



**HAL**  
open science

## Text enhancement by PDE's based methods

Zouhir Mahani, Jalal Zahid, Sahar Saoud, Mohammed El Rhabi, Abdelilah  
Hakim

► **To cite this version:**

Zouhir Mahani, Jalal Zahid, Sahar Saoud, Mohammed El Rhabi, Abdelilah Hakim. Text enhancement by PDE's based methods. the 5th International Conference on Image and Signal Processing (ICISP 2012), Jun 2012, Agadir, Morocco. 10.1007/978-3-642-31254-0\_8 . hal-01812659

**HAL Id: hal-01812659**

**<https://enpc.hal.science/hal-01812659>**

Submitted on 11 Jun 2018

**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.

# Text enhancement by PDE's based methods

Zouhir Mahani<sup>1</sup> Jalal Zahid<sup>1</sup> Sahar Saoud<sup>1</sup> Mohammed El Rhabi<sup>2</sup> and Abdelilah Hakim<sup>2</sup>

<sup>1</sup> Université Ibn Zohr, ESTA, Laboratoire Matriaoux, Systèmes et Technologies de l'information,  
B.P: 33/S, 80000 Agadir - Maroc,  
z.mahani@uiz.ac.ma ; j.zahid@gmail.com ; s.saoud@uiz.ac.ma,  
<http://www.esta.ac.ma/>

<sup>2</sup> Université Cadi Ayyad, Faculté des Sciences et Techniques - Guéliz (FSTG)  
Laboratoire de Mathématiques Appliquées et Informatique  
Bd. Abdelkrim El Khattabi , B.P. 618 Guéliz, 40000 Marrakech- Maroc  
elrhabi@gmail.com ; abdelilah.hakim@gmail.com,  
<http://www.fstg-marrakech.ac.ma/>

**Abstract.** In this work, we propose a new method to enhance text in document-image. Firstly, we introduce a classical model and a way to solve it by means of a non-convex optimization problem. So, a simultaneous estimation of the reflectance and the luminance is obtained when the non uniform illumination (also called luminance) is a smooth function and the reflectance is a function of bounded variation. We give an analyse of this problem and some conditions of existence and unicity. Then, we consider the “log” of the classical model. A new pde's model is proposed. This method is based on the resolution of an original partial differential equation (PDE) estimating the log of the luminance. We assume that the luminance is enough smooth and is the solution of a non classical second order's PDE. Then we deduce the reflectance from the estimated luminance and the acquired image. The effectiveness and the robustness of the proposed process are shown on numerical examples in real-world situation (images acquired from cameraphones). Then, we illustrate the ability of this method to improve an Optical Character Recognition (OCR) in text recognition.

## 1 Introduction

Computer vision based documents recognition could be an interesting way to dematerialize informations to manage clients and company's internal documents, offering enterprise wide fast access to business critical information while enhancing the achitecture in place. Typically the dematerialized document formats are PDF. Here, the problem could be separated as least in two steps: localization and recognition. Computer approach increases the performance of both steps, localization and recognition (see [1]). In the computer vision approaches, low cost cameras (webcam, cameraphone, ...) could introduce some distortions and noise artifacts, see [2] for an overview of document image degradation modeling.

The image restoration images by PDE's based models starts to produce signifiant results as images with the diffusion of documents to remove noise while preserving important information for readability ([3]) or to separate the back to front ink interferences ([4]). These works have shown that could be improved up to 30% recognition rate of OCR by restoring the document images.

Generally speaking, an image  $u$  could be seen as the product of a reflectance  $v$  and the illumination effect  $I$  (see [5]). This model has been used in [6,7] to restore blurred barcode signals under nonuniform illumination.

This paper provide a new method based on an anisotropic diffusion to estimate the luminance (or the log of the luminance) of a document-image in order to enhance the text. This method could be use for example in a mobile phone scanning solution or as a first step OCR processing.

