

Generic Document Image Dewarping by Probabilistic Discretization of Vanishing Points

Gilles Simon and Salvatore Tabbone

Université de Lorraine, CNRS, Inria, LORIA UMR 7503



The document image dewarping problem



↓ OCR

même nature provenant de l'abbaye de Saint-Bertin et Conservé au musée de Saint-Omer. Des études plus récentes Visent à nous en dissuader, et l'on admet actuellement Qu'aucune création mosane conservée ne conduit à la croix Solennellement consacrée à saint-Denis le jour de Pâques De 1147 par le pape Eugène III

Perspective distortion



même nature provenant de l'abbaye de Saint-Bertin et conservé au musée de Saint-Omer. Des études plus récentes viennent à nous en dissuader, et l'on admet actuellement que aucune création mosane concrète ne conduit à la croix solennellement consacrée à Saint-Denis le jour de Pâques de 1147 par le pape Eugène III.

En face des ateliers mosans, le grand centre rhéno-germanique s'affirme avec une originalité qui lui est propre. La ville, capitale d'un puissant archevêché, était le cœur de la Basse-Lobbeirie et bénéficiait d'un long et brillant passé artistique.

Vers 1150, un certain Eilbert, seigneur de Cologne, imprime son nom sur un anneau portant du trépas des Welfs dans la cathédrale de Brunswick – au Kunstgewerbe Museum de Berlin –, ainsi imagine par son dessin peintur

208 209 - LES ARCHANGES GABRIEL ET MICHAËL - SÉRAPHINS ET CHÉRUBINS. Quatre plaques en ivoire d'époque de la cathédrale de Saint-Martin. Vers 1170. Hauteur : 60 cm.



Page curl



même nature provenant de l'abbaye de Saint-Bertin et conservé au musée de Saint-Omer. Des études plus récentes visent à nous en dissuader, et l'on admet actuellement qu'aucune création mosane n'est parvenue à la croix solennellement consacrée à Saint-Denis le jour de Pâques de 1147 par le pape Eugène III.

En face des ateliers mosans, le grand centre rhéno-germanique de Cologne, capitale d'un puissant archevêché, était le cœur de la Basse-Léonhardie et bénéficiait d'un long et brillant passé artistique.

Vers 1150, un certain Eilbert, seigneur de Cologne, imprime son nom sur un autel portant du trépas des Welfs dans la cathédrale de Brunswick – au Kunstgewerbe Museum de Berlin –, ainsi qu'une par son dessin peint sur

que par son style. Celui-ci s'enracine dans la tradition romane sans se couper de Byzance. L'iconographie demeure étrangère aux spéculations théologiques chères aux Mosans. Sur une autre œuvre colongaise fameuse, la chaise de saint Hirsbert, le fondateur de l'abbaye de Deutz (vers 1105-1170), la tradition mosane est perceptible dans la technique et la couleur des encaux illustrant la vie du saint. Par contre, le milieu local impose à la chaise son architecture et sa composition. Cette création, la première de la série des grandes chaises en or qui devaient faire la gloire de Cologne, joue un rôle décisif pour conduire l'architecture romane à sa maturité.

Cologne finit par exercer la pénétration des influences architecturales dans le domaine des objets culturels et d'une manière plus générale, des objets de petit format.

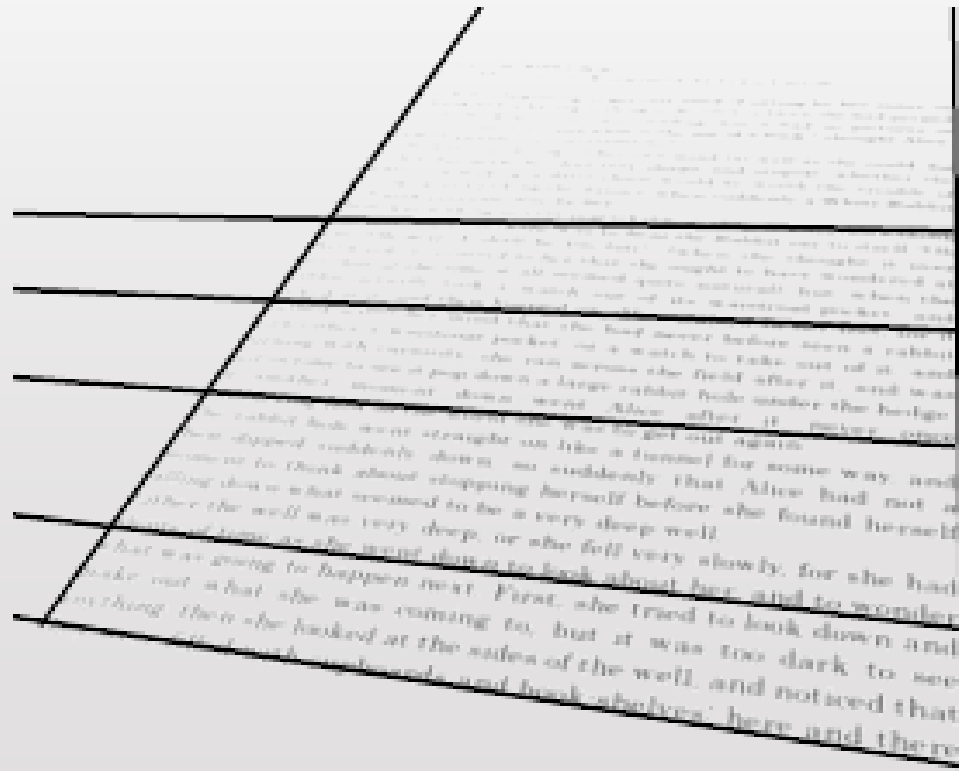
205-208 - LES ARCHANGES GABRIEL, MICHAEL, SERAPHIN ET CHERUBIN. Quatre plaques en métal décorant la chaise de saint Maurice. Vers 1170. Hauteur : 60 cm.

324



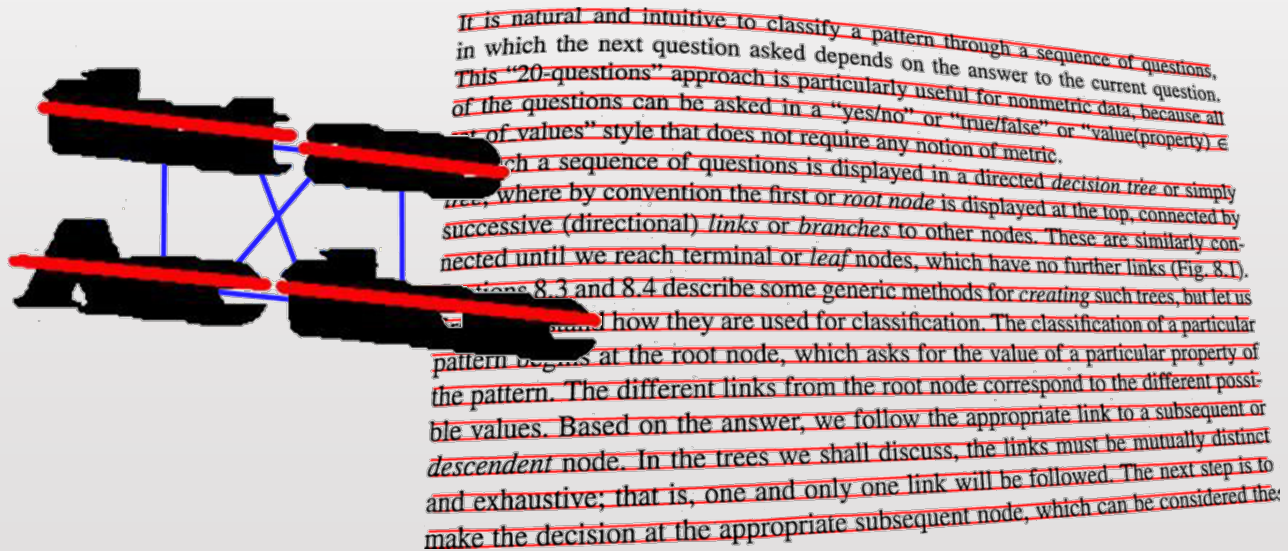
Previous methods

- Perspective distortion can be corrected by detecting vanishing points [Takezawa et al. 2016 & 2017]
- Not effective in case of curved pages



