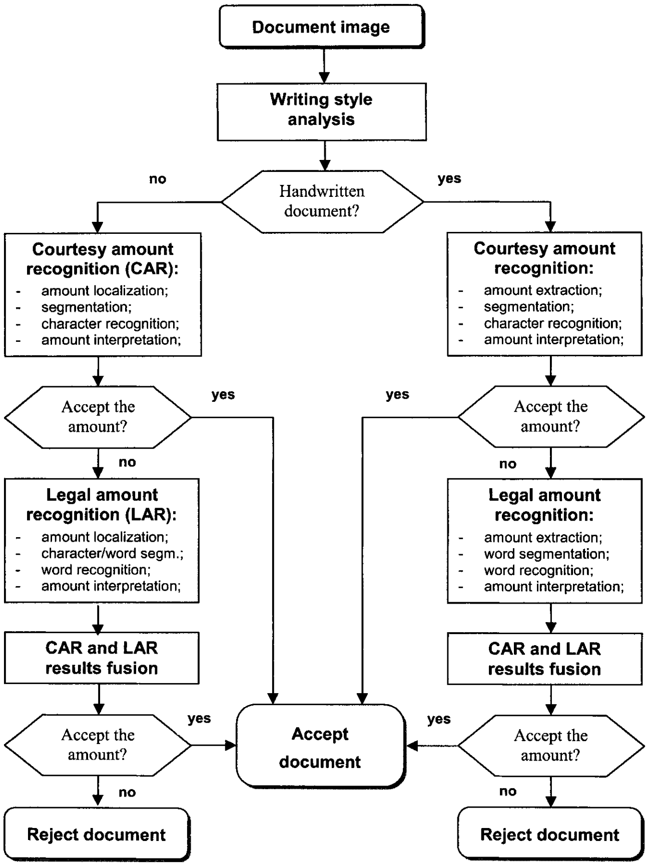


Fig. 3. General overview of the A2iA CheckReader system

1. *Hierarchical organization*, for the major part with bottom-up information flow. Starting from the pixel level, the recognition system proceeds via character parts, characters, words, to the amount level. Recognition of an element of a certain level is based on recognition results of lower-level elements, observations, and *a priori* information of a current level, and contextual information of higher level(s). Nearly all document analysis tasks are treated in the system as recognition tasks and are solved by recognition modules. For example, segmentation is considered as recognition of a correct segmentation option among several possible alternatives; decision making is approached as recognition of one variant from the set of optional decisions, etc.
2. *"Soft" recognition and decision-making.* All recognition modules of the system are soft classifiers. Each module produces not a single (hard) decision, but a list of possible decisions (candidates) ordered by decreasing confidence values. The traditional recognition rate is not a sufficient quality measure for a soft recognizer: it characterizes just the quality of the top candidate, while other members of the list remain unutilized. Instead, we use an information-based criterion of recognition quality [11, 19]. It is possible to prove that maximum information on correct classification is contained in such a list, whose members are estimates of posterior class probabilities. Thus, probabilistic candidate lists are optimal for our purposes and all recognition modules of the system return their results in this form. The quality criterion we use is the normalized quantity of information contained in the output candidate list of a recognition module (on average):

$$Q = 1 - \mathbf{E}[-\log \Pr(true\_class)]/ \log(N\_classes), \tag{1}$$

where averaging is done on a labeled data set with a known *true_class* value. Individual recognizers at all system levels, as well as the system as a whole, are characterized and adjusted by this criterion. Systematic use of soft recognition greatly increases error tolerance of the system. Only the very final decision is hard (binary): a document is either accepted with a certain amount (and score) or rejected.
3. *Modularity* and *standards* for input/output of the modules. Each recognition task in the hierarchy is isolated as much as possible and solved by a corresponding software module. Information flows between different modules are standardized: modules receive and return lists of candidates with associated probabilities.
4. *Complementarity.* In order to achieve robustness and redundancy, a given recognition task is solved by several complementary algorithms and/or with several complementary data representations. For example, from two to four OCRs are used at the character recognition level; two complementary word recognizers are used at the word level; two chains of courtesy- and legal amount recognition support the amount recognition level. The results of complementary recognizers are integrated by combining output candidate lists of individual recognizers and producing a new single output list. Two types of integrators are used in the system: the log-liner rule and a neural network [11]. The log-linear combination rule for N soft probabilistic recognizers is expressed as follows:

$$\boldsymbol{L} = \sum_{i=1}^{N} w_i \boldsymbol{L}_i + w_0 \boldsymbol{L}_0 + \boldsymbol{C},$$

where w_i are relative weights of the recognizers, $\boldsymbol{L}_i$ are vectors of their logarithmed outputs; $\boldsymbol{L}$ is the vector of logarithmed outputs of the integrator; $\boldsymbol{L}_0$ is the vector of logarithmed class priors, and $\boldsymbol{C}$ is a normalizing constant. The neural network integrator is a multi-layered perceptron with one hidden layer. It uses the Softmax output function, and is trained with the cost function that is a relative entropy of the output list; therefore, this network estimates posterior class probabilities. Both integrators maximize the same information criterion (1), thus providing maximum information in their output lists. The log-linear rule has only $N+1$ adjustable parameters $(w_0, \ldots, w_N)$ and is efficient in simple cases of weakly correlated classifiers, while the neural network is more powerful, but demands much more training data.