We develop previously an isotropic method with some relative success(see [8]). This process allows us to estimate the luminance and to enhance texts then to share or print document images. By document images, we mean all kinds of images which include some handwritten or machine printed text. The basic camera phones often produce images of lower qualities. In addition, the conditions make it difficult shooting the readability and printing of document images. Regarding the text, these conditions make it difficult to interpret the information contained in this image without proper pretreatment. We have developped this method on a server side project, named “Qipit” : Qipit is a way to capture and share written documents with your camera phone or digital camera. Handwritten notes, signed contracts, whiteboards can be transformed into clean, crisp digital copies. After a specific treatment, this method has also directly embedded on photophones (Nokia, Samsung, Iphone ...). As an example, on a Nokia N73 (3.15 megapixels, Symbian OS 9.1, S60 3rd edition), the time processing is less than 6 secondes (it depends on the choosen resolution). So, we have built an OEM business with the likes of Samsung, Sanyo and others with over 100 million copies sold and shipped factory-installed by mobile

phone manufacturers to date. In this paper, since we present a preliminary result, the time processing is not discussed.

Our methods can produce three kinds of documents: black and white, grayscale or color ones. In all cases, input images are color pictures. Producing color documents should be seen as a different workflow, involving mostly the same algorithms as in the grayscale processing.

Document images are supposed to be obtained from a mobile device - a cameraphone more exactly, but could also come from any digital camera. In the following, we will only suppose the device to be a cameraphone as this is the case where most of the problems occur. Basic cameraphones often produce document images with poor quality - meaning it should be really tedious or even impossible to be read by someone or by an OCR (Optical Character Recognition). Because of constraints design, cameraphones have some limitations which we have to take into account, if we want to produce legible documents.

The paper is organized as follows : section 2 gives a brief description of a camera lens. Section 3 presents the non uniform illumination problem and defines a global model. Then two methods are given as solutions, namely a natural criterion afterward a pde's based model. Section 4 describes the method and shows how to implement it. Finally, some numerical results illustrate this work in Section 5.

## 2 Description

Photographers control the camera and lens to expose the recording material (digital sensor or film) to the required amount of light. The controls include the focus of lens, the aperture of the lens (amount of light allowed to pass through the lens), the focal length and type of lens (macro, wide angle, or zoom), the duration of exposure (or shutter speed), the sensitivity of the medium to light intensity and color/wavelength, the nature of the light recording material, for example its resolution as measured in pixels or grains of silver halide.

Camera controls are inter-related, as the total amount of light reaching the image plane (the "exposure") changes proportionately with the *duration of exposure*, *aperture* of the lens, and *focal length* of the lens (which changes as the lens is focused, or zoomed). Changing any of these controls alter the exposure. Many cameras automatically adjust the aperture of the lens to account for changes in focus, and some will accommodate changes in zoom as well.

The duration of an exposure is referred to as *shutter speed*, often even in cameras that don't have a physical shutter, and is typically measured in fractions of a second. *Aperture* is expressed by an *f-number* or *f-stop* (derived from focal ratio), which is proportional to the ratio of the focal length to the diameter of the *aperture*.

Exposures can be achieved through widely differing combinations of shutter speed and aperture. The chosen combination has an impact on the final result. In addition to the subject or camera movement that might vary depending on the shutter speed, the aperture (and focal length of the lens) determine the *depth of field*, which refers to the range of distances from the lens that will be in focus.

## 3 The model

We suppose that the document-image  $u$  has been acquired by a cameraphone, we consider two sorts of problems due to this acquisition : image distortions and noise. Here, we adress a distortion, so-called "non uniform illumination" or variations of brihtness (see [9]). These variations render image processing difficults. One of these effects is : the shadows. In image processing, a shadow is considered as a region with low lightness and high gradients contours. So, we could separate two kinds of shadows, the own shadow and the shadows due to the acquisition. Own shadow occurs when the light hits a surface with a slope change. The brightness of pixels corresponding to the area decreases as the angle of incidence deviates from the normal of the surface. The brightness reaches its minimum when the incident light and the surface normal are orthogonal. Drop shadow occurs when the light source is obscured by an object before the light reflection on the surface. Others distortions problems like blur or warping are beyond of the scope of this paper.

In this paper, we consider an image  $u \in \mathbb{R}^{N_1 \times N_2}$  where  $N_1$  et  $N_2$  are two intergers. The non uniform illumination can be modeled as a multiplicative effect. This modelling combining the reflectance and the luminance of the image was proposed by Barrow and Tenenbaum in 1978 [10]. That said, due to various factors that may be involved in the construction of the image (the illumination of the object, the geometry of the scene acquired, the camera settings ...), such modelling is very difficult to tackle.

