Document Analysis in Postal Applications and Check Processing

21

Michel Gilloux

Contents

M. Gilloux
Seres/Docapost EBS (Groupe La Poste), Nantes, France
e-mail: mgilloux@seres.fr; michel.gilloux@laposte.net

D. Doermann, K. Tombre (eds.), *Handbook of Document Image Processing and Recognition*, DOI 10.1007/978-0-85729-859-1_22,

Abstract

The continuous improvement of general purpose document image processing and recognition techniques over the years has made possible at some points in their history the emergence of successful industrial applications. Among these, postal and check processing applications have been two of the earliest because of their economic significance, stemming from the number of items concerned, and of their particular characteristics that allowed even early document image processing and recognition techniques to produce satisfactory systems. To achieve these results, it has been necessary to integrate contributions from a majority of image processing domains, from image acquisition and preprocessing to interpretation through symbol, character, and word recognition. Through their proved return on investment, these applications have in turn highly driven and contributed to the progress of fundamental recognition techniques in many domains.

This chapter reviews their shared and distinctive characteristics, relates their history, and describes their respective state of the art through the components and techniques underlying a typical postal or check recognition system. Finally, it gives pointers to vendors and available image databases and draws some perspectives as to the development of future postal and check recognition systems.

Keywords

Address recognition • Applications • Check recognition • Postal applications

Introduction

Every day, millions of mail pieces are sent and delivered in the world. Similarly, millions of checks are written and cashed. Obviously, the usage of paper mail and of checks differs from country to country depending on history, cultural habits, and local regulations. For the United States alone, about 68 billion checks are processed per annum while postal volumes, although decreasing, were about 177 billion pieces in 2009.

Although electronic communications and payment technologies are growing rapidly, paper-based communication and payment media are still popular and rely on large-scale industrial processes to be implemented. At the involved volumes, automating further these processes remains a valuable goal with significant returns on investment since mail can be processed using automation equipment at one-tenth of the cost of manual processing.

When considering paper-based transactions, automation has basically two main aspects: (1) managing physically paper items for sorting, delivering, and archiving purposes and (2) capturing data to drive the physical manipulation of paper items and to trigger electronic transactions. For mail pieces this covers typically the reading of destination addresses to implement the sorting of mail items, while for checks this consists in particular in the decoding of the amount before crediting the payee. In both cases, automation is partly based on high-speed automatic sorting machines capable of processing more than 30,000 items per hour. The other side of this automation is the data capture involved for which document image processing and recognition techniques play a fundamental role. To bridge the gap between those two aspects, sorting machines are equipped with high-speed image acquisition devices providing multi-grey level images from every paper item processed.

In response to this huge economic challenge of data acquisition for postal items and checks, document image processing and recognition techniques have been used for several decades resulting in some of the most successful, integrated, inspiring, and profitable document image processing and recognition industrial applications. Early mono-font-printed character recognition techniques were sufficient to implement ZIP code and amount recognition and led to the first industrial applications. With the advancement of document recognition techniques, many new functions were brought into postal and check processing applications and led in turn to some of the most significant progresses in the domain by raising new problems and by funding research programs [13].

Postal and check recognition systems have in common both as a shared characteristic and as a distinctive feature with respect to other document recognition applications the necessity to appeal to a large set of different elementary recognition techniques, from image acquisition and processing to semantic interpretation, through layout recognition, text segmentation and location, printed and handwritten character recognition, word recognition, word spotting, graphics recognition, text post-processing, multilingual interpretation, system performance, and image databases.

The integration of individual document recognition techniques into industrial systems has not only been a matter of engineering but has raised new challenges of its own induced by the constraints imposed on real-life applications (e.g., computing in real-time and/or with limited computing resources, controlling system performances). Most of these new challenges have been solved by industrial vendors, but some of them have become subjects for the research community. The value added by the adaptation of basic recognition techniques and by their integration has been in fact the critical cause of the success of postal and check recognition systems.

The generalization of the use of postal and check recognition systems has also led to the extension of their typical recognition functions (ZIP codes, addresses, amounts) to new ones driven by the need to optimize their profitability and the ambition of making of paper documents a media almost equivalent to electronic ones. Profitability has been notably increased by implementing tight integration models between recognition systems and the so-called video coding man-machine interfaces used for the manual keying of non- or partly recognized data. In that domain, some by-products of the recognition are typically used to guide the human video coders in their task and thus to improve their productivity. On the other side, mail and checks have become highly technical media bearing all sorts of signs prone to document recognition techniques: bar codes, stamps, franking marks, fraud prevention marks, stickers [27], etc. Like it had happened during the development of the first postal and check recognition systems, this extended context has itself brought into the document recognition domain new problems and new techniques to solve them.

During the last period, postal and check recognition systems have also escaped from the industrial premises where they were normally used and have entered the world of individual users who may now enjoy them even on their favorite smartphone (e.g., for stamp recognition or electronic check remittance).

The history of postal and check recognition systems is not only that of two successful but specific applications of document recognition. Through the connections it shares with the development of many recognition techniques, it is the best illustration of pattern recognition put into practice.

This chapter reviews the shared and distinctive characteristics of postal and check processing applications, relates their history, and describes their respective state of the art through the components and techniques underlying a typical postal or check recognition system. Finally, it gives pointers to vendors and available image databases and draws some perspectives as to the development of future postal and check recognition systems.

Postal and Check Processing: Two Examples of the Importance of Large-Document-Volume Industrial Applications

A Common Technical Foundation

Although they do not operate on the same images and have the task of reading different type of information, postal and check recognition applications have a common technical foundation in the way they start with an image to finally output a result which is then used to perform an action [13].

Both postal and check recognition applications operate on streams of images coming from high-speed sorting machines which operate also some of the actions governed by the result of recognition. The throughput of recognition must then partly correlate to the one of the sorting machines (up to 40,000 items per hour) to absorb their inputs.

The two classes of applications have the common goal of reducing process costs by automating data entry. There is a minimal level of automation, i.e., of recognition rate, that these applications must reach to be useful and to provide a satisfying return on investment.

The recognition itself is concerned with paper documents of varying formats for which there exists no fixed a priori position for the information to be read.

This information may be expressed also in both cases under any written form: typed, handprinted, and handwritten even in cursive style. Moreover, this information is also cast into numeric or alphabetic strings and even into alphanumeric ones.

The strings and words subject to recognition are constrained by lexical, syntactic, and semantic rules which have an important role to play in the reliability of recognition.

The information to be read is also expressed under several distinct forms (the ZIP code and the city name, the numerical and letter amount) that the applications should corroborate to optimize their performance.

Finally, the results returned have to be reliable enough to be used in implementing the actions they govern (sorting a mail item, transferring an amount of money) so that recognition algorithms must meet specific constraints in terms of error rates.

All these common properties explain why postal and check recognition applications share several design principles and are made from same or related fundamental recognition techniques.

A Brief History of Postal and Check Processing Applications

Between postal and check recognition applications, the former were the first to emerge. Their advent was encouraged by the availability of the first automatic letter sorting machines. Postal address recognition applications started in the 1960s in the United States, Japan, and Europe. The first systems in use were able to recognize printed ZIP codes. In the mid-1960s, they were extended to the recognition of the entire ZIP/City/State line at the United States Postal Service.

At the end of this decade, the first systems capable of recognizing separated handprinted numbers within ZIP codes were put in application in Japan and then in Italy. These systems were successful in particular because they were concerned with a limited portion of mail for which recognition problems were constrained through the use of boxed ZIP codes. The next step was the recognition of the full address (with street name and number) for typewritten addresses which was implemented in the late 1970s in Europe, the United States, and in Japan for the Katakana and the Kanji alphabet. During the 1980s the spectrum of readable ZIP codes would be extended to touching handwritten digits. At this time, the progresses in postal applications were hampered by the limitations of computing hardware. In order to respect the real-time computing constraints, most of the systems were implemented on specialized hardware whose architecture was mimicking the structure of the algorithms they were executing.

New improvements were made when the postal operators implemented bar coding schemes that allowed associating an individual id to mail items using bar codes in a way that it became possible to run the recognition in an almost unlimited time while the item was carried from its origin sorting center to its destination. Since that time, conventional general-purpose hardware architectures have been used to implement address recognition. For example, the Remote Computer Reader (RCR) of the USPS [12, 36] was launched in the mid-1990s and allowed the recognition of full handwritten addresses, even with cursive city and street names. Beyond mere recognition, this movement was also stimulated by the availability of complete national computerized address directories. Similar projects were conducted in Europe and Japan.

With the continuous improvement of general-purpose computer power and the decrease in their cost, the real-time recognition of addresses on conventional hardware became possible at the turn of the century. Since that time, researchers and industrials have been able to focus on algorithm development without being too strongly limited by hardware constraints. During the 1990s, advantage was also taken of the evolution of address recognition systems to extend their use beyond letters to flats (mail items of typical A4 size including paper journals under plastic covers) which present additional challenges with respect to layout analysis. In the same manner, after year 2000, appeared the first address readers specialized for parcels. Since then, most of the efforts in the domain have been devoted to improving the performances of postal applications in terms of lowering their reject and error rates. In that purpose, the best results have been achieved through strategies combining several address readers from different vendors.

In the last years, the postal recognition applications have also started to extend their typical address recognition functions to the reading of other types of information carried by envelopes and in the purpose of implementing other tasks than just sorting: recognition of stamps and meter marks for fraud prevention, recognition of individual names for automatic forwarding, recognition of sender addresses for the return of undelivered items, and recognition of logos, of stickers, etc. In parallel, the vulgarization of the technology has accelerated its spreading to other countries where different languages and alphabets are in use. This internationalization of postal address recognition is also put into practice within single countries where it is applied to outbound cross-border mail.

The history of check recognition applications is to some extent a diverted branch in the history of postal applications since many vendors of check readers were spin-offs from the postal domain and were able to take advantage of the common technical foundation of the two domains to address the check market.

Check recognition applications development takes place later in time because, contrarily to mail, the value of check recognition was highly dependent on the ability to recognize handwriting. For example, IBM introduced a check deposit reader system at the end of the 1970s. But it was only during the 1980s that the technology was mature enough to provide return on investment for the users. Unlike postal address recognition, check recognition was not so much constrained by real-time computing. Since checks have built-in id numbers, the recognition of their fields could be done offline since the beginning.

The first promoters of check readers were during the 1980s and 1990s the manufacturer of check sorting machines (IBM, BancTec, NCR, and Unisys). Their systems were similar and limited to the recognition of courtesy amounts (i.e., numeric amounts). These systems were limited to the recognition of amounts written in separated digits. During the 1990s new players specialized in document recognition came into the game and brought into new technologies. At the end of the 1990s, they had improved courtesy amount recognition by accepting touching digits, and they had also introduced the recognition of legal amounts (amounts in letters) to improve system performance in a domain where error rates are critical.

As for postal applications, the last decade has seen not only progresses in terms of system performances but also the arrival of new functions like the recognition of check holder addresses, of dates, of bank accounts, and even of signatures. Again as for postal applications, the relative vulgarization of the technology has stimulated its adaptation to more and more countries, languages, and alphabets.

Common Issues in Postal and Check Processing Applications

Being real-world applications, postal and check recognition applications have also in common a number of issues that pattern recognition techniques must solve to be embedded in practical recognition systems [13, 35].

The first, and maybe the foremost, characteristic of real-world applications is that of having to deal to some extent with non-legible inputs and to remain robust to such cases. For mail and checks, a non-exhaustive list of such configurations is the following: upside down or reverted images, double feeds (image containing more than one item), images of poor quality or with highly cluttered background, etc. Recognition systems must be able to detect these configurations rather than take decisions that in such cases are by construction errors. Figures 21.1 and 21.2 show some examples of non-legible images.

