

Les bases du prétraitement des images

Véronique EGLIN
veronique.eglin@insa-lyon.fr

LIRIS, INSA Lyon

Janvier, 2011

Plan

■ Quelques bases en traitement des images

- ◆ Réduire la présence de bruit dans les images
- ◆ Filtrer les images
- ◆ Modification du contraste
- ◆ Diverses corrections
- ◆ Restaurer les images binaires
- ◆ Segmenter les formes
- ◆ Compresser les contenus

Le problème du bruit dans les images (1)

► **Bruit d'une image**

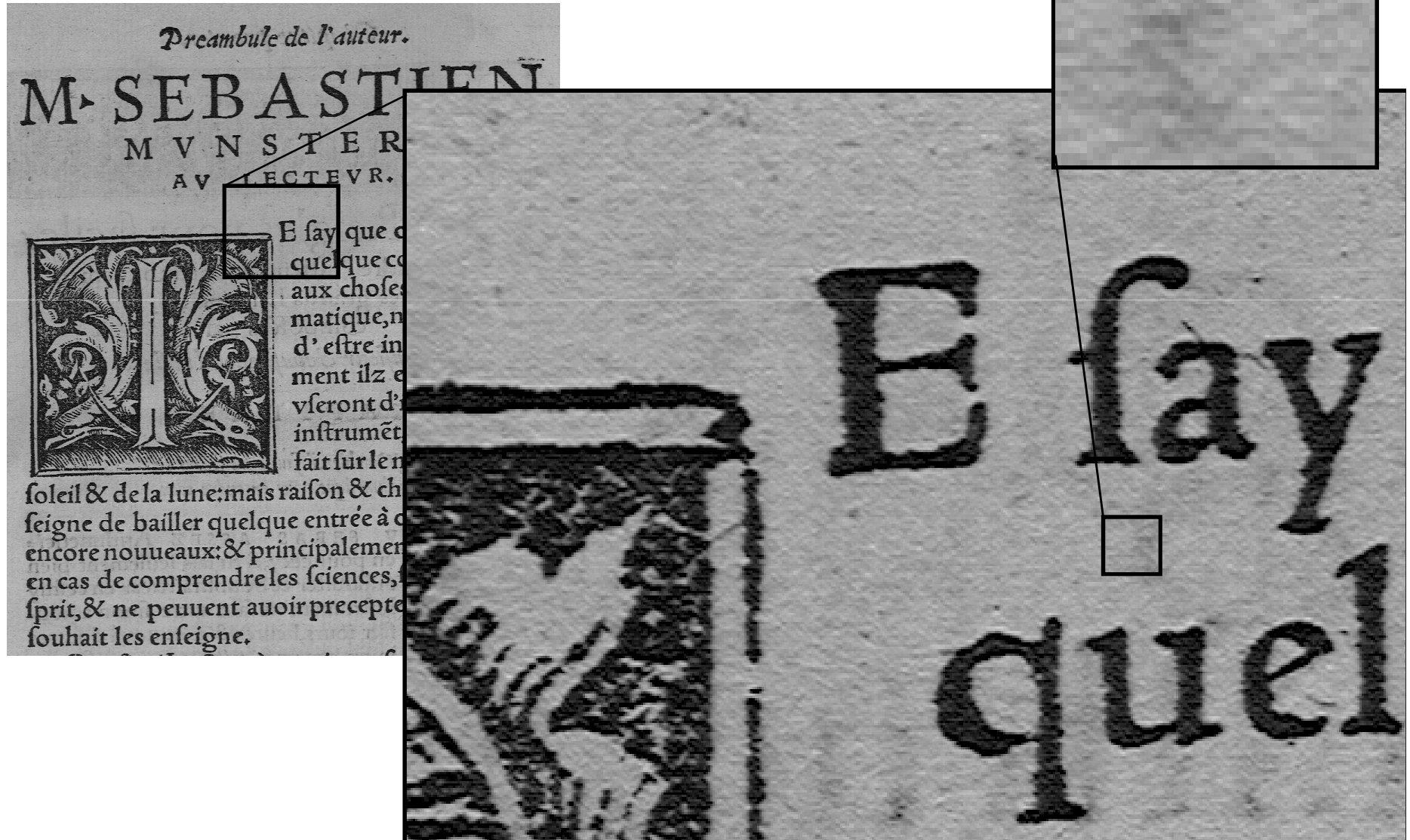
Fluctuations des couleurs sur des régions uniformes

► **Origine, au niveau de:**

- éclairage (reflets, réflexions multiples...)
- objet (rayures, déformation de la surface, rugosité)
- transmission des ondes lumineuses (pollution, chaleur)
- optique de la caméra (déformations optiques)
- capteur CCD (bruit du capteur)
- échantillonnage (discrétisation du signal continu)
- stockage de l'image quand celle-ci est comprimée (JPEG)

Le problème du bruit dans les images (2)

Bruit : tous les effets non désirés dans une image



Réduire le bruit dans les images?

- 1) Simplifier la chaîne d'acquisition
 - 2) Réduire le bruit du maillon le plus faible de la chaîne
 - 3) Filtrer l'image si pas d'autres solutions
-
- ▶ Parfois refaire la numérisation
 - ▶ Passer plus de temps à la prise d'image plutôt qu'à la restauration

Plan

- Quelques bases en traitement des images
 - ◆ Réduire la présence de bruit dans les images
 - ◆ Filtrer les images
 - ◆ Modification du contraste
 - ◆ Diverses corrections
 - ◆ Restaurer les images binaires
 - ◆ Segmenter les formes
 - ◆ Compresser les contenus

Filtrage morphologique

Morphologie mathématique est une science qui s'intéresse à **analyser les formes** dans les images, pour les caractériser, les détecter, les mesurer.

Cette science s'appuie sur la théorie mathématique **ensembliste**, la **géométrie discrète** et la **topologie**

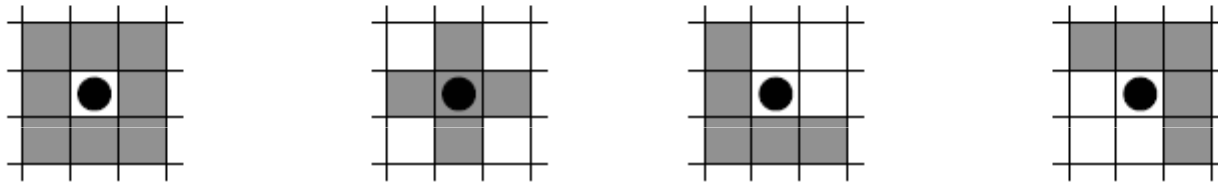
Il s'agit de comparer les objets à un autre objet de forme connue **B (élément structurant)** à l'aide de transformations dont les résultats sont de nouvelles images

Les principales opérations sont les **érosions** et les **dilatations** d'une image par un **élément structurant B**

Filtrage morphologique

- Élément structurant

- Un élément structurant est un masque binaire muni d'un point d'ancrage :



Éléments structurants

- Soient (x,y) les coordonnées d'un pixel et M un élément structurant. Alors, $M(x,y)$ représente l'ensemble des pixels qui coïncident avec les points noirs de M lorsque le point d'ancrage est superposé au pixel de coordonnées (x,y)

Filtrage morphologique

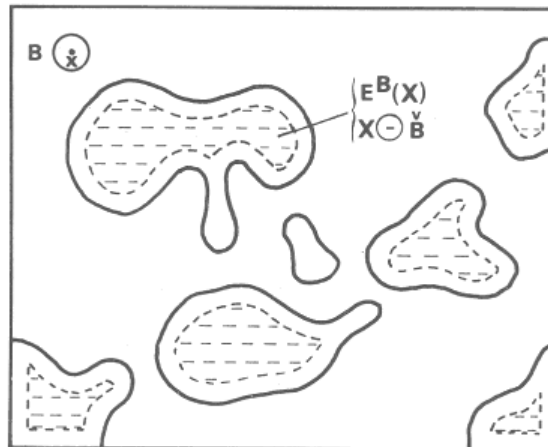
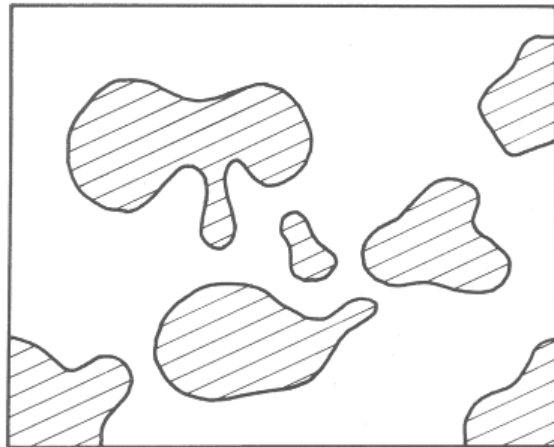
$D^B(X)$: Dilatation de X est l'image Y des points de X qui ont une intersection non vide avec B

$E^B(X)$: Erosion de X est l'image Y des points de X qui ont une intersection vide avec B

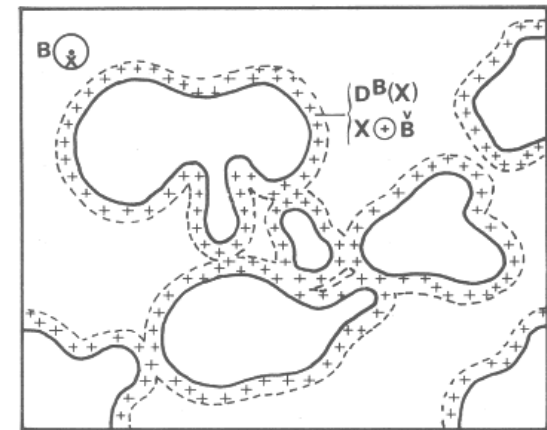
$$D^B(X) = \{X / B \cap X \neq \emptyset\} \quad E^B(X) = \{X / B \cap X = \emptyset\}$$

Erosion et Dilatation sont des opérations duales :

$$D^B(X) = \left(E^B(X^c) \right)^c$$



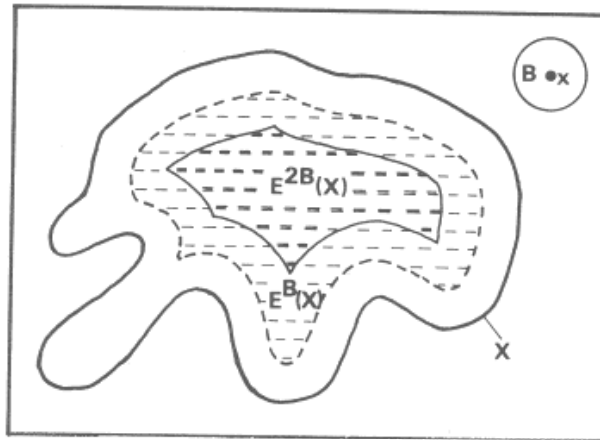
Erosion



Dilatation

Filtrage morphologique

Traitements itératifs (exemple : érosions itératives)

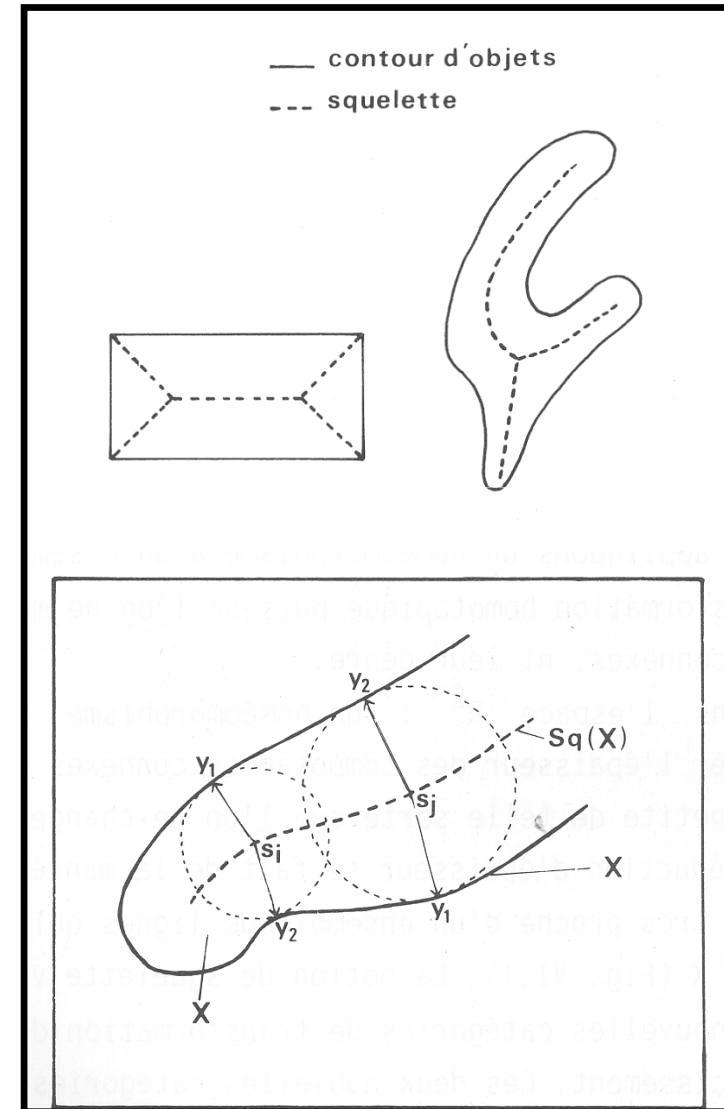


Squelettes

L'ensemble des points situés à équidistance de deux points du bord

L'ensemble des centres des boules maximales incluses dans la forme

Les squelettes sont calculées par érosion morphologique

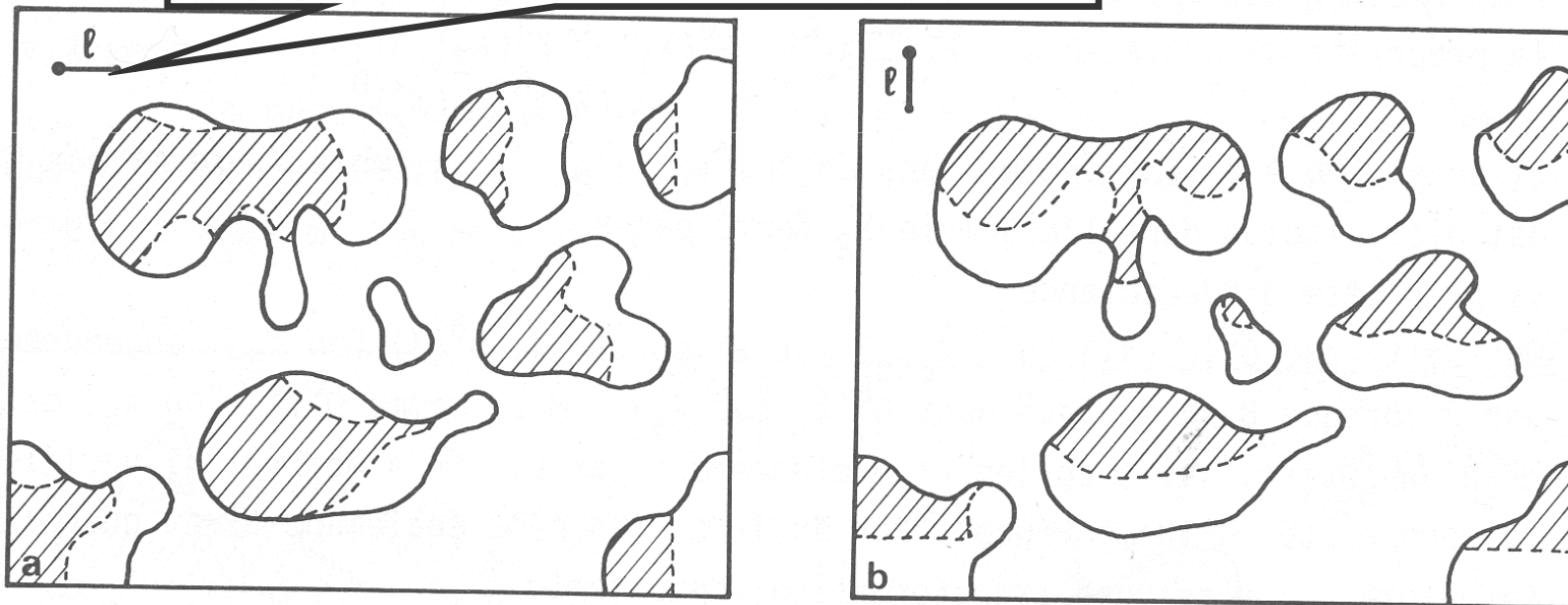


Filtrage morphologique

Changement d'élément structurant (élément structurant ligne/colonne)

La position du centre de l'élément structurant est importante pour la symétrie de la transformation

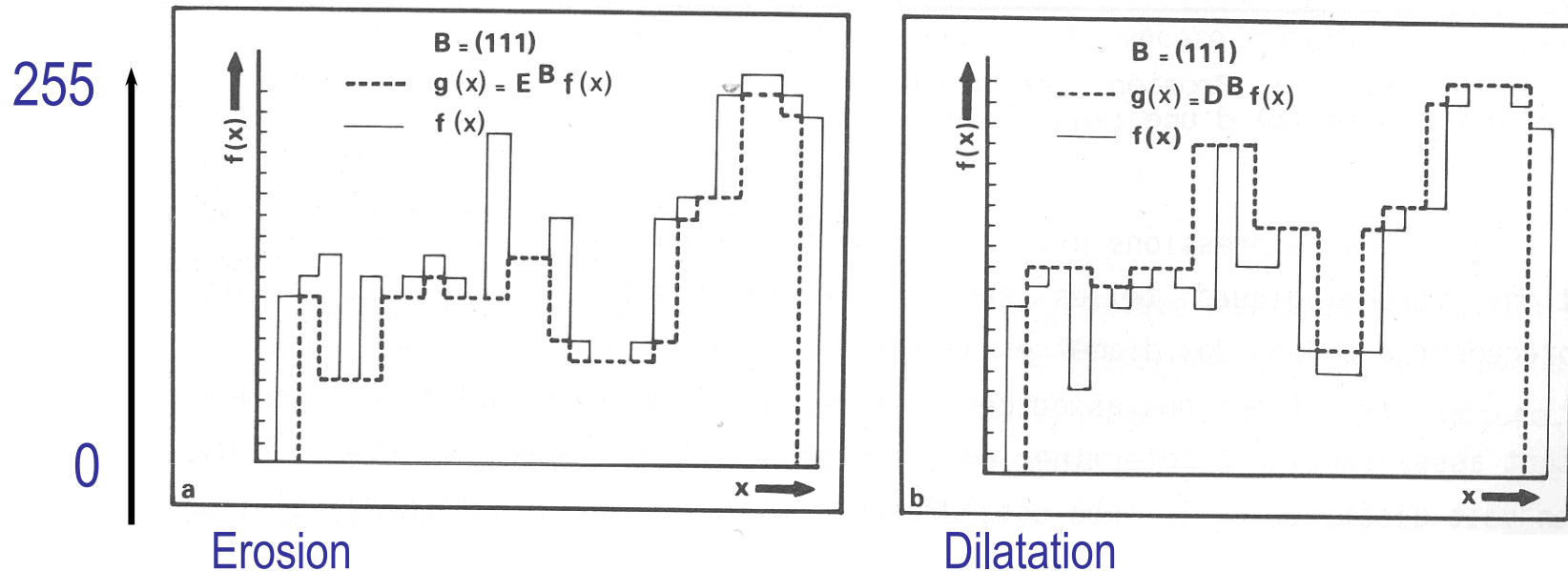
Élément structurant non symétrique



Un élément structurant 'ligne' érode les parties dont la largeur est inférieure à la taille de l'élément structurant

Filtrage morphologique

Morphologie mathématique sur les images à niveaux de gris



$$E^B (f(x)) = \text{Min} \{ f(x), x \in B \}$$

Réduit les pics, étale les vallées, homogénéise l'image, l'assombrit, étale les bords de l'objet

$$D^B (f(x)) = \text{Max} \{ f(x), x \in B \}$$

Comble les vallées, épaissit les pics, homogénéise l'image, l'éclaircit, fait disparaître les objets sombres

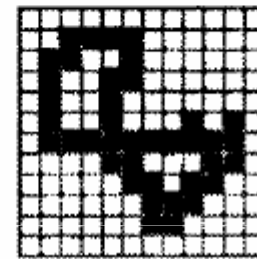
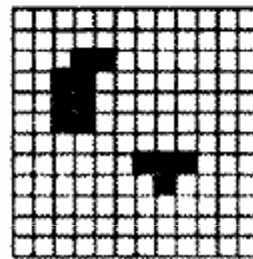
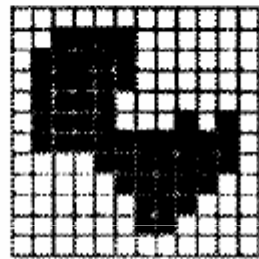
Filtrage morphologique

- Extraction du contour

L'extraction du contour d'une composante 4-connexe est obtenue par l'expression

$$\text{Contour4}(X) = X - \text{Ero}_{B9}(X)$$

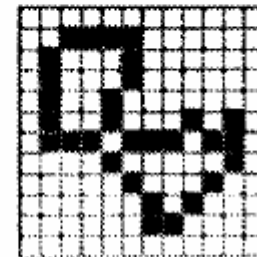
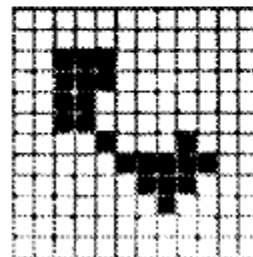
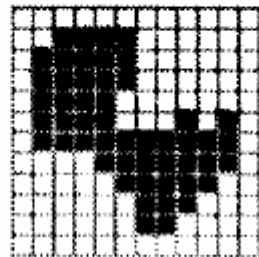
$$B9 =$$

L'extraction du contour d'une composante 8-connexe est obtenue par l'expression

$$\text{Contour8}(X) = X - \text{Ero}_{C5}(X)$$

$$C5 =$$

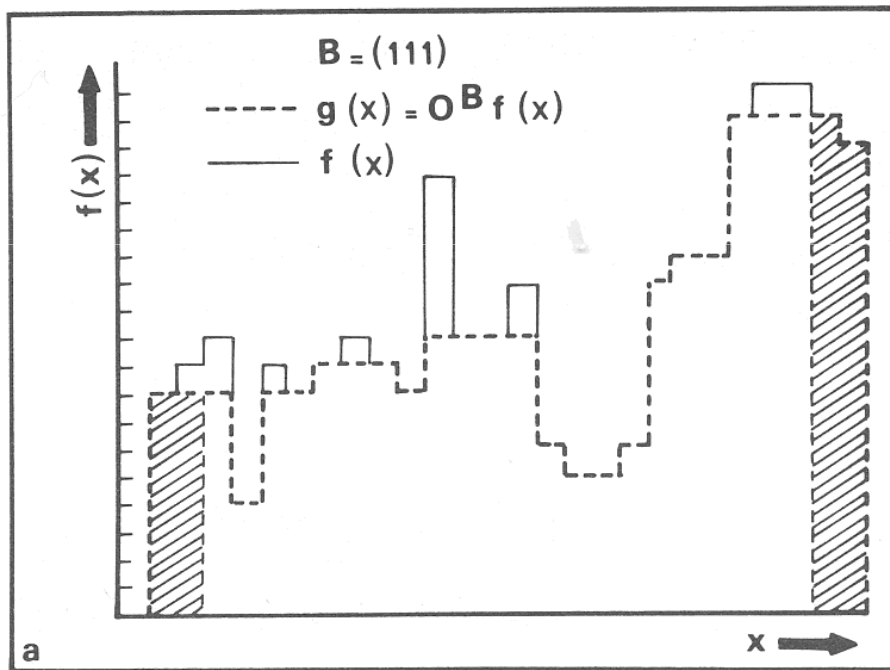



Filtrage morphologique

Morphologie mathématique sur les images à niveaux de gris

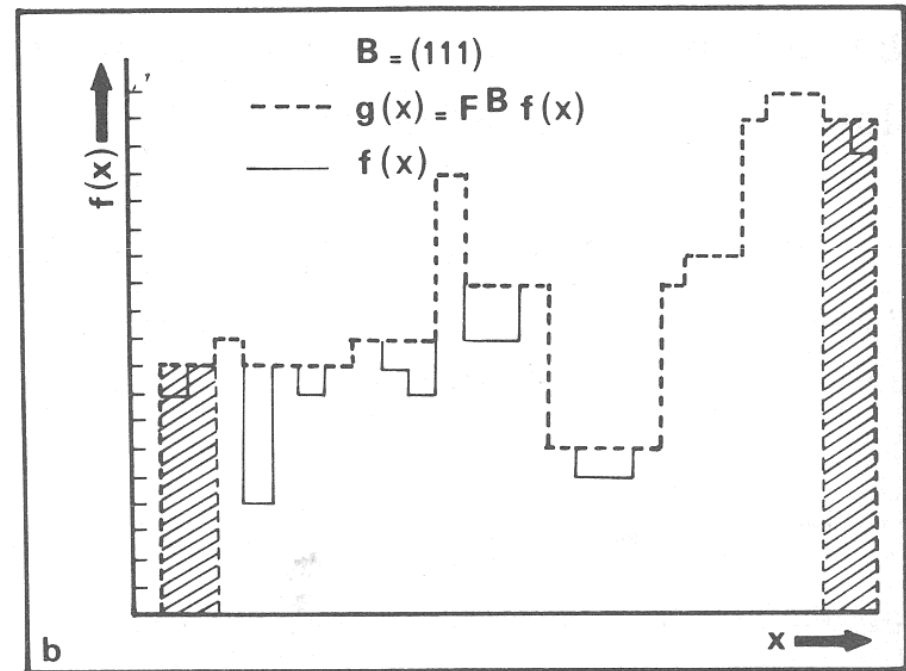
Ouverture = Erosion suivie d'une dilatation avec le même élément

Fermeture = Dilatation suivie d'une érosion avec le même élément



$$O^B(f(x)) = D^B(E^B(f(x)))$$

Lisser les contours



$$F^B(f(x)) = E^B(D^B(f(x)))$$

Fermer des contours disjoints

Filtrage morphologique

- Élimination du bruit
 - L'élimination du bruit dans une image binaire peut se faire par un opérateur morphologique d'ouverture

BG vom

AS, 11

BG vom

AS, 11

- Restauration des composantes connexes
 - La restauration des composantes connexes d'une image binaire peut être réalisée par un opérateur morphologique de fermeture

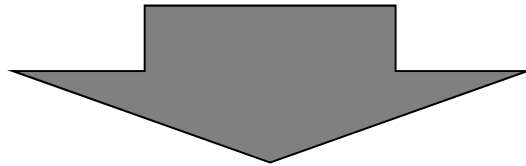
Verband
abfälle

Verband
abfälle

Filtrage morphologique

Restauration des caractères

So, we find special
term rewriting for



So, we find special
term rewriting for

Filtrage morphologique

Transformation en chapeau haute forme (Top Hat):

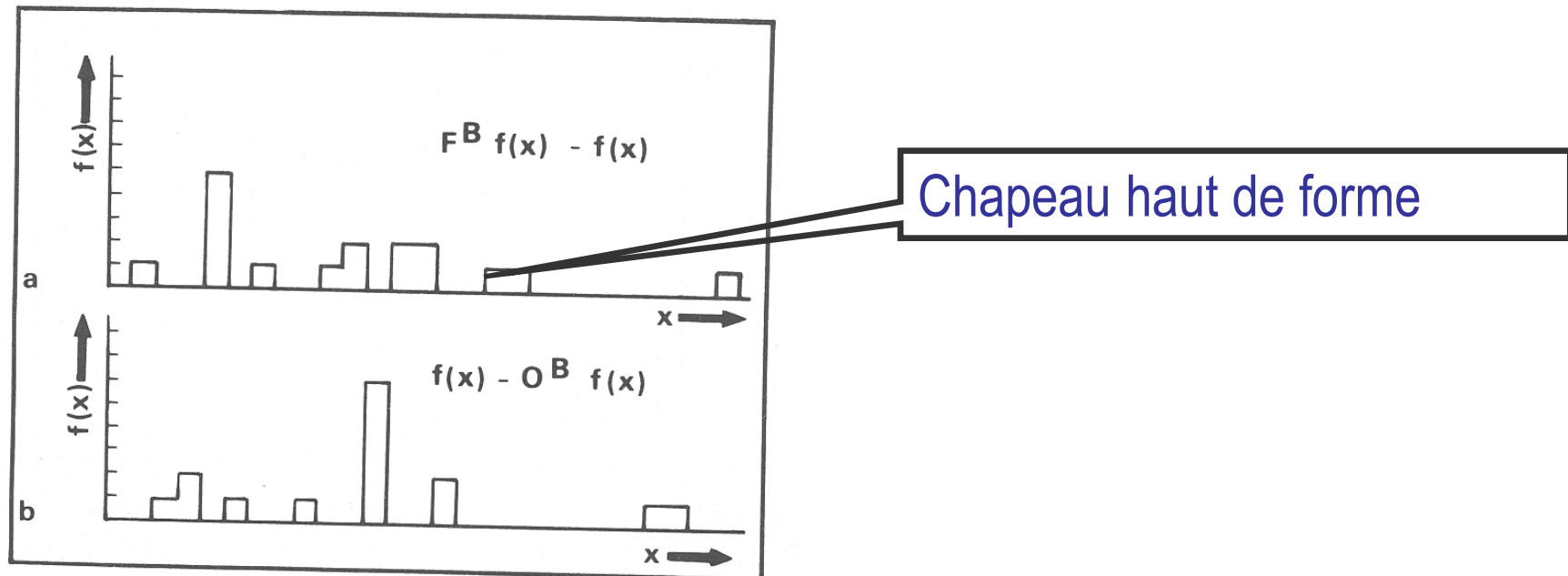
$$F^B(f(x)) - f(x) \quad \text{et} \quad f(x) - O^B(f(x))$$

-L'ouverture nivelle les crêtes étroites -

-La fermeture nivelle les vallées étroites

La Fermeture moins l'image originale met en valeur les crêtes, détecte les zones foncées de l'image (BTH)

L'image originale moins l'ouverture met en valeur les vallées (WTH)



Filtrage morphologique

Les images de texte sont des **images de traits** , elles peuvent donc être traitées efficacement par morphologie mathématique

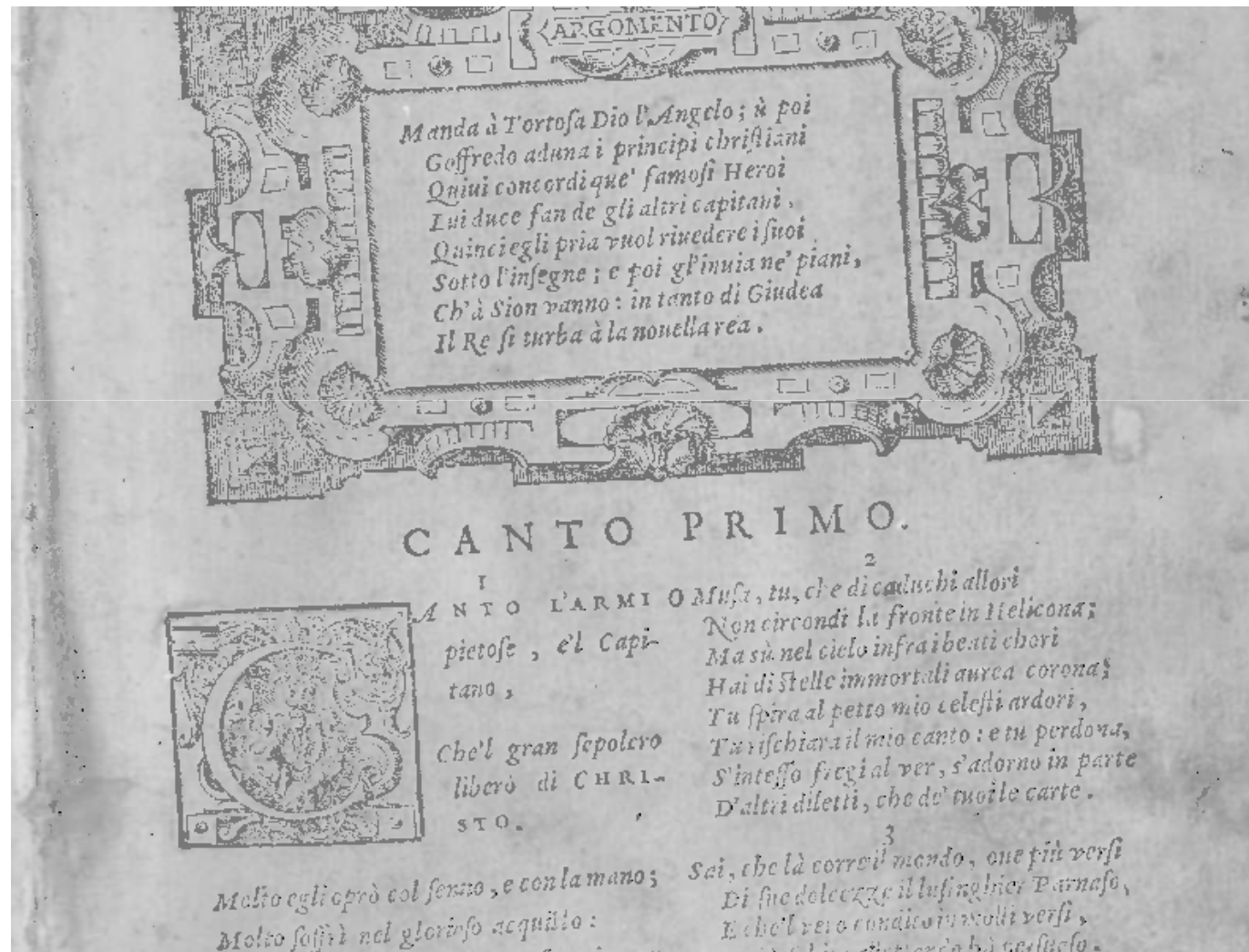


Image originale

Filtrage morphologique

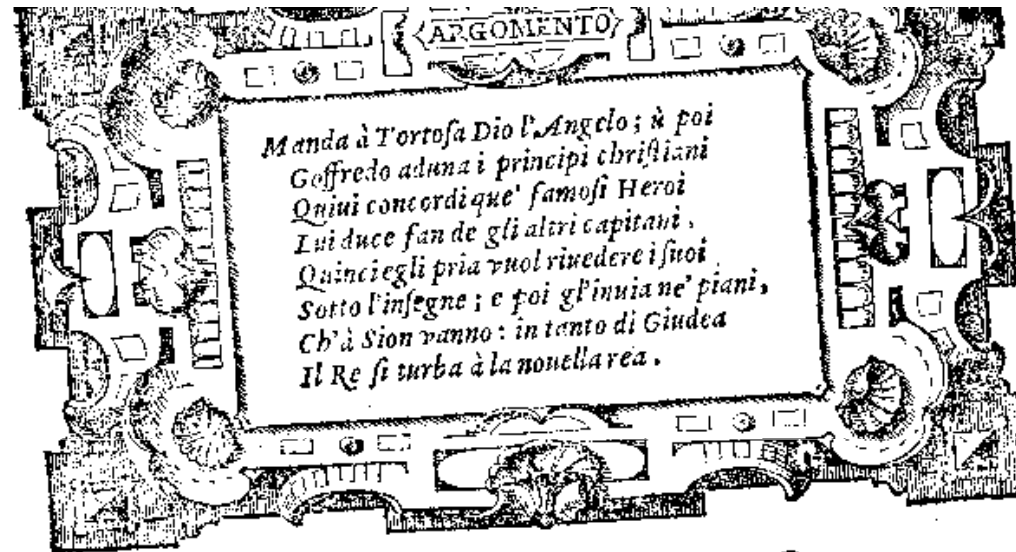
Les images de texte sont des **images de traits** , elles peuvent donc être traitées efficacement par morphologie mathématique



ouverture

Filtrage morphologique

Les images de texte sont des **images de traits** , elles peuvent donc être traitées efficacement par morphologie mathématique



CANTO PRIMO.



¹ CANTO L'ARMI O Musa, tu, che di caduchi allori
 pietose, e'l Capitano,
 Che'l gran sepolcro liberò di CRISTO.

² Non circondi la fronte in Helicon;
 Ma sù nel cielo infra i beati chori
 Hai di Stelle immortali aurea corona;
 Tu spiri al petto mio celesti ardori,
 Tu rischiara il mio canto: e tu perdona,
 S'intesso fregi al ver, s'adorno in parte
 D'altri diletti, che de' tuoi le carte.

