



# The SOLAIRE Project: A Gaze-Contingent System to Facilitate Reading for Patients with Scotomas

Emilien Tlapale, Jean-Baptiste Bernard, Eric Castet, Pierre Kornprobst

## ► To cite this version:

Emilien Tlapale, Jean-Baptiste Bernard, Eric Castet, Pierre Kornprobst. The SOLAIRE Project: A Gaze-Contingent System to Facilitate Reading for Patients with Scotomas. [Technical Report] RT-0326, INRIA. 2006, pp.16. inria-00102542v2

**HAL Id: inria-00102542**

**<https://inria.hal.science/inria-00102542v2>**

Submitted on 3 Oct 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ***The SOLAIRE Project: A Gaze-Contingent System to Facilitate Reading for Patients with Scotomas***

Émilien Tlapale — Jean-Baptiste Bernard — Éric Castet — Pierre Kornprobst

**N°0326**

Octobre 2006

Thème BIO

 ***apport  
technique***





## The SOLAIRE Project: A Gaze-Contingent System to Facilitate Reading for Patients with Scotomas

Émilien Tlapale<sup>\*</sup>, Jean-Baptiste Bernard<sup>†</sup>, Éric Castet<sup>†</sup>, Pierre Kornprobst<sup>\*</sup>

Thème BIO — Systèmes biologiques  
Projets Odyssée

Rapport technique n° 1000 — Octobre 2006 — 16 pages

**Abstract:** Reading is a major issue for visually impaired patients suffering from a blind area in the fovea. Current systems to facilitate reading do not really benefit from recent advances in computer science, such as computer vision and augmented reality. On the SOLAIRE project (Système d'Optimisation de la Lecture par Asservissement de l'Image au Regard), we develop an augmented reality system to help patients to read more easily, resulting from a strong interaction between ophthalmologists and researchers in visual neuroscience and computer science. The main idea in this project is to control the display of the text read with the gaze, taking into account the specific characteristics of the scotoma for every individual. This report describes the system.

**Key-words:** Scotomas, Dégénérescence maculaire, systems to improve speed of reading

<sup>\*</sup> INRIA Sophia Antipolis, Projet Odyssée, 2004 Route des Lucioles, 06902 Sophia Antipolis

<sup>†</sup> INCM, Equipe DyVA UMR 6193 CNRS-Université de la Méditerranée 31, chemin Joseph Aiguier 13402 Marseille cedex

## Le Projet SOLAIRE: Un Système d'Optimisation de la Lecture par Asservissement de l'Image au Regard

**Résumé :** Un problème majeur rencontré par les déficients visuels qui sont aveugles dans la portion centrale de leur champ visuel est la lecture de texte. Les systèmes actuels d'aide à la lecture ne tirent pas partie de l'essor récent des recherches et technologies concernant la vision, telles que la vision par ordinateur ou la réalité augmentée. Dans le projet SOLAIRE (Système d'Optimisation de la Lecture par Asservissement de l'Image au Regard), nous développons un système de « vision augmentée » dédié à la lecture en combinant les compétences d'ophtalmologistes, de chercheurs en neurosciences visuelles et de spécialistes de vision artificielle. L'idée maîtresse de ce projet est d'asservir l'image du texte lu à la position instantanée du regard en adaptant le traitement de l'image aux caractéristiques de la pathologie propre à chaque patient. Ce rapport en décrit la mise en œuvre.

**Mots-clés :** Scotomes, Age related macular disease, systèmes d'aide à la lecture

# 1 Motivation

A scotoma is an area or island of loss or impairment of visual acuity surrounded by a field of normal or relatively well-preserved vision.

Common causes of scotomata include demyelinating disease such as multiple sclerosis (retrobulbar neuritis), toxic substances such as methyl alcohol, ethambutol and quinine, nutritional deficiencies, and vascular blockages either in the retina or in the optic nerve. Scintillating scotoma is a common visual aura in migraine. Less common, but important because sometimes reversible or curable by surgery, are scotomata due to tumors such as those arising from the pituitary gland, which may compress the optic nerve or interfere with its blood supply. The most important cause of scotomata is probably the age related macular disease (ARMD) with 13 million cases in the United States and 1 million in France. These numbers are increasing and are expected to be multiplied by three in the coming 25 years.

A pathological scotoma may involve any part of the visual field and may be of any shape or size. A scotoma may include and enlarge the normal blind spot. Even a small scotoma that happens to affect central or macular vision will produce a severe visual handicap, whereas a large scotoma in the more peripheral part of a visual field may go unnoticed by the bearer. When the scotoma affects the fovea region, some activities such as reading become very difficult, since the zone of sharp vision is impaired.

To improve reading performance of ARMD patients, many different augmented vision systems have been designed over the last 15 years (see [12] for a review). The most important aspect of these systems is a digitalization of the text to be read. It is thus possible to apply to the text different image processing techniques which transform the initial text into an image format more susceptible to compensate partly for the detrimental effect of the scotoma. Two of these transformations are especially efficient.

