

Issues in Developing a Commercial Parcel Reading System

Michiel de Rijcke, Marija Bojovic, Wouter Homan and Martijn Nuijt
Prime Vision BV
Olof Palmestraat 16-18, 2616 LR Delft, The Netherlands
M.deRijcke@PrimeVision.com

Abstract

In this paper, we present a real time operational parcel reader from an image processing perspective. We will discuss the various difficulties that had to be dealt with to read the addresses on parcels. Main focus will be on the techniques to overcome two specific difficulties, the location of the address block and the occurrence of sender errors. The system presented was developed and implemented last year and since then it is up and running.

1 Introduction

In the area of postal automation, parcel readers are not widely spread. This is due to various reasons:

Return on investment. The financial business case of parcel readers is relatively unattractive for postal companies due to the relative low volumes of most parcel sorting centers. Besides, fully automated parcel reading is relatively expensive requiring multiple (costly) cameras and several image handling systems.

Camera technology. Automating a parcel sorting process requires state-of-the-art camera technology that is capable of high-speed capturing of constant-, high-resolution images of objects with varying heights and wildly colored backgrounds from different perspectives. For recent, this technology was not widely available.

OCR requirements. The requirements for the OCR technology for a parcel sorting system are very demanding. It should be able to handle omni directional writing, multiple font sizes, and heavy image deformations. Moreover, many parcel streams contain a large sub portion of handwriting for which the industrial read rates are still relatively low.

Image sizes. Effective postal automation requires high-resolution images of the postal objects. A typical industrial standard is a resolution of 200 dpi. Such requirements yield huge images for large objects like

parcels (more than 60 Mb). To process these high-speed is a non-trivial task.

Due to improvements of camera technology, increasing processing power and evolving OCR techniques these problems have decreased nowadays. Nevertheless, building a well performing parcel reading system is still a skill requiring challenge.

2 Layout of Danish parcels

The system presented here is designed to read addresses on parcels from Post Denmark. Therefore, it was tuned to deal with the peculiarities of this parcel stream. We highlight some of them here:

Address labels. Most of the addresses are printed or written on an address label. These labels differ depending on their origin: Post Denmark (approximately 30% of all parcels) or a customer of Post Denmark (65% of all parcels). The remaining 5% of all parcels do not have an address label. In general, address labels contain data like the parcel id, addressee, sender, special handling categories and sometimes price. Typical examples of address labels are shown in Figure 2; a customer label left and a Post Denmark label right.



Figure 1. Examples of address labels

Although the layouts and fonts of the labels have to comply a number of restrictions, they appear in many flavors.

Parcel identification by barcode. Each parcel is identified by means of a (code 128) barcode. This barcode can be part of the address label or it can be a

sticker that is put on manually. Its presence is obligatory.

Sender addresses crossed. It is obligatory in Denmark to put a cross through the sender address. This cross usually passes the city-, or street field. Occasionally a cross is missing.

Address Block Locators. A minor percentage of the address labels is equipped with a quartet of edgy address markers (see Figure 2) identifying the addressee, so-called Address Block Locators (shortly ABL).

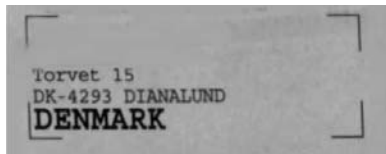


Figure 2. Address with Address Block Locators

3 Overall system

The parcel reading system that is presented in this paper is only a small part of a parcel sorting plant, including sorting machines, volume and weight measuring systems, a computer network and several management systems. In this paper, we will focus only on that part of the system that is related to the capturing and reading of images of parcels.

Starting point is a typical sorting center with several feeders, from which parcels are fed on to a big transporter loop, and with several chutes in which parcels are sorted, each chute belonging to a range of postal codes. The dimensions of parcels vary from 1 to 60 cm in height and from 14 to 120 cm in length. We used 150 dpi digital zoom cameras providing 8-bit grayscale uncompressed images.