5. *Adaptivity.* Wherever possible, variations in the data are modeled by a set of parameters that are then learned from training data. The training or adjusting of any module of the system means obtaining an extremum of quality criterion (1) in the parameter space given a labeled test data set. This enables us to adjust the system for different exploitation conditions, for example, in different countries. However, to perform this we need labeled data sets for each module to be trained.
6. *Segmentation-with-recognition.* In order to recognize a given object, it is segmented into parts that are simpler to recognize and/or have a smaller set of possible classes. Several segmentation options are usually considered. The probability that the original object belongs to any of the object classes is estimated by summing up probabilities for this class over all segmentation options:

$$\mathrm{Pr}\,(object) = \sum_{options} \left(\mathrm{Pr}\,(option) \prod_{parts} \mathrm{Pr}\,(part|option) \right).$$

   This approach assumes summation of all class probabilities coming from different segmentation options. Thus, all available information for object recognition is used. However, by the end of this process, it is difficult to trace back the origination of classification of a particular object in terms of segmentation options and object parts.

Thus, the *A2iA CheckReader*™ is essentially probabilistic system, all parts of which are developed to get a locally optimal performance according to the criterion of maximum output information. Soft decision making at all levels and use of complementary algorithms lead to high robustness and reliability of the system. At the same time, one of the main principles of system design is adaptivity of each recognition module. It enables us to obtain the maximum possible performance having enough training data. The price of this is the necessity to have not a single multilingual and country-independent system, but, in fact, a family of several recognition systems in one, each being trained on its own data or country-specific documents.

3 Country-specific problems

At first glance, it would not seem difficult to develop a "generic" system that is able to recognize amounts of payment documents from different countries (at least, the courtesy amount). However, our experience shows this is not true. There are several reasons for this at various levels of amount recognition: in different countries there are different styles of character- and amount writing/typing, different delimiters, special symbols, currency signs, etc. Surprisingly enough, one of the deepest and most important differences appears to be at the character level (for handwritten characters), even when character sets are identical.

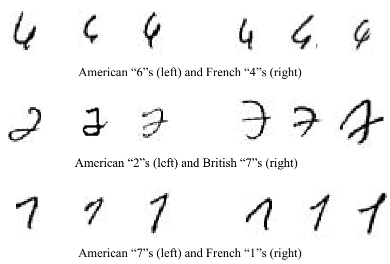


Fig. 4. Examples of different digits having similar shapes in different countries

For example, the amount 742,18 of a French check (Fig. 1) might be read as 962,78 in a US check. In US checks, it is really difficult to find a "1" written with a French-style "nose": such figures would rather be taken as "7". At the same time, an American "6" very often looks like French "4" (see Fig. 4). In Britain, there are specific "7"s and "9"s that can be recognized as "4" in France or the USA. In France, people use commas to separate centimes (cents) and points to separate thousands, while English-speaking people do reverse; besides this, the British often use dashes as a cent separator.

To demonstrate this effect "quantitatively" let us consider the character recognition level of the *A2iA CheckReader*™. The character recognizer consists of four OCRs that recognize segmented character images [19]. Then, OCR results are integrated to obtain final scores of character classes. All four OCRs are neural networks that estimate posterior class probabilities but use different feature sets:

- OCR_1 processes pixel values of reduced character image;
- OCR_2 measures various topological image characteristics;
- OCR_3 uses character contour-based features;
- OCR_4 takes into account contextual information (character position, relative sizes, etc.).

The neural network integrator combines the results of individual OCRs and produces final class probability estimates. The neural network of the integrator has the same architecture as the neural networks of OCRs.

This character recognition module was trained separately for US- and UK-systems on character images extracted from real American and British handwritten bank checks. The volume of training sets was about 90,000 and 50,000 characters, respectively. Twenty-six character classes included digits from 0 to 9, currency signs, various delimiters, and a few frequent letters that might appear in courtesy amount fields. Experiments consisted, first, in running the recognition module trained with US data on US- and UK test data, and second, in analogous running of the module trained with UK data. Two test sets contained 5,000 digits from 0 to 9 extracted from US- and UK checks. The results of experiments are shown in Tables 1 and 2.

It is obvious from Tables 1 and 2 that OCRs trained on real material originating from a particular country can be rather weak or nearly useless in recognizing text from another country, even if these countries have the

Table 1. Performance of the A2iA character recognition module on 5,000 isolated digits extracted from American checks (% of correct answers)

	OCR_1	OCR_2	OCR_3	OCR_4	Combination
Module trained on US characters	92.5	89.9	90.5	42.2	97.3
Module trained on UK characters	80.7	80.6	80.4	28.5	87.4

Table 2. Performance of the A2iA character recognition module on 5,000 isolated digits extracted from British checks (% of correct answers)

	OCR_1	OCR_2	OCR_3	OCR_4	Combination
Module trained on UK characters	90.1	92.1	89.1	40.0	98.5
Module trained on US characters	85.1	85.9	84.7	26.6	94.3

same languages and are as culturally close as the USA and the UK.

For countries with different languages this effect can be even stronger. Besides the character recognition level, similar problems appear at the word recognition and amount interpretation levels: there are different styles of amount writing/typing, different vocabularies, delimiters, special symbols, currency signs, etc. All this leads to the necessity of individual training of country-specific versions of the system, rather than one universal system for all countries. Our solution is to have a common kernel (more then 90% of the software) and special country-dependent parts that are developed and/or adjusted individually with country-specific data. This way we can get the maximum possible performance from the system and while not having to develop each country's version for too long. However, we still have to pay for this with several weeks of work to prepare a new release of the *A2iA CheckReader*™ for a new language or country with a distinctive writing style.

For the time being, there are three "main" country-specific versions of the system: French, English (for UK documents), and American (for US documents). While developing a version for a new country we first combine existing modules of the "main" versions. For example, the Canadian version is built from the modules of American and French ones. If the results of the combined version are satisfactory, it is released, otherwise we collect country-specific data and make an individual adjustment of the version with these data.