- The whole text is magnified so that the size of the words becomes more adapted to the low visual resolution of spared peripheral regions around the macula.
- The luminance contrast of the text is increased. Although these systems have improved patients' reading performance to some extent, their efficiency is still disappointing especially when considering the vast potential of recent advances in artificial vision and the ever increasing power of cheap computers.

Our project aims at developing an innovative augmented-vision reading system based on real-time gaze-contingent transformations of the image with the goal of significantly increasing patients' reading speed (compared to current systems). The use of gaze-contingent viewing is the crucial aspect of our system as it should remove the necessity for patients to use complex and time-consuming oculo-motor strategies.

In this article we describes our system, and review the algorithms which can be used at each step of the implementation. The proposed system has two components, as described in Figure 1. When the patient has digitalized his text with a webcam for example (Section 2), the digital image of the text is analyzed. In the first component, the objective is to extract

from the digital image the content, which is the text recognized (see Section 3). The second component refers to the dynamical aspect of the display. At any moment, the patient can choose a paragraph and decide to read it (Section 4). Once the patient selects the paragraph on the initial digital image, we display the recognized text under controlled conditions, with some visual enhancements (Section 5).

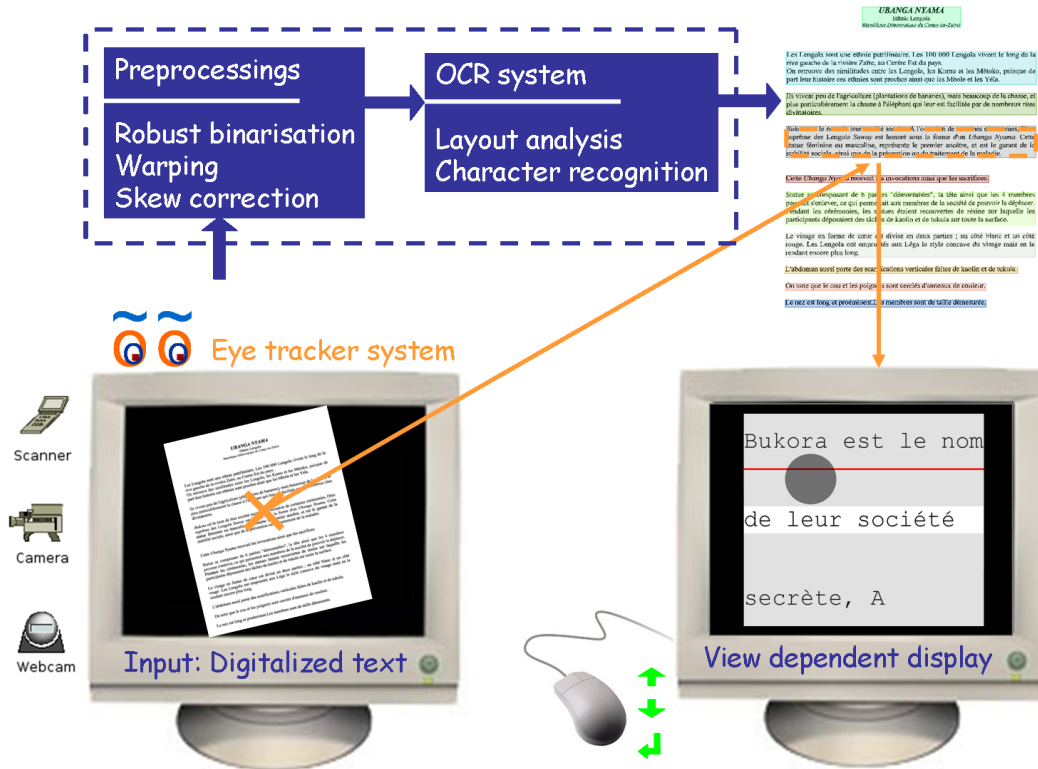


Figure 1:

When the patient has digitalized his text with a webcam for example, the digital image of the text is analyzed so that the content is extracted. The patient can then select which paragraph he wants to read, just by fixing the paragraph and a mouse click. The selected paragraph is then displayed under controlled conditions, which are dependent on the sight of the patient and the shape of his scotoma. Using for instance the mouse, the patient may scroll up and down the text, or come back to the view of the digitized text.

## 2 What is to be analyzed?

The digital image acquisition is the entry point of our system. The acquisition grabs a digital image from a physical device such as a scanner, a webcam or a digital camera. The acquisition procedure can either use a high fidelity device such as a flatbed scanner, or a lower resolution grabber like a webcam. Of course, the result may be very different, as shown in Figure 2, which will have some consequence on the type of necessary preprocessing.

Of course, the acquisition procedure has an influence on the required document analysis steps.