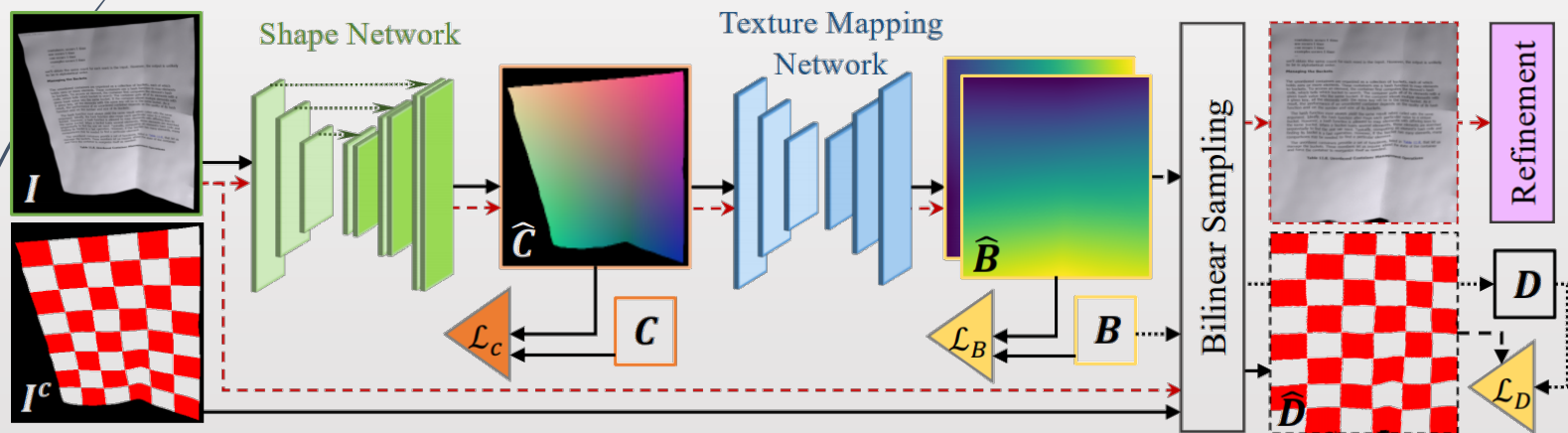
Previous methods

- Curved pages can be handled by classical image analysis (connected components, baseline fitting etc.) [Liu et al. 2015]
- Does not take into account the overall perspective of the document
- Not suited to non text regions



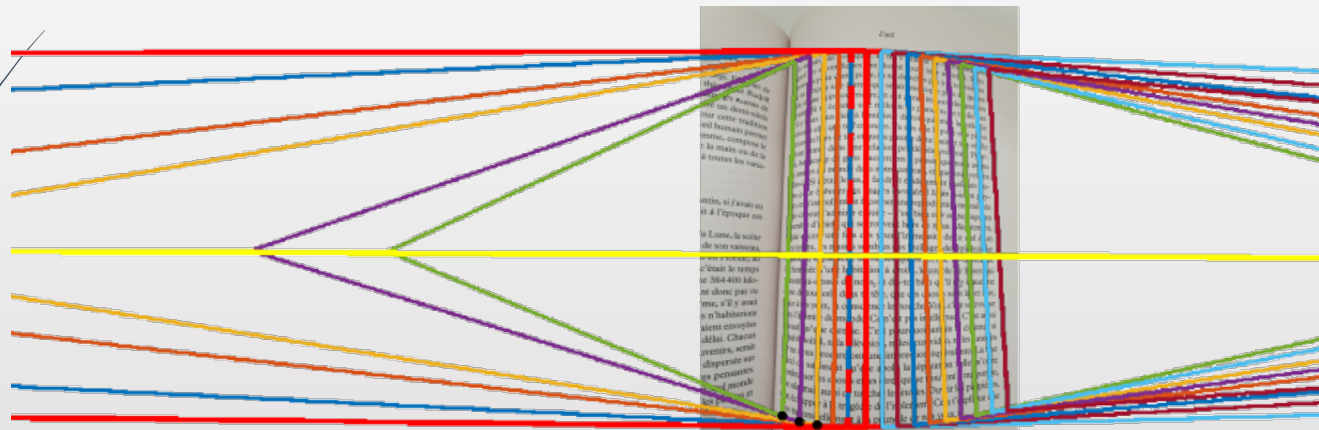
Previous methods

- Deep learning technique may solve both problems simultaneously [Das et al. ICCV 2019]
- Require an important number of learning samples
- Generally lack precision



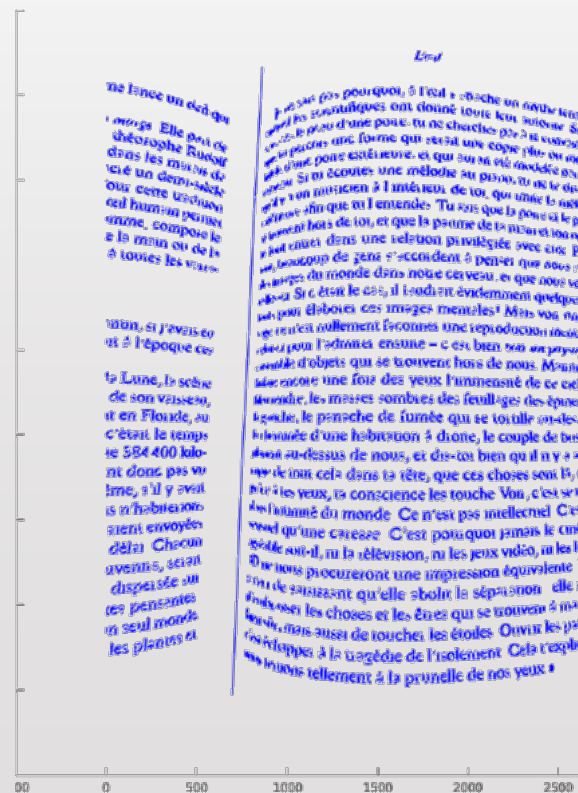
Main contribution

Under a generalized cylinder shape assumption, perspective and page curl can be dewarped simultaneously by detecting vanishing points along the "horizon line"



Algorithm

1. A line segment detector [Grompone et al. IPOL 2012] is applied to the image
 - ▶ Letter strokes as well as pictures and page boundaries provide a lot of lines, some of which are “vertical” (in 3D), others “horizontal”