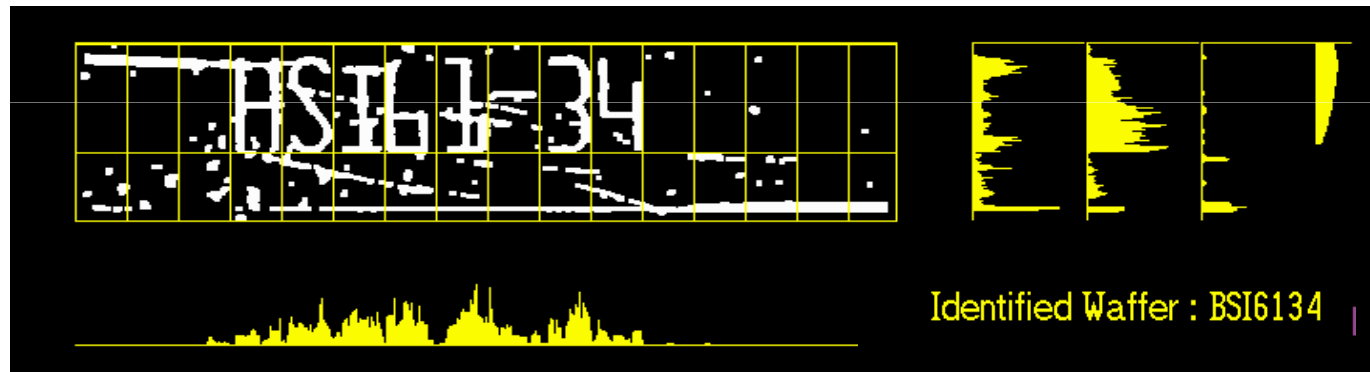
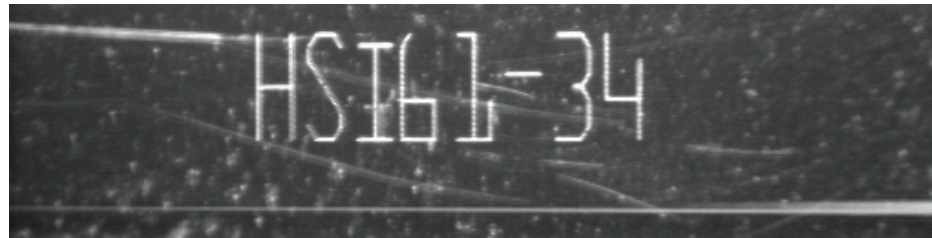
Molto egli oprò col ferro, e con la mano;
 Molto soffrì nel glorioso acquisto:

³ Sai, che là corre il mondo, one fin versi
 Di sue delezze il lusinghier Parnaso,
 E che'l vero condito in molti versi,
 E che'l vero condito in molti versi.

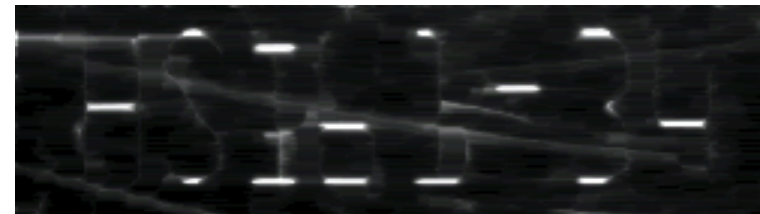
WTH = Image originale -ouverture

Filtrage morphologique

Morphologie avec un élément structurant ligne et colonne pour trouver les traits caractéristiques des caractères (caractérisation d'une forme en niveaux de gris)



Segmentation par seuillage , échec de reconnaissance

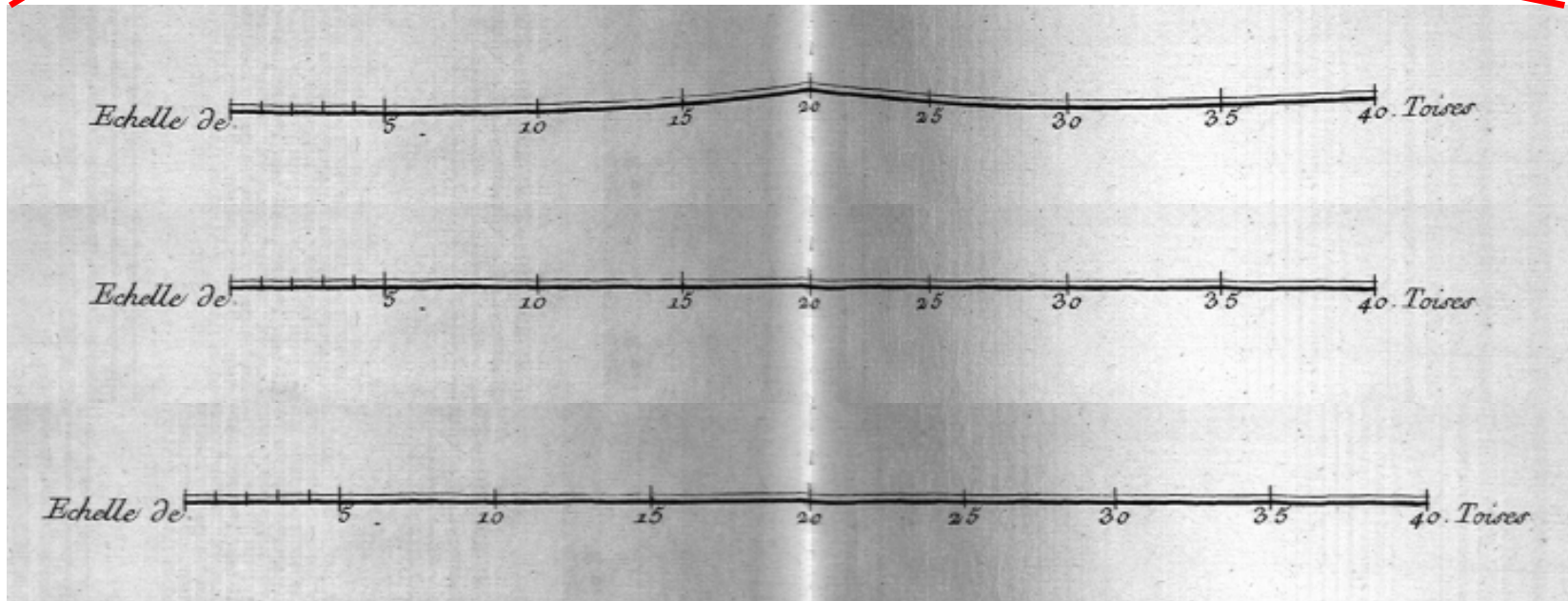
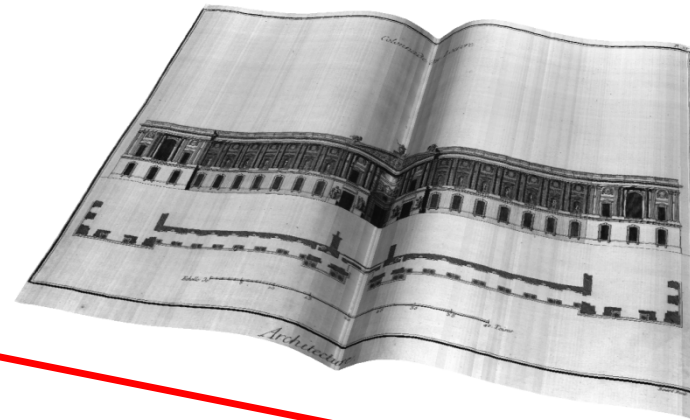
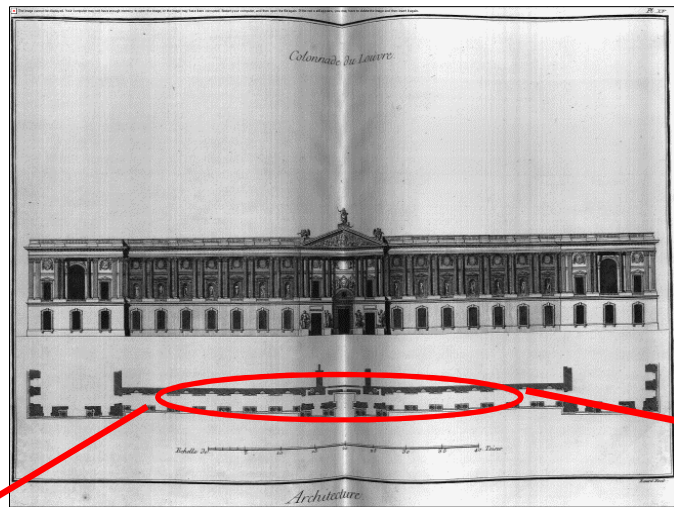


Détection des traits verticaux , Détection des traits horizontaux

Plan

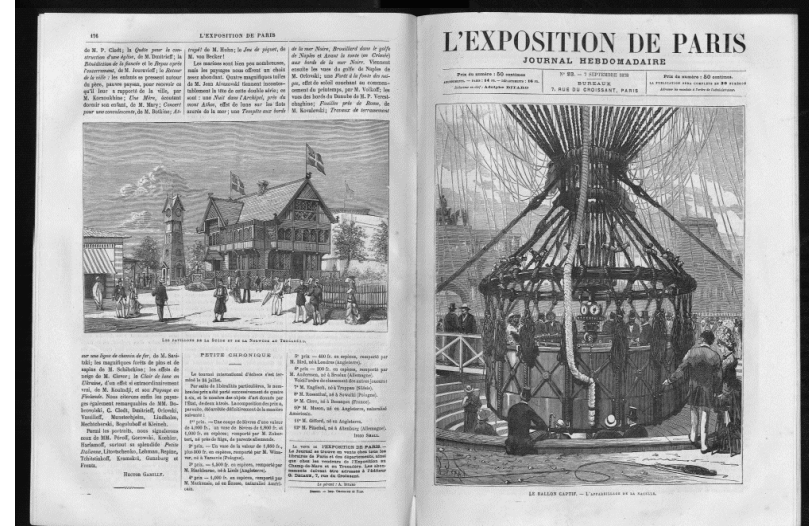
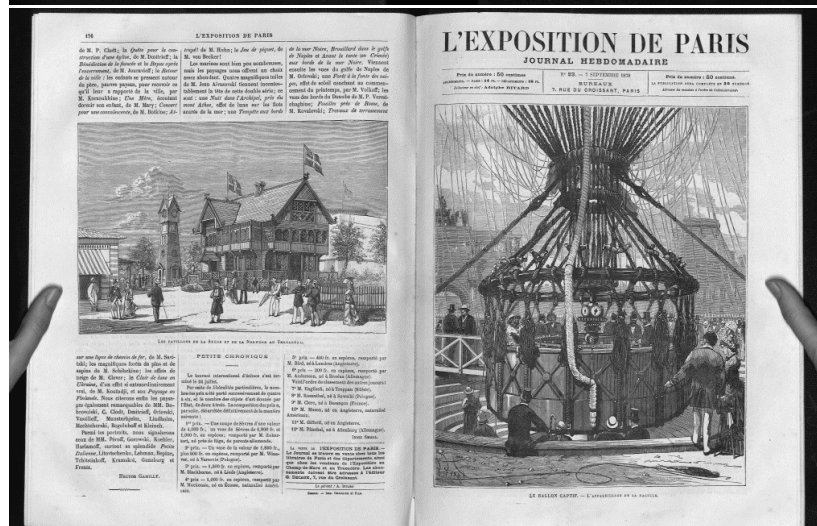
- Quelques bases en traitement des images
 - ◆ Réduire la présence de bruit dans les images
 - ◆ Filtrer les images
 - ◆ Modification du contraste
 - ◆ Diverses corrections
 - ◆ Restaurer les images binaires
 - ◆ Segmenter les formes
 - ◆ Compresser les contenus

Correction des distances



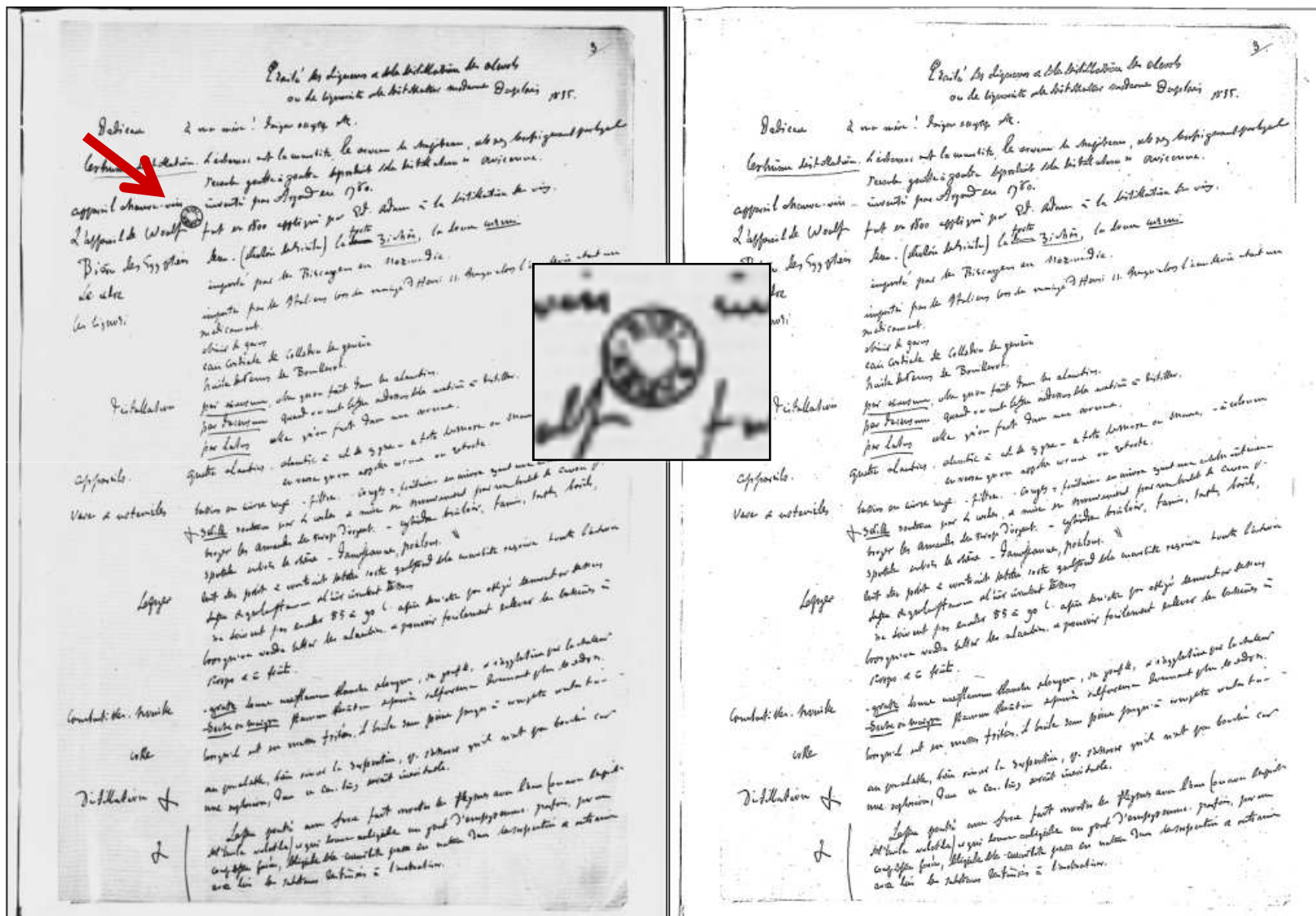
Traitements réalisés par Bookrestorer d'I2S en partenariat avec le LIRIS

L'effacement des doigts



Traitements réalisés par Bookrestorer d'I2S en partenariat avec LIRIS

La suppression des tampons



Traitement réalisé par V. Malleron, doctorant LIRIS

Correction automatique de l'inclinaison

64

en les richereux d'une sorte des impôts et en les gouvernant avec douceur.

20. Après la mort de Henry IV, Sully s'était retiré dans ses terres, et s'était consacré à la culture. Louis XIII l'ayant voulu chercher pour lui demander des avis sur les affaires, il y vint quelque avec répugnance. Les jeunes courtisans cherchaient à se trouver en ridicule sur son habillement qui n'était plus de mode, mais pour maintenir grave et par ses manières, Sully s'en aperçut et dit au Roi : « Sire, quand le soir entre bien, de glorieuse mémoire, me faisait l'honneur de me consulter sur les grandes et importantes affaires, au préalable, il faisait sortir tous les buffons et baladins de la chambre.

21. Louis XIV mourant, avoua toutes les fautes en présence de son fils, qui allait lui succéder. « Tâchez, lui dit-il, de vous souvenir de faire avec vos vassaux j'ai trop aimé la guerre, ne m'imiter pas en cela, non plus que dans les très grandes dépenses que j'ai faites. Travaillez vos sujets à l'égalité que vous donnez et faites ce que j'ai eu le malheur de ne pouvoir faire moi-même.

22. Pendant la minorité de Louis XV, en 1715, Marseille fut envahie par un terrible fléau, la peste, introduite dans cette ville par un marchand venant de Smyrne. L'imprudence des officiers de santé présida au lazaret, avait laissé entrer dans la cité, fit en peu de temps un grand nombre de victimes. On en compta bientôt jusqu'à cinq cents par jour. Mais la peste n'a pas appris tous ses secrets. Le sublime de la peste de l'époque Belzunce. On le vit alors fronder au malin les devoirs de la Religion, ravir les fidèles par son or et son langage et les consoler par ses fausses paroles.



64

en les richereux d'une sorte des impôts et en les gouvernant avec douceur.

20. Après la mort de Henry IV, Sully s'était retiré dans ses terres, et s'était consacré à la culture. Louis XIII l'ayant voulu chercher pour lui demander des avis sur les affaires, il y vint quelque avec répugnance. Les jeunes courtisans cherchaient à se trouver en ridicule sur son habillement qui n'était plus de mode, mais pour maintenir grave et par ses manières, Sully s'en aperçut et dit au Roi : « Sire, quand le soir entre bien, de glorieuse mémoire, me faisait l'honneur de me consulter sur les grandes et importantes affaires, au préalable, il faisait sortir tous les buffons et baladins de la chambre.

21. Louis XIV mourant, avoua toutes les fautes en présence de son fils, qui allait lui succéder. « Tâchez, lui dit-il, de vous souvenir de faire avec vos vassaux j'ai trop aimé la guerre, ne m'imiter pas en cela, non plus que dans les très grandes dépenses que j'ai faites. Travaillez vos sujets à l'égalité que vous donnez et faites ce que j'ai eu le malheur de ne pouvoir faire moi-même.

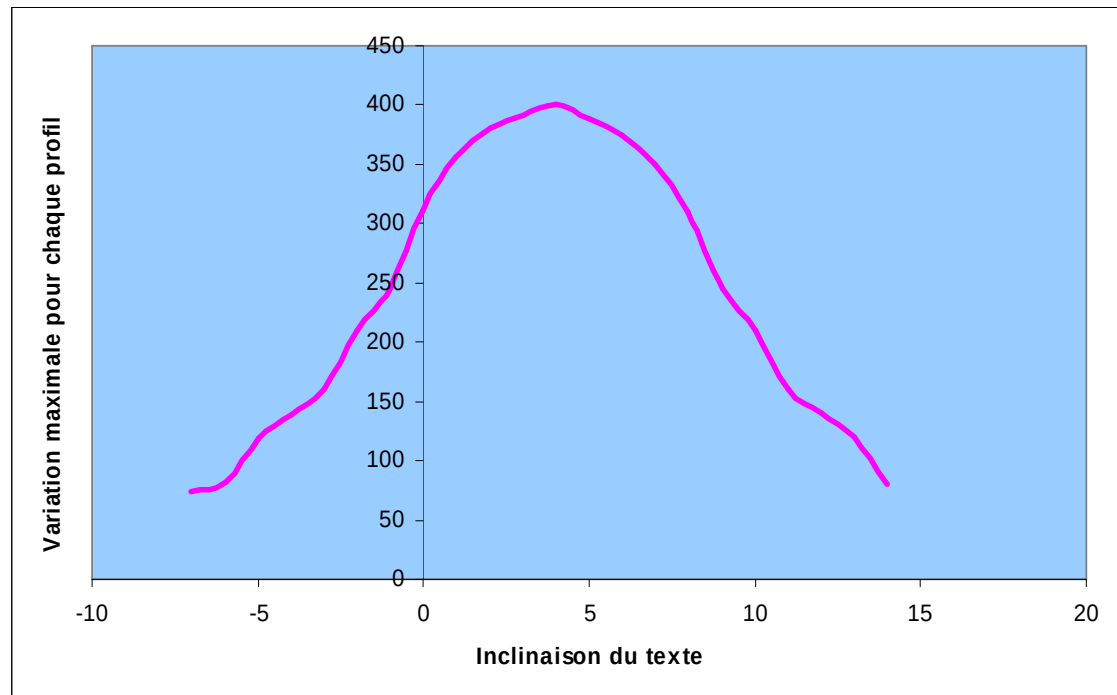
22. Pendant la minorité de Louis XV, en 1715, Marseille fut envahie par un terrible fléau, la peste, introduite dans cette ville par un marchand venant de Smyrne. L'imprudence des officiers de santé présida au lazaret, avait laissé entrer dans la cité, fit en peu de temps un grand nombre de victimes. On en compta bientôt jusqu'à cinq cents par jour. Mais la peste n'a pas appris tous ses secrets. Le sublime de la peste de l'époque Belzunce. On le vit alors fronder au malin les devoirs de la Religion, ravir les fidèles par son or et son langage et les consoler par ses fausses paroles.



Correction automatique de l'inclinaison

Une première étape : la détermination de l'inclinaison

Par l'analyse des profils de projection :

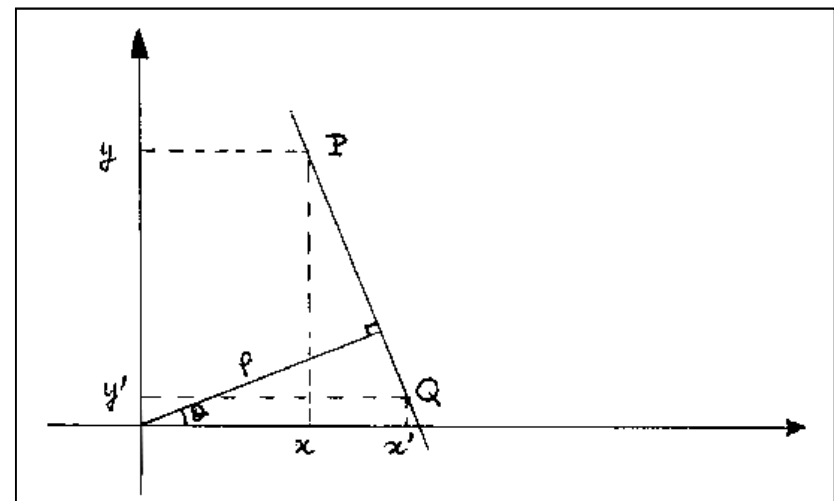


Correction automatique de l'inclinaison

Une première étape : la détermination de l'inclinaison

Par la transformée de Hough

1. Technique fondée sur le fait que le texte est composé d'entités connexes noires selon une inclinaison homogène pour l'ensemble des composantes.
2. Les lignes sont détectées en changeant **d'espace de représentation** à l'aide de la transformée de Hough, [Du73].
3. L'espace transformé est de dimension égale au nombre de paramètres (ρ, θ) , avec ρ la distance de cette droite à l'origine du repère et θ l'angle que fait la normale à cette droite avec l'axe des abscisses.



Correction automatique de l'inclinaison

Une première étape : la détermination de l'inclinaison

Par la transformée de Hough

4. Dans le plan de Hough, une **ligne sera représentée par un point**.
5. Pour définir ces lignes avec le plus de précision possible, un calcul de **moyennage** est effectué sur **les rectangles circonscrits** aux composantes pour assurer les meilleurs alignements possibles.
- . Un degré d'appartenance à une droite donnée est ensuite attribué à chaque composante. Il permet d'écarter les composantes isolées parasites non alignées.

L'angle d'inclinaison du nuage de points (ligne de texte) correspond à la valeur maximale de l'accumulateur.

Correction automatique de l'inclinaison

Application de la transformée de Hough : Méthode de Srihari 1996

Utilisation de l'espace transformé comme une matrice d'accumulation de dimensions n^2 , où n est le nombre de paramètres (ρ, θ) .

Les valeurs accumulées dans une case de la matrice correspondent à la chance d'apparition d'une droite dont les paramètres (ρ, θ) sont ceux de cette case. La (ou les) meilleure(s) direction(s) sera (seront) donc associée(s) aux maxima présents dans l'accumulateur.



Très bonne détection des régions de texte *même bruitées* quelle que soit leur inclinaison. Elles nécessitent cependant beaucoup de mémoire pour conserver les matrices de valeurs.

Correction automatique de l'inclinaison

La transformée de Hough

- Consomme beaucoup de temps : plutôt que considérer tous les **points** de l'image (les n^2 points) , on ne s'intéresse qu'aux **composantes connexes** (à leur centre, ou à tout autre point significatif de la composante)
- On peut aussi filtrer le texte et constituer des blocs de mots, plus rapide que les composantes connexes isolées

Correction automatique de l'inclinaison

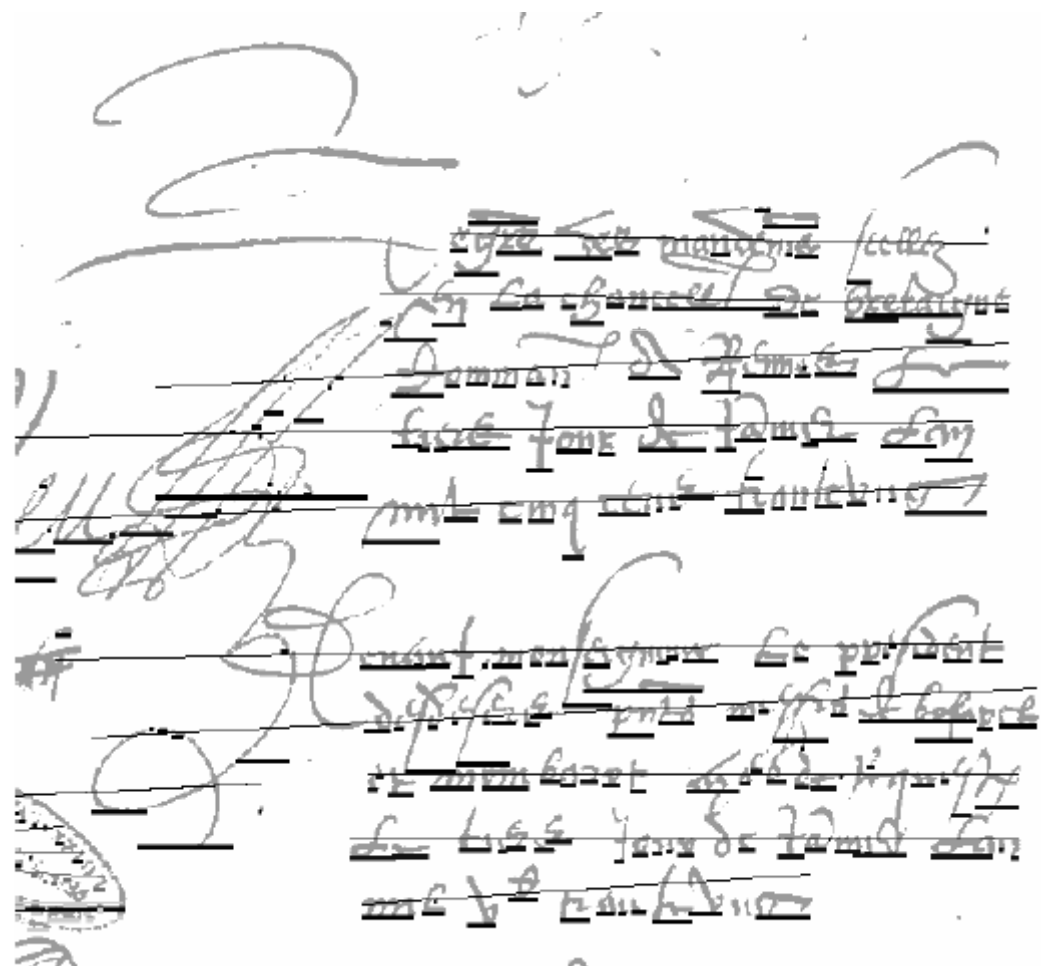
Autour de la transformée de Hough...

Présence de plusieurs inclinaisons différentes dans un même bloc

- Méthode de groupement pour le redressement des textes inclinés
- Méthode similaire à la transformée de Hough
- Chaque valeur est calculée dans un voisinage plutôt que sur l'ensemble du domaine
- Utilisation de tenseurs plus robustes que les vecteurs et scalaires
- Les angles d'inclinaison sont estimés le long des lignes inclinées par le tracé des courbes reliant les centres de chaque composante connexe
- Le champ complet des angles d'inclinaison s'obtient par interpolation de surface élastique.

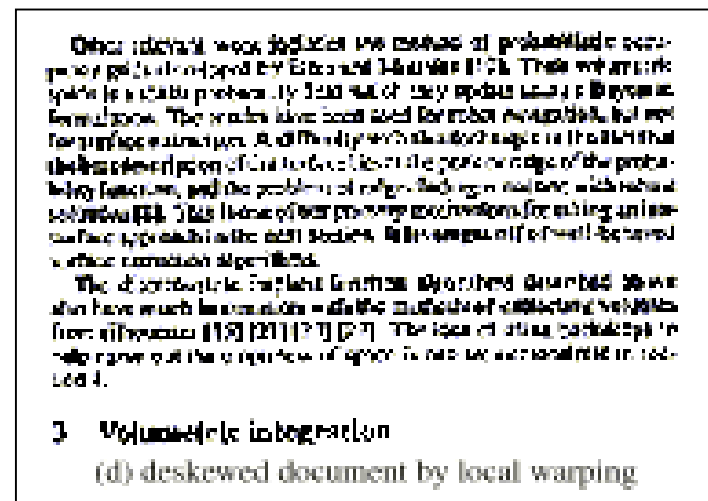
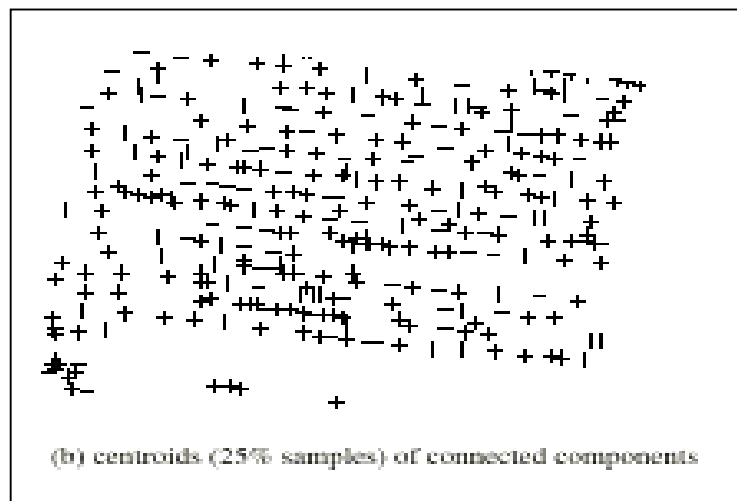
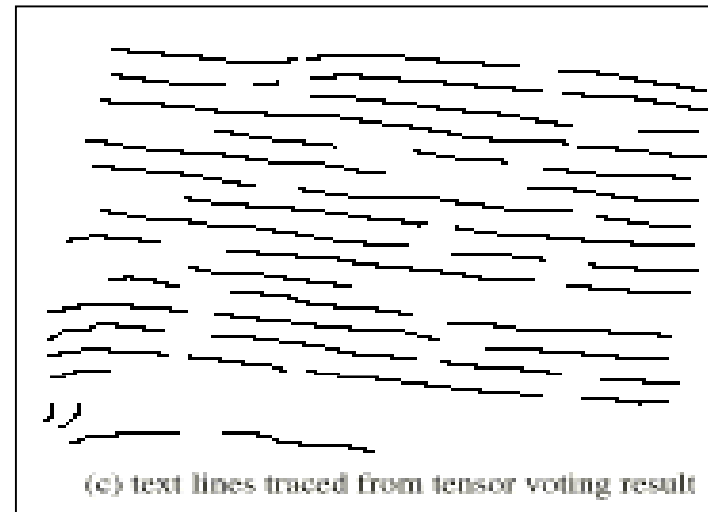
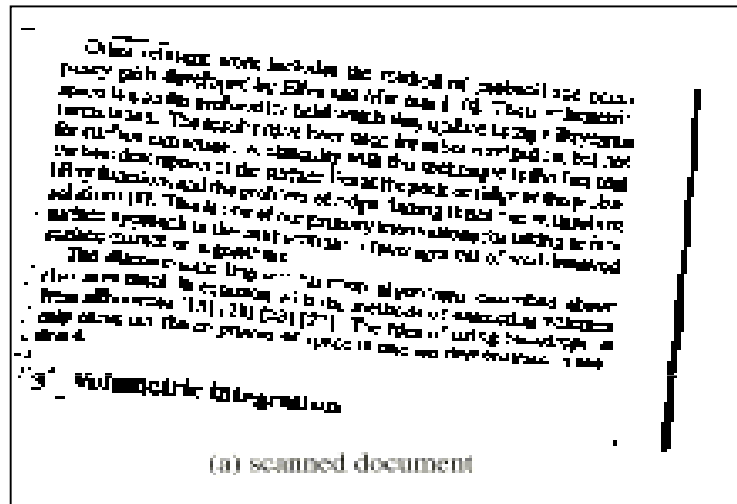
Correction automatique de l'inclinaison

La transformée de Hough



Correction automatique de l'inclinaison

Variante plus robuste: Local warping et tensor voting



Correction automatique de l'inclinaison

400

A N A

des caractères extérieurs les choses intérieurement différentes. Ces apparences sont destinées à nous servir d'étiquette pour suppléer à la faiblesse de nos sens, qui ne pénètrent pas jusqu'à l'intérieur des objets ; mais quelquefois nous nous méprenons à ces étiquettes. Il y a des plantes venimeuses qui ressemblent à des plantes très salutaires. Quelquefois nous sommes surpris de l'effet imprévu d'une cause, d'où nous nous attendions à voir naître un effet tout opposé : c'est qu'alors d'autres causes imperceptibles s'étant jointes avec cette première à notre insu, en changent la détermination. Il arrive aussi que le fond des objets n'est pas toujours diversifié à proportion de la dissimilitude extérieure. La règle de l'*analogie* n'est donc pas une règle de certitude, puisqu'elle a ses exceptions. Il suffit au dessein du Créateur, qu'elle forme une grande probabilité, que ses exceptions soient rares, & d'une influence peu étendue. Comme nous ne pouvons pénétrer par nos sens jusqu'à l'intérieur des objets, l'*analogie* est pour nous ce qu'est le témoignage des autres, quand ils nous parlent d'objets que nous n'avons ni vus, ni entendus. Ce sont-là deux moyens que le Créateur nous a laissés pour étendre nos connoissances. Détruisez la force du témoignage, combien de choses que la bonté de Dieu nous a accordées, dont nous ne pourrions tirer aucune utilité ! Les seuls sens ne nous suffisent pas : car quel est l'homme du monde qui puisse examiner par lui-même toutes les choses qui sont nécessaires à la vie ? Par conséquent dans un nombre infini d'occasions, nous avons besoin de nous instruire les uns les autres, & de nous en rapporter à nos observations mutuelles. Ce qui prouve en passant, que le témoignage, quand il est revêtu de certaines conditions, est le plus souvent une marque de la vérité ; ainsi que l'*analogie* tirée de la ressemblance extérieure des objets, pour en conclure leur ressemblance intérieure, en est le plus souvent une règle certaine. Voyez l'article CONNOISSANCE, où ces réflexions sont plus étendues.

400

A N A

des caractères extérieurs les choses intérieurement différentes. Ces apparences sont destinées à nous servir d'étiquette pour suppléer à la faiblesse de nos sens, qui ne pénètrent pas jusqu'à l'intérieur des objets ; mais quelquefois nous nous méprenons à ces étiquettes. Il y a des plantes venimeuses qui ressemblent à des plantes très salutaires. Quelquefois nous sommes surpris de l'effet imprévu d'une cause, d'où nous nous attendions à voir naître un effet tout opposé : c'est qu'alors d'autres causes imperceptibles s'étant jointes avec cette première à notre insu, en changent la détermination. Il arrive aussi que le fond des objets n'est pas toujours diversifié à proportion de la dissimilitude extérieure. La règle de l'*analogie* n'est donc pas une règle de certitude, puisqu'elle a ses exceptions. Il suffit au dessein du Créateur, qu'elle forme une grande probabilité, que ses exceptions soient rares, & d'une influence peu étendue. Comme nous ne pouvons pénétrer par nos sens jusqu'à l'intérieur des objets, l'*analogie* est pour nous ce qu'est le témoignage des autres, quand ils nous parlent d'objets que nous n'avons ni vus, ni entendus. Ce sont-là deux moyens que le Créateur nous a laissés pour étendre nos connoissances. Détruisez la force du témoignage, combien de choses que la bonté de Dieu nous a accordées, dont nous ne pourrions tirer aucune utilité ! Les seuls sens ne nous suffisent pas : car quel est l'homme du monde qui puisse examiner par lui-même toutes les choses qui sont nécessaires à la vie ? Par conséquent dans un nombre infini d'occasions, nous avons besoin de nous instruire les uns les autres, & de nous en rapporter à nos observations mutuelles. Ce qui prouve en passant, que le témoignage, quand il est revêtu de certaines conditions, est le plus souvent une marque de la vérité ; ainsi que l'*analogie* tirée de la ressemblance extérieure des objets, pour en conclure leur ressemblance intérieure, en est le plus souvent une règle certaine. Voyez l'article CONNOISSANCE, où ces réflexions sont plus étendues.

