

# Bank-check Processing System: Modifications Due to the New European Currency

N. Greco, D. Impedovo, M.G. Lucchese, A. Salzo, L. Sarcinella.  
*Consorzio Interuniversitario Nazionale per l'Informatica (CINI),  
Via G.Petroni 15/F.1, 70124 Bari – Italy  
E-mail: solidar@media.it*

## Abstract

*The introduction of a new currency in Europe has changed the way of writing both the courtesy and the legal amount on checks. This paper presents the most important modifications brought on the bank-check processing system in order to solve the related problems also by proposing the software tools that must be utilized. The Computer Aided Software Engineering tools provided by the "Khoros" system are used to support the improvement of the system prototype. A visual programming environment is used to assemble the bank-check processing system that can be easily modified and extended. The experimental results allow the adjournment of the improved system, as the modifications are introduced.*

## 1. Introduction

The rapid evolution of new technologies is yielding constant advances in bank-check processing field. The goal of any new technology in remittance processing is to reduce manual labor at every step of the process. The aim is to extend these productivity gains to the reading and to the recognition of a customer's amount check, so that they can be automatically converted into digital data for further processing. The results of more than thirty years of study and research have allowed to produce and to improve a prototype of a bank-check recognition system [1].

Bank-checks are very complex documents which differ not only in background but also in layout, depending on the issuing bank. Some pre-printed information as the guidelines, the name of the bank, the number of the check, the number of the customer account, the Bank Identification Code and the Bank Agency Code, and some user-entered information as the legal amount, the courtesy

amount, the date and the signature are present on a bank-check. In order to develop a bank-check processing system, several algorithms for document analysis and handwriting recognition must be integrated. Recently, a system for bank-check processing has been proposed. It consists of several modules, each for a specific purpose: *CMC-7 Recognition Module, Signature Verification Module, Layout Processing Module, Legal Amount Recognition Module, Courtesy Amount Recognition Module* and *Amount Validation Module*. Interested readers can find the description in the paper of ref [2].

As it is well known, a bank-check processing system needs to be continuously updated in order to adapt it to the progresses made by the scientific research.

Recently, it has been subject to important changes due to the monetary revolution which has taken place on 1 January 2002 when a new currency was put in circulation: the EURO. After a short period of dual circulation, by 28 February 2002 a complete changeover to the Euro has happened in EU, old notes and coins are withdrawn from circulation in twelve countries of Europe.

The novelty of this new currency consists on the introduction of cents, so that, since the Euro has started to circulate as official currency, it was necessary to reason in cents. Hundred cents go to make up one Euro and they must be written in a decimal form.

Now the legal amount and the courtesy amount must be accurately compiled paying attention to the introduction of a numerical centesimal part, written following some new conditions. Firstly, an Euro courtesy amount is made up of two main parts, divided only by a comma “,”. These are a whole part, consisting of an handwritten digits sequence, and a centesimal part, consisting of two handwritten digits. Like the courtesy amount, also the Euro legal amount presents some modifications due to the new way of compiling a bank-check. Actually, it is made up of a worded handwritten amount, that is the whole part, and of an handwritten digits amount, the centesimal one. These parts are divided by an handwritten slash “/”.

In consequence of these changes, the previous version of the bank-check processing system, designed

---

### Acknowledgements

This paper has been supported by the Italian Ministry "Ministero dell'Istruzione, dell'Università e della Ricerca", MIUR under law 488, Project "Rete Puglia" CINI-BA - D.M. n.600, 11-11-1999.

specifically for Italian bank-checks, must be modified since it was necessary to provide new solutions for the detection of comma, of handwritten slash and also of centesimal parts (on both the courtesy and the legal amount). These modifications to the system are made without the risk of damaging the result of research.

In order to make the process of developing and managing the enormous number of programs and applications adapted to bank-checks issued for the circulation of the new currency, it was necessary to use a specific Computer Aided Software Engineering tool and a software framework. Since the Khoros software supports the entire software developing cycle (the process of developing, maintaining, delivering and sharing the programs and applications), it is suitable for the system development and maintenance [3].