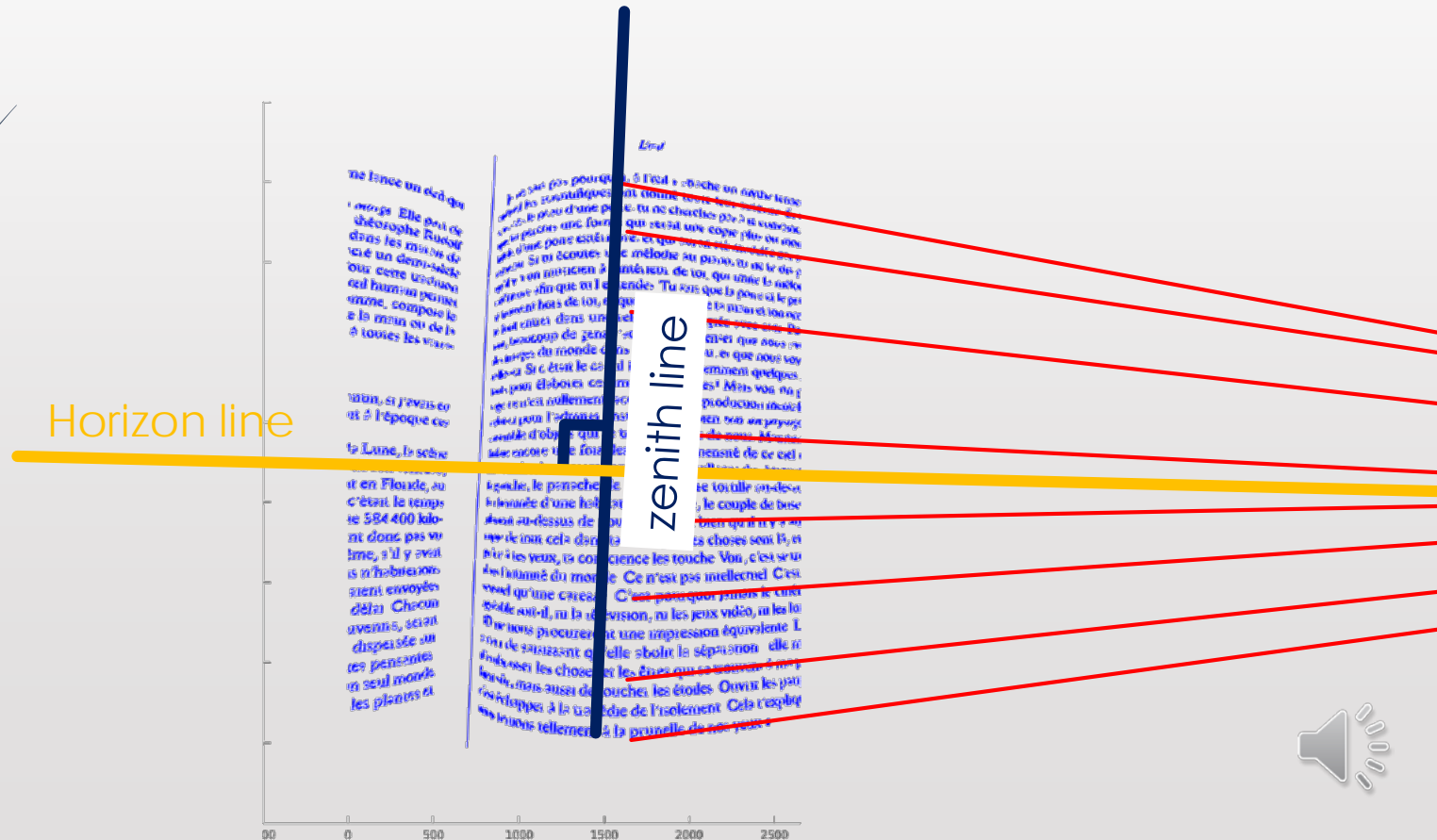
Algorithm

2. The vertical line segments allow to compute the zenith and the zenith line



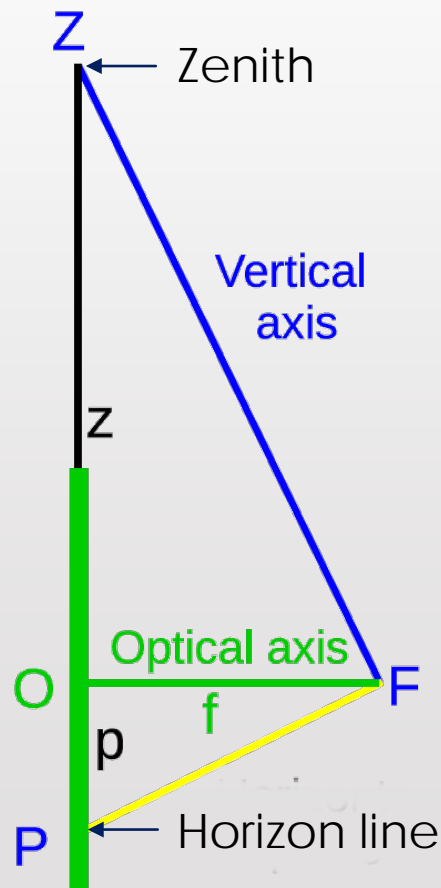
Algorithm

3. The horizontal line segments allow to compute the horizon line perpendicular to the zenith line



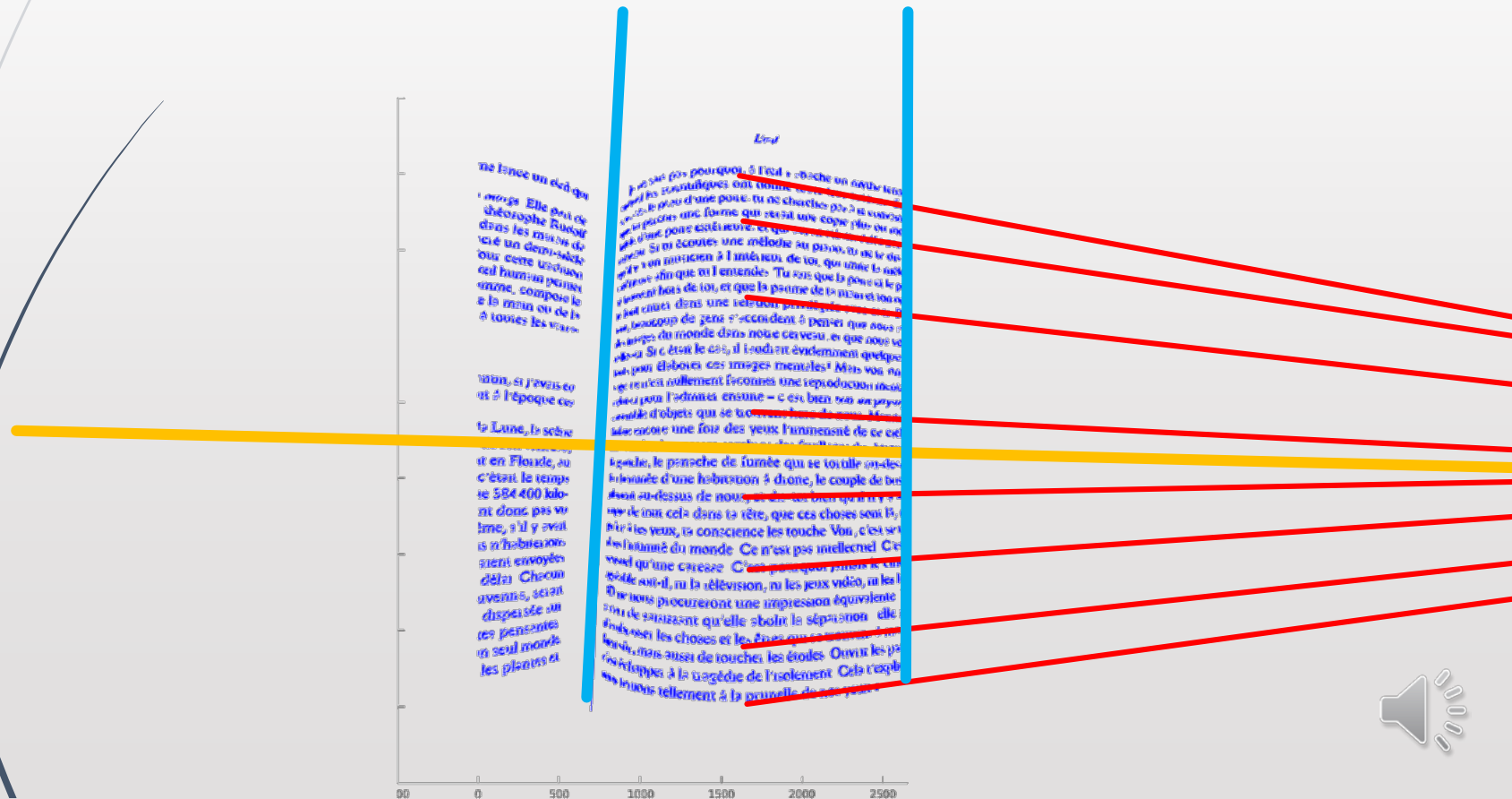
Algorithm

4. The focal length f is deduced from the positions of the zenith and horizon line relative to the principal point



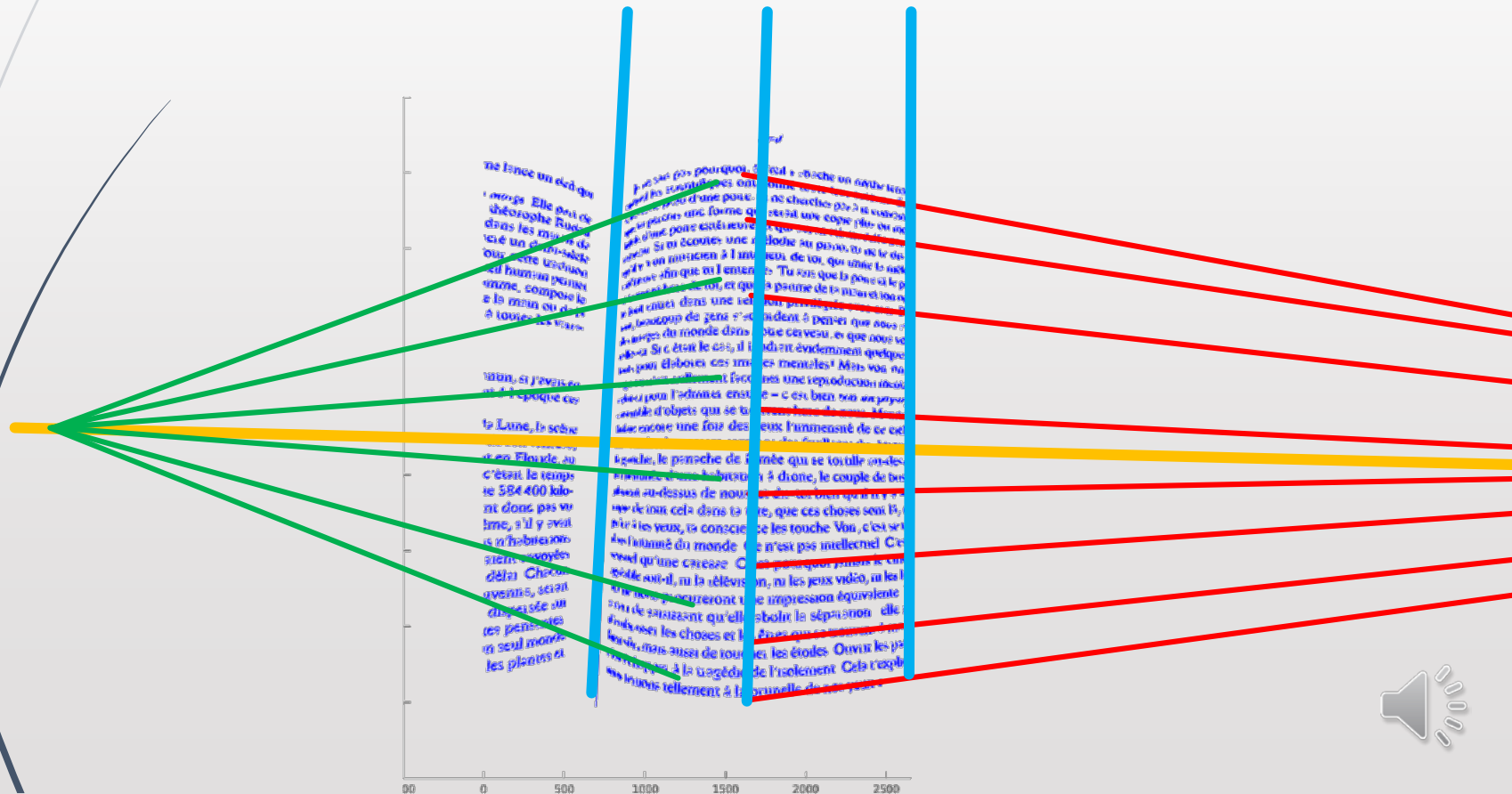
Algorithm

- 5. Vanishing points are detected along the horizon line by performing a binary-tree descent search