- The acquisition *resolution* which influence the quality of the character recognition described in section 3.4.
- The *noise* induced by the acquisition procedure.
- The *illumination* having a great importance in binarization as we will see in section 3.1.
- The document *deformation* including the skew of the document for most of the acquisition devices but also some 3D transformations for warped documents which can be corrected by algorithms such as [6, 9].

In the current system, acquisition is done on GNU/Linux via the Video4Linux or the SANE libraries. Having a TWAIN or a WIA interface would be interesting to get Microsoft Windows support.

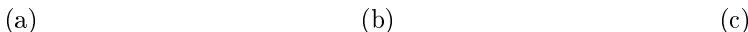
Beside the quality, the layout, i.e., the shape of a document may be very different. To illustrate this, we represent in Figure 3 some examples.

## 3 Document analysis

The front-end is very similar to current commercial optical character recognition (OCR) systems. The OCR step is mandatory since we want to display the text at a high quality to the user and transformations on images will reduce their definition (think of a zoom). Since we are under controlled condition (we can decide the mode of acquisition), and since that we do not consider handwritten text, we claim that it is not necessary to use some sophisticated OCR. In this section we explain the main steps to build an OCR:

- Image binarization. This is done via local thresholding algorithm since factors such as lighting conditions or paper quality cannot be guaranteed.
- Document layout analysis. This step will identify the parts of the document such as the paragraphs, the headings, the images and select an order to display them.
- Optical character recognition of the text contained in the different textual regions so it can be dynamically displayed to the user.





morning. In this case, we prefer the schedule with the morning in the afternoon over the schedule with the morning in the morning, assuming that this time constraint allows the morning to be scheduled both in the morning and the afternoon.

The syntax of the logic is *preliterate* in accordance with Blauhaus [56], except for the syntax of first order logic by introducing a modal operator  $\Box$  with the rule of *detachment*: if  $A$  is a formula then so is  $\Box A$ . Following the model theoretical approach as outlined by Carnap [24, 16] and [49] (p. 54), possible worlds semantics for modal logic was introduced. Carnap [56] provides a possible worlds semantics for premodal logic. Roughly, the truth value of propositional logic under interpretation  $\omega$  is *determined* in the usual way, but a formula of the form  $\Box A$  is true at a world  $\omega$  if and only if  $A$  is true in *all* worlds  $\omega'$  accessible to  $\omega$  by the relation  $\prec$ , i.e.  $\omega \prec \omega'$ . The intuitive meaning of the relation  $\prec$  is that  $\omega'$  is a *possible* world accessible to  $\omega$  at any instant of time. The relation  $\prec$  is a *strict partial order*, i.e. it is reflexive, transitive and antisymmetric. We will be interested in frames where the relation  $\prec$  is a *total order* property that if  $\omega' \prec \omega$ , then there exists a formula  $A$  such that  $\omega' \models A$  and  $\omega \not\models A$ .

[illegible]

We briefly survey the various approaches to optimisation that have been proposed in the logic programming and deductive databases literature.

(a)

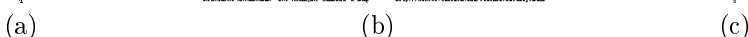


Figure 3: Some examples of document structures. (a) Single-column simple document, (b) Multi-column document, (c) Complex document with equations and images.

To appear in *ICMI 22 Workshop on Text Mining: Foundations, Techniques and Applications*

## Chen-Nan Hsu and Chien-Chi Chang

Inst. of Information Science, Academia Sinica  
Nankang 1<sup>st</sup> Sec. Taipei City, TAIWAN  
(chrisman,2001)@iis.sinica.edu.tw  
P/c: +886-2-2783790x101 Fax: +886-2-27834111  
<http://chrisman.iis.sinica.edu.tw/>

[illegible]

As approaches to text mining from semi-structured documents on the World-Wide Web, for example, this document is a collection of structured data and thus apply data mining techniques to analyze the data. This conversion involves information extraction from semi-structured documents. For example, consider a line

[illegible]

Persons work in industry careers like, e.g., [HARRIS, 1988], are usually not distinguished by their education level. However, they are distinguished by their work. Moreover, a widely used approach is to program self-referential rules in human expert systems that learn from their own experience. This is done by using neural networks in a specific fashion, just as self-referential neural networks. This requires that the system can learn the regularity of its own experience. This is done by using self-referential neural networks with their own set of weights in the  $n$ -components, based on the World Wide Web modules, *neurotic* being developed by many authors, e.g., [KLEIN, 1995], [KLEIN, 1996], [KLEIN, 1997], [KLEIN, 1998], [KLEIN, 1999], [KLEIN, 2000], [KLEIN, 2001], [KLEIN, 2002], [KLEIN, 2003], [KLEIN, 2004], [KLEIN, 2005], [KLEIN, 2006], [KLEIN, 2007], [KLEIN, 2008], [KLEIN, 2009], [KLEIN, 2010], [KLEIN, 2011], [KLEIN, 2012], [KLEIN, 2013], [KLEIN, 2014], [KLEIN, 2015], [KLEIN, 2016], [KLEIN, 2017], [KLEIN, 2018], [KLEIN, 2019], [KLEIN, 2020], [KLEIN, 2021], [KLEIN, 2022], [KLEIN, 2023], [KLEIN, 2024], [KLEIN, 2025].