Correction automatique de l'inclinaison

La méthode des moindres carrés

$$a = \frac{n \sum_{k=1}^n x_k y_k - \sum_{k=1}^n x_k \sum_{k=1}^n y_k}{n \sum_{k=1}^n x_k^2 - \left[\sum_{k=1}^n x_k \right]^2}$$

$$b = \frac{\sum_{k=1}^n y_k - a \sum_{k=1}^n x_k}{n}$$

- On cherche des droites qui approchent correctement tous les points de la fonction
- En cas de succès, l'angle d'inclinaison est déduit de ses coefficients directeurs
- Dans le cas contraire, l'espace de recherche est divisé en deux et la recherche est relancée sur chaque partie
- Ce processus dichotomique est appliqué jusqu'à ce que toutes les parties soient ~~approchées par une droite~~

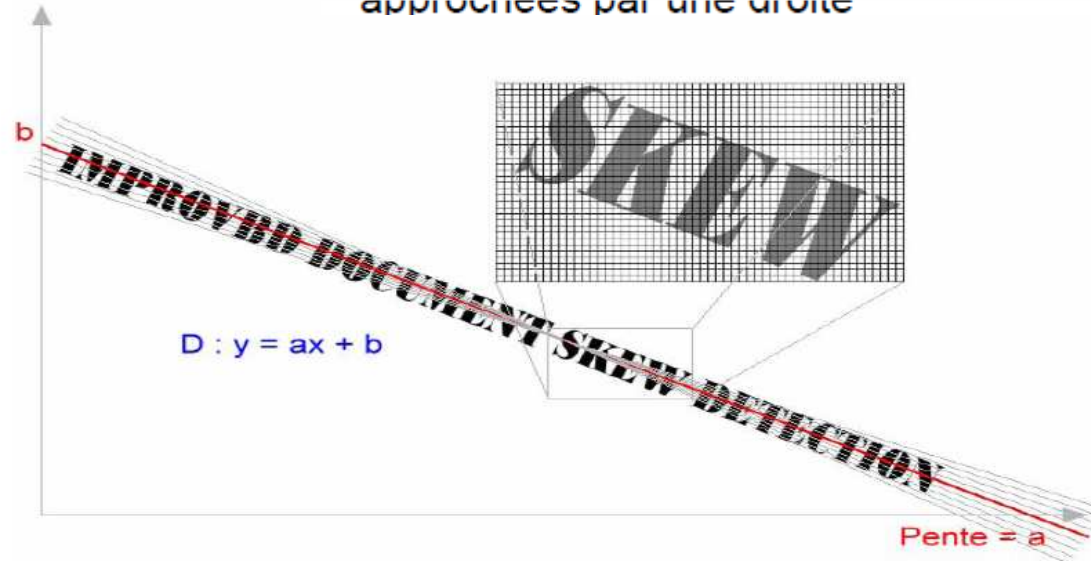


Figure 6 : Droite D des moindres carrés

Correction automatique de l'inclinaison

La méthode des moindres carrés

♦ Calcul de l'angle d'inclinaison du texte

- Chaque droite fournit un angle, on prend le plus fréquent qui doit correspondre à l'inclinaison du document
- Afin de ne pas prendre des valeurs correspondant au bruit, on considère les différentes séries de valeurs consécutives non nulles
- On calcule le critère : $l \times v$ où l = longueur de la série et v = nombre d'éléments prenant la valeur du mode de la série
- On retient ensuite la valeur dont le critère est le plus fort et on déduit l'angle d'inclinaison α

♦ Redressement

$$\begin{aligned}y' &= y \cos \alpha - x \sin \alpha \\x' &= x \cos \alpha + y \sin \alpha\end{aligned}$$

♦ Critiques

- Bons résultats si la taille des caractères est petite et si les lignes sont resserrées
- Peut conduire à des résultats erronés au bout des lignes

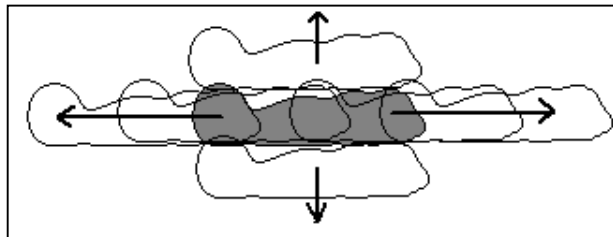
Correction automatique de l'inclinaison

L'auto-corrélation

Elle permet de mettre en évidence les périodicités dans les images par translation des différents points d'un objet dans toutes les directions.

Par exemple: translation d'une droite dans sa propre direction va conduire à un fort niveau de correspondance ($\approx$ forte valeur de l'auto-corrélation dans la direction de la droite).

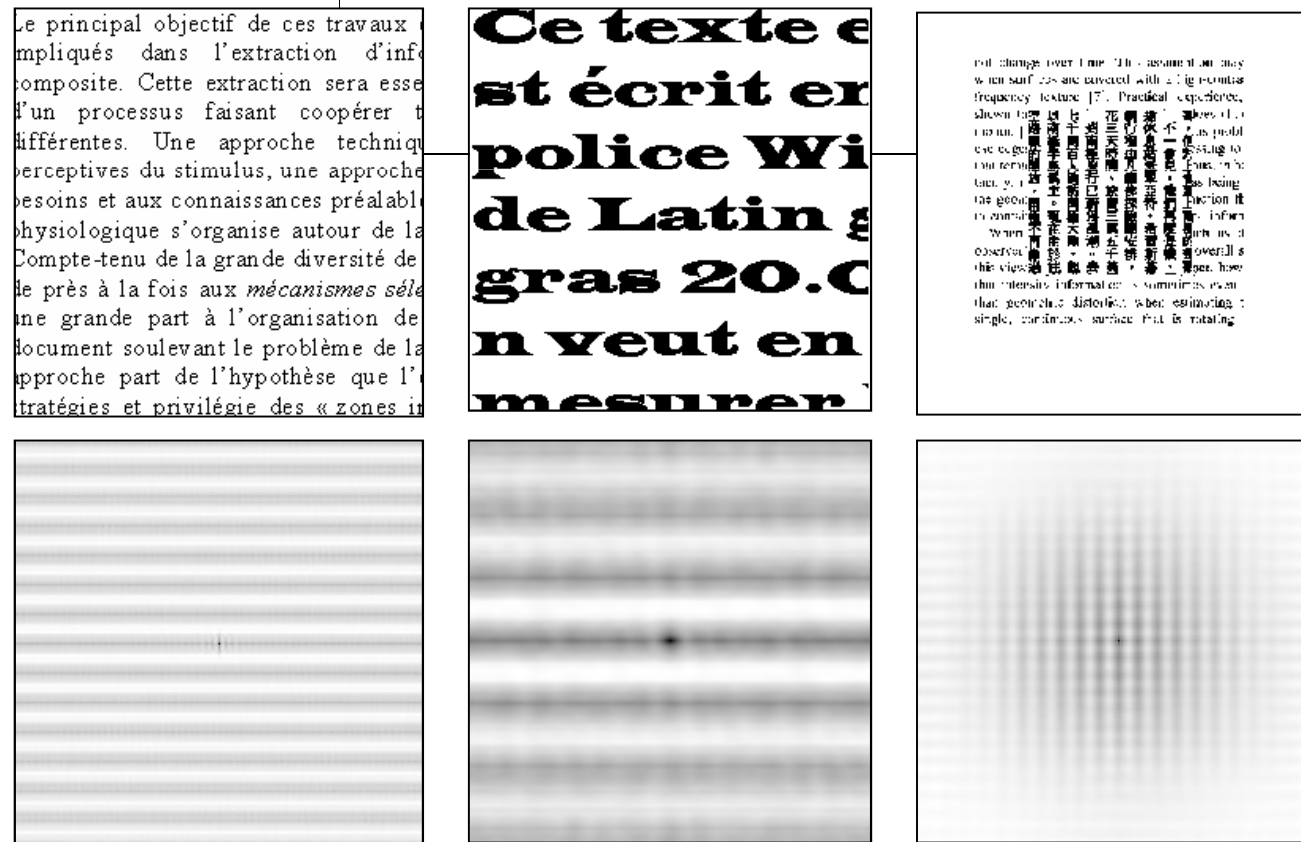
Dans la direction orthogonale à la droite, on a une très faible superposition $\approx$ très faibles valeurs de l'auto-corrélation.



Correction automatique de l'inclinaison

L'auto-corrélation

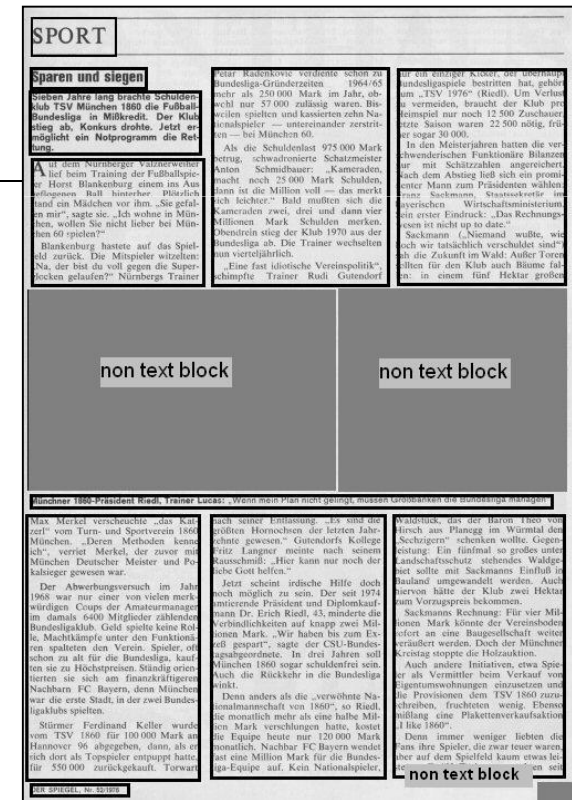
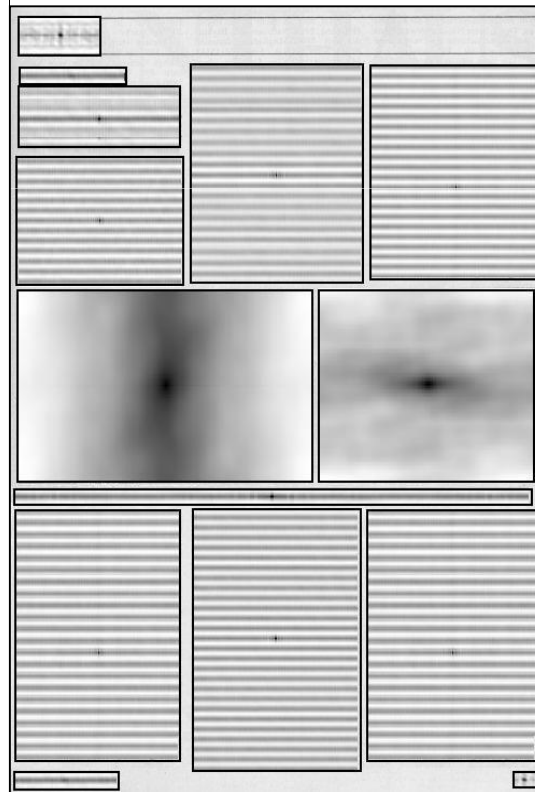
Elle permet de mettre en évidence les périodicités dans les images



Correction automatique de l'inclinaison

L'auto-corrélation

Elle permet de mettre en évidence les périodicités dans les images



Plan

- Quelques bases en traitement des images
 - ◆ Réduire la présence de bruit dans les images
 - ◆ Filtrer les images
 - ◆ Modification du contraste
 - ◆ Diverses corrections
 - ◆ Restaurer les images binaires
 - ◆ Segmenter les formes
 - ◆ Compresser les contenus

Restauration d'images binaires

Le filtre de restauration des caractères ne peut pas s'appliquer aux illustrations !

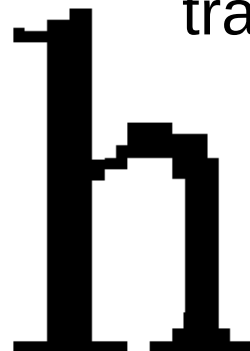
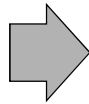
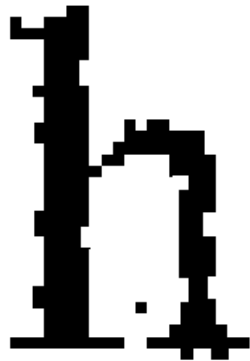
CHRI
BEAV-IE
BEAV-I
Vavaigne D



➡ Il faut séparer le texte des illustrations pour appliquer le filtre que sur des zones de texte

Restauration d'images binaires

Application de masques binaires

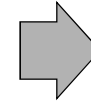
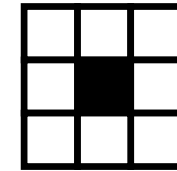


Pixels isolés

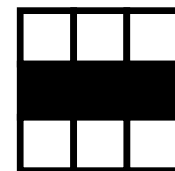
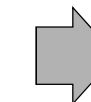
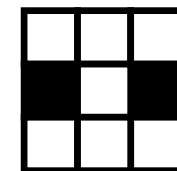
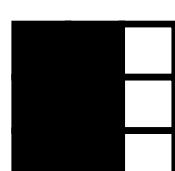
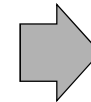
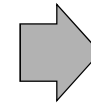
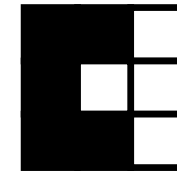
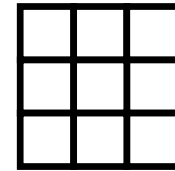
traits Verticaux

Réparation rupture

Si



alors



Etc...

Restauration d'images binaires

Le filtre de restauration des caractères ne peut pas s'appliquer aux illustrations !

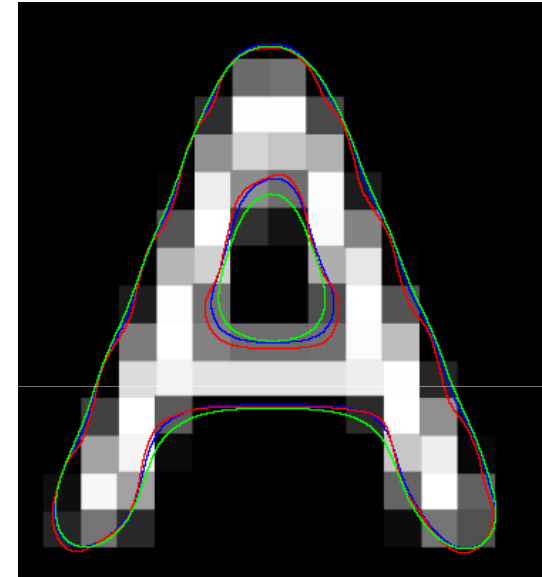
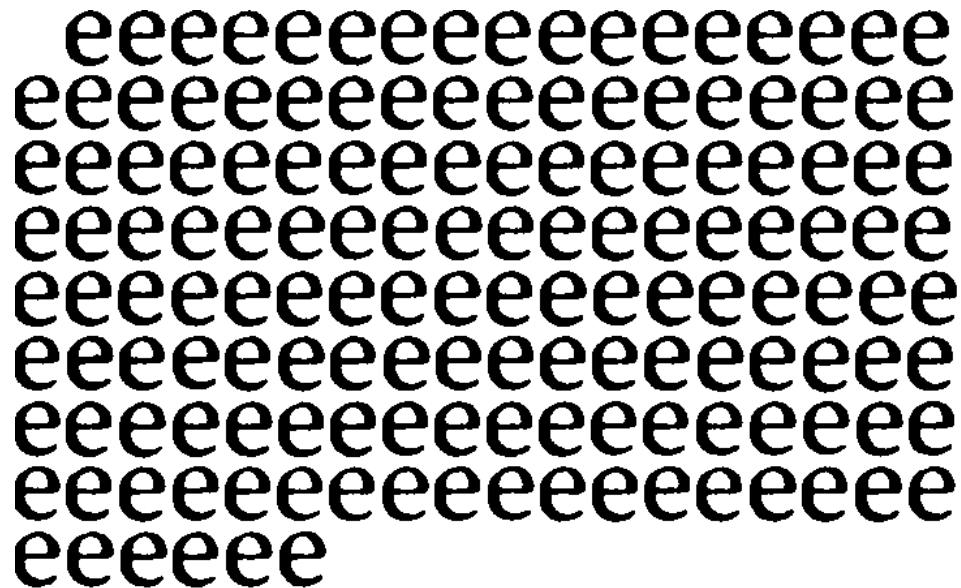
CHRI
BEA V-IE
BEA V-I
Varignon D



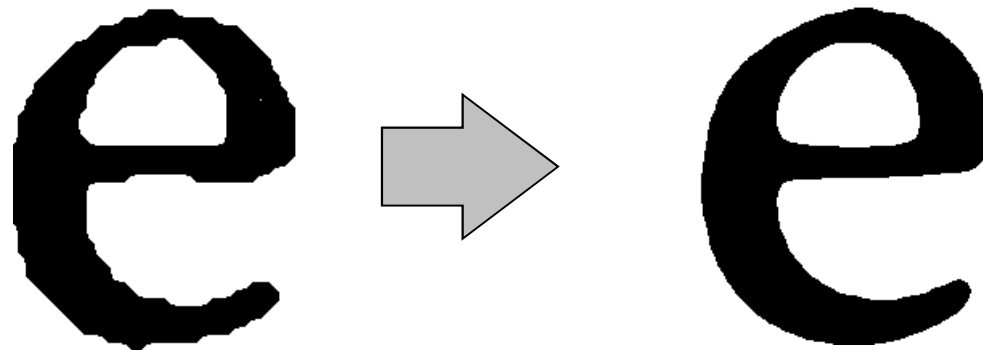
➡ Il faut séparer le texte des illustrations pour appliquer le filtre que sur des zones de texte

Restauration des caractères

Collection des formes identiques de caractères



Construction d'un modèle idéal pour chaque caractère



Remplacement des caractères par le modèle moyen

Plan

- Quelques bases en traitement des images
 - ◆ Réduire la présence de bruit dans les images
 - ◆ Filtrer les images
 - ◆ Modification du contraste
 - ◆ Diverses corrections
 - ◆ Restaurer les images binaires
 - ◆ Segmenter les formes et séparation en couches d'information
 - ◆ Compresser les contenus

Qu'est ce qu'une segmentation ? (1)

- Opération qui consiste à séparer une forme du fond



Pourquoi segmenter ?

- La très grande majorité des outils de reconnaissance utilise des images segmentées pour analyser les formes
- Pour conserver une image de taille réduite tout en conservant sa lisibilité

Qu'est ce qu'une segmentation ? (2)

- ▶ Une bonne segmentation est déjà une étape importante pour la reconnaissance !!!!
- ▶ Une segmentation est un traitement **irréversible** car c'est le résultat d'une **interprétation** suivant un **critère** et une **méthode**.
- ▶ Toute segmentation est le produit d'un choix d'une méthode et de ses paramètres !
- ▶ Plusieurs méthodes de segmentation :
 - ▶ suivant la couleur (seuillage)
 - ▶ à partir d'analyse de formes

Segmentation par seuillage

- ▶ C'est la méthode la plus simple et la plus utilisée
- ▶ Il y a une relation entre les niveaux de gris d'un pixel et son appartenance ou non à une forme



Mais cette relation n'est pas toujours évidente

Segmentation par les nuances de gris

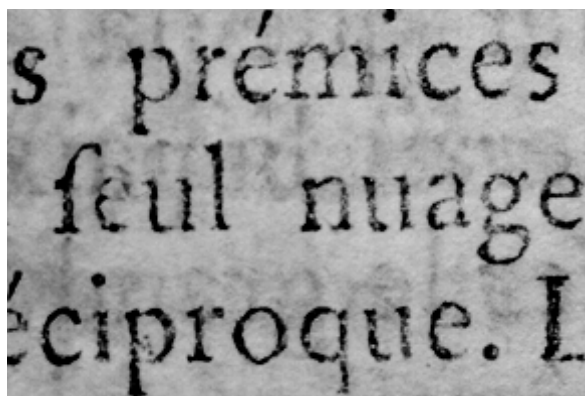
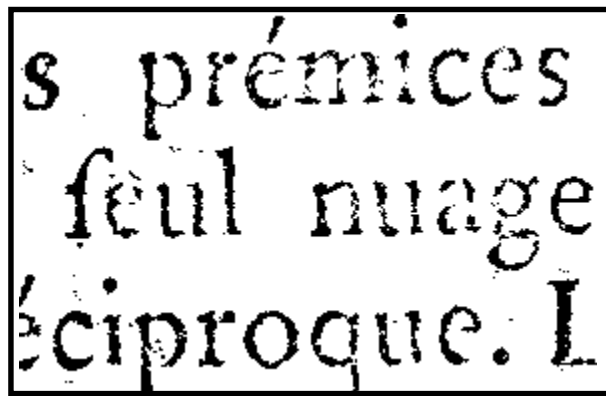
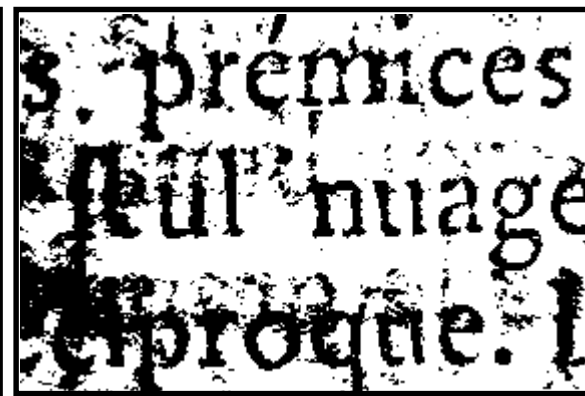


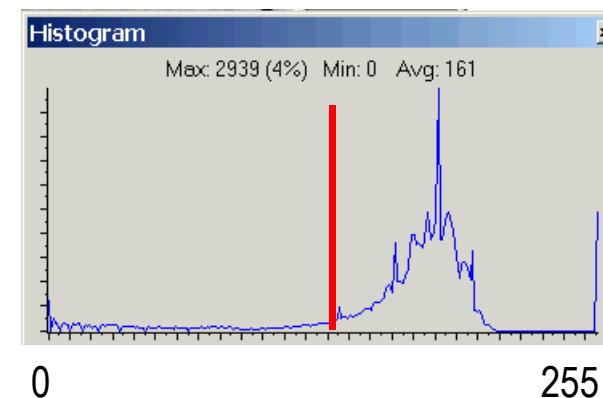
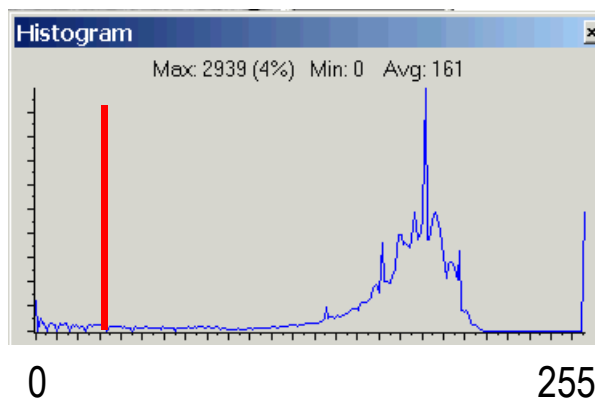
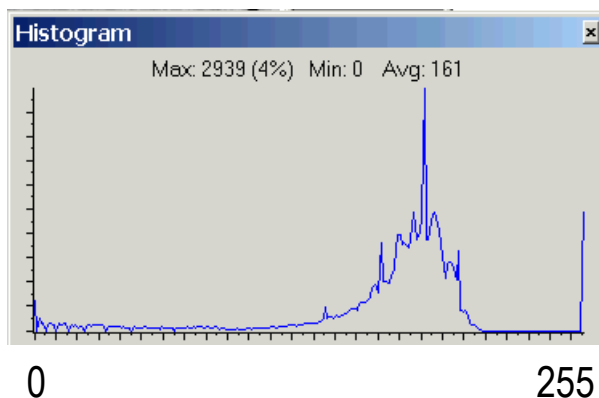
Image à niveaux de gris



seuil trop bas



seuil trop haut



Texte



Support

Certaines nuances de gris se retrouvent
simultanément dans le fond et la forme

Seuillage global vs adaptatif

Seuillage global :

si le pixel $F(x,y) > S$ alors $F(x,y) = 1$ sinon 0

Seuillage Adaptatif :

si le pixel $F(x,y) > S(x,y)$ alors $F(x,y) = 1$ sinon 0

Un niveau de seuil $S(x,y)$ pour chaque coordonnées (x,y)

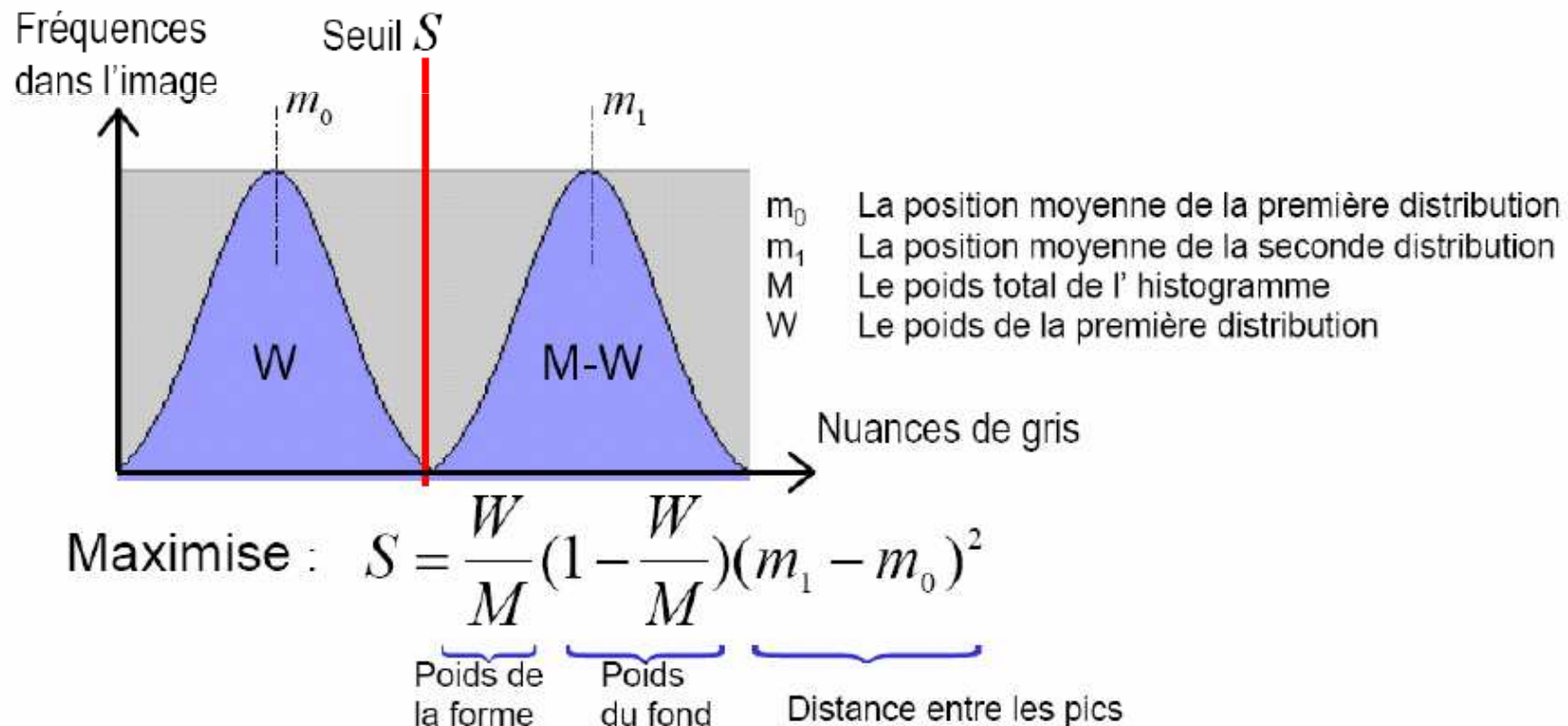
Contrairement au seuil automatique qui est calculé pour l'ensemble de l'image, le seuil adaptatif est calculé pour chaque pixel

Seuillage global vs adaptatif

Seuil Automatique

Seuil global automatique calculé à partir de l'histogramme de l'image

Méthode de Fisher. Fait l'hypothèse que l'histogramme est la mixture de deux distributions gaussiennes (lois normales)

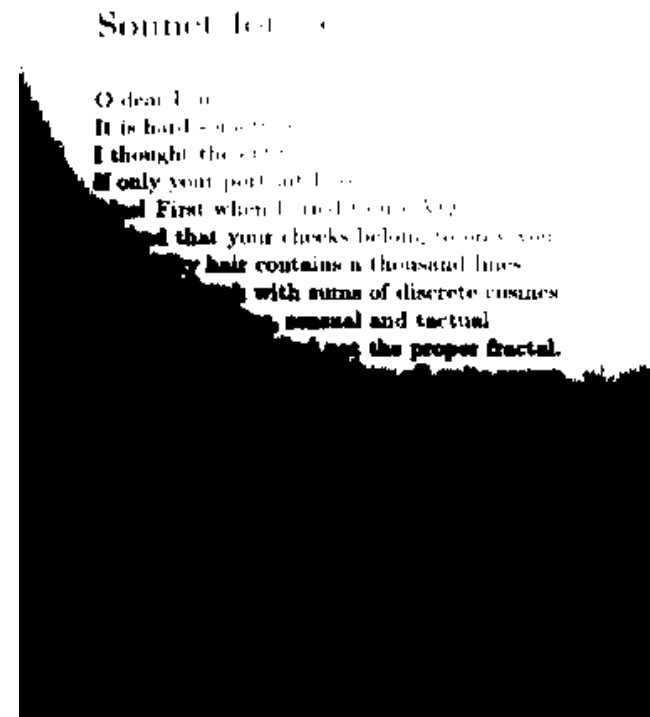
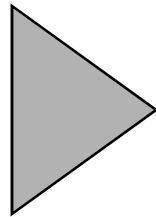
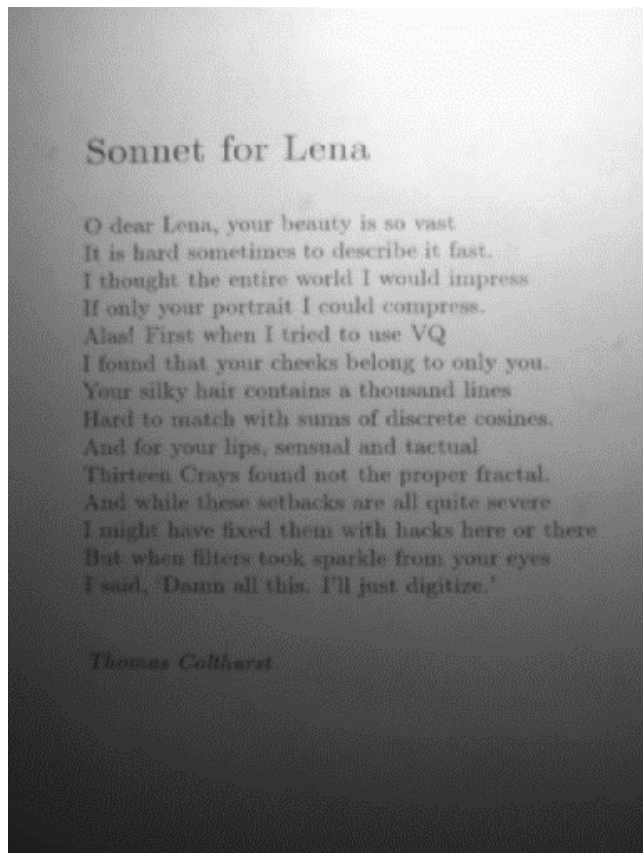


Seuillage global vs adaptatif

Méthode entropique détermine un seuil à partir de l'entropie maximale de la distributions des niveaux de gris h_i (Histogramme)

$$S = \underset{s \in [0..255]}{\text{Arg max}} \left\{ \text{Ln} \left(\sum_{i < S} h_i \times \sum_{i \geq S} h_i \right) - \left(\frac{\sum_{i < S} h_i \times \text{Ln}(h_i)}{\sum_{i < S} h_i} + \frac{\sum_{i \geq S} h_i \times \text{Ln}(h_i)}{\sum_{i \geq S} h_i} \right) \right\}$$

Inconvénient d'un seuil global



Seuillage global

♦ Méthode de OTSU

- Otsu formule le problème comme une analyse discriminante, pour laquelle il utilise une fonction critère particulière comme mesure de séparation statistique

Des statistiques sont calculées pour les deux classes de valeurs d'intensité (**fond** et **print**) séparées par un seuil intensité

On calcule les statistiques pour chaque niveau d'intensité i , c.à.d. pour tous les seuils possibles. Le niveau qui minimise la fonction critère est choisi comme seuil

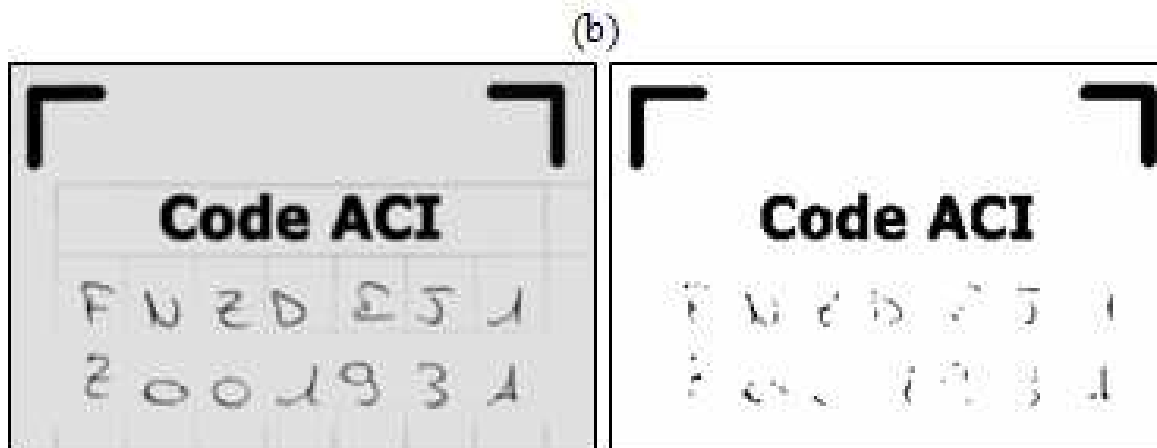
Seuillage global

♦ Méthode de OTSU (suite)

- La mesure du critère discriminant utilisée est :

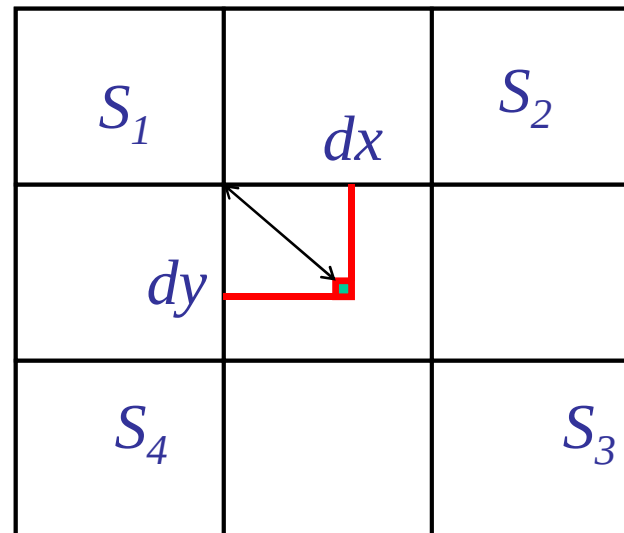
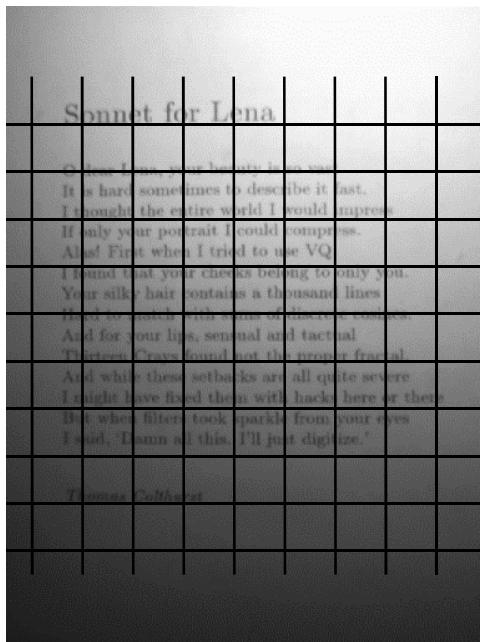
$$\frac{\sigma B_i^2}{\sigma T^2}, i=0, \dots, I-1;$$

où σB_i^2 est la variance inter-classe, σT^2 est la variance totale, et le domaine d'intensité est 0 à $I-1$. L'intensité maximisant cette fonction est le seuil optimal



Seuillage adaptatif

L'image est découpée en bloc NxN puis pour chaque bloc on calcule un seuil automatique par Fisher ou avec une approche entropique



Puis on calcul le seuil de chaque pixel avec une moyenne pondérée des seuils locaux des blocs voisins en fonction de leurs distances

$$S(x, y) = dx.dy.S_1 + dx.(1-dy).S_2 + (1-dx).dy.S_3 + (1-dx)(1-dy).S_4$$

Seuillage adaptatif

◆ Méthode de Niblack

- L'idée de la méthode est de varier le seuil dans l'image en fonction des valeurs de la moyenne locale et de l'écart type local
- Le seuil calculé pour le pixel (x,y) est :
 - $T(x,y) = m(x,y) + k*\delta(x,y)$
où $m(x,y)$ et $\delta(x,y)$ sont respectivement la moyenne et l'écart type calculés dans un voisinage local de (x,y)
- La taille du voisinage doit être suffisamment petite pour préserver les détails locaux, mais suffisamment large pour supprimer le bruit
- La valeur de k est utilisée pour ajuster la partie de l'objet «print» totale considérée comme appartenant à un objet donné

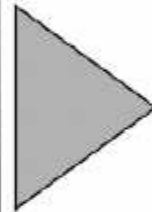
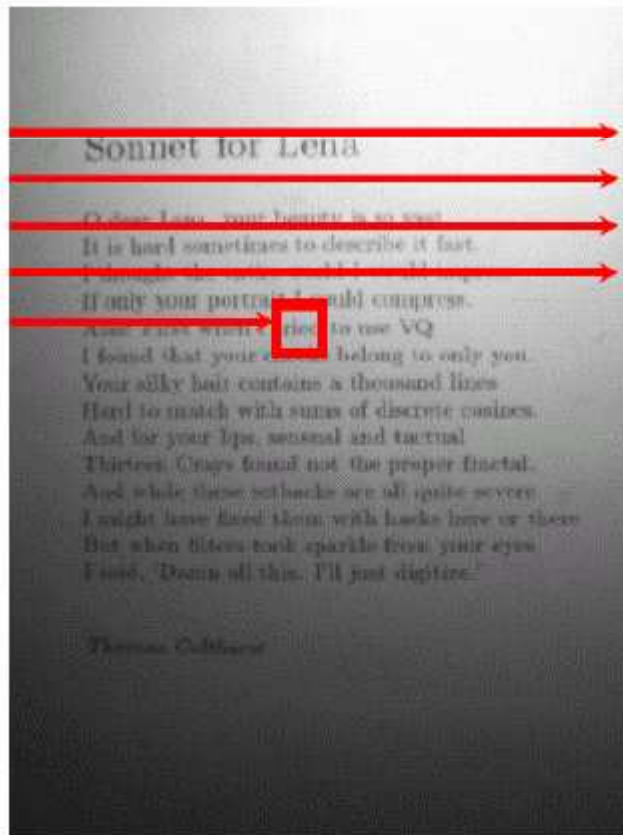
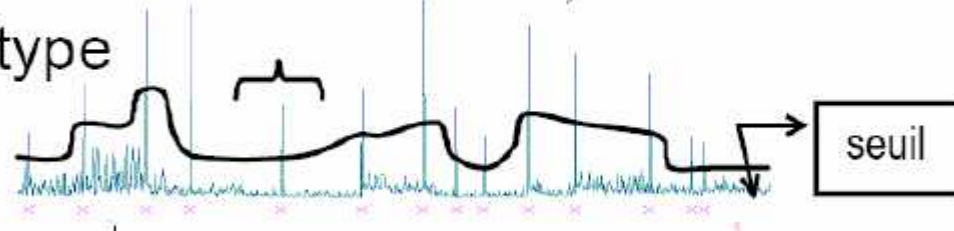
Seuillage adaptatif

- Méthode de Niblack

Exemple :

m : moyenne et σ : l'écart-type

$$S = m + k\sigma^2 \text{ avec } k = -0,2$$



Sonnet for Lena

O dear Lena, your beauty is so vast
It is hard sometimes to describe it fast.
I thought the entire world I would impress
If only your portrait I could compress.
Alas! First when I tried to use YQ
I found that your cheeks belong to only you.
Your silky hair contains a thousand hairs
Hard to match with sums of discrete cosines.
And for your lips, sensual and tactual
Thirteen Crays found not the proper fractal,
And while these setbacks are all quite severe
I might have fixed them with hucks here or there
But when filters took sparkle from your eyes
I said, 'Damn all this. I'll just digitize.'