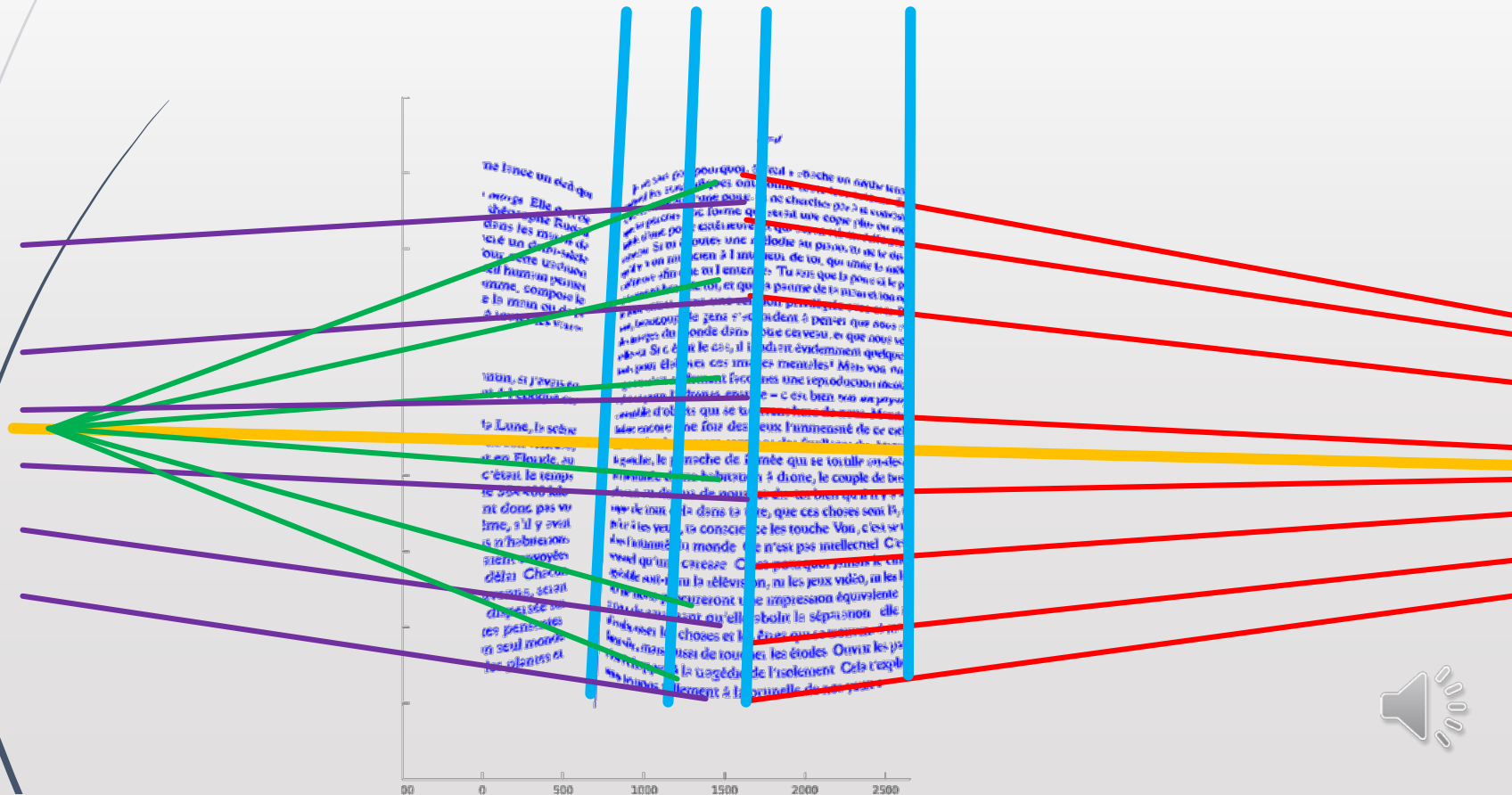
Algorithm

5. Vanishing points are detected along the horizon line by performing a binary-tree descent search



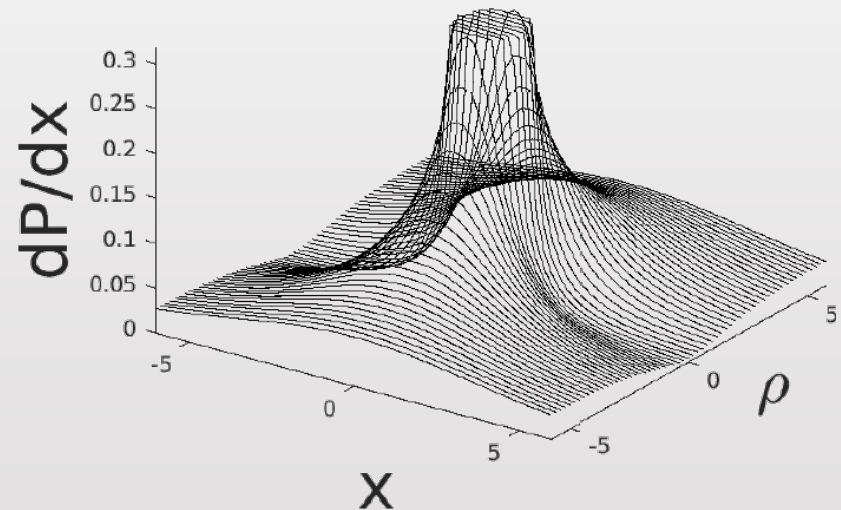
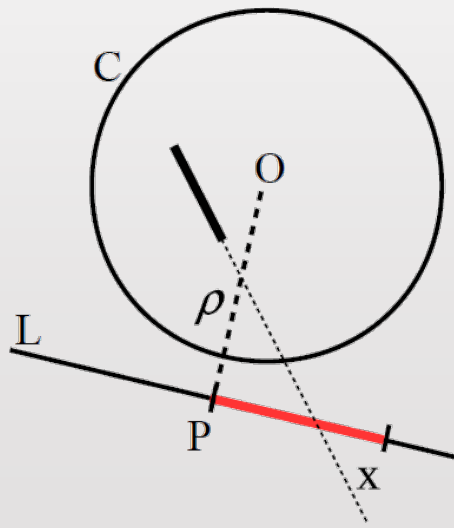
Algorithm

5. Vanishing points are detected along the horizon line by performing a binary-tree descent search based



Algorithm

5. Vanishing points are detected along the horizon line by performing a binary-tree descent search
 - ▶ A probabilistic criterion is used to decide when to stop cutting: a vanishing point is said *meaningful* if it cannot be obtained (at least once) *by chance*