[illegible]

Two factors increase the complexity of the problem. First, the graph structures can have nodes from two different sets (Kruskal and Pease, 1967; Daphin, 1987), one will restrict the degree of possible graph structures or assume that a structure is given.

and the model. As we have shown, it is simple to construct symbolic expressions for  $m$  in terms of the risk-neutral model parameters. The minimization algorithm differentiates the expressions symbolically to evaluate the gradient and Hessian along each linear  $x$ . The procedure is inexpensive since the expressions for  $d$  and  $e$  in Equations 7 and 8 have a particularly simple form.

interior of University High School in Urbana, Illinois. The superimposed lines indicate the edges the user has marked.

Figure 1 consists of two sub-images, (a) and (b), each showing a 3D perspective view of a simple house model. In image (a), a camera frustum is shown in the foreground, and several red dots representing camera poses are scattered around the house. In image (b), the same house model is shown, but the camera poses are represented by blue dots, indicating a different set of poses or a different procedure.

Given an observed edge  $u_{ij}$  the measured normal  $m^i$  to the plane

$$\mathbf{m}' = \begin{pmatrix} x_1 \\ y_1 \\ -f \end{pmatrix} \times \begin{pmatrix} x_2 \\ y_2 \\ -f \end{pmatrix} \quad (7)$$

Figure 4: The high school model, reconstructed from the *in situ* photographs of the *in situ* model. The model is composed of a series of vertical lines, each camera rotation can be viewed by the camera. The camera rotation is independent of the camera's location in space. Our method is to minimize the following objective function  $F$ , that sums the errors to which the rotation  $R_i$  violate the constraints arising from Equation 5:

Once initial estimates for the camera rotations are computed, Equation 6 is used to obtain initial estimates of the model parameters and camera locations. Equation 6 reflects the constraint that all of the points on the line defined by the tuple  $\{y, d\}$  should lie on the plane with normal vector  $\mathbf{n}$  passing through the camera center.

In the special case where all of the black relations in the model have a known notation, this objective function becomes a simple quadratic form which is easily minimized by solving a set of linear equations:



### 3.1 Binarization

The main purpose of the binarization step is to discriminate between the text and the background regions so the later processes are made easy. Indeed doing a skew detection, a layout analysis or a character recognition with a set of pixels having only two values is really more simple than having a whole greylevel set. Discriminating between text and background also means dealing with the noise and the illumination of the image. The thresholding step should also preserve the characters, no splitting or merging them, otherwise the OCR work would be very hard.

The binarization algorithms are classically classified as *global* or *local* thresholdings: for each pixel, if its value is lower than its threshold, then it is classified as text, otherwise as background. For global thresholding algorithms there is only one threshold value for the whole image, as opposed to local thresholding.

- Global algorithms [20, 25, 21] are the most simple as they need to calculate only one threshold and use it for the whole image but they are also very limited. For example if the image illumination is not uniform a certain region might have a background greylevel having the same value as the text of another region. Such a case would either cause text being marked as background or the contrary (see for example Figure 4).
- Local algorithms try avoiding this problem using a threshold specifically calculated for each pixel. This is particularly important when a device such as a camera is used since controlling the illumination or the noise is difficult.

Niblack's algorithm [24] uses the mean and standard deviation in a predetermined window centered on the processed pixel to calculate its threshold. Assuming the illumination is slowly changing in the document and so have a lower frequency spectrum than the character one, Thillou and Gosselin [29, 30] use a wavelet transform, which can be followed by color analysis, to remove the background. Gatos *et al.* [14, 15], Seeger and Dance [28] and He and Zhang [17] do a background surface estimation via a first approximated text/background segmentation and an interpolation.

Some of the local algorithms uses parameters which are intended to be set by the user, or at least by the programmer. These parameters may have an important impact on the result, a better parameter value giving better results, but it may be document-dependant. An important feature of algorithms is the *adaptivity*: the capacity to tune their parameters by themselves so the user does not need to set them. This is done, for example, by Yang and Yan [32] add adaptivity to the Kamel and Zhao [19] thresholding technique or by Bartolo *et al* [5] for the Brensen's algorithm.

Noise reduction can either be done via the binarization procedure or in a standalone step, for example using a Wiener filter [18].

We use the algorithm described by Sauvola and Pietikäinen [27] to do the binarization. It provides an amelioration to Niblack's one [24] by amplifying the contribution of the standard deviation in an adaptive way. The results contains less noise than with Niblack's algorithm. Figure 4 shows an example of result.