In 1999, Laszlo [11] has proposed a generative model of the image, based on a combinaison of Fredholm integral and a modelling of the settings of the camera. This model is very difficult to implement. Thus, it is the global illumination method [11] remains the most widely used :

$$u(x, y) = I(x, y) v(x, y) \cos_{\theta}(x, y) + b(x, y), \quad (1)$$

$u(x, y)$  is the grayscale of the pixel  $(x, y)$ ,  $I$  is the luminance or the non uniform illumination,  $v(x, y)$  the reflectance and  $\cos_{\theta}(x, y)$  the cosine of the angle between the incident light ray and the surface normal at the point of the object. In image processing, this modelling is even more simplified by integrating  $\cos_{\theta}(x, y)$  on the component  $I(x, y)$ . Thus, the final model becomes:

$$u(x, y) = I(x, y) v(x, y) + b(x, y). \quad (2)$$

This modelling of the image is far from perfect, because it does not take into account neither the problems of geometry of the object (the presence of surfaces which can create shadows on the object ...), or external factors in the formation of the image. The advantage of this simple model is to estimate the reflectance of an object from an approximation of its luminance.

The estimation of the reflectance  $v$  is crucial, because here we have an opportunity to characterize an object independently of illumination problems. Here, we suppose that the noise  $b$  is a gaussian random variable.

In the following, we suppose  $u, I$  and  $v \in L^2(\Omega)$  (where  $L^2(\Omega)$  is the square intergrable space,  $\Omega$  is the domain of the image) and the distortion is introduced by the measuring device and the conditions of acquisition.

### 3.1 Natural criterion

In this part, we consider the model defined by the equation 1. Thus and ideally, we would like to estimate both  $I$  and  $v$  up to the noise, a natural criterion could be:

$$J_1(I, v) = \frac{1}{2} \|Iv - u\|_{L^2(\Omega)}^2 + \lambda_1 R_1(v) + \lambda_2 R_2(I), \quad (3)$$

where  $\lambda_1, \lambda_2$  are two positive hyperparameters,  $R_1, R_2$  are two regularization functions allowing to control the noise and to provide some *a priori* on the pair of solution.

If we suppose that  $I$  is a smooth function and that the singularities are reported on  $v$ . A classical choice for the regularization function is :

$$R_1(v) = \int_{\Omega} |\nabla v| d\Omega, \quad R_2(I) = \int_{\Omega} |\nabla I|^2 d\Omega, \quad (4)$$

in other terms,  $v \in BV(\Omega)$  (see [12] for more details on this space) and  $I \in H^1(\Omega)$ .

Then, we are looking for:

$$(I, v) = \arg \min_{I \in H^1(\Omega), v \in BV(\Omega)} J(I, v). \quad (5)$$

**Proposition 1.** *If we suppose that the noise  $b = 0$  and if solutions of (5) exist where  $I \in ]0, 1]$  and  $v$  in  $E = \{u \in L^2(\mathbb{R}^2), u(x, y) = 0 \text{ or } u(x, y) = 1\}$ , let us consider two solutions  $(I_1, v_1)$  and  $(I_2, v_2)$  of (5) then  $v_1 = v_2$  and when  $u_1 = u_2 \neq 0$  then  $I_1 = I_2$ .*

**Proposition 2.** *If we suppose that  $I$  is a positive unknown constant in  $]0, 1]$  there exist a solution  $(I, v)$  satisfying problem (5) in  $E_A = \{v \in E, v(x, y) = 0 \forall x, y \in \mathbb{R}^2 \setminus A\}$  where  $A = [0, 1]^2$ .*

*Proof* —Under the assumption of the proposition, let  $I_j$  and  $v_j$  be a minimizing sequence of the problem (5), the sequence  $(v_j)_j$  belongs to  $E_A$  and has bounded variations. Then by a classical compactness result (see [13]) for the functions whith bounded variation and the fact that the all element of the sequence  $(I_j)_j$  belongs to a compact  $K \subset ]0, 1]$ . Then, we can extract a convergent subsequence, this subsequence converges toward  $(I_{\infty}, v_{\infty})$  with  $I_{\infty} \in K$ . As  $E_A$  is a closed subset of  $L^2(\mathbb{R}^2)$ , we have  $v_{\infty} \in E_A$ .