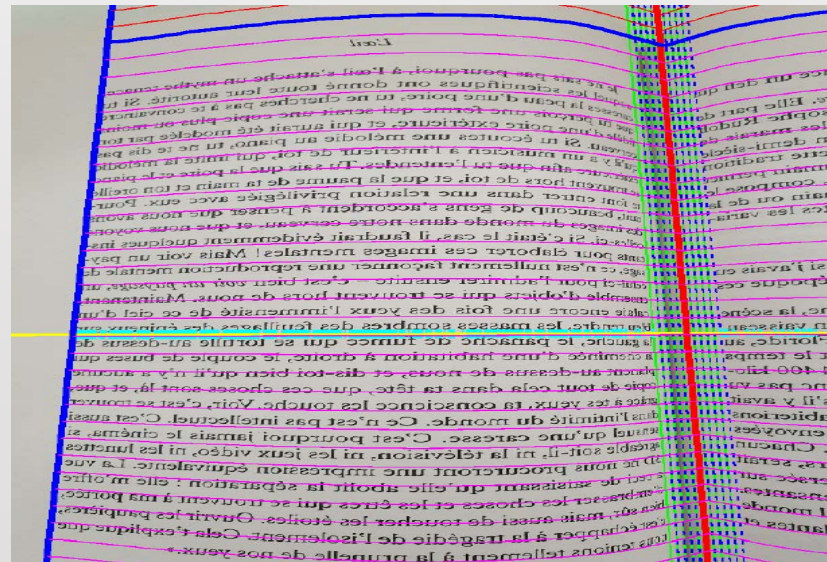
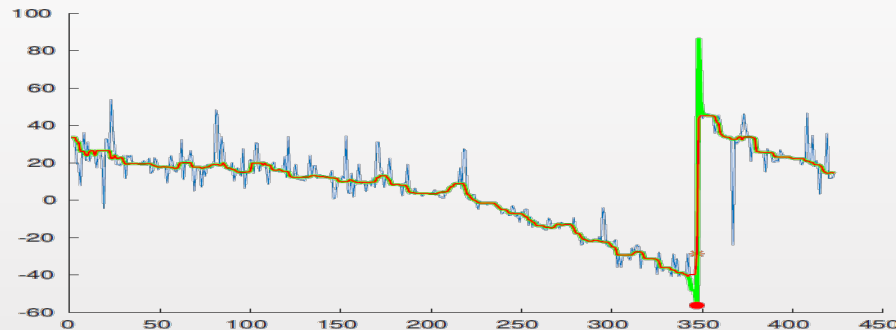


Probability distribution that a chord of circle C meets a line L at position x *by chance*



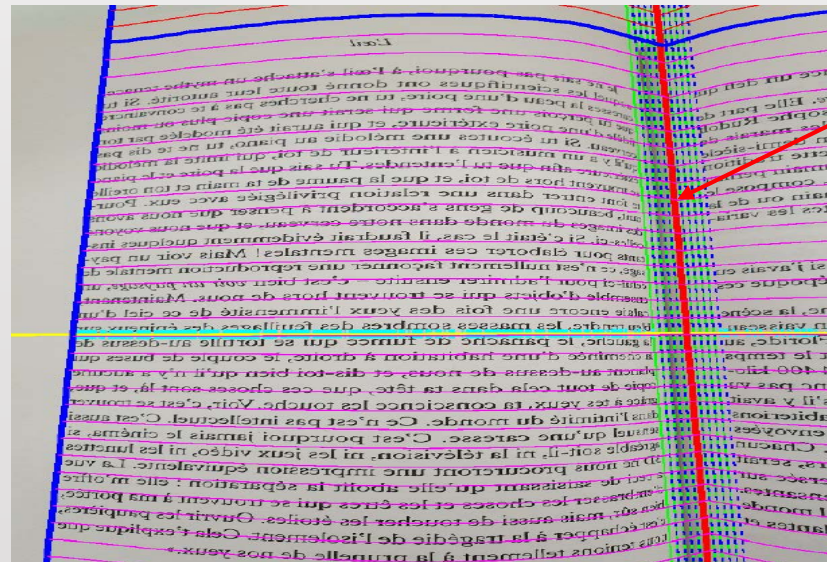
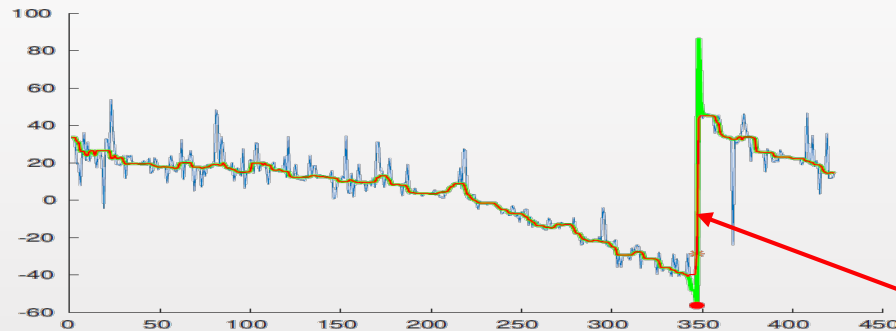
Algorithm

6. The 3D directions of the vanishing points are computed and their angles with the camera x-axis are median filtered



Algorithm

- The 3D directions of the vanishing points are computed and their angles with the camera x-axis are median filtered

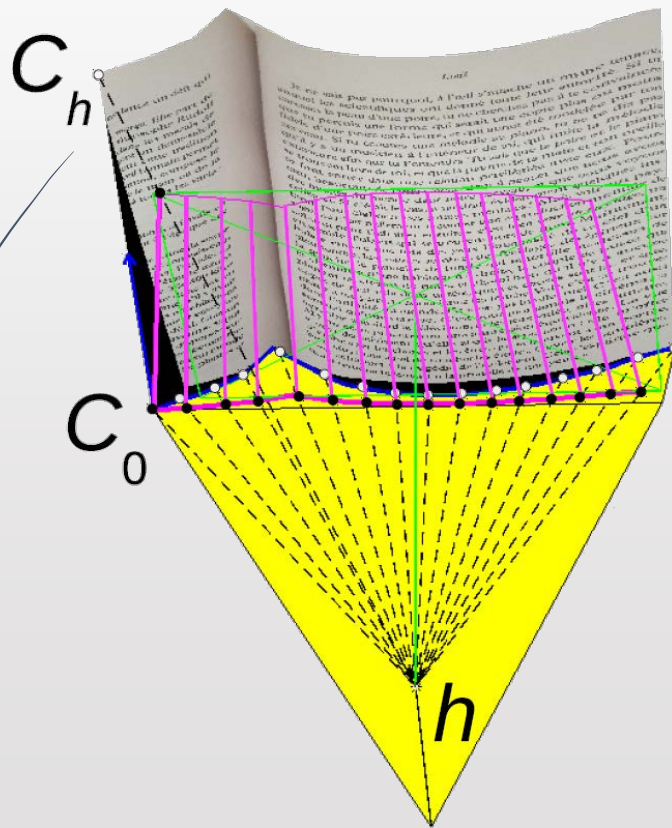


Page break



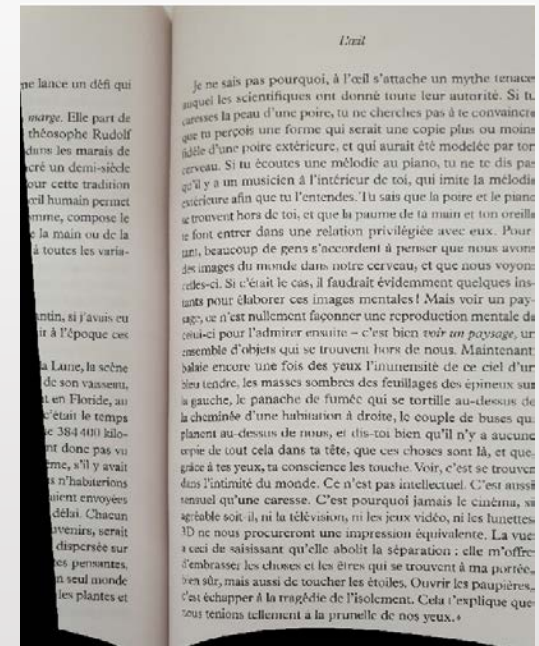
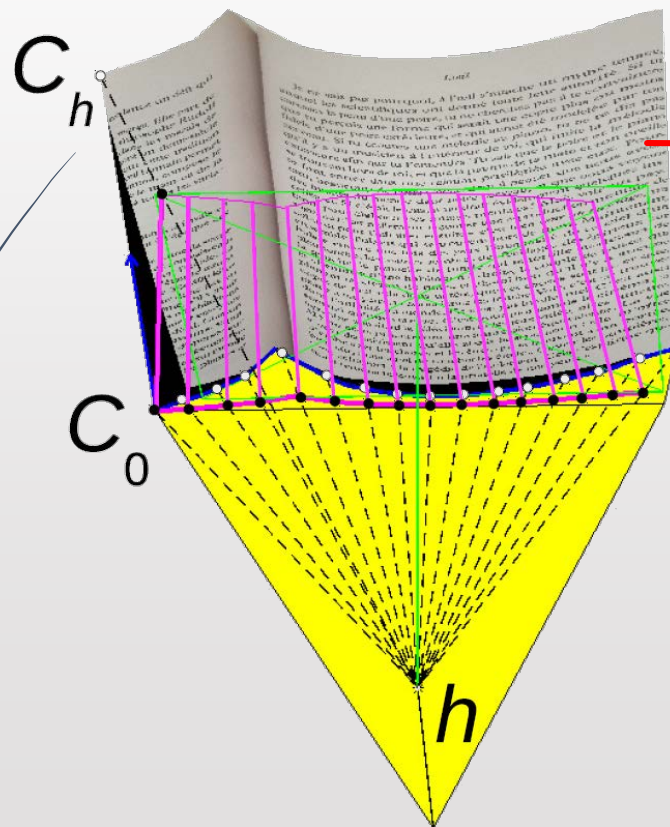
Algorithm