The preprocessing of images may be hindered by the existence of spurious graphics (stamps, diagonal lines), especially when they cross the fields to be read. Images with low contrast (especially checks filled at a cash register) are also difficult cases. Figure 21.3 shows a check filled at a cash register.

Real-world inputs have also the peculiarity of featuring printed texts as well as handwritten ones. Specific voting technologies, which are not popular as research topics within the academic community, have then to be implemented. Because the layout analysis has to face image with both printed and handwritten texts and graphics, the location of meaningful fields may not be possible without their prior recognition. The voting between alphabets and languages may also require prior recognition.

Obviously, mail and checks concern omni-scriptor applications where there is no a priori knowledge about writers, by contrast to multi-scriptor (a limited number of known scriptors) and mono-scriptor applications. The recognition itself may be omni-font, with sometimes difficult to recognize fonts, and of course omni-scriptor with touching digits and cursive style texts. Within one field, it is also possible to find both digits and letters or words.

For practical applications, one of the principal success criteria is the ability to provide an optimal return on investment. In practice, this criterion requires to respect a certain balance between performances and costs, among which is the computing hardware needed to implement the system. As a consequence, powerful but computationally complex algorithms have to be discarded.

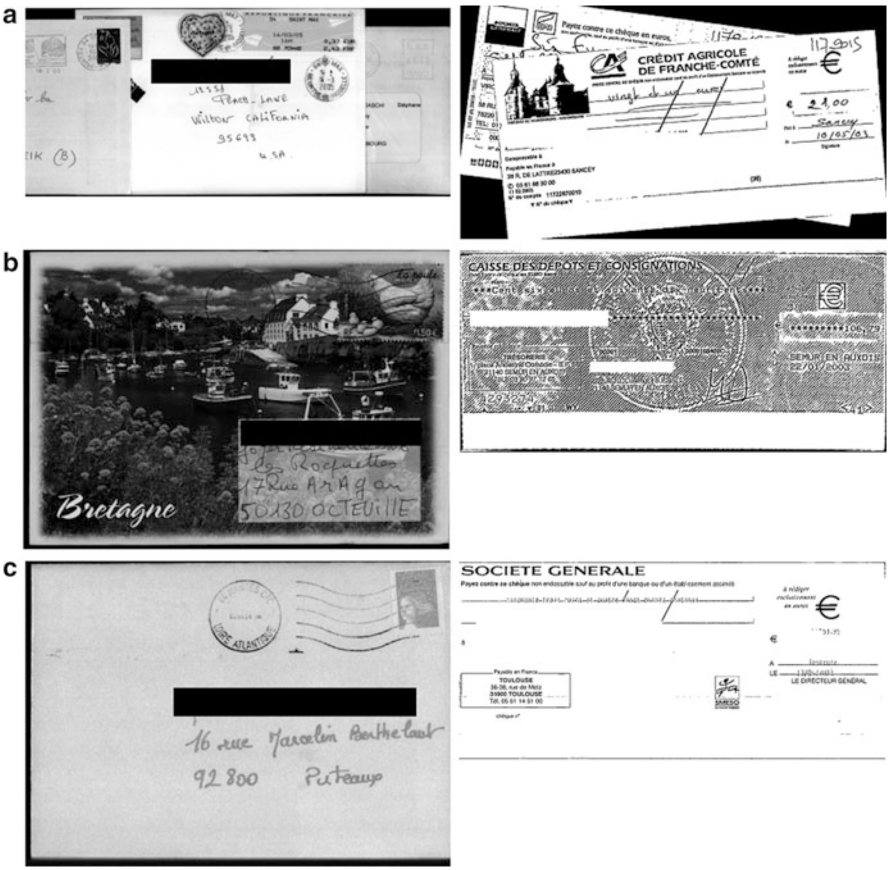


Fig. 21.1 Examples of non-legible images: (**a**) multiple items in a single image, (**b**) cluttered background, (**c**) low contrast

These applications involve also lexicons with thousands of entries (e.g., the city names in one country with all their different spellings). The recognition of handwriting within such large lexicons entails not only lower recognition rates but also computationally complex operations. The syntax of information itself is not particularly formal; many variations have to be anticipated and algorithms must be robust to non-anticipated ones.

Beyond non-legible cases and omni-scriptor issues, mail and check recognition systems are given inputs of highly variable formats. In other terms, as with omni-scriptor handwriting, the format of mail items or checks is whatever its producer wants, even for checks for which formats usually vary for a single bank. The categorization of document models is made difficult by the very large spectrum of formats and by the tiny indicia in favor of a particular pre-defined format. Therefore, technologies relying too much on mail or checks explicit models are not acceptable unless they address very frequent formats.

Fig. 21.2 Examples of non-legible images: (**a**) cross-outs, (**b**) interfering stamps

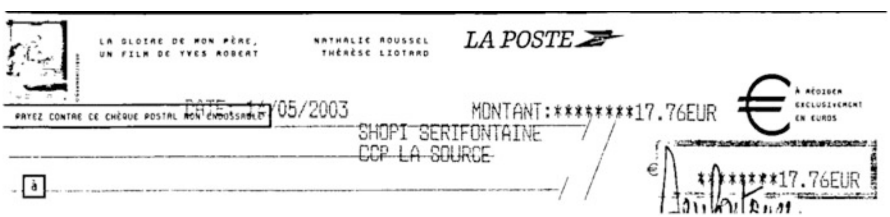


Fig. 21.3 Example of a check filled at a cash register

The semantic and contextual interpretation may also involve a large number of item values and relations between them which prevents from investigating all possibilities and requires suboptimal search strategies.

Final decision taking has to be aware of the constraints it must respect in terms of error rates. Systems must be able to take decisions but also to figure out when it is preferable not to output a decision (rejects).

For all the above reasons, postal and checks recognition systems require the integration of best quality pattern recognition algorithms but also of sophisticated and adapted strategies to overcome the limits of pattern recognition algorithms. There is in fact a trade-off between the two aspects in the sense that optimal pattern recognition algorithms would require only very simple integration. This integration strategy has also to cope with non-legible cases that are rarely studied during the initial development of pattern recognition algorithms which tends to be done over idealized cases.

The Differences Beyond Shared Characteristics

Beyond their common technical foundations, postal and check recognition applications have a number of distinct characteristics.

Postal address recognition is first concerned with the reading of information which is used straightaway to sort physically mail items. Typical delays between address recognition and sorting range from 1 to 30 s and correspond to the time taken by the mail item to go from the image acquisition device to the first sorting machine diverter. By contrast, check recognition does not require such real-time recognition because physical sorting is not directly linked to recognition. This difference puts strong constraints onto the computational complexity of algorithms used in address readers.

Another difference lies in the lexicons involved in these applications. For postal address recognition, the vocabulary corresponds to all the addresses in one particular country. In practice this means that recognition may be performed over lexicons of size of the order of magnitude of 100,000 entries (e.g., the entire city names in one country). In the case of checks, the vocabulary is limited to the words used to express amounts, typically by the number of less than 50.

Check recognition applications have the critical property of dealing with money transfer which puts strong requirements on their reliability in terms of error rates. Acceptable rates are usually of 0.1 %. Postal recognition tolerates error rates between 0.5 and 1 % because misreads have less severe consequences and are hidden behind other sources of erroneous sorting.

Although check recognition is concerned with large varieties of check formats, it is partly a form recognition application where the paper substrate is somewhat designed to facilitate recognition (e.g., complex background is usually avoided). This adaptation to recognition techniques is not complete because checks are also considered by banks as marketing items and should offer an esthetic feel in that purpose. On the other side, the appearance of postal items, at least those emitted by the general public, is not constrained.

In its usual definition, check recognition is only concerned with the recognition of an amount. By contrast postal applications, even when restricted to address recognition, have to output composite results from ZIP code to city, street name, and street number. More generally, the functional richness of postal applications lies above that of check recognition.

Postal applications tend also to become multi-language in the sense that they have to process addresses expressed in various alphabets and syntaxes within a single flow. Check recognition is rather applied to only one language and alphabet at a time although there exist applications for many different ones.

All these differences motivate the use of distinct fundamental recognition techniques or of distinct strategies for integrating those techniques in a system depending on the type of application at hand.

A Survey of Common Techniques for Postal and Check Processing Applications

A Unified Model for Postal and Check Processing Applications

Due to their common properties, the general design principles of postal and check recognition applications may be described at a certain level with the same concepts and decomposition into a set of stages.

First, they have both to preprocess the input image to make it more amenable to pattern recognition algorithms. This preprocessing may consist in removing background pictures or graphics, improving the image quality, and binarizing multi-grey level images.

Then they have to perform layout analysis to categorize the signs present in the image and to locate those which bear the information sought. During these two first steps, it may be useful to determine how the physical item is oriented in the image (upright or upside down) and whether it belongs to a preknown category with a recognizable format. The signs retained for further processing may have to be aggregated into more abstract entities or fields (address blocks, numerical or alphabetic amount). Depending on the nature of each of these aggregations, they have to be read using different and suitable sets of recognition techniques (printed vs. handwritten texts, numeric vs. alphabetic, language/alphabet). For that purpose, a voting step is performed. The results of the recognition performed on each field are then submitted to lexical and syntactic checking which has the task of filtering non-legible recognition results and which outputs a set of candidate interpretations for each field. The candidates for the different fields recognized in an image are in turn corroborated to check their relative consistency (does the ZIP correspond to the city? does the numeric amount agree with the letter amount?) and their agreement with external knowledge sources (list of possible amounts for a check remitted with a remittance slip, address database). Like lexical and syntactic analysis, this semantic or contextual checking filters out the remaining candidates and combines the values of the retained ones in a more abstract meaning (address code, canonical amount). Finally, a decision is taken in view of these retained candidates and of the estimated confidence for their recognition. The final decision may be one single interpretation or a list of candidate interpretations together with their confidence scores, or a reject code when the confidence in the best candidate itself is not sufficient for the task at hand or when it was not possible to perform the entire above process when a component step failed.

This typical bottom-up view of postal and check recognition may often be described as top-down by considering retroaction loops between the different component steps in which the results of a given stage are given back to the previous ones before they are run again. Another technique that breaks the univocal bottom-up flow of information is the use at each step of candidate sets rather than single candidates together with partial confidence scores. By using this technique, early decisions may be postponed until all the necessary information to take them has been gathered.

It is also possible to repeat the entire process using different systems from different sources operating on the same image and using modified versions of the image (perturbations) [13] before the results returned by the different processes considered as a committee of experts are combined into a unique final decision. Through this scheme, fewer rejects (no decision) may be achieved and final confidence scores may be raised.

This common technical foundation between postal and check recognition applications illustrates to the best the importance of the integration strategy of basic pattern recognition algorithms in practical applications. However, beyond this common technical foundation, postal and check recognition presents a number of differences in the recognition algorithms they use themselves.

Image Acquisition Devices

Image acquisition devices for postal items and checks are primarily high-speed scanners embedded into large-throughput sorting machines. Their technology is specifically designed to be compatible with these high throughputs and with physical items of varying mechanical properties (different sizes and thickness, reflective transparent windows, plastic covers, etc.).

These scanners typically produce both multi-grey level images, whether compressed (JPEG) or not, and binary images. Due to the considered throughputs, binarization is usually performed with specialized hardware embedding adaptive binarization algorithms [28] (see also ►Chap. 2 (Document Creation, Image Acquisition and Document Quality) in this handbook).