Thomas Callanvet

Seuillage adaptatif

◆ Méthode de Sauvola

- Niblack ne fonctionne pas sur un fond à texture claire :
 - les valeurs des éléments non désirés peuvent dépasser les seuils. De plus, l'étape de post-traitement est coûteuse
- Modification : tenir compte de la variance locale
 - Seuil proposé : $T(x,y) = m(x,y) * [1 + k * \delta(x,y) / R]$ où R est la dynamique de l'écart type
 - La multiplication des deux termes par la moyenne locale a pour effet d'amplifier la contribution de la variance de manière adaptative

◆ Méthode de Sauvola (suite)

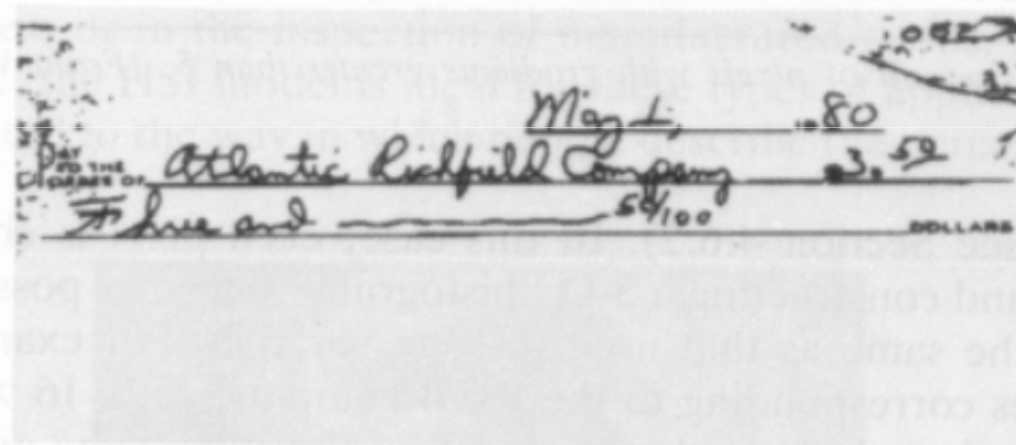
- Si on considère par ex un texte foncé sur un fond clair, mais avec du bruit, m fait décroître la valeur du seuil dans les régions du fond
- L'effet de cette méthode est d'effacer d'une manière efficace le bruit dans une image seuillée

Seuillage adaptatif: Sauvola Exemple





(a)

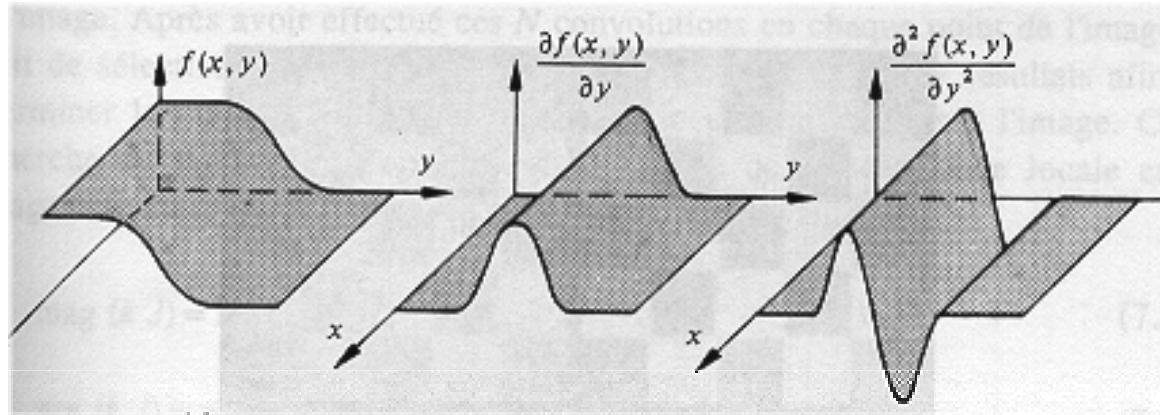


... Et il existe des scanners à seuillage adaptatif grand public
Mais qui garde leur propriété ... et leur mystère !!

Seuil à partir des contours

Une image est une fonction discrète de 2 variables $Z=f(x,y)$ où Z représente la valeur du pixel (x,y)

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}$$



Le vecteur gradient ∇f permet d'étudier les variations dans une image

Approximation grossière du gradient par dérivation

$$\frac{\partial f(x_0, y_0)}{\partial x} = \lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \cong f(x_0 + 1, y_0) - f(x_0, y_0)$$

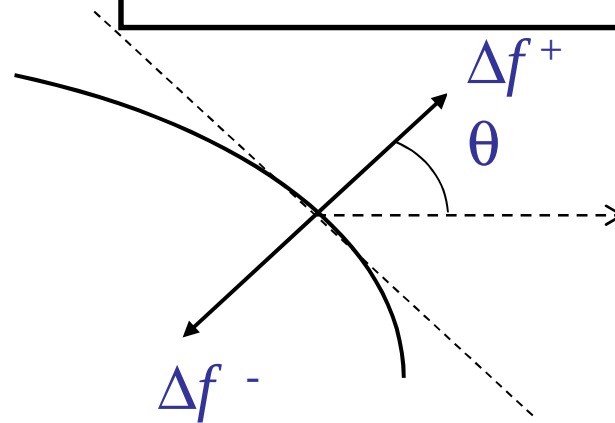
$$\frac{\partial f(x_0, y_0)}{\partial y} = \lim_{k \rightarrow 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k} \cong f(x_0, y_0 + 1) - f(x_0, y_0)$$

Détection des contours par le Gradient

Norme du gradient
(Niveau du contour)

$$\|\nabla f\| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

Orientation du gradient $\theta = \tan^{-1} \left(\frac{\frac{\partial f}{\partial y}}{\frac{\partial f}{\partial x}} \right)$

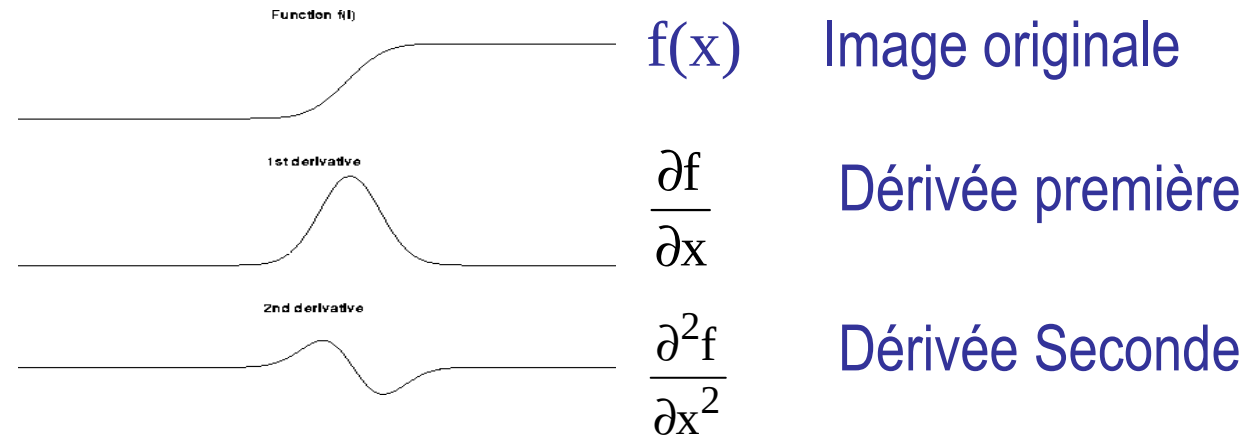


$$\|\nabla f(x, y)\| \approx |f(x+1, y) - f(x, y)| + |f(x, y+1) - f(x, y)|$$

Les contours servent à caractériser
une forme dans une image couleur
ou niveaux de gris



Détection des contours par le Laplacien



Les dérivées Secondes se calculent par la dérivation de l'image dérivé

$$\frac{\partial f}{\partial x}(x-1, y) \approx f(x-1, y) - f(x, y) \qquad \frac{\partial f}{\partial x}(x, y) \approx f(x, y) - f(x+1, y)$$

$$\frac{\partial^2 f}{\partial x^2}(x, y) \approx \frac{\partial f}{\partial x}(x-1, y) - \frac{\partial f}{\partial x}(x, y) \approx f(x-1, y) - f(x, y) - f(x, y) + f(x+1, y)$$

$$\frac{\partial^2 f}{\partial x^2}(x, y) \approx f(x-1, y) - 2 \times f(x, y) + f(x+1, y)$$

$$\frac{\partial^2 f}{\partial y^2}(x, y) \approx f(x, y-1) - 2 \times f(x, y) + f(x, y+1)$$

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \approx f(x-1, y) + f(x+1, y) + f(x, y-1) + f(x, y+1) - 4 \times f(x, y)$$

Seuillage adaptatif par les contours

$$S(x,y) = \begin{cases} 0 & \text{si } \|\nabla f\| < T \\ + & \text{si } \|\nabla f\| > T \text{ et } \Delta f > 0 \\ - & \text{si } \|\nabla f\| > T \text{ et } \Delta f < 0 \end{cases}$$

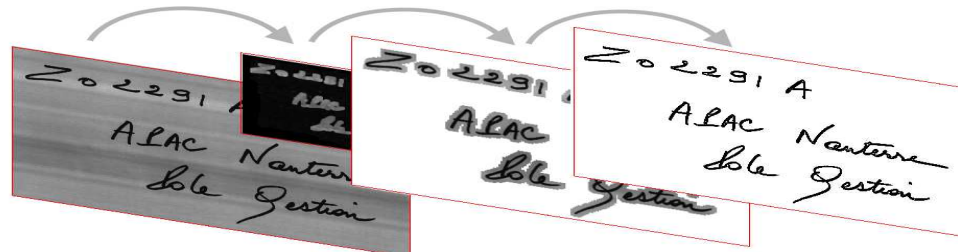
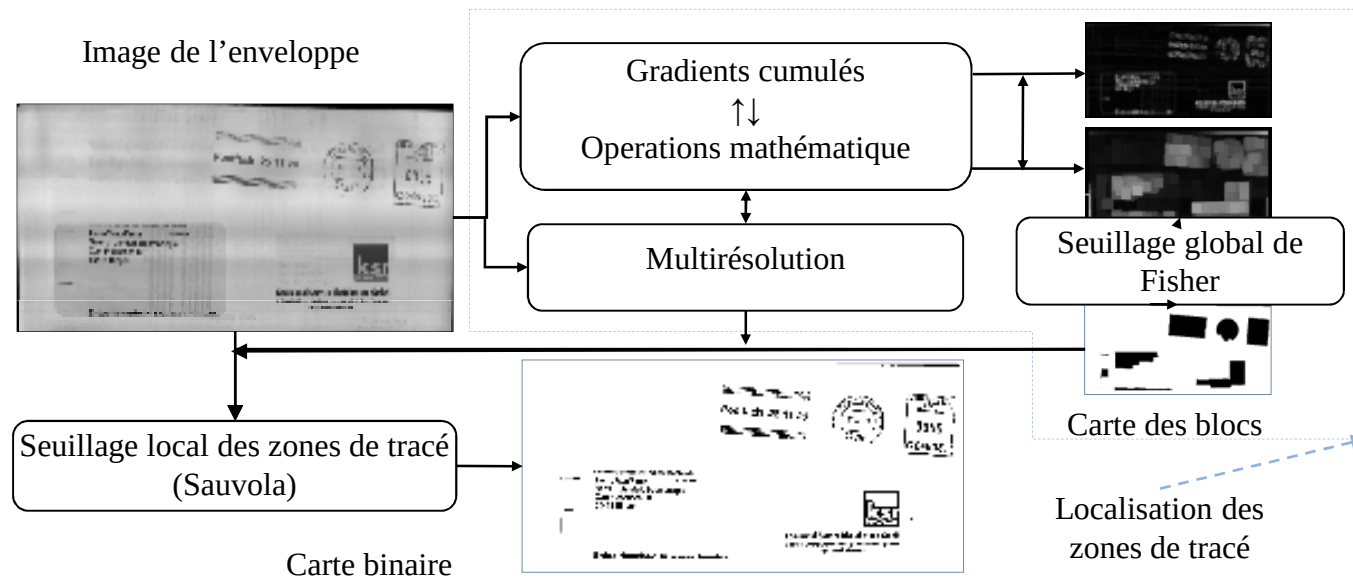
Remplir tous ce qui est entre - + ..0.. + -

[illegible]

Binarisation : seuillage mixte

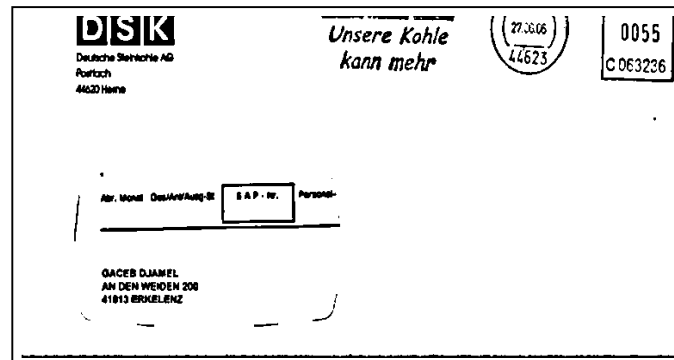
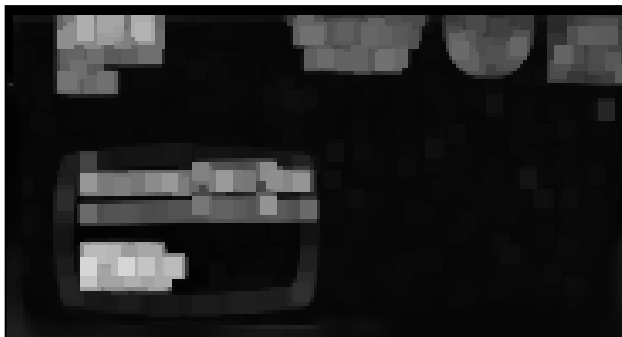
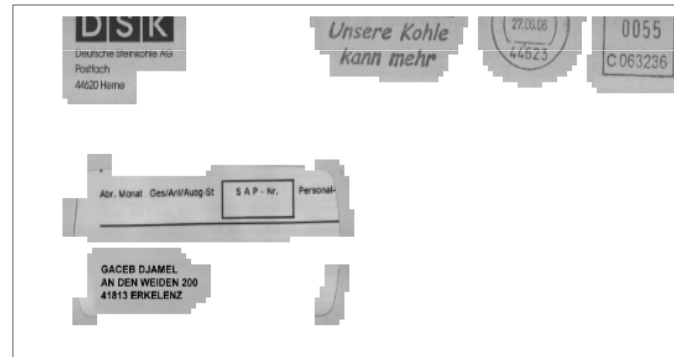
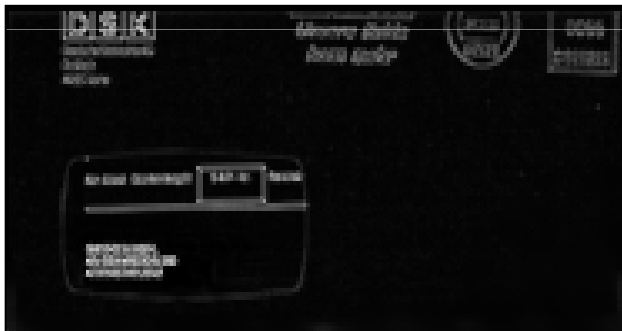
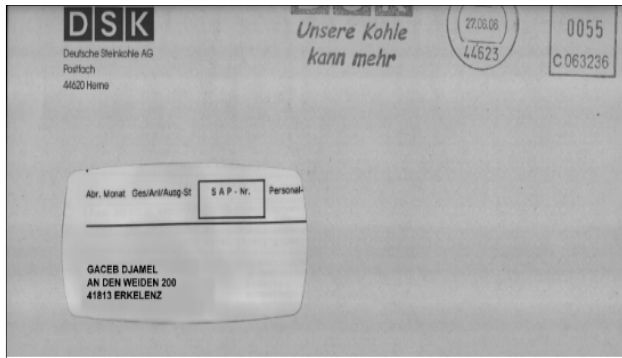
Méthode de seuillage mixte

- 1/ Localisation des zones de tracé
- 2/ Binarisation adaptative des zones de tracé



Binarisation: seuillage mixte

❑ Seuillage mixte (localisation / seuillage adaptatif)



Binarisation: seuillage mixte

❑ Qualité des caractères binaires offerte par notre méthode

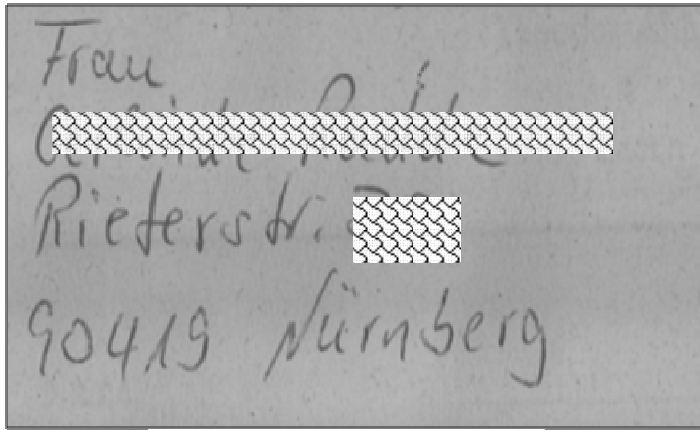
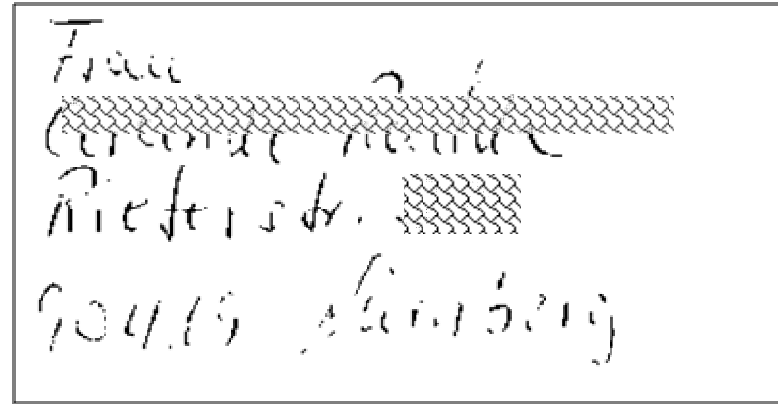
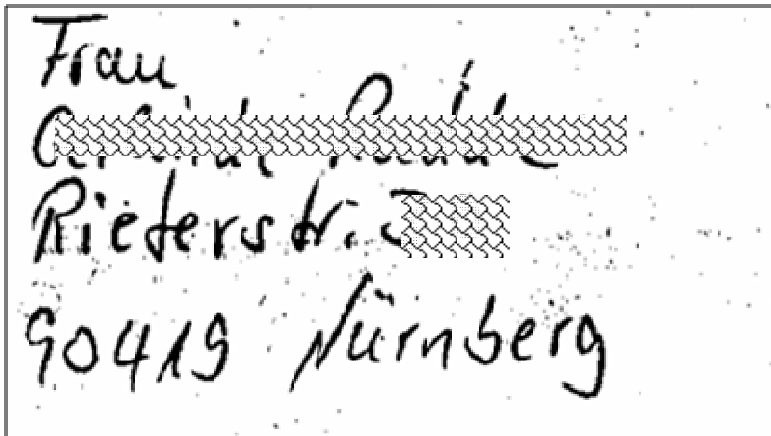


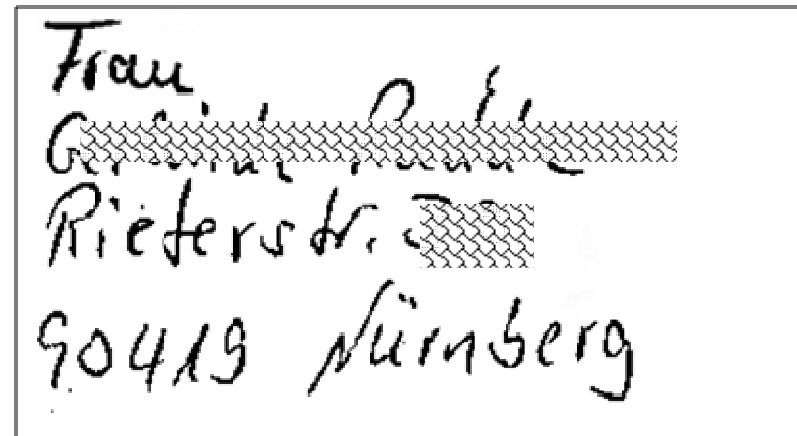
Image en niveau de gris



Méthode globale d'Otsu



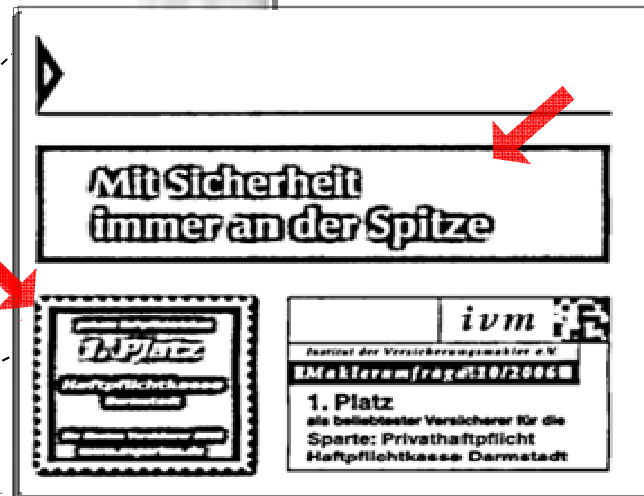
Méthode de Sauvola



Notre méthode (seuillage adaptatif / localisation)

Binarisation: seuillage mixte

- ❑ Réduire les temps de calcul des composantes connexes



Notre méthode de binarisation



Extraction des CCs

$$Temps = \frac{Temps}{3}$$

Et plus généralement sur des images couleurs?

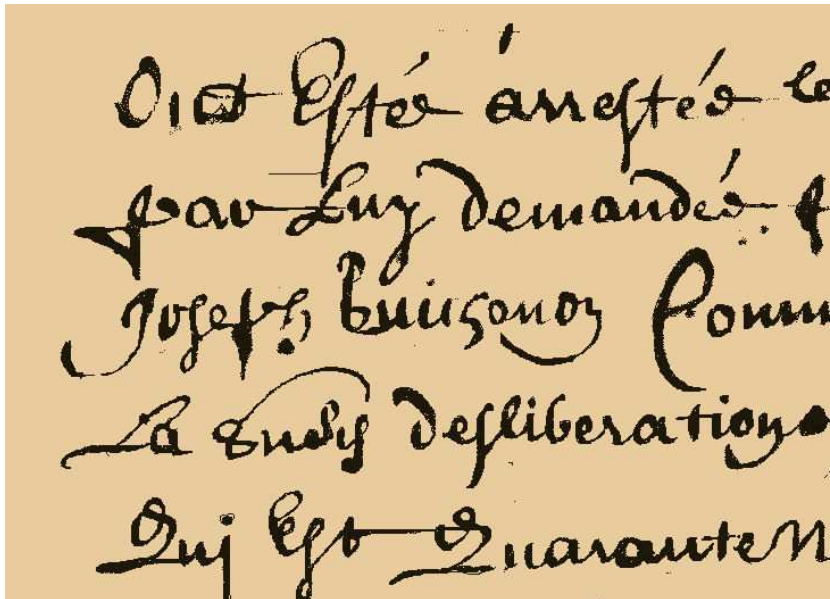
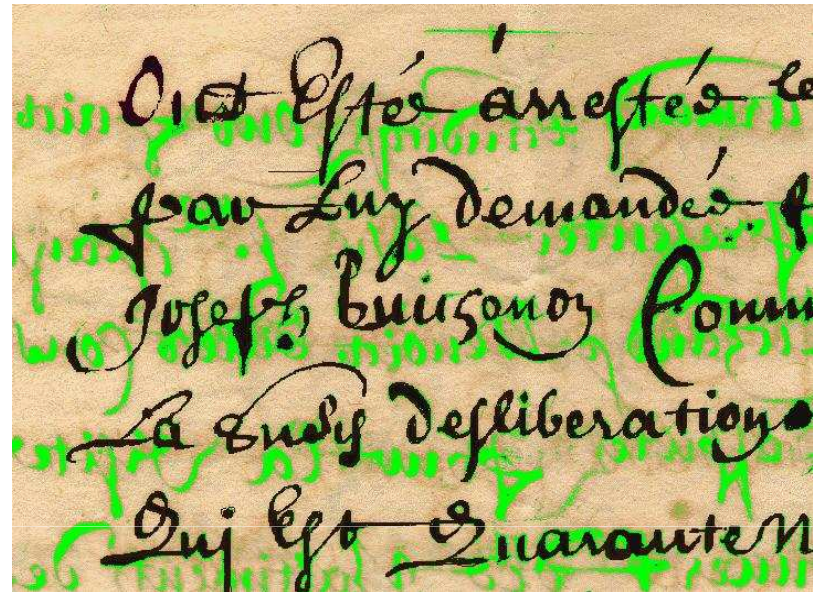
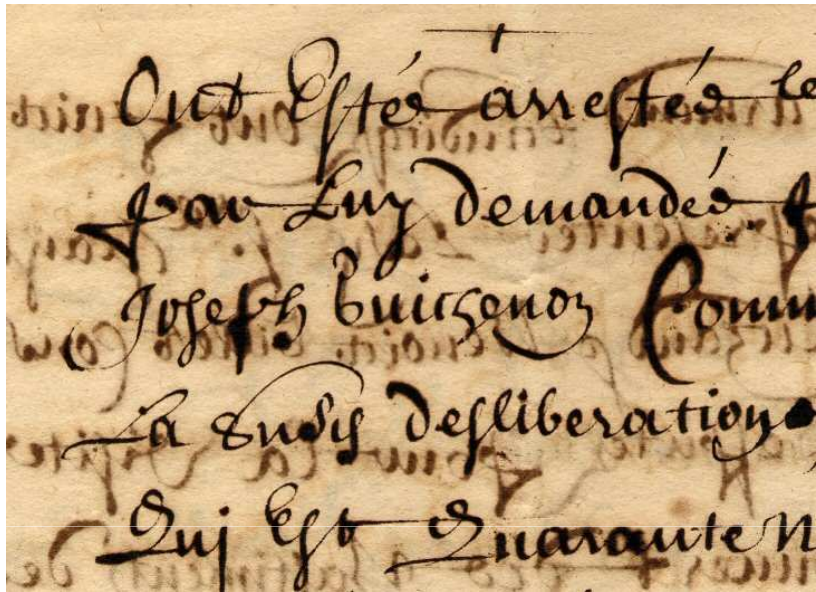


Et plus généralement sur des images couleurs?

- **La couleur des textes est une métadonnée:** *codicologie, papyrologie...*
- **Il n'existe pas encore de méthodes de segmentation d'images couleurs adaptées aux images de documents**
 - *larges images 4000x6000,*
 - *images de traits,*
 - *bruits particuliers autour des fortes variation, bruit JPEG...*
- **La segmentation globale est un échec**
- Il existe en ARD des méthodes locales de binarisation (Niblack, Sauvola...) d'images en niveaux de gris
- Proposition : Extension à la segmentation locale en plusieurs classes d'images couleurs

Segmentation recto/verso

Drira 06



Principe:

- Classification des pixels en trois classes: (1) arrière-plan, (2) texte recto, (3) texte verso.
- Technique de classification de type *k-means* et analyse récursive en composantes principales appliquées sur les données décorréées

Segmentation recto/verso

Principe de la méthode des k-means

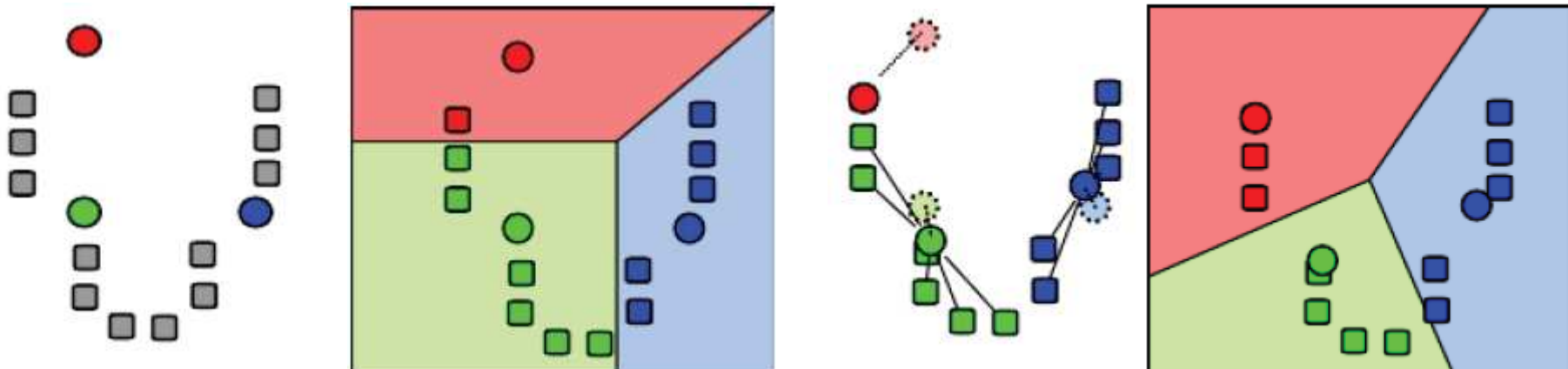
- Appliquer un classifieur non supervisé (le *k-means*) sur une fenêtre glissante pour rendre la segmentation dépendante du contexte
- Permettre au classifieur d'adapter subtilement les nuées à toute variation locale de couleurs.
- *k-means* = *classifieur* non supervisé le plus simple

Segmentation recto/verso

Principe de la méthode des k-means

L'algorithme des K-means est un algorithme simple et efficace de clustering. Son principe est le suivant :

1. choix d'une métrique pour le calcul des distances (euclidienne, manhattan)
2. définition d'un nombre k de classes sur un ensemble d'échantillons
3. initialisation aléatoire des k centres de gravité (centroïdes)
4. affectation de chaque échantillon à son centre le plus proche suivant la métrique choisie
5. calcul des nouveaux centres suivant les affectations effectuées à l'étape précédente
6. répétition des étapes 4 et 5 jusqu'à ce que la position des centres n'évolue plus



Segmentation recto/verso

Principe de la méthode des k-means

Exemple $k=5$



Image originale



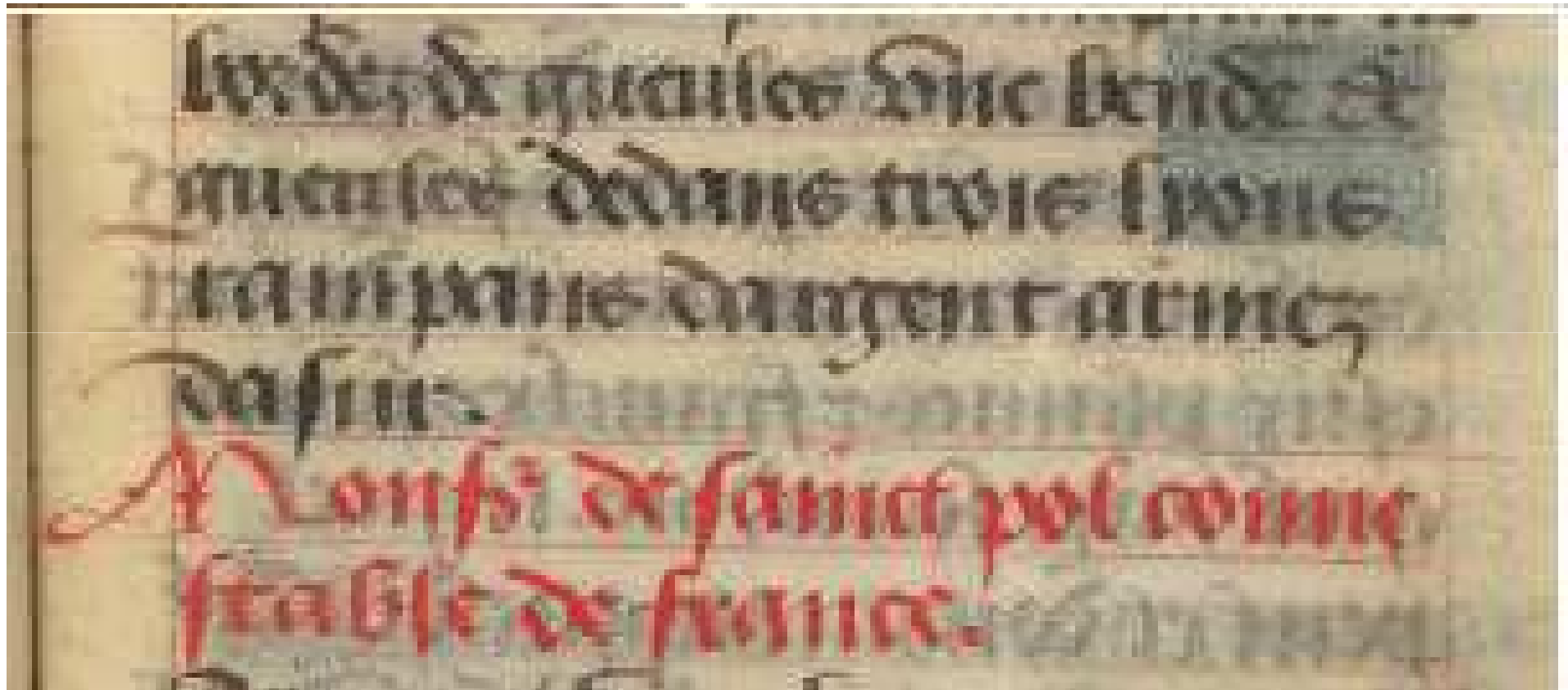
K-means sur le nvg



K-means sur la couleur

Chaque pixel est représenté par la valeur moyenne du cluster auquel il appartient

Segmentation recto/verso



Segmentation recto/verso

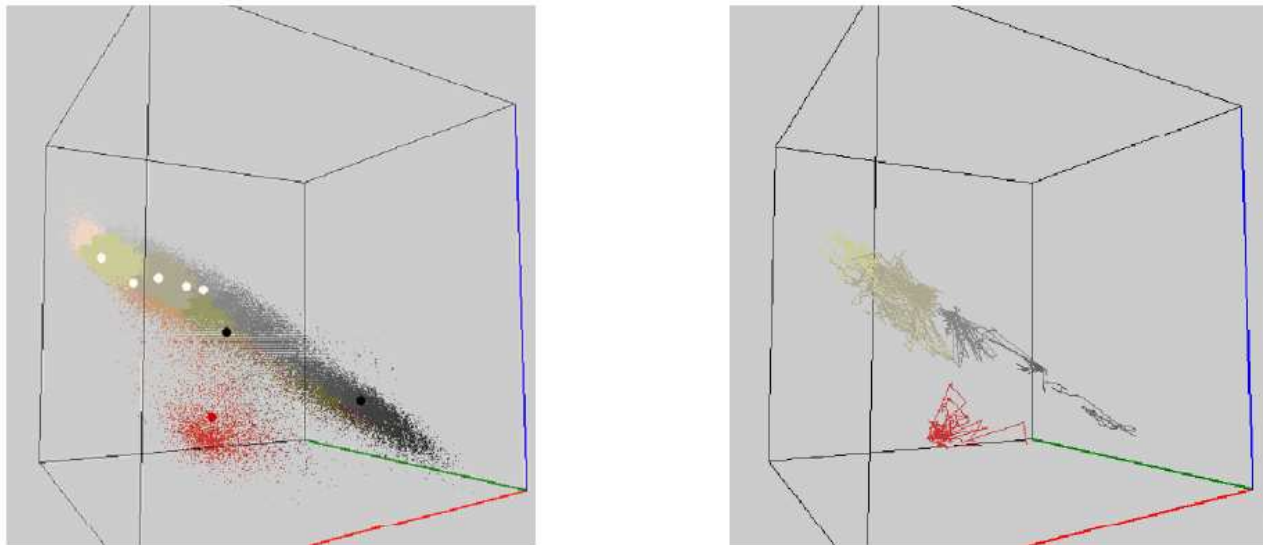


FIG. 2 – a) Couleurs dans l'espace RVB et centres des nuées,
b) Déplacement des centres des nuées le long d'une ligne.

