

Module 9 - Section 2 :

Numériser les documents

ANNE DEBANT
PATRICK PERROT



version 1

14 novembre 2011

Table des matières

1. Introduction à la numérisation	9
1.1. Définitions.....	9
1.2. La numérisation: une technique aux bases anciennes.....	11
1.2.1. La transmission des documents écrits ou dessinés.....	11
1.2.2. La transmission des photographies à distance.....	11
1.2.3. L'héritage actuel.....	12
1.3. La bonne numérisation.....	12
2. Notions fondamentales en numérisation	13
2.1. Image en pixels.....	13
2.2. Résolution.....	14
2.3. Modes colorimétriques.....	16
2.4. Analyse et typologie des documents à numériser.....	17
2.4.1. Analyse du document.....	17
2.4.2. Nature des documents déterminant la colorimétrie.....	18
3. Objectifs de la numérisation	19
3.1. Objectifs.....	19
3.2. Usage du fichier final.....	20
3.3. Portabilité.....	20
3.4. Répartition des documents numérisés selon l'usage.....	20
3.5. Usages mixtes.....	21
4. Formats d'enregistrement des images numérisées	23
4.1. Formats sans perte de données.....	23
4.2. Formats avec pertes de données.....	24
4.3. Choix des formats d'enregistrement selon l'usage.....	27
4.3.1. Création d'une photothèque numérique.....	28
4.3.2. Création d'une bibliothèque de documents à pagination multiple.....	28
4.4. Formats libres et formats propriétaires.....	31
4.4.1. Formats libres.....	31
4.4.2. Formats propriétaires.....	32

4.5. Microfilm et numérisation : des champs d'application connexes.....	32
5. Numériseurs	35
5.1. Format des originaux déterminant le numériseur.....	35
5.2. Le choix du numériseur.....	36
5.3. Types de numériseurs.....	36
5.4. Analyse des caractéristiques des numériseurs.....	37
5.5. Définition des fonctionnalités.....	38
5.6. Numériseurs pour l'acquisition mono-page.....	38
5.7. Numériseurs pour l'acquisition des documents multi-pages.....	43
5.8. La mise en place des éléments du poste de numérisation.....	49
5.9. Ségrégation des documents.....	50
5.10. Post-traitement des images.....	51
5.11. Informations textuelles dans le fichier image.....	52
5.12. Trucs et astuces.....	52
5.12.1. Documents plus grands que la vitre d'exposition.....	52
5.12.2. Numérisation d'un ouvrage «à la baguette».....	53
6. Numérisation des manuscrits	55
6.1. La nature des manuscrits.....	55
6.2. Choix d'un procédé physique pour les reliés.....	55
6.3. Choix d'un procédé physique pour les feuilles volantes.....	56
6.4. Choix du mode colorimétrique.....	56
6.4.1. Le mode bitonal.....	57
6.4.2. Le mode niveaux de gris.....	58
6.4.3. Le mode couleurs RVB.....	58
6.5. La numérisation de microformes.....	59
6.6. Importance de la taille des fichiers.....	59
7. Numérisation des photographies	63
7.1. Première approche en numérisation de photographies.....	63
7.2. Plaques de verre.....	64
7.3. Diapositives et négatifs de petit format.....	64
7.4. Photographies sur papier.....	66
7.4.1. Photographies monochromes.....	66
7.4.2. Photographies en couleur.....	67
8. Numérisation des gravures	69

8.1. Numérisation au bon mode.....	69
8.2. Cas spéciaux.....	72
9. Numérisation des imprimés	73
9.1. Contenu des imprimés, procédés d'obtention.....	73
9.2. Principe de la similitravure.....	73
9.3. Principe de l'héliogravure.....	75
9.4. Procédés de numérisation selon les objectifs.....	76
10. Numérisation des plans	89
10.1. Plans tracés.....	89
10.2. Tirages de plans.....	91
11. Solutions de numérisation en tableaux	93
11.1. Inventaire des typologies de documents.....	93
11.1.1. Documents figurés sans tramage photomécanique.....	93
11.1.2. Plans et dessins techniques.....	94
11.1.3. Photographies, plaques, films.....	94
11.1.4. Documents figurés en similitravure.....	94
11.2. Procédés de numérisation à partir de la typologie des documents.....	95
11.2.1. Manuscrits, dessins, gravures, lithos.....	96
11.2.2. Peintures et petits objets plats.....	98
11.2.3. Photographies, négatifs, diapositifs.....	98
11.2.4. Plans d'architecture et techniques.....	100
11.3. Solutions de numérisation à partir de l'usage du fichier.....	101
12. Bibliographie en ligne	105
Numérisation.....	105
Bibliothèques en ligne.....	106
13. Evaluation des connaissances	107
Galleries associées à ce module	111

Objectifs

Description du module :

Dans une politique de préservation bien conçue, la reproduction des documents, par microfilmage et numérisation, peut être considérée comme une véritable mesure de conservation préventive puisqu'elle permet d'arrêter temporairement ou définitivement la communication des documents originaux. Bien que la numérisation soit moins généralisée en raison des problèmes de pérennité des supports et des technologies, elle est particulièrement favorable à la communication et à une large diffusion du contenu des documents.

Le but du module est de :

présenter les techniques de base pour l'utilisation raisonnée des moyens de reproduction : microfilmage et numérisation.

L'apprenant doit être en mesure de :

Pour le microfilmage

- distinguer les différentes sortes de microfilms
- comprendre les différents usages du microfilms
- connaître le matériel existant pour monter un atelier de microfilmage
- programmer le microfilmage des documents
- organiser matériellement le microfilmage des documents
- conserver les microfilms

Pour la numérisation

- reconnaître la typologie des documents vis à vis du processus de numérisation
- déterminer les moyens matériels de numérisation adaptés aux documents
- déterminer les modes colorimétriques en fonction de l'objectif de la numérisation et de la typologie du document
- pratiquer la numérisation de toutes sortes de documents écrits ou iconographiques
- déterminer les formats d'enregistrement en fonction de l'usage
- adapter les sous fichiers à la diffusion en ligne et faire des assemblages de pages multiples

Positionnement :

Deuxième module du cours 4 sur les politiques et pratiques de conservation et préservation des documents.

Il suit tout naturellement le cours de gestion et traitement des archives et précède celui sur leur communication et mise en valeur.

Conseils d'apprentissage :

Prérequis :

Avant d'aborder ce module, il convient d'avoir pris connaissance des modules 2 (notions générales d'archivistique) et si possible 6 (traitement des archives définitives).

Prérequis pour la numérisation :

- connaissances de base en informatique

-
- pratique de la photo ou de l'image en général nécessaire pour les iconographies

Conseil :

Le cours ne donne pas les détails pas à pas pour la pratique du microfilmage ni de la numérisation sur tel ou tel équipement. Il faut suivre le cours par étapes et tenter d'appliquer la connaissance, à chacune de ces étapes, sur son propre équipement. Ce sont seulement les exercices pratiques qui permettent d'évaluer les difficultés de terrain en matière de reproduction.

Où aller en fin de module?

Le cours 5 sur la communication et la valorisation des archives est la suite logique.

1. Introduction à la numérisation



La numérisation des documents est un moyen de diffuser ces documents de manière immatérielle.

Au départ, il existe une image réelle, imprimée ou photographiée sur papier, c'est le document, parfois vieux de plusieurs siècles, auquel on va donner une nouvelle vie : il pourra se transmettre par des réseaux de communication (la toile «web»), et être restitué à l'autre bout de la planète sans perte de qualité.

L'exemple le plus significatif est celui des journaux. Ainsi pour l'impression quotidienne du *Times* il fallait 17 tonnes de plomb et des tonnes de papier qui étaient distribuées à New York et aux quatre coins du monde. Maintenant d'un simple clic de souris, de bon matin, le fameux journal apparaît à l'écran.

Auparavant, revoyons quelques définitions...

A. 1.1. Définitions



Avant d'entrer dans la détail de notre cours, nous devrions hiérarchiser les types d'archives *virtuelles* que nous sommes appelés à manipuler. Ces archives virtuelles sont celles générées par l'ordinateur, ce sont toutes des archives numériques.

Archives numériques

Ensemble des archives (physiques) numérisées et des archives informatiques, conservées sur un support numérique. On les appelle aussi archives électroniques. Une archive numérique est restituée de façon directe par le truchement d'un ordinateur.

On l'appelle *numérique* parce que l'ordinateur est conçu pour fonctionner avec un langage binaire caractérisé par deux chiffres 0 et 1 représentant un passage de courant ou le non passage de courant dans le processeur.

Les archives numériques contiennent deux types d'archives, à notre sens:

- les **archives informatiques** dont la source est virtuelle ou qui sont arrivées au centre d'archives uniquement sous cette forme virtuelle et dont on ne peut situer

d'original physique.

- les **archives numérisées** qui sont des copies de documents physiques existant bel et bien dans un centre d'archives

Archives informatiques

Les archives informatiques sont les documents créés à l'origine par un ordinateur ou entrés en centre d'archives sous forme numérique, par l'intermédiaire d'un système informatique de transmission à distance ou d'un enregistrement sur support informatique (CD ROM, DVD....).

Ce sont par essence des fichiers texte, des bases de données, des résultats de calculs, des données bancaires, des fichiers de Publication Assistée par Ordinateur (PAO), de Dessin Assisté par Ordinateur (DAO), etc...

Certaines de ces archives peuvent contenir des fichiers de documents numérisés, mais dans ce cas on ne retiendra que la forme sous laquelle elles sont entrées dans le centre d'archive: la forme numérique (le centre d'archives ne possède pas le document physique d'origine)

Quant aux photos issues d'appareils numériques, on les fera entrer aussi dans cette catégorie des archives informatiques, car par essence le modèle original ayant servi à les générer (une scène instantanée de la réalité) n'existe plus, et les éventuels tirages sur papier en sont les copies (et non les sources).

Terme anglais : *computer data files*

Archives numérisées

Ce sont les formes numériques d'archives physiques **reproduites** au mode image par points(pixels), et **dont il existe un original physique**.

Ces documents numériques sont obtenues par l'intermédiaire d'un dispositif appelé numériseur (ou scanner pour les anglophones), qui traduit des réalités analogiques (la lumière provenant de l'original) en signaux numériques.

Attention : Les archives filmées

Nous considérons dans cette catégorie les productions (films) contenant des images animées et des sons, avec un titre, un réalisateur, un générique et un visa d'exploitation pour certains cas: documentaires, courts et moyens et longs métrages, publicités filmées et films publicitaires...

Les techniques modernes permettent d'utiliser l'informatique pour créer, ou enregistrer ces images animées et sonores. Dans ce cas, on classera les films sur supports numériques dans les archives filmées et non pas dans les archives informatiques, ni dans les archives numérisées, car elles appartiennent à un secteur d'activité qui n'est pas celui de l'image fixe, mais bel et bien des **archives audiovisuelles**.

Bien sûr si quelques clips vidéo de quelques secondes ou quelques prises de vues documentaires se retrouvent dans un fonds d'archives informatiques, on les y laissera, sans faire de distinguo.

Dans le cours qui suit, nous allons étudier la création des **archives numérisées**, soit la génération de **copies numériques** de documents existant physiquement.

B. 1.2. La numérisation: une technique aux bases anciennes

Un appareil permet de transformer l'image de l'original en signaux électriques : c'est le numériseur.

Le numériseur est une machine qui semble nouvelle, mais ses ancêtres datent du 19e siècle.

Le principe de base de la numérisation est de transformer un document physique en impulsions électriques par l'intermédiaire d'un code de transmission.

1. 1.2.1. La transmission des documents écrits ou dessinés

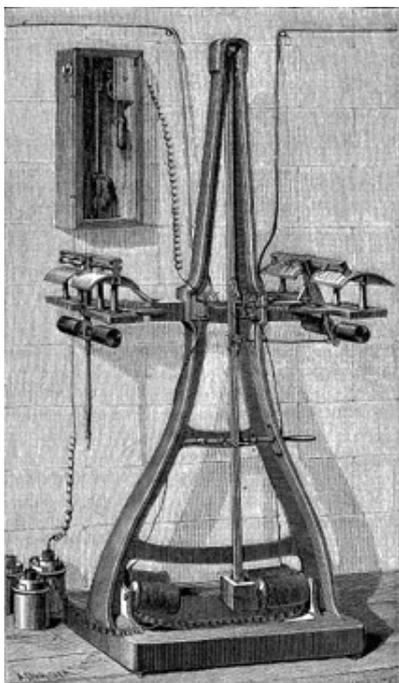
Le premier appareil capable de transmettre des images en deux valeurs, le noir et le blanc, est le télégraphe autographique (1862) inventé par l'abbé Caselli et perfectionné par Meyer. On l'appelle aujourd'hui télécopieur (fax), il servait à transmettre des textes et images que l'on traçait sur un papier couché rendu partiellement conducteur du courant par ce dessin.

Le télégraphe autographique émettait des signaux binaires qui passaient par les lignes télégraphiques (à cette époque, le téléphone n'existait pas). Les signaux binaires sont:

1 = passage du courant

0 = pas de courant.

Il est à noter que le codage binaire est le principe de fonctionnement des processeurs au cœur des ordinateurs



Il se retrouve comme procédé appliqué à la numérisation des images "au trait" que nous nommerons "bitonales" dans le cours qui suit. Le principe de Caselli est donc reconduit dans les machines actuelles.

La transformation du dessin en signaux électriques se faisant sur le pantélégraphe autographique de Caselli par un contact électrique direct, il manquait juste un perfectionnement : la cellule photoélectrique qui peut analyser la lumière réfléchie par n'importe quel original et la transformer en signaux électriques, sans contact physique.

2. 1.2.2. La transmission des photographies à distance

Dès le début du XXe siècle, des savants avaient découvert les propriétés du sélénium, point de départ des recherches sur les cellules d'analyse photographique, ce qui permit au Français Edouard Belin d'achever la mise au point en 1925 d'un appareil permettant la transmission, aussi bien par le câble que par ondes hertziennes, d'images fixes en niveaux de gris.



Grâce au "Bélinographe", la presse écrite a pu, pendant des décennies, publier des photographies d'actualité reçues instantanément du monde entier.

Le parler canadien a gardé la trace de cet ancêtre, car on appelle bélinographe le télécopieur que les français nomment volontiers "fax" ou "téléfax"

L'épilogue de cette histoire : la transmission intercontinentale des images par une ligne électrique a plus d'un siècle, mais il a fallu attendre la fin du XXe pour voir sa démocratisation!

Agrandir l'image (cf. 'Bélinographe Tpe B.E.L.1, (1928) collection F.B.' p 107)

3. 1.2.3. L'héritage actuel

Dans notre numériseur informatique, l'analyse de l'image est encore faite par un procédé dit analogique: le courant varie en fonction de l'intensité du gris, pour devenir maximal pour le blanc, au niveau des capteurs de lumière de l'appareil.

Ces signaux sont ensuite interprétés en codage binaire avant d'être envoyés à l'ordinateur.

La "discrétisation" du signal va transformer ces variations continues en une série de niveaux prédéfinis qui donneront une illusion de continuité à la restitution du document.

Ces niveaux sont caractérisés par une valeur numérique (par exemple entre 0 et 256), d'où le nom de numérisation de l'image.

Alors que par le passé, les photographies en couleurs étaient décomposées en plusieurs films en niveau de gris pour l'expédition au bélinographe, les systèmes actuels reproduisent directement en un même fichier les couleurs des objets de la numérisation. Cependant le principe reste le même, les trois ou quatre couches de couleurs sont juste rassemblées dans un même fichier.

C. 1.3. La bonne numérisation



Il faut bien avoir à l'esprit que le document numérique, comme le film ou l'enregistrement phonographique, est une reproduction de la réalité, et non la réalité.

Il peut être de qualité médiocre, telle une télécopie par exemple. Dans le domaine sonore, on pourrait comparer avec les nouvelles que l'on écoute sur un poste à transistor. La qualité sonore est médiocre, le message est cependant complet et compréhensible.

Il peut être de haute qualité au point de donner l'illusion de l'original, tout comme l'enregistrement d'un concert que l'on écoute dans un auditorium. La numérisation ne remplace pas le document d'archives, mais peut en garder une trace fidèle si on choisit bien son procédé.

On verra ci-dessous que :

- une bonne numérisation est une adéquation entre le document original, l'usage du fichier numérisé et la portabilité en réseau ;
- les objectifs de la numérisation et l'usage auquel on destine les documents numérisés sont déterminants pour le choix des procédés de numérisation.

2. Notions fondamentales en numérisation



l'archiviste.

Un certains nombres de connaissances sont indispensables pour comprendre les mécanismes de numérisation de l'image. La connaissance de ces pincipes est essentielle pour garantir la qualité du produit fini et sa portabilité (taille du fichier à diffuser en ligne)

L'archiviste a tout pouvoir de faire les choix fondamentaux pour sa numérisation, et la réussite dépend non de la performance des machines mais du discernement de

A. 2.1. Image en pixels



placement de l'image dans des compositions numériques (pages de magazines, de livres, etc.)

L'ordinateur étant une machine logique, il procède par un langage binaire 0 et 1. Il lui faudra donc décomposer l'image en éléments finis, appelés pixels (pour *Picture Element*) et affecter à chacun d'eux un code permettant de reproduire une couleur. Ce procédé est nommé discrétisation de l'image. Chaque pixel est un carré, qui se place dans une grille virtuelle définissant les limites de l'image numérique.

Ces limites sont celles d'un rectangle, permettant le



Image agrandie des pixels noirs et blancs, niveaux de gris et couleur.

Complément : Masque de détournage

Certaines images sont dotées d'un masque de détournage permettant de rendre transparentes certaines parties pour le placement sur un fond de couleur. Ce masque est un tracé vectoriel dans le monde de l'édition et est fabriqué dans les logiciels de retouche d'images professionnels. Dans la bureautique et l'imagerie "Web", on affecte à certains pixels la propriété de transparence, c'est la "couche alpha".



Personnage détourné, et converti au format Gif pour conserver cette propriété dans un navigateur web

B. 2.2. Résolution

Le nombre de pixels contenu dans cette grille, pour une dimension physique de l'original, va déterminer la **résolution** (appelée aussi définition) de l'image. Cette résolution est issue des pratiques de l'imprimerie, en pixels par pouce : ppp ou en anglais dpi pour Dot Per Inch. Plus il y aura de pixels par pouce :

- plus la finesse de la reproduction sera élevée
- plus on verra les détails
- mais aussi plus le fichier sera lourd.

Lorsque l'on double la résolution, la taille du fichier en pixels (ouvert dans une application) est multipliée par 4.

Résolutions des images en tons continus:



Aspects des résolutions 72, 150, 300 dpi pour un document à tons continus (agrandissement 4x).

Les résolutions en tons continus proposées ci-dessus sont caractéristiques des usages les plus courants: 72 dpi pour la diffusion en ligne (Web) , 150dpi pour l'impression domestique, et 300dpi pour la substitution et l'imprimerie à l'échelle originale. Les mêmes résolutions s'appliquent aux documents en niveaux de gris.

Résolution des images bitonales (au trait)



Aspects des résolutions 150, 300, 600 dpi pour un document bitonal (agrandissement 4x).

La résolution 150 dpi analogue à celle d'un télécopieur, demeure lisible pour les textes dactylographiés ou imprimés. Pour les textes manuscrits, elle peut être insuffisante

Les résolutions en mode bitonal proposées ci-dessus sont caractéristiques des usages les plus courants: 150 dpi pour diffusion en ligne, 300 dpi pour l'impression domestique ou les documents plus fins et 600dpi pour l'impression laser. L'imprimerie utilise des résolutions plus élevées: 1200 ou 2400dpi.

Choix de la résolution :

Les solutions de numérisation décrites dans ce cours sont basées sur la reproduction à l'échelle 1 des documents numérisés, ce qui est à la base du travail archivistique : restituer le document dans son intégrité, y compris le format.

C'est principalement le choix de la résolution du fichier numérique qui va déterminer sa qualité intrinsèque. Il est recommandé de numériser le document physique dans la plus grande des résolutions adaptée à l'usage que l'on en fera (par exemple celle de l'impression). Il suffira ensuite de générer des fichiers en plus basse résolution en utilisant les commandes de ré-échantillonnage des logiciels de traitement d'image pour les autres usages : cédérom éducatif, pages Internet, etc.

Les documents multi-pages en fichier PDF peuvent aussi être sous-échantillonnés en pratiquant une nouvelle impression PDF du document élaboré en haute définition. Il faudra choisir une plus basse résolution dans le réglage «compression» de l'imprimante PDF en activant l'option «sous-échantillonnage».

Le choix de la résolution est plus ou moins défini dans les divers cas envisagés dans ce cours (voir tableaux récapitulatifs), pour une première approche. On s'apercevra que le rendu final dépend aussi du type de document à numériser.

Il existe des moyens de calcul des résolutions optimales en fonction de la finesse du trait .

C. 2.3. Modes colorimétriques

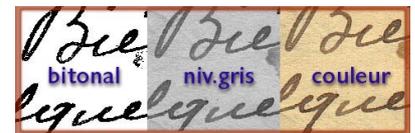


Pour le noir et blanc ou la couleur, les fichiers pixellisés peuvent être enregistrés dans différents modes. Le mode définit l'encodage que l'on va affecter à chaque pixel : il sera très simple pour un fichier représentant du noir pur et du blanc pur et complexe pour les fichiers représentant une image en couleurs: la tonalité, la luminance et la saturation de la couleur perçue par l'œil humain sont à traduire en valeurs numériques. Différents codages sont utilisés par les logiciels d'imagerie, mais la numérisation des documents

d'archives fait principalement appel à trois d'entre eux :

1. le mode bitonal ou bitmap pour la reproduction des documents et textes illustrés de gravures,
2. Le mode niveau de gris pour la reproduction des archives photographiques en noir et blanc,
3. Le mode RVB (rouge vert bleu) pour la reproduction de tous documents en couleurs.

Remarquer la persistance de la tache en mode bitonal : les densités dépassant le seuil de conversion apparaissent.



Complément : Interprétation de la réalité par le numériseur

Lors de l'acquisition de l'image du document par le numériseur, les capteurs image vont dans tous les cas analyser une image en couleurs, et le processeur interne au numériseur va transformer ce signal analogique en signal binaire compréhensible par l'ordinateur (la discrétisation).

- il fera une interprétation des couleurs en deux niveaux (noir et blanc) pour les images bitonales autour d'un *seuil* de densité c'est à dire qu'il va transformer toutes les valeurs au dessous d'une certaine densité en blanc et toutes les valeurs au-dessus de cette densité en noir. C'est le réglage pertinent de ce seuil de densité qui fixera la qualité du fichier fini.
- il fera une transformation des couleurs en niveau de gris, c'est à dire qu'il va transformer toutes les couleurs en valeurs de gris, comme le ferait une pellicule noir et blanc à la prise de vue d'une scène de la vie. Là encore, comme on règle le diaphragme sur un appareil photo, il faudra régler la densité moyenne de l'image ou le point noir/point blanc pour en restituer toutes les nuances.
- Il fera une décomposition de l'image en trois ou quatre couleurs fondamentales pour générer le fichier en couleurs, et enverra à l'ordinateur les informations traduisant les densités de chaque couleur fondamentale ou modèle mathématique représentant la couleur.

Complément : Les modes colorimétriques élaborés

Il existe plusieurs modes colorimétriques applicables aux documents en couleurs, outre le bien connu RVB (rouge vert bleu). Ils sont utilisés principalement par les imprimeurs et les créateurs d'images numériques.

- Le mode **CMJN** est utilisé par les imprimeurs pour les impressions en quadrichromie c'est-à-dire en quatre couleurs fondamentales, le cyan (proche du bleu), le magenta (rouge - rosé soutenu), le jaune et le noir. La combinaison de ces couleurs de base permet de reproduire sur le papier des images en tons continus dans un assez large spectre de couleurs.
- Le mode **LAB** (Luminosité, couleurs A et B) est utilisé par les numériseurs de haute qualité pour l'acquisition de diapositives, dont la profondeur de couleurs et les

différences de densité entre la zone la plus claire et la zone la plus foncée sont beaucoup plus élevées que celles des documents sur papier. Ce sera le mode à utiliser pour la préservation des archives photographiques susceptibles de se décolorer dans le temps. Le mode LAB sera la plupart du temps converti en modes RVB ou CMJN pour la diffusion ou l'impression de l'image.

- Le mode **couleurs indexées** à nombre de couleurs réduites (256 couleurs ou moins) sert à l'affichage sur la toile (web). L'échantillonnage des couleurs peut être adapté à l'image (256 nuances créées à partir des couleurs de l'objet numérisé) ou appartenir à une palette prédéfinie: 256 couleurs Windows system ou 256 couleurs Mac OS system. L'usage de ces palettes accélère le téléchargement; c'est surtout utile pour les très petites images comme les icônes, les vignettes...
- Le mode **BICHROMIE**, (avec ses corollaires trichrome, et quadrichrome) est une option de certains programmes professionnels permettant d'affecter directement des couleurs d'encre à un fichier à l'origine en niveau de gris. C'est ce mode que l'on utilise pour la simulation des images en sépia dans les publications. Les courbes d'encre sont réglables, ce qui permet de contrôler les nuances et tonalités de l'image.



Une image originale en niveaux de gris et son traitement en bichromie. Elle apparaît beaucoup plus attractive dans la publication, et une manipulation habile des courbes d'encre permet d'améliorer le rendu des détails.