In the recent years, check processing processes and regulations have allowed distributed check capture consisting in the acquisition of check images at the bank branches through the use of low-volume scanners.

Image Preprocessing and Enhancement

Image binarization itself is usually not sufficient to eliminate noisy graphics present in the images. Preprocessing and image enhancement algorithms are useful for eliminating noise, for example through classical line extraction algorithms [28] (see also ►Chap. 4 (Imaging Techniques in Document Analysis Processes) in this handbook).

Document Registration and Categorization

Document categorization is the process by which an application detects within an image an instance of a known document format. Whenever this happens, it is then straightforward to locate the fields in the document that the application has to recognize. There exist a number of techniques for performing document categorization (see ►Chaps. 7 (Page Similarity and Classification) and ►19 (Recognition of Tables and Forms) in this handbook). Some of these techniques require a manual pre-modeling of recurring formats through the definition of anchor points (typically salient features in an image like square angles, crossings). Other techniques are able to determine automatically these anchor points through detectors of visual vocabularies. In both cases, anchor points are used to form a signature of every image and to match this signature with a database of known document models.

Layout Analysis: Region of Interest Location

In the cases where an image does not correspond to a preregistered template, its layout must be analyzed through general principles governing the relative and individual placement of fields (see ►Chaps. 5 (Page Segmentation Techniques in Document Analysis) and ►6 (Analysis of the Logical Layout of Documents) in this handbook). Methods that discriminate texts from other types of graphics are usually employed in that purpose. Due to the severity of real-time computing constraints in such applications, strategies relying on full text recognition of the image prior to the analysis of their layout are not acceptable.

Layout analysis methods typically rely on morphological processing which take into account the typical dimensions of characters and of lines of text. In order to stay robust to the range of possible text sizes, multi-scale layout analysis is sometimes employed [30].

Once text lines have been extracted, their grouping into meaningful fields is necessary. Here again, morphological analysis techniques at the line level are the best candidates for this task. Morphological analysis relies here on the typical distance and justification of lines of text within a field and on the expected location of each field within an image (address blocks or lines of a legal amount are usually left justified; address blocks are preferably in the lower right part of an envelope). As for lines of text, multi-scale analysis is useful for providing robustness to variations in line spacing. Multi-scale (or multi-resolution) strategy is also a design principle akin to perturbation techniques where a given image is processed several times after being modified (perturbation) in different ways.

Special cases of layout analysis must take into account images with vertical or upside-down orientation. In that context, morphological techniques may be used to extract features correlated with the orientation of lines of text. The proportion of ascenders with respect to descenders or of upward with respect to downward valleys are some of the most popular features for implementing text orientation detection.

The above layout analysis is also in charge of giving to each located field a category relevant for the application. In most cases, the rough location parameters used to locate the fields are sufficient for the attribution of a category.

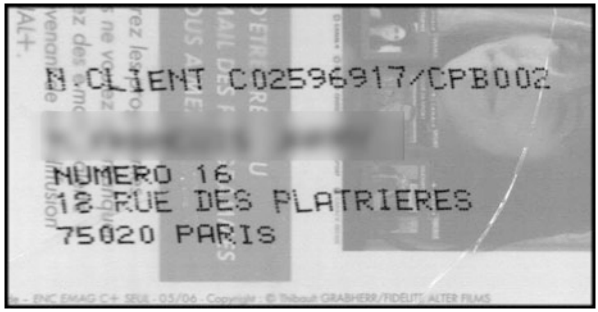


Fig. 21.4 Address printed over a plastic cover with an ink-jet printer

Voting for Image Types

Within applications dealing with printed or handwritten texts without any prior knowledge of whether a given text is printed or handwritten, a voting scheme is necessary to give a class to a field before it is recognized with methods specific to this class (see ►Chap. 5 (Page Segmentation Techniques in Document Analysis) in this handbook).

Morphological features of lines of text like their dimension, the density of black vs. white pixels, and the alignment of characters bottoms and tops are typical features used to characterize a line of text. Ad hoc or general-purpose classification methods like neural nets are used to implement a voter based on such features.

Character and Word Recognition

Character and word recognition methods obviously depend on the nature of the text to be recognized.

For printed texts, today general-purpose OCR engines (see ►Chap. 10 (Machine-Printed Character Recognition) in this handbook) are powerful enough to take into account a majority of cases. However, there exist cases for which such general-purpose tools fail because the quality or nature of printed characters is outside their domain of application. Images with poor-quality printing (e.g., characters only partially formed) or with special fonts (e.g., printings of a cash register, dot matrix printing, and ink-jet printing over plastic covers) constitute such cases. Figure 21.4 shows an address printed over a plastic cover with an ink-jet printer. More robust recognition methods have been employed for such cases which are mostly derived of methods used for handwriting recognition (e.g., Hidden Markov Models – HMM) [38].

Contrarily to conventional general-purpose OCR engines, top-down recognition strategies and the use of candidate lists require OCR recognition to output character hypotheses together with confidence scores. For that reason, OEM (original equipment manufacturer) rather than stand-alone versions of conventional OCR engines are integrated into postal and check recognition applications.

For handwritten texts, various classes of recognition methods should be applied depending on the nature of the fields they are processing.

For numeric fields like courtesy amounts and ZIP codes, segmentation-recognition methods [3] are privileged because there exist no lexicon for the considered fields which are rather respecting syntactic constraints best implemented through the segmentation-recognition principle (see ▶Chap. 11 (Handprinted Character and Word Recognition) in this handbook).

By contrast, alphabetic fields are best recognized through word recognition techniques relying on the existence of a lexicon (see ▶Chaps. 11 (Handprinted Character and Word Recognition) in this handbook and ▶12 (Continuous Handwritten Script Recognition) in this handbook). These word recognition techniques usually start by extracting features of text images and then classify the feature representation of words through a suitable classifier. Depending on the size of the lexicon associated to a given field, specific feature extraction and classification methods are used [38].

Prior to word recognition, it is often necessary to segment text into words or to input whole sentences to recognition engines. Explicit segmentation techniques relying on gaps between words or implicit ones which perform some sort of keyword recognition have been employed [25] (see also ▶Chap. 24 (Image Based Retrieval and Keyword Spotting in Documents) in this handbook).

The distinction between numeric and alphabetic texts is not always straightforward (e.g., delivery lines in addresses include both street names and street numbers; legal amounts in check may involve the cents part expressed through numbers). In such cases, word recognition may also be endowed with the capacity to detect numeric portions.

Decision Making

Decision making is basically concerned with taking into account several lists of candidates for a given field value (see ▶Chap. 6 (Analysis of the Logical Layout of Documents) in this handbook). Those lists may come from several regions of an image where a given value is expressed in different ways (e.g., numeric and alphabetic), from regions with different but correlated meanings, from different systems ran in parallel, or from different outputs of a given system after different perturbations of the original images. In all cases, the lists of candidates may be seen as the advice of different experts within a committee of experts. General methods for integrating multi-expert decisions [21] are applied to come out with a single result.

Another issue is the assessment of the reliability of the final decision in order to reject it in case it does not conform to the trustworthiness required by the application (see ▶Chap. 30 (Tools and Metrics for Document Analysis Systems Evaluation) in this handbook). General-purpose multi-expert decision methods usually embed the computation of a confidence value. However, these methods are concerned with the assessment of the ambiguity of the input data but do not perform well in cases where the input data is outside the scope of legible cases for the application (outlier rejection). Practical recognition systems take into account such cases by modeling non-legible configurations [31].

Postal Applications

Specific Issues in Postal Applications

Specific issues in postal applications start with the layout analysis of images of postal items.

The criterion used to locate and analyze address blocks is specific to postal applications. Address block location is indeed one of the most difficult subtasks in address recognition since it may even consist in locating a text within a cluttered background of texts in the case of a flat item or even to detect the address block within several views of the same object in the case of a parcel with images of its six faces. Figure 21.5 illustrates the difficulty of locating addresses on flats.

Address recognition applications are also characterized by the use of large-volume lexicons (e.g., more than 100,000 city names). Recognition methods have to be fast enough for operating in acceptable times over such large lexicons. They have also to be precise enough to return reliable results although ambiguity increases with lexicon size.

Large-volume lexicons are part of address databases which have to be matched with the address recognized on a mail item. This matching is a complex process which has to take into account the uncertainty of recognition and the discrepancies between the way an address is actually expressed on a mail item and its canonical form.

The texts to be recognized within postal applications are made of different fields written under different formats (numerals, words). The integration of several fundamental recognition techniques designed only for a specific class of images is also one of the challenges of address recognition. Figure 21.6 shows street lines with both numerals and alphabetic words.

Finally, the variants in address syntaxes are numerous and necessitate the thorough collection of all possible cases during the design of a postal recognition system.

A Survey of Specific Document Analysis Techniques Used in Postal Applications

The following sections describe the main processing steps in a postal address recognition system: address block location and extraction, ZIP code location and recognition, city name location and recognition, street name and street number location and recognition, and decision making, and the specific techniques

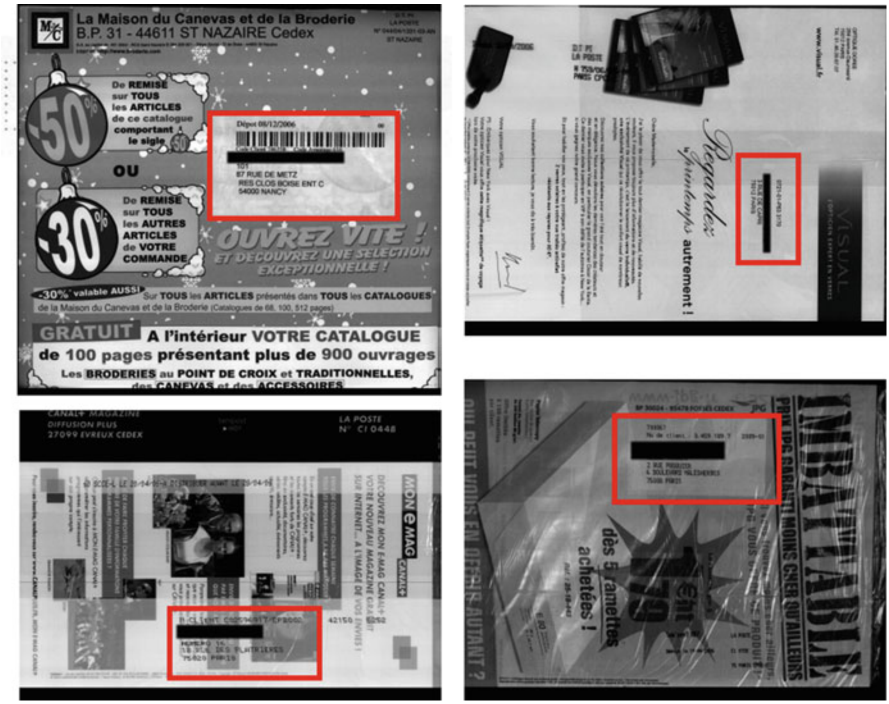


Fig. 21.5 Examples of address blocks on flats (*red rectangles* show the location of the address)

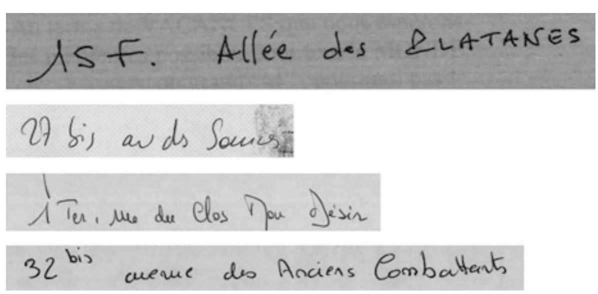


Fig. 21.6 Examples of street lines

commonly used to implement them. These steps are better understood if we describe first the overall recognition strategy for a postal address recognition system.