*Nous de saint pol come
stable de france.
Sargent ung lion rampant de g.
a double queue coronne et
arme dor.
Le duc de calabze.*

Segmentation recto/verso

Remandez En presence de sieur
Joseph Guizot, l'un des Conseillers
des Reaux. Et ce à Quoy Il
se Monteront L'era payé au
Nugot par Lesy Cindiq. Et
l'appontant Quintante L'era ser
Entre. Et leur compte signé
à la Notte. L'era Cindiq,
Blancard Et follet Secrétaire

Segmentation recto/verso

Remanded En presence de sieur
Joseph Buisson Luy des Conseillers
des Breues Et ce à Quoy Il
se Monteront Serat payé au
Nugor par Lesdiz Cindiqz Et
Vapportantz Quintaine leur Ser
Entre Et leur Comptes signé
à la Motte Tenein Cindiqz,
Blanzard Et follet Secrétaire,

Segmentation binaire – approche de Sauvola

Segmentation recto/verso

Remanded & presentee de sieur
Joseph Guichenon Luy des Conseillers
des Breues Et ce à Suoy Ih
Se & Monteront Serat gayé aux
Nugos par Ledy Sindiq &
Rapportant Quintanes Leu & era
Entre & leu compte signé
à la Motte Tenevin Sindiq,
Blancard Et Collet Secrétaire

Segmentation couleur – séparation en deux classes

La segmentation d'images par la couleur

La couleur fournit une information supplémentaire qui permet de correctement segmenter les images



Image originale

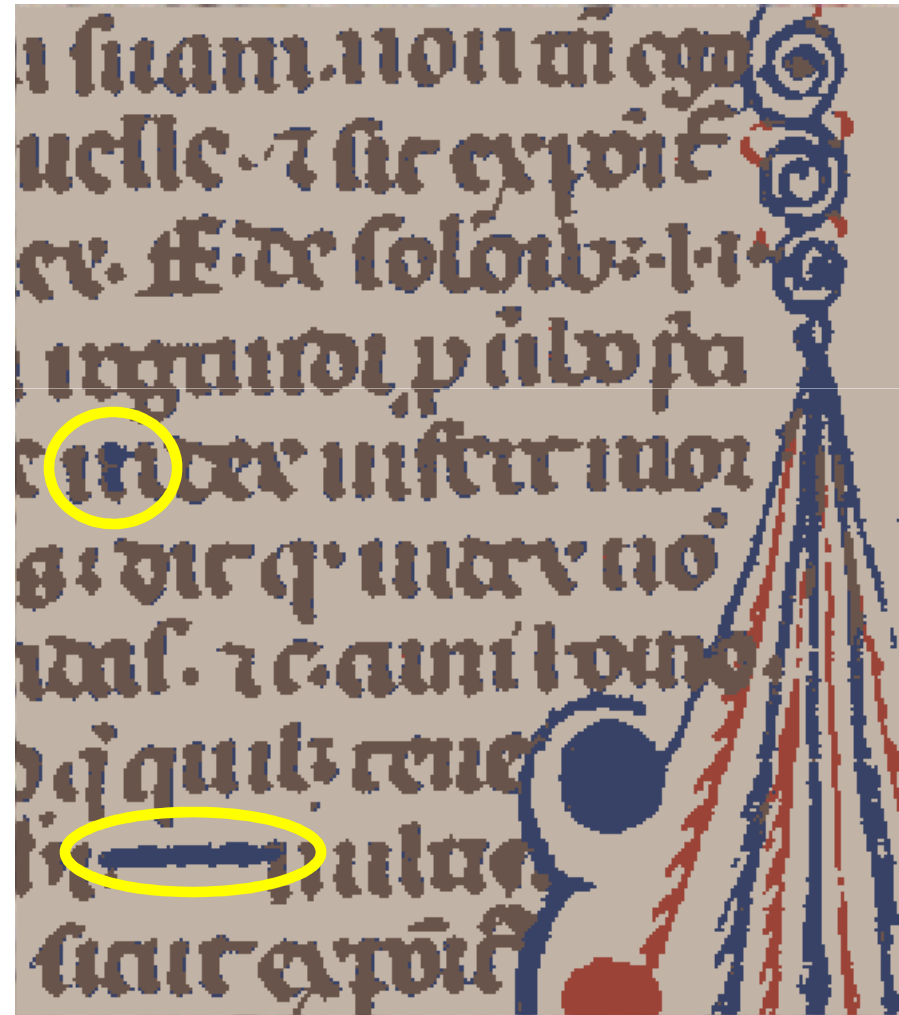


Image segmentée en 4 couleurs

من هذا القول ملكا له جناح بالمشرق والى خير
 بالمغرب ورجلاه مفروقتان في الارض السابعة
 السجلى وعنه ملتوية تحت العرش يقول الله
 الله عز وجل صل على عبدك كما صل على نبيك محمد
 فهو يصل عليه الى يوم القيامة **وروي** عنه صلى الله
 عليه وسلم انه قال يتردى على الحوض يوم القيامة
 اقوام من امر قبهم الا بكثرة صلواتهم على وعنه
 صلى الله عليه وسلم انه قال من صلى على مرته
 واحدة صلى الله عليه **عشر مررات** **في كل صلاة**
 ومن صلى على عشر مررات **اليه** **عليه** **مائة**
 مرة ومن صلى على مائة مرة صلى الله عليه الف مرة ومن
 صلى على الف مرة حرم الله جسده على النار وثبتت بالفول
 الثابت في الحيوة الدنيا وفي الآخرة عنه المسئلة وادخله
 الجنة وجاءت **طائفة** **عليه** **نور** **اليوم** **القيامة** **على** **الصرط**
مسيرته **خمسة** **مائة** **مرا** **وعنه** **الله** **بكل** **عامة** **صالحا**
 فصر في الجنة قل هذا او كثر وقال النبي صلى الله عليه
 وسلم ما من عبد صلى على اخيه حتى الطاعة تدب عن يمينه
 ولا يفسد دونه ولا يجر ولا يشرق ولا يغرب وتقول انا صلاتي

٩٤

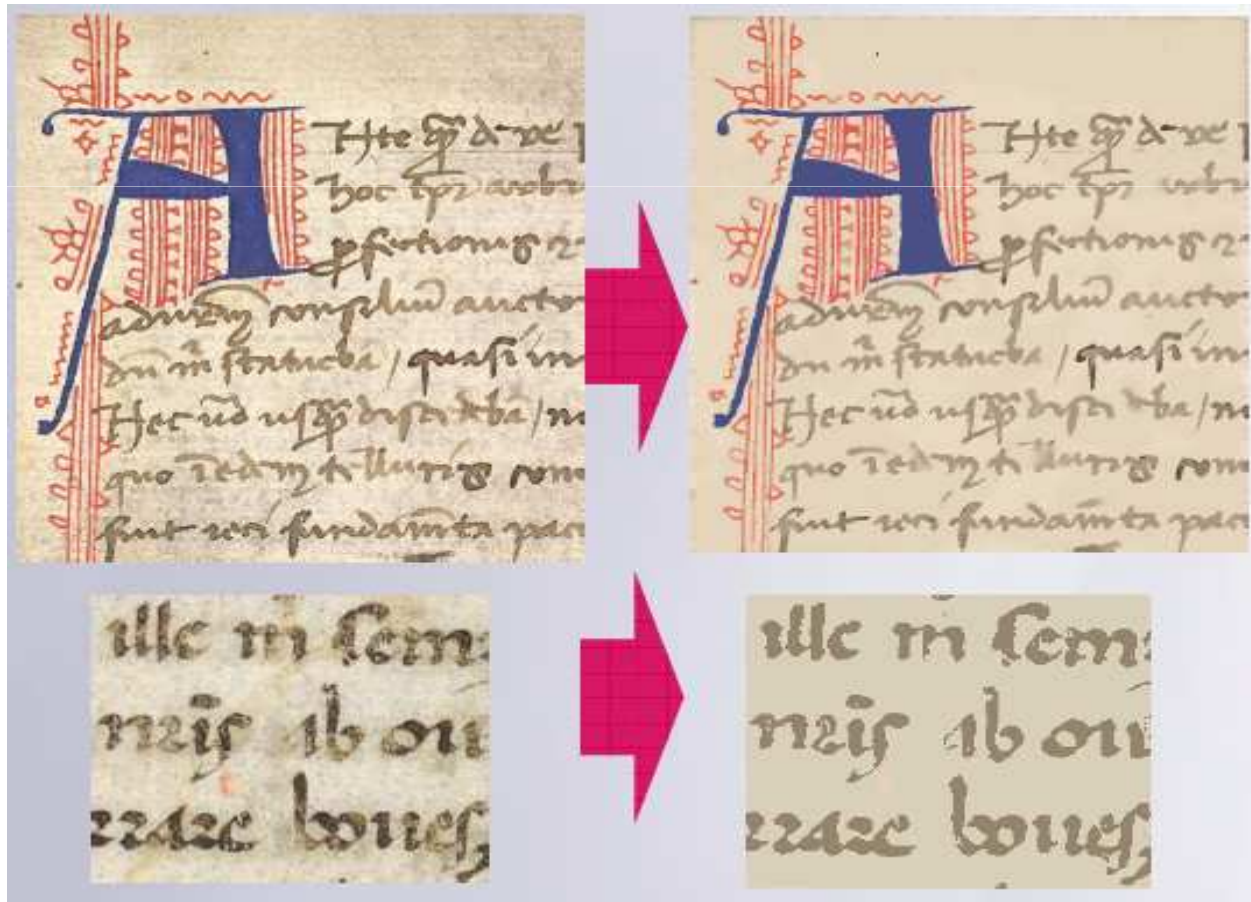
Vers une possible
 reconstruction d'un
 original trop dégradé?

Segmentation ou Reconstruction ?

Par diffusion anisotrope des gradients « couleurs »

image originale

image idéale



Segmentation ou Reconstruction ?

Couplage diffusion et segmentation couleur

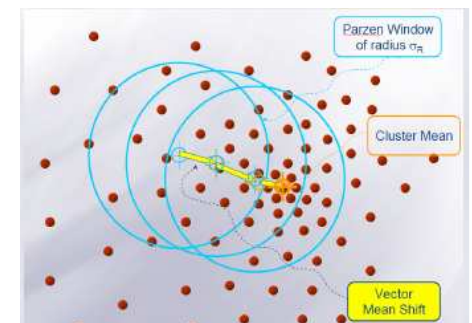
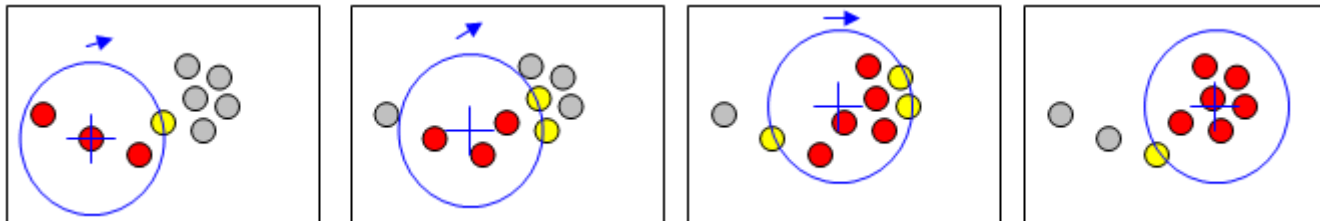
- **MeanShift:** segmentation de l'espace colorimétrique (pas de prise en compte de la proximité spatiale des couleurs dans l'espace)
- **Diffusion:** régularisation de l'image dans le domaine spatial (pas de prise en compte de l'information couleur globale)
- **Couplage:** pour une segmentation des images réduisant les défauts aux contours et arrière plan. Interaction mutuelle des informations couleur et spatiale

Segmentation ou Reconstruction ?

Principe du MeanShift: Faire converger un point vers le maximum local le plus proche :

- Choisir un point de départ P .
- Chercher l'ensemble E des points qui sont dans le voisinage de P .
- Déplacer P vers l'isobarycentre de E .
- Répéter depuis l'étape 2 jusqu'à convergence.

Les déplacements successifs vers l'isobarycentre font converger le point P vers les zones de fortes densités.



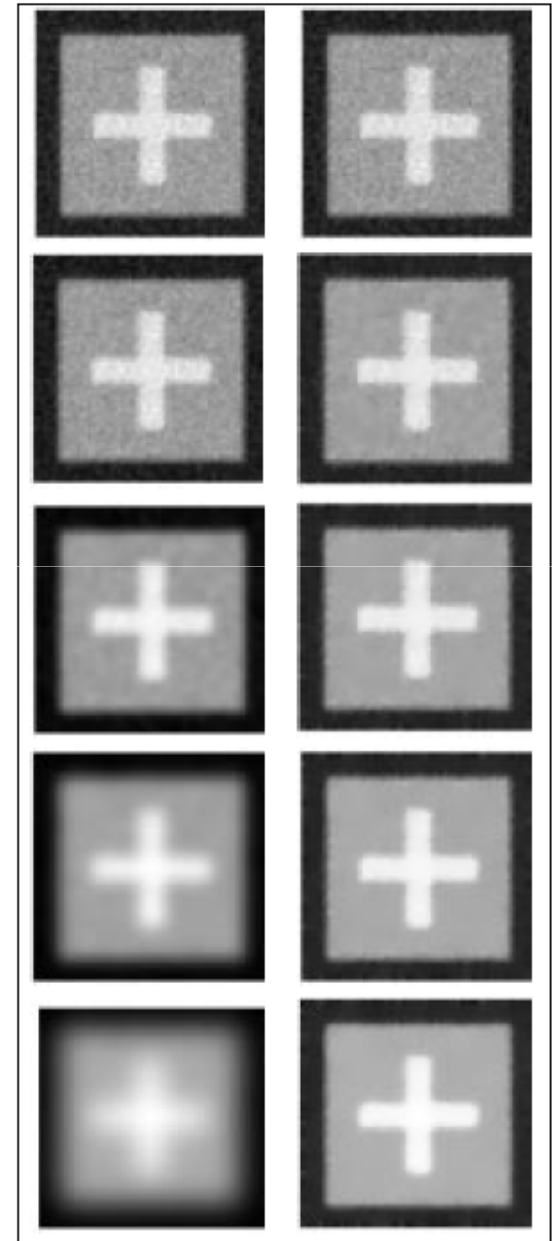
Segmentation ou Reconstruction ?

Principe de la diffusion Perona-Malik:

- Filtrage des régions étendues bruitées (filtrage gaussien) et préservation des contours (priviliégiant certaines directions)

- La différence entre le filtrage isotrope et anisotrope est lié au paramètre (K) qui contrôle le degré d'anisotropie: limiter la diffusion dans les régions à fort gradient

Résultats convolution de l'image par une Gaussienne (gauche), et filtrage anisotrope à une échelle t identique (droite).

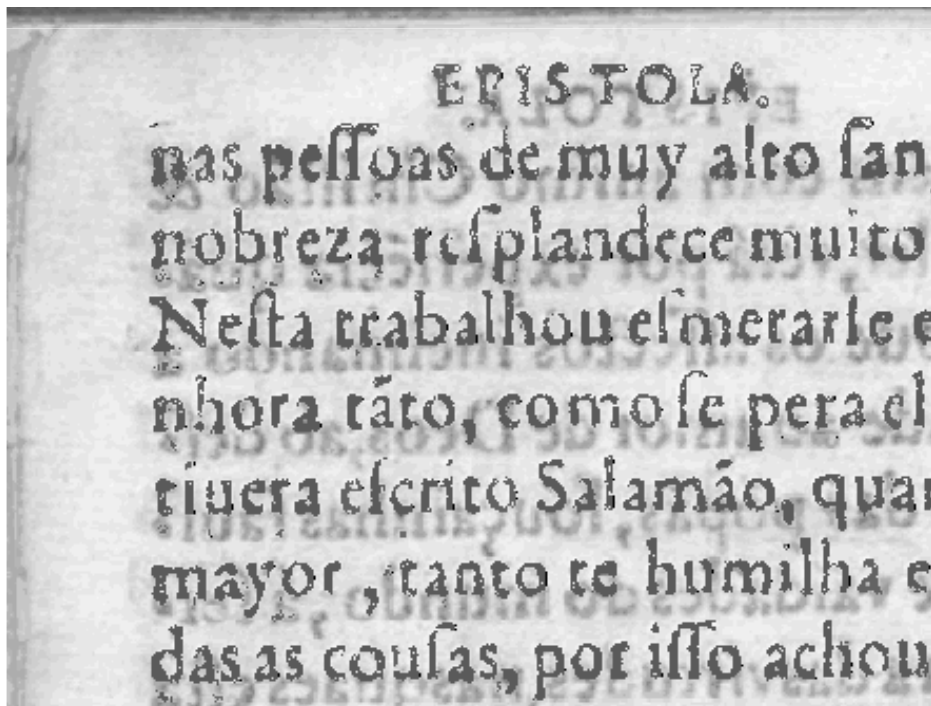


Séparation recto/verso (DEBORA)

► Critère : le taux de redondance des formes de caractères

Une segmentation inadaptée produit des caractères collés ou l'apparition du verso sur le recto ou des caractères cassés.

Une segmentation optimale correspond à un taux maximal de redondance des formes de caractères ! *Un bon critère d'évaluation de la segmentation*

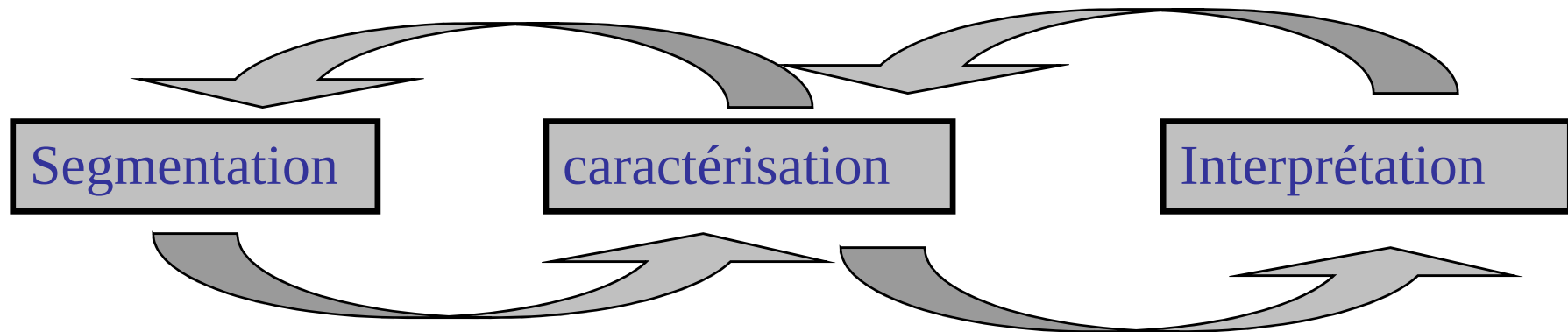


EPISTOLA.
nas pessoas de muy alto san
nobreza resplandece muito
Nesta trabalhou elmerarle e
nhora tão, como se pera el
tiuera escrito Salamão, qual
mayor, tanto te humilha e
das as cousas, por isso achou

Segmentation à partir de connaissances

Il s'agit d'utiliser des connaissances a priori pour segmenter les images de texte avec un degré plus élevé de connaissance

Il s'agit de faire un bouclage d'information entre le processus de segmentation et de caractérisation



A partir de la seule segmentation/caractérisation on peut aller plus loin vers l'interprétation

(séparation texte/graphique, texte/math...)

Bouclage entre segmentation/caractérisation/interprétation

Segmentation par caractérisation du texte

Une **zone de texte** est une région de l'image présentant une très **forte densité de traits** qui forment des **alignements** à une échelle donnée.

Problème d'échelle : comment connaître la taille des caractères a priori ?

Le problème du seuil de détection : A partir de quel seuil de densité de traits faut-il décider de la présence d'un bloc de texte ?

Le problème de l'alignement : Comment connaître a priori l'orientation du texte ? (nécessite la correction de la courbure)

Les Méthodes possibles de détection:

Filtre de texture, filtre morphologique, filtre fréquentiel, Filtre directionnel (Gabor), détection par apprentissage (Markov Random Field)...

Séparation texte/image (1)

Une zone de texte est une concentration de pixels noirs

Ce filtre réagit positivement si il existe dans un voisinage carré $\Delta x \times \Delta y$ au moins un pixel **noir**

$$\begin{cases} \tilde{G}(m,n) = 1 & \text{si } \sum_{i=0}^{i < \Delta y} \sum_{j=0}^{j < \Delta x} f(m \times \Delta x + j, n \times \Delta y + i) \neq 0 \\ \tilde{G}(m,n) = 0 & \text{sinon} \end{cases}$$

4° Comme α est $\mathcal{C}^1$ sur l'ouvert W , l'intégrale de $d\alpha$ le long d'une courbe fermée incluse dans W est nulle. Donc $\int_{C_r} \omega = \int_{C_r} \tilde{\omega} \neq 0$.

Supposons un instant que f n'ait aucun point fixe. Alors, quand le vecteur z parcourt C_r , le vecteur $g(z)$ parcourt une courbe fermée incluse dans V^* . D'après le théorème du relèvement que nous admettrons ici, il existe des applications ρ et γ définies sur $[0, 2\pi]$ telles que

$$g(r \cos \theta, r \sin \theta) = (\rho(\theta) \cos \gamma(\theta), \rho(\theta) \sin \gamma(\theta)), \text{ avec } \gamma(0) = \gamma(2\pi).$$

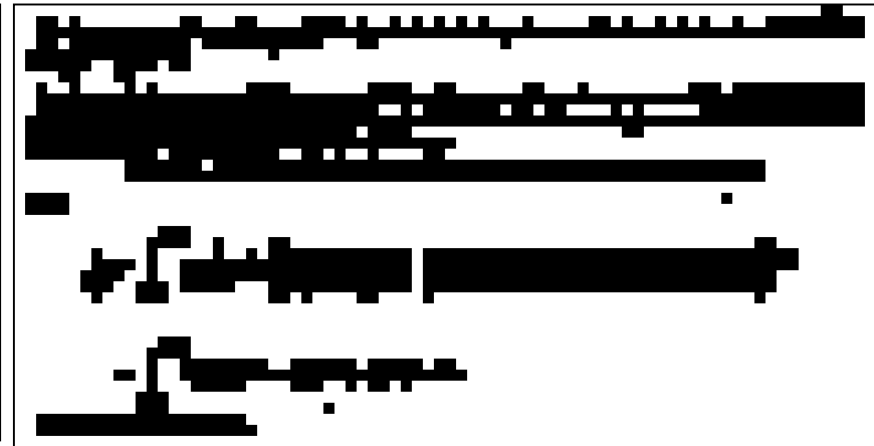
On a :

$$\int_{C_r} \omega = \int_0^{2\pi} \frac{1}{\rho(\theta)^2} \det \begin{pmatrix} \rho(\theta) \cos \gamma(\theta) & \rho'(\theta) \cos \gamma(\theta) - \rho(\theta) \sin \gamma(\theta) \gamma'(\theta) \\ \rho(\theta) \sin \gamma(\theta) & \rho'(\theta) \sin \gamma(\theta) + \rho(\theta) \cos \gamma(\theta) \gamma'(\theta) \end{pmatrix} d\theta$$

$$= \int_0^{2\pi} \gamma'(\theta) d\theta = \gamma(2\pi) - \gamma(0) = 0.$$

Il y a bien une contradiction.

Segmentation obtenue



Filtre de réponse G

Séparation texte/image (2)

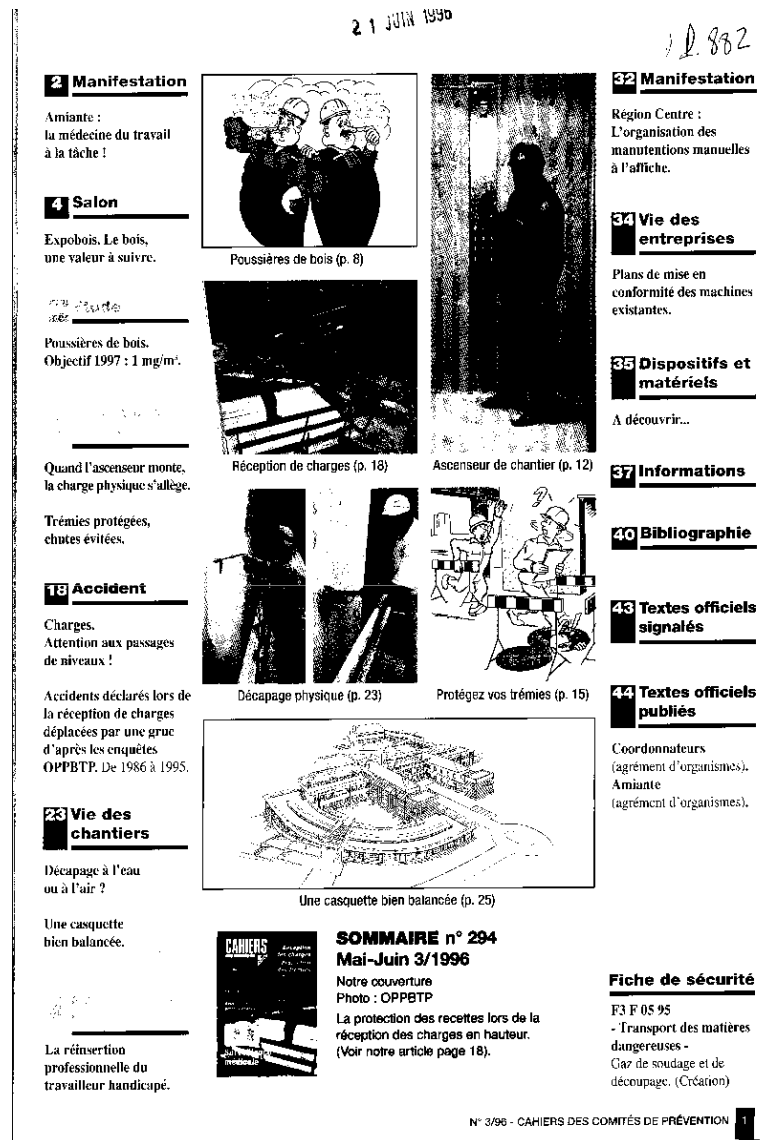
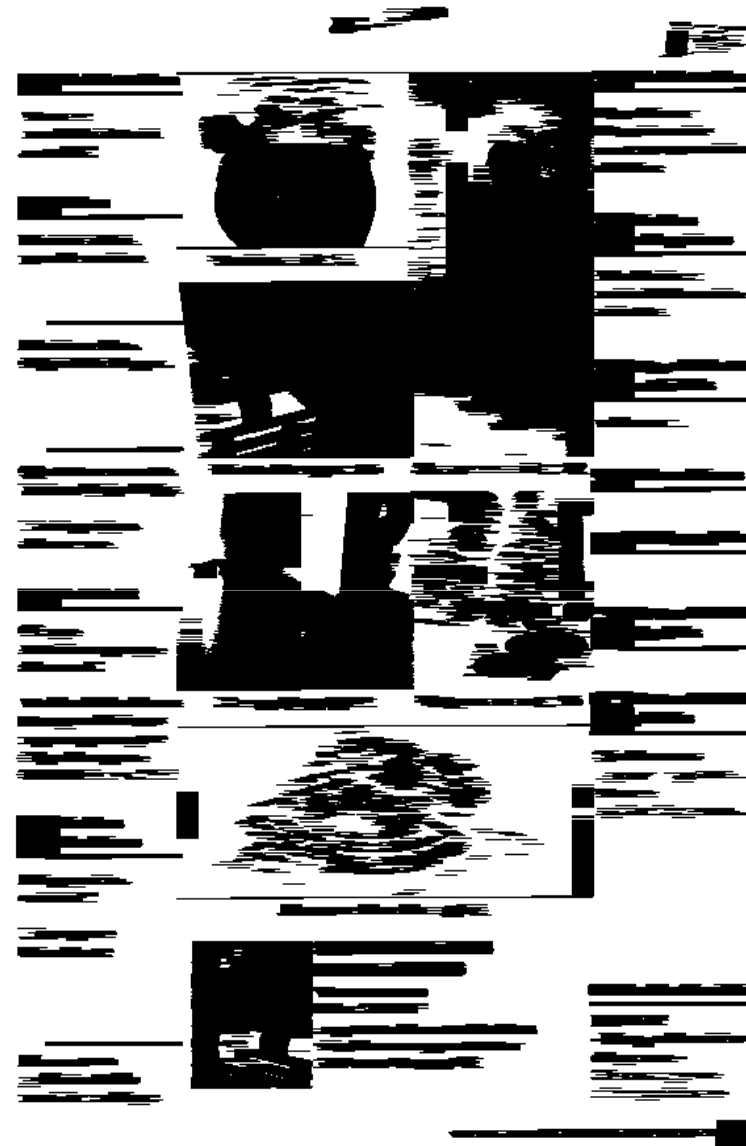
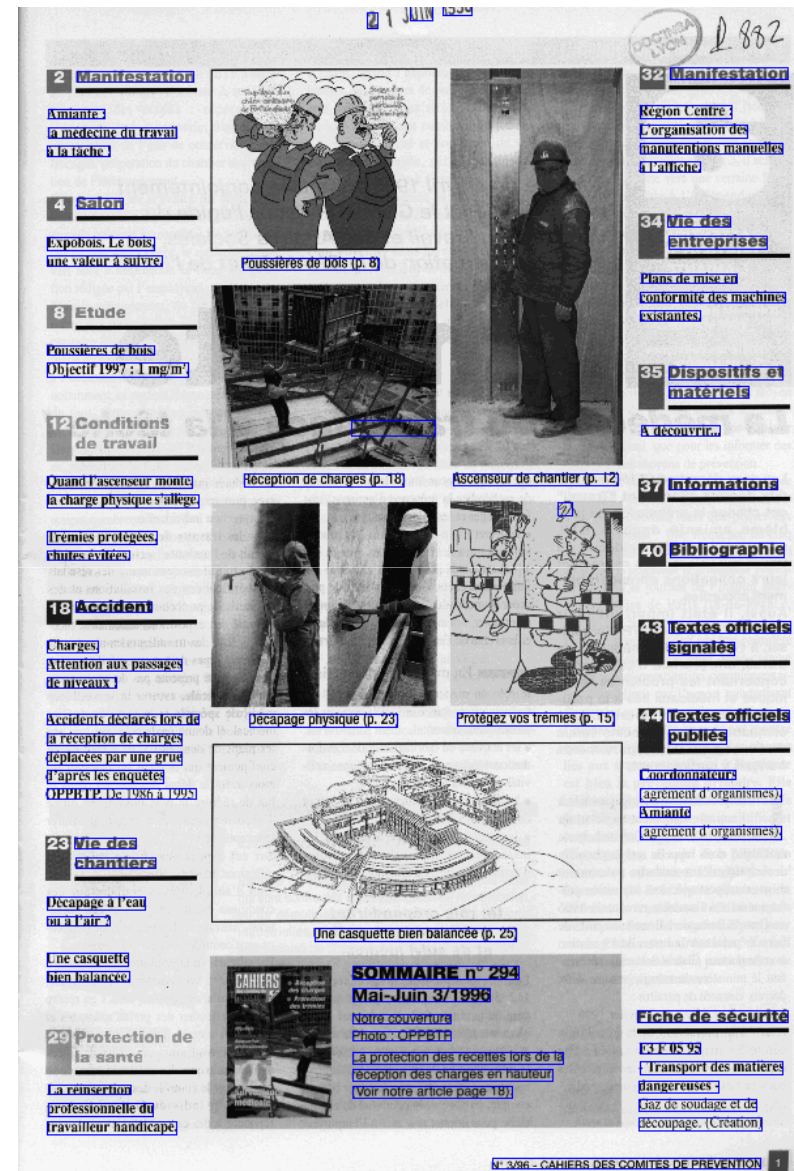


Image binaire



Résultat du filtrage

Séparation texte/image (3)



Détection des alignements par morphologie Résultat de la segmentation

Séparation texte/image (4)

21 JUIN 1996 D. 832

2 Manifestation
Amiante : la médecine du travail à la tâche !

4 Salon
Expobois. Le bois, une valeur à suivre.

8 Etude
Poussières de bois. Objectif 1997 : 1 mg/m³.

Quantité, consommer moins, la charge physique s'allège.

13 Accident
Charges. Attention aux passages de niveaux !
Accidents déclarés lors de la réception de charges déplacées par une grue d'après les enquêtes OPPBTP. De 1986 à 1995.

23 Vie des chantiers
Découpage à l'eau ou à l'air ?
Une casquette bien balancée.

La réinsertion professionnelle du travailleur handicapé.

32 Manifestation
Région Centre : L'organisation des manutentions manuelles à l'affiche.

34 Vie des entreprises
Plans de mise en conformité des machines existantes.

35 Dispositifs et matériels
A découvrir...

37 Informations

40 Bibliographie

43 Textes officiels signalés

44 Textes officiels publiés
Coordonnateurs (agrément d'organismes), Amiante (agrément d'organismes).

Fiche de sécurité
F3 F 05 95
- Transport des matières dangereuses -
Gaz de soudage et de découpage. (Création)

SOMMAIRE n° 294
Mai-Juin 3/1996
Notre couverture
Photo : OPPBTP
La protection des recettes lors de la réception des charges en hauteur. (Voir notre article page 18).

CAHIERS
des comités de prévention
du BTP
N° 3/96

Réception de charges (p. 18)

Ascenseur de chantier (p. 12)

Découpage physique (p. 23)

Protégez vos trémiés (p. 15)

Une casquette bien balancée (p. 25)

Analyse en binaire <<

21 JUIN 1996 D. 832

2 Manifestation
Amiante
la médecine du travail
à la tâche

4 Salon
Expobois. Le bois
une valeur à suivre

8 Etude
Poussières de bois
Objectif 1997 : 1 mg/m³

12 Conditions de travail
Quantité, consommer moins
la charge physique s'allège

18 Accident
Charges
Attention aux passages
de niveaux
Accidents déclarés lors de
la réception de charges
déplacées par une grue
d'après les enquêtes
OPPBTP. De 1986 à 1995

23 Vie des chantiers
Découpage à l'eau
ou à l'air
Une casquette
bien balancée

29 Protection de la santé
La réinsertion
professionnelle du
travailleur handicapé

32 Manifestation
Région Centre
L'organisation des
manutentions manuelles
à l'affiche

34 Vie des entreprises
Plans de mise en
conformité des machines
existantes

35 Dispositifs et matériels
A découvrir...

37 Informations

40 Bibliographie

43 Textes officiels signalés

44 Textes officiels publiés
Coordonnateurs
(agrément d'organismes)
Amiante
(agrément d'organismes)

Fiche de sécurité
F3 F 05 95
Transport des matières
dangereuses
Gaz de soudage et de
découpage. (Création)

SOMMAIRE n° 294
Mai-Juin 3/1996
Notre couverture
Photo : OPPBTP
La protection des recettes lors de la réception des charges en hauteur. (Voir notre article page 18).