4 System details

The main functions of both printed- and handwritten recognition chains of the *A2iA CheckReader*™ are identical:

1. Extraction of courtesy- and/or legal amount fields from the document image and cleaning them from background forms and noise.
2. Segmentation of obtained text images into graphemes, characters or words.
3. Recognition of characters or words.
4. Interpretation of amounts expressed by digits (courtesy amount) or words (legal amount).
5. Decision-making: to accept or reject the recognized amount.

Document processing begins from the localization of the courtesy amount field. There are several modules in the system designed to locate special objects in a document image. To detect the courtesy amount, the most important anchors are currency signs and special words (such as "BPF" that should precede the courtesy amount in French checks). Based on the results of the anchor search, a potential amount field is localized. Then a style-detection module analyses this field. It measures different characteristics of connected components inside the field and uses them as features to make a decision: does the field contain machine-printed amounts or is it handwritten? This decision is correct on average in 95% of cases. Depending on the detected field style, either the printed- or handwriting chain of the *A2iA CheckReader*™ is called.

Whatever the chosen recognition chain is, the first stage is courtesy amount recognition (CAR). First of all, courtesy amount objects are extracted from the localized amount field of the document image. This is one of the most informal and difficult problems due to the variety of check forms, possible check layout pictures, and suboptimal image binarization. This task appears to be easier for UK and US documents than for French ones. The reasons for this are essentially a cleaner layout and stricter position of the courtesy amount field on these checks, despite the fact that different US payment documents may have different sizes.

The extracted and cleaned courtesy amount image is segmented into characters. This task is rather easy for UK documents where all characters are usually placed on the same base line. It is more difficult for French checks and the most difficult for US checks due to the frequent presence of a "two-dimensional" fractional cent part where characters are often touching each other. Segmentation produces several segmentation options, each then being processed and interpreted separately. In each option, the courtesy amount image is partitioned into non-intersecting parts representing potential characters (Fig. 5).

Four complementary OCRs recognize these characters, then the recognition results of all OCRs are integrated producing a list of class candidates with associated probability estimates [11, 19] (see also Sect. 3). At the character level, the recognition rate varies from 97%

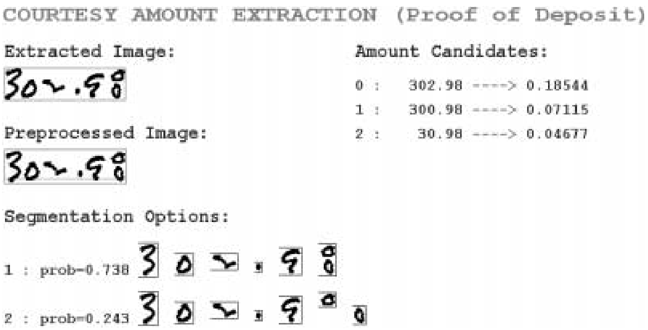


Fig. 5. Courtesy amount recognition

for the US system to 98.5% for the UK one on ten digit classes.

The amount interpretation module uses character recognition results along with character position- and segmentation information. It generates syntactically correct amount candidates (again, associated with probability estimates). At this level, probabilities of identical amounts originating from different segmentation options are summed up. The top amount candidate is correct in 75%, 77%, and 80% of cases for French-, US- and UK documents, respectively.

Courtesy amount recognition results are often reliable enough to make a decision on the check amount. Results at this level highly depend on a variety of possible amount writing styles. For example, in French checks, one and the same amount of 200 francs can be expressed as 200,00 or 200 *Frs* or 200F00 or 200^{F}00 or 200,^{f}00 or 200- and in many other ways. Probably due to this, the performance of the system on French documents is the worst at the CAR level: it equals approximately 40% if the substitution rate is fixed at 1%. On the contrary, for the same substitution rate, 50% of US documents appear to be accepted, and for British ones this figure increases up to 60%.

For checks that are not accepted at the CAR level, the legal amount recognition (LAR) is called to obtain more information for decision making (Fig. 3). LAR starts from the extraction of the legal amount field. This operation is easier for US checks, as the layout is cleaner and there is only one line of handwriting. However, in French and UK checks there can be two lines. After extraction, some preprocessing is performed, including slant correction, lowercase zone detection, and location of a cent part (Fig. 6). The latter is specific for US and UK checks while French centimes are more often expressed by words than by a numeral. The last step of preprocessing is grapheme extraction. Graphemes are elementary parts of letters; they are used further in word segmentation and recognition.

The word segmentation process produces multiple options of segmenting the legal amount image into possible words. Unfortunately, our segmentation technique developed for French checks works worse on US ones: 50% of correct segmentations in the top option instead of 60%. This happens due to the inherent tendency of people in the USA to omit gaps between words or to glue words together.

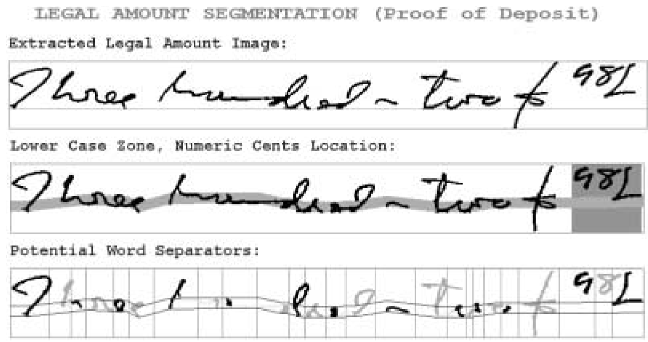


Fig. 6. Legal amount segmentation into graphemes, words, and dollar- and cent parts

Table 3. Performance of the word recognition module on 36,000 isolated words of 38 classes extracted from US checks

