HAL
archives-ouvertes.fr

Kalman Filter Contributions Towards Document Segmentation

Charles Queguiner, Jean Camillerapp, Ivan Leplumey

▶ To cite this version:

HAL Id: hal-00763395
https://hal.inria.fr/hal-00763395

Submitted on 10 Dec 2012

Kalman Filter Contributions Towards Document Segmentation

Ivan Leplumey (+), Jean Camillerapp (+), Charles Queguiner (*)

(+) INSA-IRISA 20, av. des Buttes de Coesmes - 35043 RENNES CEDEX - FRANCE
(*) IFSIC-IRISA Campus de Beaulieu - 35042 RENNES CEDEX - FRANCE
{Ivan.Leplumey, Jean.Camillerapp, Charles.Queguiner}@irisa.fr

Abstract

One way of tackling document analysis is to split the document into different layers and then to find a segmentation for each one. The straight line is often one of the basic elements of the layers making up a document. The detection of these segments is often disturbed by the superposition of different layers in the same area. Therefore we developed a segment extractor, which processes masked areas, discontinuities and even curvatures. After showing in our previous work, the interest of such a system in binary image analysis, our aim was to show the method could be generalized to grey level images, by taking into account pixel luminosity as one of the parameters. The Kalman filter method combined with an adequate control scheme, allows the pooling of local observations, when their position, width and luminosity are coherent. The proposed model gives an interresting scheme to manage drawings masking each other (eg. cross-lines).

1 Introduction

Document segmentation aims at making easier the document interpretation. Locating different kind of information in a document global contents is a part of document segmentation. Some documents, such as maps, musical scores, written checks can be seen as a superposition of multiple information layers [1]. For example, we can easily discern two layers in the image of a check : the background layer (printed paper, base lines, preprinted information) and the superimposed handwritten characters. Most people are working on binary images but, as we think the luminosity given by grey levels could improve the segmentation process, we propose to work on grey level images.

To achieve the analysis of a document, we try to separate the different components to recognize them individually. Multi layer superimpositions bring about difficulties. The "linear drawings" are often one of the background elements. They often interfere with handwriting on preprinted documents (checks, forms, ...). In our study, we tried to discover those linear drawings (also called "segments").

• "Line drawing" (also called "segment") definition

A linear drawing is defined as an alignment of pixels. Each of them being chosen as a representative point of a real observation extracted from the image (see section 3). It is necessary to define some requested properties.

• (1) meaningfulness of the neighborhood ; an alignment of isolated points is not considered as a linear drawing.

• (2) possible existence of discontinuities : it is useful to allow locally an absence of points, due to the quality or the nature of the extracted object (dotted linear drawings).
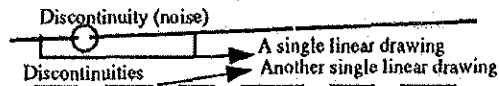

• (3) Use of the different parameters related to each representative point, such as thickness and brightness of the corresponding observation (for grey level images).
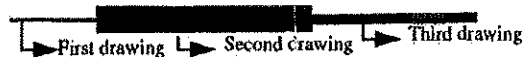
• (4) extraction of segments of different sizes, (ranging from a few points to several hundred representative points).

• (5) possibility of a drawings superimposition, e.g. :
 - staves in a musical score covered by quavers bars,
 - base lines on a written check.

This superimposition characteristic results from the document layers theory.
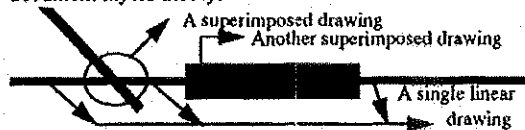
• (6) Lastly, possibility to include a possible "curvature" as a "straight segment" although it may be seen contradictory to the "linear drawing" notion. Locally, we can observe a straight line but, globally, this "straight" segment is slightly curved, and we can decide to keep this or not.

·Linear drawing extraction methods

Two methods are commonly used to solve the straight segment detection problem.

- The first : Hough transform [2] does not respect (6).
- The second : includes a thinning stage followed by an iterative polygonal matching, but does not respect (2) (5).
- That's why we propose a method which respects all the six previously described properties. It locally works on linear drawing hypotheses which are confirmed or not during the analysis of the image. The used measures do not only concern the spatial location of the drawing, but also other characteristics (thickness and grey level).