- Le mode TLS (Teinte-Luminosité-Saturation) est utilisés comme interface par des applications de bureautique, mais le fichier enregistré est un fichier RVB
- Le Web utilise des codes littéraux pour décrire les couleurs incluses ou hors de la "palette système". Dans les applications de retouche photo, ce code est indiqué dans la fenêtre de dialogue du choix de la couleur. On peut indiquer ce code directement dans l'éditeur html, par exemple pour appliquer à un fond de page la couleur de fond d'une image qui y est intégrée. Le spectre de ces couleurs est celui du RVB (écran)

D. 2.4. Analyse et typologie des documents à numériser

Ce paragraphe analyse les types de documents que l'on peut rencontrer en vue de la numérisation.

A chacun de ces documents correspond un procédé physique de numérisation et un mode colorimétrique apte à traduire ses caractéristiques.

Pour déterminer ces procédés, il convient de faire une analyse générale de l'aspect du document et de déterminer sa catégorie ou typologie.

1. 2.4.1. Analyse du document

La forme du document à numériser doit être analysée avant toute numérisation :

- son format (sa taille) déterminera le type de numériseur à utiliser,
- le nombre de pages déterminera l'automatisation du procédé,
- son aspect déterminera le mode colorimétrique à utiliser qui permettra de traduire ses caractéristiques.

Attention

Les pages des ouvrages et catalogues contiennent souvent des typologies mélangées (texte pur, gravures au trait, photos en couleur ou niveau de gris); il conviendra de numériser

chaque page dans le mode convenable, puis d'assembler ces sous-fichiers en un seul pour reconstituer le document intégral.

Etre trop systématique dans les traitements de documents multi-pages conduit soit à des numérisations trop lourdes (tout en couleurs) ou dégradantes (tout au trait).

Après avoir parcouru les pages de ce cours, vous serez à même de comprendre cet avertissement ou de comprendre pourquoi les fonds que vous avez déjà numérisés sont si encombrants...

2. 2.4.2. Nature des documents déterminant la colorimétrie

Du point de vue de la numérisation et, par conséquent, de la **colorimétrie à utiliser**, la typologie des documents se décline en grandes familles:

- manuscrits en feuilles détachées, registres,
- textes imprimés, dupliqués ou dactylographiés,
- gravures au trait, illustrations au trait, dites "estampes" dans le monde des archives,
- simili gravures des ouvrages imprimés,
- héliogravures,
- lithographies et chromolithographies,
- photographies, diapositives, négatifs sur films ou plaques, microformes,
- dessins originaux et peintures,
- plans et dessins techniques,
- documents imprimés rassemblant du texte, des gravures, et des similigravures, voire des lithographies.

Les manuscrits et textes imprimés sont relativement faciles à numériser, quel que soit le rendu que l'on ait choisi pour les restituer (trait, niveau de gris ou couleur).

On s'apercevra par la pratique que les fonds iconographiques (photos, négatifs, diapositives, plaques de verre, peintures, chromolithographies, héliographies, plans en couleurs...) sont aussi relativement aisés à numériser, car ils font appel à des modes en demi-tons, faciles à retoucher et à adapter à l'usage à l'aide des logiciels de retouche photo. (toutefois, une bonne perception des couleurs par l'opérateur est nécessaire).

Les difficultés apparaissent dès lors que l'on a affaire à des procédés de reproduction imprimés comme la gravure au trait, la simili gravure et autres procédés photomécaniques d'imprimerie pour lesquels la réussite se décide lors de l'opération de numérisation. Les spécificités relatives à ce traitement nécessitent de la pratique et de la circonspection. Les retouches ultérieures sont malaisées, voire impossibles.

Enfin les plans au trait sont assez faciles à traiter, car les systèmes de numérisation bitonale utilisés actuellement dans l'architecture et l'industrie sont parfaitement adaptés. Le principal problème que l'on risque de rencontrer est celui de leur format, lorsqu'il dépasse la largeur de passage du numériseur, car les normes définissant les formats de plans ont été vraiment appliquées à partir des années 50.

3. Objectifs de la numérisation

Nous avons appris quelques notions essentielles, notamment combien le type de document à numériser influe sur le procédé de numérisation ou la colorimétrie .

Mais bien évidemment **il n'est pas question de numériser n'importe quel document**, encore faut-il **se demander pourquoi numériser**.

Pour répondre il faut tenir compte:

- de différents objectifs,
- de l'usage du fichier obtenu,
- de la portabilité du document.

A. 3.1. Objectifs

Les objectifs de la numérisation peuvent être:

1. un moyen de sauvegarde d'originaux menacés de dégradation physique (processus de vieillissement naturel),
2. la diffusion imprimée en haute définition (édition du livre),
3. la diffusion en ligne d'un document par ailleurs correctement conservé,
4. la destruction des documents papier sans valeur archivistique ni graphique, pour réduire le volume d'archives physiques (numérisation de substitution à des fins de gestion d'archives courantes et intermédiaires).

Chacun de ces cas va mener à des approches différentes, influençant directement **la résolution** et **la fidélité** des documents numériques.

- La résolution concerne la finesse d'analyse du document, c'est-à-dire la possibilité de montrer les détails les plus fins.
- La fidélité concerne la colorimétrie et donc l'aspect du document. La dégradation de cette colorimétrie affecte la fidélité si on utilise des formats d'enregistrement avec pertes de données (réduction de la palette des couleurs).

On choisira:

- une basse ou moyenne résolution et on tolérera des pertes de données pour les objectifs des catégories 1, 2, 3 et W définies ci-dessous (voir 3.4),
- la haute résolution et une grande fidélité des couleurs pour l'objectif de la catégorie 4.

B. 3.2. Usage du fichier final

L'utilisation du fichier final est déterminante pour choisir la qualité de la numérisation, mais là encore, le discernement de l'archiviste doit s'exercer en prévoyant les utilisations futures. On ne pourra pas transformer un document de basse résolution en haute résolution, tout comme on ne peut pas éditer un CD haute fidélité avec l'enregistrement d'un concert diffusé par un poste à transistor !

C. 3.3. Portabilité

La bonne numérisation va être basée sur des formats d'enregistrement compressés dont la taille naturelle ne nécessitera pas de compression ultérieure, ni sur le disque de stockage, ni pour l'envoi en réseau.

Attention

Il est à noter que les fichiers image compressés sont parfaitement optimisés et incompressibles. Il est donc inutile de refaire une compression Zip sur des fichiers jpeg, PDF et tiff dont les options de compression ont été activées à l'enregistrement.

Il est aussi inutile de compresser les pièces jointes d'un courriel, car c'est l'application dédiée qui s'en charge avant l'envoi (compression Zip en général).

D. 3.4. Répartition des documents numérisés selon l'usage

Selon nos besoins et l'usage que nous voulons en faire, nous pouvons répartir les documents à numériser en différentes catégories, sachant que certains peuvent relever de plusieurs catégories à la fois :

- textes pouvant simplement être lus sans exigence graphique,
- textes et images pouvant être consultés sans exigence graphique,
- ouvrages ou documents en téléchargement,
- documents iconographiques numérisés pour la sauvegarde ou la reproduction imprimée,
- documents iconographiques pour la toile (Web).

Catégorie 1 : textes pouvant simplement être lus sans exigence graphique

C'est le cas typique des **documents administratifs ou d'affaires** que l'on met en ligne ou que l'on archive numériquement pour une conservation correspondant à la durée d'utilité administrative. Pour ces documents, le critère principal est la portabilité (petite taille) permettant un flux rapide dans les réseaux et la diffusion par Internet. L'impression de ces documents a l'aspect d'une télécopie en qualité normale. Il n'y a que du noir et du blanc, c'est une reproduction **au trait ou bitonale**.

Catégorie 2 : textes et images pouvant être consultés, sans exigence graphique

Ce sont des **documents** que l'on met en ligne pour permettre une consultation rapide par des internautes ou utilisateurs d'un réseau interne (intranet). On accepte une dégradation des images par la basse définition et par la compression. Ce sont des documents qui sont acceptables à l'écran mais dont l'impression est médiocre. On notera comme exemple des aperçus d'ouvrages à consulter au centre d'archives, des comptes-rendus et bulletins de liaison en ligne.

Entrent aussi dans cette catégorie les documents que vous composez avec un logiciel de

publication d'amateur auxquels sont intégrées des images de documents d'archives à titre iconographique accessoire.

Catégorie 3 : documents ou ouvrages ou en téléchargement

Certains centres d'archives ou bibliothèques mettent en ligne des archives publiques ou des ouvrages anciens libres de droits, pour permettre un téléchargement et une impression par l'internaute; une perte de définition est admise sur les images, mais leur lisibilité est conservée. C'est une situation intermédiaire entre le simple texte et la qualité impression.

Cette même technique peut être utilisée pour la mise en consultation en réseau intranet de documents dont plusieurs services ont un usage courant . Il est possible de protéger individuellement ces documents par des clés et mots de passe avec certains formats d'enregistrement.

Catégorie 4 : documents iconographiques numérisés pour la sauvegarde ou la reproduction imprimée

Les documents dont on envisage de faire une copie de sauvegarde numérique, soit parce qu'ils sont menacés de destruction physique, soit parce qu'ils se dégradent, doivent être numérisés sans perte de données, et dans les meilleurs critères de qualité applicables à leur typologie.

Pour la copie de sauvegarde des iconographies sujettes à dégradation naturelle (photographies en couleurs), on choisira des modes colorimétriques avancés pour conserver les nuances d'origine. La préservation des diapositives devrait permettre au moins l'impression sur papier en 18x24 cm.

Catégorie W : documents iconographiques pour la toile (Web)

Cette catégorie englobe les images à basse définition que l'on rend disponibles en ligne pour les internautes, l'enseignement etc...et les images utilisées comme icônes d'accès aux fichiers haute définition d'une photothèque haute résolution en ligne.

Les formats de fichier sont des formats « image » et ce ne sont pas des documents multi - pages : chacune des images est traitée à part pour l'intégration à un document final. La résolution est basse et l'on accepte des dégradations importantes par la compression. Ces fichiers sont généralement issus d'un sous-échantillonnage à partir du fichier de numérisation "source" ou "maître".

E. 3.5. Usages mixtes

Plusieurs de ces catégories peuvent être applicables simultanément à un document faisant l'objet d'usage multiples.

C'est le cas pour une exposition à thème pour laquelle on utilisera une même image dans:

- une brochure en couleurs,
- des panneaux
- un site web avec téléchargement de documents et d'images.

La numérisation se fera en une seule fois dans **les plus hauts critères de qualité**, puis on génèrera des fichiers de moindre qualité pour les diffusions en ligne, par l'intermédiaire d'un logiciel de traitement des images. Cette opération s'appelle le sous-échantillonnage de l'image.

Attention

Il est cependant inutile de faire de la sur-qualité dans la numérisation des documents, car les espaces disque sont inutilement dévorés et le processus se trouve alourdi pour un résultat invisible. Chaque fois qu'on double la résolution, on multiplie la taille du fichier par quatre. Il est inutile de numériser, par exemple, les photos dans une résolution plus de deux fois supérieure à la taille du grain argentique.

4. Formats d'enregistrement des images numérisées

Lors de l'acquisition avec le numériseur, l'image est entrée dans l'ordinateur avec la qualité maximale que l'on aura définie lors de cette acquisition.

Ces images doivent ensuite être enregistrées pour être stockées sur un disque dur, un CD-ROM, un DVD ou un serveur pour la consultation en ligne.

Les formats d'enregistrement vont déterminer la taille (ou poids) du fichier, sa portabilité en ligne et son universalité de reconnaissance par les applications.

Il existe des formats d'enregistrement sans réduction de qualité de l'image, (sans pertes de données) et d'autres avec réduction de qualité (perte de données).

En aucun cas les formats engendrant des pertes de données ne doivent être utilisés pour la sauvegarde d'archives en voie de détérioration physique.

Attention !

Lorsque l'on acquiert une image il est conseillé de créer un fichier dit "fichier source" ou "fichier maître" qui sera conservé dans un format n'entraînant pas (ou peu) de pertes de données.

Ce sont des copies avec pertes de données qui seront à placer en ligne, afin de limiter les temps de téléchargement.

A. 4.1. Formats sans perte de données

Les formats sans pertes de données sont **à utiliser pour constituer la banque de fichiers source**, par exemple **dans le cas des photothèques**.

Ils sont pourvus d'algorithmes de compression, qui réduisent la taille du fichier enregistré, mais le restituent sans perte de données, ni en couleurs, ni en pixels.

De par leur nature même, ils seront utilisés dans tous les cas pour les fichiers en mode bitonal.

Le format **TIFF** (*Tagged Image File Format*) est plus particulièrement utilisé pour les documents mono-pages, les photothèques.

Le format **PDF** (*Portable Document Format*) est plus particulièrement utilisé pour les documents multi-pages ou nécessitant une protection contre certaines utilisations.

Complément : Le format TIFF

Ce format (Tagged Image File Format - extension .tif) est le premier venu dans le temps des formats universels.

Il permet une compression des fichiers par l'algorithme "LZW" ou "ZIP" sans dégradation de la qualité des couleurs ni des niveaux de gris.

Pour la compression des images bitonales, il utilise le moteur CCITT qui s'avère le plus performant dans cette situation.

C'est par essence le **format recommandé pour la sauvegarde des archives**, car il conserve la gamme complète des couleurs quels que soient le mode colorimétrique et le niveau d'encodage (de 1 à 14 bits par couche).

Le format TIFF supporte les modes colorimétriques bitmap bitonal), niveau de gris et les couleurs RVB, CMJN, et LAB.

Nous recommandons l'usage du TIFF:

- dans tous les cas pour les images au trait dites bitonales (bitmap) (gravures, textes imprimés),
- dans la plupart des cas pour les images en niveau de gris (photographies en noir et blanc),

pour les images couleur destinées à la conservation et à l'imprimerie (reproductions de photos, peinture, documents en couleur).

Les versions récentes du format TIFF ont une option pages multiples exploitable par les systèmes de reproduction, photocopieurs numériques, mais pas par les applications courantes.

Complément : Le format PDF avec compression non dégradante

Le format PDF est une enveloppe autour des formats d'images pixellisées cités dans ce cours: il peut contenir des images avec ou sans pertes de données, selon le moteur de compression que l'on choisira dans la boîte de dialogue.

Il est capable comme le tiff d'enregistrer les modes colorimétriques bitonal, niveau de gris et couleurs, dans les encodages supportés par le Tiff.

Mais il peut aussi enregistrer des images vectorielles au format .eps (*encapsulated Postscript*) issues des logiciels de création vectorielle ou des images multicouches (bichromies) pour la préresse, ainsi que des pages composées contenant du texte, des images pixellisées et des images vectorielles.

Pour enregistrer des images sans perte, on choisira entre les compressions CCITT ou ZIP pour les fichiers bitonaux

Attention !

Une image au format PDF ne pourra être intégrée à un document électronique que si l'application hôte est munie des traducteurs (filtres d'importation) adéquats. C'est le cas des applications de préresse.

Les applications bureautiques ne peuvent intégrer ces fichiers sans une conversion préalable dans une application de traitement d'image pixellisée, car ils sont dépourvus de fonctions de traitement des objets vectoriels PostScript.

B. 4.2. Formats avec pertes de données

Certains formats d'enregistrement entraînent des pertes de données. Le moteur de compression utilisé par ces formats réduit le nombre de couleurs mais pas le nombre de

pixels. La compression est destinée à réduire la taille des fichiers pour accélérer le téléchargement et la taille des documents finis.

Le format **JPEG** est destiné à la diffusion d'images à l'aide un taux de compression variable. On choisit un bon rapport qualité-taille pour l'usage que l'on veut faire du fichier. Il est capable de millions de couleurs.

Le format **GIF** est typiquement conçu pour la toile, mais peut rendre des services pour le traitement de documents monochromes. Il est limité à 256 couleurs.

Le format **PNG** a aussi été conçu pour la toile, bien qu'une nouvelle version propose une compression sans perte de données. Le **PNG 8** (8 bits=256 couleurs) est équivalent au GIF, le **PNG 24** a les fonctionnalités du TIFF et du JPEG, mais ces formats sont encore assez méconnus.

Le format **PDF** avec perte de données est un format de diffusion acceptant des documents multi-pages et / ou des documents que l'on veut protéger contre certaines utilisations.

Complément : Le format JPEG

Le format JPEG (*Joined Photographs Expert Group* - extension .jpg) a été créé pour permettre la diffusion sur toutes les plates-formes et en ligne des images numériques en couleur. Son algorithme permet une compression progressive en 10 à 12 niveaux selon les logiciels, avec une dégradation invisible (niveau 12) à importante (niveau 0). Ce format ne supporte pas le mode bitonal (bitmap).

Le format JPEG supporte les modes niveau de gris, couleurs RVB et CMJN.

Nous recommandons l'usage du JPEG:

- pour les images destinées au téléchargement, pour lesquelles on veut régler la compression afin d'optimiser la taille, en mode niveau de gris et RVB,
- pour les images en basse définition destinées au web,
- pour les images que l'on veut envoyer par courriel (mail), dans la mesure où l'on maîtrise le rendu des couleurs lors de la compression,
- pour la bureautique.

La compression par le format JPEG réduit le **nombre des couleurs** de pixels, mais ne **réduit pas le nombre des pixels** formant l'image.

Les niveaux de qualité «maximale» n'altéreront pas visiblement la qualité de l'image, cependant les nuances de couleurs seront réduites, ce qui peut être gênant pour le traitement ultérieur de l'image en retouche ou interprétation artistique (travail des couleurs, ou agrandissement par rééchantillonnage).

Les niveaux de qualité «supérieure» commencent à altérer visiblement l'image si on fait un agrandissement: Il apparaît des «plaques» de même couleur très gênantes dans les dégradés (ciels) et dans les reflets sur les objets lisses.

Les niveaux de qualité «moyenne» accentuent les défauts précédents et font apparaître du «sable» à la frange des zones de contraste.

Les niveaux de qualité «basse» font apparaître un genre de tissage de panier en accentuant les défauts cités précédemment.

Attention !

Ce niveau de compression ne peut être utilisée que pour de l'affichage de page Internet.



Niveaux de dégradation des images JPEG

Le procédé de compression consiste à créer des plages de pixels de couleurs approchantes et d'en faire une moyenne.

La figure ci-dessus montre un agrandissement des pixels de l'image: dans les faits, ces phénomènes sont beaucoup plus petits et souvent peu discernables à l'écran, par contre ils transparaissent à l'impression.

Complément : Le format GIF

Le format **GIF** est réservé à l'usage sur le web. Il sous-échantillonne les couleurs à 256 niveaux, mais peut produire des fichiers plus lourds que le jpeg. En revanche, il présente la possibilité de définir des zones transparentes, ce qui est utile pour le détourage d'objets présentés en pages web. Il n'a pas d'intérêt direct pour les activités de numérisation des archives, mais seulement pour le webmestre chargé de leur diffusion Internet. L'enregistrement des numérisations initiales se fera dans les formats tiff ou jpeg, et une exportation sera faite vers le gif. Des «plug-ins» d'optimisation existent aussi pour ce format.



A noter: cette figure contrairement à celle du format jpeg n'a pas été agrandie, car son usage est réservé au navigateur web, et dans ce cas, c'est l'illusion qui compte !

L'image au format gif avec la palette de couleur «système» accélère le téléchargement au détriment de la justesse de l'image. Avec la palette de couleurs «adaptive» l'image est à peine plus longue à télécharger, mais plus précise. Le Gif a un rendu assez inattendu sur des documents monotones comme les reproductions de manuscrits. L'exemple ci-dessus est flatteur car c'est une prise de vues en haute définition à 300dpi, et les grandes plages d'une même couleur sont inexistantes. Le rendu du papier reste cependant très approximatif.

Complément : Le format PNG

Le format **PNG** (prononcer ping) est un format non propriétaire qui conquiert peu à peu le monde du web, grâce à ses qualités intrinsèques: il occupe 5 à 25% de place de moins qu'un GIF et avec une qualité supérieure. Le png peut évoluer dans deux mondes: du 256 couleurs analogue au gif et dans peu de temps il sera appliqué aux milliards de couleurs (64 bits) applicable à l'analyse des diapositives. C'est dans cette gamme qu'il conviendrait aux numérisations d'archives, en alternative au format tiff.

Son algorithme de compression n'altérera pas les images (compression sans pertes) dans les niveaux de qualité les plus hauts et de ce fait il pourrait tendre à supplanter le jpg. Il faudra encore attendre un peu pour le trouver avec ces fonctionnalités sur toutes les applications en matière de retouche photo. Mais là encore, l'optimisation d'un fonds numérique nécessitera des connaissances de la part de l'opérateur quant au choix du taux de compression.

Par contre, le format jpeg permet des compressions plus importantes, et une lecture plus

universelle, dans le monde du Web.

Complément : Le format PDF avec compression JPEG

Il existe des extensions (plug-ins) à installer dans vos logiciels de retouche d'image qui rendent lisse et sans «sable» ni «paniers» les compressions JPEG, dans une taille inférieure à la compression classique (par exemple: Boxtop pro JPEG de Boxtopsoft). Leur fenêtre de dialogue montre en détail la texture de l'image et l'utilisateur peut pousser son taux de compression en fonction du résultat. Cette méthode peut être appliquée pour la conservation de grands fonds iconographiques pour gagner un espace disque conséquent par rapport au TIFF ou à une compression à taux fixe prédéterminé. Cependant cette solution ne peut être mise en œuvre que par un opérateur très averti et distinguant correctement les couleurs. Il peut être nécessaire de pratiquer un test de détection du daltonisme.

Complément : Amélioration des rendus du JPEG

Il existe des extensions (plug-ins) à installer dans vos logiciels de retouche d'image qui rendent lisse et sans «sable» ni «paniers» les compressions JPEG, dans une taille inférieure à la compression classique (par exemple: Boxtop pro JPEG de Boxtopsoft). Leur fenêtre de dialogue montre en détail la texture de l'image et l'utilisateur peut pousser son taux de compression en fonction du résultat. Cette méthode peut être appliquée pour la conservation de grands fonds iconographiques pour gagner un espace disque conséquent par rapport au TIFF ou à une compression à taux fixe prédéterminé. Cependant cette solution ne peut être mise en œuvre que par un opérateur très averti et distinguant correctement les couleurs. Il peut être nécessaire de pratiquer un test de détection du daltonisme.

C. 4.3. Choix des formats d'enregistrement selon l'usage

Nous avons vu ci-avant combien le type de document à numériser influe sur le procédé de numérisation ou colorimétrie .

Nous avons dit que pour choisir le procédé adéquat il faut tenir compte de l'usage que l'on veut faire du fichier .

Ce chapitre fait un inventaire des **règles de base à respecter pour réaliser une photothèque numérique** (caractérisée par des **fichiers individuels**) et **une bibliothèque numérique** (caractérisée par des **fichiers à pages multiples**).

- **Définition de la photothèque numérique** : dans notre usage, on appellera **photothèque toute collection de documents à page unique**, soit
 - les photographies sous toutes leurs formes,
 - les cartes postales,
 - les dessins, tableaux, gravures,
 - les registres ou documents qui sont décomposés en pages uniques servies par une base de données,
 - les plans et dessins techniques.

La numérisation se fera dans un format image natif comme le tiff ou le jpeg pour chacun des documents et individuellement.

- **Définition de la bibliothèque numérique** : la bibliothèque numérique comprendra les documents multi-pages que l'on téléchargera par sections ou entiers, soit
 - les ouvrages et catalogues,
 - les actes,
 - les registres traités en sections,

- les thèses, rapports et autres documents multi-pages,
- certains plans techniques multi-folios,
- dans certains cas, les séries de cartes postales (dont l'original est monté en accordéon).

La numérisation se fera dans un format d'assemblage comme le PDF.

1. 4.3.1. Création d'une photothèque numérique

Il est possible de créer une photothèque éditée dans un format image. Il convient d'adapter le choix du format à l'usage que l'on souhaite en faire: sauvegarde ou mise en ligne sur la toile.

Dans une photothèque, il y aura un fichier par image, et l'on pourra télécharger ce fichier pour l'intégrer à une composition d'imprimerie, une page web, un fond d'écran.

Il est important d'enregistrer les fichiers dans des formats courants et reconnus par tous les logiciels "de base" du monde Apple et du monde PC.

Les formats les plus courants comprennent des algorithmes de compression permettant de réduire la taille du fichier enregistré sur le disque.

La compression peut s'accompagner d'une perte de données par le truchement de la réduction du nombre des couleurs existant lors de l'acquisition par le numériseur. (voir ci-avant dans la description des formats)

a) Création de la photothèque de sauvegarde ou "source"

Pour cette raison, il est conseillé de constituer une **photothèque de base** dont les éléments sont enregistrés dans un format **sans perte de données** (Tiff par exemple).

Cette photothèque n'est pas obligatoirement en ligne; elle peut être stockée sur un disque dur, des DVD et chacune des images source en **haute définition** qui la composent ; elle peut être mise à disposition seulement **à titre payant**. Les fichiers devraient contenir des indications sur les droits de reproduction pour l'édition imprimée (enregistrer ces informations dans l'image).

b) Création de la photothèque en ligne

A partir de cette photothèque originale, on va générer des **copies à basse résolution**, auxquelles on appliquera des **compressions importantes**, pour accélérer le téléchargement. Les formats seront le Jpeg, le gif, voire le PDF si l'on veut protéger le document contre certains usages.

La **qualité** de ces copies doit être suffisamment **médiocre** pour empêcher leur utilisation dans le monde de l'édition imprimée.

Ces copies destinées au web réduisent pour la plupart le nombre des couleurs ; cette perte de données est irrécupérable à la réouverture du fichier.

On peut intégrer à ces images le texte du copyright, en indiquant la provenance, le nom du centre d'archives, etc.

Dans cette catégorie d'images, on générera aussi une série d'icônes ou vignettes à placer dans la page web qui donne accès aux images en basse définition et/ou au bouton de commande pour une image haute résolution.