Overall Strategy

A postal address is defined through different components whose form and possible values belong to sets of different complexities and sizes. The methods suited to recognize each of them vary accordingly in their performance and in their need for computing resources. Moreover, the components are tied in various ways which reduce the potential value for a given one based on another one. For both reasons, the recognition strategy of an address recognition system is generally organized to implement the recognition of the components in an order that goes from the most general to the most specific (i.e., from the less constrained component, e.g., the ZIP code, to the most constrained one, i.e., the one whose value depends the most of the others ones, e.g., the street name which is constrained by the ZIP code and the city) and from the most easily recognized (digits strings) to the less easily recognized (words in large vocabularies). This strategy suffers from exceptions when the first components in this order list are not recognized or are missing. In such cases, recognition has to start with the most difficult task, and the recognition techniques applied are chosen for their capacity to face that difficulty within the limits of the constraints imposed to the overall address processing, especially in terms of computing time.

There are two classes of strategies which form the basis of an address recognition system and which are mixed at a certain degree depending on what is found on the image being processed (for the sake of clarity, we present them below in the simplified case of a ZIP code-city-street name-street number address. More complex cases appear when considering the destination country, the state, the extended ZIP code, the addressee name, etc., but without changing the philosophy of this overall strategy):

1. Progressive refinement strategy:
   (a) Locate and recognize the ZIP code.
   (b) Locate and recognize the city name within the cities compatible with the ZIP code.
   (c) Locate and recognize the street name within the street names of the recognized city.
   (d) Locate and recognize the street number within the range of allowed street numbers in the recognized street.
2. Bottom-up strategy:
   (a) Locate and recognize the ZIP code.
   (b) Locate and recognize the city name.
   (c) Locate and recognize the street name.
   (d) Locate and recognize the street number.
   (e) Integrate recognition hypotheses into a single interpretation based on the consistency constraints between ZIP, city, street name, and number.

The progressive refinement strategy is best suited for addresses with difficult to recognize components, in particular handwritten addresses. It is in particular useful for progressively reducing the size of lexicons among which recognition takes place up to the point where these sizes are compatible with available technologies. On another hand, progressive refinement strategies are not robust to recognition errors or even absence of recognition results in early stages since these early stages condition all the following ones.

The bottom-up strategy is more adapted to addresses with perfectly or almost perfectly recognized components for which the main difficulty is address interpretation rather than recognition, in particular printed addresses.

In the extreme, a simple and naïve implementation of the bottom-up strategy aims at extracting from the mail item image a string of ASCII characters corresponding to the address text and then to pass this string to a general-purpose ASCII address interpretation stage such as those used by direct marketing address management companies or even made available in map services on the Internet like Google maps. Although it may work for some simple cases, this naïve strategy, and in general most bottom-up strategies, fails on more difficult cases because on the one hand they assume that it is possible to recognize an address as any other text, which is highly difficult for ordinary handwriting, and, on a second hand, because they overlook some of the information available in the original image as, for example, all the recognition errors which are not taken into account by general-purpose ASCII address interpretation software. The only merit of the naïve implementation of the bottom-up strategy is the ease with which running systems may be quickly assembled by stacking general-purpose modules.

More complex implementations of the bottom-up strategy are also useful for bypassing the errors or the difficulties that may occur in early stages in progressive refinement strategies. For example, addresses with erroneous ZIP codes require the recognition of city names without prior knowledge of the ZIP code. Only bottom-up strategies will start by recognizing city names, but the price to pay for that is a larger lexicon size and thus a more difficult recognition task.

In practice progressive refinement and bottom-up strategies should be associated to get the best from each.

Another dimension for analyzing the architecture of address recognition systems is the distinction between handwritten and printed addresses. In the case of addresses printed with normal quality, address recognition reduces to address interpretation since recognition is easy so that the input data is equivalent to an ASCII text.

By contrast, printed addresses with low quality may be assimilated to handwritten addresses so that all techniques described below apply to both cases.

Address Block Location

Various address block location techniques have been proposed for postal applications [14, 30]. In their majority, they rely on morphological analysis techniques used to group characters into words and lines, then lines into texts and to decide whether a text is a good candidate for being an address block. Some specific techniques are also used which take advantage of a number of peculiarities of mail items. Special marks for explicit address location have been proposed and suggested to large-volume mailers, sometimes with the help of incentives. Akin to marks are the various boxing systems used in different countries. The shape of boxes, or their predefined location, is a cue to locate further the address block. Opportunistic marks may also be made of bar codes often included within addresses. Transparent address

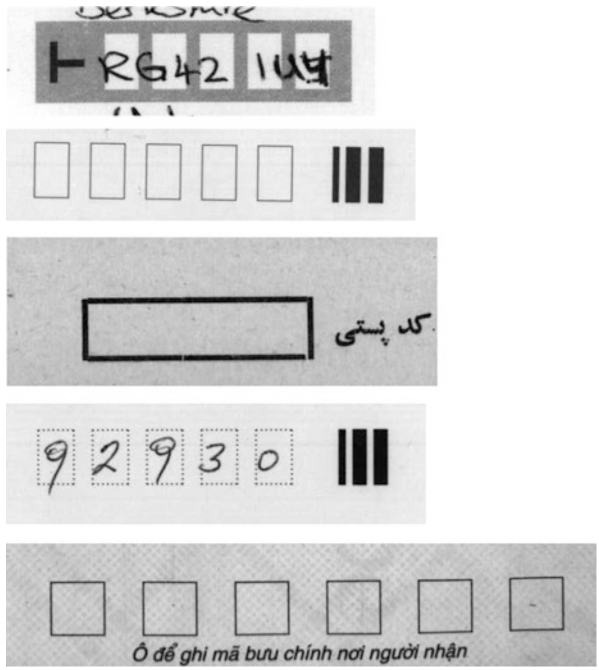


Fig. 21.7 Address block locators and ZIP code boxes

windows may be also subject to specific detection algorithms. Figure 21.7 shows different address block locator marks and ZIP code boxes.

ZIP Code Recognition

ZIP codes are usually placed in the leftmost or rightmost position within the last line of an address block. Therefore, typical ZIP code location and recognition methods, especially under progressive refinement strategies, consist in attempting a recognition in these expected locations. However, the exact position of the ZIP code may not be limited by a sufficient gap between the ZIP and the generally following or preceding city name. The recognition method used must then be robust to this uncertainty. The same type of difficulty is increased for bottom-up strategies which are particularly interesting when the ZIP is not in its normal position and which have to guess this position from the shape of words isolated during address block location.

ZIP codes can be considered as character strings. They are thus amenable to recognition methods specific to character strings.

In the particular case of ZIP codes, simple string recognition strategies have been used because in many cases ZIP codes are written with cleanly separated characters and also because the number of characters in a ZIP is usually known in advance. Within such strategies, characters positions are first isolated and then submitted to individual recognition in a purely sequential manner.

However, in many real cases, ZIP codes are written with touching characters and/or in an unknown number. General character string recognition techniques are then required.

There exist commonly two classes of general methods for the recognition of ZIP code strings: those with explicit character segmentation and those with implicit segmentation.

Explicit character segmentation-based methods implement the segmentation-by-recognition principle in which potential segmentation configurations are validated or invalidated by character recognition.

Common methods for detecting potential segmentation configurations rely of heuristic rules which recognize known configurations of touching characters. For any particular ZIP code recognition task addressing a certain character alphabet and given ZIP codes syntax and writing habits, specific rules have to be crafted in view of a large number of examples. Therefore, the sets of such rules change whenever the ZIP code system changes so that they cannot be described in general.

Symbol recognition also inherits from general-purpose methods for character recognition like statistical or neural network-based classifiers and adapted feature sets. Since ZIP codes are generally composed of digits and/or letters, they do not require other than by exception specifically designed recognizers adapted to other symbols than digits or letters. However, recognition engines for digits or for letters trained on images from other tasks than the particular ZIP codes to recognize shall suffer from degraded performances because the shape of digits and letters within ZIP codes is often influenced by the fact that they appear within a ZIP code. For example, the most frequent ZIP codes tend to end with several "0" written in smaller sizes than the rest of the ZIP and often merged by ligatures or by overlapping. In such cases, specific recognizers trained on corresponding examples and using adapted features are used.

Figure 21.8 shows an example of the recognition of a five-digit ZIP code based on the segmentation-by-recognition principle.

City Name Recognition

City name recognition in postal application is a task directly dependent of the overall strategy which is implemented.

With progressive refinement strategies, city name recognition is guided by prior ZIP code recognition and has then to operate over a limited-size lexicon (typically of size 10–100) made of names of cities consistent with ZIP code recognition hypotheses. For that specific case, different word recognition methods in the literature [39] have been used in industrial systems with the most popular being analytical word recognition based on individual character recognizers and holistic word recognition based on Hidden Markov models. Both classes of methods have their strength and weaknesses for city name recognition and usually they are mixed into real systems. The fact that city names are often written in handprinted style is an argument in favor of the use of analytical methods.

In the case of bottom-up strategies, city name recognition is confronted to large lexicons of typically 10,000–100,000 entries. Not all word recognition methods are

Fig. 21.8 ZIP code recognition based on the segmentation-by-recognition principle

tractable for such a task because of their computing time, which may be linear in the number of lexicon entries, and because recognition must be of high performance to isolate the correct answer within thousands of candidates. In that case, analytical methods are preferred, which combined to tree organizations of lexicons, achieve high speed and precise recognition especially when names are written in separated handprinted letters.

Figure 21.9 shows an example of the recognition of a city name using discrete word representation features and a Markov model for scoring.

Street Name Recognition

Street name recognition has to be distinguished from city name recognition because it has to take into account not only words but whole phrases made of several words with different roles and variations (numbers, street types, names, titles, abbreviations). Another difference lies in the fact that lexicon sizes for street name recognition may range up to thousands of entries (the number of street names in a large city) independently of the chosen overall strategy.

Fig. 21.9 City name recognition using discrete word features

Input city name image

Potential character segmentation

Feature representations

Feature set1: `–s –s –@ –s –n –s –n –u –# –s –s –u –n`

Feature set2: `d3 u3 U3 N2 U3 S2 N6 V1 T3 d3 L1 E2 E3`

Candidate city names and scores

54300 LUNEVILLE	4. 93e-26
52300 FRONVILLE	4. 20e-29
54300 BIENVILLE	4. 40e-30
REJECT	3. 08e-34

Final decision

54300 LUNEVILLE

Two classes of specific recognition strategies are implemented for street name recognition.

Keyword-based strategies perform recognition of words rather than on whole phrases or lines. Here again, both analytical and holistic word recognition methods may be used and implemented in a very direct way on the word images isolated during address block location. However, keyword-based strategies suffer from several shortcomings. First, since they rely on word segmentation hypotheses, they may miss true word positions, e.g., when words in the street name are written without gaps. Secondly, they do not take into account relative word positions which constitute a useful cue for street name recognition.