7. The filtered bottom curve then the full document are reconstructed in 3D

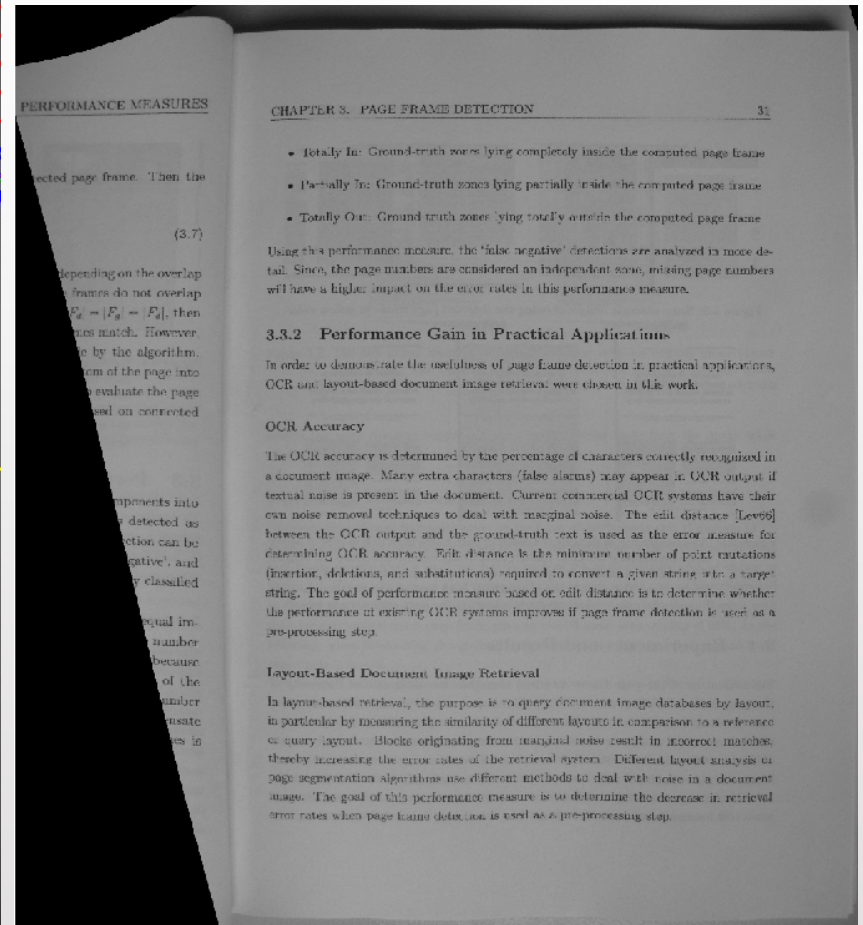
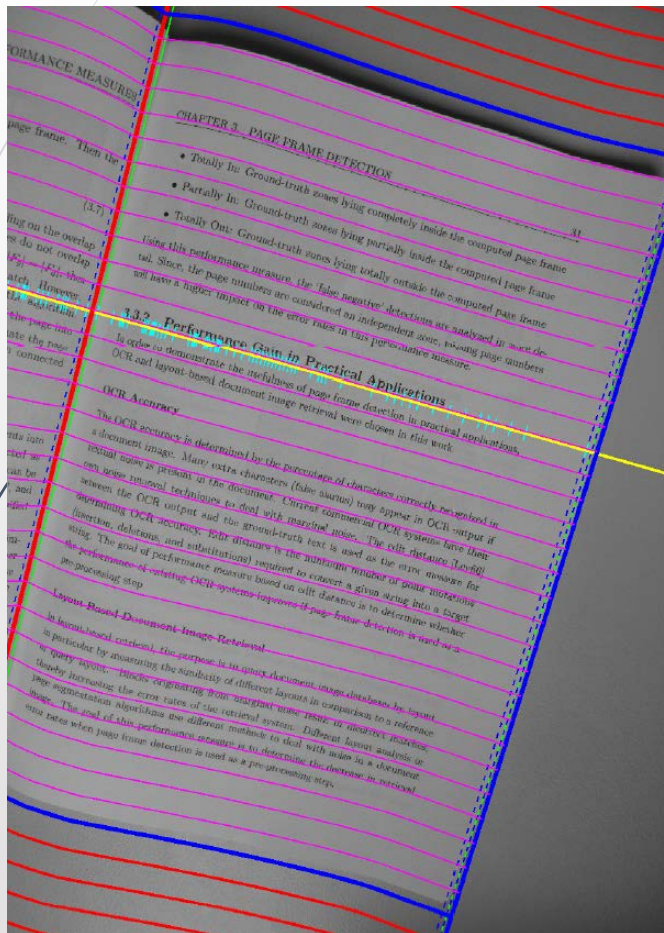


Algorithm

8. The image is back-projected onto the 3D model which finally only has to be unfolded to get the dewarped image



Some qualitative results

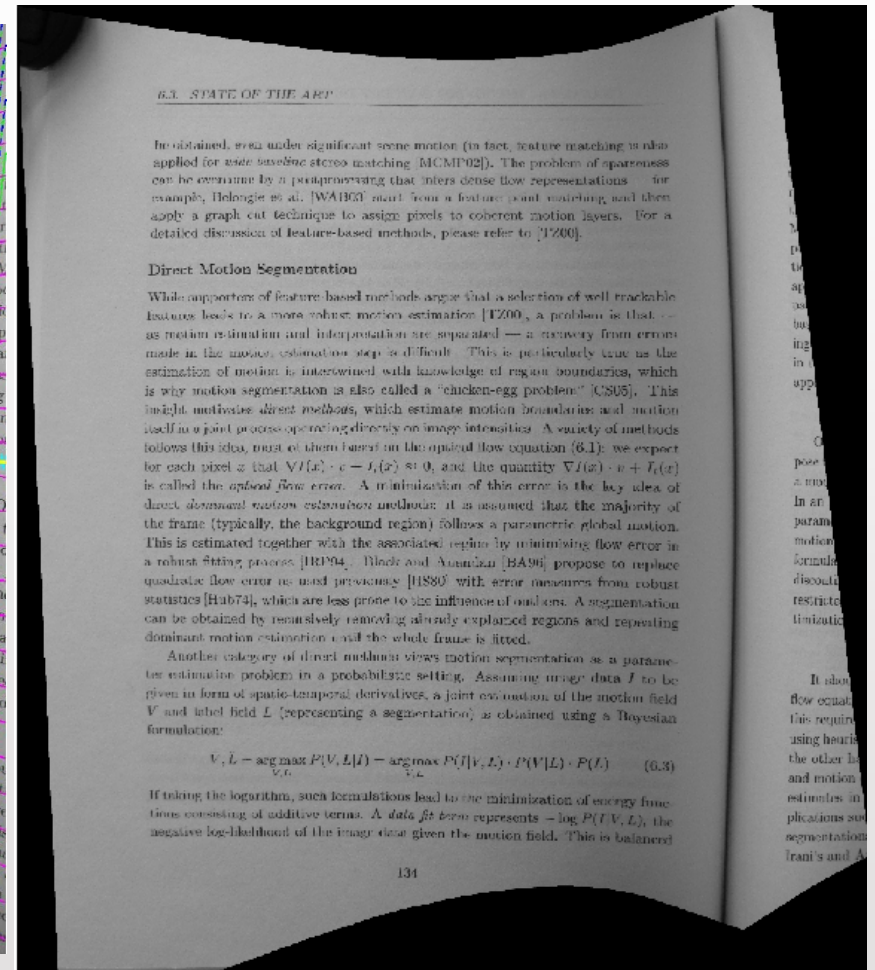
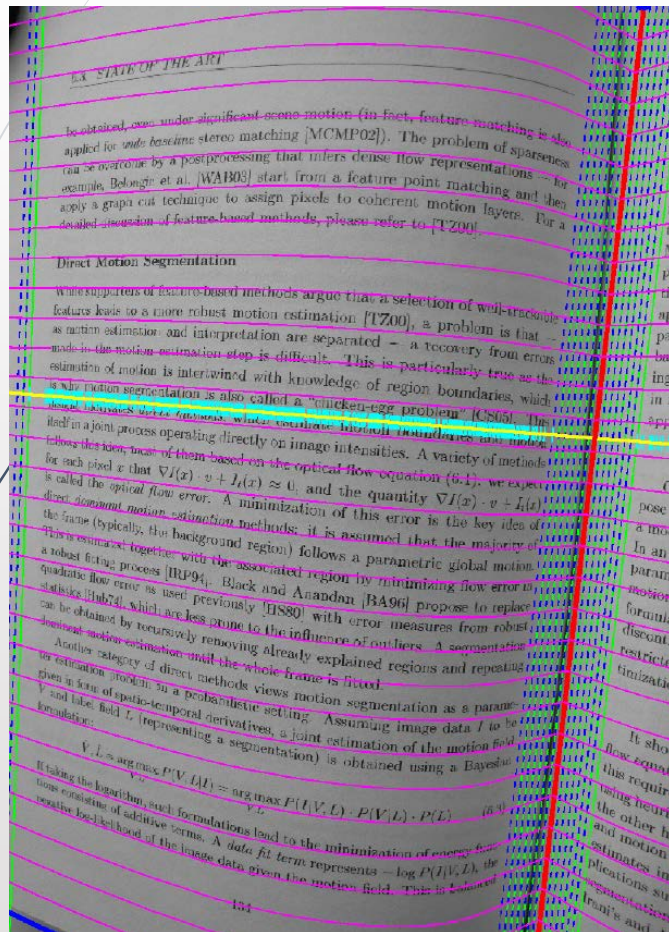