2. 4.3.2. Création d'une bibliothèque de documents à pagination multiple

Une bibliothèque rassemblera une série de documents manuscrits, registres ou imprimés que l'on décide de diffuser en ligne.

Mises à part les photographies ou iconographies faites pour un produit fini (l'imprimé) qui doivent être d'une haute définition, la numérisation des documents d'archives sert généralement simplement à une consultation à distance. La numérisation se fera donc dans des formats avec pertes de données pour accélérer le téléchargement.

Dans un système informatique, les formats courants d'enregistrement d'images permettent d'éditer une seule image par fichier. Comme notre numérisation source sera faite d'un fichier image par page numérisée, il faudrait diffuser une suite de 250 fichiers pour un ouvrage de 250 pages, ce qui est en pratique inutilisable.

Pour rendre la consultation possible et aisée, il est **indispensable de rassembler ces images primaires**.

Il existe deux solutions:

- Créer une base de données capable d'assembler les pages et de les fournir en un bloc à l'utilisateur.

Il existe une multitude de solutions de bases de données, mais elles sont pour la plupart prisonnières de l'environnement du centre d'archives ; certaines nécessitent même l'achat ou le téléchargement d'un progiciel de lecture pour l'exploitation à distance. Certaines de ces bases de données mettent les fichiers natifs dans un "paquet" (exemple le format ZIP) que l'on doit "déballer" à la réception. Au final, le chercheur aura encore une foule de fichiers non rattachés qu'il devra ouvrir un à un.

- Utiliser un format permettant l'assemblage direct des pages et supporté par un logiciel lecteur standard, multi plates-formes et gratuit.

Ce format universel pour la gestion du multi pages, c'est le format Adobe PDF (Portable document Format) qui permet l'assemblage de documents divers dans un même fichier. Certains logiciels de base de données sont capables de faire cet assemblage en partant de pages individuelles ou de chapitres au format PDF, de sorte que l'internaute reçoit un fichier monolithique contenant toutes les pages qu'il a demandées. La diffusion en PDF est donc la plus adéquate pour les documents à pages multiples.

Attention

Le format PDF n'est plus considéré par les spécialistes comme un format propriétaire car les clés de développement complètes sont distribuées par Adobe.

Les trois paragraphes qui suivent décrivent la méthodologie pour constituer un document multi-page en ligne:

- *l'acquisition des pages,*
- *l'assemblage de ces pages,*
- *l'habillage du document et sa protection,*
- *la recopie du document en basse définition pour sa diffusion en extranet,*
- *notion sur les robots d'assemblage.*

4.3.2.1. Numérisation des éléments à assembler

Lorsque les pages d'un document sont à assembler en un fichier PDF, il est recommandé de les enregistrer dans ce format avant assemblage.

a) Enregistrement de fichiers image au format PDF

Le format PDF est l'un des formats d'enregistrement des logiciel professionnels de retouche d'image, et dans ce cas, on peut enregistrer l'image directement au format PDF en vue de l'assemblage dans un fichier multi-pages.

L'enregistrement avec pertes de données est proposé pour augmenter la portabilité: on choisira l'option compression jpeg et le niveau de compression jugé nécessaire pour la diffusion en ligne.

Le format PDF permet aussi de conserver les documents sans perte de données, en sélectionnant le procédé de compression des images lors de l'enregistrement (ZIP ou LZW), mais dans ce cas, la taille des fichiers est plus importante.

b) Adjonction de pages de commentaires

On peut avoir à compléter un document multi-pages par des commentaires, un résumé, la cote, la description...

La génération d'un fichier PDF pour une application informatique courante (comme Word) se fait par le truchement d'une "imprimante virtuelle" dite "Acrobat distiller" qui est accessible à travers la boîte de dialogue d'impression, lorsqu'on a installé le logiciel Adobe Acrobat complet.

Dans ce qui suit, nous parlerons "d'impression PDF" pour la génération des fichiers à partir d'une application informatique quelconque ou du logiciel Acrobat lui-même.

4.3.2.2. Assemblage des fichiers au format PDF

Le format PDF (Portable Document Format) créé par Adobe est actuellement le seul capable de mélanger tous les types de fichiers numériques et de restituer dans n'importe quel environnement (Mac, PC, Unix, OS2) le document dans son aspect exact, en conservant le placement, les caractères et les couleurs d'origine.

C'est un format «propriétaire» pour l'assemblage des documents (il faut acquérir le logiciel Adobe Acrobat complet), mais est **libre en mode lecture seule** par le truchement du logiciel Acrobat Reader installé d'origine dans la plupart des systèmes d'exploitation actuels ou disponible sur Internet, ainsi que dans les CD-ROMs de magazines d'informatique. Le format PDF est un standard reconnu et qui perdurera du fait même de son origine — le monde de l'imprimerie — et de son implantation dans toutes les plateformes. Par ailleurs, Adobe livre les codes de développement pour générer soi-même du PDF au sein d'une application.

Le format PDF supporte l'assemblage en un unique fichier des documents au format PDF quelle que soit leur taille de restitution (format A4, mélangé avec du A3, des cartes postales, des affiches de grand format...), quel que soit leur mode colorimétrique (bitonal, niveaux de gris, couleur RVB, CMJN...), quelle que soit leur résolution et quelle que soit leur provenance (image provenant de numérisation, document de prépresse ou de bureautique exporté en PDF).

Les fonctionnalités permettent de recadrer les pages une à une ou en bloc, d'orienter les pages, de remplacer des pages, d'ajouter des commentaires au document sous forme de "post-it" virtuels.

Le point de vue de l'éditeur Adobe l'archivage au format pdf

4.3.2.3. Liens dynamiques et protection individuelle

Liens dynamiques du document PDF

Acrobat permet d'ajouter des liens dynamiques (URL ou hypertexte) vers les pages du document lui-même, vers un site internet, ou vers un autre document placé sur un serveur local ou distant.

Protection du document PDF

Le format PDF permet d'inclure au document des protections partielles ou totales. La copie du document hors de son serveur d'origine ne déverrouille pas ces protections, le document garde donc à tout moment et en tout lieu sa protection originale.

La mise en service d'un mot de passe peut verrouiller l'accès complet au document, et un autre mot de passe peut verrouiller individuellement les fonctions de la liste ci dessous :

- l'impression
- la copie d'éléments texte et images
- la modification des pages ou du contenu texte
- l'adjonction d'annotations.

La mise en place du mot de passe déclenche le cryptage en 40 ou 128bits RCA du

document, et le protège efficacement contre les piratages informatiques.

Cette disposition est intéressante pour **protéger les fonds d'archives à un niveau juste nécessaire**, sans pénaliser les utilisateurs, ni favoriser les plagiaires.

4.3.2.4. Recopie des documents en basse résolution

C'est aussi avec l'imprimante PDF que l'on peut générer en une seule opération un document à basse résolution pour le web en faisant une réimpression PDF du document source en haute résolution: il suffira de régler les propriétés de l'imprimante PDF pour la basse résolution désirée (livre électronique ou écran) pour obtenir instantanément un document léger et transportable.

Si l'on veut garder **quelques pages** seulement de ce document **en haute définition**, il suffira de les copier du document source et de les réintégrer dans le document allégé par la commande de remplacement de pages.

La réimpression en PDF ne détruit pas les liens hypertexte que l'on a pu placer.

4.3.2.5. Assemblage par un robot en ligne

Le découpage et l'assemblage d'un document peut se faire sur demande de l'utilisateur, de la page n à la page n+x, par l'intermédiaire d'un robot d'assemblage (exemple: Bibliothèque Nationale de France - Gallica; voir liens et exemples à la fin de ce module).

Le développement de cette solution nécessite des compétences informatiques assez importantes.

On peut créer un assemblage de fichiers image à travers une banque de données et déclencher l'impression finale à travers une simple imprimante Acrobat Distiller en ligne.

Le meilleur robot sera celui qui assemble directement des pages PDF pour constituer le document particulier. Pour créer ce robot, il faudra utiliser les codes source Adobe.

D. 4.4. Formats libres et formats propriétaires

Formats libres, formats propriétaires: que choisir ?

Avant de numériser un fonds, il est nécessaire d'effectuer un choix des formats d'enregistrement pour s'assurer de la pérennité de ce fonds.

Dans ce qui suit, les avantages et inconvénients de chacun sont énoncés.

1. 4.4.1. Formats libres

Les formats les plus courants sont libres, c'est à dire ne sont pas propres à une application et ne donnent pas lieu à paiement de redevance.

Ces formats «libres» ont des codes sources publics (dont JPEG, TIFF, PDF) que chacun peut intégrer dans une application spécifique, comme une base de données à éditeur intégré. Les formats les plus connus sont très complets, lisibles par toutes les applications et ...au point.

Attention

Il existe depuis l'avènement du système LINUX des applications gratuites qui travaillent sur ces formats reconnus ou génèrent eux-mêmes des formats libres. Dans ce second cas, les formats ne passent pas à travers les filtres des autres applications et deviennent de fait "prisonniers" de l'application.

Tant que faire se peut, on choisira des applications libres générant des formats libres mais

reconnus.

2. 4.4.2. Formats propriétaires

Les formats «propriétaires» sont ceux développés par une marque et dont les codes source ne sont pas publics. En principe, il n'est pas possible de modifier ou d'adapter soi-même ces formats pour son usage propre.

Bien souvent ces formats ne sont lisibles que par la seule application d'origine. Certains formats propriétaires sont cependant lisibles par des applications connexes ou faisant partie du package de l'éditeur original.

Attention

Certains vendeurs de **systèmes de numérisation clés en main** proposent des systèmes «propriétaires» à des coûts relativement élevés. Souvent ces systèmes n'utilisent que des formats standard (voire libres) cachés par une clé qui empêche la lecture sur d'autres applications. L'utilisateur est cantonné dans des schémas qui ne correspondent pas forcément à ses besoins. On ne choisira dans ces systèmes que ceux qui sont capables de fournir des fichiers aux formats standard énoncés ci-dessus, et on fera un test d'intégration des fichiers générés dans des applications reconnues comme Microsoft Word, Adobe Pagemaker, Quark X'press, Adobe In Design et, bien sûr, on essaiera de les ouvrir dans Adobe Photoshop ou Graphic Converter pour savoir s'ils sont corrects.

Il faut aussi avoir un accès direct et individuel à chacun des fichiers pour pouvoir les copier et les sauvegarder sans avoir à passer par la base de données du système: en effet en cas de bogue dans la base de données, l'ensemble des fichiers peut être perdu et tout le travail de numérisation à recommencer!

Un système de numérisation clés en main peut aussi créer des fichiers PDF pour la diffusion en pages multiples. Pour connaître les propriétés exactes d'un fichier PDF, il suffit de les afficher par Fichier > Informations sur le document du logiciel Adobe Acrobat Reader. Il faut prêter attention aux propriétés de la protection, car si l'automate introduit un mot de passe aléatoire, il vous sera impossible de déverrouiller ces fichiers ultérieurement pour les assembler, les annoter, les modifier.

Attention

Il existe des formats « prisonniers » qui ne sont lisibles que par une seule application, par exemple certains logiciels de numérisation d'archives ou certains éditeurs de logiciel de retouche photo livrés avec les numériseurs. Les images enregistrées dans ces formats peuvent devenir irrécupérables en cas d'altération de l'application originale ou de changement de système d'exploitation.

E. 4.5. Microfilm et numérisation : des champs d'application connexes

Aussi incroyable que cela paraisse, la volatilité des supports de stockage informatique ont donné une nouvelle vie aux procédés de micrographie (microfilmage-microfichage). Les microfilms ont une durée de vie supérieure (et prouvée) aux supports informatiques.

Quand la microforme constitue une archive informatique

Plus on dématérialise les documents, plus le flux d'information virtuelle s'intensifie, plus on ressent le besoin de sécuriser les informations importantes.

La micrographie est à la conquête d'une nouvelle application : **la sauvegarde informatique**.

Le microfilm "numérique" (entendre par là un microfilm argentique qui a été impressionné par un laser numérique) s'impose désormais comme le moyen le plus sûr d'archiver à long terme des informations provenant de systèmes informatiques (les systèmes de banque, de gestion de caisse...). La matérialisation de ces informations les met à l'abri de l'effacement par inadvertance, de l'incompatibilité informatique et de la fragilité des supports d'enregistrement informatiques.

Ce champ d'application couvre les documents juridiques et légaux, les formulaires, documents sociaux, plans et documents au trait, mais ne s'applique pas bien aux documents contenant des photographies, de par la nature même de la microfiche.

L'information sur un microfilm est garantie pour un siècle, les laboratoires de R&D KODAK travaillent actuellement sur un support argentique de 500 ans d'archivage.

Ces sauvegardes sont effectuées à travers des traceurs micrographiques écrivant sur les microformes normalisées.

On assiste donc à un enregistrement simultané sur deux supports de conservation:

- le support informatique pour la diffusion, l'accès rapide, que l'on pourrait assimiler à une archive tampon,
- le support d'archivage permanent (de sauvegarde) sur des microfiches argentiques.

Attention

Les documents informatiques stockés en microfiches perdent leur nature même de document numérique, car ce sont seulement des **images** du document qui sont enregistrées sur la microfiche. La renumérisation de la microfiche ne redonnera pas vie instantanément à un document texte ni à une base de données!

Numérisation des microfiches

Inversement, les archives permanentes enregistrées sur microfiches, sont maintenant numérisées pour permettre leur exploitation en ligne. Les plans techniques, les registres, les formulaires ayant fait l'objet d'un microfilmage "remontent à la surface" par le biais de ce procédé.

Le point de vue de l'archiviste

Il faut respecter scrupuleusement les normes de microfilmage lors de la sauvegarde de documents informatiques sur les films, en particulier respecter la taille des caractères pour assurer la lisibilité. Alors que certains documents ont été composés avec des caractères ne satisfaisant pas les normes de microfilmage (le caractère Times par exemple, installé par défaut par Microsoft), il sera peut-être nécessaire de faire un pré-traitement de mise en conformité-conversion des textes en polices conformes aux prescriptions de microfilmage (Helvetica ou Arial par exemple), ou changer de format de microfiche pour assurer la lisibilité.

5. Numériseurs

Les numériseurs vont faire l'objet de notre attention dans ce chapitre.

Le numériseur lui-même ne peut pas dans tous les cas être utilisé seul, il faut lui adjoindre un ordinateur pour l'enregistrement et le post-traitement des images.

Les formats conseillés ci-avant franchissent toutes les plateformes (Mac, PC, Unix), le choix du poste de travail se fera dans l'un de ces trois mondes. Il est à noter que l'environnement Mac est spécialement conçu pour le traitement d'images, c'est pourquoi il est omniprésent dans le monde de l'imprimerie et de l'audiovisuel.

A. 5.1. Format des originaux déterminant le numériseur

Il faudra en premier lieu faire un récolement des fonds et d'établir un programme de la numérisation. En effet, il est possible d'étaler l'investissement en commençant par numériser les originaux les plus demandés, puis procéder à un complément d'équipement au fur et à mesure de l'avancement dans les fonds. Il sera utile :

- de définir les priorités de numérisation des documents selon leur typologie, leur état et leur risque de détérioration,
- d'évaluer la fréquence de consultation des documents selon la demande du public,
- d'évaluer les courants et la fréquence d'utilisation des fichiers (le téléchargement, la consultation à l'écran, l'impression...) relatifs à une typologie d'archives,
- de déterminer les besoins exacts en qualité, si la numérisation n'est pas une sauvegarde, mais un moyen de diffusion d'archives correctement conservées.

A partir de ces données et des formats physiques des documents, le choix du numériseur se fera en fonction de l'occurrence des situations.

Exemple

Un fonds comprend 200 mètres linéaires de plans A2, A1 et A0, 50 mètres de photos d'art sur plaques négatives 13x18 et 18x24 cm, 300 000 pages A4 de documents divers en bon état et non fragiles, 1m linéaire de négatifs 24x36 et 25 m linéaires de photos diverses et cartes postales, quelques registres et quelques ouvrages.

La demande annuelle est de 250 reproductions de plaques photo, 50 plans, 300 documents et 75 photos sur papier.

a) Il faudrait acheter en premier un numériseur à plat avec un dos lumineux capable de reproduire les photos sur plaques; il sera dans ce cas au format A4 car il n'y a pas dans ce fonds de formats nécessitant une plus grande glace d'exposition. Il pourra aussi traiter les documents papier (en moyenne un par jour) par ce biais et les photos sur papier. Les plans seront confiés à un reprographe.

b) Un numériseur à plans pourra être acquis en second lieu pour satisfaire aux demandes de numérisation de plans.

c) Le petit nombre de négatifs 24x36 ne justifie pas l'achat d'un numériseur à négatifs (il vaut le prix de 3000 retirages sur papier)

S'il devient nécessaire de numériser les 300 000 formats A4, il faudra faire l'acquisition d'un numériseur automatisé passant les documents par un chargeur automatique et délivrant directement les fichiers assemblés.

B. 5.2. Le choix du numériseur

La gamme des appareils d'utilisation courante existant sur le marché est adaptée à la numérisation des originaux actuels. Ces appareils permettent de numériser tout ou presque jusqu'au format A3, dans tous les modes colorimétriques. Les numériseurs d'imprimeurs peuvent aller jusqu'au format A2, et on peut facilement numériser en mode bitonal jusqu'au format A0 avec des appareils destinés aux bureaux d'études.

Au-delà de cette gamme, il est nécessaire d'acquérir des numériseurs étudiés spécialement pour le monde des archives. Quelques revendeurs présentent les gammes de produits complètes, et une assistance technique sur site (entretien des machines de numérisation). C'est ce service qui coûte le plus cher dans leur proposition, mais c'est aussi la clé d'une numérisation sûre.

Lorsque les originaux sont plus grands que le format A3, il existe des numériseurs de grand format, soit conçus comme les numériseurs à plans (l'original défile entre des rouleaux, et doit être en papier souple), soit conçus spécialement pour la numérisation des livres, en couleurs ou en monochromie. Ces appareils permettent de numériser des formats jusqu'à A2, A1 et A0, et comportent pour certains une option livre demi-ouvert pour ne pas abîmer les reliures.

Ces appareils délivrent généralement des fichiers aux formats tiff - jpeg, et il est nécessaire de pratiquer un post-traitement pour l'assemblage des pages, si l'on traite des registres ou ouvrages reliés.

C. 5.3. Types de numériseurs

Il existe plusieurs familles de numériseurs :

a) **les numériseurs à plat** pour documents opaques et transparents dont il existe des versions domestiques, fonctionnant en tous modes colorimétriques; ils vont du format A4/letter américain au format A3 et jusqu'au format A2 dans le monde de l'imprimerie;

b) **les numériseurs automatisés** passant les documents par un **chargeur** automatique et fabricant directement des fichiers multi-pages (tous modes colorimétriques); ils sont au format A4/letter américain;

c) **les numériseurs pour diapos et négatifs** de format inférieur à 5x5 cm, tous modes colorimétriques;

d) **les numériseurs à livres** jusqu' au format A0, au mode bitonal, et niveau de gris, et couleur pour certains; leur profondeur de champ permet de traiter des objets non plans;

e) **les numériseurs à plans monochromes** jusqu'au format A0 , à défilement, qui

acceptent seulement des documents souples; ils délivrent des fichiers au mode bitonal, parfois en niveau de gris;

f) **les numériseurs à documents en couleur** jusqu'au format A0 , à défilement, qui acceptent seulement des documents souples; ils délivrent des fichiers au mode bitonal, en niveaux de gris ou en couleurs;

g) **les numériseurs à tambour** pour les diapos de format supérieur à 5x5cm tous modes colorimétriques; ces appareils très précis et coûteux sont dédiés à l'usage des agences de publicité et grandes imprimeries; nous les citons pour mémoire, leur emploi est exceptionnel dans le monde des archives.

D. 5.4. Analyse des caractéristiques des numériseurs

Les principales performances des numériseurs sont exprimées par:

- La résolution optique, caractéristique du nombre de cellules du capteur CCD (et non de sa résolution logicielle qui est le résultat d'une extrapolation). La résolution optique est indiquée par le constructeur. La résolution optique pratique devrait être 3 fois l'épaisseur des traits les plus fins des gravures à analyser, soit en pratique au minimum 1200 dpi par 1200 dpi.
- La résolution optique des numériseurs à négatifs ou des numériseurs à plat que l'on utiliserait pour des négatifs de moyen format doit être plus élevée, car, dans tous les cas, on fera un agrandissement de l'image de l'ordre de 7 à 10 fois pour l'utilisation. La définition optique minimale devrait être de 2400 dpi x 2400 dpi.
- La capacité à **distinguer et restituer les couleurs**. La profondeur d'analyse des couleurs est indiquée par le constructeur en bits (nombre de signaux d'entrée utilisés pour modéliser la couleur d'un pixel). On trouve des numériseurs de 24 à 42 bits, soit de 16 millions à des milliards de couleurs possibles pour chaque pixel. Une profondeur de 24 bits, soit trois fois 8 bits en RVB et 256 niveaux par couleur, convient à la reproduction des documents d'archives opaques. Il est nécessaire de numériser à 42 bits uniquement pour les diapositives ou les tableaux.

Mais la véritable distinction des couleurs n'est pas quantifiée dans les notices, c'est le logiciel de conversion qui fait la différence (voir ci-dessous).

- La capacité à restituer les **différences de densité**, par exemple lors de l'analyse des diapositives qui ont de grands contrastes entre les zones les plus claires et les zones les plus foncées. Le constructeur annonce ce rapport par un nombre suivi de D (par exemple 4,2 D pour un numériseur de haut de gamme). Cette caractéristique est essentielle en archives si on veut explorer les détails des plaques photo en verre (en noir et blanc) qui présentent d'importants contrastes à cause de l'épaisseur de leur émulsion.

Attention : Le piège des performances «commerciales»

Les fabricants de numériseurs ont tendance à mettre en avant les critères de résolution de leurs appareils, voire une vitesse d'acquisition record. Ce sont des pièges commerciaux.

La véritable qualité du numériseur est celle du logiciel de conversion et d'acquisition, qui va du pire au meilleur selon le fabricant, et ceci à un même niveau de prix. La conversion règle la justesse des couleurs, et corrige point par point les défauts inhérents aux capteurs et à la lampe de l'appareil.

Lorsqu'on veut choisir un numériseur de moyenne gamme, on s'orientera vers un appareil conçu par un fabricant de matériel photographique, plutôt que vers un fabricant de matériel bureautique qui va favoriser la rapidité au détriment de la justesse des couleurs. Dans cet ordre d'idées, on s'orientera vers un banc d'essai mené par un magazine de photographie ou d'imprimerie, et non vers celui d'un magazine d'informatique.

E. 5.5. Définition des fonctionnalités

Pour permettre de travailler efficacement, un numériseur doit comporter :

- un réglage de dureté d'image (ou contraste ou gain ou accentuation) à l'acquisition,
- un outil permettant de ré-étalonner la courbe de couleurs en fonction de l'usure de la lampe ou un compteur d'heures de fonctionnement de la lampe,
- une sortie SCSI ou FireWire (et non USB, car c'est trop lent pour les fichiers à haute définition),
- une compatibilité Mac / Unix ou PC si nécessaire,
- des chartes d'étalonnage, une compatibilité Colorsync de Apple pour contrôler effectivement la couleur à l'écran et à l'impression, ceci pour accorder sa chaîne graphique avec celle des imprimeurs,
- un réglage fin ou manuel de la fonction de détramage des similigravures.

Complément : Fonction avancée de détramage

La plupart des constructeurs proposent une fonction détramage, mais il existe deux moyens pour parvenir à ce résultat :

- La correction logicielle qui applique un flou sur l'ensemble de l'image. Ce moyen est celui des numériseurs de supermarché, et donne des résultats décevants. N'achetez pas ce genre d'appareil.
- Le détramage par numérisation séquentielle qui a pour principe de numériser plusieurs fois une zone de l'original, en déplaçant la grille virtuelle des pixels de l'image et en en faisant une moyenne. On reconnaît aisément cette fonction à l'usage, car la numérisation dure 2 à 3 fois plus longtemps qu'en mode normal, et quelquefois le numériseur « bourdonne » pendant l'opération (l'attelage optique mobile avance et recule d'un demi-pixel pour effectuer ses deux passes d'échantillonnage).

F. 5.6. Numériseurs pour l'acquisition mono-page

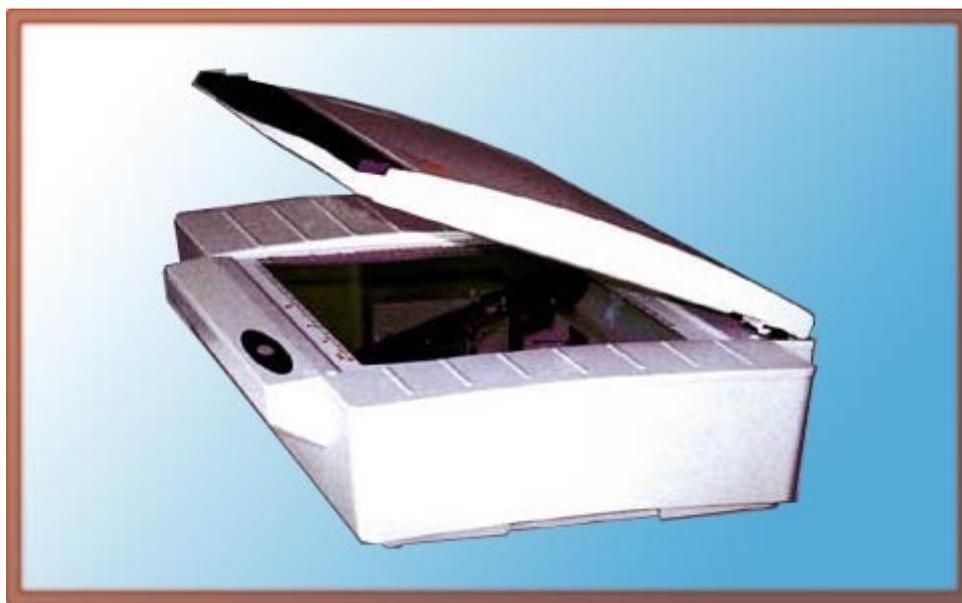
Cette famille de numériseurs sera essentiellement compétente pour ce que nous avons appelé la "photothèque" constituée de documents plats et mono-page.