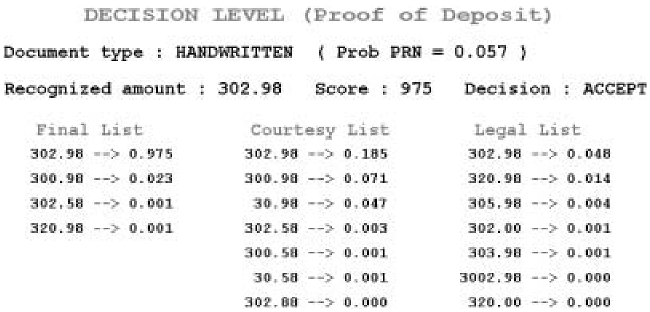
Markov Word Recognizer	Holistic Word Recognizer	Combination of both
81%	70%	89%

Two complementary word recognizers process possible words. The first one is based on holistic features, the second uses hidden Markov models [1]. Word recognition results are integrated with a neural network integrator and then presented as a list of word class candidates with associated probabilities. At the word level we obtain from 85% to 93% top candidate recognition rates for 27–40 word classes. Table 3 presents the results for words from US personal checks.

The amount interpretation module uses word recognition results along with word segmentation information. It generates syntactically and semantically correct amount candidates associated with probability estimates. The numerical cent/centime/pence part of the legal amount is recognized separately: it is segmented into characters and processed similarly to courtesy amount. At this level, probabilities of identical amounts originated from different segmentation options are summed up. The top candidate is correct in 55%–70% of cases.

CAR and LAR results are joined with a log-linear integrator to get the final list of amount candidates (Fig. 7). The top candidate from this list can be either accepted or rejected – this is the only hard decision of the system. To make this decision, a set of properties of the final list is analyzed. Then a special neural network estimates the probability that the top candidate is the true check amount. If this value is greater than a predefined threshold, the check is accepted, otherwise it is rejected. Varying the threshold, one can obtain a flexible exchange of recognition errors with rejections. The use of LAR increases system performance by 15–25% compared with the CAR-alone variant.

The above examples considered handwritten documents. The printed document processing chain is similar to that of the handwritten one. The main difference in approaches to processing printed documents and handwritten ones is in the location of the amount zone. For a handwritten document, the amount zone is unique, and



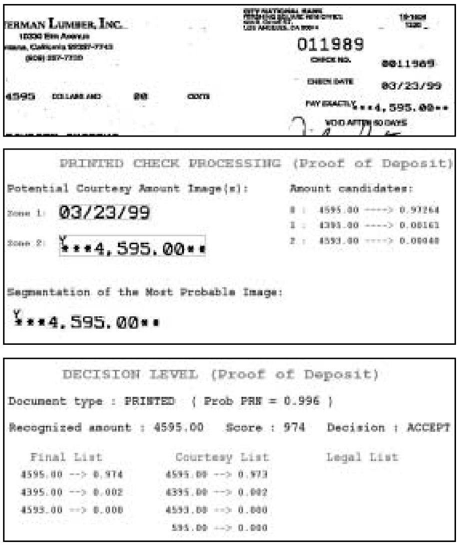
DECISION LEVEL (Proof of Deposit)

Document type : HANDWRITTEN (Prob PRN = 0.057)

Recognized amount : 302.98 Score : 975 Decision : ACCEPT

Final List	Courtesy List	Legal List
302.98 --> 0.975	302.98 --> 0.185	302.98 --> 0.048
300.98 --> 0.023	300.98 --> 0.071	320.98 --> 0.014
302.58 --> 0.001	30.98 --> 0.047	305.98 --> 0.004
320.98 --> 0.001	302.58 --> 0.003	302.00 --> 0.001
	300.58 --> 0.001	303.98 --> 0.001
	30.58 --> 0.001	3002.98 --> 0.000
	302.88 --> 0.000	320.00 --> 0.000

Fig. 7. Decision making by fusion results of courtesy- and legal amount recognition



PRINTED CHECK PROCESSING (Proof of Deposit)

Potential Courtesy Amount Image(s):

Zone 1: 03/23/99

Zone 2: ***4,595.00**

Amount candidates:

0 : 4595.00 ----> 0.97264
1 : 4395.00 ----> 0.00161
2 : 4593.00 ----> 0.00048

Segmentation of the Most Probable Image:

***4,595.00**

DECISION LEVEL (Proof of Deposit)

Document type : PRINTED (Prob PRN = 0.996)

Recognized amount : 4595.00 Score : 974 Decision : ACCEPT

Final List	Courtesy List	Legal List
4595.00 --> 0.974	4595.00 --> 0.973	
4395.00 --> 0.002	4395.00 --> 0.002	
4593.00 --> 0.000	4593.00 --> 0.000	
	595.00 --> 0.000	

Fig. 8. Recognition of a machine-printed document

several segmentation options of this zone are considered, while for printed documents several potential amount zones are extracted, and a single segmentation of each zone is analyzed (Fig. 8). As for results at different levels, they are more homogeneous for different countries than that of the handwritten chain. The only big exception (and the main difficulty) is a number of French personal checks printed at supermarkets. On these checks the amount could appear nearly anywhere, it might be not separated from other data, and the quality of printing can be really poor (Fig. 9). On average, from 60 to 75% of printed items can be accepted at the CAR level. The LAR part can bring 5–10% more. This is lower then the LAR impact on the handwritten chain, as not all printed documents contain a legal amount part.