By opposition, phrase-based strategies attempt to match the whole street name to candidates in the lexicon. Doing this they not only rely on content word recognition, but also they take advantage of other words in street lines like street type and article or prepositions. In that goal, methods based on a flexible alignment between images and lexicon entries are particularly suitable. All the recognition techniques using the framework of Hidden Markov models have this property and many have been used in industrial systems. The interest of phrase-based strategies may be lowered by their computational complexity and then the related computing time. Heuristic search strategies and tree-organized lexicons are used to solve these issues. Figure 21.10 depicts an example of the recognition of a street name using a keyword-based strategy.

Input street name image

Potential character segmentations

Feature representations

Feature set1: `–s –# –s –s –s –# –s –s –n –s –u –s –n`

Feature set2: `L1 y2 R2 u2 u3 e1 c1 a1 E1 dN4 m6 o1 t1`

Candidate street types, positions and scores

RUE	.... RRuuuue_...................	1.05e–31
LA	LLa_............................	1.69e–33
AVENUE	avvveennuuuue_..................	4.32e–34

Candidate street key words and scores

DE_CARNOT	1.09e-30
CARNOT	2.79e-32
DES_CARMES	1.51e-33
LES_CARMES	1.39e-33
BACCARAT	6.47e-35
LEOMONT	7.83e-37

...

Candiate full street names and scores

RUE_CARNOT	1.01e-21
RUE_LEOMONT	2.83e-26
REJECT	2.88e-26

Final decision

RUE_CARNOT

Fig. 21.10 Street name recognition with a keyword-based strategy

Street Number Recognition

Street number recognition is similar to ZIP code recognition except for the number of digits, which is not known in advance; for the precise position, which is not clearly marked with respect to the street name in the same line; and for the composition of the number, which may not include only digits (e.g., commas, dashes).

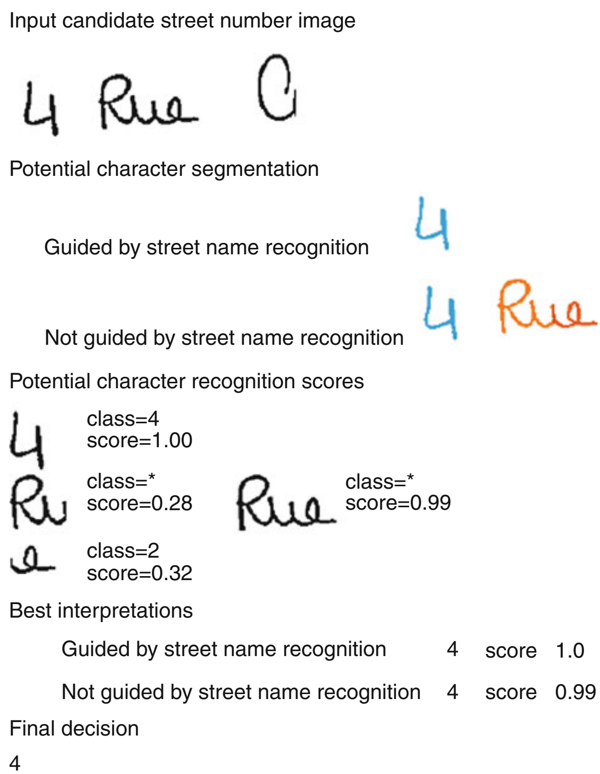


Fig. 21.11 Street number recognition

There exist two main classes of strategies for solving the above issues.

General-purpose character string recognition methods are sometimes applied to the expected position of the street number and augmented with additional character recognizers. In particular, the additional recognizers are in charge of detecting street number limits (e.g., commas) whenever the position is not precisely known.

Another class of strategies relies on phrase-based street name recognition. Here it is assumed that the flexible alignment between street names and street line images has a by-product which is the exact position of the street number within the line.

Figure 21.11 describes an example of the recognition of a street number based on the segmentation-by-recognition principle.

Lexicon Representation

Postal recognition applications have also the particularity of relying on large-volume address databases with tight interaction with pure recognition techniques. Various methods which organize these interactions have been proposed and applied. They generally rely on heuristic search techniques of the artificial intelligence domain [22,26]. Completely bottom-up strategies (i.e., text recognition then address matching) are in general not as powerful as the former methods.

Table 21.1 State-of-the-art recognition rates in postal recognition applications

		Letters (%)	Flats (%)	Parcels (%)
ZIP code	Printed	>97	>90	>70
	Handwritten	>85	>80	>40
Street name and number	Printed	>95	>90	>60
	Handwritten	>65	>60	>20

Included in the relations between recognition and address databases are the algorithms used for recognizing words within large-volumes lexicons. All postal applications embed in that respect a specific compiler whose role is to transform the original address database into a representation suitable for fast recognition (this step is known under the name of Address DataBase Generation – ADBG). Tree-type representations and associated heuristic search techniques are often implemented [39].

The recognition of composite fields in addresses (ZIP code line with ZIP code, city and state, delivery line with street name, street number) requires also specifically designed methods which are able to recognize a word or phrase within a larger text and which are capable of discarding noisy signs as well as numerical or alphabetic ones. Keyword handwriting recognition techniques have been implemented in address readers.

State of the Art in Postal Applications

Due to the differences in alphabets, languages, address syntax, address databases, and potential numbers of distinct outputs (e.g., number of possible ZIP codes) from country to country, it is not possible to give common state-of-the-art performance figures for all postal recognition applications [11, 12, 17, 32–34, 36, 37, 41, 42]. These figures also vary depending on the category of mail items considered: letters are easier to recognize than flats which are easier to recognize than parcels.

However, Table 21.1 gives the typical recognition rates achieved in today applications (at least with Latin alphabets) which are usually reached with an error rate around 1 %.

The figures demonstrate that there is still room for improving recognition technologies in postal applications.

Reference Datasets

A number of image databases issued from postal recognition applications have been published and have become for some of them almost standards, not only for the evaluation of full applications but also for the testing of fundamental techniques. Among these are the Cedar database from the State University of New York in Buffalo (http://www.cedar.buffalo.edu/Databases/), the MNIST database of handwritten digits (http://yann.lecun.com/exdb/mnist/), the MNIST database of handwritten upper-case letters (http://publications.idiap.ch/index.php/publications/show/709), and the KAIST-POWORD Handwritten Postal Hangul Word Image Database (http://ai.kaist.ac.kr/Resource/dbase/Address/POWORD.htm).

Vendors

Postal recognition systems are designed and marketed by a number of specialized companies throughout the world among which the most important are given in Table 21.2.

Check Processing Applications

Specific Issues in Check Processing Applications

Through their particular characteristics, check recognition applications present a number of specific challenges for pattern recognition techniques.

Layout analysis in check recognition applications is facilitated by the design of checks which makes them special types of forms. However, this design has to reconcile the interest for easy to recognize documents with the task of conveying the best vision of a bank. The latter often gets priority so that check forms are encumbered by textured backgrounds and preprinted texts likely to appear behind the fields to be recognized. The separation of such graphics from useful text must be done prior to recognition.

In the useful fields themselves, one may find non-numeric signs (slashes, dots, commas, currency sign, currency name, pound signs, fraction bars, etc.) often used to prevent the fraudulent modification of amounts. The elimination or interpretation requires their specific recognition. Figure 21.12 shows examples of courtesy amounts with non-numeric signs included.

Courtesy and legal amounts feature all sorts of syntactic and lexical variants specific to the expression of amounts and within that context particular to each alphabet and language. The collection of all possible options is one of the necessary tasks in the design of a check recognition system. Figures 21.12 and 21.13 show examples of various amount syntaxes.

In order to implement amount recognition, and in particular legal amount recognition, one has to design ways of controlling digit or word recognition by the collected syntax. For courtesy amount recognition, this is made more difficult by the degrees of freedom in the vertical positions of graphemes which somewhat makes that recognition akin to that of mathematical formulas.

Finally, the high requirements in terms of low error rates compel the decision stages to output very accurate estimates of recognition reliability in the form of confidence values.

Table 21.2 Indicative list of vendors of postal recognition software and systems

Company	Website/comments
A2iA	www.a2ia.com Based on his experience as a pioneer in check recognition software, A2iA has developed general-purpose handwriting recognition software engines which have been integrated by system integrators into postal address recognition systems in association with technologies from other vendors in countries like France, Germany, and the UK
Bell & Howell	www.bellhowell.net Bell & Howell, LLC is a provider of solutions and services for paper-based and digital messaging. Bell & Howell provides address recognition systems adapted to presorting of outgoing mail for mailing companies in the United States
Elsag	www.elsagdatamat.com/Automazione.htm Elsag is a provider of global solutions for postal sorting including sorting machines, information systems, image acquisition systems, hardware and software for address recognition and envelope interpretation, video coding systems. Elsag has provided address reading systems in many countries including Italy and France
Lockheed Martin	www.lockheedmartin.com/capabilities/it/ Lockheed Martin is an integrator of address recognition systems associating its own technologies to the technologies of other vendors. Lockheed Martin is in particular the integrator of the address readers used by the USPS and in the UK
NEC	www.nec.com/global/solutions/postal/content/systems.html NEC is a provider of global solutions for postal sorting including sorting machines, information systems, image acquisition systems, hardware and software for address recognition and envelope interpretation. NEC has provided address reading systems in many countries including Japan, Finland, and Asian countries
Parascript	www.parascript.com Parascript is a provider of document recognition software. As part of this activity, Parascript has developed an address recognition software which is integrated by system integrators in countries like the USA and the UK
Prime Vision	www.primevision.nl Prime Vision, a former subsidiary of the Dutch national postal operator. Prime Vision was founded from the Netherlands National Postal and Telecommunications Corporation (PTT) research department. Prime Vision develops document recognition solutions in the domains of Logistics, Banking, Fraud prevention, and Regulation enforcement
	Prime Vision has implemented solutions from Asia Pacific to Australia, the Middle East, and North America
RAF Technology	www.raf.com RAF Technology, Inc. is a worldwide leader in advanced pattern and image recognition, intelligent information extraction, and data verification solutions for Courier, Express, Postal, Warehouse, Transportation, and Field Service, Enterprise, and Government agencies
	RAF provides flexible software solutions that enable organizations to more effectively process mail

(continued)

Table 21.2 (continued)

Company	Website/comments
Seres	www.seres.fr Seres, a subsidiary of the French La Poste, is a provider of general-purpose document recognition systems and software. Seres has developed the primary address recognition software used at La Poste
Siemens	www.mobility.siemens.com/mobility/global/EN/LOGISTICS/POSTAL-AUTOMATION/Pages/postal-automation.aspx Siemens is a provider of global solutions for postal sorting including sorting machines, information systems, image acquisition systems, hardware and software for address recognition and envelope interpretation. Siemens has provided address reading systems in many countries including Germany, the USA, and recently in Arabic countries
Solystic	www.solystic.com Solystic is a provider of global solutions for postal sorting including sorting machines, information systems, image acquisition systems, hardware and software for address recognition and envelope interpretation
Toshiba	www3.toshiba.co.jp Toshiba is a provider of global solutions for postal sorting including sorting machines, information systems, image acquisition systems, hardware and software for address recognition and envelope interpretation. Toshiba has provided address reading systems in many countries including Japan, France, and Asian countries
Vitronic	www.vitronic.de Vitronic is a provider of machine vision systems and software in industry, logistics, and traffic control.
	Vitronic has developed address recognition systems including image acquisition devices, hardware, recognition software, and video coding systems adapted to parcel sorting

A Survey of Specific Document Analysis Techniques Used in Check Processing Applications

The following sections describe the main processing steps in a check processing system, namely, field location and extraction, courtesy amount recognition, legal amount recognition, and decision making, and the specific techniques commonly used to implement them. These steps are better understood if we describe first the overall recognition strategy for a check recognition system.