Following these propositions, it could be interessant to add a constraint to project the solution  $v$  as closed as possible to  $E$ . A natural candidate function could be :

$$g : [m, M] \rightarrow \mathbb{R}^+ \\ x \mapsto \frac{(M-x)(x-m)}{M-m}$$

where  $m$  is the minimum value of the acquired image  $u$  and  $M$  is its maximum value.

So, we could penalized the criterion defined in 5 as follow:

$$J_2(I, v) = \frac{1}{2} \|Iv - u\|_{L^2(\Omega)}^2 + \lambda_1 R_1(v) + \lambda_2 R_2(I) + \lambda_3 G(v), \quad (6)$$

where  $\lambda_3 > 0$  and  $G(v) = \int_{\Omega} |g(v)|^2 d\Omega$ .

Then, the researched solution is:

$$(I, v) = \arg \min_{l \in H^1(\Omega), w \in BV(\Omega)} J_2(l, w). \quad (7)$$

The Euler-Lagrange equations associated to the problem 6 are:

$$\begin{aligned} \frac{\partial J}{\partial v}(I, v) &= I(Iv - u) - \lambda_1 \operatorname{div} \left( \frac{\nabla v}{|\nabla v|} \right) + \lambda_3 (M + m - 2v) \frac{(M-v)(v-m)}{(M-m)^2} = 0, \\ \frac{\partial J}{\partial I}(I, v) &= v(Iv - u) - \lambda_2 (\Delta I) = 0. \end{aligned} \quad (8)$$

### 3.2 A direct PDE based method

The objective of this approach is to extract the reflectance  $v$  from the non uniform illumination  $I$ . If we neglect for the moment the noise.

The model 2 becomes:

$$\begin{aligned} u(x, y) &= I(x, y) v(x, y), \\ \log(u(x, y)) &= \log(I(x, y) v(x, y)), \\ \log(u(x, y)) &= \log(I(x, y)) + \log(v(x, y)), \end{aligned} \quad (9)$$

**Remark 1** *The grayscale image could be shifted by an offset to avoid the null values.*

We are thus reduced to find  $\tilde{I}$  and  $\tilde{v}$  satisfying:

$$\tilde{u} = \tilde{I} + \tilde{v} \quad (10)$$

So, as  $\tilde{I}$  is assumed smooth, we will find a way to approach  $\tilde{I}$  regardless of  $\tilde{v}$ , then deduce  $\tilde{v}$  by substaction with  $\tilde{u}$ .

Here, we suppose that  $\tilde{I}$  is the solution of:

$$\begin{cases} w_t = s \max(0, s\Delta^A w), \\ \frac{\partial w}{\partial n} = 0 \text{ on } \partial\Omega, \\ w(t=0) = u, \end{cases} \quad (11)$$

where  $s = 1$  if the grayscales of the text are on average “smaller” than the background ones and  $s = -1$  else.  $\Delta^A$  is defined as:

$$\Delta^A := \operatorname{div} \left( \varphi'(|\nabla u|) \frac{\nabla u}{|\nabla u|} \right). \quad (12)$$

Now, we introduce the orthonormal set  $(\tau, n)$ , where  $n$  is defined by:

$$n(x) = \frac{\nabla u}{|\nabla u|}.$$

The vector fields  $\tau$  et  $n$  are respectively tangent and normal to the level curves (isocontours) of  $u$ .

Then (12) becomes:

$$\Delta^A = \left( \frac{\varphi'(|\nabla u|)}{|\nabla u|} u_{tt} + \varphi''(|\nabla u|) u_{nn} \right), \quad (13)$$

where :

$$\begin{aligned} u_{tt} &= \tau^t \nabla^2 u \tau = \frac{1}{|\nabla u|^2} (u_{x_1}^2 u_{x_2 x_2} + u_{x_2}^2 u_{x_1 x_1} - 2u_{x_1} u_{x_2} u_{x_1 x_2}) \text{ and} \\ u_{nn} &= n^t \nabla^2 u n = \frac{1}{|\nabla u|^2} (u_{x_1}^2 u_{x_1 x_1} + u_{x_2}^2 u_{x_2 x_2} + 2u_{x_1} u_{x_2} u_{x_1 x_2}). \end{aligned}$$

$u^t$  is the transpose of  $u$ ,  $\nabla^2 u$  is the hessian,  $u_{x_i}$  and  $u_{x_i x_j}$ ,  $i, j \in \{1, 2\}$  designe the first and second order partial derivative of  $u$ . This allows to separate both directions of diffusion:  $\tau$  et  $n$ .