Furthermore, the Khoros visual programming environment, *Cantata*, interprets the visual program dynamically. So that, the system prototype can be modified when required. It allows to minimize the side effects of maintenance operations of the system prototyping phase, to produce a system with the highest quality software designed to deliver optimum performance and accuracy with characteristics like software modularity and reuse and to rapidly provide new solutions in prototyping phase. The Khoros programming tools enforce these characteristics making easier the updating operation. The "toolbox-oriented" view of software development supports not only the creation of powerful collections of algorithms, which are always available and reusable within the environment, but it also supports the co-operative and distributed designing of bank-check processing system. For more information, readers can find Khoros organization and characteristics in the manual on ref [4]. In this paper an important modify brought on the bank-check processing system is proposed for solving the problems due to the Euro circulation. Most of the running algorithms were immediately improved and other new applications had to be developed therefore they would be placed/replaced within the system. The Euro bank-check processing prototype consists of two main phases. Firstly, a set of algorithms, collected into toolboxes for a specific application domain, was created or modified (i.e. handwritten slash detection on legal amount recognition module and comma detection on courtesy amount recognition module) in order to individuate and to check the centesimal parts on both the courtesy and the legal amount. Then these toolboxes were assembled in the visual programming environment provided by Khoros, by dynamically selecting and connecting the programs, using an engineered approach [5] to design an high modular system for bank-check processing. The novelty consists of an up to date system.

This paper allows its adaptability to the new currency circulation, since the modifications are introduced.

In Section 2 the overview of the modified modules for the Euro bank-check processing system is reported.

The description of system processing phases is shown in Section 3. Finally, the experimental results are reported in Section 4.

## 2. Overview of the modified modules of the bank-check processing system

The "Khoros" software has been used to fast create and modify the bank-check processing system.

As from 1 January 2002 Euro notes and coins have been gotten and the old notes and coins have been withdrawn within a short time. An Euro amount is written in a decimal form, consisting of a centesimal part.

Nowadays, to compile a bank-check has become more difficult because, for the first time, the centesimal currency part is introduced on both the legal and the courtesy bank-check amount. Written in the centre of the check as in the past, now the legal amount is composed of a worded handwritten whole part and of a numerical centesimal one, divided only by an handwritten slash "/".

Also the courtesy amount has been modified with the introduction of a numerical centesimal part in addition to the pre-existing numerical whole part, separated only by a comma ",". In consequence of this new way of compiling both the legal and the courtesy amount on a bank-check, it becomes necessary to provide new solutions for comma, handwritten slash and centesimal parts detection.

The data, the signature and all the other pre-printed layout zones now are still the same as they were in the old bank-check layout. So that, the modules for *layout analysis*, *image preprocessing*, *digit recognition* and *combination methods* remain unchanged in the prototype phase of the updated bank-check processing system, only grouped within each of these toolboxes, divided into subcategories according to their specific functions.

In the following a short description of the main modifications made on the Legal Amount and Courtesy Amount Modules is reported. The modifications start with a phase of programs development within the toolboxes. The tools Composer and Guise are used in this phase to create or to modify the programs described in the next modules. By using the Craftsman tool, the Legal Amount Recognition Module needed some changes. Composer and Guise are intensively used to edit, to compile and to modify the operators and their GUI visualization.

Firstly, the centesimal part is splitted by using algorithms belonging to the local and to the global approaches, respectively called *Drop\_Falling* and *Upper\_Lower\_Contour*. In the local approach a cutting line, extracted pixel-by-pixel, is detected; in the other

approach candidate splitting points are detected by the analysis of the whole pattern [6]. Each approach extracts all the connected components, that subsequently, are recognized by using some statistical algorithms with different classification techniques [5, 7]. The final decision is obtained by combining the output of several classifiers [8, 9, 10, 11]. It returns the recognized centesimal part. This module is also able to detect the handwritten slash, that divides the legal amount into two parts (the worded whole part and the centesimal one), allowing the selection of the zone between these two legal amount parts. After a training phase, the seven statistical classifiers, proposed for the classification of  $\{0, \dots, 9\}$  digits, now are suitable for recognizing an eleventh class, the handwritten slash class “/”. Each algorithm begins from a features extraction reported in the following.

In *Template Matching* technique the character itself is used as a feature vector. It is a diagonal straight line passing on the opposite vertices of the minimal box.

In the *Histograms* method, the number of detected foreground pixels is almost the same along the horizontal and vertical scanning lines.

In the *Crossing Line* method, the number of intersections between the foreground pixels and the scanning lines is almost the same along the horizontal and vertical directions.