Overall Strategy

Beyond general-purpose corroboration methods, check recognition may take advantage of the fact that courtesy and legal amounts on a given check have normally the same value. The overall recognition strategy commonly uses this property to implement courtesy amount recognition and/or legal amount recognition based on

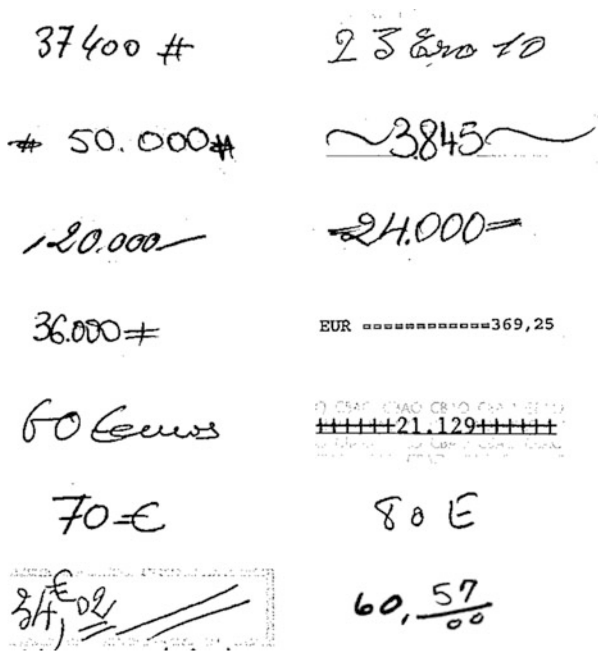


Fig. 21.12 Courtesy amounts with non-numeric signs included

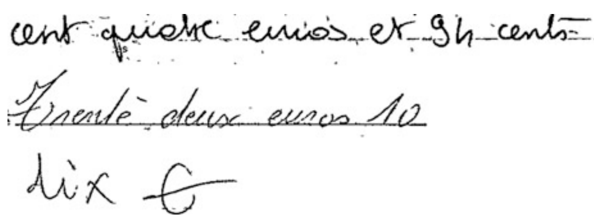


Fig. 21.13 Legal amounts syntaxes

the information acquired in a bottom-up manner from the image of the check. The strategy has also to make the best compromise between performance (recognition rate and error rate) and computing resources. Since courtesy amount is often recognized faster than legal amount (the image of a legal amount contains more signs than the equivalent courtesy amount) and since courtesy amount fields are less prone to noisy background due to typical check form design, common strategies tend to privilege courtesy amount recognition over legal amount recognition.

The property that makes the legal amount directly equivalent to the courtesy amount within a given check also dictates strategies where legal amount recognition is performed in a verification mode over the hypotheses coming from the courtesy amount recognition.

Therefore, a typical strategy is the following:

1. Perform courtesy amount recognition.
2. If the confidence value of the best candidate from courtesy amount recognition is high enough, then this candidate is the final result.

3. Else, if the list of candidates from courtesy amount recognition is not empty,
   (a) Perform legal amount recognition in verification mode
   (b) Integrate courtesy amount and legal amount recognition candidates.
   (c) The best candidate after integration is the final result.
4. Else, if the list of candidates from courtesy amount recognition is empty,
   (a) Perform legal amount recognition in recognition mode.
   (b) Perform courtesy amount recognition in verification mode.
   (c) Integrate courtesy amount and legal amount recognition candidates.
   (d) The best candidate after integration is the final result.

Note that with such a general strategy, the legal amount may not be checked when the process ends in step 2. Although the compliance with local regulations may require to check the value of the legal amounts, practical implementations may take the risk of not doing this checking to privilege the recognition rate, and thus the level of automation, based on the fact that most of the time the legal amount is the same as the courtesy amount. A common refinement consists in trying to recognize the legal amount when the risk is high, that is, when the recognized value for the courtesy amount exceeds some threshold.

Field Location and Extraction

Field location and extraction is the process by which the different fields in a check image are automatically located and isolated from the rest of the check image before being processed during their corresponding recognition steps.

Among the specific recognition techniques specifically used in check recognition applications [18], those in charge of the location and the extraction of fields may take advantage of the regularities of check forms.

A commonly used set of methods for field location and extraction rely on check templates registration [29] where a number of reference images of blank checks are stored as templates and compared to the processed image. Once the processed imaged has been identified to a given template, the location of the different fields is directly derived from the field positions associated to the template.

Matching of reference images to the current image may be based on direct pixel-to-pixel comparison or on general-purpose form recognition techniques which define templates through a set of characteristic zones (called anchors) and which match these anchors to similar region in the processed image. Typical anchors are shapes like lines ends and crosses, logos, characters, and words.

By matching anchors and not just pixels, it is possible to create robustness to small deformations of the image during its acquisition which excludes perfect alignment with the model. In that case, the relative position of anchors in the model and in the image is the base to the computation of deformation parameters like slant, translation, and scaling. Given these deformation parameters, the actual location of fields in the image is adapted from their canonic location in the model.

In practice, there exist hundreds of different check models due to the variety of banks and even to the variants of check forms for one particular bank at one time and also over time. This is why another set of common techniques do not rely on explicit models but try to adapt automatically to any type of check by just considering the constant and common properties of check forms and of fields in these checks.

In that respect, courtesy amounts may be often located and extracted by searching for a currency sign in a restricted part of the image [10, 24]. The courtesy amounts/location is then defined as the region of a given size typically on the right of the currency sign. In a similar way, legal amounts, dates, and account numbers can be located by looking for horizontal guidelines in specific position and with specific size.

For check forms for which there are no guidelines or whose guidelines disappear during image acquisition (i.e., guidelines printed in light colors or made of dots), another set of techniques try to locate directly courtesy amounts by detecting line of characters in typical location for courtesy amounts and with typical relative distance and justification. In that purpose character lines are revealed by applying morphological processing methods like smearing which tend to group characters and words in a line as a rather long and horizontal group of connex pixels.

Once located, the fields have to be extracted from the rest of the image by keeping only their component pixels and getting rid of pixels coming from the check form itself.

Location methods based on templates usually perform field extraction by subtracting the recognized blank template to the image to be processed. Here again, this works only for ideal cases without any deformation.

In more realistic situations, general-purpose line and box extraction are used to remove the shapes coming from the check form in the located zones without any knowledge of the particular template in the image.

After field location and extraction, the fields are prepared for being recognized through their respective recognition technique.

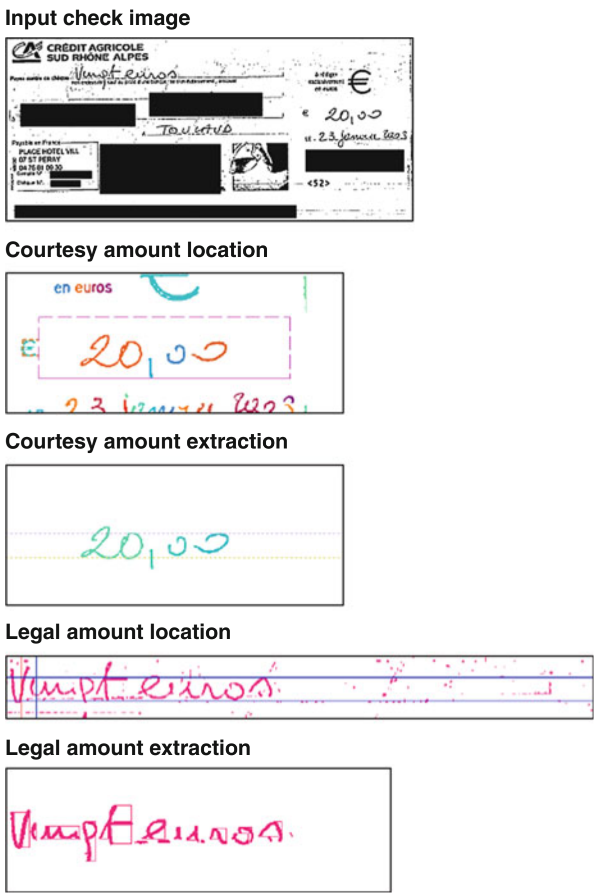
Figure 21.14 demonstrates the extraction of courtesy and legal amounts on a given check.

Courtesy Amount Recognition

Numerical (courtesy) amounts belong to the class of symbol strings respecting a certain number of syntactic constraints as opposed to words or phrases formed over a given vocabulary. Therefore, courtesy amount recognition inherits of the corpus of general methods designed to recognize symbol strings. In the specific case of courtesy amounts, these general methods are specialized by using particular rules for explicit segmentation and for character or symbol recognition and are augmented with techniques for taking into account syntactic constraints, the only approximate 1D character of courtesy amount strings, for finally interpreting a set of symbols as a numeric amount, and for providing a set of scored candidates for integration in the overall decision making strategy.

There exist commonly two classes of general methods for the recognition of symbol strings: those with explicit character segmentation and those with implicit segmentation.

Fig. 21.14 Extraction of courtesy and legal amounts



Explicit character segmentation-based methods implement the segmentation-by-recognition principle in which potential segmentation configurations are validated or invalidated by symbol recognition.

Common methods for detecting potential segmentation configurations rely of heuristic rules which recognize known configurations of touching symbols. For any particular check recognition task addressing a certain currency, a certain country, a certain alphabet or language, and certain courtesy amount-forming habits, heuristic rules are established from the analysis of large sets of representative examples so that there exists no general purpose set of rules efficient across all task configurations.

Symbol recognition also inherits from general-purpose methods for character recognition like statistical or neural network-based classifiers and adapted feature sets. Since courtesy amounts are often composed not only of digits but also of all sorts of non-digit signs (start and end strokes, currency symbol, currency name, comma, etc.), specific recognition techniques are also applied for non-digit signs [16]. They may be implemented by extending the digit recognizers with the considered classes, or they may rely on specific detectors with adapted feature extraction and binary classifier. For example, linear strokes at the start and end of a courtesy amount may be recognized through specific features like the aspect ratio of the symbol or its second-order moments. Commas are also often detected through their absolute size and their relative position to the rest of the field. Specialized recognizers are also often used to take into account difficult segmentation configurations specific to the case of courtesy amounts. For example, touching double zeros which appear often within amounts may be recognized as a single symbol rather than being explicitly segmented. A last class of specialized recognizers may be used to detect others types of touching symbols configurations which should not be classified as a particular meaningful class but should rather be discarded as garbage. Adapted features like the size and the aspect ratio of the number of black-white or white-black transitions are used along with general-purpose classifiers to cope with these specific sub-images.

Basic segmentation-by-recognition methods assume that they process one-dimensional symbol string. In the case of courtesy amounts, they have to be augmented with the capacity of dealing with almost 2D configurations to take into account, for example, currency signs place over a comma or fraction parts in an amount. Common techniques used in that respect extend segmentation-by-recognition to operate over recognition graphs or lattices and not just strings.

In a difficult task like courtesy amount recognition, it is important to take advantage of any a priori over the possible interpretations of a field. Syntactic constraints are such an a priori and are thus commonly embedded into courtesy amount recognition. These constraints are typically modeled as regular expressions or finite-state automata which constrain the paths allowed within the recognition hypothesis lattice. It is also common to incorporate into the syntactic constraints the interpretation rules which are used to transform an amount expression into a numeric value. The most frequent and natural way for implementing interpretation rules consists in building these rules in the syntactic constraints as variable extraction in the case of regular expressions or as two-level finite-state automata.