Our method included two levels. The first level contained a prediction/verification tool : the Kalman filter [3]. The second level was a control layer permitting to interpret the expected cases (discontinuities, crossings). In a similar way, Jean-Pierre Veran [6], also used the Kalman filter, to follow roads in aerial grey level images.
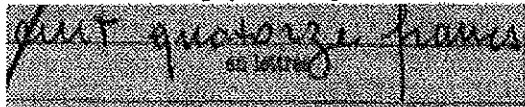

Fig. 1 : Example of a grey level image

2 Kalman filtering

What we were trying to detect was a series of nearly horizontal aligned observations (base lines on checks).

2.1 General principle of prediction/verification

The basic hypothesis is that any object can be characterized by a state c_k varying in relation to a scale, usually time. The method allows to follow the evolution of this state to estime it and also to predict it. For that purpose the method used a recursive linear estimator which according to the known past of a state e_k, can predict its evolution.

e_k = linear estimation ($e_{k-1}, e_{k-2}, e_{k-3}, \dots, e_0$).

After predicting this stage, the next step is to verify it by examing the actual current state of the object and correcting the estimator by taking into account the error made between the prediction and the measured observation.

Accepting errors and correcting the estimator by taking those errors into account gives this principle all his power to manage the state of an object. However, it implies knowing how the state of the object can vary with time, i.e. having a model of this object.

2.2 Kalman filter equations

The Kalman filter formalizes accurately the prediction/verification principle. This is a recursive linear estimator which offers the advantage of taking into account calculated errors between measurements and estimations.

Those errors are automatically incorporated into the model to improve future estimations.

This method uses two models [5] :
- a state model S, varying with time and

- a model of measure X, deducted from S.

Both the models S and X are represented by a vector. S cannot be directly known and we are trying to estimate it. Evolution state equation :

$$S(k) = A(k-1).S(k-1) + W(k-1) \qquad (e1)$$

where W stands for the model error, characterized by a zero mean and a variance Vw. A is a matrix characterizing the evolution of the vector S.
Measure equation :

$$X(k) = C(k).S(k) + N(k) \qquad (e2)$$

where C is a matrix used to deduct the measure vector X from the state vector S, and where N stands for the measure noise with a zero mean and a variance Vn.

Besides, it is important to note that W and N are supposed to be uncorrelated.

To find the Kalman filter equations, see [5].

3 Observation extracted from an image

To apply this theory, it is necessary to define what kind of object we are studying. We have to define its state, and its evolution model in the course of time (here columns of the image). We are searching for "observation" alignments in the columns. These pooled observations will define the linear drawings that are the objects we are looking for in the document image. The state of an object will be determined from parameters, chosen for the object model definition.

First of all, we had to define the observations looked for in the image. Vincent Poulain d'Andecy in [4] defined observations in binary images, we propose to extend this definition to grey level images.

It must be remembered that we were trying to extract horizontal base lines in grey level images of checks (Fig. 1). Those lines can be considered as dark thin objects laid on a clear background. This characteristic determined the observation model we were going to search for in the image. A cross-section of a thin horizontal object has a typical luminosity profile pattern, that we tried to find in the image matrix. We sought in each column, a sequence of points of decreasing luminosity, followed by another point sequence of increasing luminosity. Several parameters, which corresponded to the luminosity and position in the column of these remarkable points of the observation, were associated to each detection and recorded as follows :

- begin point of decreasing luminosity (bln, blumi)
- medium point (mln, mlumi) : last point of decreasing luminosity and first point of the increasing luminosity
- last point of the increasing luminosity (lln, llumi).

As the two point positions, bln and lln, were too noise sensitive, we added two new points (mi1ln and mi2ln) which had a better position stability. Those points are located respectively half-way down and half-way up the luminosity curves (see Fig. 2). They corresponded fairly well to points selected by edge detectors.
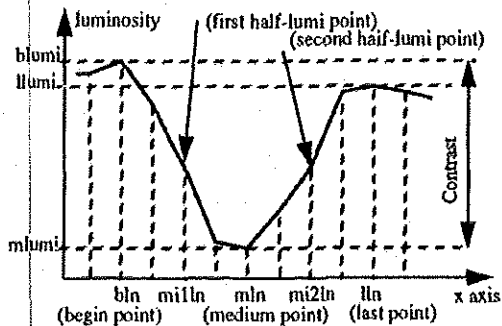
Fig. 2 : An observation (luminosity curve profile) and the eight associated parameters