Complément : Numériseurs à plat

Les numériseurs à plat pour documents opaques

Ces numériseurs sont du type le plus commun. Une lampe fluorescente éclaire l'original opaque et l'image est captée à travers un système optique par des capteurs CCD, délivrant un signal en mode analogique. L'image est ensuite numérisée par une conversion des signaux analogiques en pixels numériques. Et c'est dans cette opération que se décide la qualité du numériseur. Le logiciel de conversion, qui gère à la fois les couleurs et le suréchantillonnage (génération de plusieurs pixels à partir d'une information unique provenant du capteur) et le sous-échantillonnage (génération d'un pixel à partir de plusieurs pixels générés par le numériseur).

Les plus courants acceptent le format A4 (210x297) ou légal américain (215 x355 mm). Les numériseurs « pros » acceptent le format A3 (297x 420mm)

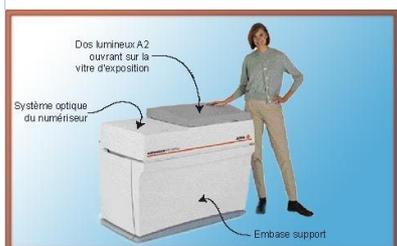


Numériseur A3 à plat avec zoom optique interne

Les numériseurs à plat pour documents opaques et transparents

La plupart des numériseurs à plat dignes de ce nom ont un dos lumineux pour numériser les plaques photo et les négatifs. Mais contrairement à ce qu'annoncent les notices publicitaires des appareils domestiques, ces appareils ne sont pas aptes à numériser directement les négatifs de petit format (24x36 mm) car leurs capteurs à 1200 dpi optiques n'ont pas une définition suffisante. Les numérisations à 9600 dpi (voire plus) qu'ils annoncent sont des extrapolations se traduisant par une image aux contours incertains. Seuls les numériseurs (très peu d'exemplaires) qui comportent un équipement de zoom à deux positions (un objectif est intercalé entre le négatif et le capteur) sont effectivement capables de numériser à la fois les négatifs et positifs de petit format (24x36 mm), les moyens format (6x6 et 6x9cm) et les grands formats (plaques 9x13 cm, 13x18 et 18 x 24cm).

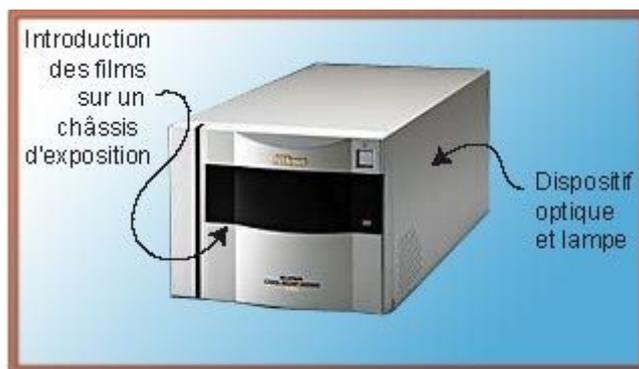
Les numériseurs d'imprimerie acceptent le format A2 ou plus, mais la taille des fichiers issus de la capacité de ces machines conduit à utiliser des espaces disques aussi démesurés que la vitre d'exposition !



Ce type de numériseur permet les acquisitions de la plus haute qualité.

Complément : Numériseurs à négatifs de petit format

Si l'on veut numériser une grande quantité de négatifs 24x36 mm il vaut mieux acquérir un numériseur dédié à cette fonction. Mais il faut garder à l'esprit que cette machine ne pourra pas traiter les négatifs et diapos au-delà de 5x5 cm, donc tous **les moyens formats** cités ci-dessus -les plus courants en iconographie industrielle et institutionnelle- **ne passent pas**.



Numériseur à diapos 5x5 et négatifs de petit format

Complément : Numériseurs à livres ouverts

Ces numériseurs appartiennent typiquement au monde des archives et des bibliothèques et sont construits de façon analogue à un banc de microfilmage de registres: le livre ou l'original de grande dimension (A2 généralement, A1 maximum) est posé sur un plateau. La caméra est placée au-dessus et une source de lumière disposée sur la colonne éclaire l'objet à numériser. L'épaisseur de l'objet peut atteindre 200 mm. Ces appareils évitent de détériorer les reliures et ont des options adaptées pour le recadrage, l'élimination des bords et le redressement de l'image. Certains ont une option d'assemblage de pages en PDF, ce sont eux qui seront notre meilleur choix.

Attention : ces appareils sont sensibles à la lumière environnante, et certains d'entre eux n'ont pas un rendu très contrasté. La définition est limitée à 300 dpi (ce qui suffit amplement) et ils ne peuvent numériser que les documents opaques.

Certains peuvent être munis de plateaux compensateurs, à l'instar des caméras de microfilmage.

Ils peuvent être utilisés en autonome (impression directe depuis le numériseur, en libre-service des bibliothèques) ou être reliés à un poste de travail sur lequel on pratiquera le post-traitement.



Un numériseur de livres à colonne, conçu comme un banc de microfilmage

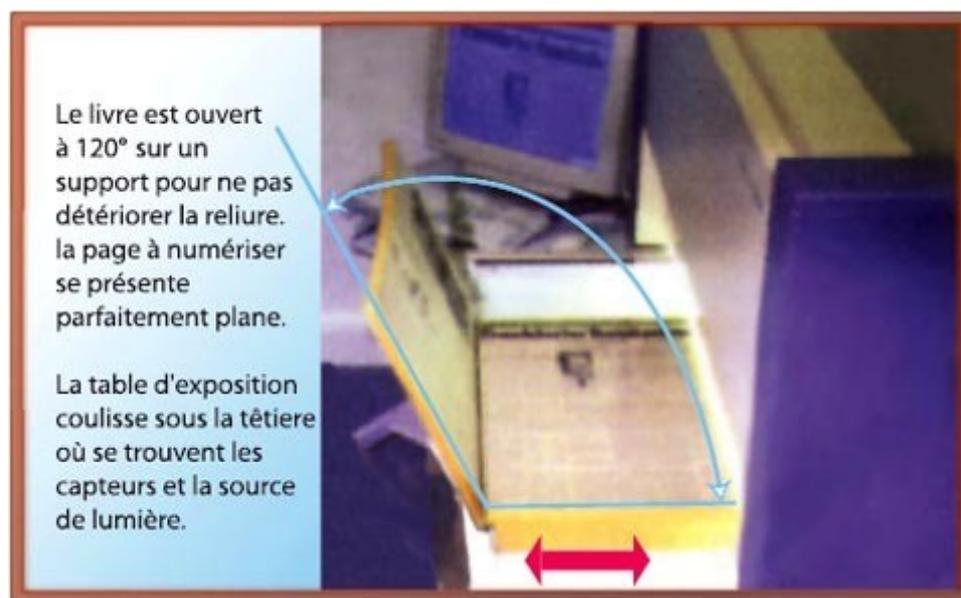
Complément : Numériseurs à plateau de grand format

Ces appareils de taille impressionnante (de prix non moins impressionnant) permettent de reproduire les plans en couleurs, les livres ouverts, les objets d'épaisseur pouvant aller jusqu'à 200 mm. Ils couvrent le format A1 ou A0 selon les modèles. L'objet à numériser est placé sur un plateau qui avance et présente l'original en défilement continu à trois capteurs linéaires pour une définition de 300 à 400 dpi. Le logiciel de reconnaissance mémorise le format et délivre l'image aux dimensions finies même si le livre est placé approximativement. Les fichiers en niveaux de gris ou en couleurs seront sujets à un post-traitement pour l'assemblage des pages dans un fichier unique.



Ce principe existe aussi pour le format A0

Sur certains de ces appareils, un support permet de présenter le livre ouvert à 120°, ce qui permet de numériser une seule page à la fois, **mais bien à plat et sans détériorer la reliure**

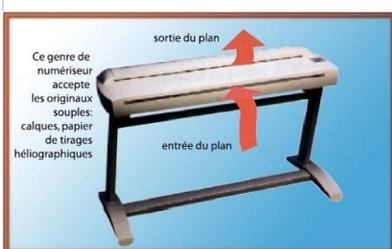


Numérisation d'un livre sans détériorer le dos de la reliure

Complément : Numériseurs à plans

Les numériseurs à plans font généralement partie d'une chaîne de reproduction des plans «au format» et sont flanqués d'une imprimante de grand format A0 + et d'une plieuse des tirages effectués. Cependant, la plupart des revendeurs acceptent d'installer le numériseur uniquement, avec une interface pour un poste de travail PC ou Mac.

Les numériseurs à plans acceptent seulement des originaux souples qui défilent devant les capteurs CCD. Il n'est pas possible d'exposer des livres, des originaux rigides ou cartonnés. Les calques et les papiers peuvent passer dans ce transporteur. Les originaux mécaniquement très fragiles sont à placer dans des chemises transparentes en « Mylar » qui sont fournies par le constructeur. La largeur admissible du plan est généralement de 914mm, mais la longueur peut souvent dépasser les 1180 mm du format A0. Si l'on a tirages de plans de largeur supérieure à l'ouverture du numériseur, on peut éventuellement les remettre dans leur pli pour les numériser en 2 fois, puis assembler les fichiers dans un logiciel de traitement d'image.



Ce type de numériseur peut délivrer des fichiers en couleurs

Complément : Format de sortie des numériseurs de grand format

Le format de sortie est généralement du Tiff à 300 ou 600 dpi, en mode bitonal pour les numériseurs à plans (architecture et industrie) et en couleurs pour les numériseurs de graphistes et/ou d'archivistes.

Généralement, il y a possibilité de faire une sortie en Tiff multi-pages sur les numériseurs à plans, et l'assemblage s'arrête lorsque l'utilisateur déclare le changement de travail.

Attention !

Le Tiff multi-pages n'est pas reconnu par les applications courantes lors d'une importation

de fichier. En effet, le filtre reconnaît la première page uniquement ou refuse la conversion. En revanche, il est facilement compris par une imprimante de photocopieur muni d'un logiciel de chargement direct.

G. 5.7. Numériseurs pour l'acquisition des documents multi-pages

Il existe des solutions adaptées à la numérisation automatisée des documents, applicables à des documents à pages multiples.

Ce sont physiquement soit :

- un chargeur automatique à adapter sur un numériseur à plat,
- des numériseurs à chargement automatique dédiés à cette fonction,
- des numériseurs de bureau avec export à travers un courriel (e-mail),
- des numériseurs faisant partie d'un photocopieur numérique, avec une option d'exportation du fichier en réseau ou vers un poste de travail associé.

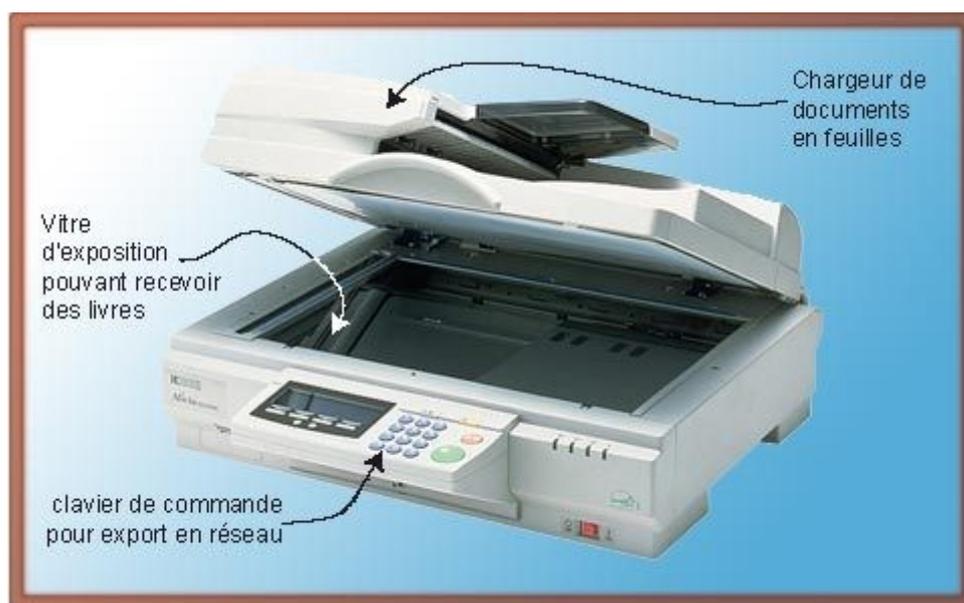
Complément : Numériseur couleur avec chargeur multi-pages

Un adaptateur-chargeur d'originaux est proposé comme accessoire de certains numériseurs à plat de moyenne gamme. Il permet de numériser des originaux en série, mais le rendement est faible, car les numériseurs de haute qualité sont assez lents, et convertissent généralement l'acquisition en haute définition vers un fichier à plus basse définition.

Ces machines ont évolué vers une génération de numériseurs conçus pour l'acquisition rapide. Ils sont équipés d'un module réseau qui permet d'expédier des fichiers en ligne vers le poste de traitement. Ils peuvent alors être partagés entre plusieurs utilisateurs. Au niveau de la conception, ce sont des modules de numérisation de copieurs couleur.

Plutôt que de réadapter un chargeur automatique sur un ancien numériseur, il vaut mieux acquérir une machine actuelle, dont la rapidité (et la précision) sont supérieures.

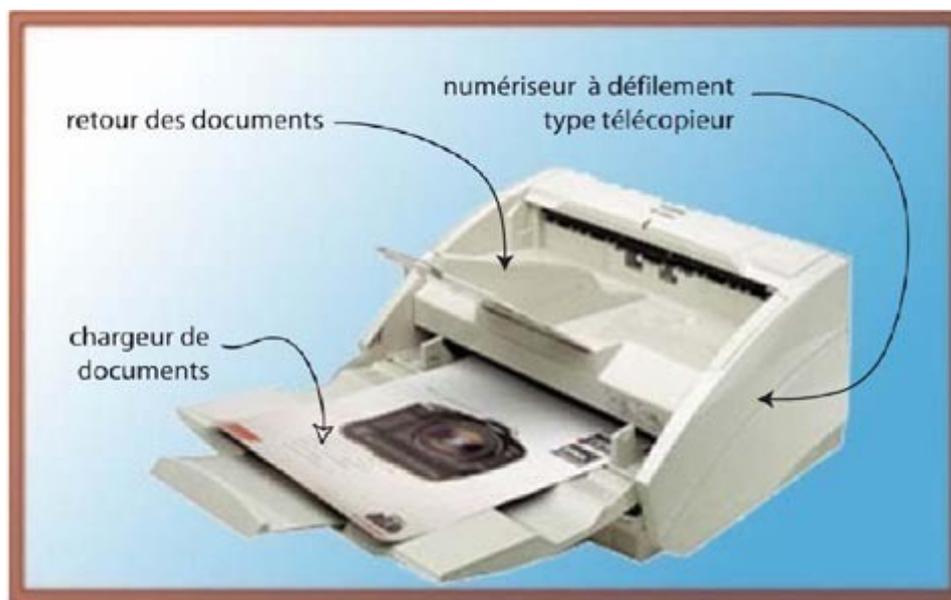
Ces numériseurs génèrent des fichiers «mono modaux», c'est à dire au mode choisi par l'utilisateur: soit du bitonal, soit du niveau de gris, soit de la couleur.



Numériseur doté d'un chargeur automatique de documents. (appareil de bureautique)

Complément : Numériseur monochrome à chargement automatique

Ce genre de numériseur se présente comme un module de télécopie, à défilement du document autour d'un tambour en verre à l'intérieur duquel se trouvent la lampe d'exposition et les capteurs. Ce dispositif est assez rapide, mais est conçu pour la numérisation des documents d'affaires, en mode bitonal. On n'attendra pas de ce genre d'appareil des acquisitions de haute qualité pour la restitution fidèle des archives. Les fichiers sont transmis à un poste de travail associé ou à une messagerie électronique.



Numériseur multi-pages pour les documents d'affaires.

Complément : Numériseur multi-pages à modes combinés

Des fabricants ont eu l'idée d'installer dans un petit numériseur multi-pages, une messagerie permettant d'utiliser le numériseur comme une machine autonome, et d'envoyer les fichiers par courriel à un destinataire préselectionné. Ces fichiers peuvent aussi être envoyés en télécopie. On comprendra aisément que cet appareil est un hybride entre le télécopieur et le numériseur de bureau. Son principal avantage est d'être accessible à tout le monde, sans monopoliser un ordinateur pour l'acquisition des images, et surtout de générer des documents en PDF en différents modes colorimétriques.

La numérisation peut se faire avec le chargeur automatique, ou en plaçant le document sur la vitre d'exposition qui est plane. L'utilisateur se voit proposer divers modes d'acquisition :

- mode «texte» générant un fichier bitonal à basse-moyenne résolution (catégorie 1- texte lisible qualité télécopie en mode fin)
- mode «photo» générant un fichier couleurs RVB à basse-moyenne résolution (catégorie 2)
- mode «document en couleurs» qui active un module d'analyse et mélange les modes bitonal et RVB dans un même document PDF. En fait, le fichier est en 2 couches : le fond en mode bitonal et les images en couleurs posées par-dessus, et cachant les «fantômes» noirs issus de la numérisation bitonale des photos. (catégorie 2).

La numérisation se fait en une passe unique et est très rapide, les photos sont habilement détramées (surtout grâce à la basse définition du numériseur). Des fichiers multi-pages sortent assemblés au format PDF, ou au format tiff multi-pages en modes photo et texte.

Les fichiers réalisés sont alors prêts à être expédiés par courriel en composant l'adresse du correspondant. Ils se présentent en pièces jointes dans le message. La réception et la

lecture sont possibles quelle que soit la plate-forme (Mac, Unix ou PC).

A l'arrivée, il ne reste qu'à contrôler le document, assembler éventuellement des séries de pages et le placer dans le stockage de masse pour une diffusion en ligne.

C'est une solution fonctionnelle pour la numérisation en quantité de documents **à consulter en ligne**, mais c'est **insuffisant en termes de résolution pour conserver une image fidèle des documents en voie de dégradation**.



Le placement manuel sur la vitre d'exposition se fait en soulevant le couvercle.

Complément : Numérisation à travers un photocopieur numérique

Il est possible de numériser des documents en haute définition (600 x 600 dpi au moins) en utilisant le plateau de numérisation d'un photocopieur noir et blanc récent. Cet appareil va générer des fichiers «mono modaux» c'est à dire soit du bitonal au trait sur toute la surface pour un mode «texte» du copieur, soit des modes niveaux de gris en mode «photo» du copieur, mais ne mélangeront jamais les genres. La sortie est généralement un fichier tiff option multi-pages ou un fichier PDF multi-pages.

Les numériseurs faisant partie d'un copieur couleur, vont générer des fichiers en couleurs. Là encore, l'image sera «mono-modale» pour les textes et les images et on pourra régler le cadrage à un format prédéfini d'impression du copieur. Le fichier final est souvent un tiff multi-pages, ou un PDF de qualité moyenne pouvant être maximale si l'on n'active pas une compression jpeg excessive.

Beaucoup de ces copieurs offrent à présent cette option numérisation vers le réseau ou vers une adresse E-mail. L'avantage principal est la rapidité de numérisation, mais l'exportation vers un poste de travail via un réseau peut être des plus acrobatiques. Il faut assister à une démonstration réelle et tester la machine sur son propre réseau avant d'acheter ce genre de système.

Il y a souvent besoin d'un post-traitement pour nommer les fichiers et /ou les assembler en PDF pour la diffusion.



Copieur numérique connecté

Parmi les numériseurs pour documents multi-pages, il y a des **numériseurs spécifiques pour les microformes. A quoi cela sert-il?**

Il est étonnant que l'on numérise des microformes, qui sont considérées comme des supports de sécurité fiables, pour les transférer sur un support dont la pérennisation est encore problématique (compatibilité des encodages, durée de vie limitée des CD-ROM).

En fait, ce transfert se fait actuellement dans l'industrie pour diffuser des documents d'usage commun à plusieurs groupes de travail qui sont éloignés dans l'espace.

Il se fait également dans les services d'archives pour diffuser plus largement des microfilms de sécurité ou de complément.

Mais en aucun cas, il ne faut détruire les microfiches après numérisation, car leur nature leur permet de traverser le temps, au moins au titre d'archives historiques dans le premier cas et de support de sécurité dans le second.

Complément : Remarques sur les numériseurs de microformes

Les microformes (microfiches, microfilms, cartes à fenêtre) peuvent être numérisées sur des appareils spécialisés. Les visionneuses actuelles sont équipées d'un écran de visualisation optique et en outre d'un module de numérisation pour l'export du document sur un réseau .

Certaines machines sont dédiées toutes entières à la numérisation pure et dure et sont à connecter à un poste de travail standard, généralement un PC.

Ces numériseurs sont délivrés avec un logiciel d'acquisition et de retouche "fonctionnelle" de l'image:

- redressement des vues mal positionnées
- élimination des bords noirs
- correction automatique et/ou manuelle des niveaux

L'acquisition peut se faire en niveaux de gris ou en bitonal; on choisira le mode le plus adéquat pour chacun des originaux. Le mode niveau de gris ne sera appliqué qu'aux documents montrant de grandes zones sombres ou des traits ténus qu'il n'est pas possible de traiter en mode bitonal.

On peut effectuer la numérisation à partir des microformes argentiques ou des copies en diazo (violet) si celles-ci ont été conservées dans de bonnes conditions.

La résolution présélectionnée dans la machine suit les recommandations ISO prescrites pour les microformes.

La définition indiquée par le logiciel d'acquisition, par exemple 300dpi pour un document A4, est celle du document restitué à sa taille d'impression, c'est pourquoi il faut bien prendre garde au taux de réduction appliqué lors du microfilmage pour ne pas sous-échantillonner ni sur-échantillonner le fichier.

Complément : Numériseur de cartes à fenêtre

Les cartes à fenêtre sont numérisées à l'aide d'un système à transport dynamique. Les cartes disposées dans l'introducteur à gauche sont transférées vers l'unité de lecture, puis ressortent dans le réceptacle à droite. Le logiciel de pilotage permet de prédéfinir la taille de restitution du document, le mode colorimétrique, la résolution à la taille d'impression. Il y a possibilité de traitements par lots. Certains logiciels lisent les cartes à fenêtre perforées et prennent ce texte comme nom du fichier. Les formats Tiff, jpeg et PDF sont proposés par la plupart des fabricants.



Les cartes entrent à gauche et sortent à droite de manière automatique. Les cartes qui posent problème sont rejetées dans le réceptacle latéral, pour traitement manuel par l'opérateur

Complément : Numériseur de microfiches

Le numériseur de microfiches fonctionne à l'aide d'une caméra dynamique qui scrute les images une par une et les met à disposition de l'opérateur dans un logiciel de tri et d'adressage. Il reste à les nommer et les retoucher éventuellement.

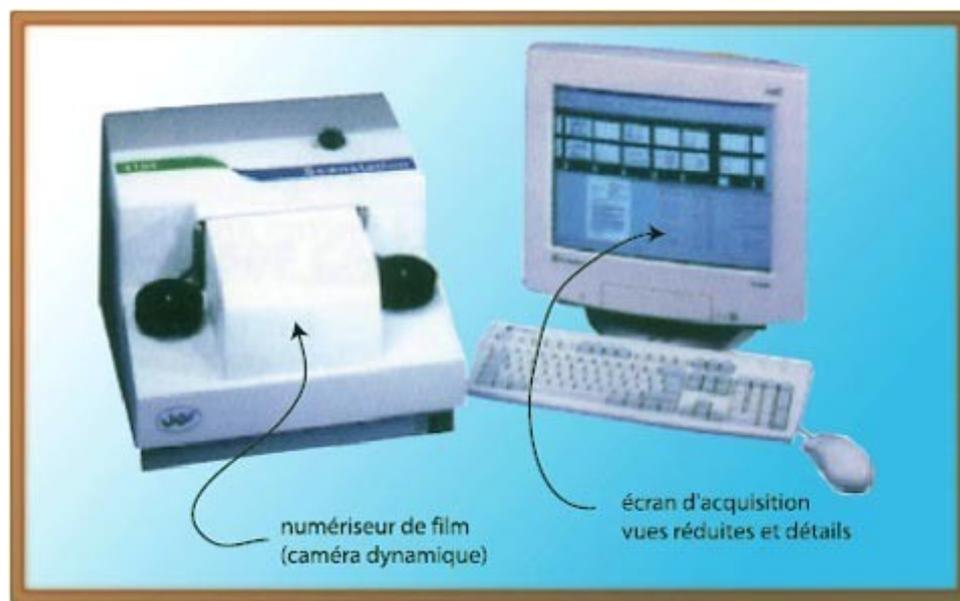
L'automatisation de ce genre d'appareil et l'homothétie des documents à traiter permet de le confier à du personnel peu qualifié.



