

Scan de document-images dans des conditions difficiles d'acquisition

Étude proposée par le département IMI, le CERMICS et la société 6perform.

Encadrants :

Mohammed EL RHABI (département IMI),
Gabriel STOLTZ (CERMICS),
et **Philippe LUTMAN** (société 6perform).

Sommaire

1 But : transformer votre photophone en scanner mobile	1
2 Description	1
3 Le modèle	2
4 Un critère naturel	3
5 Une autre méthode	3
6 Résultats attendus	3

1 But : transformer votre photophone en scanner mobile

Le logiciel développé permettra le scan d'une image de document (manuscript ou imprimé) acquise par un photophone ou une caméra numérique. Les applications sont nombreuses : par exemple l'utilisateur final pourra copier des tableaux (dans cette étude des tableaux blancs), des notes de cours, factures/devis ... Ainsi, le photophone serait aussi un scanner mobile!

L'utilisateur pourra alors stocker le document ou le partager par courriel ou via des réseaux sociaux. Bien évidemment, il existe déjà des logiciels embarqués sur mobile proposant ce genre d'application. Le vôtre devra se distinguer de (ou au moins égal à) l'existant!

Au delà de la qualité du scan, un autre des grands challenges de votre application est sa capacité à répondre en quasi-temps réel.

2 Description

Pour l'utilisation souhaitée, la luminosité, la taille et la nature des documents sont très variables, les images obtenues subissent un certain nombre de distorsion. Ainsi, les conditions de prise de vue rendent difficiles la lisibilité et l'impression de ce genre d'images de document.

Ce projet consiste en l'évaluation de nos derniers travaux concernant le réhaussement du texte dans l'image obtenue à partir de la caméra d'un photophone. Nous nous limiterons à la distorsion due à la luminosité (illumination) variable et aux documents ne contenant que du texte. Ce traitement permet le réhaussement du texte dans une image de document, mais il peut être aussi une aide à la reconnaissance automatique de caractères.

Le chapitre 3 rappelle le modèle. Puis, nous introduirons deux méthodes d'approximation. Nous terminerons par quelques illustrations et résultats numériques attendus.

3 Le modèle

On suppose que l'image de document u a été acquise par la caméra d'un téléphone portable, les problèmes liés à cette acquisition sont (il y en a évidemment d'autres dont on ne tient pas en compte ici) :

1. Une atténuation provenant du fait que la lumière traverse de manière variable l'image, ce qu'on appelle « illumination variable » (voir [1, 2, 3, 5, 24, 7]). Un effet indésirable de l'illumination : les ombres. On distingue deux types d'ombres, les ombres propres et les ombres portées :
 - Une ombre propre se produit lorsque la lumière arrive sur une surface avec un changement de pente. La luminosité des pixels correspondant à cette zone diminue à mesure que l'angle d'incidence s'écarte de la normale de la surface. Cette luminosité atteint son minimum lorsque l'angle entre la lumière incidente et la normale de la surface est égale à 90° .
 - Une ombre portée se produit lorsque la source de lumière est occultée par un objet avant la réflexion de la lumière sur la surface. En traitement d'images, une ombre est considérée comme étant une région à faible luminosité et ayant des contours à fort gradient.
2. Le bruit dû au capteur de mesure (les capteurs de la caméra n'étant pas suffisamment éclairés).

Dans ce document, nous considérerons une image $u \in \mathbb{R}^{N_1 \times N_2}$ où N_1 et N_2 sont deux entiers naturels.

Le modèle, que nous considérons ici, se base sur la séparation des deux composantes de l'image : la luminance I et la réflectance v . La modélisation de l'image par combinaison de la réflectance et de la luminance a été proposée par Barrow et Tenenbaum en 1978 [8]. Ceci dit, en raison des différents facteurs qui peuvent entrer en jeu pour la construction de l'image (l'illumination de l'objet, la géométrie de la scène acquise, les paramètres de la caméra...), l'analyse de cette modélisation reste très difficile à traiter.

