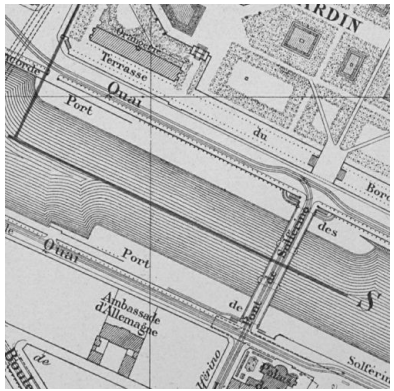
Linear Object Detection in Document Images using **Multiple Object Tracking**

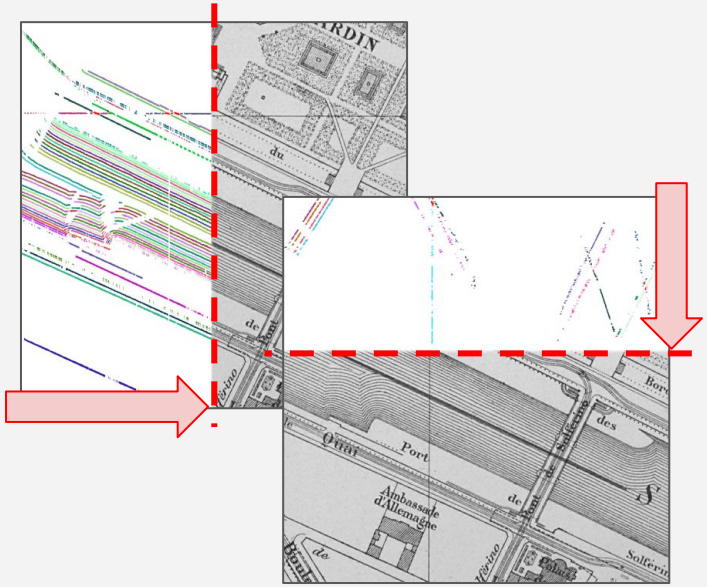
P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*

ICDAR'23
Poster Session 2
Wednesday, Aug. 23

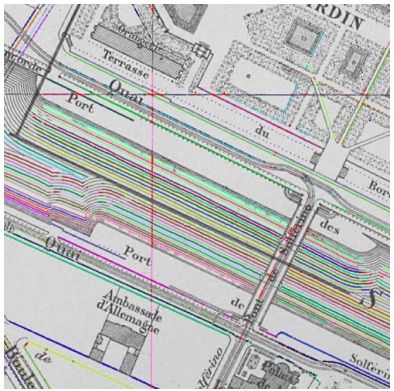
Input:
document image



Processing:
1 horiz. scan + 1 vert. scan



Output:
pixel-level instance segmentation of complex linear objects

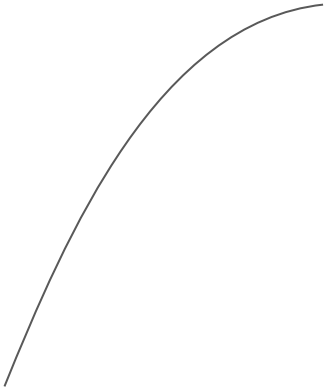


Public code & data

https://github.com/**EPITAResearchLab**/**bernet.23.icdar**

P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*
ICDAR'23
Poster Session 2
Wednesday, Aug. 23

Existing approaches fail with linear objects which are



curved

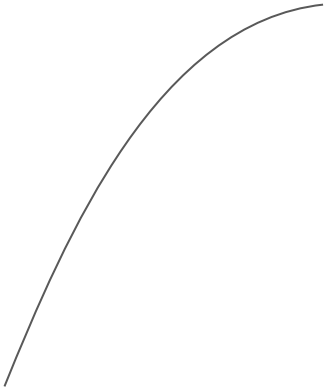
Public code & data

https://github.com/**EPITAResearchLab**/**bernet.23.icdar**

Linear Object Detection in Document Images using **Multiple Object Tracking**

P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*

Existing approaches fail with linear objects which are



curved

erased

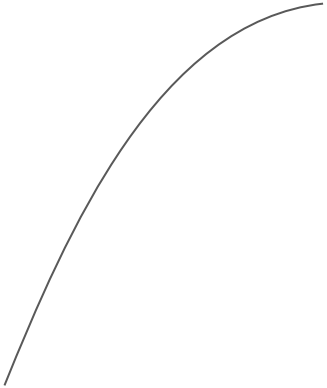
Public code & data

https://github.com/**EPITAResearchLab**/**bernet.23.icdar**

Linear Object Detection in Document Images using **Multiple Object Tracking**

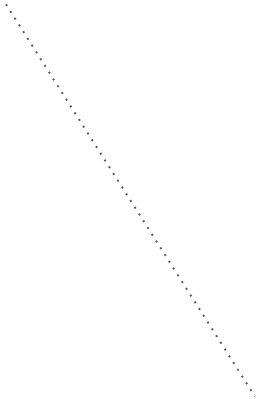
P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*