Implicit segmentation-based methods do not explicitly define segmentation configurations but derive segmentation as a by-product of recognition. Within that class, Hidden Markov Model (HMM) techniques are the most employed for courtesy amounts. Here the modeling of a courtesy amount is no longer based on the manual identification of heuristic rules but rather on the automatic training of the recognition model over a set of representative examples. In order to take into account the peculiarities of courtesy amounts, general-purpose Hidden Markov Model methods have also to be augmented with syntactic constraints, semantic interpretation, and 2D aspects. A common way to embed these issues in an HMM consists in extending and composing the basic finite-state automaton equivalent to the HMM with other automata modeling the syntax, the interpretation, and the 2D lattice of a courtesy amount.

Industrial courtesy amount recognition systems try also to take advantage of the orthogonality of different individual recognition methods to improve their overall performance. Various strategies are employed in that purpose. Within the class of explicit segmentation-by-recognition techniques, several symbol recognizers may be associated. It is also possible to combine an implicit segmentation recognizer with an explicit one. The perturbation principle strategy is also commonly used with perturbations slightly modifying the image of the amount by performing arbitrary skew corrections, for example.

Be it based on explicit or implicit segmentation, courtesy amount recognition has to come out with a set of candidate amounts together with confidence values adapted to the overall decision strategy. These confidence values have to take into account not only the symbol recognition scores but also the syntax and numeric interpretation of the amount and potentially some priors about the amount value. Many type of scoring principles are used in that purpose like neural network integrators or Bayesian combination rules.

The above is valid for courtesy amount recognition applied in pure recognition or in verification mode. In the particular case of verification mode, the syntactic constraints and the scoring priors are modified to only accept the candidate amounts input to the verification.

As an illustration of the above principles, Fig. 21.15 demonstrates the steps in the recognition of a courtesy amount.

Legal Amount Recognition

As a first approximation, legal amounts belong to the class of words or phrases formed over a given vocabulary. Therefore, legal amount recognition inherits of the corpus of general methods used for the recognition of words and phrases.

There exist two main approaches to recognize words and phrases which are both used for legal amount recognition: analytical and holistic recognition. In the case of legal amount recognition, because the vocabulary of words is of limited size, hybrid analytical and holistic methods are also employed.

Analytical recognition explicitly segments words and phrases into units (typically characters, graphemes, words) whose individual recognition is then combined at the word or phrase level. Different general-purpose methods are used in the context of legal amount recognition: Hidden Markov Models with explicit grapheme segmentation using various grapheme feature extraction schemes, hybrid Hidden Markov Model-neural nets techniques which use neural net recognition as a feature extractor, and edit-distance combined with character recognizers.

Holistic recognition methods do not explicitly segment words and phrases but extract global features (e.g., ascenders, descenders, loops, pixel densities within grids) and match these features with words in the lexicon through a number of classification schemes: e.g., k-nearest neighbors, multilayer perceptrons, and ad hoc distances. These methods become applicable in the context of legal amount recognition because the size of the vocabulary is typically less than 50 units. Thus, recognizing a word among 50 using only rough global features is acceptable. Holistic recognition methods present the advantage of being able to take into account contextual effects in the shape of letters within words [4].

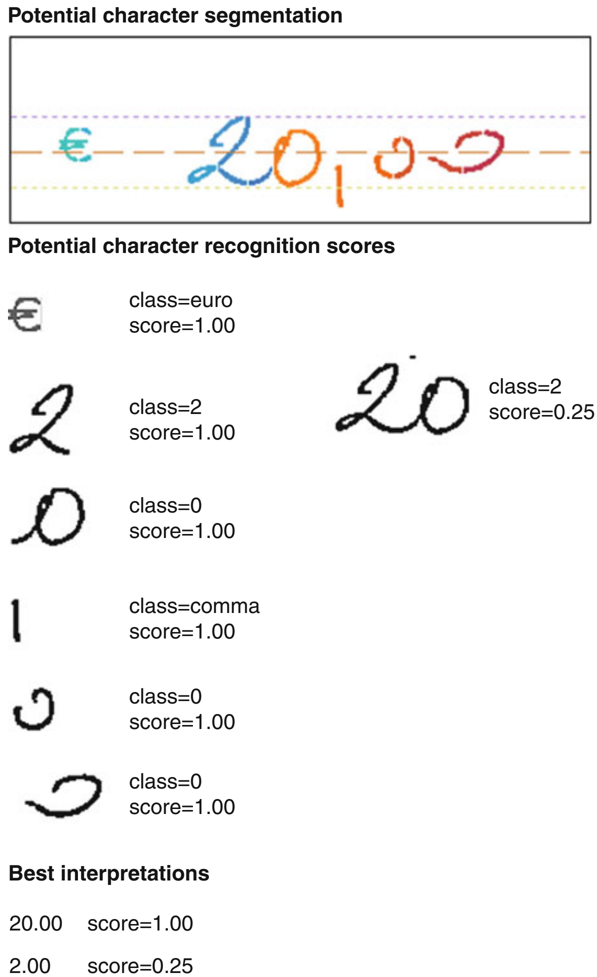


Fig. 21.15 Several steps in the recognition of a courtesy amount

The moderate size of the lexicons also justifies the fact hybrid analytical-holistic methods have been used where analytical features designed at the grapheme or character level are viewed as holistic features.

Word recognition techniques, either analytical or holistic, have to produce a list of candidate words together with their confidence values. These confidence values depend on the particular recognition method used and can be accordingly interpreted as distances, probabilities, and likelihoods. For integration purposes, they are often normalized within a 0–1 scale and all confidence values sum up to 1. As for courtesy amount recognition, perturbation and multi-recognition engine strategies are commonly applied. In that case, confidence values coming from different recognizers are integrated into a single one through general-purpose integrators.

For legal amount recognition, recognition may take place at the word level or at the phrase level. When done at the word level, recognition has to be preceded by a word segmentation step which separates words based on gaps between them.

Since this segmentation may not be 100 % reliable because some writers tend to omit word gaps, words are commonly organized into a lattice of segmentation hypotheses. This lattice is in turn used as the input to a syntactic analysis typically based on a finite-state automaton version of the syntax of legal amounts. This analysis has also to produce the numeric interpretation of the legal amount through standard syntactic analysis techniques.

In many cases, legal amounts include decimal parts expressed in digits rather than in words as they should be. To take into account such cases, a recognizer equivalent to that used for courtesy amounts but with parameters suited to decimal values is applied. Due to its potential cost in terms of computing resources used, it is important to decide when to apply it. A rejection scheme based on a threshold over confidence values at the word recognition level plus additional value like the position in the legal amount image is normally used for that purpose.

The above is valid for legal amount recognition applied in pure recognition or in verification mode. In the particular case of verification mode, the syntactic constraints are modified to only accept the candidate amounts input to the verification.

Figure 21.16 illustrates these principles in the recognition of a legal amount guided by the courtesy amount recognition hypotheses.

Decision Making

Decision making is the step which integrates candidates coming from both the courtesy and legal amount recognition within the overall recognition strategy described above. The result of the integration may be a single amount whose confidence is high enough to be the only interpretation of the image, or a list of candidate amount together with their confidence value which may then be used in a higher-level decision process like the balancing of a set of checks whose sum of amounts is known from outside, or a rejection decision when the list of candidates is empty.

Here again general integration schemes are used to integrate the candidate list coming from the courtesy and legal amounts [19].

State of the Art in Check Processing Applications

The performance of check recognition applications is highly dependent on the country where it is put into practice because checks forms and the way they are filled up differ strongly from one country to another [5, 8, 10, 15, 20, 23, 40, 43]. Moreover, the range of possible amounts depends on the currency and on local habits about what checks are used to pay for. However, practical systems with read rates as high as 75 % at 0.1 % error rates have been reported. In some applicative contexts, the potential amount for a check is known prior to recognition because it has been read on a remittance slip or has been keyed on an automated teller machine. It may also be constrained rather than precisely known in applications where only a limited number of amounts are possible or when a check is remitted within a batch whose total value is known. In those cases, recognition performances at 0 % reject rate are the figures to be considered. Recognition rates higher than 90 % have been reported in such contexts.

Fig. 21.16 Legal amount recognition and decision

Input line image

Potential word segmentations

PW1: 

PW2: 

PW3: 

Word re

PW1:	word=VINGT	score = 1.67e-11
	word=DOUZE	score = 5.93e-16
PW2:	word=EUROS	score = 6.45e-12
	word=DEUX	score = 9.40e-17
PW2:	word=SOIXANTE	score = 8.78e-27
	word=QUARANTE	score = 3.45e-29

Courtesy amount hypotheses

20.00 score=1.00
2.00 score=0.25

Legal amount hypotheses (guided by courtesy amount)

VINGT_EUROS score = 1.08e-22
DEUX_EUROS score = 2.50e-34

Final decision and score

20.00 score=0.999

Reference Datasets

Few image databases have been made public for check recognition. Most of the existing databases have been kept secret by banks and vendors due to private confidentiality issues. Among the few available ones, one can cite the database by the University of Bari, Italy [9], and the CENPARMI Arabic checks dataset [1].

Table 21.3 Indicative list of vendors of check recognition software and systems

Company	Website
A2iA	www.a2ia.com A2iA is a leading developer of natural handwriting recognition, Intelligent Word Recognition (IWR) and Intelligent Character Recognition (ICR), technologies and products for the payment, mail, document, and forms processing markets. A2iA is a pioneer in check processing. Its software technology is deployed through system integrators
	A2iA's technology is available in 6 language versions: English, French, German, Italian, Portuguese, and Spanish and 23 country-specific technology versions
Kappa Image	www.kappaimage.com Kappa Image delivers fraud detection software and services. Kappa, currently protecting over ten million accounts globally, is a software solution providing automated forgery and counterfeit detection from check or giro images
Mitek Systems	www.miteksystems.com Mitek Systems is currently a provider of mobile imaging solutions. Mitek Systems has long been an industry leader and innovator in image analytics and pattern recognition. Mitek Systems delivers a software suite of applications and tool kits that focus on the areas of Payments, Document identification, Signatures, Fraud, and Data Capture. Mitek Systems suite is used in Mitek Systems mobile check deposit product
Orbograph	www.orbograph.com/car-lar-check-recognition-software.htm Orbograph is a provider of recognition-centric services and software for check processing in the financial industry and end-to-end electronic solutions in healthcare revenue cycle management. Orbograph associated its own check recognition technology with technologies from other vendors
Parascript	www.parascript.com Parascript is a provider of software and technologies for document recognition and other intelligent pattern recognition applications like medical imaging. Parascript was a pioneer in check processing software. Its technology was first created in the 1980s at the Soviet Academy of Sciences working on a theory that a computer could actually read handwriting.
	Parascript has extended its check recognition technology to international, non-US markets to read digitized check images specific to a particular country. Parascript offers versions for Argentina, Australia, Brazil, Canada, Chile, France, India, Malaysia, Portugal, Puerto Rico, and the UK
Seres	www.seres.fr Seres is a subsidiary of the French La Poste. Its check recognition technology is used internally in the La Poste Group

Vendors

Check recognition systems are designed and marketed by a number of specialized companies throughout the world among which the most important are given in Table 21.3.

Conclusion

Postal and check recognition systems are two of the most successful and significant pattern recognition applications through their economic worth and technical complexity. Millions of mail items and checks are processed every day in the world using such systems and millions of work hours are saved. These applications embed a large array of pattern recognition technologies encompassing most of the research topics in the domain. Through the hard challenges that postal and check recognition systems design have raised, many technologies have evolved and many new ones have emerged. Postal and check recognition systems illustrate how fundamental recognition methods have to be adapted to tackle real-life problems. They also prove the critical importance of the integration of basic techniques in the success of a real-world application.