En 1999, Laszlo [9] a proposé un modèle génératif de l'image, basé sur la combinaison de l'intégrale de Fredholm et d'une modélisation des paramètres de la caméra. Cette modélisation reste assez complexe à mettre en œuvre. La modélisation par réaction diffuse de l'image reste la plus utilisée :

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

avec $u(x, y)$ est le niveau de gris du pixel (x, y) , avec x et y les coordonnées d'un pixel de l'image. I est l'amplitude de l'illumination (en fait la fameuse « atténuation » - dépendante de la source de lumière), $v(x, y)$ la réflectance (dépendante de la nature intrinsèque de la surface de l'objet) et $\cos\theta(x, y)$ est le cosinus de l'angle entre le rayon lumineux incident et la surface normale au point de l'objet. En intégrant ce dernier dans la luminance I , le modèle devient :

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

Cette modélisation de l'image est loin d'être parfaite, car elle ne prend pas en compte ni les problèmes liés à la géométrie de l'objet (présence de surfaces qui peuvent créer des ombres sur l'objet...), ni les facteurs externes à la formation de l'image. L'avantage de cette modélisation simple est de pouvoir estimer la réflectance d'un objet à partir d'une approximation de sa luminance. L'estimation de la réflectance v est très importante, car nous avons ici une possibilité de caractériser un objet indépendamment des problèmes liés à l'illumination. Le bruit de mesure b est une variable aléatoire que l'on supposera gaussienne.

Dans la suite, on supposera aussi que u, I et $v \in L^2(\Omega)$ (Ω un ouvert borné définissant le domaine de l'image) et que la dégradation est introduite par l'appareil de mesure et les conditions d'acquisition.

4 Un critère naturel

Dans cette partie, nous allons considérer le modèle défini par l'équation 2. Ainsi idéalement, on aimerait estimer à la fois I et v au bruit près par la minimisation d'un critère qui semblerait « naturel » :

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

où λ_1 et λ_2 sont deux paramètres strictement positifs. R_1 et R_2 sont deux fonctions de régularisation.

1. Si l'on suppose que I est « régulière » et v à variations bornées (voir [10]), quelle type de régularisation peut-on proposer ?
2. Ces fonctions définies, proposer un algorithme permettant une estimation du couple (I, v) .
3. Comme on cherche à réhausser un document contenant essentiellement du texte, peut-on en tenir compte dans le critère ?

5 Une autre méthode

L'objectif de cette approche est d'extraire la réflectance à partir de l'illumination variable. Négligeons pour le moment le bruit.

Le modèle 2 : $u(x, y) = I(x, y)v(x, y)$. En utilisant la fonction logarithme, le modèle multiplicatif devient additif :

$$\log(u(x, y)) = \log(I(x, y)) + \log(v(x, y)).$$

Remarque L'image u peut être éventuellement décalée d'un off-set pour éviter les valeurs nulles.

Si on pose $\tilde{u} = \log(u)$, on est donc ramené à chercher $\tilde{I} = \log(I)$ et $\tilde{v} = \log(v)$ vérifiant : $\tilde{u} = \tilde{I} + \tilde{v}$.

Comme précédemment \tilde{I} est supposé suffisamment régulière, nous allons chercher un moyen de l'approcher indépendamment de \tilde{v} , puis déduire \tilde{v} par soustraction.

Si le niveau de gris du texte est plus « sombre » que le niveau de gris du fond (qu'est ce que cela veut dire ? et pourquoi cette hypothèse ?), pour $T > 0$, on peut supposer que \tilde{I} est la solution stable de l'équation :

$$\begin{cases} w_t(x, y, t) &= [\Delta w(x, y, t)]^+, \text{ dans } \Omega \times (0, T), \\ \frac{\partial w}{\partial n}(x, y, t) &= 0, \text{ sur } \partial\Omega \times (0, T), \\ w(x, y, 0) &= \tilde{u}(x, y), \text{ dans } \Omega \end{cases} \quad (4)$$

1. Les questions d'existence et d'unicité ne seront pas abordées ici.
2. Proposer un schéma numérique pour le calcul de la solution.
3. La dépendance par rapport aux paramètres discrétisations ?
4. Le temps de calcul ?
5. Facultatif : la stabilité ? convergence ?

6 Résultats attendus

Il n'empêche: ces 1 910 nouvelles places doubleront presque l'offre existante de parkings. La preuve qu'ici, contrairement à de nombreuses villes, on ne fait pas la chasse aux voitures. Bien au contraire: « Nous sommes une ville où l'on ne veut pas s'attaquer aux voitures », argumente de façon assez désarmante Marie-Christine Jung. Il est vrai que Boulogne-Billancourt héberge le siège social de Renault, très gros contributeur du budget municipal par le biais de la taxe professionnelle...