The *Contour Slope* and the *Modified Contour Slope* extract the features basing their observations on the character contour. Necessarily, the handwritten slash contour pixels appear to be labeled with two opposite Freeman’s chain codes, denoting their scanning direction.

The *Regions* algorithm, that uses a zone technique for extracting the number of foreground pixels contained in small regions of 3\*3 pixels, has a concentration of background pixels in the regions along the straight diagonal.

In the *Characteristic Loci* method, on the basis of the slash foreground pixels, two main adjacent background areas are symmetrically labeled.

After the features extraction, each algorithm applies a specific decision rule in the classification phase in order to achieve the recognition of the handwritten slash. The final result is obtained by combining the output of these classifiers.

The Legal Amount Recognition Module also effects the off-line worded handwritten recognition [12, 13]. The related algorithms remain the same as they were in the old bank-check processing system, with a simple modification done on the legal amount lexicon. All contextual knowledge derived from lexical and syntactic information is still used in order to increase the recognition rate. Obviously the type of information to be exploited is strictly related to the bank-check issuing country. This means that the "legal amount recognition" toolbox must be specialized for new currency circulation

of bank-checks in order to handle national specific lexicon and grammar. In fact, any legal amount can be obtained by combining certain basic words according to specific updated syntactic rules [12]. For instance, 48 words are sufficient in each Italian amount (Figure 1). Please note that “unmiliardo” and “miliardi” terms are useless in bank-check amount.

uno, due, tre, quattro, cinque, sei, sette, otto, nove, dieci, undici, dodici, tredici, quattordici, quindici, sedici, diciassette, diciotto, diciannove, venti, ventuno, ventotto, trenta, trentuno, trentotto, quaranta, quarantuno, quarantotto, cinquanta, cinquantuno, cinquantotto, sessanta, sessantuno, sessantotto, settanta, settantuno, settantotto, ottanta, ottantuno, ottantotto, novanta, novantuno, novantotto, cento, mille, mila, unmilione, milioni.
--

**Figure 1. Example of updated Italian legal amount lexicon.**

At the end the LAR Module returns the entire legal amount recognition results.

Essential modifications are conducted also on the Courtesy Amount Recognition Module. The *Upper\_Lower\_Baselines* operator detects the upper and the lower baselines through the analysis of an horizontal histogram. More specifically, let  $H[i] \ i=1, \dots, M$  be the horizontal histogram of the word bitmap, let  $H[j] = \max_{i=1, \dots, M} H[i]$  be the maximum projected

value and  $A\_H = \frac{1}{M} \sum_{i=1}^M H[i]$  be the mean projected

value. By scanning the horizontal histogram from the j-th row to the M-th row in the top-down direction, the lower baseline (LBL) is the first row for which the following condition holds:

$$H[LBL] \leq A\_H.$$

Similarly, by starting from the j-th row and scanning the horizontal histogram in bottom-up direction, the upper baseline (UBL) is the first row for which the following condition holds:

$$H[UBL] \leq new\_A\_H$$

$$\text{where } new\_A\_H = \frac{1}{LBL} \sum_{i=1}^{LBL} H[i].$$

After detecting the reference lines, the digits image is divided into three zones: the upper zone, medium zone and the lower zone. The image is mainly contained within the medium zone. Subsequently, this module extracts all blocks, named connected components, by means of the analysis of handwriting contour proposed by Sobel. It is also suitable for detecting the precise position of comma in the courtesy amount. For this purpose each connected

component is analysed. Candidate commas are all the connected components laying under the upper baseline. The possible comma has  $d_{\max} = \max\{d_i\}$ , where  $d_i$  = distance between the upper baseline and the highest foreground pixel of the  $i$ -th connected component.

The seven algorithms for digit recognition [7] have been trained in order to recognize also the comma. The result of the comma recognition is obtained by using a multi-expert system, that dynamically chooses the classifiers to be combined with the *Majority Vote* combination method [9]. Once the comma has been found, it keeps the whole digit amount definitively separate from the centesimal one. These two parts are put through a segmentation phase and then through a recognition phase by combining the seven classifiers output in the same multi-expert system [8, 9, 10, 11]. The connected components rejected at this level are put into a second segmentation level by using the *Drop Falling* and the *Upper Lower Contour* algorithms [6]. These are the same procedures as they were in the old version of the bank-check processing system [2]. At the end, the CAR Module returns the recognition of the entire courtesy amount.