0

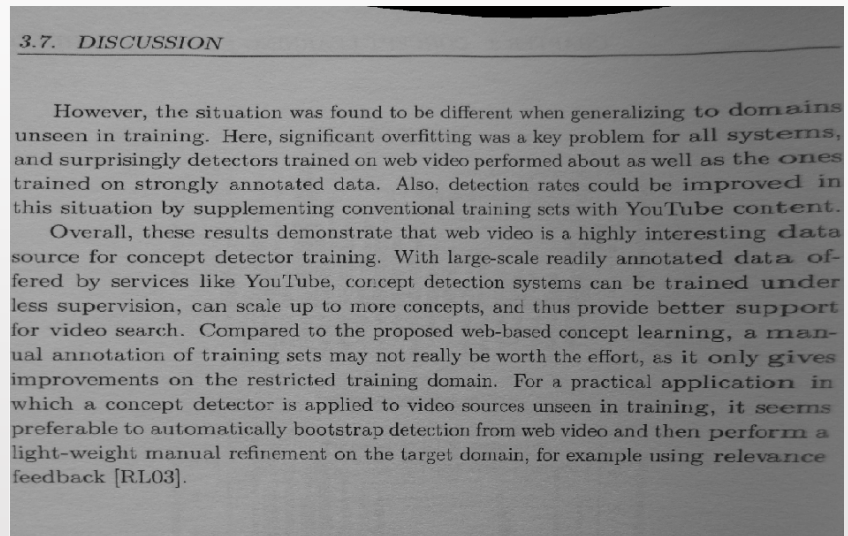
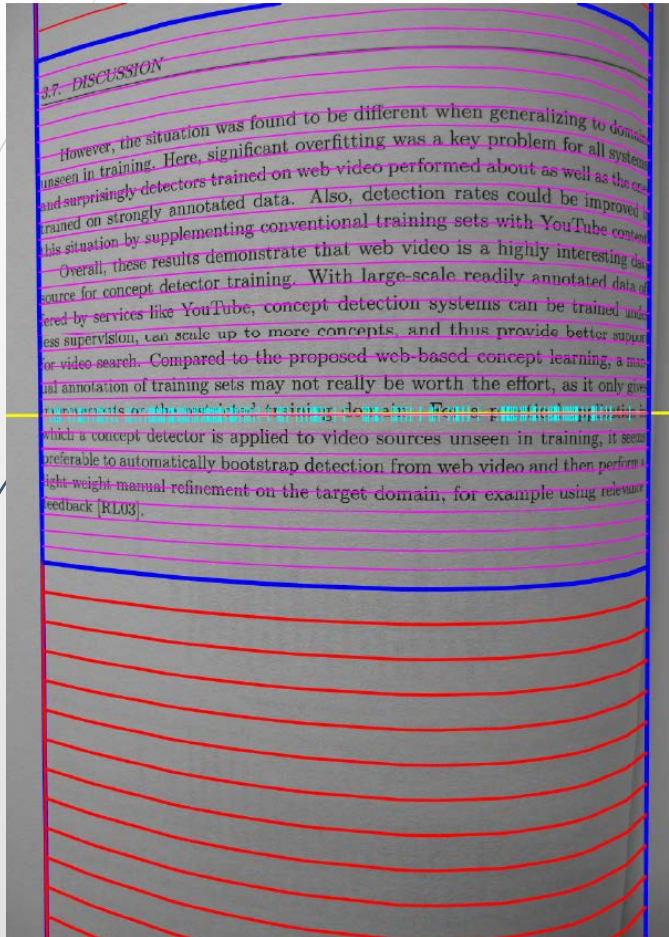
302

302

Some qualitative results



Some qualitative results

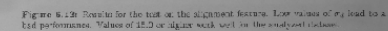


126

172

180



[illegible]

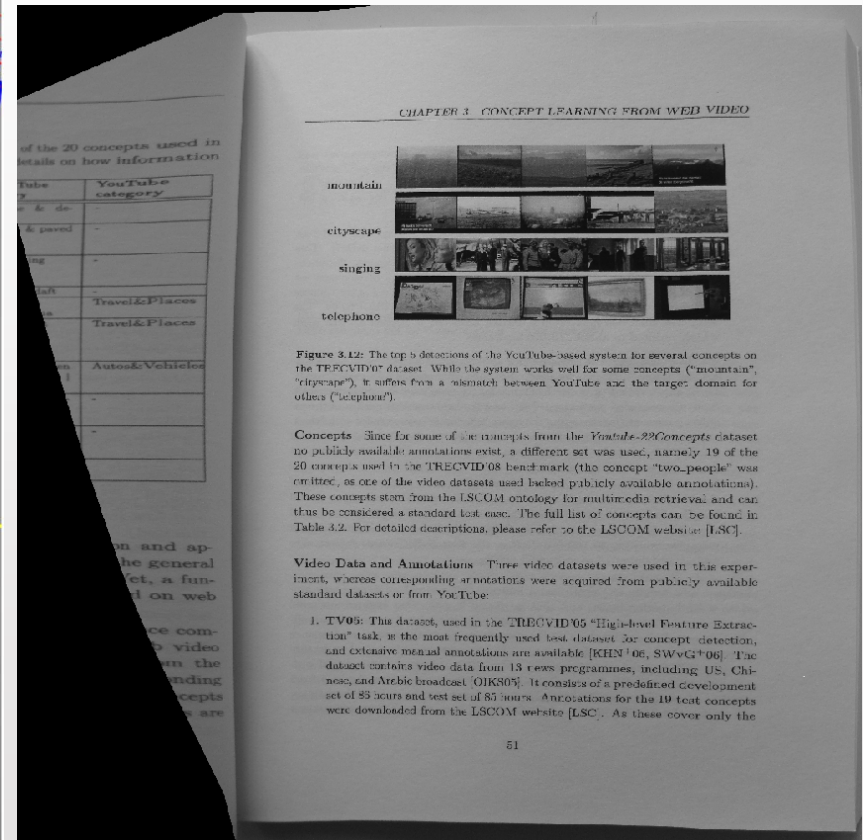
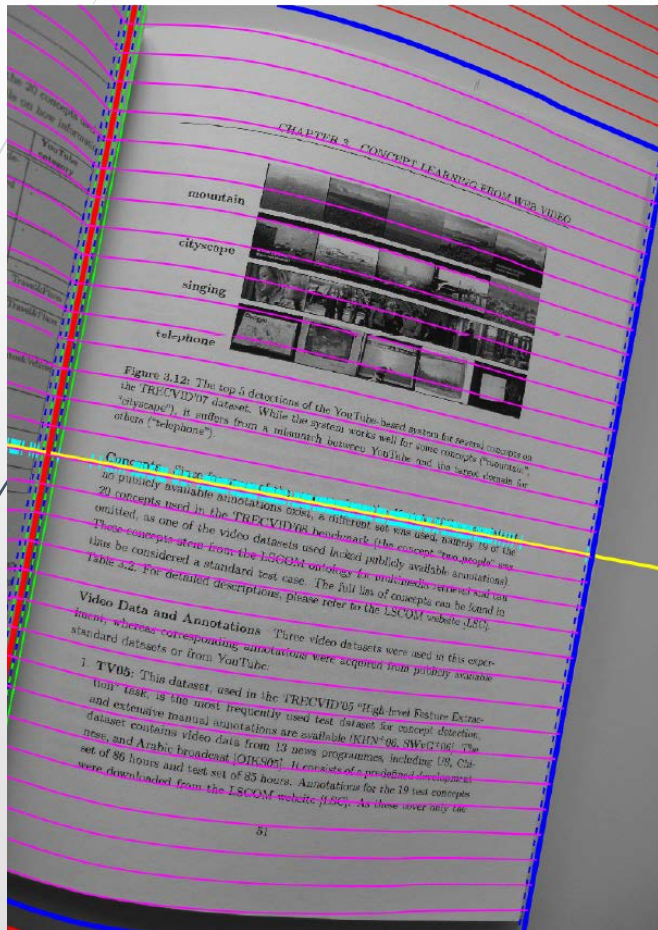
In contrast to the previous results on the slow angle-based approach, the results on the TP300 dataset are better than for the FPC300 dataset. This is likely due to the increased amount of control over the alignment in the manual forgery process, than in the case of the two-pass print forgeries, where the alignment is only partially controllable by adjusting the document page in the scanner tray accordingly.