La mairie annonce néanmoins son intention de favoriser le vélo. Un « plan vélo » est en cours d'élaboration. Il devrait offrir aux habitants des cycles en libre service. « Pourquoi pas ? admet l'écologiste Michel Goddart. Mais il n'y a toujours pas de vraies pistes cyclables à Boulogne, ni sur la route de la Reine ni sur l'avenue du Général-Leclerc, ni sur les quais. Un seul petit tronçon existe, avenue Robert-Schuman. Mais il a suffi d'une pétition de 500 parents qui voulaient pouvoir garer leur 4 x 4 à côté de l'école pour que la mairie réduise la piste cyclable à une voie. » Consciente de la difficulté, la mairie veut avancer pas à pas: « Il faut d'abord commencer par créer le besoin de vélos sur des trajets identifiés comme la piscine, le marché, le centre-ville... » En clair: on prête des vélos aux Boulonnais, puis on regarde quels parcours ils empruntent pour décider où tracer les pistes cyclables. Reste à savoir quand... Pour

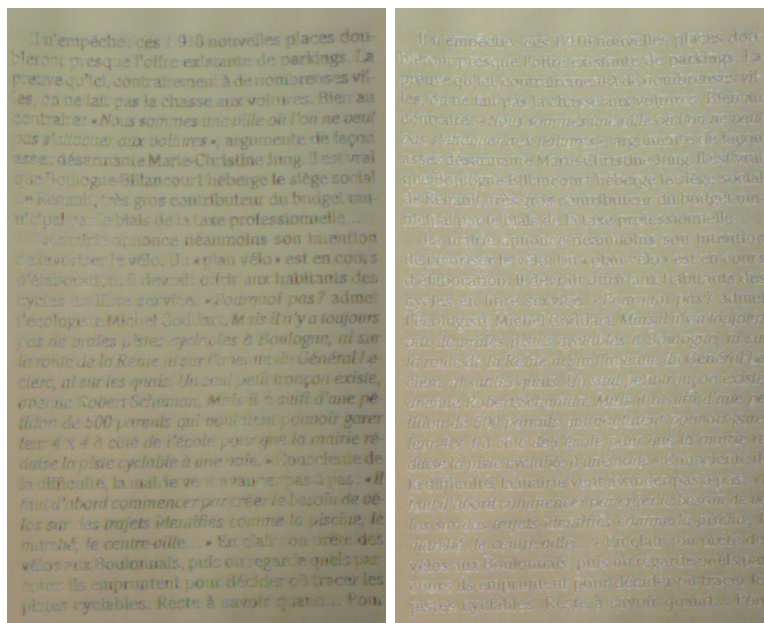
Il n'empêche: ces 1 910 nouvelles places doubleront presque l'offre existante de parkings. La preuve qu'ici, contrairement à de nombreuses villes, on ne fait pas la chasse aux voitures. Bien au contraire: « Nous sommes une ville où l'on ne veut pas s'attaquer aux voitures », argumente de façon assez désarmante Marie-Christine Jung. Il est vrai que Boulogne-Billancourt héberge le siège social de Renault, très gros contributeur du budget municipal par le biais de la taxe professionnelle...

La mairie annonce néanmoins son intention de favoriser le vélo. Un « plan vélo » est en cours d'élaboration. Il devrait offrir aux habitants des cycles en libre service. « Pourquoi pas ? admet l'écologiste Michel Goddart. Mais il n'y a toujours pas de vraies pistes cyclables à Boulogne, ni sur la route de la Reine ni sur l'avenue du Général-Leclerc, ni sur les quais. Un seul petit tronçon existe, avenue Robert-Schuman. Mais il a suffi d'une pétition de 500 parents qui voulaient pouvoir garer leur 4 x 4 à côté de l'école pour que la mairie réduise la piste cyclable à une voie. » Consciente de la difficulté, la mairie veut avancer pas à pas: « Il faut d'abord commencer par créer le besoin de vélos sur des trajets identifiés comme la piscine, le marché, le centre-ville... » En clair: on prête des vélos aux Boulonnais, puis on regarde quels parcours ils empruntent pour décider où tracer les pistes cyclables. Reste à savoir quand... Pour

Il n'empêche: ces 1 910 nouvelles places doubleront presque l'offre existante de parkings. La preuve qu'ici, contrairement à de nombreuses villes, on ne fait pas la chasse aux voitures. Bien au contraire: « Nous sommes une ville où l'on ne veut pas s'attaquer aux voitures », argumente de façon assez désarmante Marie-Christine Jung. Il est vrai que Boulogne-Billancourt héberge le siège social de Renault, très gros contributeur du budget municipal par le biais de la taxe professionnelle...