The term  $\frac{\varphi'(|\nabla u|)}{|\nabla u|}u_{tt}$  defines the diffusion in direction  $\tau$ , while  $\varphi''(|\nabla u|)u_{nn}$  defines the diffusion in the direction  $n$ .

Thus, we have some qualitative requirements on the function  $\varphi$ : an isotropic diffusion on homogeneous region (part) of the image,  $\varphi'(0) = 0$ . We have also  $\lim_{t \rightarrow 0} \frac{\varphi'(t)}{t} = \varphi''(0) > 0$ .

On the edge of the image, where we have a high gradient, we prefer a tangential diffusion and not in its transverse direction. So, we demand:  $\lim_{t \rightarrow +\infty} \frac{t\varphi''(t)}{\varphi'(t)} = 0$ .

An example of function satisfying these requirement is:

$$\varphi(t) = \sqrt{1 + t^2}, \quad (14)$$

If  $\tilde{I}$  is estimated then we can deduce  $\tilde{v}$  :

$$\tilde{v} = \frac{(1+s)}{2} \max(\tilde{u}) - s \left| \tilde{I} - \tilde{u} \right|, \quad (15)$$

## 4 Discretization

In the discrete form, an image is composed of a set of pixels indexed by  $(i, j)$ ,  $1 \leq i \leq N$ ,  $1 \leq j \leq M$ .  $u = (u_{i,j})_{1 \leq i \leq N, 1 \leq j \leq M}$  belongs in  $X$ , where  $X = \mathbb{R}^{N \times M}$ . The space  $X$  is equipped with the euclidian inner scalar product:

$$\forall u, v \in X, \langle u, v \rangle_X = \sum_{i=1}^N \sum_{j=1}^M u_{i,j} v_{i,j}.$$

By a minor abuse of the notation, we state for  $X^m$ , where  $m \geq 1$ , the space  $(\mathbb{R}^m)^{N \times M}$ . The gradient of  $u \in X$ , written  $\nabla u$  belongs to  $X^2$  and could be defined by several manners. One of them consists to set  $\nabla u = (g^{(1)}, g^{(2)})$  with:

$$g_{i,j}^{(1)} = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N, \\ 0 & \text{si } i = N. \end{cases} \quad g_{i,j}^{(2)} = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } j < M, \\ 0 & \text{si } j = M. \end{cases} \quad (16)$$

This div operator is defined in  $X^2$  to  $X$  as the adjoint operator of  $-\nabla$ . So, for all  $p = (p^{(1)}, p^{(2)}) \in X^2$ , we have:

$$\forall z \in X, \langle \text{div} p, z \rangle = -\langle p, \nabla z \rangle.$$

In the case where the gradient is given by (16), one can prove that

$$(\text{div} p)_{i,j} = (\text{div} p)_{i,j}^{(1)} + (\text{div} p)_{i,j}^{(2)}, \quad (17)$$

with

$$(\text{div} p)_{i,j}^{(1)} = \begin{cases} p_{i,j}^{(1)} - p_{i-1,j}^{(1)} & \text{if } 1 < i < N, \\ p_{i,j}^{(1)} & \text{if } i = 1, \\ -p_{i-1,j}^{(1)} & \text{if } i = N. \end{cases}$$

$$(\text{div} p)_{i,j}^{(2)} = \begin{cases} p_{i,j}^{(2)} - p_{i,j-1}^{(2)} & \text{if } 1 < j < M, \\ p_{i,j}^{(2)} & \text{if } j = 1, \\ -p_{i,j-1}^{(2)} & \text{if } j = M. \end{cases}$$

We state for all  $u \in X$ ,

$$\Delta u = \text{div}(\nabla u). \quad (18)$$

and

$$\Delta^A u = \text{div} \left( \varphi'(|\nabla u|) \frac{\nabla u}{|\nabla u|} \right) \quad (19)$$