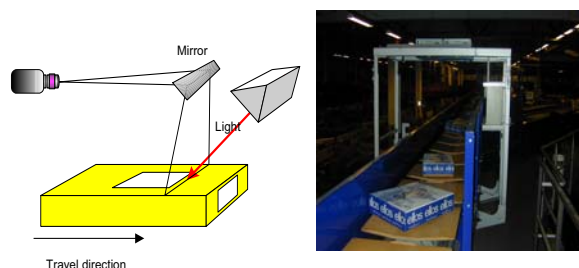


Figure 3. Images of the sorting center

Once a parcel is on the transporter, images are taken from 5 sides of the parcel by 4 different cameras (front & top, back & top, left & right side). Next, the barcode is read (by the camera) and identified. Parcels whose destination is already known (linked to the barcode) are sorted out and OCR engines process the images of the remaining parcels. If the OCR engines do not give a read result for a parcel, its images are

passed to an operator at a video coding system. Parcels that could not be read by this operator are dumped in a special chute for further manual inspection.

4 Addressing OCR difficulties

Designing an OCR system often comes down to solving difficulties due to the peculiarities of the parcel stream. In our case, the main difficulty was to locate the address exactly, despite the standardized use of address labels. This was caused by the following reasons:

Sender addresses. Danish parcels contain multiple addresses: two on the address label (of the sender and addressee) and possibly others elsewhere on the parcel. Usually at most one of the sender addresses is crossed. This introduces a tedious and ill posed problem: how to avoid reading a sender address? In many postal reading applications, the sender problem is tackled by using geometrical features (shape and size) of address blocks, the relative position of the various candidates, the position upon the object and the presence of certain keywords or logos. Unluckily, these features are not suitable for this application.

Label layout variation. Given an address label, it is not straightforward how to determine its layout. There are a very large variety of address labels and layouts, and they do not have a unique property for identifying them. In fact, since each customer may (re-)design its own label not all actually used layouts were known in advance. Also, form identification methods based on keyword spotting (like for instance [1]) did not seem to be the right approach.

Address segmentation. Address labels are generally full of text, lines and logos, frustrating the finding and segmentation of an address. Real difficulties appear when such label-background needs to be removed, like when the address mixes with lines from the bounding box or with preprinted words (see Figure 4).

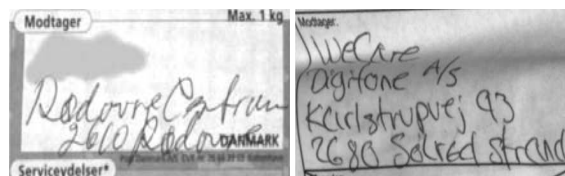


Figure 4. Segmentation challenges

Non-intended label usage. A non-negligible part of all senders print/write the address outside the address label, on a separate sticker or on the carton itself. In that case, the address label is used for cross-referencing or is left empty. Although our techniques are able to detect a considerable portion of these addresses, distinguishing them from sender addresses is even a more difficult task.

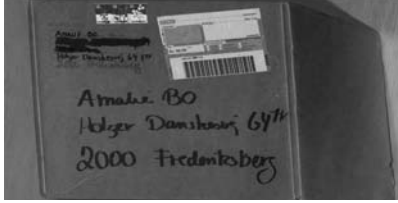


Figure 5. Address not on address label

5 Outline of parcel reading engine

The main task of the parcel reader engine was to determine the parcels destination. Depending on a configurable range of postal codes, this should be the destinations postal code or its PAIN. Here, a PAIN is a 9-digit combination of the (4-digit) postal code and personal identification code of the addressee. In most cases, there is a one-to-one correspondence between PAIN and address, except if multiple PAIN's share the same address, like for instance companies in a big shopping center.

The functional outline of the parcel reader engine was as follows:

- I. Read and locate the barcodes.
- II. Rotate and binarize address snippets.
- III. Find the Address Block.
- IV. Read the Address Block.
- V. Analyze the read result.

The process steps III to V are passed during several iterations until an acceptable read result (PAIN or postal code) is found.

Next, we discuss each step in some detail:

Reading and locating the barcode is essential for the sorting process, since the barcode uniquely identifies a parcel. Furthermore, the barcode contains strong information about the location and orientation of the address block. In fact, the vast majority of these barcodes is situated in parallel or orthogonal to the address block and, moreover, both barcode and address block are close, being part of the same address label. Since the images are huge and rotation and binarization are expensive pixel-wise operations, knowledge of the position and angle of the barcode is used to minimize the region of interest, herewith reducing the total processing time of an image.