The success and popularity of postal and check recognition primarily implemented in the industrialized countries has recently triggered the development of new systems designed for more and more countries. Through this, a large number of different languages have become subjects for pattern recognition research.

Although they have reached a high level of performance, postal and check recognition systems are still subject to improvement and will benefit in the next years from new technologies. Among these progresses, the application of address recognition to parcels – the only class of paper documents at large whose volume is increasing due to the Internet-based commerce – which has already started has the largest potential for performance increases [6, 27]. Another critical issue for the improvement of address recognition systems is the development of more complete reference address databases. Systems that will learn new addresses from the images they process are a research topic.

While the attractiveness of paper-based communication is steadily decreasing, vendors and users of postal and check recognition systems have been looking for ways of improving the value of paper documents by endowing them with new high-tech functions. In that aim, pattern recognition technologies have a prominent role to play and have actually been exploited.

In the mail domain, postal operators are attempting to support their development by marketing new services based on the recognition of mentions other than addresses on mail items. These new functions concern the prevention of fraud with the automatic control of meter marks and of stamps, the tracking-tracing of individual items, the determination of individual addressees for mail forwarding, the recognition of logos for data mining purposes, the delivery of mail under electronic form, the recognition of stickers, the recognition of stamps for advertisement, etc.

The contents of mail item themselves are becoming a new target for pattern recognition applications with the arrival of mailroom solutions [13].

In the check domain, many security issues motivate the recognition of other fields that amounts like payee bank account numbers, dates, and payer address. New regulations like Check 21 in the United States have also raised the need for new functions like the remote deposit of checks from mobile devices like smartphones.

Beyond their inherent benefits, all these new services are likely to reactivate the virtuous cycle in which postal and check recognition applications will raise new challenges from which the research community will provide them with new answers.

Cross-References

►Analysis of the Logical Layout of Documents
►Continuous Handwritten Script Recognition
►Document Creation, Image Acquisition and Document Quality
►Handprinted Character and Word Recognition
►Image Based Retrieval and Keyword Spotting in Documents
►Imaging Techniques in Document Analysis Processes
►Machine-Printed Character Recognition
►Page Segmentation Techniques in Document Analysis
►Page Similarity and Classification
►Recognition of Tables and Forms
►Tools and Metrics for Document Analysis Systems Evaluation

References

1. Al-Ohali Y, Cheriet M, Suen C (2000) Database for recognition of handwritten Arabic cheques. In: Proceedings of the 7th international workshop on frontiers in handwriting recognition, Amsterdam, pp 601–606
2. Bayer T (2005) Method and device for sorting parcels. US Patent 6,888,084
3. Casey RG, Lecolinet E (1996) A survey of methods and strategies in character segmentation. IEEE Trans Pattern Anal Mach Intell 18(7):690–706
4. Castro M et al (2005) A holistic classification system for check amounts based on neural networks with rejection. Pattern Recognit Mach Intell Lect Notes Comput Sci 3776:310–314
5. Cheriet M, Al-Ohali Y, Ayat N, Suen C (2007) Arabic cheque processing system: issues and future trends. In: Chaudhuri BB (ed) Digital document processing advances in pattern recognition. Springer, London, pp 213–234
6. de Rijcke M, Bojovic M, Homan W, Nuijt M (2005) Issues in developing a commercial parcel reading system. In: Proceedings of the eighth international conference on document analysis and recognition, Seoul, pp 1015–1019
7. Desprez O, Caillon C, Miette E (2010) Method of processing postal items including management of digital fingerprints of the postal items. US Patent 7,674,995
8. Dimauro G, Impedovo S, Pirlo G, Salzo A (1997) Automatic bankcheck processing: a new engineered system. Int J Pattern Recognit Artif Intell 11(4):467–504
9. Dimauro G et al (2002) A new database for research on bank-check processing. In: Proceedings of the eighth international workshop on frontiers in handwriting recognition 2002, Niagara-on-the-Lake, pp 524–528
10. Dzuba G, Filatov A, Gershuny D, Kil I, Nikitin V (1997) Check amount recognition based on the cross validation of courtesy and legal amount fields. Int J Pattern Recognit Artif Intell 11(4):639–655
11. El-Yacoubi M, Gilloux M, Bertille J-M (2002) A statistical approach for phrase location and recognition within a text line: an application to street name recognition. IEEE Trans Pattern Anal Mach Intell 24(2):172–188. Table of contents archive

12. Filatov A, Nikitin V, Volgunin A, Zelinsky P (1999) The AddressScript recognition system for handwritten envelopes. In: DAS'98 selected papers from the third IAPR workshop on document analysis systems: theory and practice, Nagano. Springer, pp 157–171
13. Fujisawa H (2008) Forty years of research in character and document recognition-an industrial perspective. Pattern Recognit 41(8):2435–2446
14. Gaceb D (2008) Improvement of postal mail sorting system. Int J Doc Anal Recognit 11(2): 67–80
15. Gorski N (2001) Industrial bank check processing: the A2iA CheckReaderTM. Int J Doc Anal Recognit 3(4):196–206
16. Impedovo D, Greco N, Lucchese MG, Salzo A, Sarcinella L (2003) Bank-check processing system: modification due to the new European currency. In: Proceedings of the 7th international conference on document analysis and recognition – ICDAR'03, Edinburgh, pp 343–347
17. Ishidera E, Nishiwaki D, Yamada K (1997) Unconstrained Japanese address recognition using a combination of spatial information and word knowledge. In: Proceedings of the 4th international conference on document analysis and recognition, Ulm, 18–20, 1997, p 1016
18. Jayadevan R, Kolhe SR, Patil PM (2011) Automatic processing of handwritten bank cheque images: a survey. Int J Doc Anal Recognit (IJDAR) 15(4):1–30
19. Kaufmann G, Bunke H (2000) Automated reading of cheque amounts. Pattern Anal Appl 3:132–141
20. Kim G, Govindaraju V (1997) Bank check recognition using cross validation between legal and courtesy amounts. Int J Pattern Recognit Artif Intell 11(4):657–674
21. Kittler J, Hatef M, Duin RPW, Matas J (1998) On combining classifiers. IEEE Trans Pattern Anal Mach Intell 20(3):226–239
22. Lee CK, Leedham CG (2004) A new hybrid approach to handwritten address verification. Int J Comput Vis 57(2):107–120
23. Leroux M, Lethelier E, Gilloux M, Lemarie B (1997) Automatic reading of handwritten amounts on French checks. Int J Pattern Recognit Artif Intell 11(4):619–638
24. Liu K, Suen CY, Cheriet M, Said JN, Nadal C, Tang YY (1997) Automatic extraction of baselines and data from check images. Int J Pattern Recognit Artif Intell 11(4):675–697
25. Marti UV, Bunke H (2001) Text line segmentation and word recognition in a system for general writer independent handwriting recognition. In: 6th international conference on document analysis and recognition, Seattle, pp 159–163
26. Mercier D, Cron G, Deneux T, Masson M-H (2009) Decision fusion for postal address recognition using belief functions. Expert Syst Appl Int J 36(3):5643–5653
27. Miletzki U (2008) Genesis of postal address reading, current state and future prospects: thirty years of pattern recognition on duty of postal services. In: Proceedings of the 14th ACM SIGKDD international conference on knowledge discovery and data mining, Las Vegas, pp 5–6
28. Niblack W (1986) An introduction to digital image processing. Prentice-Hall, Englewood Cliffs
29. Okada M, Shridhar M (1997) Extraction of user entered components from a personal bank check using morphological subtraction. Int J Pattern Recognit Artif Intell 11(5):699–715
30. Palumbo PW, Srihari SN, Soh J, Sridhar R, Demjanenko V (1992) Postal address block location in real time. Computer 25(7):34–42
31. Pitrelli JF, Subrahmonia J, Perrone MP (2006) Confidence modeling for handwriting recognition: algorithms and applications. Int J Doc Anal Recognit 8(1):35–46
32. Roy K, Vajda S, Belaid A, Pal U, Chaudhuri BB (2005) A system for Indian postal automation. In: Proceedings of the eighth international conference on document analysis and recognition, Seoul, pp 1060–1064
33. Setlur S, Lawson A, Govindaraju V, Srihari SN (2001) Large scale address recognition systems truthing, testing, tools, and other evaluation issues. Int J Doc Anal Recognit 4(3):154–169
34. Seyedin SA, Tabatabaey Mashadi N, Seyedin SMM (2008) Classifying envelopes using machine vision in reading & processing farsi (Persian) handwritten addresses. In: 5th Iranian conference on machine vision and image processing MVIP 2008 in Persian, Tabriz, Iran, pp 1–7

35. Srihari SN (2000) Handwritten address interpretation: a task of many pattern recognition problems. Int J Pattern Recognit Artif Intell 14(5):663–674
36. Srihari SN, Kuebert EJ (1997) Integration of hand-written address interpretation technology into the United States postal service remote computer reader system. In: Proceedings of the 4th international conference on document analysis and recognition, Ulm, pp 892–896
37. Vajda S et al (2009) Automation of Indian postal documents written in Bangla and English. Int J Pattern Recognit Artif Intell 23(8):1599–1632
38. Vinciarelli A (2002) A survey on off-line cursive script recognition. Pattern Recognit 35(7):1433–1446
39. Vinciarelli A, Bengio S, Bunke H (2004) Offline recognition of unconstrained handwritten texts using HMMs and statistical language models. IEEE Trans Pattern Anal Mach Intell 26(6): 709–720
40. Wang D (2009) Study on information fusion based check recognition system, cutting-edge research topics on multiple criteria decision making. Commun Comput Inf Sci 35:384–391
41. Wang C, Hotta Y, Suwa M, Naoi N (2004) Handwritten Chinese address recognition. In: 9th international workshop on frontiers in handwriting recognition, Kokubunji, pp 539–544
42. Xue G, Lianwen J (2009) SwiftPost: a vision-based fast postal envelope identification system. In: Proceedings of the 2009 IEEE international conference on systems, man and cybernetics, San Antonio, pp 3268–3273
43. Yu ML, Kwok PCK, Leung CH, Tse KW (2001) Segmentation and recognition of Chinese bank check amounts. Int J Doc Anal Recognit 3(4):207–217

Further Reading

In this chapter postal mail and check processing applications have been demonstrated as drawing their success and characteristics primarily from a clever integration of a number of different document processing techniques. Therefore, readers interested by more detailed descriptions of the involved techniques should refer to many other chapters in this handbook and in particular those cited in the Cross-References section below.

Postal mail and check processing systems are today the focus of two niche markets where a small set of commercial companies have entered into a strong competition. As a consequence, few technical details will be found in the documentations made available by system providers which of course aim to keep secret the distinctive features of their proprietary technologies. However, consulting recurrently these documentations may at least give some hints as to the progress of systems performances and as to the evolution of their functional coverage. There also exist cases where disclosure is compatible with ownership protection especially in the case of patent publication. Interested readers will find detailed descriptions of innovative methods and functions in published patents (e.g., [2, 7]). However, due to the characteristics of patent regulations in the concerned countries, one is more likely to find in this type of source descriptions of novel functionalities rather than detailed document processing methods.

In some other cases (languages and countries not yet addressed by commercial system providers or research teams using postal and check applications as an exemplification of their specific research interests), detailed system descriptions or pointers to their component techniques will be found in research papers published in major journals and conferences having document processing or neighboring domains as one of their focus (e.g., [5, 8, 20, 32, 37, 40, 42]).