CAHIERS
des comités de prévention
du BTP
N° 3/96

Réception de charges (p. 18)

Ascenseur de chantier (p. 12)

Découpage physique (p. 23)

Protégez vos trémiés (p. 15)

Une casquette bien balancée (p. 25)

Analyse en niveaux de gris

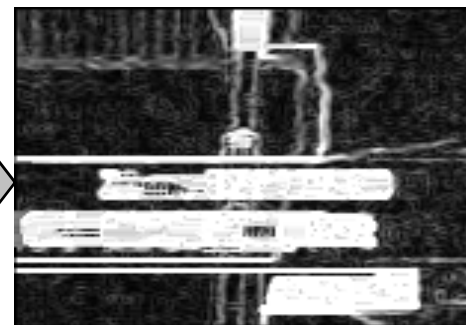
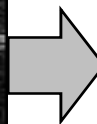
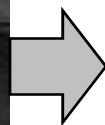
Séparation Texte/image (5)

Ce Filtre répond à une concentration de fortes variations de l'intensité lumineuse de l'image dans un voisinage V

$$G(x_0, y_0) = \iint_{V(x_0, y_0)} \left| \frac{\partial f(x, y)}{\partial \vec{v}} \right| dx dy$$

- ▶ Les fortes variations sont mesurées par les **dérivées partielles**
- ▶ Faible complexité de calcul, vitesse élevée de traitement (10")
- ▶ La taille du voisinage est définie par l'échelle d'analyse.
- ▶ Le résultat du filtre est normalisé par la surface du voisinage puis comparé à un seuil global.
- ▶ Pour les images numériques, on effectue une double sommation de dérivées d'ordre 1 dans la direction de l'orientation du texte

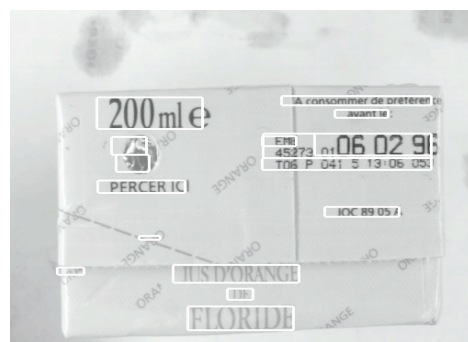
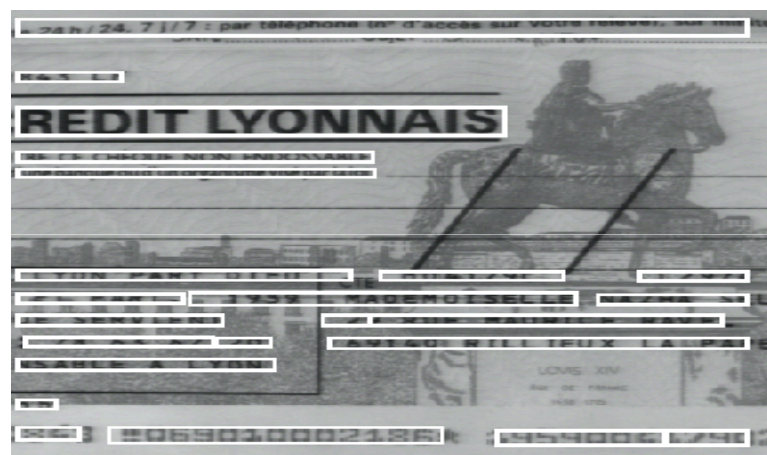
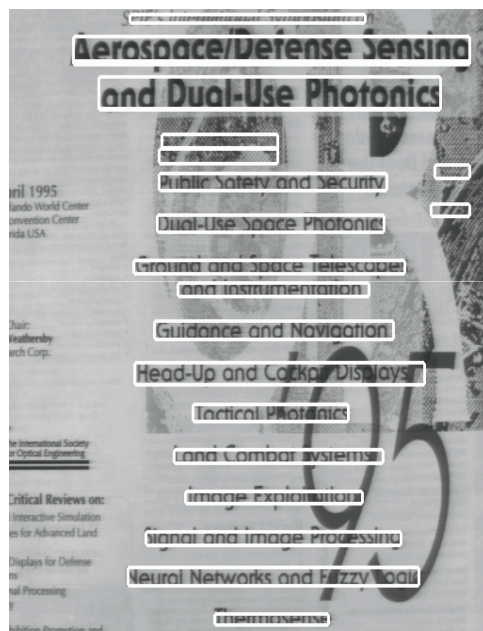
$$G(x_0, y_0) = \frac{1}{\text{Card}(V(x_0, y_0))} \sum_{(x, y) \in V(x_0, y_0)} \sum_{\vec{v}} |f((x, y) + \vec{v}) - f(x, y)|$$



Texte

Contours (dérivées)

Intégration des dérivées



Séparation Texte/image (6)

Nécessité de corriger le résultat du filtrage par morphologie pour reconstituer les lignes (filtre de détection d'alignements)

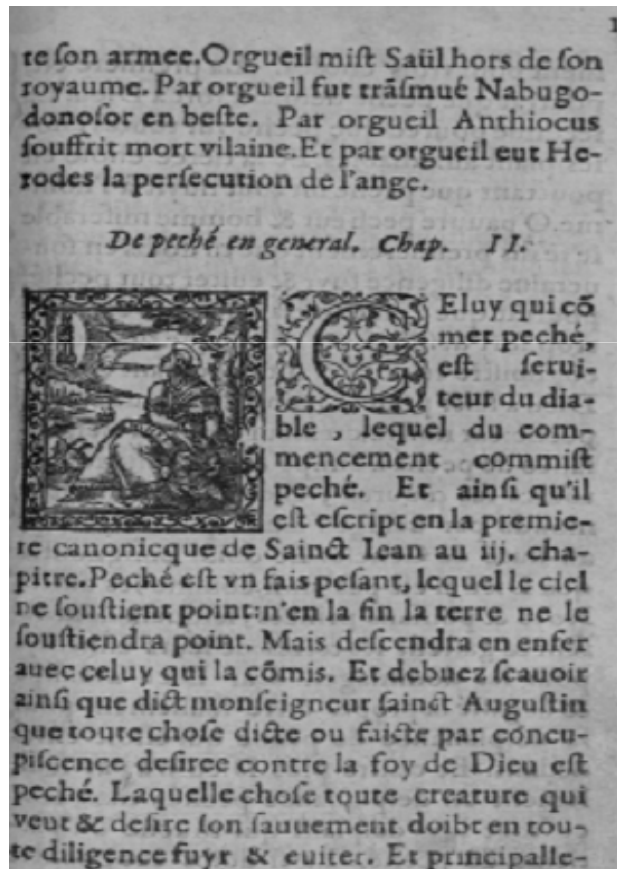
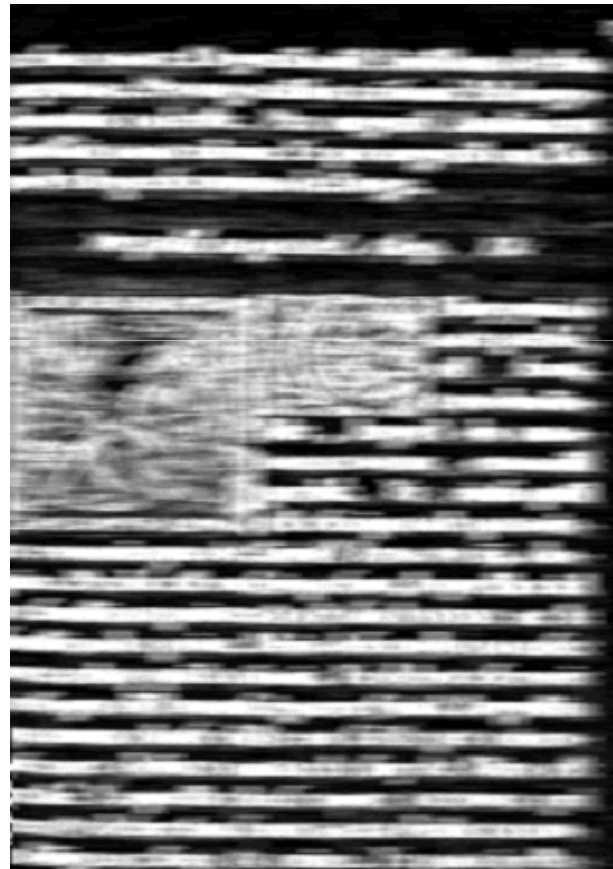
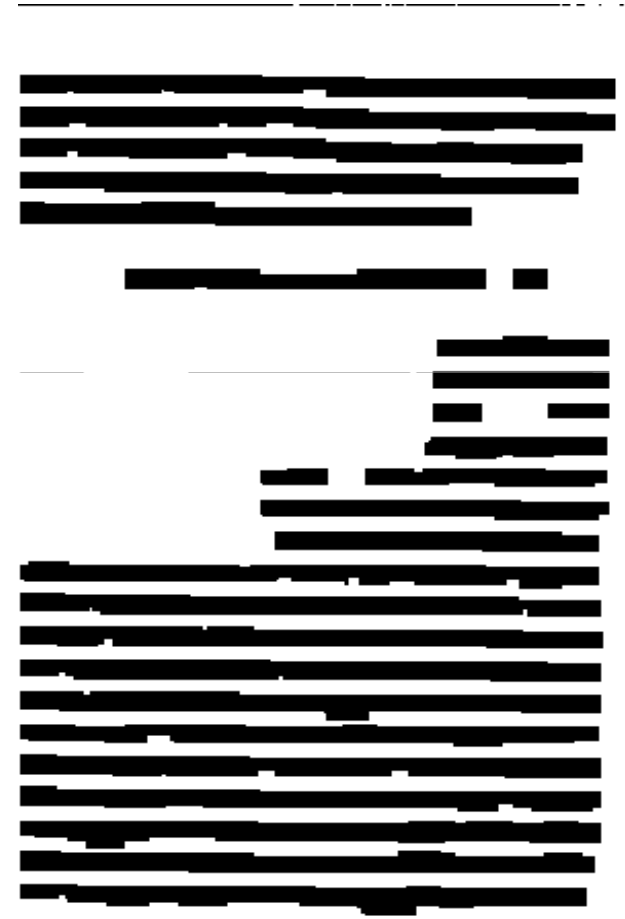


Image originale



Résultat du filtre G



Détection des lignes
par morphologie

Séparation Texte/image (7)

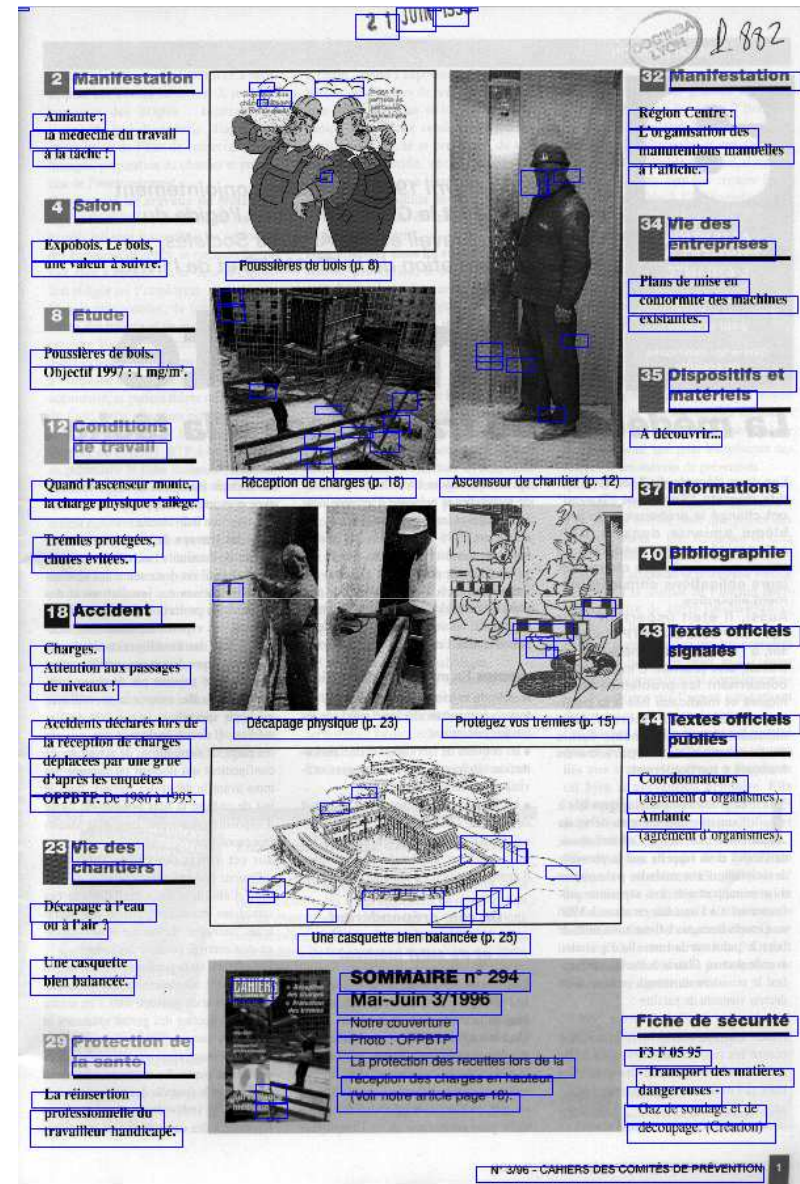
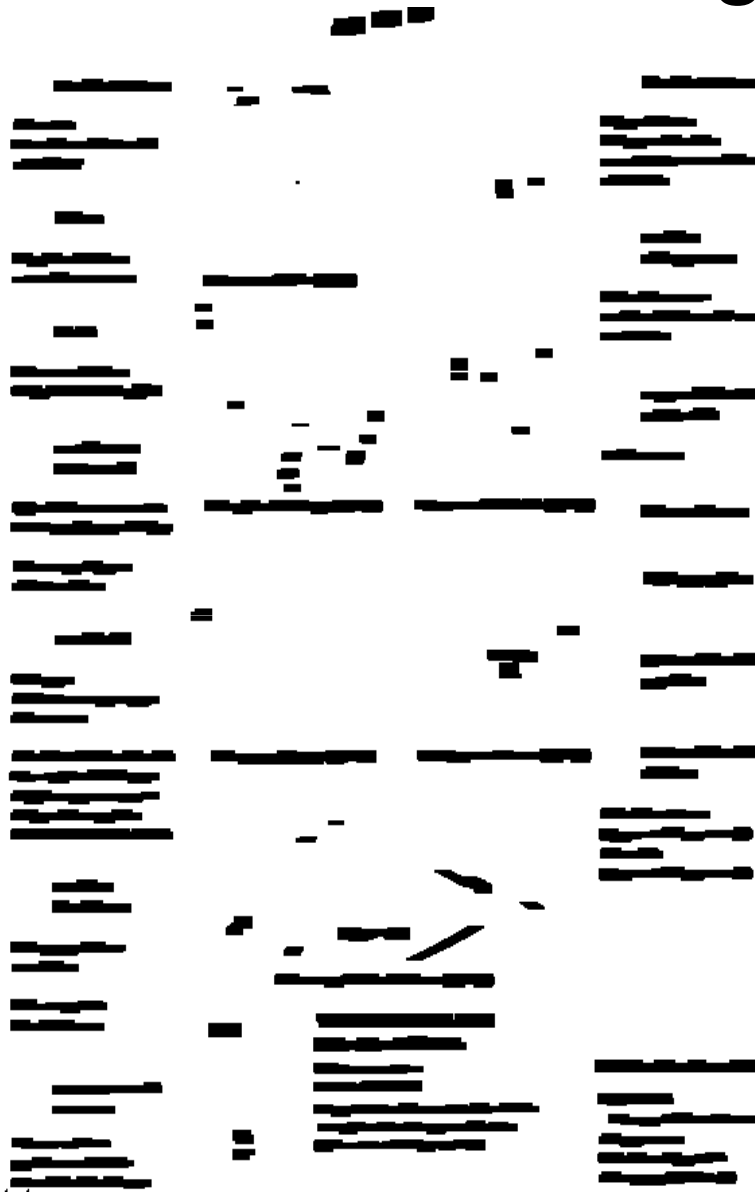


Image originale



Résultat du filtre

Séparation Texte/image (8)



Détection des alignements par morphologie Résultat de la segmentation

Séparation Texte/image : les limites (9)

Limitations: Un critère de segmentation à partir des seules informations issues de l'image n'est pas toujours suffisant !

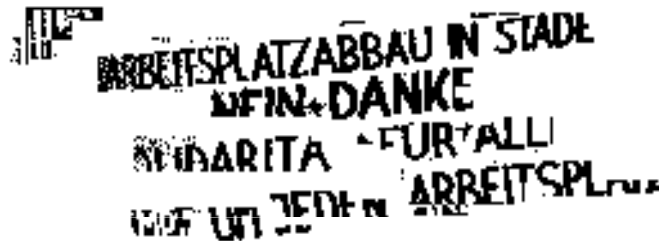
► Trouver d'autres critères sémantiquement plus évolués

SPECIAL REPORT: No one wants to raise taxes. Governments around the world are finding it easier to legalize gambling—and rake in the proceeds. Casinos and lotteries are booming everywhere. But can states control the betting fever they've unleashed? **Page 46**



EUROPE: The dream of monetary union, once scheduled for the end of the decade, is fading fast, even among those ministers who support it **Page 20**

SPECIAL REPORT: No one wants to raise taxes. Governments around the world are finding it easier to legalize gambling—and rake in the proceeds. Casinos and lotteries are booming everywhere. But can states control the betting fever they've unleashed? **Page 46**

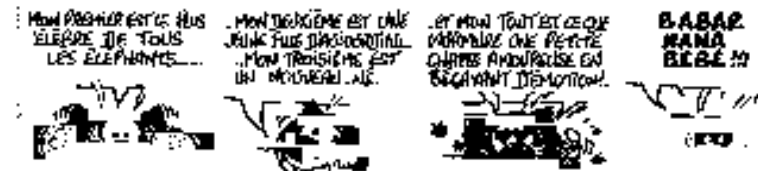


EUROPE: The dream of monetary union, once scheduled for the end of the decade, is fading fast, even among those ministers who support it **Page 20**

Les facéties de Barnabé



Les facéties de Barnabé

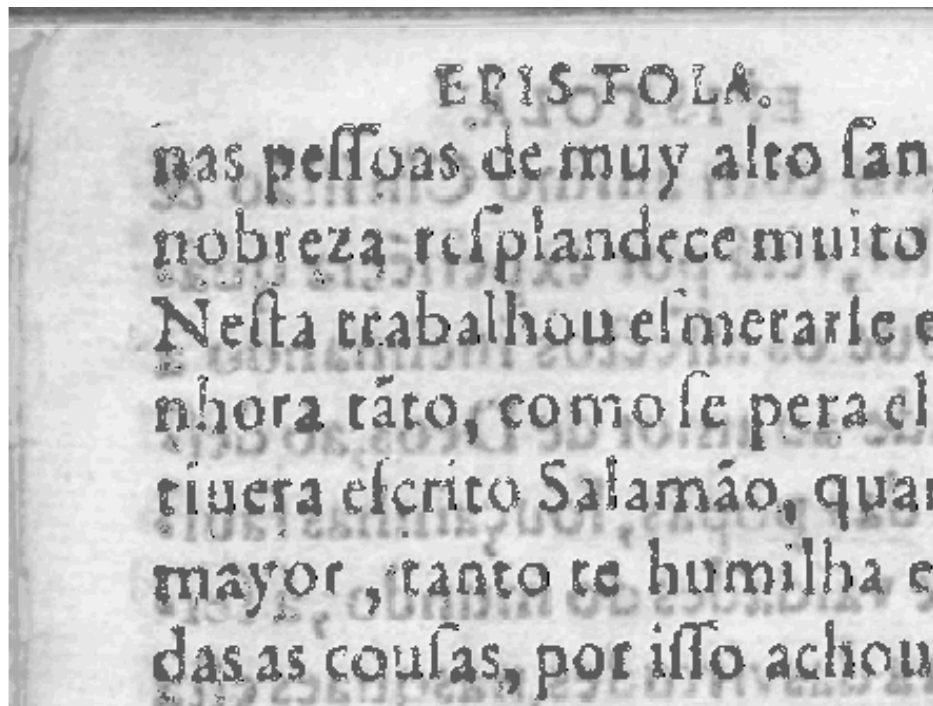


Segmentation à partir de l'interprétation

► **Critère d'extraction** : le taux de redondance des formes de caractères

Une segmentation inadaptée produit des caractères collés ou l'apparition du verso sur le recto ou des caractères cassés.

Une segmentation optimale correspond à un taux maximal de redondance des formes de caractères !



EPISTOLA.
nas pessoas de muy alto san
nobreza resplandece muito
Nesta trabalhou esmerar-se e
nhora tão, como se pera el
tiuera escrito Salamão, qual
mayor, tanto te humilha e
das as cousas, por isso achou

Plan

- Quelques bases en traitement des images
 - ◆ Réduire la présence de bruit dans les images
 - ◆ Filtrer les images
 - ◆ Modification du contraste
 - ◆ Diverses corrections
 - ◆ Restaurer les images binaires
 - ◆ Segmenter les formes
 - ◆ Compresser les contenus

Stockage & compression des images

- ▶ Le problème du volume de stockage des images,
- ▶ Difficultés de transmission sur la bande étroite du réseau.

Quelques chiffres ...

Un livre de 500 pages occupe :

- ❑ 1 Mo en texte intégral
- ❑ 3,5Go (6 CDRoms ou 1DVD) en niveaux de gris
- ❑ 39 heures pour charger un ouvrage complet à 56Kb/s

Une page en 400 dpi sans compression

- ❑ en 16 millions de couleurs représente 40 Mo
- ❑ en 256 couleurs ou a niveaux de gris 10 Mo
- ❑ en binaire 1,5 Mo
- ❑ 16 millions de couleurs comprimés JPEG 0.5..4 Mo

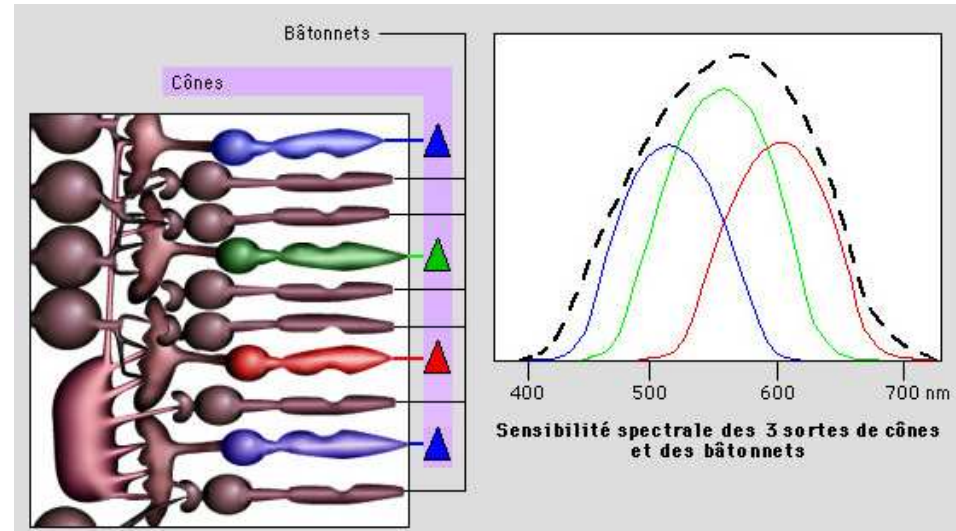
Réduction de l'information guidée par la perception (1)

- ▶ Pour la lecture sur écran, on utilise les caractéristiques de la perception humaine pour réduire la quantité d'information sans gêner le confort visuel.
- ▶ On peut ainsi réduire l'information couleur ou l'information haute fréquence (compression JPEG) sans nuire à la lisibilité

Mais cette perte d 'information peut avoir des conséquences catastrophiques pour l'interprétation de l'image par un logiciel !

Réduction de l'information guidée par la perception (1)

► Nous ne distinguons en moyenne pas plus de 32000 couleurs et 64 niveaux de gris



➡ L'image peut être simplifiée en abaissant le nombre de couleurs ou en privilégiant la luminance sur la chrominance



16 millions de couleurs



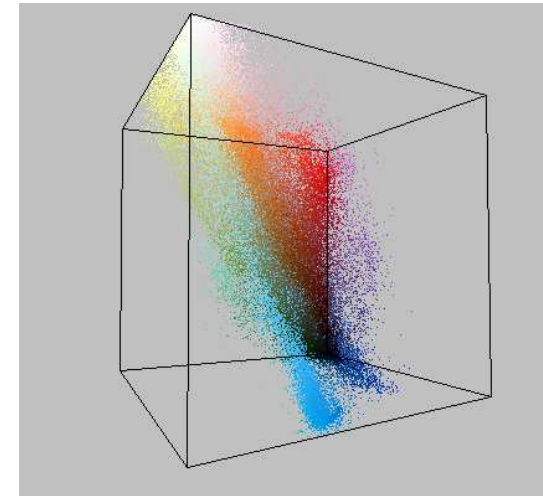
256 couleurs

Réduction du nombre de couleurs

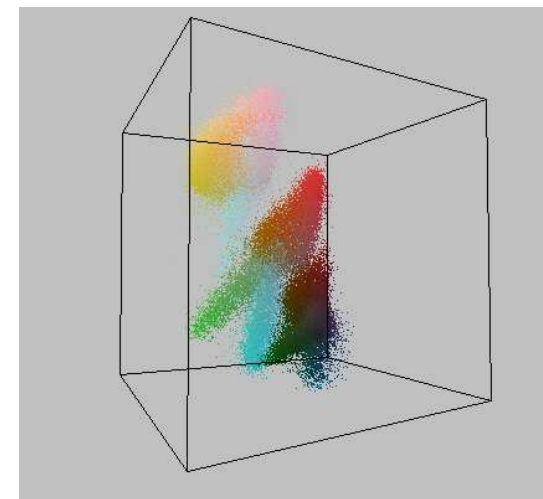
Réduction aux K couleurs les plus fréquentes



155494 couleurs utilisées sur
16 millions de couleurs possibles



155494 couleurs



65535 couleurs

Réduction du nombre de couleurs (1)



Pour une
image couleur
à palette

$F(x,y)$ = indice

Palette[indice]=
(R,V,B)

Application directe aux cartes graphiques de PC

16 principales couleurs conservées dans une palette



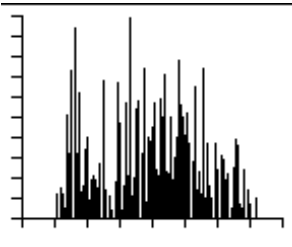
Réduction du nombre de couleurs (2)

$L=(R+V+B)/3$ mais pour la vision humaine $L=0.3R+0.6V+0.1B$

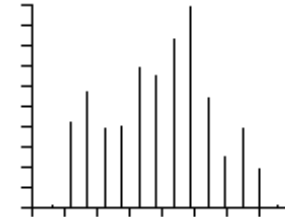
Possibilité de requantifier sans impression visuelle de dégradation



138 couleurs utilisées



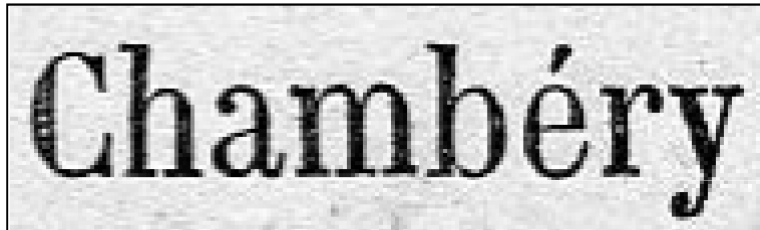
14 couleurs utilisées



JPEG: Le fléau des documents numérisés

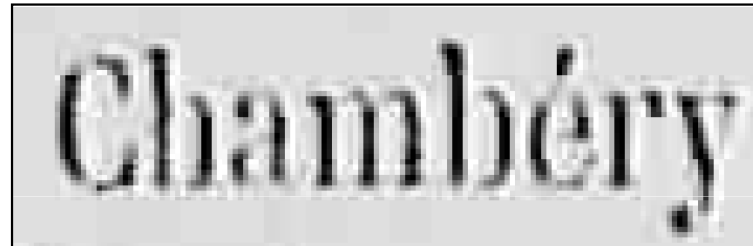
- ▶ JPEG est un format de compression adapté aux images naturelles et non aux images de textes !

Image originale



Non comprimé : 800 ko

Image JPEG 90% qualité



JPEG : 500 ko

- ▶ JPEG : Compression inefficace pour les Images de documents
- ▶ Taux de compression 1:2 à 1:20 moyenne $\approx$ 1:10
- ▶ Pour les images naturelles (1:5..1:500)

La compression JPEG, précisions... (1)

C'est une méthodes de compression « **psychovisuelle** » avec perte basée sur la vision humaine

- ❑ Réduction de l'information de couleur
- ❑ Réduction des informations spatiales inutiles ou redondantes

Perte d'information non perceptible

Perte d'information irréversible pour l'analyse des images

Perte d'information variable suivant un facteur de qualité Q

Q=100% (meilleure qualité Tx de compression faible 1:2)

Q=70% (Qualité moyenne Tx de compression optimal 1:10)

Q=40% (Qualité Faible Tx de compression élevé 1:15)

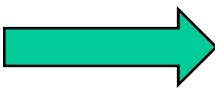
La compression JPEG, précisions... (2)

Historique

- Fin 80: CCITT et ISO => une norme internationale pour la compression d'images fixes
- Création du groupe JPEG (Joint Photographic Experts Group)

Le principe général

- La matrice des pixels est décomposée en blocs de 8x8 pixels qui vont tous subir le même traitement: transformée DCT.



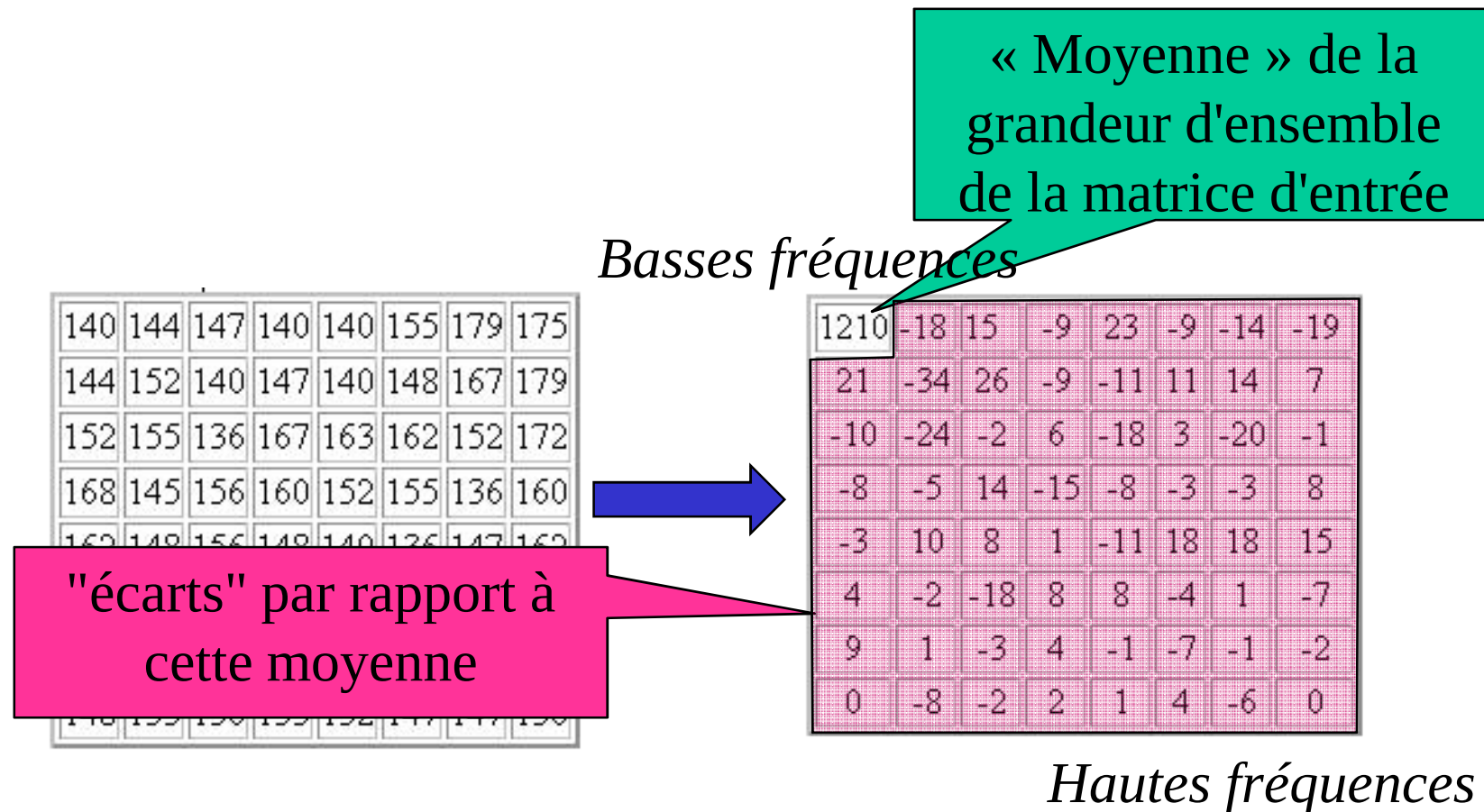
140	144	147	140	140	155	179	175
144	152	140	147	140	148	167	179
152	155	136	167	163	162	152	172
168	145	156	160	152	155	136	160
162	148	156	148	140	136	147	162
147	167	140	155	155	140	136	162
136	156	123	167	162	144	140	147
148	155	136	155	152	147	147	136

1210	-18	15	-9	23	-9	-14	-19
21	-34	26	-9	-11	11	14	7
-10	-24	-2	6	-18	3	-20	-1
-8	-5	14	-15	-8	-3	-3	8
-3	10	8	1	-11	18	18	15
4	-2	-18	8	8	-4	1	-7
9	1	-3	4	-1	-7	-1	-2
0	-8	-2	2	1	4	-6	0

La compression JPEG, précisions... (3)

La transformée DCT

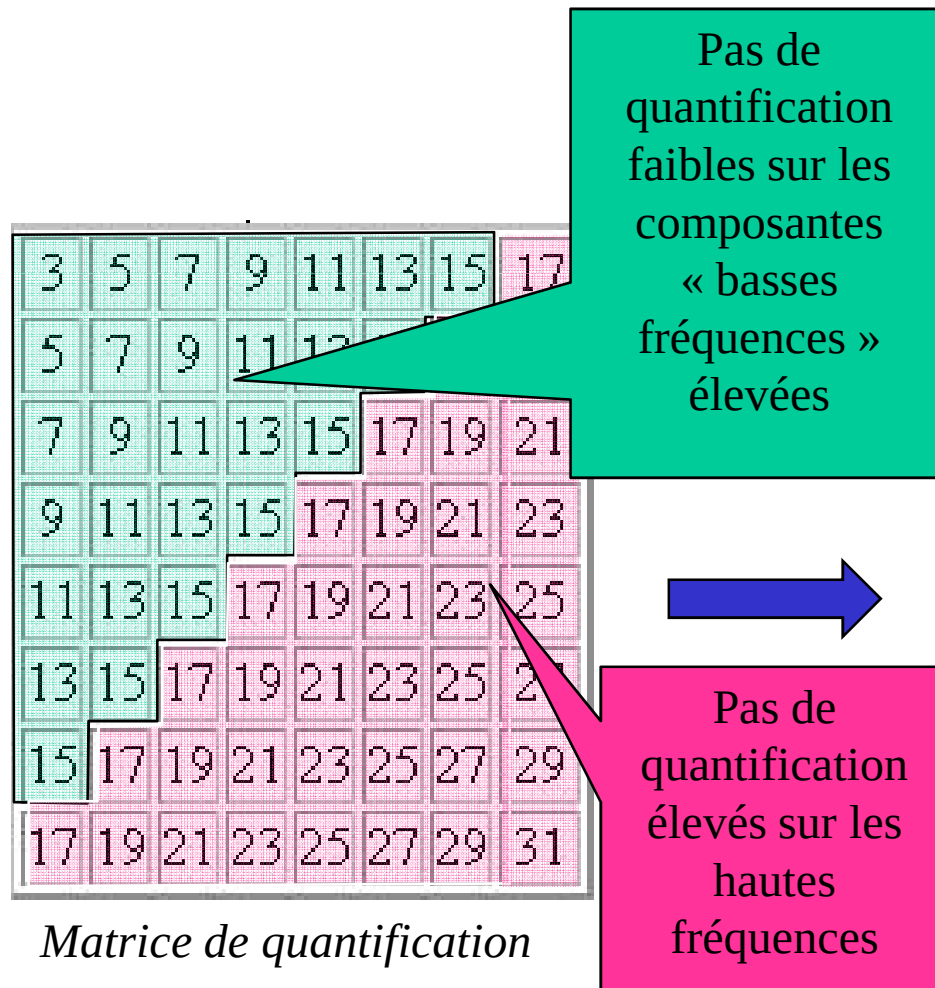
Elle permet de décrire chaque bloc en un graphique de fréquences (correspondant à l'importance et à la rapidité d'un changement de couleur) et en amplitudes (qui est l'écart associé à chaque changement de couleur) plutôt qu'en pixels et qu'en couleurs comme c'est le cas normalement.



La compression JPEG, précisions... (4)

La quantification par application d'une matrice de coefficients

Décrit chaque bloc en un graphique de fréquences (correspondant à l'importance et à la rapidité d'un changement de couleur) et en amplitudes (qui est l'écart associé à chaque changement de couleur).



-Permet de choisir un pas de quantification important pour certaines composantes jugées peu significatives visuellement.