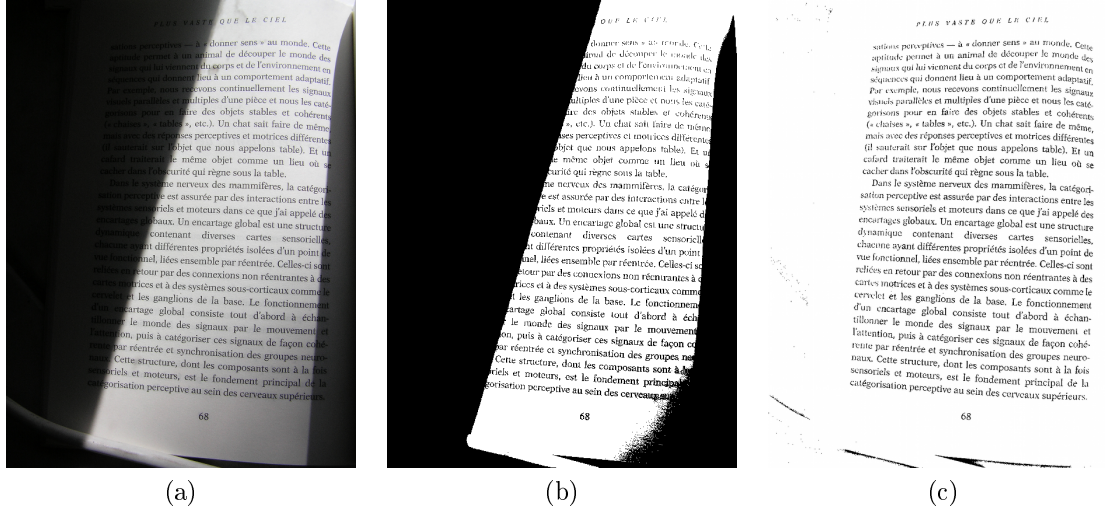


Figure 4: Results of the binarization step. (a) Input image with uneven illumination. (b) Simple thresholding (no optimal threshold can be found). (c) Local thresholding via Sauvola and Pietikäinen's algorithm [27].

### 3.2 Skew detection

Most of layout analysis and OCR algorithms are very sensitive to skew. So it is necessary to correct it before proceeding to the following steps.

Some skew algorithms depends on the document structure which can be simple or more complex including multiple columns and images. Figure 3 shows some examples from a simple document to more complex ones.

Let us review some classical approaches:

- *Projections profiles* can be used [26] to calculate the skew angle of simple documents. As they do projection for the whole angle search range, they are very time consuming.
- Gatos *et al.* [13] do a *cross-correlation* of the black pixels across two or more scanlines to detect the image skew.
- The *Hough transform* is a very adequate tool to recognize lines in an image but, due to his computational inefficiency, it is not directly applied to all the pixels in the image [2, 33].
- *Connected component grouping* can be done as a technique by its own [4] or to ameliorate another one [2].

In this article, we chose the Gatos algorithm [13]. From our experiments, we observed that neither projection profiles nor cross-correlation analysis work perfectly for complex

documents. However, the real advantage of cross-correlation algorithms, is that they are very fast. We implemented the Avila and Lins's algorithm [4] which groups similar neighbor connected components in order to form lines. The angle of every line built this ways is then calculated *via* the least squares calculation and incremented to the angles histogram. The best coarse angle wins, and a fine angle with 0.1 degree is searched in the fine angle histogram. This algorithm works on complex documents at any skew angle and is parameterless. Computation time is reported in the table 1, the test machine is a 2.2GHz Athlon64.

Image size	542×849	1712×2288	1712×2288	1712×2288
Components	1140	2822	5760	6430
Time (s)	0.01	0.15	0.14	0.13

Table 1: Computational cost of the Avila and Lins's skew detection algorithm, including connected component analysis.

### 3.3 Layout analysis

Document layout analysis can be done using either a top-down or a bottom-up strategy. Top-down strategies [1, 10, 16, 22] start with the whole document and try to split it recursively at each step. For example we can do some vertical and horizontal projections and split where the biggest whitespace is, stopping the process at a certain threshold space value. Lee and Ryu [22] combine this with a texture analysis.

Bottom-up strategies [3, 8] combines simple elements of the document, generally the connected components, into bigger and bigger groups according to some criteria such as the average element size in a group.

Current strategy is a bottom-up connected component grouping. The connected components are grouped to regions if they are in the same space area and if they are of the same dimensions. A recursive function is then applied to fusion the existing regions until no change occurs.

### 3.4 Character recognition

As we mentioned in the introduction, our system needs an OCR and several commercial products are available. In this work, we have used some free software such as **ocrad** and **gocr**. When the system will be validated, i.e., in term of improvement of speed of reading, we plan to improve that part by using some commercial software, or by combining several of them (as done in [7]) for a better performance.