Then, from the definition of the divergence, we keep in this discretete form:

$$\forall u, v \in X, \langle \Delta u, v \rangle = -\langle \nabla u, \nabla v \rangle = \langle u, \Delta v \rangle. \quad (20)$$

Then, it follows the natural algorithm from 7, 17 and 18:

**Data:**  $u$  the acquired image  
**Result:**  $I$  the non uniform illumination and  $v$  the reflectance  
**initialization :** Given  $\varepsilon_1 > 0, \varepsilon_2 > 0, \lambda_1, \lambda_2$  and  $\lambda_3 > 0, I^0 = \frac{u}{\max(u)}, v^0 = u$  and  $\mu > 0$  adequately chosen  
**do**  
 • update  $I$  and  $v$ :  

$$v^{p+1} = v^p - \mu \left( I^p (I^p v^p - u) - \lambda_1 \operatorname{div} \left( \frac{\nabla v^p}{|\nabla v^p|} \right) + \lambda_3 (M + m - 2v^p) \frac{(M - v^p)(v^p - m)}{(M - m)^2} \right)$$

$$I^{p+1} = I^p - \mu (v^{p+1} (I^p v^{p+1} - u) - \lambda_2 \Delta I^p)$$
**until**  $\|v^{p+1} - v^p\| < \varepsilon_1$  &  $\|I^{p+1} - I^p\| < \varepsilon_2$   
 $v = v^{p+1}, I = I^{p+1}$

**Algorithm 1:** Model 7

Now, if we use an explicit scheme in time and the discretization 19 for the anisotrope laplacian, we can derive the following algorithm from 11:

**Data:**  $u$  the acquired image,  $s = 1$  or  $-1$   
**Result:**  $v$  the reflectance  
**initialization :** Given  $\varepsilon > 0, \tilde{I}^0 = \tilde{u} := \log(u)$  and  $dt > 0$  adequately chosen  
**do**  
 • update  $I$  :  

$$\tilde{I}^{p+1} = \tilde{I}^p + dt \max(0, \Delta^A I^p)$$
**until**  $\|\tilde{I}^{p+1} - \tilde{I}^p\| < \varepsilon$   
 $\tilde{I} = \tilde{I}^{p+1}$  deduce  $\tilde{v}$  from 15 and  $v = \exp(\tilde{v})$ .

**Algorithm 2:** Model 11

## 5 Numerical results

In this section, we present some simulation results comparing boths models. We show the ability of the proposed algorithms to successfully estimate the reflectance and the non uniform illumination and we compare also their performance by means of a text recognition software. As we don't control the recognition software errors, we present only how our process can improve this software. In other words, we show the recognition of the acquired image afterwards the recognition of the results obtained by the algorithms 1 and 2.

In the example 1, the original image is acquired from a Sony Ericksson K800i (3.2 megapixels) while images in Examples 2 and 3 are obtained from an Iphone 3GS (3 megapixels).

In these tests, the parameters of the algorithm 1- criterion 7 are taken equal to  $\lambda_1 = \lambda_2 = 0.1, \lambda_3 = 0.01$  and  $\varepsilon = 0.00001$ . For algorithm 2- model 11, we take  $s = 1, dt = 0.5$  and  $\varepsilon = 0.00001$ .

Figure 1 shows the estimated non uniform illumination and the estimated reflectance obtained from both algorithms. In this example, The image is affected by its own shadow. The resulting estimations are correct and quite similar.

In figures 2 and 3, the images are acquired in a severe rough environment. These tests are very interesting because the images contain receipts. If we could recognize the contents of the receipts. The consumer can then follow his budget in a detailed way. Then, we can propose an alternative solution to barcode scanning. This solution is more tracktable since we have only one shot and one scan when the barcode solution needs to decode the barcode for each product. Moreover, barcodes acquired from photophones have also distortion problems that we must take into account during the decoding process (see [7,6,14]).

In figure 2, the lightning conditions are very low and we deal with distortions which are beyond the scope of this work (blur and noise). Thus, here, we are on a "edge case" for both algoritms. As we can see, the text recognition software recognize some texts of the estimated reflectance for both algorithms whereas this software does not recognize any text from the original image. Both results are correct but the result of algorithm 2- model 11 seems more suitable if we would like to read or to print the result. In figure 3, the image are affected by two kinds of shadows : its own shadow and a drop shadow. In this test, the algorithm 2- model 11 gives the best result.