For process step II, we used two binarization techniques; one global and one Laplacian based, local binarization. Although this last technique is more time-consuming it made a major contribution to the read performance.

In step III, we applied several different determination techniques to locate the Address Block. We refer to section 6 for more details.

Our proprietary recognition engine, tuned to the syntax and writing characteristics of Danish post, is

applied in step IV, the reading of the address block. We will not discuss this part of the system in detail, but merely consider it to be a black box.

In the last step V of the OCR engine, we tried to get rid of the errors that were made by the previous steps of the engine. The vast majority of errors were due to the reading of the sender block. In section 7, we will discuss how this error source is diminished.

6 Techniques for finding the address

To optimize the read performance, we sequentially applied four different address block location techniques. To save time, methods are only applied if no previous address block location technique led to a reading result. All methods were trained and tested on an image test deck containing approximately 16.000 images of Danish parcels. Each method works on contours of connected components and use the location of the identification barcode as a point of reference. Each method is discussed in some detail.

6.1 Expanding technique

The first and most successful technique we used was to search for address blocks in predefined snippets of the image (hot spots), where the snippets are in fixed position with respect to the parcel identification barcode. For each snippet an address block is searched in the following manner:

- Sort the snippets in the order of the coverage percentage of the contours within the snippet. Reject snippets that are too empty.
- Search within each (remaining) snippet for the maximal concentration of contours. Then, starting from this point iteratively construct a rectangle by enlarging it in both vertical and horizontal direction until the coverage of contours that are added drop below a certain threshold.
- Accept the constructed rectangles as address block candidates only if they fulfill certain geometrical features (left alignment, shape of the hull of the contours, length to height ratio). Reject all others.

Hotspots were constructed manually evaluating the position (relative to the barcode) of the address blocks within the test deck. Both frequently occurring precise as bigger areas were used. The number of hotspots was reduced by afterwards removing hotspots that did not contribute to the read performance significantly.

6.2 Text clustering and classification

The second address block location technique is a contour clustering approach, resembling the method

presented in [2]. Following the order of ‘covered’ snippets as determined by the ‘Expanding method’, we do the following respectively:

- Using a standard text clustering technique determine candidate text clusters within each snippet.
- Classify each of these clusters as being an address block, by means of a trained process (for instance by a Neural Net) with features like the shape, position and weight of the clusters.
- Choose the best (e.g. highest confidence) cluster that is classified as an Address Block for further processing.

The classifier from the second stage of this method was trained on our image test deck.

6.3 Direct Postal Code detection

The third method tries to determine the address block location by a direct search for the postal code. This ‘Postal Code-detection’ is a combination of a free-orientation contour filter, searching for machine printed digits, and a clustering step searching for a fixed pattern of adjacent digits. Here, the counter filter uses parameters (boundaries) that are trained on the test deck. Once the postal code is found the address block is easily determined. For this particular application, detection technique was capable to locate approximately 75% of all machine printed postal codes. Drawback of this method is that it locates also the sender blocks effectively.

6.4 Search for Address Block locators

Last technique to determine the address block location was a contour filtering technique to search for the ABL edges. Within a restricted area with respect to the location of the barcode, it searches for a regular pattern of four of these edges. The parameters used were chosen heuristically because the limited number of examples within the test deck. The detection rate was approximately 70% with a false alarm rate of 1%. The method turned out to be vulnerable to image deformations and noise. Since the low occurrence of ABL’s, this technique was only applied to a special range of address labels.

7 Dealing with sender addresses

In case of Danish parcels the distinction between sender and addressee turned out to be an elementary problem. We therefore applied various techniques in trying to avoid them.

7.1 Cross detection

Once an address block is read, the presence of a cross is evaluated. This is done in the following way:

- Calculate the x- and y-derivative of the image using the Sobel Kernel [3].
- From every point with non-negligible gradient construct a line orthogonal to the direction of the gradient. Extend it in both directions if the gradients at the extremes of the line are equally orthogonal to the direction of the line. Connect lines that are close and in the same direction. If a line is sufficiently large add it to the list of lines.
- If two lines cross close to a postal code and their angle is not close to 90 degrees (address boxes usually are not square), then a cross is found and the address block result is considered to be a sender block, so rejected.

7.2 Sender word spotting

A second technique to reduce the number of sender blocks is to search for forbidden words (like ‘From’ and ‘Sender’) in a narrow region above the address block. To save processing time, this technique is applied only when an address was read successfully.