La mairie annonce néanmoins son intention de favoriser le vélo. Un « plan vélo » est en cours d'élaboration. Il devrait offrir aux habitants des cycles en libre service. « Pourquoi pas ? admet l'écologiste Michel Goddart. Mais il n'y a toujours pas de vraies pistes cyclables à Boulogne, ni sur la route de la Reine ni sur l'avenue du Général-Leclerc, ni sur les quais. Un seul petit tronçon existe, avenue Robert-Schuman. Mais il a suffi d'une pétition de 500 parents qui voulaient pouvoir garer leur 4 x 4 à côté de l'école pour que la mairie réduise la piste cyclable à une voie. » Consciente de la difficulté, la mairie veut avancer pas à pas: « Il faut d'abord commencer par créer le besoin de vélos sur des trajets identifiés comme la piscine, le marché, le centre-ville... » En clair: on prête des vélos aux Boulonnais, puis on regarde quels parcours ils empruntent pour décider où tracer les pistes cyclables. Reste à savoir quand... Pour

(a) Example 1 : original image (b) Text enhancement : estimated reflectance; model 2 (c) Text enhancement : estimated reflectance; model 4



(d) Text enhancement : estimated non uniform illumination; model 2 (e) Text enhancement : estimated non uniform illumination; model 4

FIGURE 1 – Original image acquired from a Sony Ericsson K800i

Références

[1] A. S. Georghiadis, D. J. Kriegman, and P. N. Belhumeur, *Illumination cones for recognition under variable lighting : Faces*, In CVPR '98 : Proceedings of the IEEE Computer Society

Conference on Computer Vision and Pattern Recognition, page 52, Washington, DC, USA, 1998. IEEE Computer Society.

- [2] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, *From few to many :illumination cone models for face recognition under variable lighting and pose*, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23(6) :643660, Jun 2001.
- [3] 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.
- [4] 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.
- [5] Tammy Riklin-Raviv and Amnon Shashua, *The quotient image :class based recognition and synthesis under varying illumination conditions*, cvpr, 02 :2566, 1999.
- [6] A. Shashua and T. Riklin-Raviv, *The quotient image : class-based re-rendering and recognition with varying illuminations*, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23(2) :129139, Feb 2001.
- [7] Haitao Wang, Stan Z. Li, Yangsheng Wang, and Weiwei Zhang, *Illumination modeling and normalization for face recognition*, In AMFG '03 : Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures, page 104, Washington, DC, USA, 2003. IEEE Computer Society.
- [8] H.G. Barrow and J.M. Tenenbaum, *Recovering intrinsic scene characteristics from images*, In CVS78, pages 326, 1978.
- [9] Szirmay-Kalos Lili, $\frac{1}{2}$ szli, $\frac{1}{2}$, *Monte-Carlo Global Illumination Methods - State of the Art and New Developments*, SCCG'99, Invited talk 1999.
- [10] 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.
- [11] I. Nwogu, Z. Shi and V. Govindaraju, *PDE-Based Enhancement of Low Quality Documents*, In Proceedings of the Ninth international Conference on Document Analysis and Recognition - Vol.01, pp.541-545 (2007).
- [12] 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.
- [13] 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).
- [14] B. K. Horn, Robot Vision, MIT Press, 1986.
- [15] 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.
- [16] 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.
- [17] Antoine Martin and Martin Lefebure, *Digitized document transmitting method for e.g. fax machine, involves extracting information relative to document, and reconstituting document by using information, filing document by considering information in reconstituted document*, Realeyes3D SA, patent, <http://www.prior-ip.com/application/33411968/>.