## 4 Navigation: Choosing what we want to read

To facilitate the navigation, we propose to use the results given by the layout analysis stage (Section 3.3). The patient has a so-called "Navigation image" as shown in Figure 5, where each paragraph is colored randomly, which gives an idea of the structure very quickly.

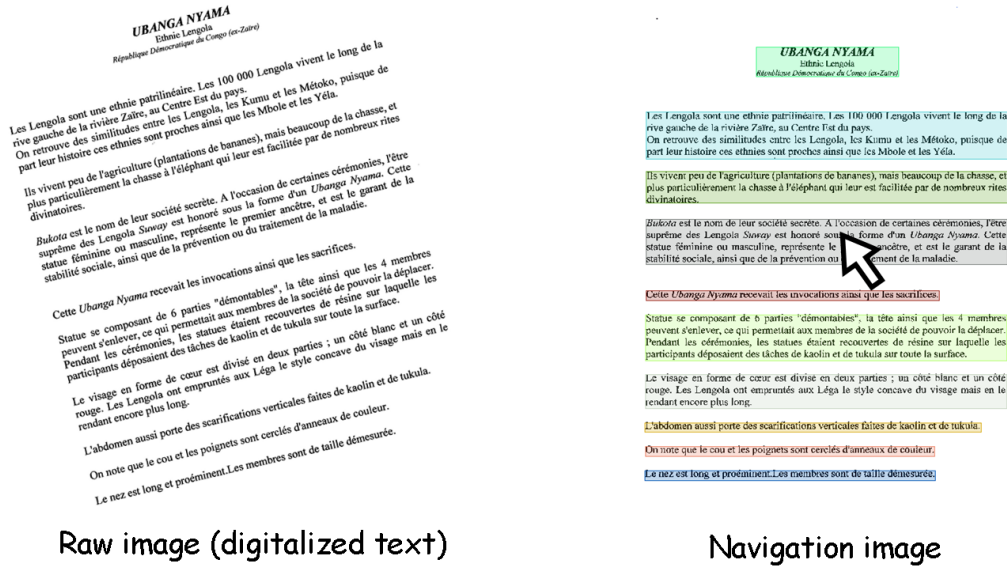


Figure 5: The navigation image is obtained from the raw image after preprocessing detailed in section 3. The structure of the paragraphs is highlighted with different colors. The patient just need to select which paragraph he wants to read.

**Remark** Let us mention the complex case in some newspapers. In fact, complex structures are not easily managed by people with scotomas. They need to find the next paragraph at the end of a column which can be just below the current one, at the right-hand side just next to it or at the right-hand side but on top of the page. In fact if the distance between columns or paragraphs is too small it can even be difficult for people with scotoma to select and read them. Linking columns together requires a semantic analysis of the text itself, in order to know when a column is continued. Such methods exist but they have not been implemented here.

## 5 The gaze-contingent display

Once one paragraph has been selected, we present the characteristics of the text displayed to the patients, where an example is given in Figure 6. In Section 5.1, we discuss the geometry of the text. In Section 5.2 we propose some text enhancements, dependent on the gaze of the patient. To know at every moment the position of the scotoma. We use the **EyeLink II** system to get the eye fixation point on the screen. This section presents various possibilities, which seem reasonable, but whose impact has not been evaluated at the moment. Ongoing psychophysical validations will validate or invalidate the different features.

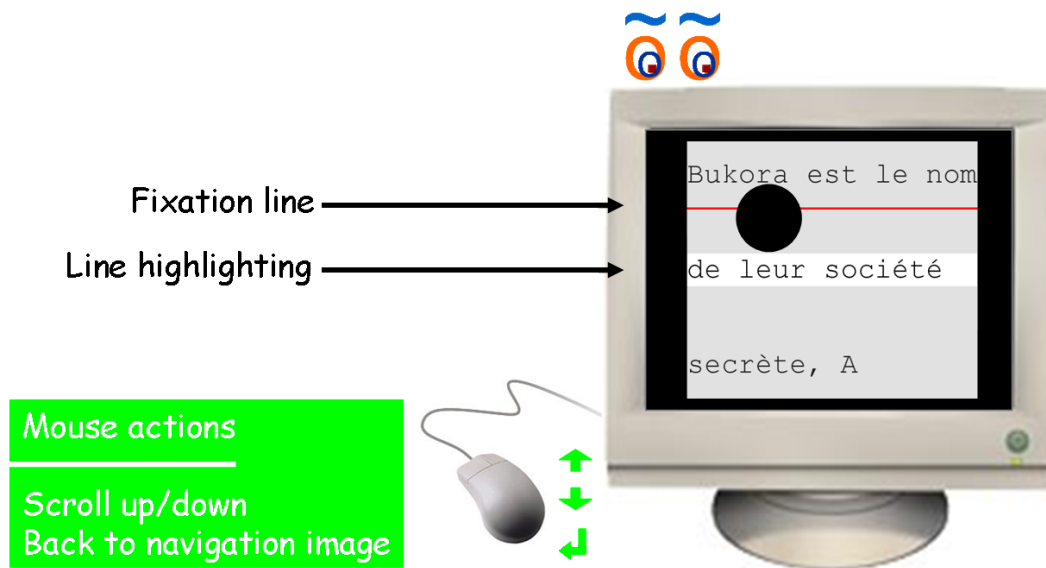


Figure 6: Gaze-contingent display: We show here how the display appears for the patients, with the different enhancements proposed in Section 5. The black circle represents the scotoma of the patient. With the mouse, the patient can scroll up and down the text, or come back the navigation image (see Figure 5 to select another paragraph).