Numériseur de microfiches

Complément : Numériseur de microfilms

Le numériseur de microfilms fonctionne par un transfert automatique des images vers un logiciel de tri et d'adressage. Il reste à les nommer et les retoucher éventuellement. Le film avance devant une camera de prises de vues qui acquiert les images alors que le film est immobile devant la fenêtre d'exposition. Les films 16mm et 35 mm sont traités par le même appareil, avec une permutation d'objectif.



Numériseur de microfilms

H. 5.8. La mise en place des éléments du poste de numérisation

Il est important de constituer un poste de numérisation cohérent, pour éviter la fatigue de l'utilisateur, les difficultés de manipulation et les gestes inutiles.

Le numériseur à plat, s'il est du type «à poser sur le bureau» doit rester le véritable centre du système; les autres éléments seront considérés comme périphériques.

Le numériseur à livre ouvert est à disposer dans un lieu où l'éclairage est modéré et ne varie pas, ou n'est pas sujet à des taches de lumière (rayons de soleil, lumières fugitives lors de l'ouverture de la porte...)

L'ordinateur sera placé au plus près du numériseur si le transfert se fait par une liaison SCSI, car la longueur de ce câble est limitée d'origine. Une solution consiste à placer l'unité centrale sur une table accolée ou sous la table du numériseur. La liaison FireWire n'impose pas cette sujétion.

L'opérateur devrait disposer d'un écran cathodique 17 pouces au minimum, un écran 21 pouces est recommandé. Les écrans plats de bureautique ne conviennent pas à la numérisation et à la retouche des images en couleurs.

Si le fonds d'archives comprend des originaux de grande taille que l'on voudrait numériser par défaut sur un numériseur à plat (des ouvrages, des gravures), il est nécessaire de constituer de part et d'autre du numériseur un plan d'appui des originaux, pour la partie qui n'est pas placée sur la vitre d'exposition, cela pour permettre une exposition dans de bonnes conditions de fixité et sans froisser les documents. C'est un plan d'appui en stratifié, lisse.

Pour assumer les tâches de numérisation, il est nécessaire de disposer :

- du numériseur et de son logiciel d'acquisition,
- d'un logiciel sérieux de retouche d'image (Adobe Photoshop domine le marché et sa version "Photoshop éléments", peu coûteuse peut suffire à un poste d'acquisition),
- d'une unité centrale conçue pour le traitement d'images et capable de lire, d'exporter et de dialoguer dans tous les mondes: Mac et Unix pour les imprimeurs, PC pour les autres utilisateurs;
les unités Mac conviennent à cet usage, alors que les unités PC ne sont pas performantes pour le traitement d'image et ne permettent pas d'enregistrer sur des supports physiques UNIX ou Mac (disquette CD-ROM, disque dur extractible, disque magnétique...); le Mac lit et enregistre sur tous les supports;
- si on traite des images en couleur depuis un numériseur A3, une mémoire vive de 1GigaOctet procure un bon confort d'utilisation et permet de faire fonctionner successivement plusieurs filtres de post-traitement.

Il est recommandé d'utiliser un autre poste pour faire office de serveur Web des images numérisées. Dans ce cas un PC convient, car les arcanes de la gestion de serveur PC sont partagées par de nombreux webmestres.

Les autres actions comme la mise en banques de données, la compilation de fichiers PDF, peuvent être faites sur un poste aux performances moyennes, de n'importe quel environnement.

I. 5.9. Ségrégation des documents

Ségrégation des documents mono-page

Afin de travailler d'une manière efficace et sériee, il est conseillé de rassembler autant que possible les originaux en suivant une typologie colorimétrique.

Exemple

On se proposera de numériser en bloc toutes les gravures d'un fond artistique, puis toutes les sanguines, puis toutes les aquarelles, etc...

Afin de travailler d'une manière efficace et sériee, il est conseillé de rassembler autant que possible les originaux en suivant une typologie colorimétrique. Par exemple, on se proposera de numériser en bloc toutes les gravures d'un fond artistique, puis toutes les sanguines, puis toutes les aquarelles, etc...

La sérialisation permet de rester dans les réglages d'acquisition du numériseur et de gagner du temps de manipulation. Certains numériseurs proposent l'acquisition par lots, sans quitter l'application-pilote du numériseur.

Ségrégation des pages d'un document relié

Avant de numériser un document relié rassemblant plusieurs types de pages, il est nécessaire de faire une reconnaissance du contenu pour adapter le mode colorimétrique à chaque page. On choisira le mode bitonal pour les pages de textes et gravures et le mode niveaux de gris ou couleur pour les pages contenant des illustrations en tons continus.

Si l'on acquiert ces pages en série, il est bon de numériser en premier toutes les pages d'un même type colorimétrique que l'on rassemble ensuite dans un fichier multi-pages (au format PDF par exemple).

Ensuite, on numérisera les pages «atypiques» par familles de modes, voire de résolutions si l'on veut affiner l'une ou l'autre des pages montrant de fins détails, pour les intégrer ensuite au fichier principal.

Dans le cas où les pages sont gérées individuellement et distribuées automatiquement par une base de données, on procèdera aussi par groupes de typologie, ne serait-ce que pour assurer la régularité du traitement colorimétrique (on conserve à l'écran une page de initiale et on ajuste les densités et tonalités de toutes les autres sur cette référence).

J. 5.10. Post-traitement des images

Travail des images

La numérisation n'est pas forcément finie dès la sortie du numériseur. Il peut être nécessaire de pratiquer un post-traitement de l'image, pour atténuer les défauts de numérisation:

- redresser l'image,
- enlever les bords noirs,
- retoucher la colorimétrie pour améliorer le rendu et rester fidèle au document original que l'on a numérisé,
- appliquer un masque de détournage,
- etc.

Le post-traitement inclut aussi la génération des fichiers à basse résolution pour la diffusion en ligne.

Le post-traitement doit être confié à un personnel plus qualifié que celui pratiquant l'acquisition brute. Pour le post-traitement il est recommandé d'acquérir un logiciel exécutant des macrocommandes préenregistrées qui regroupent une série d'actions à déclencher automatiquement à partir d'un clic de souris (comme par exemple Photoshop).

On peut même automatiser le traitement par lots (par exemple le sous-échantillonnage des fichiers de diffusion en ligne), ce qui permet de déclencher ces travaux la nuit, sans occuper inutilement le poste et laisser l'opérateur dans des tâches ineptes et répétitives.

Identification des images

Il est préférable de donner des titres clairs aux fichiers, incluant la cote d'archivage plutôt que des numéros d'ordre sans signification. La longueur des titres dans les systèmes actuels comprend au moins 35 caractères.

Ne pas laisser d'espaces pour la mise en ligne sur un serveur internet ou intranet.

Ne pas utiliser de caractères accentués pour traverser tous les systèmes d'exploitation de tous les pays sans changer le nom du fichier.

Aperçus (*thumbnails*)

Il faut toujours activer la fonction «générer l'aperçu» des fichiers image car une petite icône apparaîtra immédiatement dans les fenêtres des gestionnaires d'images, ou bases de données spécialisées sans perte de temps pour la générer lors de la consultation (pour le Mac, le système affiche toujours l'aperçu dans les fenêtres de gestion des fichiers).

K. 5.11. Informations textuelles dans le fichier image

Il est possible d'inclure à l'image, dans les formats TIFF et JPEG, des textes de description de l'image, sa cote en archives, son auteur, sa date d'édition, le copyright, etc... à partir des logiciels de retouche professionnels. Ainsi cette image, même détachée du contexte de la notice de son serveur peut encore «raconter son histoire» si on ouvre la fenêtre « informations » dans le logiciel de retouche ou une visionneuse équipé de la fonction "fetch".

Ces informations sont enregistrées dans des zones qui sont exploitables aussi par des logiciels de base de données spécialisés que l'on trouve dans le commerce dits logiciels de

«fetch». Ils sont utilisés par les journaux et agences de presse au niveau mondial, pour le choix et l'échange des photos d'actualité en ligne.

Il est aussi possible d'insérer un copyright dans l'image. A ce moment, elle ne pourra être ouverte que si les droits de diffusion sont acquittés ou déclarés. Ce copyright nécessite l'abonnement à un gestionnaire délégué des droits d'auteur et ne s'applique guère au monde des archives.

L. 5.12. Trucs et astuces

Il arrive que l'on ait à utiliser un matériel de numérisation en dehors de sa destination première, surtout pour les cas où il n'est pas économiquement envisageable d'acquérir des machines coûteuses pour des opérations ponctuelles.

1. 5.12.1. Documents plus grands que la vitre d'exposition

Parfois, quelques documents à numériser vont être plus grands que la vitre du numériseur à plat. Il est très simple de pallier cet inconvénient pour peu que l'on puisse couvrir le document en deux ou trois prises de vues.

Il suffit de placer le document dans un sous-verre démontable maintenu par des clips métalliques. Ce sous-verre va matérialiser une référence de position et maintenir le document parfaitement à plat. L'épaisseur de la vitre ne gêne pas la plupart des numériseurs dont la profondeur de champ est au moins de 3 à 4 mm. Les numériseurs professionnels sont à mise au point automatique ("auto focus").

Le dos du sous-verre doit comporter des repères délimitant la surface utile du numériseur, moins les marges de recouvrement, pour permettre ses positionnements successifs.

Il suffit de numériser l'image en déplaçant le sous-verre sur un guide fixé sur un côté du numériseur: ce guide assure le parallélisme des prises de vues successives, ainsi il sera toujours possible de les raccorder exactement dans le logiciel de retouches d'images. Au cours de ces numérisations, il faudra bloquer les fonctions automatiques du numériseur, pour obtenir toutes les images au même contraste et à la même densité.

Lorsque toutes les prises de vues sont exécutées, il faut créer dans le logiciel de retouches d'images un document vide à la taille du document fini, puis y coller une par une les images d'acquisition. Chacune des images importées se trouvera sur un calque distinct, il suffira de déplacer ces calques pour assurer la coïncidence des parties rapportées.

Lorsque les calques sont calés, et les zones de transition plus ou moins retouchées, aplatir l'image (fusionner les calques) et l'enregistrer au format désiré.

2. 5.12.2. Numérisation d'un ouvrage «à la baguette»

Lorsque l'on ne dispose pas d'un numériseur à «livre ouvert» et que quelques ouvrages sont à numériser, il faudra se résoudre à faire cette acquisition sur un numériseur à plat.

Il est difficile de numériser les livres car la reliure ouverte conduit à un gauchissement des pages qui apparaissent inclinées sur le document final. Par ailleurs, le placement sur la vitre est hasardeux et le décalage d'une page à l'autre peut être important.

Pour réduire ce phénomène, il suffit de fixer un objet cylindrique, peint en blanc sur la vitre du numériseur, à l'endroit où l'on veut placer la reliure (un gougeon de menuisier, un tube de cuivre peint...).

Ce dispositif permet:

- de placer le livre dans un position répétitive d'une page à l'autre,
- de bien orienter les pages et de les maintenir parallèles au cadre de numérisation,
- d'épargner la reliure en évitant son sur-écartement.

La charnière du livre ouvert se centre sur le cylindre et force plus ou moins les pages à s'ouvrir de manière régulière et normale à la reliure, sans la casser. Il faut adapter le diamètre du guide à la profondeur du pli de l'ouvrage (on change le guide au cours du traitement). On obtient de très bons placements surtout si la vitre du numériseur est traitée aux silicones pour être bien glissante (bombe silicones auprès du technicien photocopieur). L'ombre entre le papier et le guide est ensuite éliminée par recadrage, ou tout simplement on règle le cadrage d'acquisition pour cacher l'accessoire.

Cette technique n'est toutefois pas applicable systématiquement aux registres, car les écritures s'étendent souvent jusqu'à la charnière de la reliure, et la baguette cacherait une partie de l'information.

On peut aussi faire cette acquisition sur une baguette de section triangulaire peinte en blanc, ce qui diminue l'ombre, mais guide moins bien la page.

6. Numérisation des manuscrits

Les manuscrits, registres d'état civil, lettres et actes sont numérisés pour permettre une consultation en ligne.

Les principales vocations de la numérisation sont :

- donner accès à un grand nombre de chercheurs à distance (les généalogistes, les historiens...),
- préserver les registres de trop nombreuses manipulations qui dégradent les reliures,
- préserver les papiers pelures ou papiers anciens fragilisés par le temps.

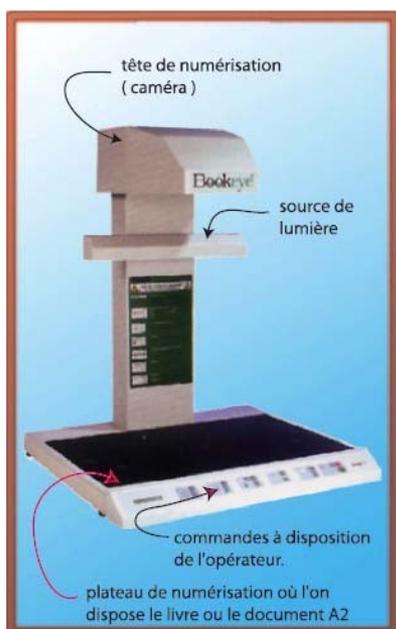
Ces vocations sont complémentaires.

A. 6.1. La nature des manuscrits

Entrent dans la catégorie des manuscrits vis à vis de la numérisation:

- les registres d'état civil, d'enregistrement, les chronos, les cahiers de bord, les comptes-rendus de séances, etc qui sont effectivement manuscrits.
- les actes notariés, les contrats, les testaments même dactylographiés...
- les lettres, les carbonés ou originaux de correspondances dactylographiées...
- et enfin les formulaires répétitifs complétés à la main: dans ce cas, on s'attachera au rendu des écritures manuscrites, la partie imprimée sera de fait rendue avec une qualité satisfaisante.

B. 6.2. Choix d'un procédé physique pour les reliés



Les registres étant fragiles, il sera nécessaire de procéder à l'aide d'un numériseur à caméra, qui permet l'acquisition en livre ouvert. L'opérateur n'a qu'à tourner les pages pour réaliser les prises de vues successives. La productivité est optimale et l'original est préservé du retournement nécessaire avec un numériseur à plat.

Dans ce cas, les pages seront présentées en face à face, paire à gauche, impaire à droite. Dans le cas où l'on voudrait décomposer le registre en pages individuelles, il serait nécessaire de faire un post-traitement informatique pour les séparer.

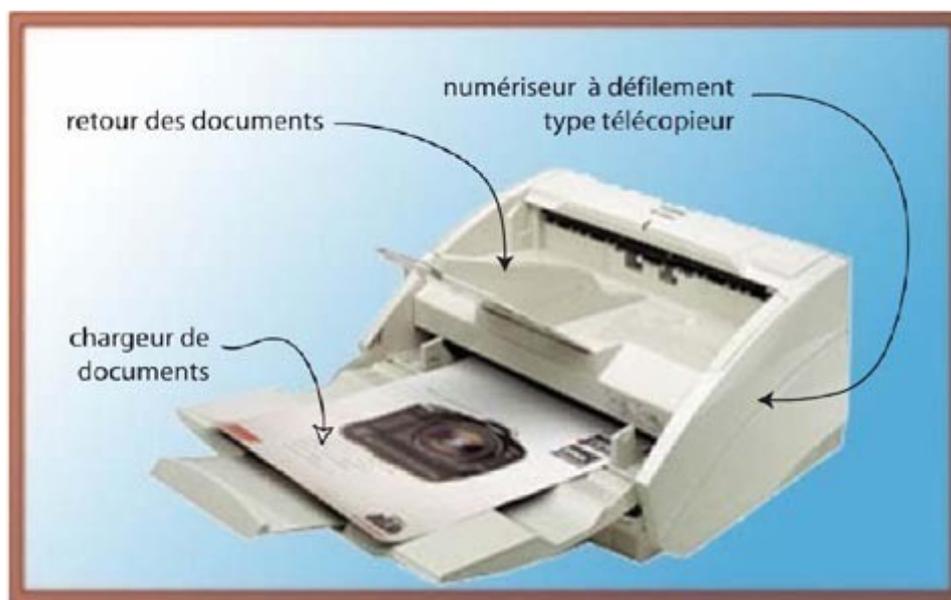
Le numériseur est généralement fourni avec un logiciel qui redresse l'image et efface les bords par des commandes simples voire automatisées (voir au chapitre numériseurs pour les détails techniques).

C. 6.3. Choix d'un procédé physique pour les feuilles volantes

Le numériseur à plat est recommandé pour les feuilles "volantes" : il permet de bien orienter l'original sur la vitre d'exposition.

Pour les pages écrites recto verso, il est nécessaire de placer derrière la page à numériser un fond en carton noir, si le couvercle du numériseur n'est pas déjà noir (la plupart des numériseurs professionnels ont un dos d'exposition noir, hormis les dos lumineux). Le noir "absorbe" l'écriture au verso tout en obscurcissant légèrement le fond de l'image à numériser. Il suffira d'augmenter légèrement le contraste de l'image numérisée pour effacer ce fond et bien ressortir l'écriture du recto.

Si l'on veut automatiser le processus pour un grand nombre d'originaux, on peut utiliser un chargeur automatique de documents, mais cette opération n'est possible que si les documents sont d'un grammage suffisant et exempts d'agrafes, de déchirures, plis etc.



Exemple de numériseur multi-pages à utiliser seulement si les originaux sont en bon état.

D. 6.4. Choix du mode colorimétrique

Le document manuscrit peut être acquis en trois modes différents :

- Le mode bitonal
- Le mode niveaux de gris
- Le mode couleur



L'exemple ci-contre contient du véritable manuscrit et de l'imprimé.

1. 6.4.1. Le mode bitonal

C'est le mode le plus adapté aux registres, tant que les écritures sont bien nettes et contrastées. Il convient aux encres noires et sépia foncé, mais peut se révéler dégradant si des parties du texte sont pâles (les écritures pâles vont disparaître si elles sont en-dessous du seuil de conversion).

Si l'acquisition est faite par défaut sur un numériseur à plat, des ombres dans la zone de reliure peuvent être traduites par de grandes taches noires. Ce problème ne survient pas avec les numériseurs à colonne (la lumière vient d'en haut et élimine les ombres).

Sur la figure ci-dessous, la tache d'humidité au-dessus du « i » est partiellement traduite en points noirs. Le mode bitonal sera à éviter pour des registres très altérés. Il convient de bien régler le seuil de conversion (niveau de traduction à partir d'une certaine densité en noir) pour assurer un bon rendu au document : pas de lettres mangées, pas d'apparition des écritures du verso ni grain du papier.



Le mode bitonal

En mode bitonal, on choisira une définition suffisante pour permettre la lecture du manuscrit:

- la numérisation en 200 dpi suffit à une écriture assez grosse (équivalente à 10-12 points),
- si les écritures sont fines, il sera nécessaire de numériser en 300 ou 400 dpi pour conserver tous les détails.

Dans tous les cas, **ces fichiers seront enregistrés au format TIFF** (Tagged Image Format File); c'est dans ce mode et avec ce format que l'on obtiendra les fichiers les plus petits pour la mise en ligne.

2. 6.4.2. Le mode niveaux de gris

Le mode niveaux de gris permet de capturer les nuances du document, les densités de l'encre, mais traduit aussi les taches et ombres en niveaux de gris plus ou moins élevés, ce qui peut gêner la lecture.

- La numérisation en 150 dpi suffit pour une écriture assez grosse
- Pour des écritures plus fines, une numérisation en 200 ou 250 dpi est suffisante
- Il n'est pas nécessaire d'aller au-delà de 300dpi, car à ce niveau de résolution, même le grain du papier est rendu dans le fichier



Le mode niveaux de gris

Le fichier sera enregistré en jpg si l'on veut obtenir une compression élevée des données. Il est à noter que même si le Jpeg crée du "sable" autour des écritures, elles restent lisibles.

En pratique, il n'est pas intéressant de numériser en niveaux de gris, car les algorithmes de compression sont peu performants sur ce genre d'images. Aussi surprenant que cela paraisse, la compression d'un document en couleurs au format jpeg donne des fichiers de taille identique, voire plus petits que leur équivalent en noir et blanc.

La lecture en niveau de gris est moins aisée qu'en mode bitonal, mais les dégradations du JPEG ne compromettent pas notablement la lecture.

3. 6.4.3. Le mode couleurs RVB

Le mode couleurs permet aussi de capturer les nuances du document, les densités de l'encre, mais traduit les taches et piqûres du papier dans leur couleur d'origine, ce qui facilite la lecture.

Le plus grand avantage de ce mode est de pouvoir obtenir des **fichiers plus petits que ceux en niveaux de gris**, avec une compression Jpeg dans les qualités moyenne et supérieure. En qualité basse, les tailles sont presque identiques.

- La numérisation en 150 dpi suffit pour une écriture assez grosse
- Pour des écritures plus fines, une numérisation en 200 ou 250 dpi est suffisante
- Il n'est pas nécessaire d'aller au-delà de 300 dpi, car à ce niveau de résolution, même le grain du papier est rendu dans le fichier



Le mode couleurs RVB

Il est à noter que même lorsque le Jpeg crée du "sable" autour des écritures, elles restent nettes, et le fichier à taille égale est plus lisible que celui en noir et blanc.

E. 6.5. La numérisation de microformes

Les microformes sont photographiées en noir et blanc. Pour la sauvegarde des registres, on utilise des films à bas contraste, pour pouvoir enregistrer les nuances de l'écriture, ainsi que les parties insuffisamment encrées. Il en résulte une image sombre et peu contrastée.

Lors de la numérisation, il est souvent nécessaire de pratiquer une retouche des niveaux et contrastes, pour rendre le document plus lisible. Certains numériseurs proposent des filtres qui corrigent automatiquement la densité de l'image pour aller du blanc pur au noir profond (au lieu de se limiter entre gris 20% et gris 60% comme beaucoup de microfiches).

Cependant la solution la plus rationnelle consiste à convertir ou à créer le fichier en mode bitonal. A cet effet, on utilisera avec avantage les commandes de réglage du seuil de conversion à partir de l'histogramme des densités que l'appareil présente dans l'un de ses modes de fonctionnement.

Les numériseurs à microfiches proposent des fonctions expertes pour rendre les fichiers en mode bitonal à "diffusion d'erreur", c'est à dire de nuages de points noirs sur fond blanc, à utiliser dès que les fonds sont colorés inégalement, le contraste insuffisant ou les écritures ténues. Le fichier produit est un peu plus volumineux qu'un fichier bitonal pur et dur, mais le rendu des demi-teintes et traits pâles est remarquable.

Si l'appareil n'est pas capable de faire cette bitonalisation à diffusion de points, il sera nécessaire de pratiquer un post-traitement dans un logiciel de retouche photo, en deux phases:

- accroître le contraste de l'image pour détacher les écritures,
- convertir en mode bitonal en diffusion d'erreur ou bitonal pur.

Prêter attention à **ne pas générer un fichier en niveaux de gris à 2 couleurs, mais un vrai fichier bitonal**, c'est-à-dire ne pas oublier de changer de mode (de niveaux de gris à

bitonal) après avoir effectué les retouches.

F. 6.6. Importance de la taille des fichiers

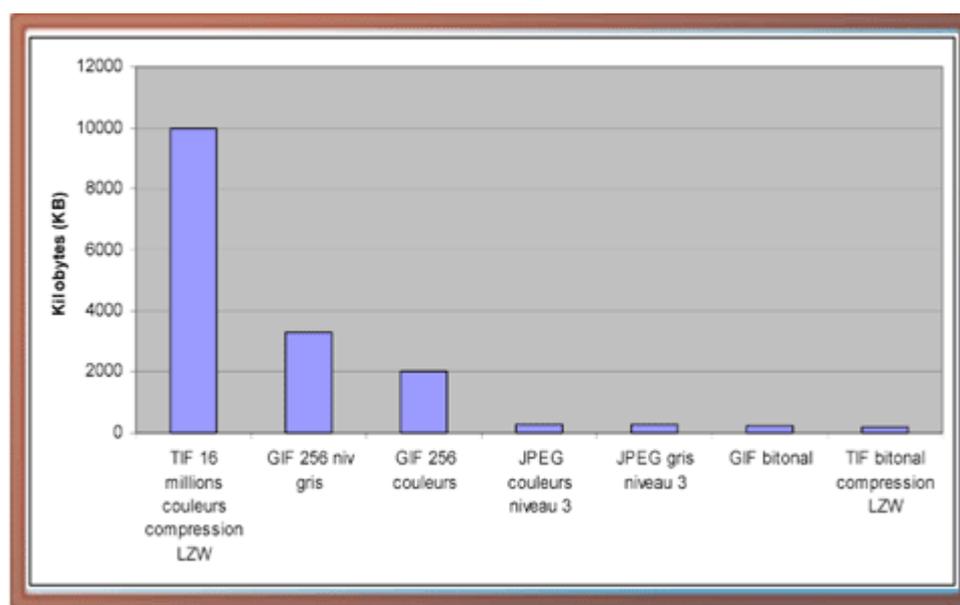
Si l'on fait l'inventaire des formats courants que l'on peut utiliser pour enregistrer une page manuscrite, on aboutit à un histogramme tel que celui ci-dessous.

La taille des fichiers est très importante lorsque l'on engage des numérisations en grande série. Les histogrammes ci-dessous démontrent de manière percutante que :

- **Notre meilleur choix sera: le TIF bitonal en premier lieu** (dans cette analyse la résolution en mode bitonal a été choisie au double de celle en niveau de gris ou en couleurs).
- Le Jpeg en couleurs sera recommandé pour les autres cas. Le jpeg en niveau de gris sera réservé à la numérisation des microfiches si on ne peut vraiment pas les transformer en mode bitonal.