### 3. New Bank-check Processing System

The new prototype for bank-check processing system has been easily assembled within the Cantata visual-programming environment. For each module included in the system and for each processing phase contained within it, the most suitable glyphs were dynamically selected from those available in the toolboxes and were connected each other using an engineered approach to design an high modularity prototype. Here, the main system prototyping phases are tidily illustrated.

Firstly, the Euro bank-check is scanned at 200 dpi with 16 gray levels. Just like in the old bank-check system version, in the *CMC-7 Recognition Module* immediately the Character Magnetic Code (CMC-7) is located and recognized. Independently on the issuing bank, the CMC-7 code lays on the left lower corner of the bank-check and it is a machine printed numeric string where each digit is composed of seven vertical segments not equally distant. It codifies information about the issuing bank, the bank-agency and the check number. The bank identification code is used to retrieve from a bank reference database the information about the bank-check layout, like the position of the user-entered fields for courtesy amount, legal amount and signature. Furthermore the check number is used to retrieve from the customer reference database information useful for the signature verification phase.

The *Signature Verification Module* remains unchanged. It continues to perform the verification of the customer signature [14]. Like before the new currency

circulation, also in the *Layout Processing Module* several pre-processing phases are performed on the courtesy amount, on the legal amount and on the signature fields. These phases include filtering, guideline removal, noise reduction, slant detection and correction, pattern normalization etc [15].

The purpose of the modified *Legal Amount Recognition Module* remains the off-line handwritten amount recognition [16, 17]. Generally, in order to recognize the handwritten cursive word part many approaches are used. They can be classified on the basis of the presence or absence of the segmentation phase into segmentation-free methods, segmentation-based methods and perception-oriented methods [12]. The segmentation-free methods recognize the word as a whole by extracting a description from the whole input word and searching for the most similar reference word in the lexicon. The segmentation-based methods look for the best match between consecutive sequences of primitive segments and word's letters extracted by a segmentation algorithm. The perception oriented approaches are based on the idea that a cursive word generally consists of a periodical or quasi periodical part called "regular part" and of a non periodical part called "singularities". In order to achieve the highest recognition rate, all three approaches are available within the Legal Amount Recognition Module.

In order to recognize also the centesimal handwritten part, some algorithms for amount segmentation and digits recognition are used. The result of the recognition of the entire legal amount is used as "a priori knowledge" in the Courtesy Amount Recognition Module. This last carries out the image segmentation returning several connected components. Recognized the comma and established its position in the courtesy amount, then it returns the results of the recognition of both the centesimal and the whole handwritten numeral parts by using a multi-classifier system. Finally, the *Amount Validation Module* checks the consistency of the legal and courtesy entire amount recognition and the genuineness of the signature and it rejects the bank-check if one of the two following events occurs:

- the legal and the courtesy amounts are inconsistent;
- the legal amount OR the courtesy amount OR the signature are rejected.

### 4. Experimental results

The new Bank-check Processing System is tested on 1600 bank-checks samples extracted from a database recently developed by us. The results of the experiments on Legal Amount Recognition Module and Courtesy Amount Recognition Module are reported in the following. Each updated module is tested on the related

non-standard input data. The performances of LAR Module are evaluated in terms of handwritten slash recognition rate. It is actually obtained with the use of a multi-classifier system. The results of the experiment show that the recognition rate is 94% when the digit recognition algorithms are combined by means of the *Majority Vote* combination method. The performances of CAR Module are evaluated in terms of comma recognition rate. The results of this last experiment show that the recognition rate is 97% combining the digit recognition algorithms by means of the *Majority Vote* combination method. The inclusion of these two extra characters hasn't modified the results of the standard characters recognition.

## 5. Conclusions

In this paper some modifications are brought forward the bank-check processing system prototype, for the circulation of a new European currency. The "Khoros" system is used to modify a previous version of bank-check processing system, designed specifically for Italian bank-checks. Its CASE tools are used to support the modifications phase. Some modifications are made in order to detect both the handwritten slash and the comma, leaving unchanged the remaining parts of the old bank-check processing system. A set of toolboxes collecting algorithms for image pre-processing, digit recognition, classifiers combination and legal amount recognition is used again in the prototyping phase. The system is dynamically assembled again in a visual programming environment, in order to correctly process both the Euro courtesy and legal amount, filled in bank-checks. Since the development environment based on the Khoros system is oriented to an intensively software reuse and no compiling is required for dynamically connected operators, the updated bank-check processing system continues to be easily maintained and extended.