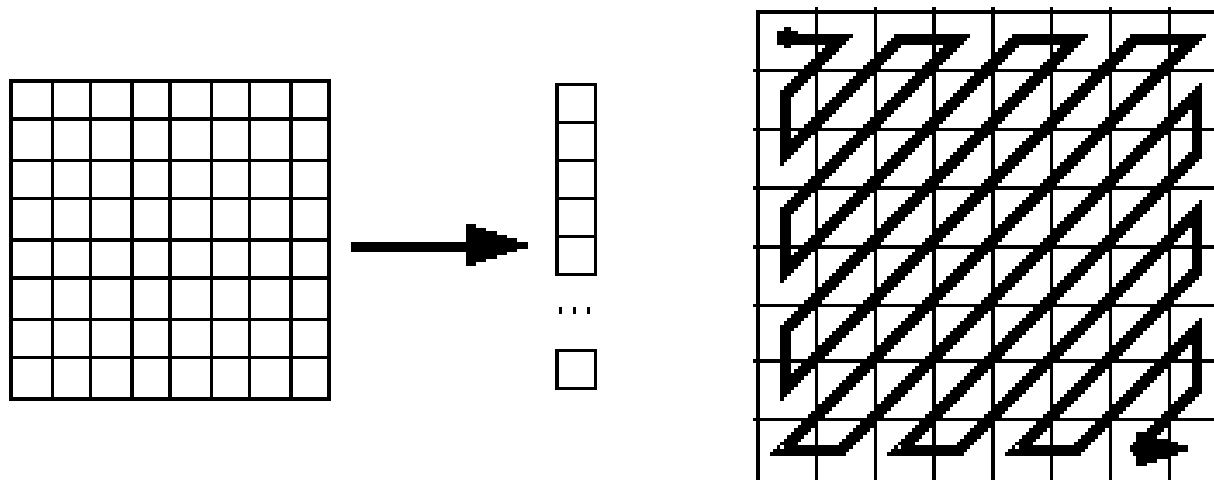
403	-4	2	-1	2	-1	-1	1
4	-5	3	-1	-1	1	1	0
-1	-3	0	0	-1	0	-1	0
-1	0	1	-1	0	0	0	1
0	1	1	0	-1	1	1	0
0	0	-1	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

DCT déquantifiée

La compression JPEG, précisions... (5)

L'encodage

- Parcours des éléments des basses fréquences aux hautes
- Beaucoup de hautes fréquences nulles dans la matrice DCT: engendre de longues suites de 0 consécutifs.
- Les suites de valeurs nulles sont simplement codées en donnant le nombre de 0 successifs.
- Les valeurs non nulles seront codées en utilisant une méthode statistique de type Huffman.



La compression JPEG, précisions... (6)

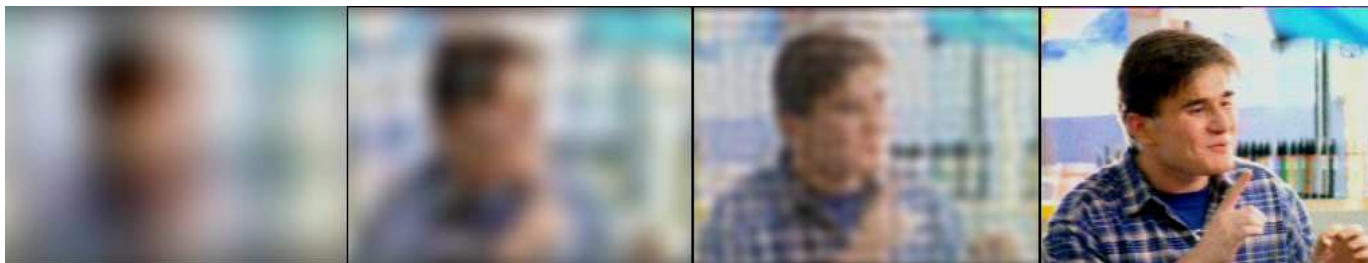
- Suppression de l'information inutile pour la vision : les HF



80% de perte d'information

94% de perte d'information

Conservation des fréquences principales



20 coefficients

60...

200...

Image originale

Résultat de la compression JPEG

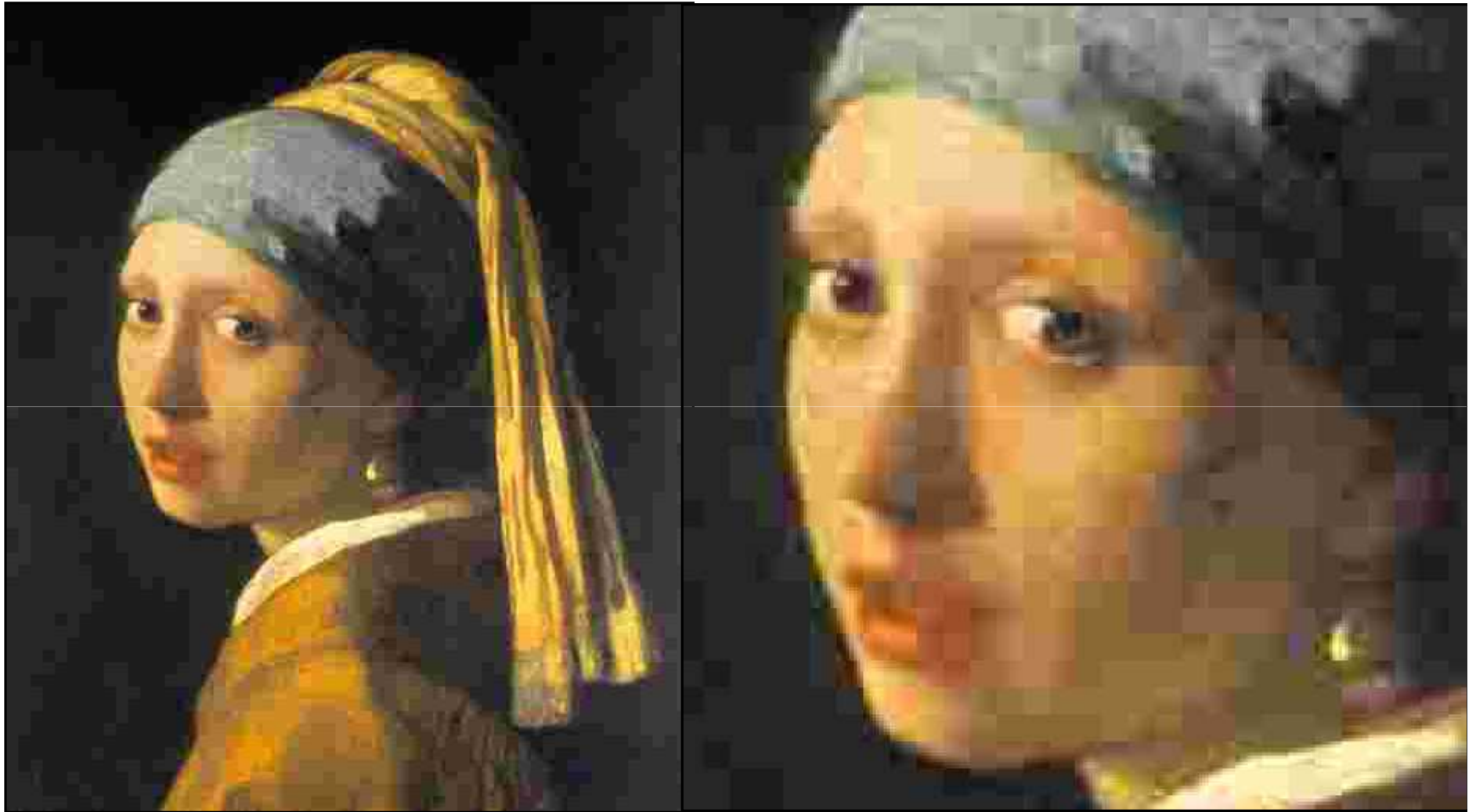
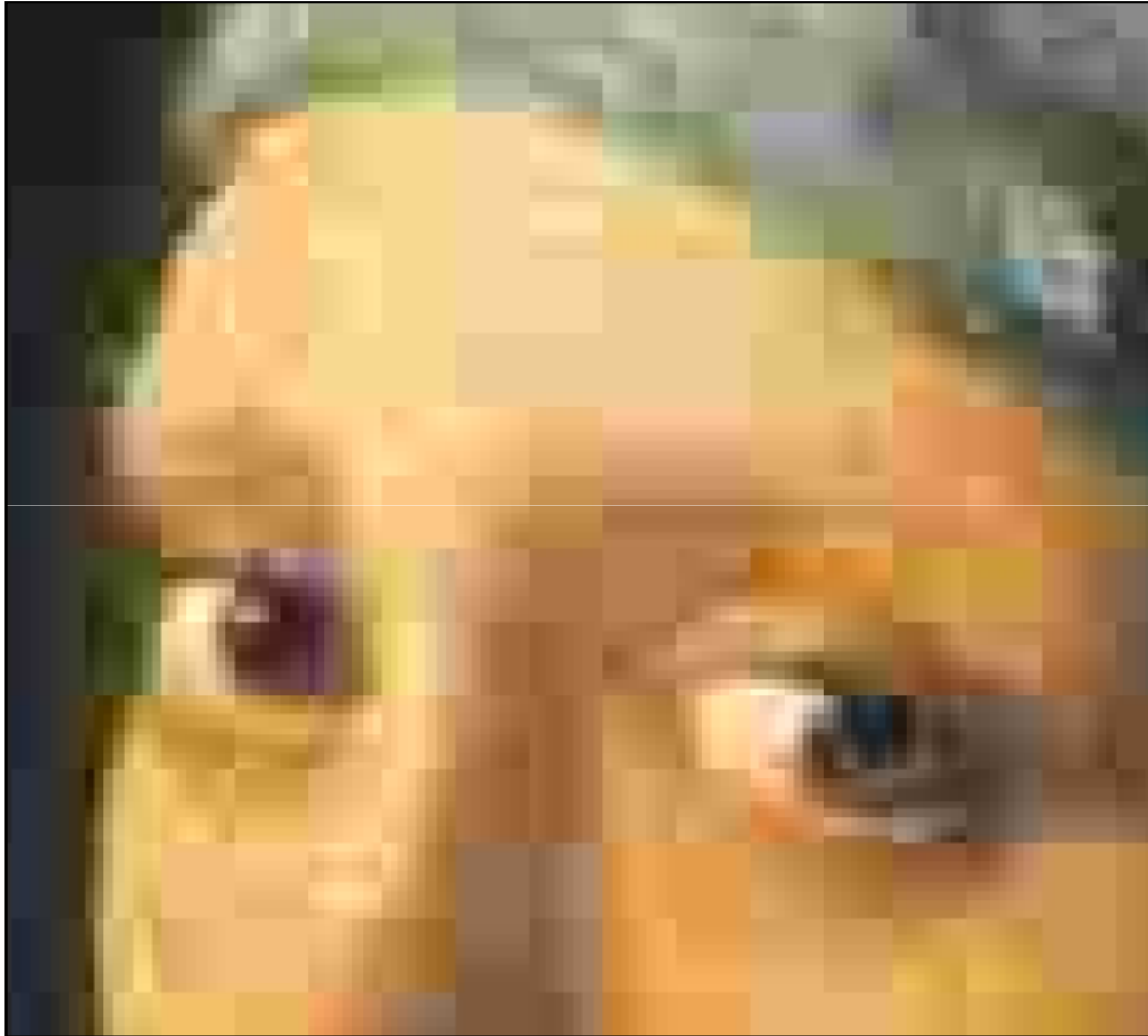


Image jpeg avec une compression 90% de qualité (bonne qualité)

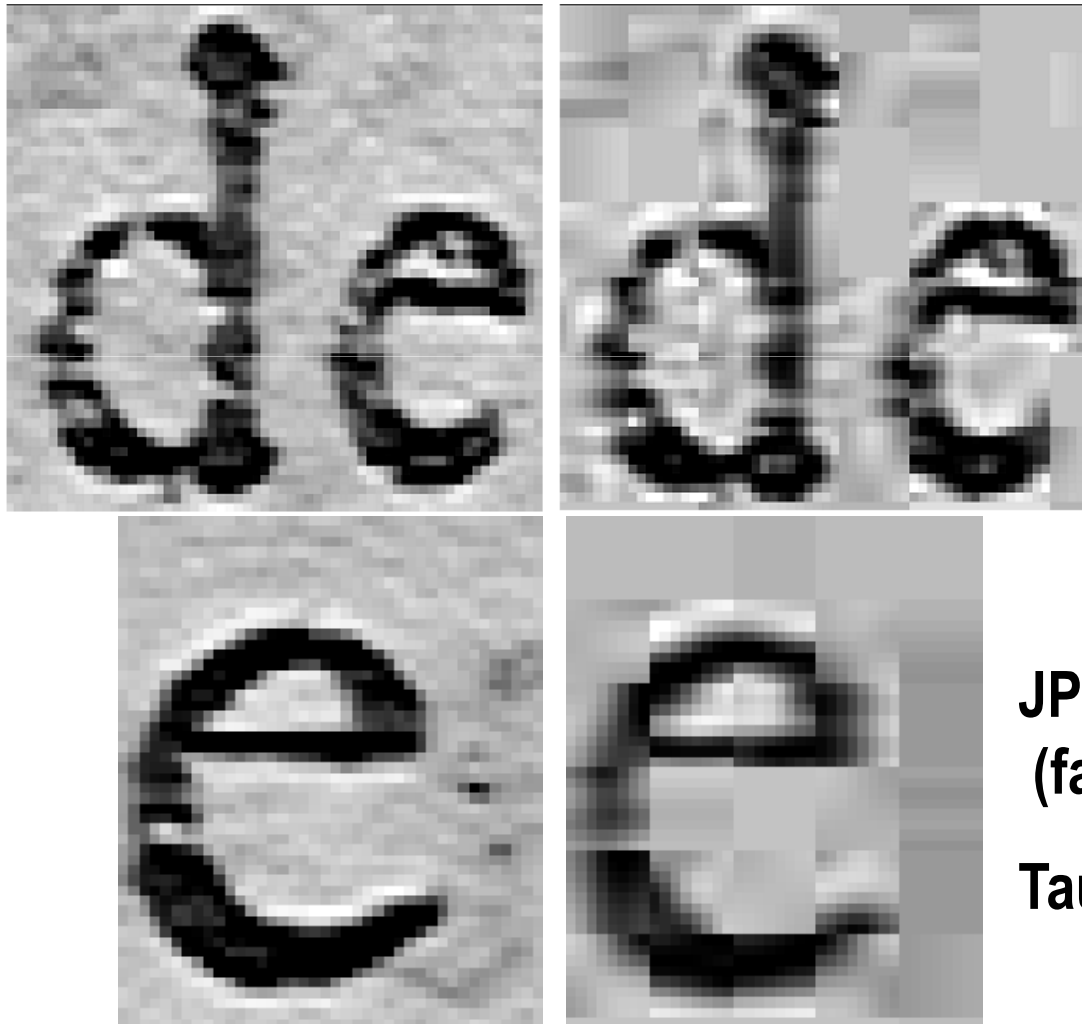
Résultat de la compression JPEG



Et pourtantles défauts sont visibles après agrandissement !

Effets de la compression JPEG sur les traits

- Perte irréversible d'information
- Déformations des traits
- Effets de blocs
- Modification des couleurs



JPEG n'est pas adapté à la complexité des traits des caractères et à leurs localisations aléatoires dans une image de texte

JPEG Q=70%
(facteur Q par défaut des logiciels)

Taux de compression 1:15

Fonctionnalités de JPEG2000

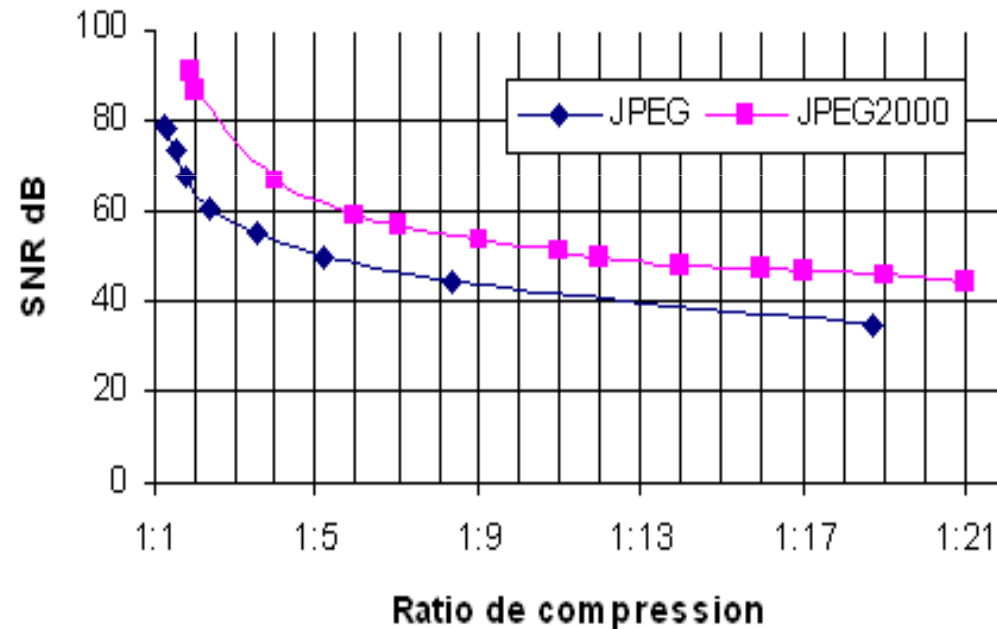
- Définition de régions d'intérêt



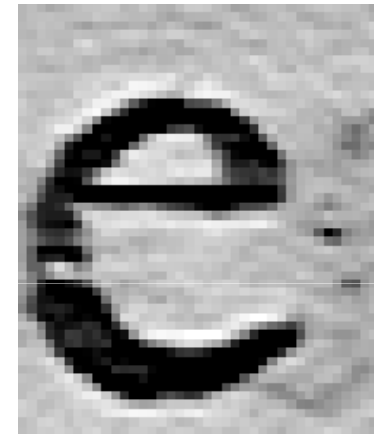
- Multi-précision (affichage progressif)
- Droits de propriété intellectuelle
- Stockage d'information de description en vue de l'indexation des images

Améliorations JPEG 2000

JPEG2000 avec sa compression en ondelettes est plus efficace que JPEG, mais les taux de compression sont de 1:10 sur les images de texte



Original



JPEG 1:12

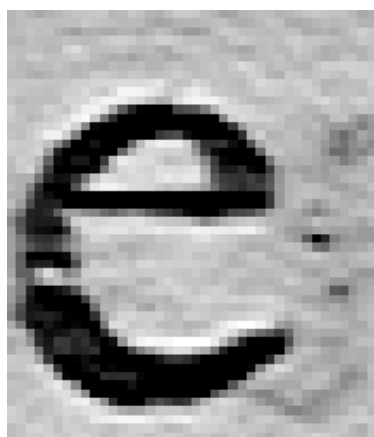
JPEG2000 1:12

Comparaison Jpeg vs Jpeg2000



JPEG 2000 1:90

JPEG 1:90



Original



JPEG 1:12



JPEG2000 1:12

La compression adaptée : DjVu

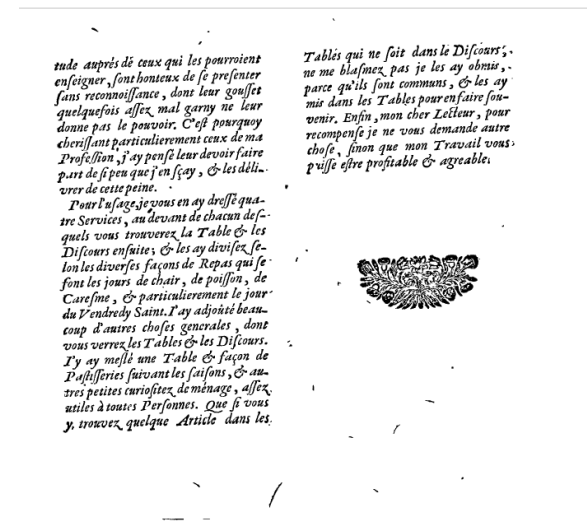
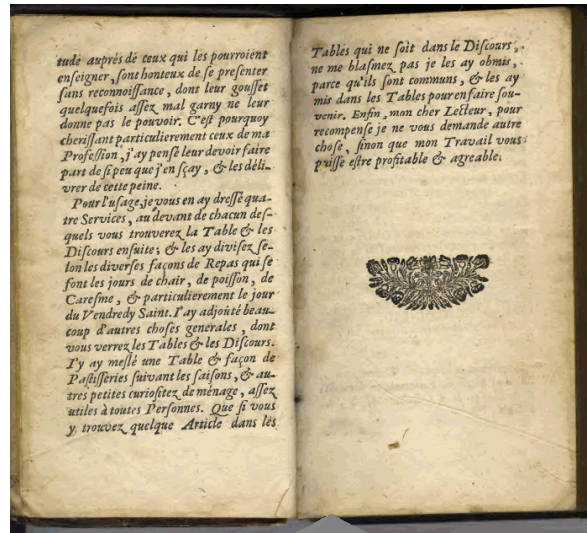
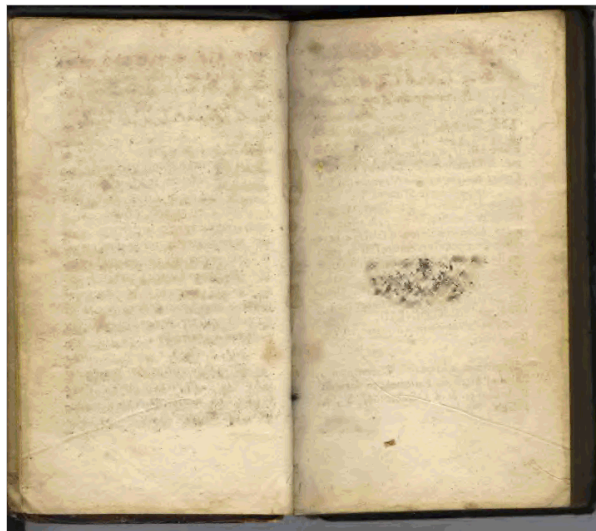
Séparation de l'image en plusieurs couches :

► l'avant-plan (traits)

► l'arrière-plan (image du support papier)

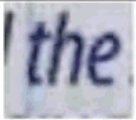
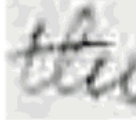
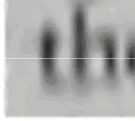
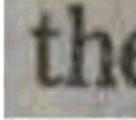
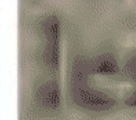
compression de chaque plan avec une méthode adaptée

Compression entre 1:20 et 1:500 sur les documents numérisés !



Exemple DjVu 1:200

Différences visuelles entre formats

Image Originale	JPEG, 300dpi,	JPEG, 100dpi,	Ondelettes 300dpi,	Compression forme/fond séparé
				
20640K	292K 70:1	50K 412:1	61K 338:1	52K 396:1
				
9534K	116K 82:1	17K 560:1	20K 476:1	19K 501:1
				
22013K	383K 57:1	41K 536:1	43K 511:1	38K 579:1
				
12990K	250K 51:1	38K 341:1	42K 309:1	40K 324:1
				
12128K	206K 58:1	35K 346:1	39K 310:1	37K 327:1
				
31059K	388K 80:1	77K 403:1	78K 398:1	73K 425:1

Compression DEBORA > 1:100

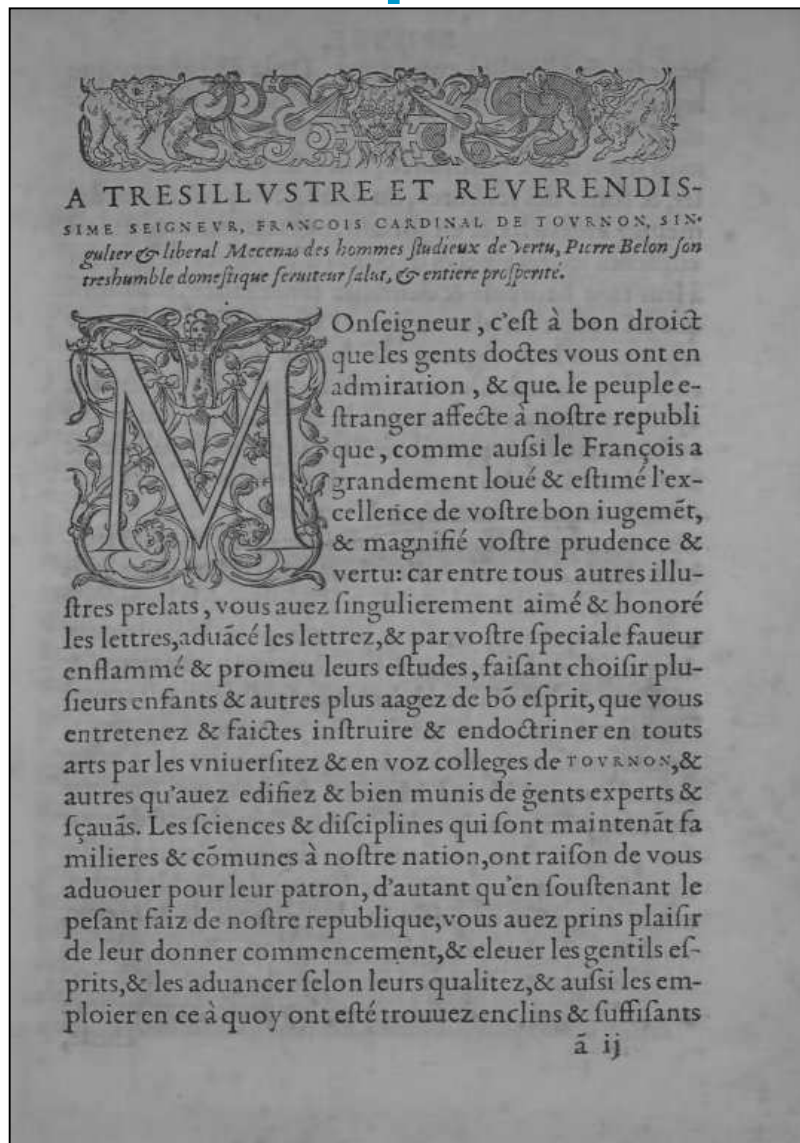
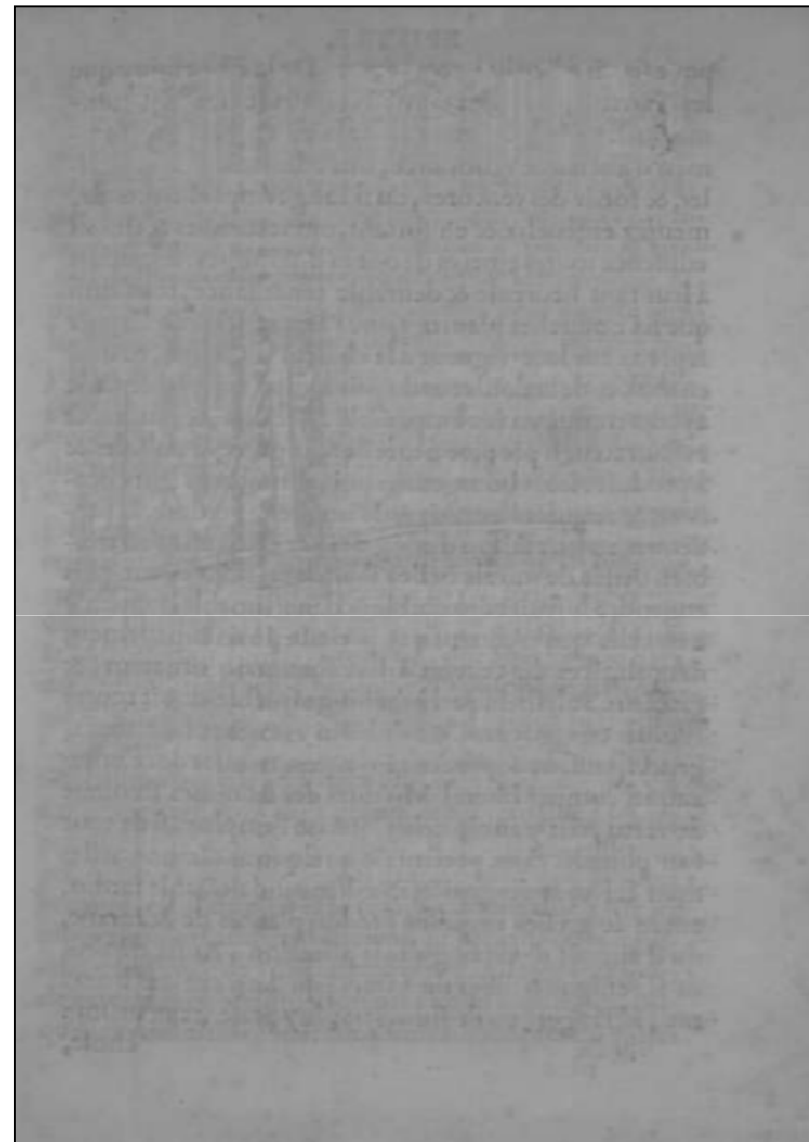


Image Originale (8 Mo)
4000x3000 en niveaux de gris



Compression du support
papier par JPEG (10 Ko)

Compression DEBORA > 1:100

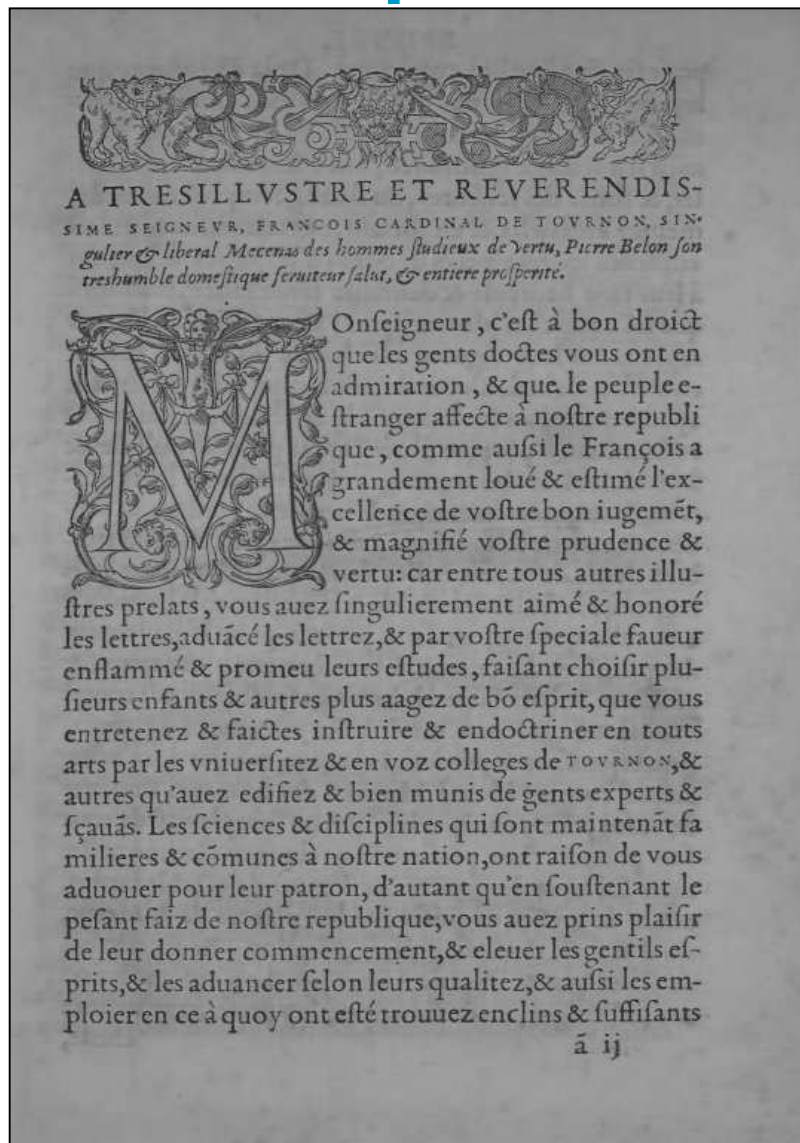
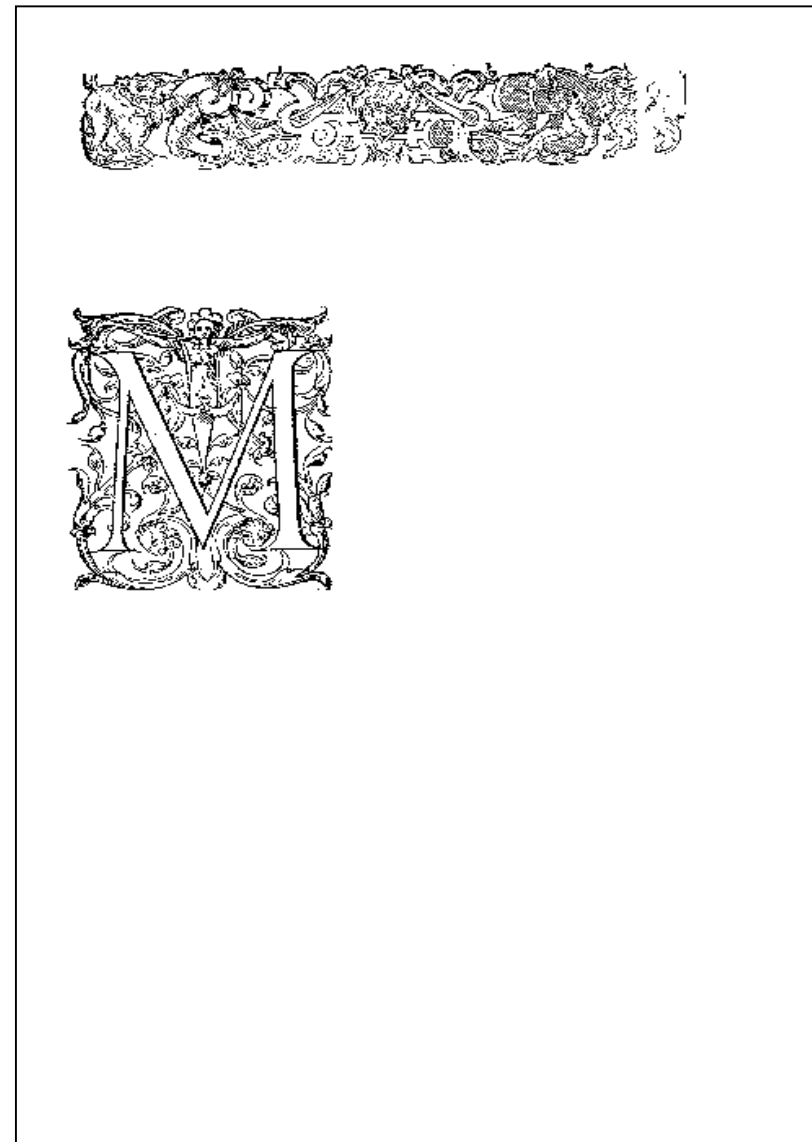


Image Originale (8 Mo)
4000x3000 en niveaux de gris



Compression des parties
graphiques CCITT G4 (20 Ko)

Compression DEBORA > 1:100

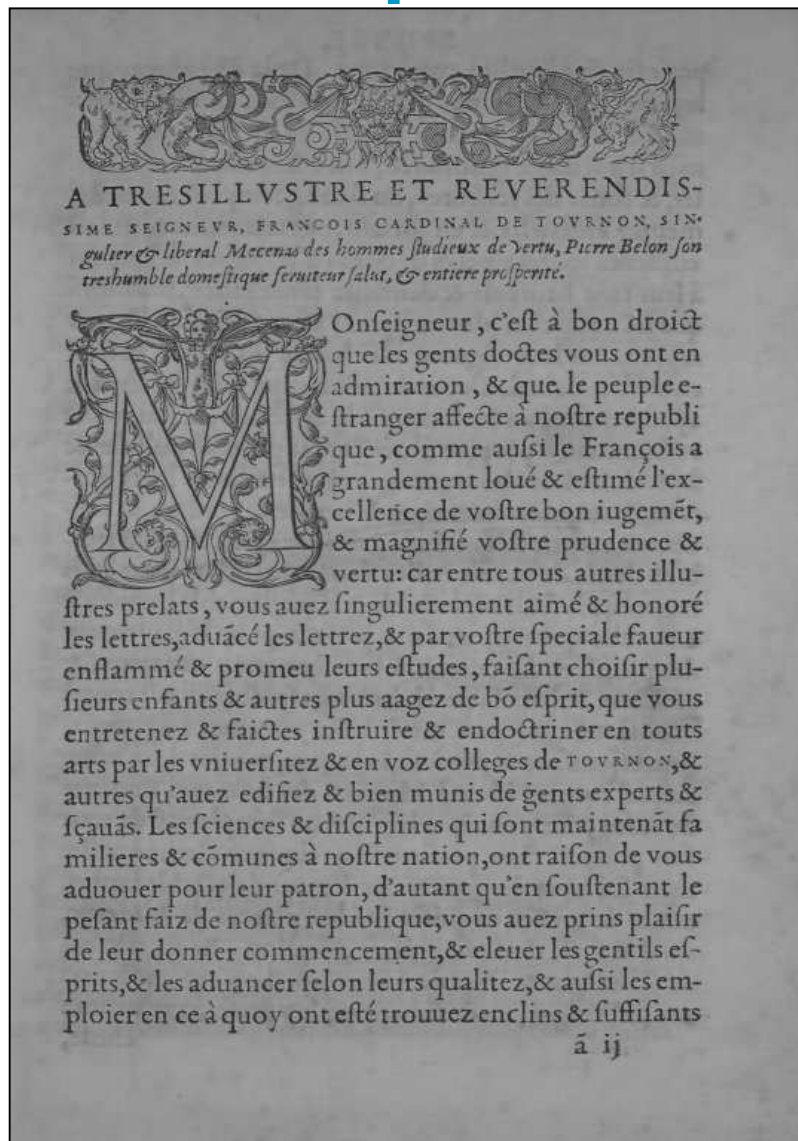
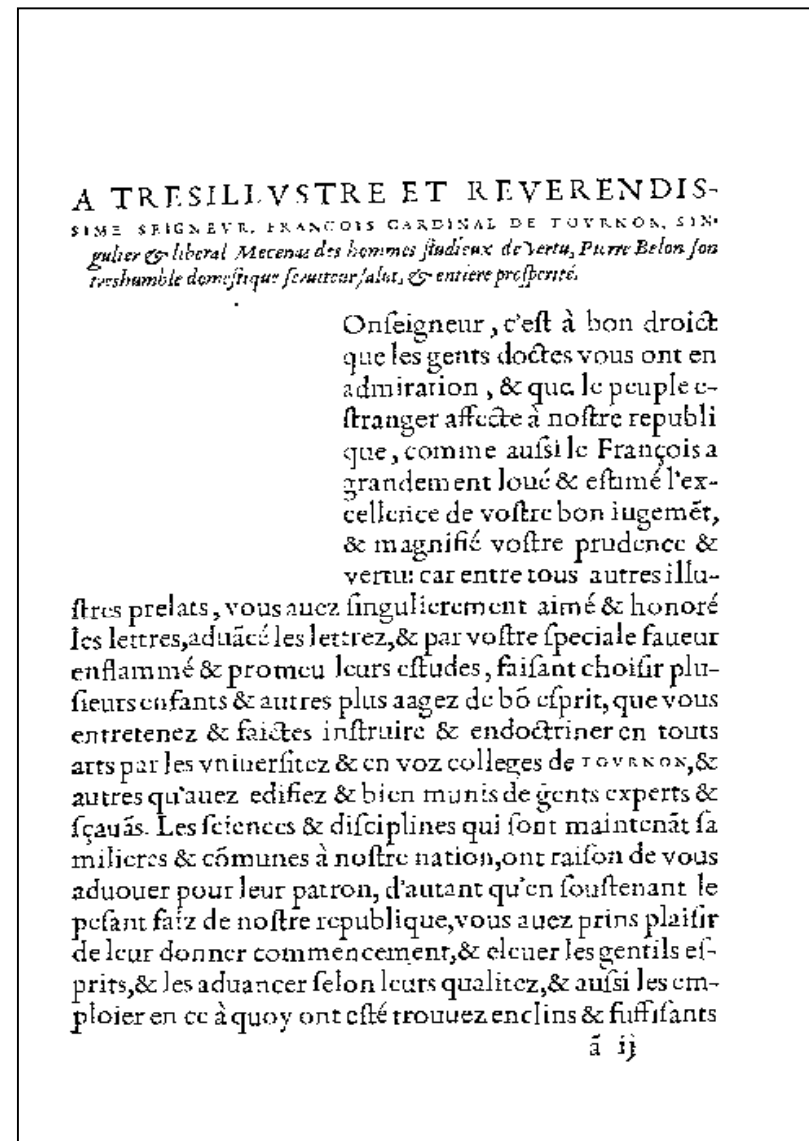