5 Experimental results

The more we work with real documents, the more we understand that it is very difficult to report the general performance of an ICR system. There are many factors affecting recognition and error rates even for large sets of dozens of thousands of documents. Among these factors it is worth mentioning:

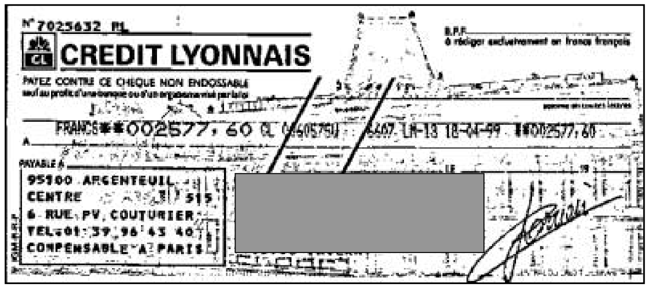


Fig. 9. A French check filled at supermarket

– various document scanners that produce images of different quality and resolution;
– different fractions of miscellaneous documents in a document flow (printed/handwritten, business/personal, checks/non-checks, etc.);
– different amount statistics (percentage of "round" and "non-round" amounts, distribution of amount values);
– local specifics in the style of writing or in document style (e.g, background pictures);
– different quality of writing depending on time and/or day of document issuing (e.g, there is a certain drop in the recognition rate on public holidays).

Due to these reasons, the results of the *A2iA CheckReader*™ reported below should be considered as country-, and site specific, despite the fact that they have been obtained on rather large test sets of real-life data. However, these results reflect the typical performance of the system in particular countries.

There are rather few test results of engineering check recognition systems published in the literature and reporting system performances [2, 9, 14, 18, 31]. Certainly, it would be more correct to compare results of altered systems if they were tested on the same data. Unfortunately, it is very difficult to carry out common tests of check recognition software. On the one hand, banks and other institutions involved in payment document processing are nearly paranoiac about the confidentiality of the document images they provide for training/testing ICRs (and they are probably right in this); on the other hand, companies developing or integrating check recognition systems do not demonstrate the results of their comparative cross-tests as they can disclose commercially valuable information. Because of this, we demonstrate below just the "absolute" results achieved by the *A2iA CheckReader*™. However, to our knowledge, these results are the best reported so far in the bank check recognition domain.

Hereinafter, the main system performance characteristics are defined as follows:

$$Read = \frac{\text{The number of accepted documents}}{\text{Total number of processed documents}} \times 100\% \,;$$

$$\textit{Recognition} = \frac{\begin{array}{l}\text{The number of documents accepted}\\ \text{with correct amounts}\end{array}}{\text{Total number of processed documents}} \times 100\% \,;$$

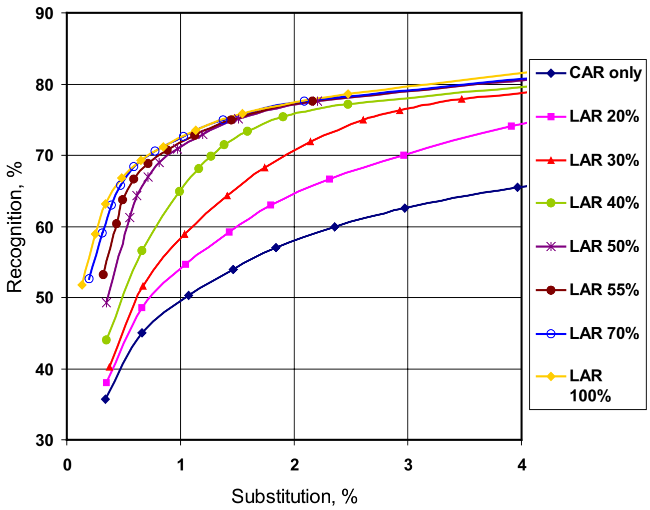


Fig. 10. Handwritten check recognition (CAR + LAR) with different thresholds of LAR calls

Error

$$= \frac{\text{The number of documents accepted with incorrect amounts}}{\text{Total number of processed documents}} \times 100\%\,;$$

Substitution

$$= \frac{\text{The number of documents accepted with incorrect amounts}}{\text{Total number of accepted documents}} \times 100\%\,.$$

We would like to stress that *Substitution* is a more suitable characteristic of an ICR performance than the frequently used *Error*. When a human operator is replaced by an ICR system, it is ICR *Substitution* adjusted equal to that of a human operator. Then *Read* demonstrates a fraction of primary data entry automatically made by an ICR. However, only *Recognition* reflects real labor economy. Sometimes it is also called the "killing rate" of an ICR, as it shows a percentage of documents processed fully automatically and never seen by humans.

Figure 10 presents recognition results on a test set of 35,000 *handwritten* US checks extracted from a real document stream. This example demonstrates how legal amount recognition helps to improve the system performance. CAR is called on each check, and LAR is called on those checks where the CAR score is less than a predefined threshold. As can be seen from Fig. 10, it is practically useless to call LAR for more than 60% of checks (on average) if *Substitution* is fixed at 1% – our target level that is typical for human operators.

Other experiments demonstrate performances of the system on a mixed document flow (handwritten and machine-printed checks, deposit tickets, etc.) typical for various countries. In these experiments the system automatically distinguished printed documents from handwritten ones and then a corresponding recognition chain processed each document.