- [18] A. S. Georghiades, D. J. Kriegman, and P. N. Belhumeur, *Illumination cones for recognition under variable lighting : Faces*, In CVPR '98 : Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, page 52, Washington, DC, USA, 1998. IEEE Computer Society.
- [19] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, *From few to many :illumination cone models for face recognition under variable lighting and pose*, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23(6) :643660, Jun 2001.
- [20] 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.
- [21] Tammy Riklin-Raviv and Amnon Shashua, *The quotient image :class based recognition and synthesis under varying illumination conditions*, cvpr, 02 :2566, 1999.
- [22] A. Shashua and T. Riklin-Raviv, *The quotient image : class-based rerendering and recognition with varying illuminations*, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23(2) :129139, Feb 2001.
- [23] HaitaoWang, Stan Z. Li, YangshengWang, and Weiwei Zhang, *Illumination modeling and normalization for face recognition*, In AMFG '03 : Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures, page 104, Washington, DC, USA, 2003. IEEE Computer Society.
- [24] H.G. Barrow and J.M. Tenenbaum, *Recovering intrinsic scene characteristics from images*, In CVS78, pages 326, 1978.
- [25] Szirmay-Kalos László, *Monte-Carlo Global Illumination Methods - State of the Art and New Developments*, SCCG'99, Invited talk 1999.
- [26] 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.
- [27] Giusti E., *Minimal Surfaces and Functions of Bounded Total Variation*, Monographs in Mathematics vol 80, Boston, MA : Birkhauser, 1984.
- [28] M. El Rhabi and G. Rochefort *Method of restoring a blurred image acquired by means of a camera fitted to a communication terminal*, Realeyes3D SA, patent : <http://www.wipo.int/patentscope/search/en/W02009112710>



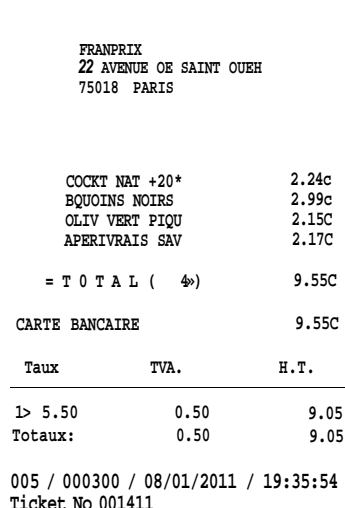
(a) Example 2, original image



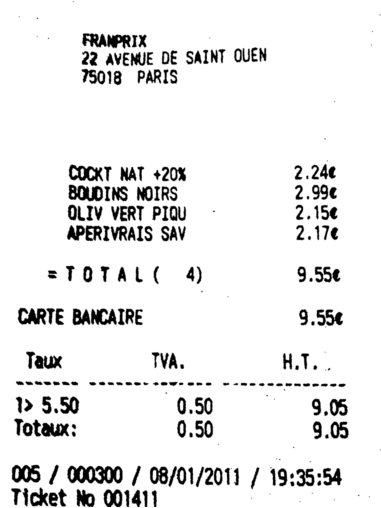
(b) Recognition by a classical OCR software; original image



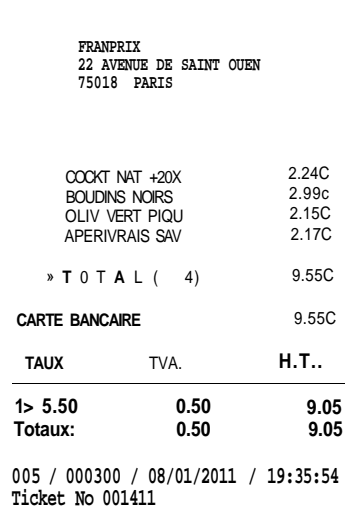
(c) Text enhancement : estimated reflectance; model 2