## 6. References

- [1] S. Impedovo, P.S.P. Wang and H.Bunke, eds., *Automatic Bankcheck Processing*, World Scientific Publ., Singapore, 1997.
- [2] S.Impedovo, A.Salzo, "A New Prototype of Bank-Check Processing System", Proc. of 4th IAPR International Workshop of Analysis Systems, Rio de Janeiro, 2000, pp.349-360.
- [3] J. R. Rasure and K. Konstantinides, "The Khoros software development environment for image and signal processing", In IEEE Transactions on Image Processing, Vol. 3, 1994, pp.243-252.
- [4] Khoral Research Inc., *Toolbox Programming Manual*, 1998.
- [5] G. Dimauro, S. Impedovo, G. Pirlo and A. Salzo "Automatic Bankcheck Processing a New Engineered System", *Automatic Bankcheck Processing*, S. Impedovo, P.S.P. Wang and H.Bunke, eds., World Scientific Publ., Singapore, 1997, pp. 467-503.
- [6] G. Dimauro, S. Impedovo, G. Pirlo, "Segmentation of Numeric Strings", Proc. of Third International Conference on Document Analysis and Recognition, Montreal, Aug. 14-16 - 1995, pp. 1038-1041.
- [7] C.Y. Suen, M. Berthod, S. Mori, "Computer Recognition of Handprinted Characters. The state of the art", Proc. of IEEE, 68 (4), 1980, pp. 469-483.
- [8] L. Xu, A. Krzyzak, C.Y. Suen, "Methods of Combining Multiple Classifiers and Their Applications to Handwritten Recognition", IEEE Trans. Syst. Man and Cybern., Vol. 22, N. 3, 1992, pp. 418-435.
- [9] S. Impedovo, A. Salzo, "Evaluation of Combination Methods", Proc. of Fifth International Conference on Document Analysis and Recognition, Bangalore, India, 1999, pp. 394-397.
- [10] S. Impedovo, A. Salzo, "A new evaluation method for expert combination in multi-expert system designing". Proc. of the First International Workshop on Multiple Classifier Systems (MCS2000), Lecture Notes in Comp. Science (1857), Eds. J.Kittler and F. Roli, Springer Verlag, 2000, pp. 230-239.
- [11] Y.S. Huang and C.Y. Suen, "A method of Combining Multiple Experts for the Recognition of Unconstrained Handwritten Numerals". IEEE Trans. On Pattern Analysis and Machine Intelligence, vol. 17, no.1, 1995, pp.90-94.
- [12] V. Di Lecce, G. Dimauro, A. Guerriero, S. Impedovo, G. Pirlo, A. Salzo, "A New Hybrid Approach for Legal Amount Recognition". Proc. of IWFHR7, Amsterdam, Sept. 11-13, 2000, pp.199-208.
- [13] J.C. Simon, "Off-line Cursive Word Recognition", Proc. of IEEE, Vol. 80, No. 7, 1992, pp. 1150-1161.
- [14] G. Dimauro, S. Impedovo, G. Pirlo and A. Salzo "A Multi-Expert Signature Verification System for Bankcheck Processing". In *Automatic Bankcheck Processing*, S. Impedovo, P.S.P. Wang and H.Bunke, eds., World Scientific Publ., Singapore, 1997, pp. 365-382.
- [15] G. Dimauro, S. Impedovo, G. Pirlo and A. Salzo "Removing Underlines from Handwritten Text: An Experimental Investigation", In *Progress in Handwriting Recognition*, S. Impedovo and A.C. Downton Eds., World Scientific, 1997, pp .497-501.
- [16] T. Steinherz, E. Rivlin, N. Intrator, "Offline cursive script word recognition- a survey". International Journal on document Analysis and recognition (IJ DAR), Volume 2, Number 2/3, December 1999, pp.90-110.

[17] G. Dimauro, S. Impedovo, G. Pirlo and A. Salzo, "An Advanced Segmentation Technique for Cursive Word Recognition". In Advances in Handwriting Recognition S.W. Lee Ed., World Scientific Publ. , Vol. 34, 1999, pp. 255-264.