Figures 11 and 12 present recognition results on data sets from French and Canadian sites. It is interesting to see how different are the recognition rates on subsets

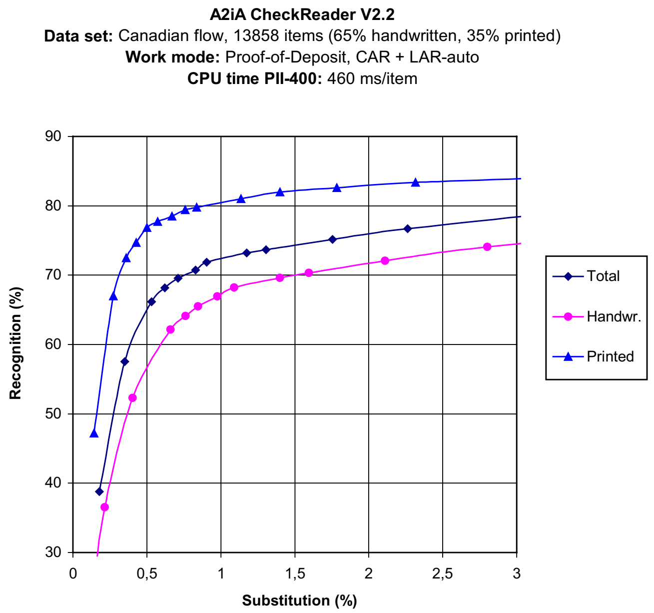


Fig. 11. Recognition of printed- and handwritten documents of Canadian document flow

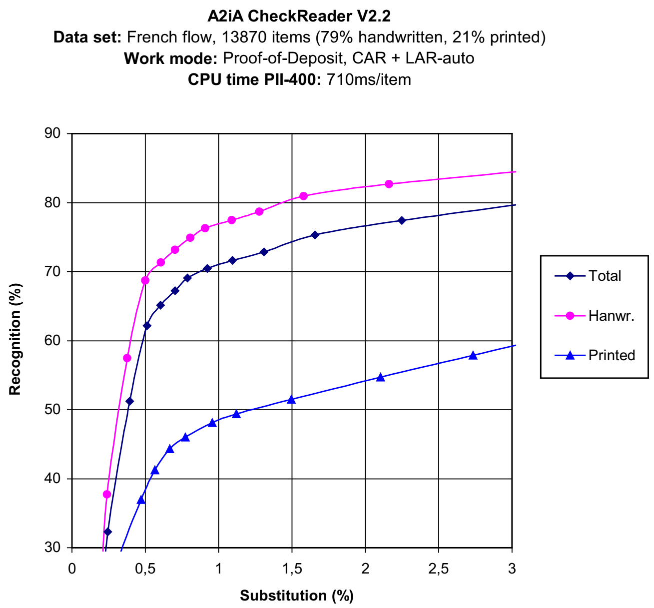


Fig. 12. Recognition of printed- and handwritten documents of French document flow

of printed documents. A poor performance on French printed documents is explained by a rather big fraction of "supermarket" checks in the data set (see Fig. 9). As has been mentioned these items are especially difficult for automatic reading. In Canada, as in many other countries, printed payment documents are of relatively good quality, so the printed chain is usually superior in performance then the handwriting chain.

Figure 13 presents performances of the system measured on data sets from six different countries. It can be observed that variation in performance between country-specific versions achieve nearly 20%. All these versions recognize the legal amount expressed either in English

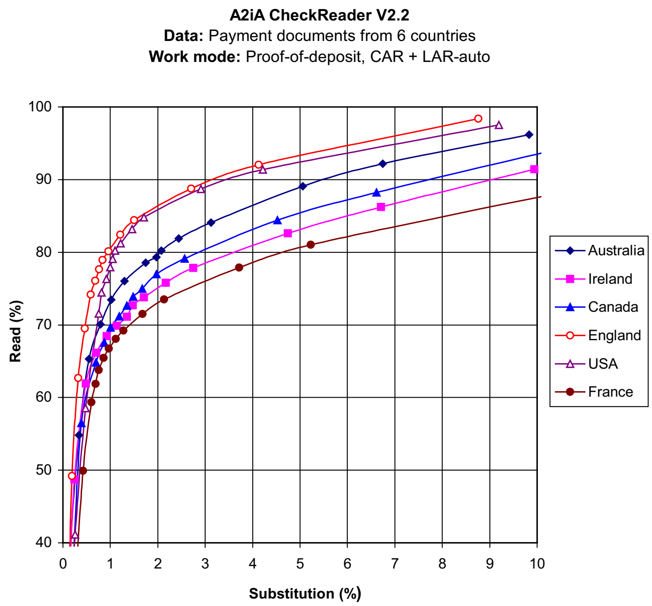


Fig. 13. Performances of country-specific versions of the *A2iA CheckReader*™

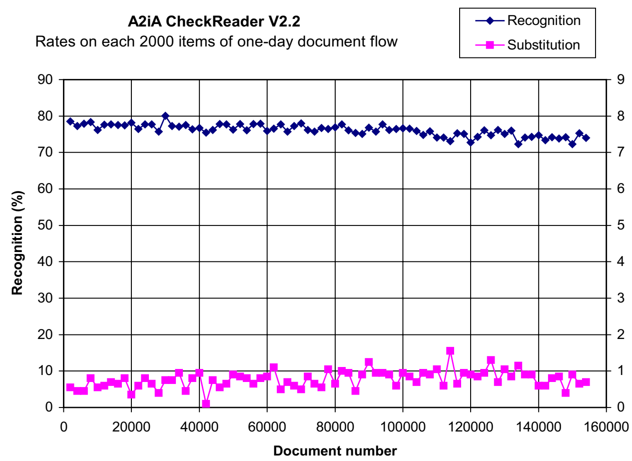


Fig. 14. Performances of the *A2iA CheckReader*™ averaged on each 2,000 items of one-day document flow

or French except the Canadian one which is bilingual and recognizes legal amounts written in both languages. An appropriate language hypothesis is chosen according to the language of key words detected in the layout of the recognized document (e.g, words *amount/montant*, *pay/payez*, *of/de*, etc.)