(d) Recognition by a classical OCR software; estimated reflectance; model 2



(e) Text enhancement : estimated reflectance; model 4

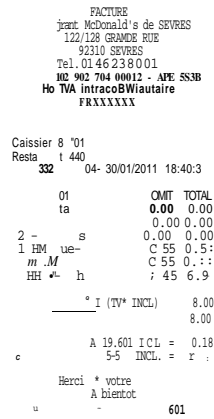


(f) Recognition by a classical OCR software; estimated reflectance; model 4

FIGURE 2 – Original image acquired from an iphone 3GS : first test



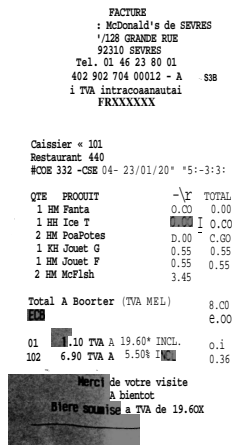
(a) Example 3, original image



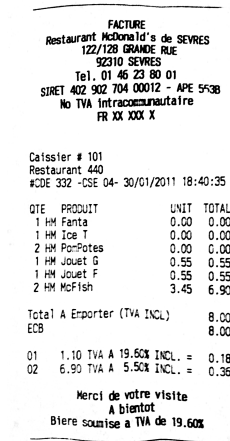
(b) Recognition by a classical OCR software; original image



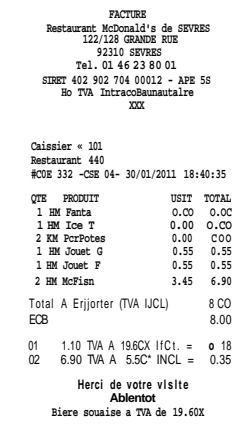
(c) Text enhancement : estimated reflectance; model 2



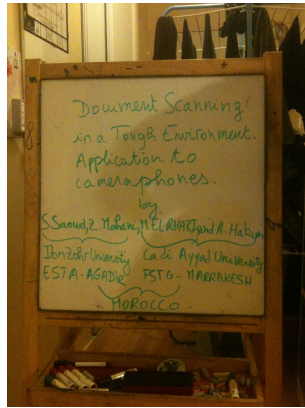
(d) Recognition by a classical OCR software; estimated reflectance; model 2



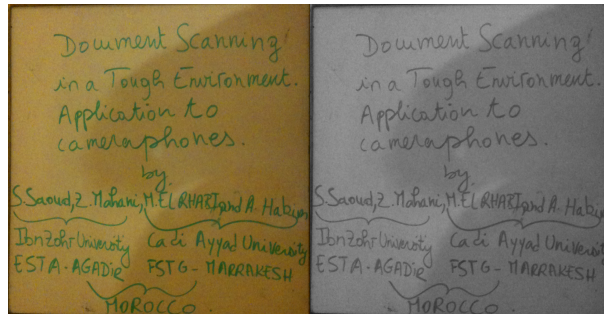
(e) Text enhancement : estimated reflectance; model 4



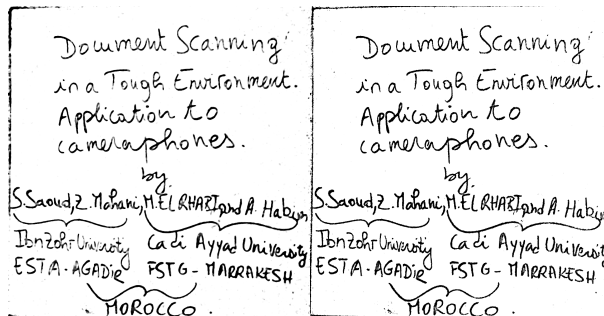
(f) Recognition by a classical OCR software; estimated reflectance; model 4



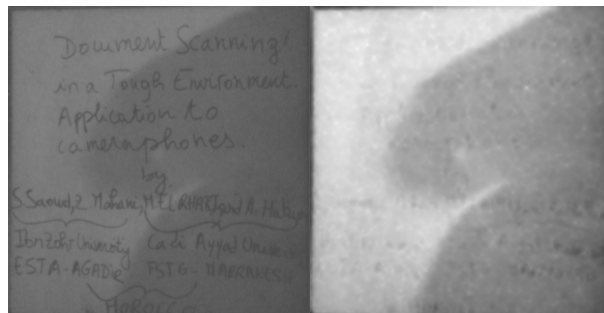
(a) Example 4 : original Whiteboard-image



(b) The (original) cropped image (c) The (original) cropped image in grayscale



(d) Text enhancement : es-estimated reflectance; model 2 (e) Text enhancement : es-estimated reflectance; model 4



(f) Text enhancement : es-estimated non uniform illumination; model 2 (g) Text enhancement : es-estimated non uniform illumination; model 4

FIGURE 4 – Original image acquired from an Iphone 3GS : Whiteboard test