### 5.1 Text characteristics

#### Text size

Experiments show that we can increase the text size until a point from which the zoom is of no utility. This point is called the critical print size (CPS). Critical Print Size. This CPS depends on the size of the scotoma. Using this CPS creates a good compromise between the need for increased size of the text and the need for a sufficiently large viewing area.

## Spacing between lines

In some printed documents, particularly in newspapers, the text is very compact in order to save some space. This is a real issue for patients with scotoma because, if a line is too close to another one, it can induce the patient to commit mistakes such as mixing the two lines when reading. One simple solution is to add more space between the lines. The effect of interline spacing is currently being evaluated by Éric Castet and Jean-Baptiste Bernard at [Institut des Neurosciences cognitives de la Méditerranée \(INCM\)](#).

## 5.2 Features of the gaze-contingent display

### Line highlighting

To help the user focus on the line he intends to read, we propose to increase the relative line contrast. To do so, we diminish the contrast of the remaining of the text, except of the current line.

### Fixation line

Patients with a scotoma cannot use their entire fovea to extract information from a text. They must use perifoveal zones of the retina. Therefore, their gaze and their scotoma come at the periphery of the location they want to analyze. Attention is then directed to “interesting” location.

Patients usually adopt (with external help or not) precise nonfoveal retinal regions for this analyze of “interesting location”. These retinal regions, relative to scotoma, are termed the *preferred retinal locus* or PRL [11, ?]. For example, some patients use a “left PRL” if they always read by directing attention to the left of the scotoma.

Psychophysical experiments showed [31] that many patients place their scotoma above the line of the text in order to direct their attention under the gaze position.

In our program, we therefore draw a horizontal line which is located above the line of text the reader is trying to read. This horizontal line should help the user by inducing saccades directed above the line of text, so that attention is focused on the text itself. It remains to be tested whether the system developed here is able to improve reading performance. If this is the case, we can imagine that this method of “up-reading” could be an interesting way of reading for patients with scotomas.

### Text deformations

The idea is to unmask in real-time the part of the text which is hidden by the scotoma.

One important point of these deformations is that they should not destroy the horizontal spatial continuity of the text. For example, we don’t want a letter to be splitted in two portions, neither to split a word randomly.

We know the importance of peripheral vision for normal readers [23]. With these deformations, we assume that patients with scotoma keep the same use of perifoveal vision as in reading without scotoma (fovea exactly on the text).

Because attention is often directed to the bottom of the fixation, we can deform portions of the text under the scotoma.

The primary kind of text deformation is *text unmasking*: shifting portions of the displayed text when the scotoma is over them to make them visible to the user. Deformations can also be more complex along a curve, which represent the scotoma shape (see for example Figure 7). Zooming or putting in bold font can be also used to emphasize the part of the text which is deformed.

The efficiency of these unmasking techniques for reading with scotoma will be soon evaluated by Eric Castet et allet@tokeneonedotat Institut des Neurosciences Cognitives de la Méditerranée (INCM).



Figure 7: Example of gaze contingent displays with text deformations

## 6 Conclusion

In this report we propose a visual help system for patient with central scotoma.

Our system uses new ideas inspired from psychophysical results and can be used for testing visual parameters in psychophysical experiments, for enhancing the reading speed of patient and for visually teaching patients ways of reading with a scotoma.

Testing is our current focus. Targeting real-time for all the processing will be our next step.

## References

- [1] O.T. Akindele and A. Belaïd. Page segmentation by segment tracing. In *Document analysis and recognition*, pages 341–344, 1993.