Image Originale (8 Mo)
4000x3000 en niveaux de gris



Compression du texte par
redondance de formes (10Ko)

Compression DEBORA > 1:100

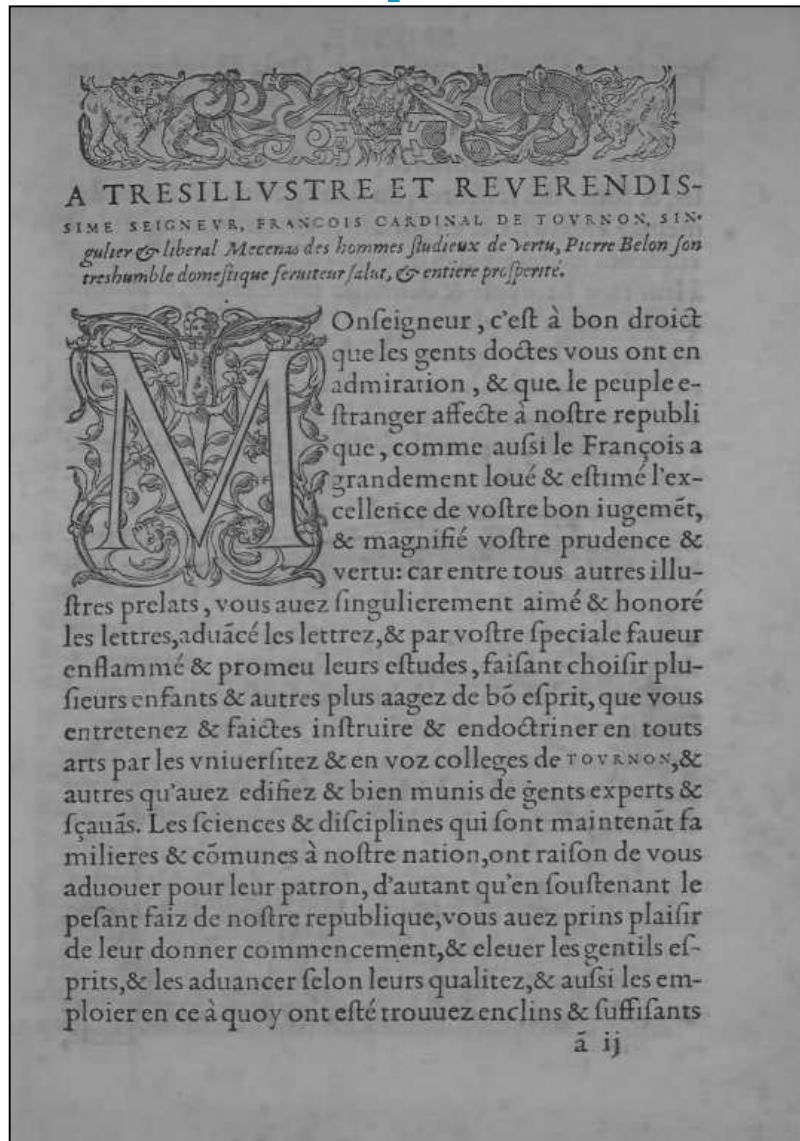
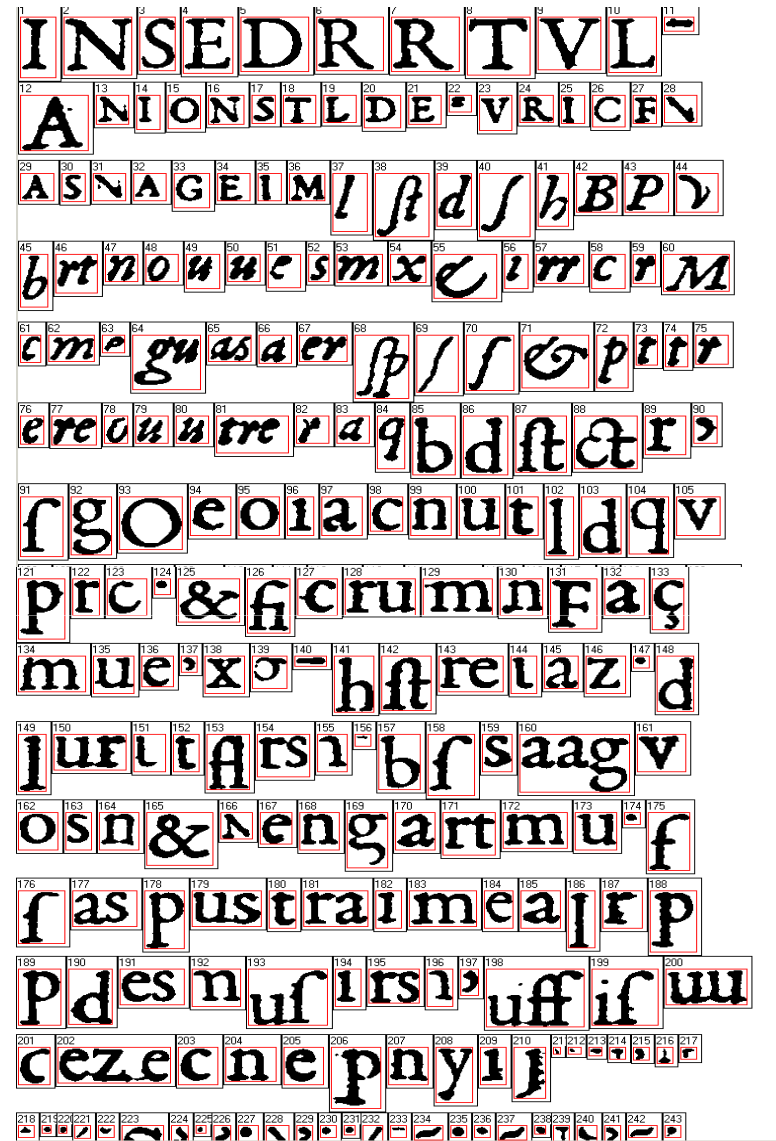


Image Originale (8 Mo)
4000x3000 en niveaux de gris



243 formes différentes de
caractères : 90% de redondance

Compression DEBORA > 1:100

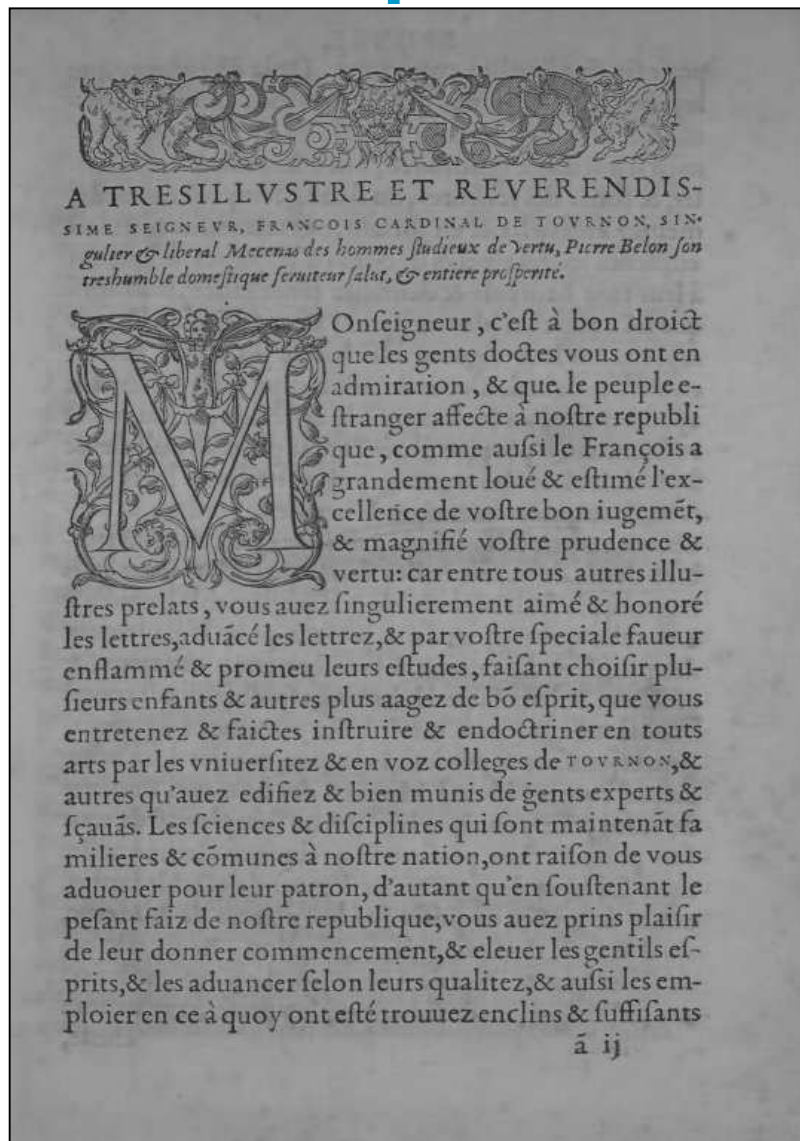
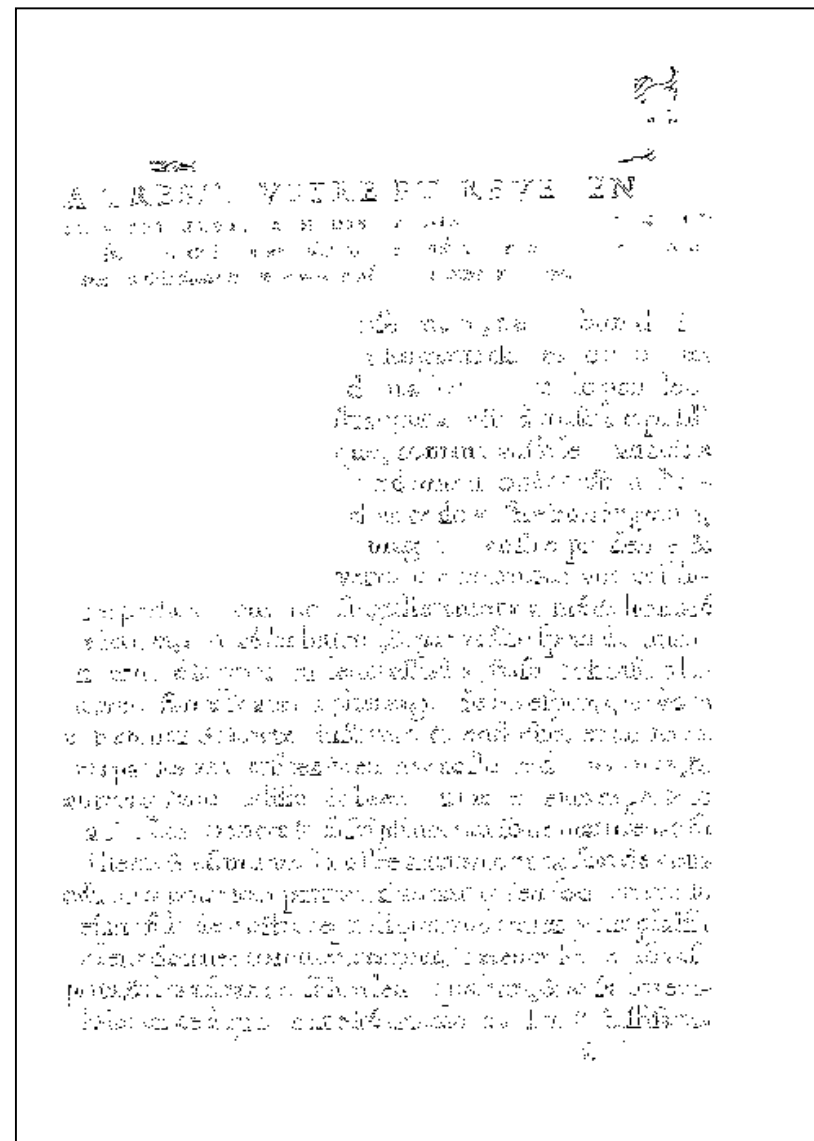


Image Originale (8 Mo)
4000x3000 en niveaux de gris



Compression sans perte des
résidus (40 Ko)

Compression DEBORA > 1:100

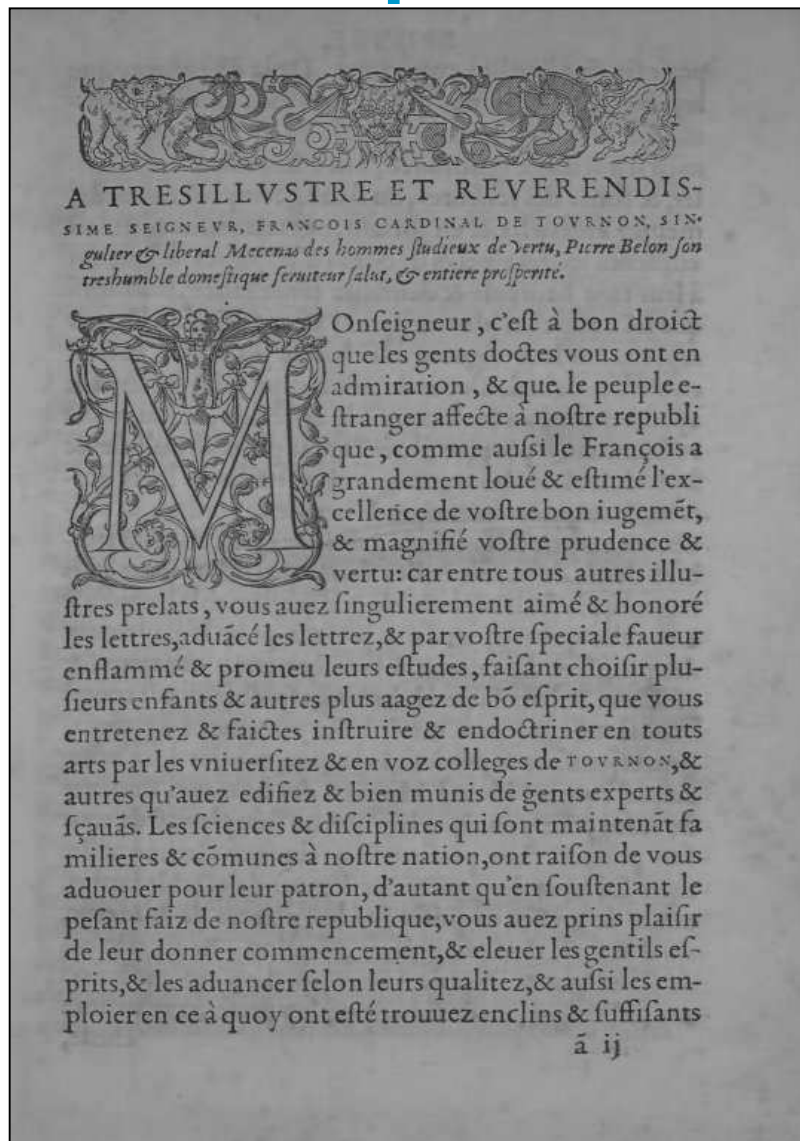


Image Originale (8 Mo)
4000x3000 en niveaux de gris

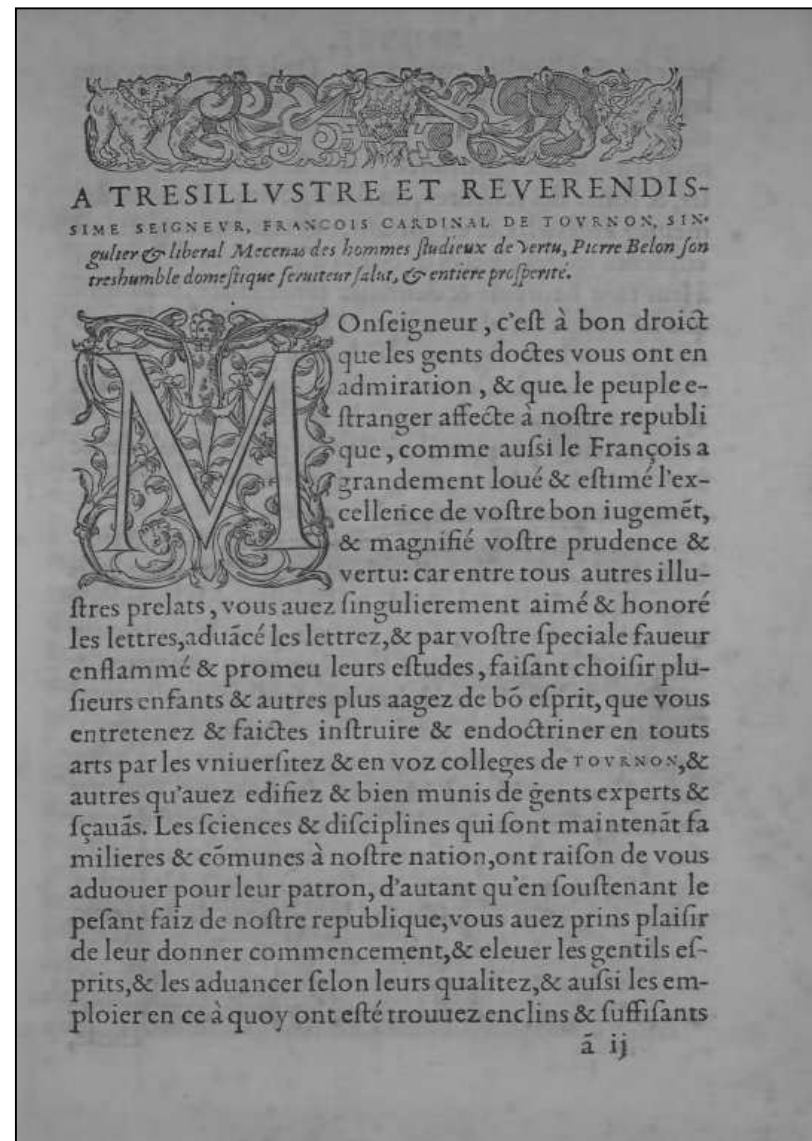


Image compressée (80 Ko)
4000x3000 en niveaux de gris

Compression DEBORA

Mécanisme de compensation

Objectif : obtenir une image exacte par rapport à l'original

Image binaire du texte original

l'opulence & grandeur de son royaume:

Image binaire du texte décomprimé (par redondance de formes)

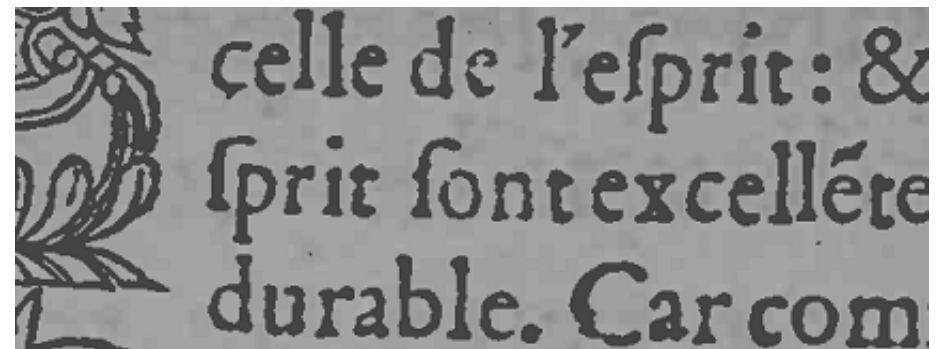
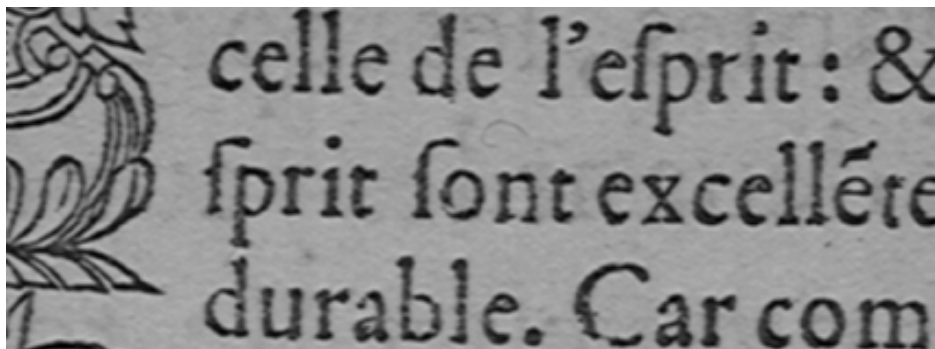
l'opnlence & grandeur de son royaume

Masque de compensation (différence) pour retrouver les formes exactes

l'opulence grandeur de son royaume

Image originale

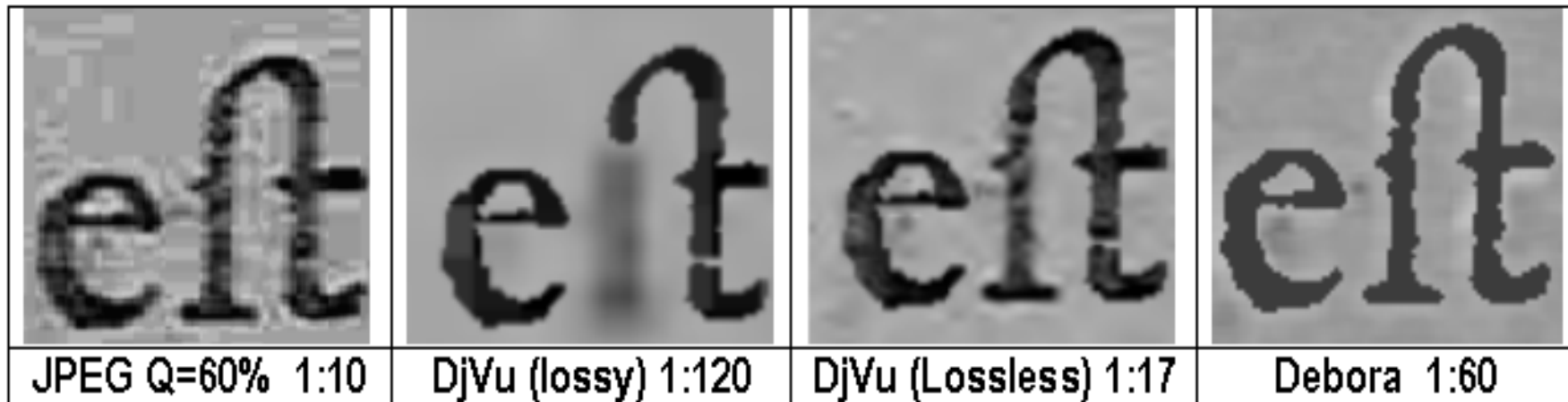
Image comprimée non compensée



Résultats de la compression avancée

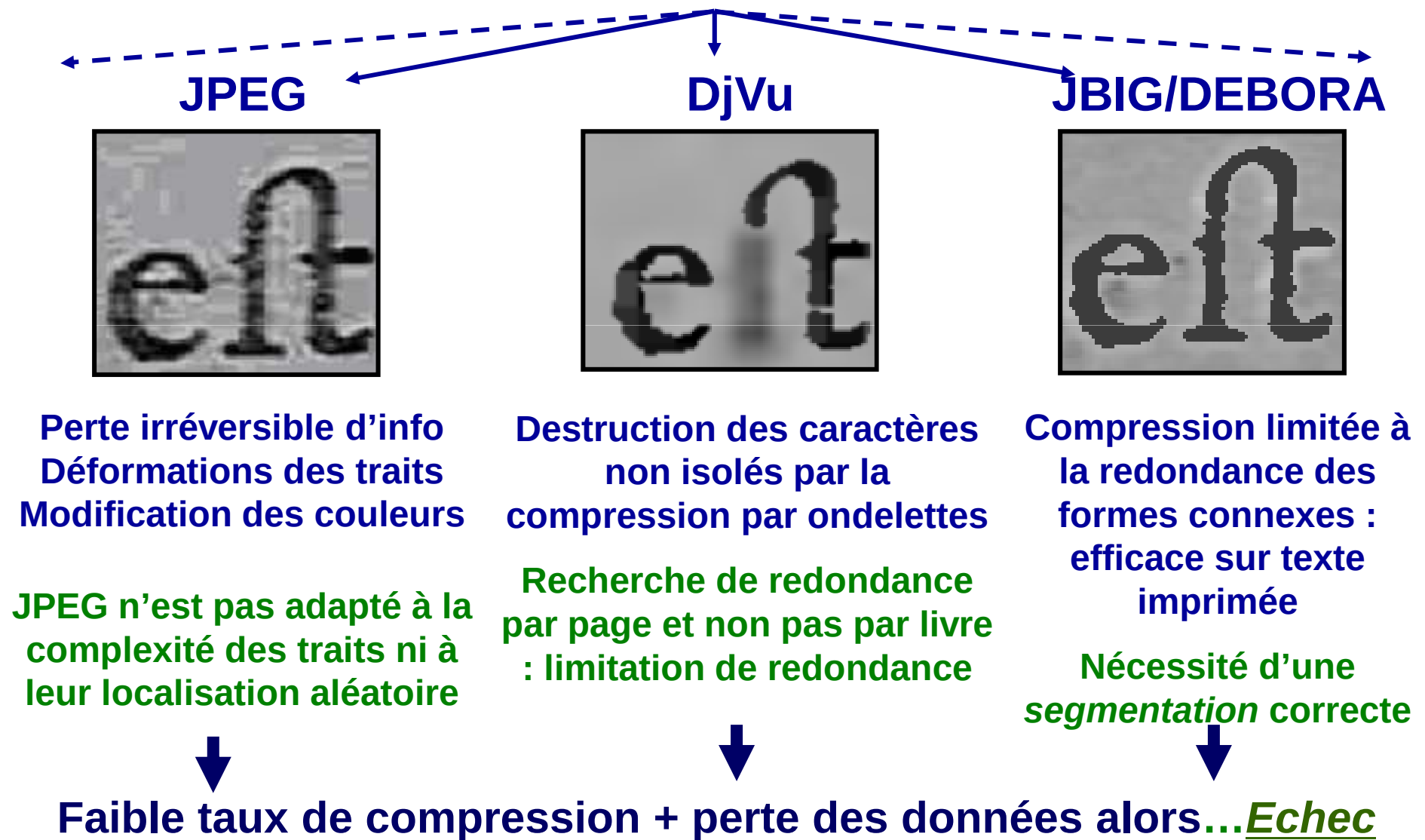
- la compression **agressive** de DjVu (1:120)
- la compression **non agressive** de DjVu (1:17)
- La compression DEBORA (1:50..1:100)
- La compression JPEG (1:2...1:12)

La compression avancée des images de textes est en moyenne 10 fois plus efficace que JPEG sans les artéfacts



Vers de nouvelles approches de compression

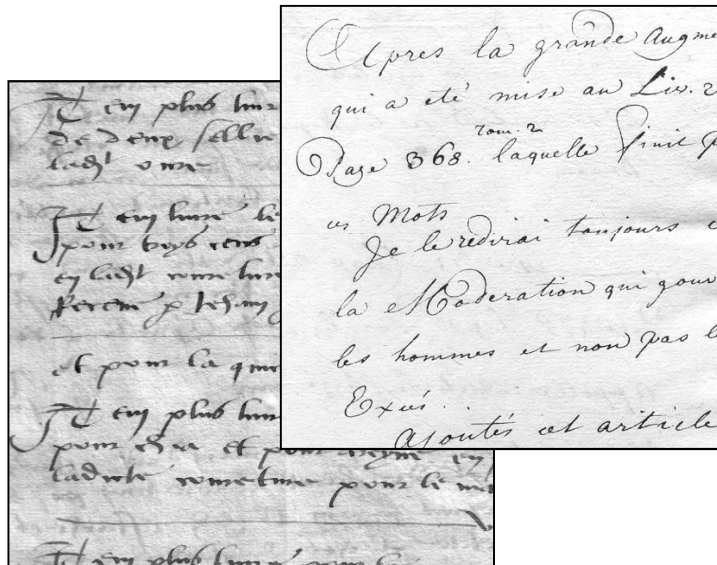
Bilan sur les formats traditionnels



Vers de nouvelles approches de compression

Le problème des images manuscrites

La difficulté de segmentation, l'absence de redondance de formes d'objets connexes et la manque de similarités entre les formes en présence

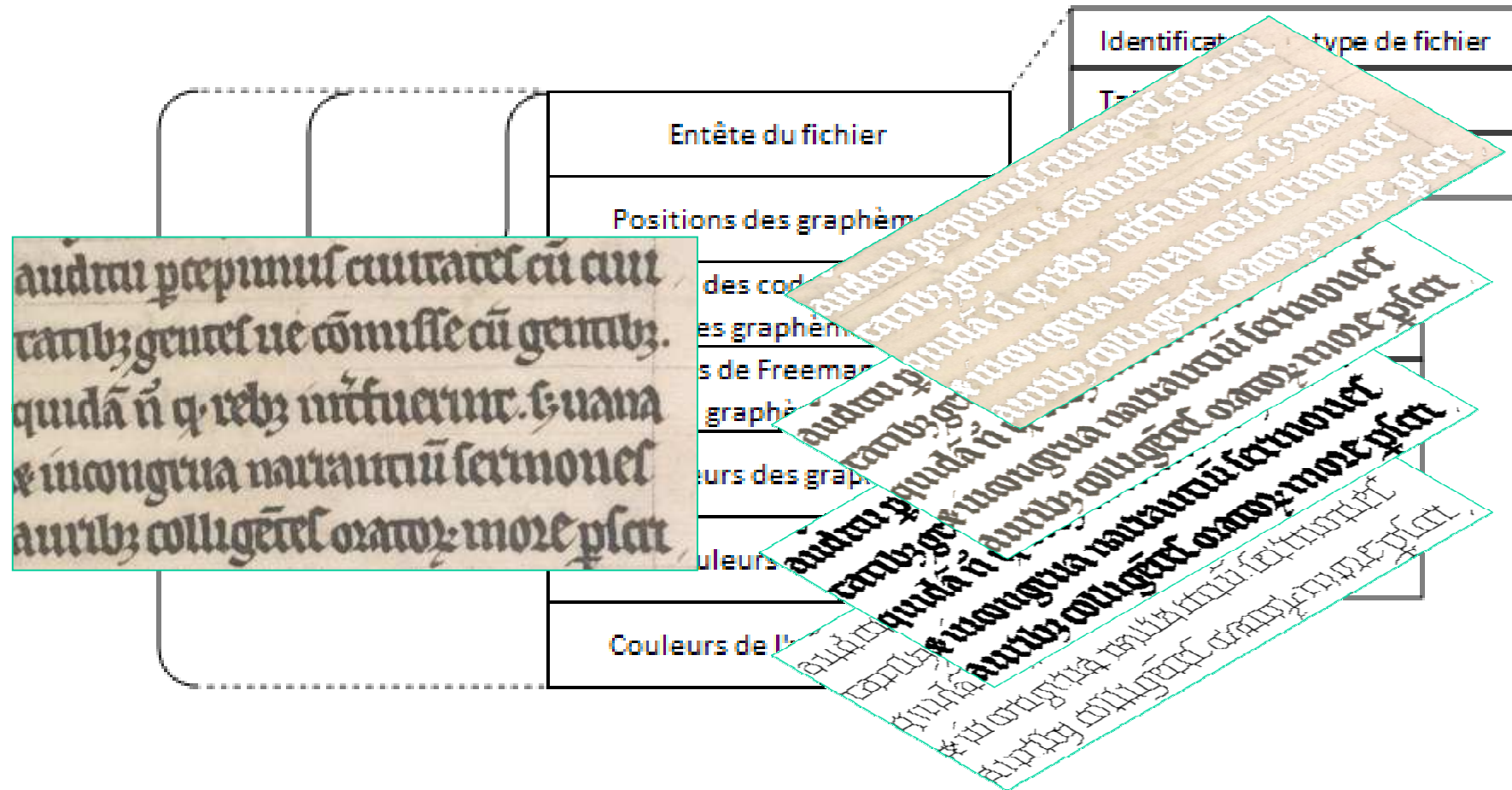


**Nécessité d'une
compression adaptée au
contenu de l'image**

La solution exploitée

S'orienter vers une nouvelle méthodologie basée sur la recherche de similarités partielles de formes adaptée aux caractéristiques des tracés manuscrits

Vers de nouvelles approches de compression



Formats d'images

	Sans Compression	Avec Compression	
		Sans Pertes	Avec Pertes
Images	BMP, PGM, PPM, RAW, TGA ...	GIF, PCX, PNG, TIF, ...	JPEG JPEG2000

Quelques formats d'image pour le stockage

TIFF « Tagged Image File Format » : Format privilégié de la numérisation

- Seul format adapté à la complexité des documents
- Format tagué (liste de tags) ouvert, libre de droit, gérant le multi-page, différentes qualités d'images et supportant différentes compression avec perte (JPEG) et sans perte (CCITT, ZIP...)

PNG “portable network graphics”

- Gamma correction, gestion de la transparence
- Faible compression sans perte de type ZIP

BMP “Bitmap” (Format propriétaire Windows)

- 8 - 24 bit couleurs, pas de compression
- Pas de transparence, pas de système CMYK
- Valable sur plateforme Windows seulement

GIF

- 8 bit couleurs, Compression sans perte (variante LZW)
- Adapté aux logos et boutons pour Internet + GIF animé
- problème de copyright par CompuServe

Quelques formats d'image pour la diffusion

NE JAMAIS UTILISER COMME FORMAT DE STOCKAGE !
SEULEMENT POUR LA DIFFUSION OU L'ECHANGE ET L'EDITION

JPEG “Joint Photographic Expert Group”

- Compression avec perte pour la diffusion seulement des images à niveaux de gris ou en couleur

JPEG2000 (une avancée dans les normes JPEG)

- Compression avec ou sans perte
- Compression avec une meilleure qualité visuelle
- Tatouage des images, descriptions des métadonnées...

PDF “portable document format” (remplaçant de EPS)

- Gestion d'hyperliens, gestion mode texte/image
- Copyright Adobe
- Format d'impression $\neq$ Format d'édition ou format d'image

PSD “Adobe Photoshop”

- Conserve toutes les informations pour l'édition et le traitement (sélections, calques...) = Format d'édition des images

Quelques formats de compression adaptés

NE JAMAIS UTILISER COMME FORMAT DE STOCKAGE !
SEULEMENT POUR LA DIFFUSION OU L'ÉCHANGE ET L'ÉDITION

DJVU Format de compression des images de documents

- Compression avec perte pour la diffusion seulement des images à niveaux de gris ou en couleur
- Pour les images binaires, une compression de type JBIG est appliquée
- Pour les images qui ne sont pas des images de texte, une compression de type JPEG2000 est appliquée
- Possibilité de lier texte/images (résultat d'OCR dans l'image)
- Le plugin pour les navigateurs sur Internet est complet (zoom, transmission progressive..)

JBIG (Joint Bilevel Image Group) format ouvert

- Compression avec et sans perte des images binaires adaptée aux images de textes
- Format de compression adapté aux contenus des images
- Pas d'encodeur efficace libre de droit mais 2 à 5 fois plus efficace que le CCITT G4, le LZW et le TIFF.

A la recherche du format idéal de données

Besoins : Rendre le document « vivant »

Quel est le meilleur format pour : Rechercher, Modifier, Annoter, Echanger, et indexer des documents numérisés ?

Un livre numérisé et entièrement rétro-converti contient une masse dense d'informations hétérogènes :

- Images
- transcriptions,
- différentes éditions (diplomatique, critique..),
- les annotations des auteurs et des experts,
- liens Images&textes,
- liens structure/images,
- liens vers des commentaires et des sites web...
- structure physique,
- structure logique,
- méta-données du livre,
- méta-données sur le contenu des pages,