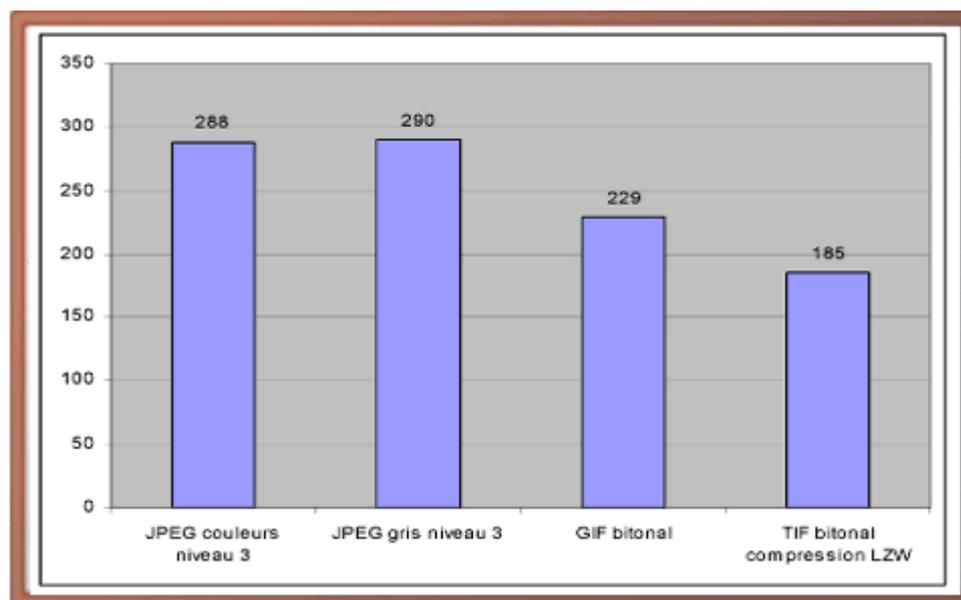
Analyse comparée des tailles de fichiers :

Un même fichier d'un document manuscrit a été converti dans tous les formats recommandés dans ce cours. La taille de départ 10 000 Kbytes, est celle du fichier Tiff en couleurs, que l'on enregistrerait pour une préservation intégrale des caractéristiques du document et sans pertes de données. Cette taille donne aussi une idée de volume d'une **photothèque** de préservation (que l'on enregistre en Tiff) par rapport à une **bibliothèque** de documents à consulter **en ligne** (tous les autres formats).



Analyse comparée des tailles de fichiers

A la lecture de l'histogramme ci-dessus, on découvre une disproportion étonnante entre les modes d'enregistrement. Pour la mise en ligne, on va donc s'intéresser aux quatre derniers.

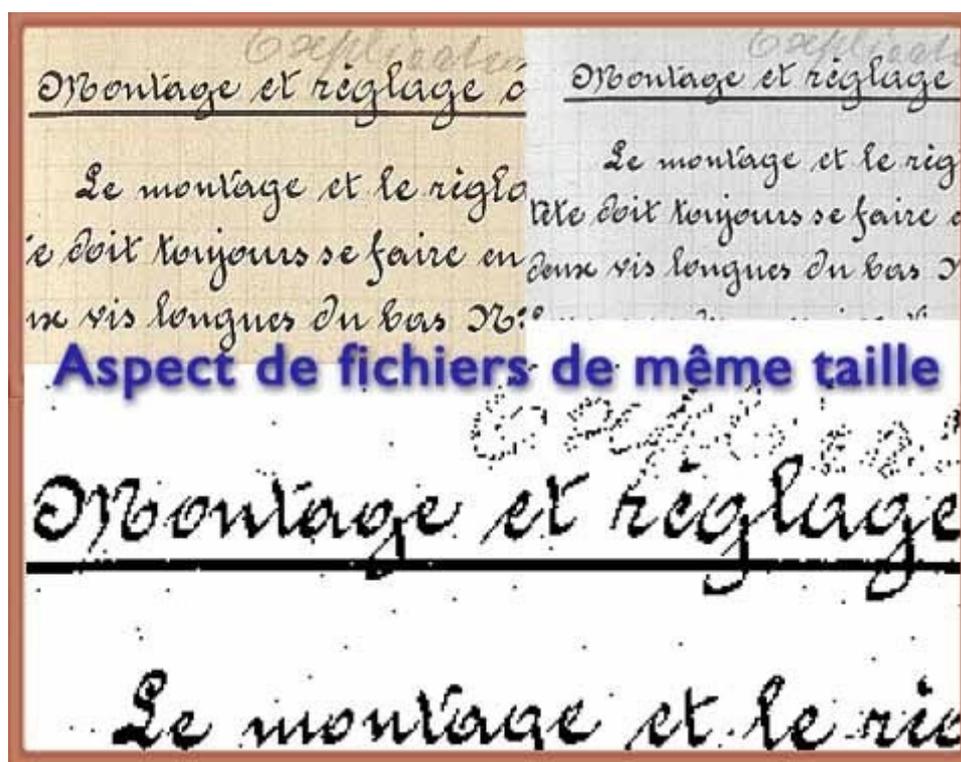


Les modes d'enregistrement

Ce deuxième graphe montre qu'à un niveau de compression élevé (3 ou basse qualité), le jpeg a la même taille en couleur et en niveaux de gris.

L'enregistrement des fichiers au mode bitonal en TIFF est plus performant qu'en GIF.

Pour donner une illustration comparée de ces modes concurrents, proposons nous de définir les aspects des fichiers à *Isotaille d'enregistrement*, c'est à dire la taille réelle en pixels de fichiers ayant la même *taille d'enregistrement après compression*.



Aspect de fichiers de même taille

La simulation ci-dessus tient compte des caractéristiques suivantes:

- Les fichiers couleur et niveaux de gris sont compressés en jpeg aux niveaux de qualité moyenne ou basse (entre 3 et 5 sur 12 niveaux de qualité)
- Dans ces conditions le fichier en niveaux de gris est juste un peu plus grand que

celui en couleurs, donc pour une même taille on aura moins de pixels (ce que montre la figure)

- Le fichier bitonal offrira 4 fois plus de pixels pour cette taille, mais se montrera moins performant sur les textes pâles (indication au crayon en haut de l'image)

Pour toutes ces raisons, on peut **choisir des modes différents de colorimétrie pour traiter une même typologie d'archives** : c'est le contraste des écritures qui nous guidera vers le mode adéquat.

7. Numérisation des photographies

Les centres d'archives conservent de plus en plus de photographies. Ces photographies sont de plus en plus demandées par les chercheurs pour illustrer les bulletins historiques et les traités. La photo est un témoin important du XXe siècle et les publications électroniques donnent des moyens rapides de les divulguer avec simplicité.

Les photographies sur papier, les plaques, les diapositives sont numérisées dans plusieurs buts:

- donner accès à un grand nombre de chercheurs à distance en mettant en ligne des copies à basse résolution,
- préparer les documents à être intégrés dans des publications imprimées,
- préserver les archives photographiques de trop nombreuses manipulations qui les dégradent,
- constituer une sauvegarde (au moins transitoire) des documents en photo couleur dont la dégradation est inéluctable.

Ces vocations sont elles aussi complémentaires.

A. 7.1. Première approche en numérisation de photographies

Il est recommandé de procéder à la numérisation avec une résolution et dans un format d'enregistrement qui préservent les qualités originales de la photographie.

Cette numérisation en très haute qualité restera en principe dans le serveur du centre d'archives, ou sur des supports de sauvegarde physiques comme le CD-ROM, le DVD...

Ce sont des copies en basse définition (catégorie 2) ou des vignettes qui sont placées en ligne, sur le site internet. Les documents en haute définition sont à distribuer à titre onéreux, et le diffuseur doit être identifié pour respecter la législation sur les droits d'auteurs (et reverser les droits à l'ayant-droit, s'il existe encore).

Ces copies en basse définition sont obtenues par un sous-échantillonnage, c'est à dire une réduction du nombre des pixels constituant le fichier original par le truchement d'un logiciel de retouche photo. La résolution de l'image sera donc réduite pour ces copies du fichier maître. La réduction de la résolution est irréversible.

B. 7.2. Plaques de verre

Les négatifs sur plaque de verre peuvent être numérisés directement avec un numériseur équipé d'un dos lumineux pour les transparents.

La résolution à choisir est de 300 dpi à la taille naturelle de la plaque. Si l'on veut agrandir le cliché à la reproduction (par exemple pour des plaques de taille égale ou inférieure à 9x13 cm), on numérise avec une résolution supérieure, tel que décrit ci-dessous pour les diapositives et négatifs.

Manipulations

Il faut dépeussier délicatement la plaque avant numérisation, mais se garder de toute autre retouche qui pourrait lui être fatale dans des mains inexpertes.

Attention

Le meilleur moyen de casser une plaque de verre, c'est de la poser sur la vitre d'un numériseur!

En effet:

- lorsque la plaque est posée directement sur la vitre du numériseur, l'air entre la plaque et la vitre se raréfie et il peut en résulter un effet ventouse dû à la pression atmosphérique;
- lorsque l'on veut relever la plaque, on risque de la casser, car elle est plaquée sur le numériseur par l'action de la pression atmosphérique.

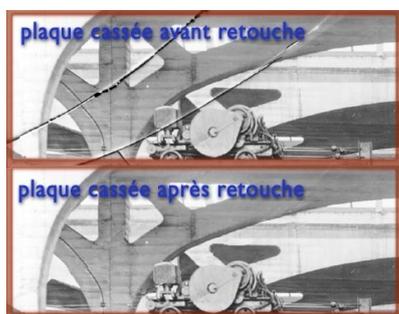
Il peut se former en outre des spires de Newton dans la zone de contact des verres (taches irisées).

Pour pallier ces inconvénients:

- disposer sur deux bords de la plaque des bandelettes de cartoline pour ménager une lame d'air entre le verre du numériseur et la plaque,
- reposer délicatement le couvercle lumineux au-dessus de la plaque et ne pas appuyer,
- utiliser la partie de la bandelette qui dépasse pour soulever la plaque.

Retouches des défauts techniques

Si l'on décide de restaurer le cliché, les rayures, taches, cassures du verre et autres défauts sont éliminés dans le logiciel de retouche. La figure ci-dessous montre un exemple de retouche numérique.



Retouche d'une plaque cassée et altérée à la cassure (fonderie SACM à Mulhouse, future faculté de droit et sciences humaines)

Cela a l'air simple, cependant, ce n'est pas à la portée des débutants.

C. 7.3. Diapositives et négatifs de petit format

Les négatifs et positifs (diapositives) de moyen et petit format (6x9 cm, 6x6 cm, 24x36mm...) sont à acquérir avec un numériseur apte à ces tâches, c'est à dire ayant une

résolution optique supérieure ou égale à 2400 dpi et une profondeur d'analyse de 42 bits au moins dans le cas des diapositives en couleur.

Les négatifs et positifs photographiques sur film nécessitent une numérisation à une résolution supérieure à celle appliquée à la taille d'impression. Pour ces documents, il conviendra de calculer le rapport d'agrandissement nécessaire et le multiplier par la résolution recherchée à la taille d'impression pour en déduire la résolution nécessaire lors de la numérisation.

Exemple

Je veux reproduire un paysage industriel dans une revue de patrimoine:

- l'image sur le négatif mesure 24 x 36 mm,
- la photo occupera une largeur de 180 mm dans l'ouvrage.

Il faut déterminer le rapport d'agrandissement : il sera de $180\text{mm} / 36\text{ mm} = 5$ fois.

En pratique, le numériseur proposera des résolutions prédéterminées qui correspondent à un fonctionnement optimal de son système de conversion, on choisira la résolution immédiatement supérieure, en l'occurrence dans l'exemple ci-dessus 1600 ou 1800 dpi.

Certains numériseurs professionnels de haut de gamme permettent d'entrer directement la taille de restitution de l'image dans une boîte de dialogue et effectuent le ré-échantillonnage à l'aide d'une courbe dite «bi cubique».

Complément : Diapositives en couleur

Les diapositives en couleurs sont à acquérir avec un numériseur ayant une profondeur d'analyse de 42bits, soit 14 bits par couche au lieu de 8 bits pour les traditionnels fichiers RVB. En outre ce numériseur doit avoir la résolution optique suffisante.

Cette analyse nécessite l'acquisition du fichier maître dans un mode colorimétrique qui ne réduit pas l'échantillonnage des couleurs ni la gamme des densités, c'est à dire en lab, (Luminosité, couleurs a et b). Pour conserver l'intégralité de la gamme chromatique, il faudra enregistrer aux formats Tiff ou Encapsulated Postscript (.eps) avec compression LZW ou ZIP ,sans dégradation de l'image.

Complément : Négatifs en noir et blanc

Pour la numérisation des négatifs en noir et blanc, le mode colorimétrique niveau de gris sera préféré au mode couleur, et le post-traitement consistera à :

- inverser l'image pour obtenir un document positif si le numériseur n'a pas cette fonction
- régler les niveaux ou courbes gamma pour obtenir une gamme complète du blanc pur au noir profond. Cette retouche est en quelque sorte la traduction informatique du métier de tireur photo, qui choisissait la gradation du papier en fonction de la vigueur du cliché, et appliquait des recettes "maison" pour mettre en valeur certains détails de la scène. On peut appliquer bêtement un filtre automatique de rendu, mais pour obtenir un bon résultat, ce travail nécessite plus ou moins des prédispositions artistiques ou de la pratique en photographie argentique.

L'enregistrement du fichier maître se fera au format Tiff compression LZW ou ZIP, sans dégradation de l'image

Complément : Négatifs en couleur

La numérisation du négatif en couleur se fait à travers le filtre du numériseur, en choisissant

comme entrée, le négatif couleur. Ce filtre spécial élimine le fond orangé et inverse automatiquement les couleurs complémentaires du négatif pour obtenir directement un tirage positif de l'image à l'écran.

Le post-traitement consistera à régler les niveaux ou les courbes gamma pour obtenir une photographie dont les couleurs sont vraisemblables. Là encore, il est nécessaire d'avoir un certain professionnalisme pour pratiquer ce traitement. Lors du traitement d'une série de clichés, il est recommandé :

- soit d'enregistrer la courbe de correction appliquée à la première image, et de l'appliquer à toutes les images prises dans les mêmes conditions,
- soit de garder à l'écran la première image retouchée, et appliquer les autres corrections par comparaison visuelle

L'enregistrement du fichier maître se fera au format Tiff compression LZW ou ZIP, sans dégradation de l'image

D. 7.4. Photographies sur papier

Les photographies sur papier sont acquises avec un numériseur à plat.

On peut procéder à des numérisations individuelles (une par une) ou placer un lot de photos sur la vitre du numériseur et faire une seule acquisition. Il faudra ensuite copier-coller les zones de l'image globale pour créer des fichiers individuels.

Tout comme pour les documents, on regroupera les photographies par lots de typologie et d'aspect similaire.

On ne mélangera pas des clichés ayant de grandes différences de contraste ou de densité en une même prise de vue: la cellule de mesure du numériseur devrait faire un moyen terme entre des clichés très différents et aucun ne serait traité correctement.

1. 7.4.1. Photographies monochromes

Les photos monochromes sont issues d'un négatif noir et blanc.

Il existe deux catégories de tirages d'après ce négatif :

- les tirages en noir et blanc qui se traduisent en niveaux de gris,
- les tirages noir et blanc qui ont ensuite été virés en couleur: une seule couleur dont l'intensité varie et parfois un mélange de noir dans les ombres et de couleur dans les tons moyens (dont nous avons décrit l'équivalent numérique: la bichromie).

Les photos en noir et blanc contemporaines (après 1945) sont généralement à acquérir en niveaux de gris, le virage étant passé de mode.

Certains tirages noir et blanc plus anciens présentent des tons chauds (papiers au bromure) qui font partie du message à transmettre. Les noirs ne sont pas noir pur mais brun très foncé, les gris sont légèrement beiges. Certains papiers présentent des zones irisées bleuâtres là où les sels d'argent sont les plus denses (photos du début du XXe siècle).

La traduction de ces clichés en niveaux de gris va enlever une "partie de la vérité" et pourra être ressentie comme une altération du document. On choisira dans ce cas une numérisation en couleurs; les seules retouches porteront sur la densité du cliché au cas où le numériseur n'aurait pas rendu correctement l'étagement des noirs.

Enfin les photos virées en sépia, virage à l'or, bleu cyanhydrique, seront traitées en couleur pour préserver leur aspect d'origine.



Les photos anciennes sont parfois tirées sur des papiers ayant subi un "virage". De gauche à droite: virage en noir gravure, virage sépia, virage cyan, ou "au prussiate" (princesse Potochka, collection Braun au CERARE à Mulhouse).

L'enregistrement du fichier maître se fera au format TIFF compression LZW ou ZIP, sans dégradation de l'image.

2. 7.4.2. Photographies en couleur

Les photos en couleur contemporaines présentent pour certaines des dégradations dues au vieillissement des copulants. Une dominante rouge-violette apparaît sur l'ensemble de l'image, les verts et les bleus s'affaiblissent

Pour transmettre correctement le message d'origine (qui est rappelons-le un message contemporain) il serait souhaitable de retoucher, ou au moins de rattraper les couleurs pour obtenir un cliché "présentable".

Complément : La retouche couleur

Il est à noter que ce travail est à faire dès que possible sur les fonds de photos en couleurs, car le processus de vieillissement est inexorable, et s'accélère après 20 ans de conservation à température ambiante.



La retouche de cette photo en couleurs, même si elle n'a pas restitué totalement les couleurs d'origine, est cependant nécessaire pour retrouver lisibilité et pertinence d'usage

Dans le cas où l'on posséderait les négatifs de ces clichés en couleurs, un retraitage de ces négatifs sur papier peut donner des résultats d'excellente qualité. D'une part la performance des analyseurs couleur a fortement progressé, et avec eux les colorants des papiers, d'autre part, les négatifs se conservent mieux que les papiers car leur support en acétate n'est pas imbibé de produits chimiques participant à la dégradation des copulants (rinçage insuffisant des épreuves).

L'acquisition se fera avec un numériseur à plat, et dans le mode RVB, car le spectre colorimétrique des tirages en couleurs ne dépasse pas celui du CMJN (Cyan Magenta Jaune Noir) utilisé en imprimerie. (Le spectre RVB est globalement plus étendu que le spectre CMJN)

La retouche se fera en utilisant des filtres qui détectent et éliminent les dominantes ("cast") Un filtre niveaux automatiques peut donner de bons résultats car il détecte le bruit de fond (une dominante uniforme) sur les composantes de l'image et l'élimine de fait en redynamisant les niveaux de chacune des couches colorimétriques.

L'enregistrement du fichier maître se fera au format Tiff compression LZW ou ZIP, sans dégradation des couleurs.

8. Numérisation des gravures

Ce chapitre décrit les règles nécessaires à l'exécution des travaux de numérisation des gravures (eaux-fortes ou estampes et similaires) ainsi que les pièges à éviter.

La numérisation des gravures est de pratique facile dès que l'on a appréhendé le procédé de traduction des traits en mode bitonal.

A. 8.1. Numérisation au bon mode

Attention

Il est important de bien lire ce chapitre, car le traitement des gravures est très mal pratiqué par 90% des éditeurs et imprimeurs : ils confondent plus ou moins volontairement le mode bitonal (trait) et le mode niveaux de gris (photo). On comprendra pourquoi en lisant ce qui suit !

La méthode de numérisation décrite ci-dessous est élaborée pour un résultat sans altération de la gravure, notamment pour la reproduction en ouvrages imprimés. Pour la mise en ligne de documents destinés à l'impression domestique (**catégorie 3**), la résolution choisie pourra être arbitraire (voir tableaux des solutions de numérisation) et les gravures seront dégradées partiellement dans leur finesse et dans leur tonalité, du fait de l'épaississement ou de l'effacement partiel des traits.

Les gravures sont des dessins faits de lignes au trait noir sur fond de papier (assimilable à du blanc) qui présentent des zones de haute densité (traits très gros et proches) et des zones de faible densité (traits très fins et espacés).

La numérisation se fera prioritairement en mode **bitonal** (bitmap) et non en niveaux de gris comme on le voit couramment (erreur fatale et commune). Le mode niveaux de gris est satisfaisant pour une lecture seule à l'écran, mais lors de l'impression, les traits de la gravure sont mangés par la trame de l'imprimante et, ce qui est plus grave, la gravure n'est pas numérisée correctement, car les zones franches entre le noir et le blanc ont été transformées en flous gris, ce qui est une **altération de la réalité**. Par ailleurs, le fichier en niveaux de gris a une taille plus importante qu'un fichier en mode bitonal à plus haute résolution, tous deux décrivant un document de la même dimension physique !

Gravure acquise en mode bitonal et gravure en niveaux de gris. Le niveau de gris altère les traits de la gravure dans le fichier enregistré.



Complément : Réimpression des gravures

L'acquisition en mode bitonal semble moins fine, car l'affichage à l'écran est synthétisé par groupement de pixels noir et blanc, qui apparaissent découpés en escaliers.

Cependant les pilotes d'imprimantes et flasheuses d'imprimerie lissent les «escaliers» pour donner un aspect de trait continu et net. Reproduites de cette manière, les gravures retrouvent toute leur précision et leurs détails subtils.

Si en revanche on choisit un mode de reproduction en demi-tons (niveaux de gris), l'imprimante ou la flasheuse vont transformer le fichier en une similigravure photographique. A ce moment, ce sont les points de la trame qui vont représenter tant bien que mal les fins traits de la gravure. La figure ci-dessous montre le résultat très agrandi, et on peut juger de la dégradation des traits d'origine.



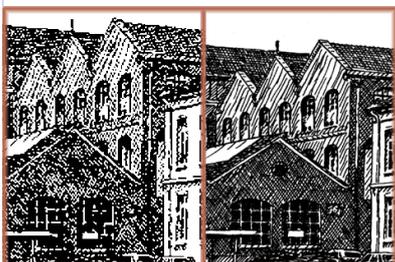
Résultats de l'impression à partir d'un fichier bitonal et à partir d'un fichier en niveaux de gris. La gravure traitée en simili est complètement dénaturée

Complément : Réglage de la résolution

Pour numériser correctement une gravure et restituer les traits dans leur intégralité, il convient de régler la résolution en dpi par rapport à l'épaisseur du trait le plus fin. On peut déterminer pour une série de gravures de la même provenance une finesse limite du trait, par sondages à l'aide d'un compte-fils gradué en centièmes de mm ou en millièmes de pouces.

La définition sera correcte, si le trait le plus fin est reproduit sur 3 pixels de largeur. Une approche pratique consiste à faire des essais de numérisation sur une zone à traits fins, puis d'agrandir à l'écran pour constater si la règle des 3 pixels est atteinte (le trait ne sera pas interrompu et sa largeur variera entre 2 et 3 pixels aux endroits les plus ténus) . Cette approche sera faite chaque fois que l'on voudra optimiser la taille du fichier pour le stockage.

Lorsque la gravure est destinée à l'impression offset, la méthode est de systématiquement numériser à 2400 dpi, ce qui évite les tâtonnements et les conversions ultérieures.



A gauche, une résolution insuffisante(72dpi): les traits de la gravure sont hachés, et les détails sont occultés.

A droite, une numérisation à 400dpi, résolution suffisante pour rendre les traits de l'original.

Complément : Réglage des densités

Pour une série de gravures d'un ouvrage, il est nécessaire de faire des essais préalables, car la difficulté est d'obtenir une gravure dont le rendu est complet sans apparition des mouchetis provenant du grain du papier. Le numériseur va générer une image à seuil de densité, c'est à dire qu'il va interpréter les nuances de l'original soit en noir, soit en blanc (0 ou 1).

- Si le seuil de densité moyenne est trop décalé vers le bas, le grain du papier va être interprété en noir et un mouchetis couvrira toute la gravure. Par contre les traits les plus fins seront rendus sans interruption, et même épaissis. Cet épaississement se ressentira aussi dans les zones sombres, et les traits serrés seront transformés en un bloc noir sans détails.



Gravure numérisée avec un seuil de conversion trop bas: le grain du papier apparaît et les détails sont bouchés.

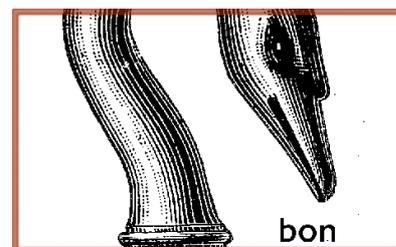
- Si le seuil de densité moyenne (conversion) est trop haut, les traits fins vont disparaître et la gravure perdra ses nuances



Les traits qui ne sont pas d'un noir absolu sont dégradés et amincis.

- La numérisation correcte sera obtenue en réglant manuellement le seuil de conversion et en procédant à des essais successifs.

Le seuil de conversion correspond à la densité moyenne des traits de la gravure : les zones ténues sont traduites en noirs, alors que le grain du papier est correctement éliminé.



La numérisation correcte d'une gravure laisse toutefois apparaître de petits points isolés correspondant à des défauts du papier ou à des petites taches d'encre. Ils sont à enlever à la gomme numérique, si l'on veut incorporer la gravure à un ouvrage.

Possibilités du numériseur

Il est important lors de l'achat d'un numériseur de s'assurer que le seuil peut être réglé manuellement à travers le logiciel de capture (ce qui est le cas de tous les numériseurs professionnels).

B. 8.2. Cas spéciaux

Il est parfois nécessaire de numériser une gravure en niveaux de gris :

- a. lorsqu'elle doit être réduite en taille pour une impression domestique,

- b. si le rapport de taille d'impression (en ouvrage) est inférieur à 50%.
- c. lorsqu'elle est à diffuser sur la toile ou pour consultation à l'écran uniquement.