- [2] A. Amin and S. Fischer. A document skew detection method using the hough transform. *Pattern Analysis & Applications*, 3:243–253, 2000.
- [3] A. Antonacopoulos, B. Gatos, and D. Bridson. Icdar2005: Page segmentation competition. In *Document Analysis and Recognition*, pages 75–79, 2005.
- [4] B.T. Avila and Lins R. D. A fast orientation and skew detection algorithm for monochromatic document images. In *Proceedings of the 2005 ACM symposium on Document*, 2005.
- [5] A. Bartolo, K. Camilleri, J. Borg, and P. Farrugia. Adaptation of brensen’s thresholding algorithm for sketched line drawings. In *EUROGRAPHICS Workshop on Sketch-Based Interfaces and Modeling*, pages 81–90, 2004.
- [6] M.S. Brown and W.B. Seales. Image restoration of arbitrarily warped documents. *Patterns Analysis and Machine Intelligence*, pages 1295–1306, 2004.
- [7] H. Cecotti and A. Belaïd. Hybrid ocr combination approach complemented by a specialized icr applied on ancient documents. In *Document Analysis and Recognition*, volume 2, pages 1045–1049, 2005.
- [8] M. Chen, X. Ding, and J. Liang. Analysis, understanding and representation of chinese newspaper with complex layout. In *Internation conference on image processing*, volume 2, pages 590–593, 2000.
- [9] K.B. Chua, L. Zhang, Y. Zhang, and C.L. Tan. A fast and stable approach for restoration of warped document images. In *Document analysis and recognition*, volume 1, pages 384–388, 2005.
- [10] L. Cinque, S. Levialdi, A. Malizia, and F. De Rosa. Dan: An automatic segmentation and classification engine for paper documents. In *Proceedings of the Fifth IAPR International workshop on document analysis systems*, pages 491–502, 2002.
- [11] RW. Cummings, SG. Whittaker, GR. Watson, and JM. Budd. Scanning characters and reading with a central scotoma. *Am J Optom Physiol Opt*, 62:833–843, 1985.
- [12] G. Dagnelie. *Virtual technologies aid in restoring sight to the blind*, chapter 15, pages 247–271. IOS, 2001.
- [13] B. Gatos, N. Papamarkos, and C. Chamzas. Skew determination and text line position in digitized documents. *Pattern Recognition*, 30:1505–1519, 1997.
- [14] B. Gatos, I. Pratikakis, and S.J. Perantonis. An adaptive binarization technique for low quality historical documents. In *Lecture Notes in Computer Science*, volume 3163, pages 102–113, 2004.
- [15] B. Gatos, I. Pratikakis, and S.J. Perantonis. Adaptive degraded document image binarization. *Pattern Recognition*, 39:317–327, 2005.

- [16] J. Ha, R.M. Haralick, and I.T. Phillips. Recursive x-y cut using bounding boxes of connected components. In *Proceedings of the Third International Conference on Document Analysis and Recognition*, volume 2, pages 952–955, 1995.
- [17] L.W. He and Z. Zhang. Real-time whiteboard capture and processing using a video camera for teleconferencing. In *Acoustics, Speech and Signal Processing*, volume 2, pages II/1113–II/1116, 2005.
- [18] A.K. Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall, 1989.
- [19] M. Kame and A. Zhao. Extraction of binary character/graphics images from grayscale document images. In *CVGIP: Graphics Models Image Process.*, volume 55, pages 203–283, 1993.
- [20] J.N. Kapur, S.Y. Sahoo, and A.K.C. Wong. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision Graphics Image Process.*, pages 273–285, 1985.
- [21] E. Kavallieratou. A binarization algorithm specialized on document images and photos. In *Document analysis and recognition*, 2005.
- [22] S.-W. Lee and D.-S. Ryu. Parameter-free geometric document layout analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 23(11), 2001.
- [23] G.W. McConkie and K. Rayner. The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics*, 17:578–586, 1975.
- [24] W. Niblack. *An Introduction to Image Processing*, pages 115–116. Prentice-Hall, 1986.
- [25] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. Systems Man Cybernet*, pages 62–66, 1979.
- [26] W. Postl. Detection of linear oblique structures and skew scan in digitized documents. In *Proceedings of the 8th International Conference on Pattern Recognition*, pages 687–689, 1986.
- [27] J. Sauvola and M. Pietikäinen. Adaptive document image binarization. *Pattern Recognition*, 33:225–236, 2000.
- [28] M. Seeger and C. Dance. Binarizing camera images for ocr. In *Document Analysis and Recognition*, pages 54–58, 2001.
- [29] C. Thillou and B. Gosselin. Robust thresholding based on wavelets and thinning algorithms for degraded camera images. In *Proceedings of ACIVS*, 2004.
- [30] C. Thillou and B. Gosselin. Segmentation-based binarization for color degraded images. In *Proceedings of ICCVG*, 2004.

- [31] M. Varsori, A. Perez-Fornos, AB. Safran, and AR. Whatham. Development of a viewing strategy during adaptation to an artificial central scotoma. *Vision Research*, 44(23):2691–705, oct 2004.
- [32] Y. Yang and H. Yan. An adaptive logical method for binarization of degraded document images. *Pattern Recognition*, 33:787–807, 2000.
- [33] B. Yu and A.K. Jain. A robust and fast skew detection algorithm for generic documents. *Pattern Recognition*, 29:1599–1629, 1996.



---

Unité de recherche INRIA Sophia Antipolis  
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes  
4, rue Jacques Monod - 91893 ORSAY Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique  
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

---

Éditeur  
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)  
<http://www.inria.fr>  
ISSN 0249-0803