The last experiment presents the behavior of the system during a one-day session at one of the UK production sites (Fig. 14). From this chart it is clear that even at one and the same site the performance of the system can vary quite a bit (+/− 5%). There is a rather obvious trend in performance that somehow degrades by the end of the day. This is a common effect for many sites where documents are processed in the first-in-first-out mode: "Morning" documents appear to be easier to recognize then "evening" ones – this might be connected, for example, with a different repartition of personal (mainly handwritten) and business (mainly machine-printed) checks through a day.

For the time being (October 2000), the *A2iA CheckReader*™ is in use at more than 100 large sites (banks, clearing centers, service bureau, etc) in France, the USA, the UK, Ireland, and Australia. The system is also used in many cash machines (ATMs) and bank branches for on-the-spot recognition of small document quantities. Currently, the total number of daily processed documents is above 10 millions items, and this number is further increasing.

6 Conclusions

We have presented the *A2iA CheckReader*™ – a commercial bank check recognition system designed to process payment documents written/printed in French or English. Recognized documents include bank checks, deposit slips, money orders, cash tickets, etc. The system has two main recognition chains: one for courtesy- and legal amount recognition on handwritten documents, and the other for amount recognition on printed document. The system performs automatic localization, extraction, and cleaning of information fields in a document image; segmentation of extracted text image into characters and words; recognition of characters and words; interpretation of amount and decision making to accept or reject the document. Average recognition rates in real production varies from 65 to 85% with an error rate fixed at the level of a human operator.

Despite the achieved performance and commercial success, the system is still under development both in extensive and intensive ways. New techniques are going to be implemented to improve recognition- and decrease error rates. New countries and languages will be added to the system to broaden the geography of its use. The next "main" country-specific version of the system will be Portuguese, designed for recognition checks from Brazil and Portugal. Versions of the system for Canada and Hong Kong are currently under testing.

In addition, in the near future, the *A2iA CheckReader*™ will be able to recognize other important information fields in payment documents such as dates, beneficiary and payee names, the payee's address and signature, etc.

References

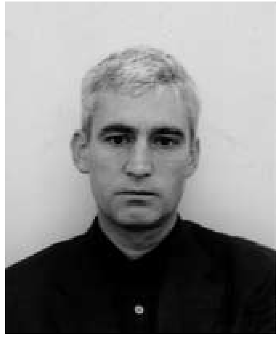
1. E. Augustin, O. Baret, D. Price, S. Knerr: Legal amount recognition on French bank checks using a neural network-hidden Markov model hybrid. In: S.-W. Lee (ed.) Advances in Handwriting Recognition. World Scientific, Singapore, 1998, pp. 81–90
2. S. Impedovo, P. Wang, H. Bunke (eds.) Automatic Bankcheck Processing. World Scientific, Singapore, 1997
3. N. Bartenek: The role of handwriting recognition in future reading systems. Proc. 5th Int. Workshop on Frontiers in Handwriting Recognition, University of Essex, UK, 1996, pp. 147–165
4. T.M. Breuel: Design and implementation of a system for recognition of handwritten responses on US census