An inspection of the errors made by the system showed the following reasons for failure:

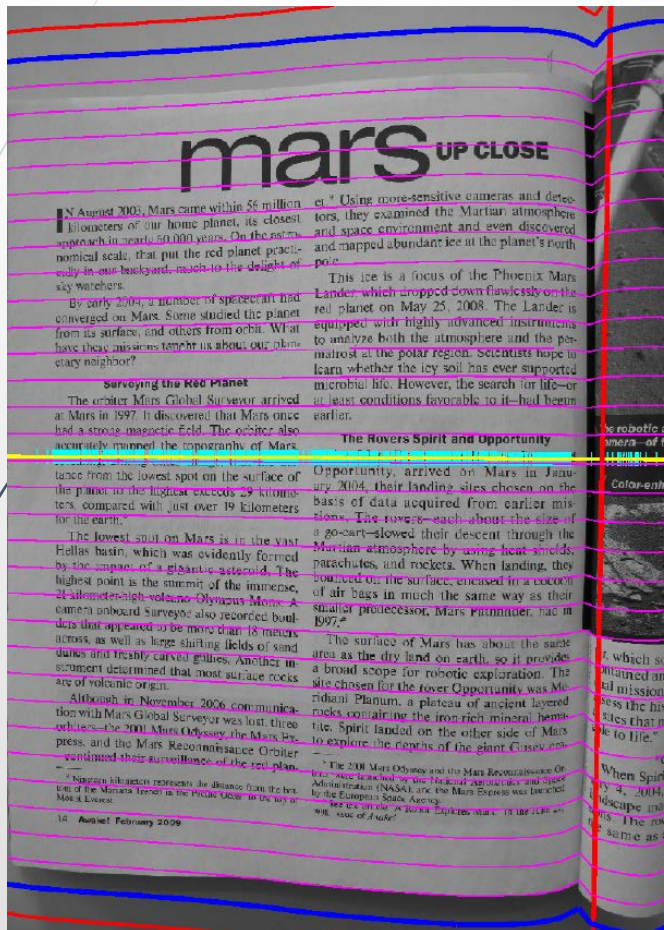
- typographic enhancement problems, as e.g. italics at the start of the line or different characters that match outside the line boundaries, as e.g. an italic 'f' at the start of the line; for oblique correction *recessive* characters (e.g. 'f') at the start of the 'normal' text-line borders, leading to false positives as can be seen in Figure 6 (4b).
- wrong alignment lines: some more indented text-lines are present than regular ones or if the number of folded text-lines is higher than the number of original text-lines, the wrong alignment lines may be found leading to



Some qualitative results



Some qualitative results



445

508

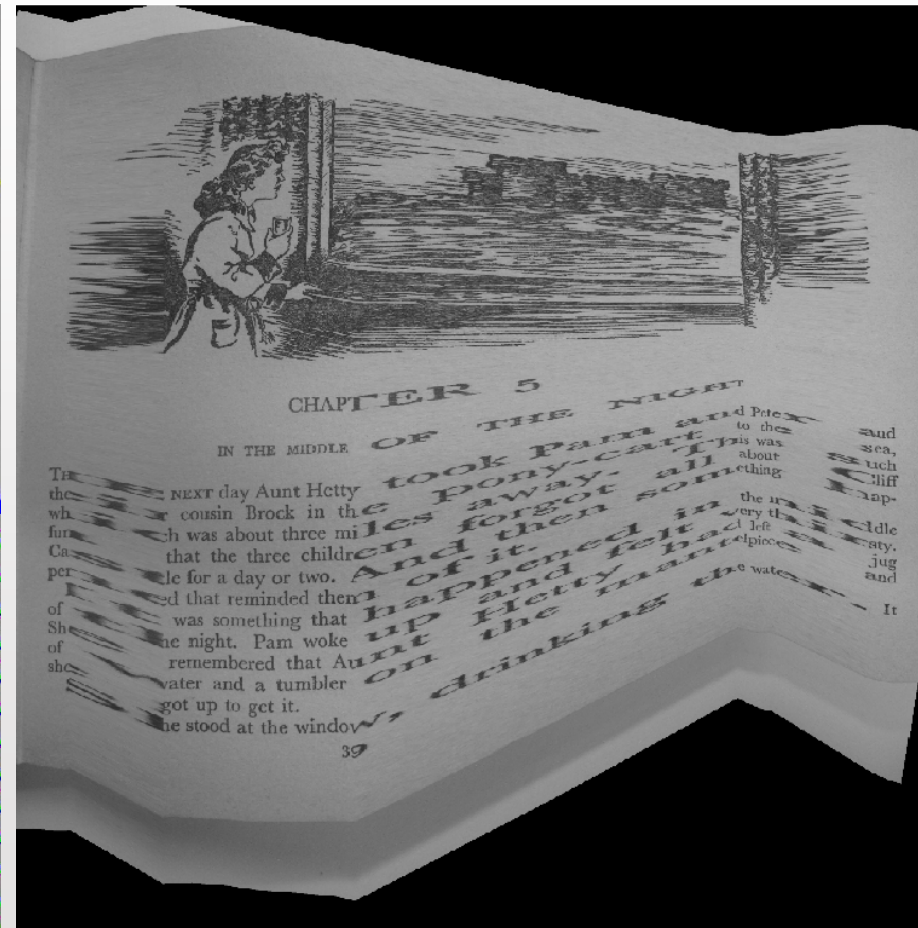
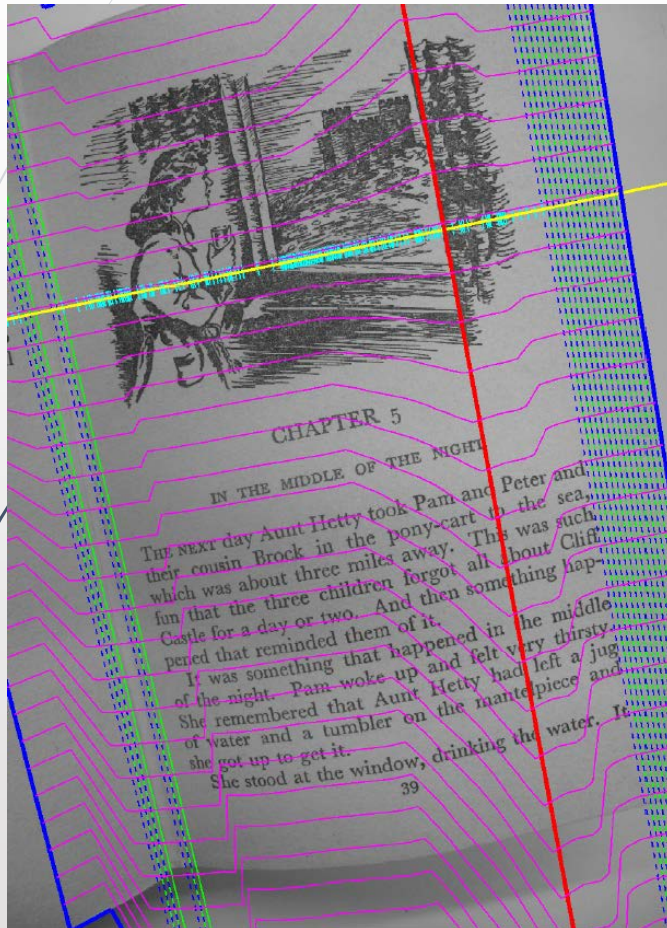
508

Quantitative results

- Recovered words on the IUPR ICDAR dataset

Our method	> 90%
[Takezawa 2016 & 2017]	~ 50%
[Chen et al. 2010]	~ 30%

A failure case



Conclusion

- Unlike many other dewarping methods ours does not require any text extraction, line segmentation or image binarization
- A generic method and interesting perspectives in areas other than document analysis