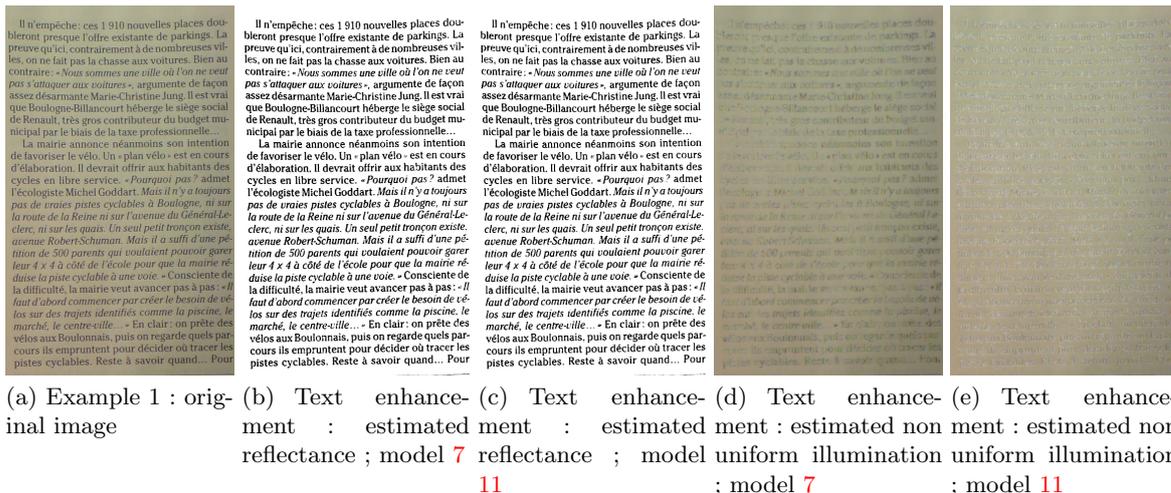


Fig. 1. Original image acquired from a Sony Ericsson K800i

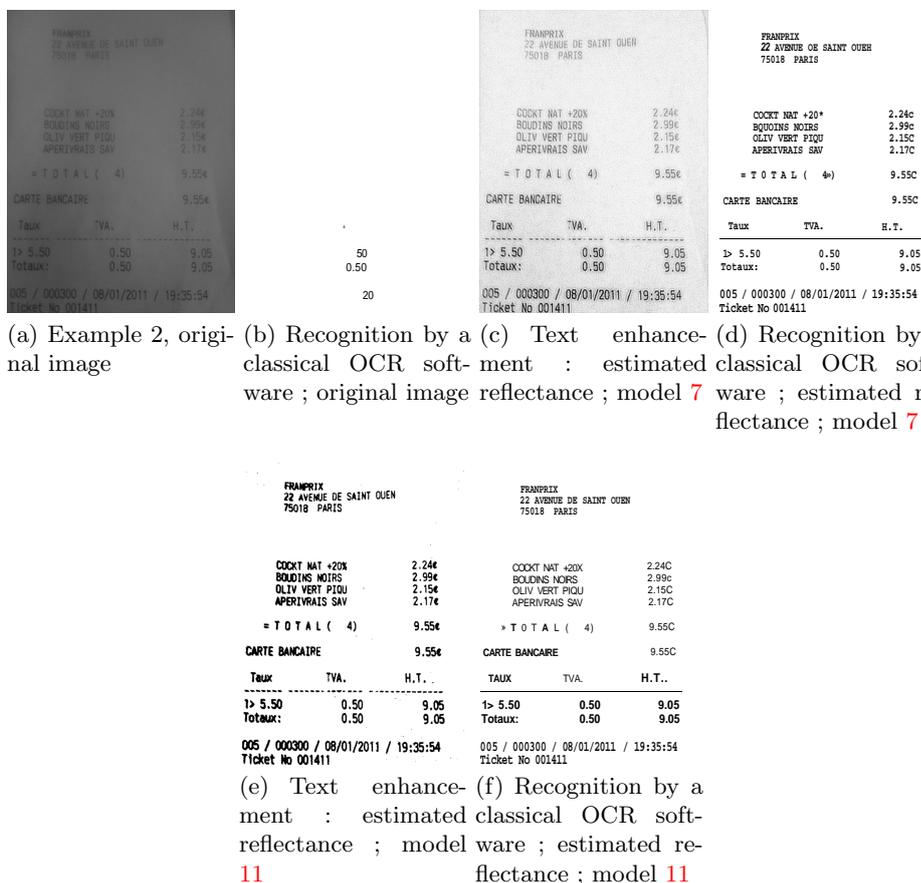


Fig. 2. Original image acquired from an iPhone 3GS: first test

## 6 Conclusion

Two methods for enhancing text in image-documents are proposed. The first is based on an highly non convex optimization problem while the second is based on a direct PDE's resolution. So far we have shown that the methods are robust to low lightning conditions. The second method based on an anisotropic pde's model provides the best result and is easier to implement. Moreover, the resulting algorithm depends only on few fixed parameters. However, this model needs some theoretical investigations and a rigorous analysis (conditions of existence, unicity, numerical analysis of the associated algorithm . . . ) that deserves further investigations. Based



(a) Example original image (b) Recognition by a classical OCR software (c) Text enhancement by a classical OCR software ; reflectance (d) Recognition by a classical OCR software ; reflectance (e) Text enhancement by a classical OCR software ; reflectance ; model 7 (f) Recognition by a classical OCR software ; reflectance ; model 11

**Fig. 3.** Original image acquired from an iphone 3GS: second test

on the resolution of the associated global illumination model 2, both methods allow to estimate the luminance in a robust and reproducible way. In particular, this approach has been successfully applied to improve an OCR software to recognize texts from real images taken from photophones even in tough environment.

## References

1. J. Wu and T. Caelli, *Model Based 3D object localization and recognition from a single intensity image*, In Krzyzak Adam, Kasvand Tony, and Suen Y. Cheng, editors, *Computer Vision and Shape Recognition*, pages 21-67, 1988.
2. Henry S. Baird, *The State of the Art of Document Image Degradation Modeling*, In Proc. of 4 th IAPR International Workshop on Document Analysis Systems, Rio de Janeiro, pp 1-16, 2000.