On peut :

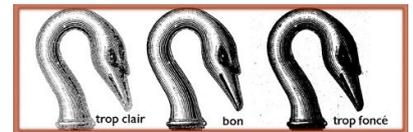
1. soit acquérir directement la gravure en niveaux de gris,
2. soit convertir un fichier bitonal en niveaux de gris, en réduisant la résolution pour l'adapter à l'usage final

Ce sont les seuls cas pratiques de transformation du fichier bitonal en niveaux de gris. Il est à noter que cette transformation, avec un rééchantillonnage du nombre de pixels formant l'image, peut se faire dans un logiciel de retouche photo, muni de convertisseurs performants.

La figure ci-dessous montre le rendu des précédentes gravures après transformation en niveaux de gris et réduction de taille en pixels.

On s'aperçoit en outre que si la première acquisition en mode bitonal n'est pas correcte, la conversion en niveaux de gris ne rattrapera pas les défauts de densité.

Gravures du cygne converties en niveaux de gris pour diffusion à l'écran : les défauts inhérents à l'acquisition ne pourront pas être corrigés par ce biais.



9. Numérisation des imprimés

Les imprimés traités dans ce chapitre sont les publications contenant des textes, des photographies en similigravure, des dessins, issus de divers procédés d'impression dits photomécanique.

Ne sont pas concernés par ce chapitre, les imprimés administratifs ou formulaires, qui ne comprennent que du texte et des filets. Ceux-ci sont à traiter comme les manuscrits.

Les ouvrages qui sont faits de gravures au trait et de texte(sans similigravures) sont à numériser en suivant les règles applicables aux gravures.

Comme le traitement des imprimés est assez particulier, il est nécessaire de comprendre dans un premier temps leur procédé d'élaboration, pour bien les appréhender lors de la numérisation.

A. 9.1. Contenu des imprimés, procédés d'obtention

Les imprimés contiennent du texte et des images en «similigravure», reproduites à partir des photographies par des procédés dits «photomécaniques». Leur nature a été transformée pour s'adapter à la création d'une forme imprimante, c'est-à-dire que les valeurs tonales variant en continu ont été transformées en points élémentaires de nature bitonale, à travers une trame : c'est la similigravure qui est à la base de toutes les reproductions imprimées modernes.

Nous verrons que cette similigravure va imposer des choix lors de la numérisation. Il n'est pas toujours possible d'acquérir ces originaux en mode bitonal, aussi faudra-t-il «détramer» certaines images photographiques imprimées en similigravure.

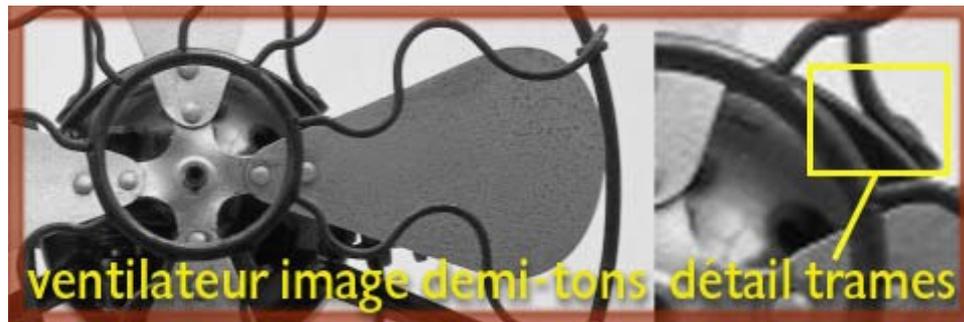
Avant de classer les imprimés, il est indispensable de comprendre le principe de la similigravure pour bien évaluer les difficultés qu'elle apportera dans la numérisation.

B. 9.2. Principe de la similigravure

La similigravure est nécessaire à la confection d'une forme imprimante appelée communément le cliché. C'est cette forme imprimante qui va reporter l'encre sur le papier. Comme cette forme est encrée de manière uniforme par la rotative, il est nécessaire de fragmenter l'image en points de différentes grosseur afin de rendre les nuances de densité

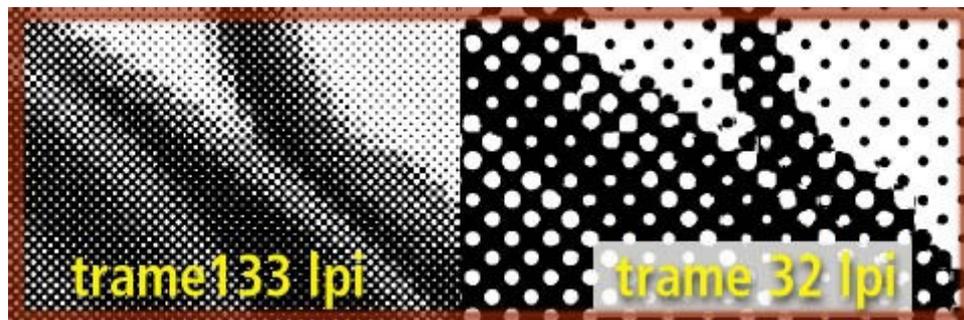
de l'image. Ces points sont répartis de manière régulière selon une trame, dont le nombre de point par pouce constitue la linéature. Dans les zones claires, les points sont petits et de fait espacés, donnant l'illusion d'un gris clair; dans les zones foncées les points sont gros et se recouvrent, donnant l'illusion du noir dense.

On l'a appelée similigravure, car elle a remplacé les gravures au trait faites à la main par des artistes graveurs au moyen de photographies traitées par ce procédé photomécanique, peu avant 1900. L'image tramée était obtenue en projetant l'image en niveaux de gris sur des émulsions à haut contraste à travers des plaques de cristal empreintes de micro prismes. Ensuite, le cliché en *simili* était obtenu par une attaque à l'acide des zones non exposées, d'une plaque de zinc photosensibilisée.



Principe de la similigravure (1)

Pour reproduire le ventilateur ci-dessus dans un document imprimé, on partira d'un fichier en niveaux de gris à 300 dpi. Il sera converti en points (au mode bitonal) pour l'impression. Le carré délimitant le "détail trames" sur cette figure correspond aux reproductions tramées montrées ci-dessous



Principe de la similigravure (2)

La figure ci-dessus montre l'agrandissement de deux trames d'imprimerie appliquées à la reproduction du détail ciblé sur la figure ci-avant. La trame 133 lpi s'applique à un magazine, la trame 32 lpi à une affiche de grande taille.

"lpi" veut dire 'Line per Inch' c'est à dire lignes par pouce. Ce terme provient du procédé de fabrication des trames en cristal, qui étaient rainurées pour former les pyramides condensant la lumière.

Complément : Interaction entre la trame et les pixels

La presse d'imprimerie procède en quelque sorte en toutes circonstances en mode bitonal et la «flasheuse» génère des images bitonales (noir 100% et blanc 100%) qui sont reportées sur un film, puis sur la plaque offset.

Pour les images en couleurs il faut générer quatre plaques que l'on encrera avec des encres dites primaires, le Cyan, le Magenta, le Jaune et le Noir (CMJN).

Le résultat d l'impression est une image à tons non continus, faite de points alignés selon la linéature (ou fréquence) de la trame. Lors de la numérisation, le numériseur va superposer

sa propre trame (celle des pixels formant l'image en acquisition), résultant en un moirage disgracieux, comme lorsque l'on superpose deux épaisseurs de l'étoffe d'un rideau.



Moirage

Lors de la consultation à l'écran, en changeant d'échelle, il se produit des moirages dûs à l'interférence entre la trame originale et la grille des pixels de l'écran.

Lors de la réimpression, il se produit le même phénomène entre la trame originale du document et la trame de l'imprimante qui reproduit le document.

C. 9.3. Principe de l'héliogravure

La reproduction des documents à tons continus peut aussi se faire par le truchement de l'**héliogravure**, du grec Hélios pour le soleil = procédé utilisant la lumière.

Le résultat imprimé est une **héliographie**.

Alors que la similitravure employée pour la plupart des procédés typographiques ou offset, procède de la transformation de tonalités progressives en valeurs bitonales, l'héliogravure est basée sur une gravure en creux dont on a réglé la profondeur par un procédé d'insolation suivi d'une attaque à l'acide. L'encre est alors distribuée d'une manière progressive donnant l'illusion d'une continuité parfaite des tons. C'est la quantité d'encre déposée sur le papier qui fixe la densité de l'image. Chacun des réservoirs d'encre formés sur le cylindre d'impression héliographique est cloisonné par les filets de la trame d'exposition.

Il n'est quasi pas possible de distinguer la trame d'une héliogravure, tant la linéature est fine. Si la fluidité de l'encre est bien choisie par rapport au papier, l'illusion de continuité est parfaite, car l'encre fuse et recouvre les zones blanches laissées par les cloisonnements.

La plupart des cartes postales sont imprimées en héliogravure, ainsi que les livres de photographie artistique, les beaux catalogues, et les magazines du début du XXe siècle aux années 80. Beaucoup de ces reproductions sont faites en utilisant des encres de couleur sépia, bleu hirondelle, vert américain, etc...

L'héliogravure est toujours pratiquée en imprimerie, c'est le procédé utilisé pour les journaux à très grand tirage comme les magazines féminins, les programmes télé... Mais la monochromie a cédé la place à la quadrichromie.

Les héliographies sont à numériser sans détramage, ainsi l'on conserve la finesse des détails. Beaucoup de photographies d'ouvrages historiques sont issues d'agrandissements de cartes postales décrivant avec une grande finesse les détails de la photo originale.



Image en héliogravure agrandie, acquise sans détramage. Le rendu est quasi-photographique. Les publications en héliogravure et cartes postales utilisaient des encres de couleur tel que montré dans la partie droite de cette image.

(doc Arthaud- l'Alsace par Hélène Mégret et Robert Minder- impression Braun à Mulhouse)

D. 9.4. Procédés de numérisation selon les objectifs

Les ouvrages les plus anciens sont les plus faciles à reproduire, en effet ils comprennent des textes, et des images en gravure sur métal ou sur bois, le tout en noir sur fond papier. Dans ce cas, la numérisation bitonale s'applique de facto, en ayant soin de régler la définition de manière à sauvegarder les traits les plus fins (**catégorie 3**), à éliminer la couleur et la texture du papier, si celles-ci ne concernent pas l'objectif de la sauvegarde

Lorsque l'on va numériser des imprimés «contemporains» en grande quantité, il faudra plus ou moins faire un choix quant au procédé pour les pages contenant à la fois du texte et des **images en similigravure** (décrivant des photographies en demi-tons):

- a) soit on choisit le procédé bitonal à basse résolution et les images en simili gravure sont altérées, voire sacrifiées si elles sont fines (**catégorie 2**),
- b) soit on choisit le procédé bitonal à haute résolution et les images en simili gravure sont reproduites avec leur trame originale, mais le volume du fichier augmente et la consultation détaillée des gravures nécessite leur impression (**catégorie 3**),
- c) soit on choisit un procédé en niveaux de gris ou en couleurs avec un détramage des images en similigravure, et le texte apparaîtra un peu flou, mais la consultation à l'écran restera aisée (**catégorie 2**).

Complément : Procédé bitonal

Dans le procédé bitonal appliqué à des originaux comprenant et/ou des gravures et des similigravures à trame grosse (moins de 100 lpi ou trame inférieure à 100 lignes par pouce), on peut prendre le parti de reproduire le tout en 300 dpi ou plus. C'est infiniment plus simple pour l'opérateur à partir du moment où les essais de densité sont effectués. Il est à noter que dans ce cas, la consultation à l'écran n'est pas «agréable», mais l'impression est de la meilleure qualité pour chacune des imprimantes utilisées. La photographie cependant peut montrer des contrastes variables et un aspect granuleux sur les imprimantes à basse définition (moins de 600dpi).



Numérisation bitonale en 300 dpi d'un catalogue de facture soignée : la lecture est aisée, les images sont un peu dégradées, voire moirées, mais acceptables.



Numérisation bitonale en 600 dpi : les images sont plus fines et le document est reproductible dans de bonnes conditions, même en imprimerie de type "bulletin d'histoire locale"

Attention !

Sur les exemples ci-dessus, les images ne sont pas réellement au mode bitonal, il s'agit d'une simulation permettant d'apprécier des différences, comme si l'on lisait le document réimprimé, sans s'approcher pour voir les détails.

Complément : Numériser des pages entières en tons continus

La numérisation des pages entières en niveaux de gris avec détramage conduit à rendre un peu flous les caractères du texte, les filets et gravures au trait, au profit du rendu des photos.

Ce procédé a l'avantage d'obtenir assez facilement une page apte à l'impression domestique, mais parfois dans un volume supérieur à celui d'un fichier bitonal. Il faut veiller à détramer efficacement les images, car ce sont les points de la trame qui augmentent la taille du fichier (leur description est plus compliquée qu'une étendue de gris uniforme)



Aspect légèrement flou d'une page numérisée en niveaux de gris avec détramage global

Rappel !

Il n'est pas nécessaire de détramer les ouvrages en héliogravure. L'acquisition est fait en mode niveau de gris sans autre artifice qu'une bonne balance des luminosités.

Complément : Choix avancés concernant les numérisations d'imprimés

Avant de se lancer dans les détails exécutoires d'une numérisation d'imprimés, il est nécessaire de se rappeler :

- que la numérisation de l'imprimé sera toujours un compromis entre la qualité et le temps ou l'espace disque que l'on veut y consacrer,
- que la numérisation des imprimés est très techniques (sa réussite se fait au moment même de la numérisation: on veillera donc à la confier à du personnel compétent).

Pour la reproduction de pages composites

Il est possible de reproduire assez simplement une page dans un bulletin d'information, à l'aide d'un fichier photo unique, comme indiqué ci-dessous.

- 1. Acquérir la page entière en mode niveau de gris, sans détramage et en 300dpi: les textes apparaissent assez nets, les photos apparaissent non détramées.



Page brute de numérisation : la photo est moirée, le fond est gris, les caractères sont "mous" - Cette page au format tiff avant retouche fait 2 mégaoctets et, après retouche seulement, 700Koctets. La grande taille du fichier de la page

brute est due à la description complexe de la trame et du fond en papier.

- 2. Ouvrir le fichier dans un logiciel de retouche, sélectionner uniquement l'image (ou les images) et appliquer un filtre de détramage en post-traitement sur ces zones.

Intervertir la sélection et augmenter le contraste ou changer les niveaux sur le texte: le fond en papier disparaît ou s'éclaircit et les caractères gagnent en vigueur.

La page est maintenant prête à être intégrée dans une composition ou en ligne, de manière simple. Mais il faut avoir à l'esprit que l'impression de ce document se traduira par un tramage des textes par l'imprimante ou la flasheuse de l'imprimeur.

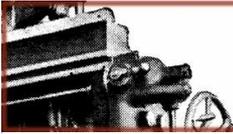


Page retouchée : la photo est détramée, sa densité est adaptée, le fond est clair et les caractères bien nets

Pour la reproduction fine en ouvrages imprimés

Pour la réédition d'ouvrages anciens en imprimerie, il faudra assurer une reproduction des textes et gravures sans trame, et des photographies (similigravures) avec une trame. IDEM On procède par une bitonalisation intégrale des documents sources, puis par la méthode des "deux couches" si les similigravures ne se traduisent pas bien à la réimpression.

- 1. Acquérir les pages entièrement au mode bitonal en haute définition (1200 à 2400 dpi), la photographie apparaît avec la trame d'origine dans ce premier fichier qui sera utilisé pour la restitution des textes, gravures et filets. Si l'original a été imprimé sur un papier glacé ou couché de bonne qualité, on pourra conserver ce document dans son intégralité en bitonal et le reproduire en offset tel quel, si les points de la trame sont bien formés et nets. Dans le cas contraire, on devra passer à la seconde phase: transformation de la photo tramée en niveaux de gris.



des cha
faibles p

Première couche de la page acquise en mode bitonal: les caractères qui vont être conservés pour la reproduction sont bien nets et seront reproduits sans tramage

Acquérir à part les similigravures en mode niveau de gris avec un détramage et pratiquer une correction globale de l'image en matière de contraste, niveaux, etc.

- 2. Les fichiers en niveaux de gris sont ensuite replacés par dessus la page bitonale, dans un logiciel de composition (prépresse). La réserve de l'image en niveau de gris est très légèrement agrandie pour couvrir le fantôme en mode bitonal, ou celui-ci est effacé dans le fichier et remplacé par un cadre fin plus petit que l'image, juste pour l'aide au positionnement. Ces images en niveaux de gris peuvent être retravaillées et même leur rendu amélioré par rapport à l'original pour augmenter la lisibilité.



des cha
faibles p

La page finie : la photo est détramée, sa densité est adaptée, le fond est clair et les caractères seront imprimés au mode bitonal avec la définition maximale de la photocomposeuse

Attention !

L'acquisition d'un document au mode bitonal est relativement aisé et peu coûteux, par contre le procédé des deux couches, nécessitant des retouches et de la mise en page reviendra plus cher, et de plus en plus cher selon le nombre des photos en similigravure. L'imprimeur devra en outre être très compétent (inutile de demander ceci à une imprimerie de labeur).

10. Numérisation des plans

Les plans en centre d'archives sont décomposés en deux grandes familles:

- les tracés,
- les tirages.

On appelle **tracés** tous les plans "originaux" qui ont été élaborés par le dessinateur, sur des supports divers: le papier fort, le papier entoilé, le calque.

Les plans modernes issus de la DAO (Dessin Assisté par Ordinateur) sont imprimés sur du papier ordinaire par des imprimantes à jet d'encre ou à poudre. On les placera dans la catégorie des tracés originaux.

On appelle **tirages**, les épreuves diazoïques ou de Van Dycke exécutées par un procédé de reproduction héliographique. Le papier sensible est insolé (lampe à ultra-violets) à travers le tracé original sur son calque ou papier mince. Le papier sensible est ensuite révélé pour obtenir un tirage destiné à l'utilisateur de l'information.

Le contre-calque retouché à l'encre de chine est à considérer comme un tirage.

Les tirages diazoïques se dégradent dans le temps, sous l'action de la lumière. Le palissement est irréversible.

Les tirages cyano-fer (ferricyanure) fond bleu, trait blanc sont aussi sensibles à la lumière, mais se régénèrent dans le noir et en présence d'humidité.

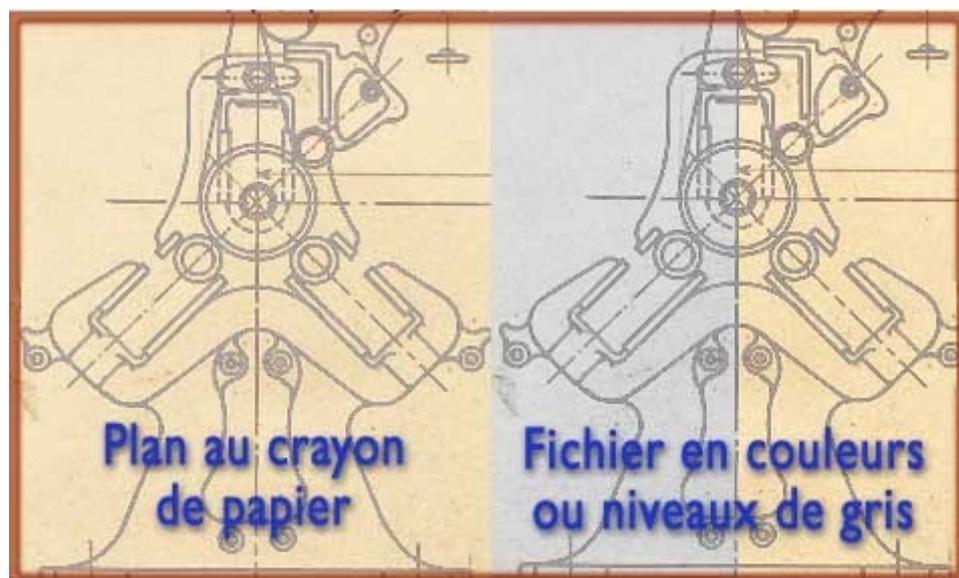
A. 10.1. Plans tracés

Domaine de définition

On appelle "tracés" tous les plans qui ont été élaborés par le dessinateur, sur des supports divers: le papier fort, le papier entoilé, le calque.

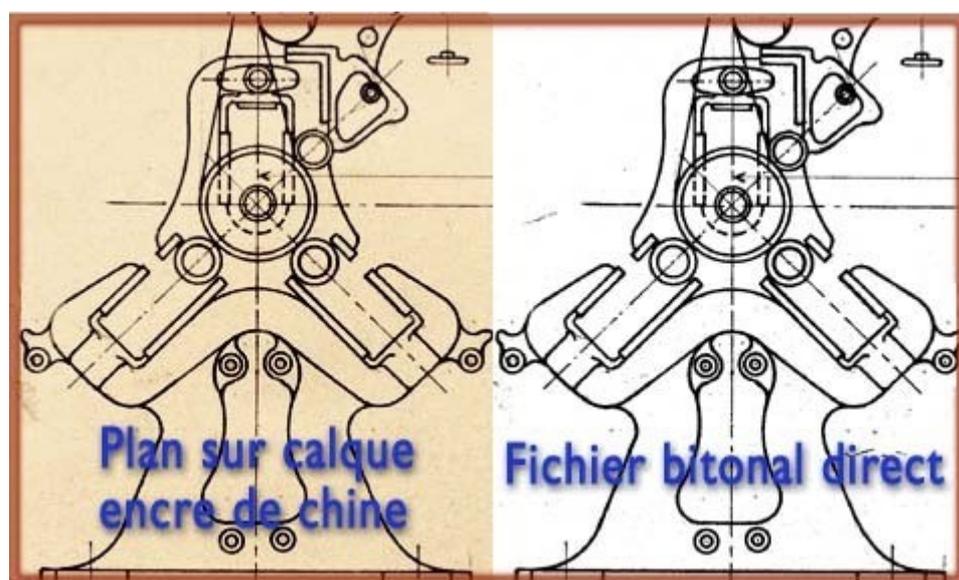
Plans tracés au crayon

Ces plans sur papier ou sur calque ont des nuances délicates et des traits très fins. La numérisation se fera en tons continus. Le mode couleurs sera préférable, car il permet de mieux discriminer le trait du papier, surtout s'il est jauni.



Plans tracés à l'encre de chine.

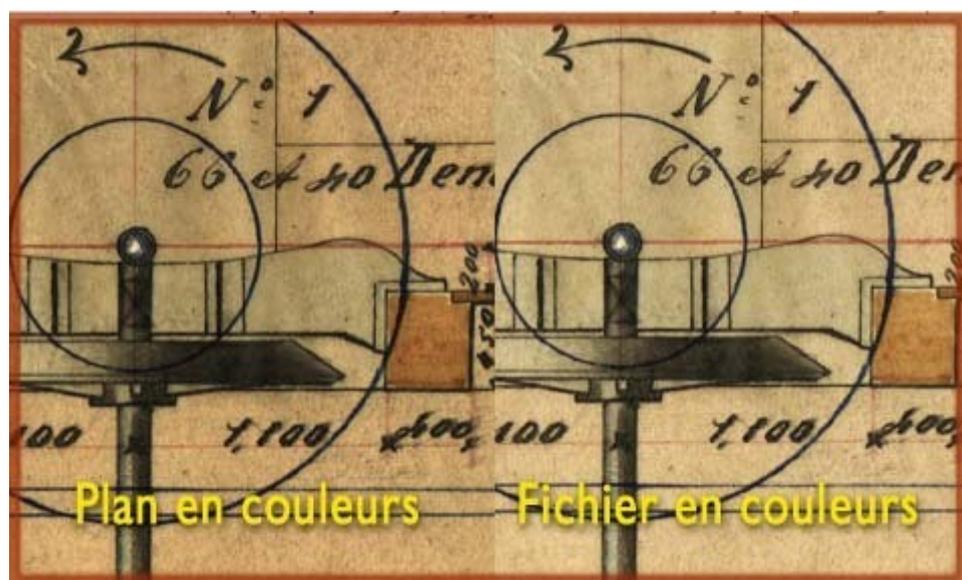
Les plans à l'encre de chine, sur calque le plus souvent, ou sur papier, sont particulièrement adaptés à la traduction par un fichier bitonal. L'acquisition est aisée, mais la fragilité du calque desséché peut poser des problèmes de manipulation (cassure des plis éventuels).



Plans coloriés

Les plans coloriés sont réalisés sur papier fort ou entoilé. Ils représentent le plus souvent des ensembles architecturaux ou de machines.

Les plans coloriés sont à acquérir en couleurs.



Plan et fichier en couleurs

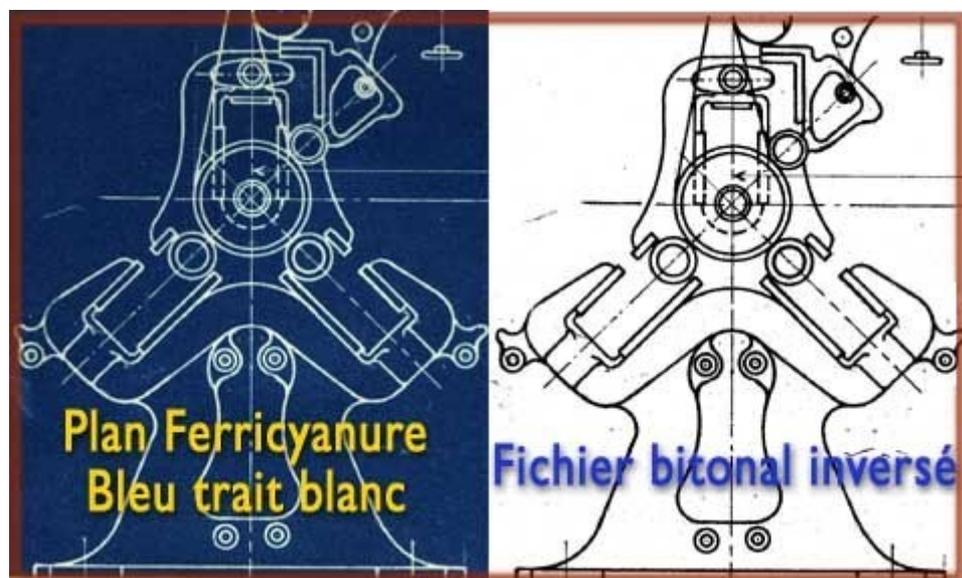
B. 10.2. Tirages de plans

Tirages au ferricyanure ou aux sels d'argent

Les plans au ferricyanure sont bleus, les traits apparaissent en blanc.