As a result, an observation was defined by eight parameters. Dependiny on the rest of the application, they are not all necessary. The technique we used only divided one column of an image into a succession of intervals that could be parts of a linear drawing. We noticed that a great number of those observations corresponded to noise in the image, that's why a selection process was necessary. To achieve that we associated to each observation, we calculated a contrast (max (blumi, llumi) - mlumi). Every observation with a contrast below a threshold was discarded.

The observations corresponding to noise that were not eliminated by the previous thresholding were noncritical for the method. The obtained results are robust to this threshold usage. The final decision to consider an observation as interesting or not was based on a broader context of histeresis thresholding. Figure 3 shows the observations (Fig 3 b, in black) obtained from the original image (Fig 3 a) with a contrast threshold of 30 grey levels.
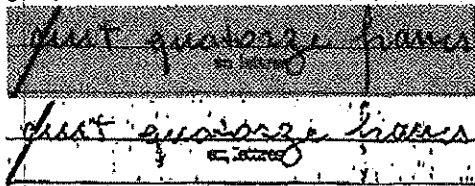

Fig. 3 : a- Original image, b- extracted observations.

The "linear drawing" object we were looking for, was finally characterized by four parameters :
• position (in the column)
• slope (computed along the drawing)
• thickness (calculated with the observations)
• luminosity (given by the observations).

A first step consits in extracting observations for a column k of the image. A second step consists in a prediction from existing objects hypotheses. Finally, a difficult problem remains to be solved : the choice of the actual observations which will be used for verification purposes. One of the Kalman filter's assets is to automatically carry out parameters to make this choice.

4 Application to segment detection : state and measure equations

As we are looking for straight lines in an image, following section 3 conclusions, we propose a state vector (quadruplet) S = (position, slope, thickness, luminosity).
The state equation showing the object evolution then is :

$$
\begin{bmatrix} S_{po(k)} \\ S_{sl(k)} \\ S_{th(k)} \\ S_{lu(k)} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} S_{po(k-1)} \\ S_{sl(k-1)} \\ S_{th(k-1)} \\ S_{lu(k-1)} \end{bmatrix} + \begin{bmatrix} W_{po} \\ W_{sl} \\ W_{th} \\ W_{lu} \end{bmatrix} \quad (e1)
$$

$$S(k) = A(k-1) . S(k-1) + W(k-1)$$

The measurement equation matches the studied observation and the state model of the segment at instant k. Hence the following equation :

$$
\begin{bmatrix} X_{po(k)} \\ X_{th(k)} \\ X_{lu(k)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} S_{po(k)} \\ S_{sl(k)} \\ S_{th(k)} \\ S_{lu(k)} \end{bmatrix} + \begin{bmatrix} N_{po} \\ N_{th} \\ N_{lu} \end{bmatrix} \quad (e2)
$$

$$X(k) = C(k) . S(k) + N(k)$$

The vectors W and N are only characterized by a priori fixed variances.

Final equations can easily be obtained by using matrices A and C in the general equations of the Kalman filter [5]. We shall not develop the choice of error parameters which are linked to both what we are looking for and what we are measuring. Thus, if we decide to refuse a "curved line segment" (property (6)), we choose a position model with errors (W_{position}, W_{pente}) valued close to zero.
The tool we are presenting, allows us, if a linear drawing has been expected up to column (k-1) to :
• predict the characteristics of the observation expected for column k ;
• evaluate associated error predictions, depending on model errors and observation errors (image noise) ;
• take into account this observation (matched with a segment) for the calculations concerning next columns.

5 "Segment hypothesis" control scheme

The segment hypothesis control scheme consists of a loop, marked by the column number in the image. This loop contains the four following actions :
• for each hypothesis, a prediction stage , updating each Kalman filter (one per segment hypothesis) ;
• a matching stage, also called verification stage, matching hypothesis and observations ;
• a new hypothesis creating stage (when an observation didn't match an existing segment hypothesis) ;
• a hypothesis stopping stage (when no observation matched one hypothesis) ; to keep it for future interpretation or to erase it when not sufficiently confirmed.