7.3 Distrust single address lines

Sender addresses that are on customer-made labels or preprinted on the parcel often exist of a single line or a partial address (postal code and city). Furthermore, sender crosses often leave the postal code in tact but downgrade the readability of the street. Therefore, we decided to reject all addresses from which only the bottom line (Postal code and city) is read reliably, giving a considerable reduction of the sender error. However, since it also downgrades our reading capabilities, we combined this rejection mechanism with an addressee word detection technique. In fact, in case of a one line reading results, we search for certain addressee words (like ‘To’ or in our case ‘Modtager’ or ‘Til’). If such word is found, the result is accepted otherwise it is rejected. The application of latter technique resulted in an absolute read rate improvement of 1%.

7.4 Position-based rejection

Since the ‘Postal Code detection’ technique disregards all contextual information in an image, it can be a source of sender errors. To avoid these errors, we rejected postal codes with a too large distance of (the center of) the postal code to (the center of) the barcode or for which the line Postal Code to barcode

makes a too small angle with the barcode (see Figure 6). Doing so effectively reduced the error caused by the ‘Postal Code detection’ technique with 86%.

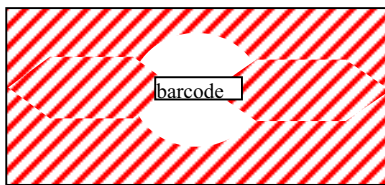


Figure 6. Forbidden area

7.5 Performance rejection techniques

We grouped these four sender block rejection techniques in two categories:

- *Sender Detection*: applying cross detection and sender word spotting.
- *Geometry*: distrust single address lines and apply position-based rejection.

‘Sender Detection’ methods are more time-expensive trying to determine if something is a sender block; the Geometry methods are faster, but use global features of error causing sender blocks. We measured the effectiveness of these methods on a set of 5850 parcels. Note that combining the rejection techniques has a certain synergy; some sender errors can only be removed applying both.

Table 1. Sender error countermeasures

	# Read	# Errors	%
Both countermeasures	3216	58	1,8
Only geometry	3272	104	3,2
Only sender detection	3486	115	3,3
No countermeasures	3533	142	4,0

The countermeasures still leave room for improvement, sometimes sender errors pass these rejection mechanisms because of the administration of the exact position of addresses, keywords and crosses is sloppy or disturbed by noise.

The application of rejection techniques causes also a loss of yield in case the rejection mechanism abusively rejects the address block:

Table 2. Inappropriate rejections

	Error %
False Sender Detection alarms	0,6
False Geometry alarms	6,7

Table 2 shows that by reducing the number of false ‘Geometry’ alarms will yield in a considerable read rate improvement.

8 System performance

In practice, the parcel reading system we have developed reads approximately 56% of the parcels fully automatically (partly PAIN, partly postal code). Focusing on the postal code only read rates above 60% are measured.

The throughput of the system is 4s per image on a single, 3 Ghz Pentium IV. Due to the architecture of the sorting process this is a reasonable performance. Much processing power is consumed by the various intelligent functionalities, as introduced in this paper, and the handling of the very large images.

Most OCR rejects are related to deformation problems (hard to solve) or to poor handwriting. In the latter case, the address block location is often the main problem; since the address box of many labels is rather tight separating it from the label is non-trivial.

9 Conclusions

In this paper, we presented a design of a system for reading addresses on parcels in Denmark. We illustrated the main document processing difficulties that needed to be addressed and tackled. During the development of the system various techniques were applied to locate, read and remove sender blocks. The system is operational in practice with great customer satisfaction, processing approximately 25 million parcels per year. Read rate improvements and error reductions are possible and expected in next system releases.

10 References

- [1] H. Sako, M. Seki, N. Furukawa, H. Ikeda and A. Imaizumi, “Form Reading based on Form-type Identification and Form-Data Recognition”, Proceedings of ICDAR 2003, Edinburgh, 2003, pp.926-930.
- [2] V. Govindaraju and S. Tulyakov, “Postal address block location by contour clustering”, Proceedings of ICDAR 2003, Edinburgh, 2003, pp.429-432.
- [3] Gonzalez, R., R. Woods, *Digital Image Processing*, Addison Wesley, 1992.