Existing approaches fail with linear objects which are



curved

erased



dashed

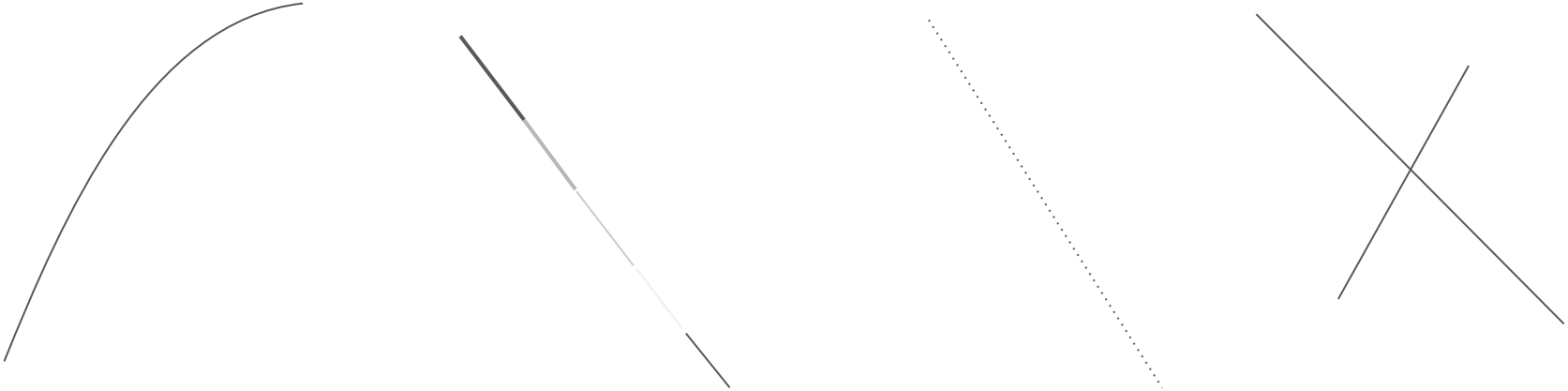
Public code & data

https://github.com/**EPITAResearchLab**/**bernet.23.icdar**

Linear Object Detection in Document Images using **Multiple Object Tracking**

P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*

Existing approaches fail with linear objects which are



curved **erased** **dashed** **intersecting**

Public code & data

https://github.com/**EPITAResearchLab**/**bernet.23.icdar**

Linear Object Detection in Document Images using **Multiple Object Tracking**

P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*

ICDAR'23
Poster Session 2
Wednesday, Aug. 23

Except one forgotten approach…

ICPR'94 ↓

Kalman Filtering For Segment Detection : Application To Music Scores Analysis

V. Poulain d'Andecy, J. Camillerapp, I. Leplumey

IRISA/INSA, 20 Avenue des Buttes de Coësmes- 35043 Rennes Cedex - France
Tel : 99-28-64-00,e-mail : Jean.Camillerapp@irisa.fr

Abstract

Many symbols in music scores are linear segments. In this context, we designed an extractor of segments. It is robust towards problems of quality within binary images (scale factor, curvature, bias and noises). It is based on Kalman filtering technic. By splitting music scores into layers of detectable symbols and by applying methodically to the defined layers both this extractor and simple rules of classification for the detected segments, we were able to recognize staves, stems, slurs, beams, bar lines, black note heads and then quarters and note groups.

1: Introduction

Some documents such as technical documents (industrial drawing, maps, ...) are seen as a superposition of graphics and text layers [1]. A music score is a technical document in which many authors [2] have discerned two layers : a graphic one carrying the staves and a text one containing all the other symbols. This approch is logical because of the structure of these documents which are a set of staves on top of which symbols are written. So, music scores analysis starts by the processing of these two layers.

To achieve that, lots of technical processes have been used : projections [5][4], Hough transforms [8], white area tagging [6], identification of staves by Line Adjacency Graphs parsing [3], pattern matching [2], etc... Partial results have been obtained by combining these methods.

Furthermore, the suggested methods are rarely robust towards the defects within binary images, i.e. biais, curvatures and noises (black or white interference pixels). These defects can be caused by acquisition, but they are often due to the low quality of the processed document. Therefore it seems important to devise a method which takes care of these problems.

We look for a musical score analysis method whose purpose is both the homogenization of the analysis technic and the enhancing of their robustness. In order to obtain this we have noted like Kato [5] that musical scores often have linear structures. For example, staff lines and stems are segments. Considering the detection of this kind of symbols we are able to describe a musical score such as four detectable layers :

- "long" horizontal segments (staves, beams and slurs)
- "long" thin vertical segments (stems and bars)
- black note heads (short and thick segments)
- all the other symbols.

We propose to use a Kalman filter as a segment detector [7]. It is able to separate these layers. After introducing quickly the detector, we will deal with the segmentation and the classification of the objects of the first three layers that we have defined above.

2: Segments Detection

An ideal segment could be defined like a succession of connected run-lengths (a run-length is a set of connected pixels of same color within a column or a line) which have the same color, the same thickness and from which the middle points of the run-lengths are on a line segment (figure 1).