3

5.1 Matching observations/ segment hypotheses
•Principle

The most difficult point of the control scheme was matching working hypotheses with current column observations. For each hypothesis we had to find, the corresponding observation, if it exists, among all those extracted from the image column. A way to do this was to calculate for each observation, its distance from all hypotheses and, to choose those within a good distance range.

Knowing $\sigma_{Thickness}$, $\sigma_{Luminosity}$ and $\sigma_{Position}$, given by Kalman filter's covariance matrix [5], allowed us to use the Mahalanobis distance using the three parameters : thickness, luminosity, position. The χ^2 test, used with three degrees of freedom, allowed us to define a distance value for the matching. However, we had to cope with three cases :

  • one observation matched one hypothesis ;
  • many observations matched one hypothesis ;
  • no observation matched a hypothesis.
•Multiple matching

In the case of multiple observations matching one hypothesis, we choose to match it with the nearest observation in terms of Mahalanobis distance. This is also known as NNSF filter (Nearest-Neighbor Standard Filter).
•Absence of matching

When no observation existed within the given distance range : was it a total absence of observations in the predicted area, or was it only because we were getting through a superimposed object ?

In the case of actual absence of observations, our choice was to add the predicted observation to the Kalman filter, delaying the decision for a certain time.

The second case, getting through a superimposed object, had to be better defined. The semantics of the image helped us to explain this. We were working on check images where the handwriting is darker than the preprinted base lines are. When those two kinds of information superimpose, only the handwriting is apparent. The luminosity parameter of the observation extracted from the image could be very different from the predicted one, so there was no observation matching, despite the fact that the position parameter was "good". That's why we considered that this real observation was crossed by "another" hypothesis :

  • if the real observation corresponded to the predicted ;
  • if the thickness was larger ;
  • if the luminosity was darker.

This crossing idea is fundamental since it allows us to "wait" before discarding or terminating a segment hypothesis. As a matter of fact, when a clear linear drawing was covered by handwriting, we waited until we found an observation which matched or did not match before deciding the hypothesis future. Moreover, this idea can help us to improve image segmentation by providing us some explanations about specific luminosity profile (when two straight lines cross each other). We could then predict the luminosity profiles of several hypotheses that cross each other and then deduct associated observations.

5.2 Creation of a segment hypothesis

The conditions to create a new segment hypothesis are simple : every observation being unmatched with an existing hypothesis generates a new hypothesis. Knowing that a Kalman filter is associated with each hypothesis, it is necessary to use a stopping strategy for those new hypotheses otherwise the required computing time will rise.

5.3 End of a segment hypothesis
•Principle

To prevent an increase in computing time, it is necessary to keep only promising hypotheses, those that still matched observations, and to get rid of the others.
•Lack of observation matching for a certain time

Hypothesis that do not match observations for a certain time will be discarded. The future of a hypothesis depends on how long it remains active. If its life was long enough, the hypothesis is stored in a data structure for later interpretation, if not it is "dropped" except of course if the hypothesis is in a crossing state.
•Hypothesis initiation

To better control the "living" hypotheses, we used an induction threshold (minimum two first real observations).
• Hypothesis synonymy

When two different hypothesis matched the same observation, the program affected this observation to those two different hypotheses. If this was the case for a certain time, we used the Mahalonobis distance to check they were not similar. If the distance is small enough we considered those two hypotheses as synonyms. The crossing case, involving two straight lines having different directions, was sorted out by comparing their respective slope. The observation could be linked to both hypotheses, but the slope would keep its Mahalonobis distance above a threshold. This allowed those two different hypotheses to survive.

6 Results

The obtained results show the method's interest. Figure 4 shows extracted linear drawings (longer than 10 pixels) with the described Kalman filter based on position, thickness and luminosity coherence. It is obvious that the two base lines of the check are detected, though they are superimposed with checkbars and handwriting.