forms. Proc. Document Analysis Systems 94, Kaiserslautern, 1994, pp. 109–134
5. R.G. Casey, E. Lecolinet: A survey of methods and strategies in character segmentation. IEEE Trans. PAMI, 18(7): 690–706, 1996
6. M.Y. Chen, A. Kundu, J. Zhou: Off-line handwritten word recognition using a hidden Markov model type stochastic network. IEEE Trans. PAMI, 16(5): 481–496, 1994
7. G. Dimauro, S. Impedovo, G. Pirlo, A. Salzo: Automatic bankcheck processing: a new engineered system. In: S. Impedovo, P. Wang, H. Bunke (eds.) Automatic Bankcheck Processing. World Scientific, Singapore, 1997, pp. 5–41
8. A.C. Downton, R. Tregidgo, C. Leedham, Hendrawan: Recognition of handwritten British postal addresses. In: S. Impedovo, J.-C. Simon (eds.) From Pixels to Features III. North-Holland, Amsterdam, 1992, pp. 129–144
9. G. Dzuba, A. Filatov, D. Gershuny, I. Kil, V. Nikitin: Check amount recognition based on the cross validation of courtesy and legal amount fields. Int. J. Pattern Recognition Artif. Intell. 11(4), 1997
10. M. Gilloux: Hidden Markov models in handwriting recognition. In: Fundamentals in Handwriting Recognition. NATO ASI series F, vol. 124, Springer, Berlin Heidelberg New York, 1994, pp. 264–288
11. N. Gorski: Practical combination of multiple classifiers. In: A. Downton, S. Impedovo (eds.) Progress in Handwriting Recognition, World Scientific, Singapore, 1997
12. N. Gorski, V. Anisimov, E. Augustin, O. Baret, D. Price, J.-C. Simon: A2iA Check Reader: a family of bank check recognition systems. Proc. 5th Int. Conf. on Document Analysis and Recognition, Bangalore, India, 1999, pp. 523–526
13. D. Guillevic, C.Y. Suen: Cursive script recognition applied to the processing of bank checks. Proc. 3rd Int. Conf. on Document Analysis and Recognition, Montreal, Canada, 1995
14. G. Houle, D. Aragon, R. Smith, M. Shridar, D. Kimura: A multi-layered corroboration-based check reader. Proc. Document Analysis Systems'96, Malvern, USA, 1996
15. J.J. Hull, T.K. Ho, J. Favata, V. Govindaraju, S. Srihari: Combination of segmentation-based and holistic handwritten word recognition algorithms. In: S. Impedovo, J.-C. Simon (eds.) From Pixels to Features III. North-Holland, Amsterdam, 1992, pp. 261–272
16. R.A. Jacobs: Methods for combining experts' probability assessments. Neural Comput., 7: 867–888, 1995
17. G. Kaufmann, H. Bunke: Error localization and correction in check processing. In: S.-W. Lee (ed.) Advances in Handwriting Recognition. World Scientific, Singapore, 1999
18. S. Kelland, S. Wesolkovski: A comparison of research and production architectures for check reading. Proc. 5th Int. Conf. on Document Analysis and Recognition, Bangalore, India, 1999, pp. 99–102
19. S. Knerr, V. Anisimov, O. Baret, N. Gorski, D. Price, J.-C. Simon: The A2iA Intercheque system: courtesy amount and legal amount recognition for French checks. Int. J.Pattern Recognition Artif. Intell. 11(4): 505–547, 1997
20. A. Koerich, L. Lee: Automatic storage, retrieval and visualization of bank check images. Proc. 5th Int. Conf. on Document Analysis and Recognition, Bangalore, India, 1999
21. A. Kornai, K. Mohiuddin, S. Conell: Recognition of cursive writing on personal checks. In: A. Downton, S. Impedovo (eds.) Progress in Handwriting Recognition. World Scientific, Singapore, 1997
22. E. Lethelier, M. Leroux, M. Gilloux: An automatic reading system for handwritten numeral amounts on French checks. Proc. 3rd Int. Conf. on Document Analysis and Recognition, Montreal, Canada, 1995, pp. 92–96
23. V. Di Lecce, A. Dimauro, A. Guerriero, S. Impedovo, G. Pirlo, A. Salzo: A new hybrid approach for legal amount recognition. Proc. 7th Int. Workshop on Frontiers in Handwriting Recognition, Amsterdam, Holland, 2000, pp. 199–208
24. R.P. Lippmann, M.D. Richard: Neural network classifiers estimate Bayesian a posteriori probabilities. Neural Comput., 3: 461–483, 1991
25. U. Miletzki, T. Bayer, H. Schafer: Continuous learning systems: postal address readers with built-in learning capability. Proc. 5th Int. Conf. on Document Analysis and Recognition, Bangalore, India, 1999, pp. 329–332
26. J.-V. Moreau: A new system for automatic reading of postal checks. In: S. Impedovo, J.-C. Simon (eds.) From Pixels to Features III, Frontiers in Handwriting Recognition. North-Holland, Amsterdam, 1992, pp. 171–185
27. S. Shetty, M. Shridhar, G. Houle: Background elimination in bank checks using greyscale morphology. Proc. 7th Int. Workshop on Frontiers in Handwriting Recognition, Amsterdam, Holland, 2000, pp. 83–91
28. J.-C. Simon: Off-line cursive word recognition. Proc. IEEE, 80(7): 1150–1161, 1992
29. S. Srihari: Recent advances in off-line handwriting recognition at CEDAR. Proc. 5th Int. Workshop on Frontiers in Handwriting Recognition, University of Essex, UK, 1996, pp. 1–15
30. O. Trier, A.K. Jain, T. Taxt: Feature extraction methods for character recognition - a survey. Pattern Recognition, 29(4): 641–662, 1996
31. X. Ye, M. Cheriet, C.Y. Suen, K. Liu: Extraction of bankcheck items by mathematical morphology. Int. J. Doc. Anal. Recognition. 2(2/3): 53–66, 1999

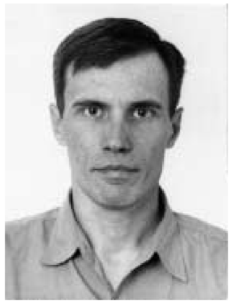
Nikolai Gorski graduated from the Leningrad Polytechnical Institute in 1976, and received a Ph.D. in Engineering Cybernetics in 1981 from the Leningrad Electrotechnical Institute. In 1989 he received his Doctor of Science degree from the Russian Academy of Sciences. In 1976 he joined the St. Petersburg Institute for Informatics and Automation where he worked as a leading researcher. He joined A2iA in 1992 as a specialist in Probalistic Information Processing. Currently he is a project leader responsible for performances of main A2iA products. He published more than 70 scientific papers and 6 monographs. He is an associate editor of the Pattern Recognition Journal. Interests of Dr. Gorski include data and image analysis and in particular practical questions of industrial document analysis and recognition.

Valeri Anisimov graduated from the Leningrad Univerity in 1984 and received his Ph.D degree from the St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences in 1993. In 1987 he joined the St. Petersburg Institute for Informatics and Automation working in the area of image processing and pattern recognition. Since 1996 he has been with A2iA (France), where he is now a research engineer. His current research interests include hierarchical multilevel systems for image processing and analysis, cursive handwritten recognition, visual perception and image understanding.

Emmanuel Augustin graduated from the school Ecole Supérieure de Physique et Chimie Industrielle de Paris, (GEIP) in 1998; he will defend his thesis in 2001 from University Paris 5. He joined A2iA in 1996, where he works on cursive word recognition for checks and documents. His interests are fields of statistical learning and document analysis.



Olivier Baret graduated from the Ecole Polytechnique in 1985 and received his Ph. D. in 1991 from the University Pierre et Marie Curie (Paris 6). From 1986 to 1991, he was a researcher in the LAFORIA laboratory (University Paris 6) where he worked on handwritten word recognition. In 1991, he joined A2iA as a research engineer working on handwriting recognition of cheques and postal addresses. Since 1999, he has been technical Vice President of A2iA.



Sergey Maxsimov graduated from Higher Naval College of Radio-Electronics in 1983 after what he was an engineer of digital control systems. He began research and development activity in 1991 as a developer of software for specialized control systems based on personal computers used in the navy. In 1994 he worked in the geo-information systems area where he developed visualization of electronic marine maps. Now his main direction of research is the recognition of printed texts and special signs.