3. F. Drira, F. Le Bourgeois and H. Emptoz, *Document images restoration by a new tensor based diffusion process : Application to the recognition of old printed documents*, 10th International Conference on Document Analysis and Recognition (ICDAR09), Barcelone, pp. 321-325, 2009.
4. R. F. Moghaddam and M. Cheriet, *RSLDI: Restoration of single-sided low-quality document images*, *Pattern Recognition, Special Issue on Handwriting Recognition*, no. 42, pp. 3355-3364 (2009).
5. B. K. Horn, *Robot Vision*, MIT Press, 1986.
6. J. Kim and H. Lee, *Joint nonuniform illumination estimation and deblurring for bar code signals* *Optic Express*, Vol. 15, issue 22, pp. 14817-14837, 2007.
7. L. Dumas, M. El Rhabi and G. Rochefort, *An evolutionary approach for blind deconvolution of barcode images with nonuniform illumination* IEEE Congress on Evolutionary Computation, pp. 2423 - 2428, 2011.
8. Antoine Martin and Martin Lefebure, Realeyes3D SA, patent, <http://www.prior-ip.com/application/33411968/>.
9. Ralph Gross and Vladimir Brajovic, *An image preprocessing algorithm for illumination invariant face recognition*, In 4th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), Springer, June 2003.
10. H.G. Barrow and J.M. Tenenbaum, *Recovering intrinsic scene characteristics from images*, In CV578, pages 326, 1978.
11. Szirmay-Kalos Laszlo, *Monte-Carlo Global Illumination Methods - State of the Art and New Developments*, SCCG'99, Invited talk 1999.
12. A. Cohen and D. Wolfgang and I. Daubechies and R. DeVore, *Harmonic analysis of the space BV*, *Rev. Mat. Iberoamericana*, 2003, volume 19 (1) pp 235-263.
13. Giusti E., *Minimal Surfaces and Functions of Bounded Total Variation*, Monographs in Mathematics vol 80, Boston, MA: Birkhauser, 1984.
14. M. El Rhabi and G. Rochefort, Realeyes3D SA, patent : <http://www.wipo.int/patentscope/search/en/W02009112710>