Fig. 4 : The detected linear drawings.
The detected observations, crossed by several hypotheses

are shown in white (superimposed with the original image) (Fig. 5). In those regions segmentation will be improved

Fig. 5 : Observations (white) crossed by many hypotheses

The following figures show the obtained layers partition (notes, quaver bars, ...) from a musical score binary image : The stave lines have been removed (original Fig. 7a) :
* by the common technique in two stages.
    - firstly locating the different lines.
    - secondly removing the observations with the right location and the right thickness, independently of other context (defaults : see circles on Fig. 6a).
* by our cooperative technique between low and high levels. The results show a good shape along the edges.
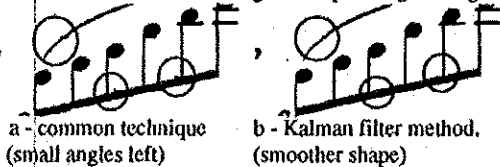
a - common technique   b - Kalman filter method.
(small angles left)   (smoother shape)

Fig. 6 : The image without the stave lines.

On the following image, the quaver bar, part of a layer, has been removed, without distortion for the stave lines, thanks to the proposed Kalman filter method (Fig. 7b).
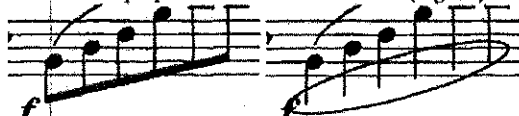
a- original image   b- without the quaver bar.

Fig. 7 : The image without the quaver bar.

The last figure shows our grey level check image, after removing the base lines. We can notice that the oblique line is not erased. The crossing zones are conserved with some aspect defects as square edges instead of round edges (see circles in Fig. 8).
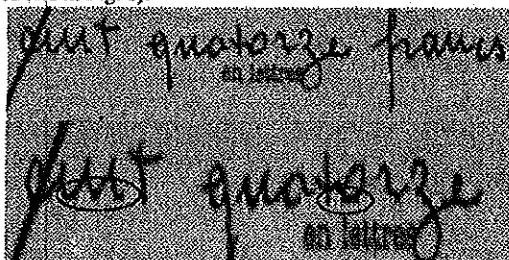
a- Same size as Fig. 1   b- Zoom

Fig. 8 : The check after elimination of the two base lines

Computing time for this image (730 columns x 100 lines) is about half a second. The average number of linear drawings hypotheses on this image is four per column.

Conclusion

The presented method whose purpose is to extract linear drawings has some characteristics that can be considered as drawbacks :
* the results are sensitive to the scanning orientation ; They can present some small "pixel level" differences.
* the detection depends on the orientation of the linear drawings (best with slope between +45° and -45°) : Nearly vertical linear drawings cannot be detected by a column oriented scan. A line oriented scan solve this.
* the method uses a great number of parameters ; Nevertheless, it is very easy to use and also very robust. As a matter of fact, the error parameters for the model and for the observations (position, thickness, lumi) are managed by the model controlling the Kalman filter.

We consider the following points as advantages :
* The method can easily accept new parameters (see section 4).

The luminosity and thickness were added to our initial Kalman filter model (for binary images). Thanks to good parameter usage, this method is very flexible. It can even accept some slight curves in detected linear drawings, by relaxing the model error parameter.
* the method manages to widen its decision context. By integrating compatible observations provided during the scan (see section 5), there is a cooperation between local analysis (real observations) and global analysis (segment hypothesis). This double reading level makes the forceful point of the method.
* this method can be used for grey level images as well as for binary images.
* moreover, when working with grey level images, our method (see section 3) does not imply a binarisation of this image to extract or remove some layers.

Bibliography

[1] Baird, Bunke, Yamamoto (Eds), *Structured Document Analysis*. Springer-Verlag, (1991).

[2] D. H. Ballard, C. M. Brown,. *Computer Vision*. Prentice-Hall, (1982).

[3] T. E. Fortmann, Y. Bar-Shalom, *Tracking and data association*.. Mathematics in science and engineering, Vol. 179, Academic Press, (1988).

[4] V. Poulain d'Andecy, J. Camillerapp, I. Leplumey, *Kalman filtering for segment detection : application to music scores analysis*.. 12th ICPR, Computer Vision and Applications, Jerusalem, (9-13 october 1994) Vol. 1, pp. 301-305.

[5] H. W. Sorenson., *Kalman filtering : Theory and Application*.. IEEE Press, (1985).

[6] J-P. Veran, *Suivi de routes dans une image aerienne par filtrage de Kalman*.. Rapport technique de l'Ecole Nationale Supericure des Telecommunications, Departement Images, Paris, (march 1993).