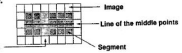


Figure 1 : example of segment

The method we introduce will be robust because it allows detection of segments bent by :

- bias, the line of the middle points can have any slope;
- weak curves, this line can oscillate;
- variable thickness of the run-lengths which are included in a segment. This thickness can be equal to zero to tolerate some cuts.
- and detected segments will have any thickness to insure independence towards scale factors.

ICDAR'95 ↓

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 at 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 interesting scheme to manage drawings masking each other (eg. cross-lines).

1 Introduction

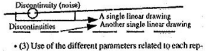
Document segmentation aims at making easier the document interpretation. Locating different kind of information in a document global contents is a part of document segmentation. Some documents, such as maps, musical scores, written checks can be seen as a superposition of multiple information layers [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").

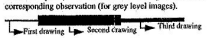
"Line drawing" (also called "segment") definition

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).

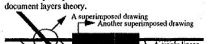


- (3) Use of the different parameters related to each representative point, such as thickness and brightness of the corresponding observation (for grey level images).

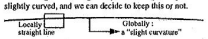


- (4) extraction of segments of different sizes, (ranging from a few points to several hundred representative points).
- (5) possibility of a drawings superimposition, e.g. :
  - staves in a musical score covered by quavers bars,
  - base lines on a written check.

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.



Public code & data

https://github.com/**EPITAResearchLab**/**bernet.23.icdar**

P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*

Contributions

1. **Public implementation** of the original approach

Public code & data

https://github.com/**EPITAResearchLab**/**bernet.23.icdar**

P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*

Contributions

1. **Public implementation** of the original approach
2. Broader, **extensible framework** based on **Multiple Object Tracking**

Public code & data

https://github.com/**EPITAResearchLab**/**bernet.23.icdar**

Contributions

1. **Public implementation** of the original approach
2. Broader, **extensible framework** based on **Multiple Object Tracking**
3. Extensive **benchmark**

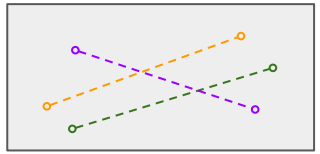
Public code & data

P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*

Contributions

1. **Public implementation** of the original approach
2. Broader, **extensible framework** based on **Multiple Object Tracking**
3. Extensive **benchmark**

   Task 1: Coarse vectorization
   MOT >> SotA training-free approaches



Public code & data

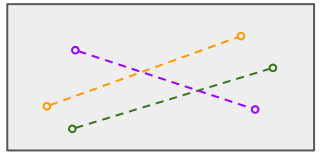
https://github.com/**EPITAResearchLab**/**bernet.23.icdar**

P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*

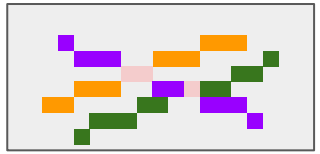
Contributions

1. **Public implementation** of the original approach
2. Broader, **extensible framework** based on **Multiple Object Tracking**
3. Extensive **benchmark**

Task 1: Coarse vectorization
MOT >> SotA training-free approaches



Task 2: Instance segmentation
comparison of tracking algorithms



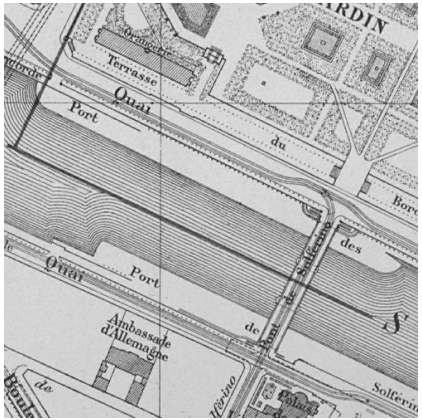
Public code & data

https://github.com/**EPITAResearchLab**/**bernet.23.icdar**

Linear Object Detection in Document Images using **Multiple Object Tracking**

P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*

ICDAR'23
Poster Session 2
Wednesday, Aug. 23



Public code & data

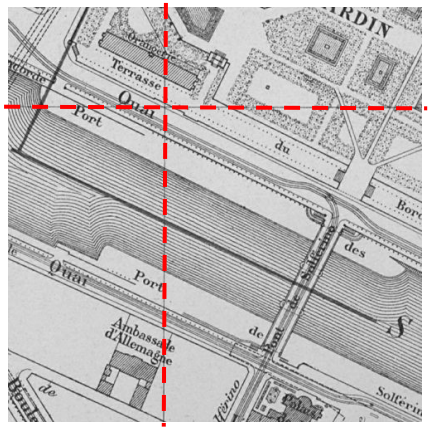
https://github.com/**EPITAResearchLab**/**bernet.23.icdar**

Linear Object Detection in Document Images using **Multiple Object Tracking**

P. Bernet, J. Chazalon, E. Carlinet, A. Bourquelot and E. Puybareau — *EPITA Research Lab (LRE), France*

ICDAR'23
Poster Session 2
Wednesday, Aug. 23

detect linear objects
then **rotate**



detect linear objects
then **georeference**

Public code & data

https://github.com/**EPITAResearchLab**/**bernet.23.icdar**