Les plans aux sels d'argent ont un fond brun et des traits blancs.

On procède à une acquisition bitonale, puis le fichier est inversé pour obtenir un plan que l'on peut imprimer facilement.

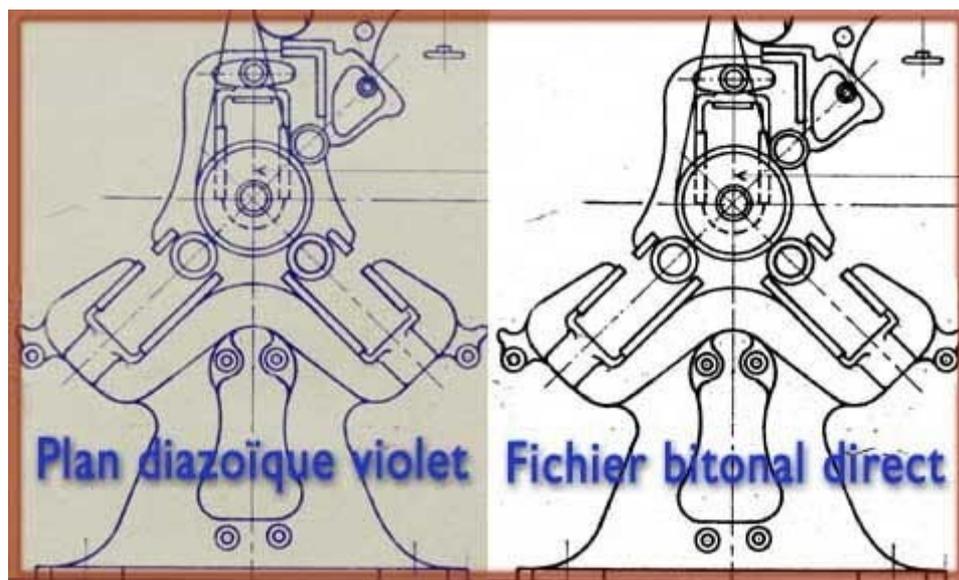


Tirages au ferricyanure

Tirage diazo violets

Les plans diazo violets sont à traiter en bitonal.

On passera au fichier à tons continus si le fond est foncé (tirage sous-exposé ou traits ténus) afin de restituer les détails incertains.

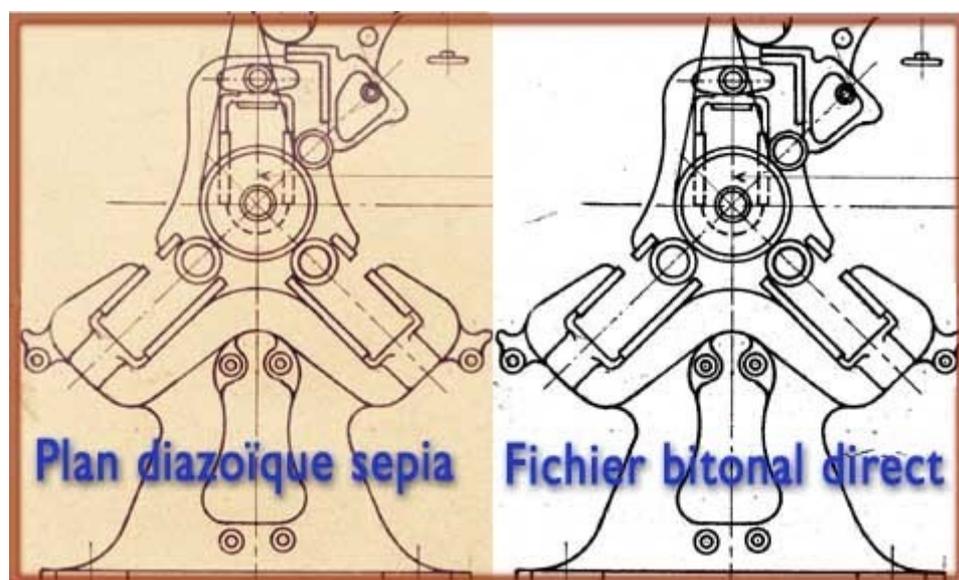


Tirage diazo

Tirages diazo sepia et contrecalques

Les "contre-calques" diazoïques et certains tirages sur papier présentent un aspect trait brun sur fond papier, avec du tain, car on a tendance à les sous exposer pour ne pas perdre les traits fins.

Avec un bon réglage de seuil bitonal, on arrive à les traduire "au trait". Dans le cas où certains plans de la série montreraient un contraste insuffisant (contre-calque de Nième génération ou tracé au crayon), on les traiterait en tons continus comme les plans au crayon de papier.



Tirages diazo sepia

11. Solutions de numérisation en tableaux

La section qui suit décrit en tableaux des « recettes » pour la numérisation de divers documents.

Dans un premier temps, ces tableaux analysent les solutions de numérisation à partir de la typologie du document : qualités à préserver, moyen de numérisation.

Dans un second temps, des solutions pratiques incluant formats et résolutions sont proposés pour les applications les plus courantes, de la préservation à la diffusion en ligne, toujours en partant du type de document.

A. 11.1. Inventaire des typologies de documents

Les tableaux ci-dessous sont basés sur un inventaire typologique dressé dans un centre d'archives historiques et industrielles qui donne une idée de la variété des documents, avec des propositions de solution de numérisation pour la reproduction sans altération, à l'échelle 1. Il n'est pas exhaustif, d'autres cas peuvent être identifiés. On cherchera à les rattacher à l'une ou l'autre famille.

Les documents iconographiques appartiennent à deux grandes familles:

- les images sans tramage photomécanique,
- les reproductions en similitravure.

Nous commencerons par un inventaire typologique des documents, à titre d'introduction aux solutions de numérisation.

En effet, la bonne numérisation se définit à partir du document et non à partir d'une habitude ou d'une machine disponible.

1. 11.1.1. Documents figurés sans tramage photomécanique

Ces documents peuvent être de première génération ou sont ceux qui n'ont pas été reproduits par une similitravure ou un procédé photomécanique qui en change la nature.

Ce sont essentiellement :

- les peintures, dessins et oeuvres d'art,
- les gravures sur bois (xylographies),
- les gravures à l'eau-forte, au burin (dites aussi estampes),

- les lithographies (elles ne sont pas tramées et ont été directement dessinées par l'auteur sur la pierre),
- les photographies sur papier, négatifs et positifs photographiques,
- les héliogravures, car leur trame est invisible et elles donnent l'illusion de tons continus (beaucoup de cartes postales sont imprimées avec cette technique),
- les calcographies (offset), les ouvrages imprimés au trait ou avec de grands aplats sans trame,
- les tirages de plans qui sont réalisés par contact et donc gardent la finesse du trait, sans artifice pour le transformer sur une forme imprimante.

2. 11.1.2. Plans et dessins techniques

Les plans originaux au lavis ont pu être entièrement dessinés sur un support en papier fort (entoilé).

Certains ont été coloriés sur un tirage héliographique à trait noir.

Les plans d'exécution sont originalement tracés sur du papier calque.

Les calques sont ensuite reproduits par contact sur du papier héliographique appelé couramment "papier diazo".

Les plans peuvent être tirés sur :

- papier au ferro-prussiate ou cyanotype à trait blanc sur fond bleu (le plus commun dans les anciens plans),
- papier cyano-fer (donne un trait bleu sur fond blanc; contemporain du cyanotype),
- papier négatif aux sels de fer et d'argent (brun) ou épreuve Van Dyke,
- papier héliographique ou diazotype (ce papier donne un trait marron ou violet ou noir sur fond blanc; ils sont les plus communs dans les plans d'après 1945).

Afin de préserver leur intégrité, on s'attachera à accorder la résolution à la finesse de l'image, c'est à dire à l'épaisseur des traits les plus fins.

Les contretypes ou tirages de plans dessinés au crayon présentent un fond très coloré, il sera souvent nécessaire de numériser ces plans en niveaux de gris pour distinguer les traits.

3. 11.1.3. Photographies, plaques, films

Les négatifs et plaques de verre ont plus ou moins des formats normalisés, car ils devaient entrer dans les appareils de prise de vue de différentes marques:

- formats des négatifs ou diapos : 4x4 cm; 6 x 4,5 cm ; 6x6 cm ; 6x9 cm pour les films en rouleaux;
- formats 9x13, 13x18 et 18x24 cm pour les planfilms;
- formats 9x13, 13x18 et 18x24, et 30x40 cm pour les plaques de verre.

On peut rencontrer de plus grands formats pour les négatifs ou positifs qui sont générés au cours d'un procédé photomécanique de traitement de l'image (photogravure en imprimerie) ou pour des plaques très anciennes.

Les papiers reprennent les mêmes formats que les plaques, car les tirages se faisaient par contact, mais il existe une foule de formats divers créés à la demande en coupant le papier : à une époque où les agrandisseurs n'existaient pas, ce procédé permettait de recadrer la photographie.

4. 11.1.4. Documents figurés en similligravure

Ces documents ont été transformés par un procédé de photogravure, destiné à interpréter les valeurs de demi-tons en points noirs ou blancs, pour satisfaire aux exigences de la forme

imprimante.

Ce sont en substance:

- les illustrations contenues dans des ouvrages, bulletins d'information, magazines et journaux,
- les illustrations soignées des catalogues de fabricants,
- certaines images reproduites sur papier couché, avec clichés de similitude en zinc.

Rappel : Attention !

On n'inclut pas dans cette famille, les héliogravures bien qu'elles soient issues d'un procédé photomécanique, mais n'utilisant pas la technique du point noir.

B. 11.2. Procédés de numérisation à partir de la typologie des documents

Les tableaux ci-après proposent des modes colorimétriques selon la nature des documents et une proposition de choix des moyens physiques de numérisation. Les formats d'enregistrement pour les fichiers maîtres seront à choisir dans le paragraphe suivant «solutions de numérisation à partir de l'usage».

1. 11.2.1. Manuscrits, dessins, gravures, lithos

Nature du document	Qualité principale à préserver	Mode conseillé	Moyen d'acquisition
texte manuscrit sur papier ordinaire	le texte lui-même	bitonal	numériseur multi-pages générant des fichiers PDF finis
texte manuscrit sur beau papier ou calligraphie artistique	la beauté du texte et du papier	couleur RVB	numériseur à plat pour les pages destinées à l'édition
ouvrages, grands livres, registres	le texte lui-même	bitonal	numériseur spécial à livre ouvert, ou à défaut, numériseur à plat.
dessins originaux au trait d'encre monochrome, sur papier blanc (dessins techniques ou artistiques)	précision du trait et fort contraste. Le papier n'est pas considéré comme faisant partie du dessin dans ce cas	bitonal	numériseur à plat jusqu'au format A3, et A2 pour certains cas. Numériseur à plans pour les formats supérieurs
dessins comprenant des aplats en couleur et des traits à l'encre ou plans coloriés (architecture)	conserver la netteté du trait et la franchise des couleurs	couleur RVB	numériseur à plat jusqu'au format A2; photographie ou numériseur de grand format au-delà
dessins au crayon, au fusain	préserver les nuances des ombres et traits tenus ainsi que le grain du papier	niveaux de gris ou couleur RVB	numériseur à plat jusqu'au format A2; numériseur de grand format au-delà
lithographies monochromes	conserver la finesse des points d'encre composant l'image. Des essais unitaires sont à faire pour ajuster le seuil d'acquisition bitonale	bitonal, ou niveau de gris si on ne maîtrise pas la technique du seuil	numériseur à plat jusqu'au format A2, numériseur de grand format au-delà
gravures au trait : tirages sur papier des eaux-fortes, gravures sur cuivre	conserver la finesse des traits et respecter les densités résultantes, sans garder trace du papier (ni des taches)	bitonal	numériseur à plat jusqu'au format A2 numériseur de grand format au-delà
gravures au trait : tirages sur papier des eaux-fortes,	conserver la gravure avec les défauts et le grain	couleurs RVB	numériseur à plat jusqu'au format A2

gravures sur cuivre	du papier, mais en altérant la finesse du trait à l'impression		numériseur de grand format au-delà
dessins à la sanguine, au pastel	conserver les nuances des ombres et traits ténus, et conserver la texture du papier	couleur RVB	numériseur à plat jusqu'au format A2 numériseur de grand format au-delà

2. 11.2.2. Peintures et petits objets plats

Nature du document	Qualité principale à préserver	Mode conseillé	Moyen d'acquisition
Aquarelles et gouaches sur papier à grain, aquatintes, collages	Conserver les détails, les nuances et le grain du papier	Couleur RVB	Numériseur à plat jusqu'au format A2
Peintures à l'huile, toutes techniques	Conserver les détails et l'impression de relief malgré la brillance du vernis	Couleur LAB, car le spectre des colorants peut dépasser le spectre RVB	Numériseur à plat jusqu'au format A2 numériseur de grand format au-delà
Peintures acryliques	Conserver les détails et la vivacité des couleurs	Couleur LAB, car le spectre des colorants peut dépasser le spectre RVB	Numériseur à plat jusqu'au format A2 numériseur de grand format au-delà
Petits objets plats : broderies, échantillons textiles, sceaux...	Conserver les détails et parfois la vivacité des couleurs pour les échantillons	Couleur RVB ou couleurs LAB si on constate un affaiblissement de tons	Numériseur à plat, avec autofocus ou mise au point de l'optique.

3. 11.2.3. Photographies, négatifs, diapositifs

Nature document	du	Qualité principale à préserver	Mode conseillé	Moyen d'acquisition
Photographies noir et blanc sur papier, simple retouchée à l'aérographe.	en sur ou à	Conserver l'ensemble des fins détails constituant l'image. Il est conseillé de numériser les petits formats (tirage contact) en 600 dpi	Niveaux de gris	Numériseur à plat.
Photographies monochromes sur papier en virage sépias, bleus, rouge, or..	sur	Récupérer les nuances de l'image originale. Il est parfois nécessaire de retoucher ces clichés pour les rendre plus lisibles	Couleur RVB	Numériseur à plat.
Photographies couleurs	en	Les photos en couleur dégradées sont parfois à retoucher pour diminuer les dominantes dues au vieillissement	Couleur RVB	Numériseur à plat.
Photographies négatives positives sur plaques de verre ou sur film 9x13 cm et au-delà, monochromes	ou sur	Prêter attention à reproduire une gamme complète du noir parfait aux hautes lumières en adaptant la courbe gamma à la densité de la plaque	Niveaux de gris	Numériseur à plat avec dos lumineux.
Photographies négatives sur films en bandes (16 mm, 24x36mm, 4x4 cm	sur films	Prêter attention à reproduire une gamme complète du noir parfait aux hautes lumières en adaptant la courbe gamma à la densité du négatif	Niveaux de gris ou couleur	Numériseur à négatifs ou adaptateur sur numériseur à plat à haute performance
Photographies négatives sur films en bandes rollfilms 6cm) noir et blanc et au-delà	sur films	Prêter attention à reproduire une gamme complète du noir parfait aux hautes lumières en adaptant la courbe gamma à la densité du négatif	Niveaux de gris ou couleur	Adaptateur du numériseur à plat à haute performance ou numériseur à tambour
Diapositives 24x36mm, 4x4 cm, sous cache 5x5 cm		Numériser en profondeur 42 à 64 bits pour préserver	Couleur LAB	Numériseur à négatifs, ou numériseur à

	toutes les nuances et contrastes		tambour
Diapositives moyen format et ektachromes d'imprimerie	Numériser en profondeur 42 à 64 bits pour préserver toutes les nuances et contrastes	Couleur LAB	Numériseur à tambour ou numériseur haute performance avec adaptateur

4. 11.2.4. Plans d'architecture et techniques

Nature du document	Qualité principale à préserver	Mode conseillé	Moyen d'acquisition
Dessins d'architecture au trait noir et trames collées	Conserver la finesse du trait et la densité de la trame	Bitonal	Numériseur grand format pour plans. (appareils de bureautique pour l'industrie)
Plans originaux sur calque ou sur papier, trait noir à l'encre, fond blanc et/ou tirages de bonne qualité et haut contraste.	Conserver la finesse du trait	Bitonal	Numériseur grand format pour plans. (appareils de bureautique pour l'industrie)
Tirages de plans à trait coloré (bleu-violet-brun-noir) sur fond coloré, manque de contraste	Il faudra chercher à chaque plan un seuil «idéal» pour éliminer partiellement le fond, sinon acquérir en niveaux de gris. (l'acquisition en couleurs peut se faire sur un numériseur spécialisé)	Bitonal	Numériseur grand format pour plans. (appareils de bureautique pour l'industrie). Certains appareils ont des commandes adaptées pour la « récupération » des originaux fades
Tirages de plans à trait blanc sur fond coloré (bleu-brun-noir)	Il faut rechercher le seuil pour ne pas «manger» les traits	Bitonal, inverser les tonalités pour obtenir des traits noirs sur fond blanc	Numériseur grand format pour plans. (appareils de bureautique pour l'industrie)
Tracés originaux de plans colorés au lavis technique, ou tirages colorés	Conserver les nuances des couleurs, le grain du papier, la finesse du trait noir	Couleur RVB	Numériseur grand format couleur pour plans ou livres. (appareil spécialisé pour la numérisation des archives)

C. 11.3. Solutions de numérisation à partir de l'usage du fichier

Le tableau ci-après propose un ensemble de modes colorimétriques, de résolutions et d'applications en partant de l'usage d'un type de document. Il conduit à la reproduction des documents à leur échelle d'origine (sauf pour les diapos et négatifs), des fichiers sous-échantillonnés étant à créer pour la mise en ligne. Les résolutions proposées sont adaptées aux documents dits «ordinaires», c'est à dire les publications imprimées courantes, les documents manuscrits, dactylographiés et les iconographies d'usage général. Les iconographies scientifiques sont à traiter au cas par cas avec une finesse améliorée (biochimie-spectres physiques...)

Il convient pour les gravures, les textes et les manuscrits, d'adapter la résolution à l'original, pour que le trait le plus fin ait une épaisseur d'au moins 2 pixels en modes niveaux de gris ou couleur et 3 pixels en mode bitonal.

Nature du document objectif	du et	Résolution recommandée du fichier d'utilisation	Description et de	Mode colorimétrie	Format d'enregistrement mono – page
Gravures ou similigravures à préserver, fichier source (catégorie_4)	ou à	1200 dpi à la taille d'impression	Mode bitonal	(Bitmap)	.tif avec compression LZW dans tous les cas
Images & documents en demi-tons à préserver, fichier source (catégorie_4)	& en à	300 dpi à la taille d'impression	Couleurs RVB ou niveaux de gris		.tif avec compression LZW dans tous les cas
Images d'un original en diapositive couleurs (pour préservation contre la dégradation) ou pour publication imprimée (catégorie_4)	d'un en	en moyenne 2400 dpi / dia 24x36 pour les iconothèques de sauvegarde ou selon la taille d'impression dans l'ouvrage	Couleurs LAB		.tif avec compression LZW dans tous les cas
Textes imprimés à lire en ligne (catégorie_1 et catégorie_2)	à lire en ligne et	150 dpi	Mode bitonal	(Bitmap)	.tif avec compression LZW dans tous les cas
Textes manuscrits à lire en ligne (catégorie_1)	à lire en ligne	200 à 400 dpi selon la finesse d'écriture	Mode bitonal	(Bitmap)	.tif avec compression LZW dans tous les cas
Gravures + textes en ligne pour impression domestique (catégorie_3)	+ textes pour	400 ou 600 dpi selon la finesse de gravure	Mode bitonal	(Bitmap)	.tif avec compression LZW dans tous les cas
Gravures pour l'imprimerie (catégorie_4) et reproduction de similigravures pour réimpression	pour	1200 ou 2400 dpi selon la finesse de l'ouvrage final	Mode bitonal	(Bitmap)	.tif avec compression LZW dans tous les cas
Fac-similés de documents monochromes pour consultation en ligne (catégorie_2)	de pour	150 dpi	Image en demi-teintes noir et blanc	Mode niveau de gris (256 niveaux)	.tif avec compression LZW ou .jpg pour l'insertion des sous fichiers dans

			une page Html (catégorie_W)
Photographies à imprimer dans une publication (catégorie_4)	300 dpi à la taille d'impression	Image en demi-teintes noir et blanc Mode niveau de gris (256 niveaux)	.tif avec compression LZW ou .jpg pour l'insertion des sous fichiers dans une page Html (catégorie_W)

12. Bibliographie en ligne

Ce chapitre n'est en rien une bibliographie exhaustive , elle est constituée de retours d'expériences servant à donner quelques conseils pratiques.

A. Numérisation

Une association de digitaliseurs pour la généalogie: utilisation du "bookeye"

<http://www.rod-neep.co.uk/index.htm>

Robotisation de la numérisation des documents avec un numériseur à livres équipé d'un tourne-pages, par l'université de Stanford

<http://www-sul.stanford.edu/depts/diroff/DLStatement.html>

Numériseurs à plat

- Doc. de numériseurs à plat pour les documents opaques et les films, profondeur d'analyse 48 bits et 4,2D (pour les archives photographiques)

<http://www.qubyx.com>

Numériseurs de livres sans contact

- Doc. du fabricant original du numériseur à livre ouvert "bookeye"

<http://www.imageware.de/en/>

ce numériseur est distribué sous beaucoup de marques diverses d'un pays à l'autre, voici une adresse où le même numériseur s'appelle "indus"

<http://www.indususa.com/bookscanners.php>

- Doc. du numériseur qui tourne les pages automatiquement, fabriqué en Suisse

<http://www.4digitalbooks.com/>

- Doc. de numériseurs à livre ouvert : Zeutschel

<http://www.zeutschel.de/>

- Doc. d'un numériseur à livre ouvert Minolta PS7000

<http://www.dpsmicro.com/brochures/digital/ps7000.pdf>

- ¹ <http://www.konicaminolta.fr/>

- Doc. de numériseur de livres couleur Digibook 6002

http://www.i2s-bookscanner.com/en/products_SUPRASCAN.asp

1 - <http://www.konicaminolta.fr/>

- Doc. de numériseur de livres couleur digibook 10 000
http://www.i2s-bookscanner.com/en/products_SUPRASCAN.asp
Distributeurs de matériel d'archivage
 - Spigraph, distributeur de numériseurs spécialisés pour l'archivage (France et Europe)
<http://www.spigraph.fr/>
 - Calligraphy, matériel de numérisation d'archives (Suisse)
<http://imprimerie-calligraphy-sierre.ch/>
- Bibliographie pour les pratiques de la numérisation (en Anglais)
 - Flatbed Scanner Review (Flaar), un banc d'essai des numériseurs et des publications relatives à l'imagerie de grand format
<http://www.flatbed-scanner-review.org/index.html>
 - Comment choisir un scanner (bibliographie américaine)
<http://www.scansoft.com/scannerguide/resources/books.asp>

B. Bibliothèques en ligne

Bibliothèques en ligne (exemples)

- Bibliothèque numérique Gallica de la Bibliothèque nationale de France pour accéder à la bibliothèque de Gallica, activer l'adresse ci-dessous:
<http://gallica.bnf.fr/>

et orienter la recherche sur l'auteur "Louis Figuiet", par exemple La bibliothèque vous propose de télécharger tout ou partie des ouvrages de l'auteur; e sont des ouvrages contenant des gravures au trait et du texte (catégorie 3).

Pour voir directement un document numérisé en mode bitonal, activer l'adresse ci-dessous. Le frontispice apparaît avec du fond, car les premières pages d'un ouvrage sont plus exposées au jaunissement que les pages internes (présence d'oxygène). Utiliser les boutons d'accès direct aux pages de l'ouvrage pour apprécier différents aspects de la numérisation bitonale.

<http://gallica.bnf.fr/Search?ArianeWireIndex=index&q=Louis+Figuiet&p=1&lang=fr>

- Plans, cadastre

Pour accéder à des plans numérisés en couleurs, activer l'adresse ci-dessous:

<http://www.cg78.fr/archives/home.html>

Le navigateur montre des réductions du plan à titre de première approche (catégorie 5) et permet le téléchargement en plus haute définition (catégorie 3) sans que ce soit un fichier assimilable à une copie de sauvegarde (de par sa taille, on ne pourrait le mettre en ligne).

- Cartes postales

Pour accéder à des documents numérisés pour un usage domestique, activer l'adresse ci-dessous

<http://web-2.cg37.fr/dipwebv4/adcpst/template/index.htm>

Le Conseil général de la Touraine vous propose à travers une banque de données de choisir des cartes postales anciennes: elles sont présentées en vignettes pour l'identification; vous pouvez télécharger gratuitement un fichier à basse définition pour la lecture à l'écran (catégorie 2) sous le nom anglais "e-card"; l'impression (médiocre) est possible en réduisant la taille de moitié; la reproduction de ces cartes à haute définition (catégorie 4) pour réimpression dans un ouvrage (ou pour voir les détails!) est disponible à titre payant.

13. Evaluation des connaissances

Avez-vous bien compris tout ce qui vient de vous être enseigné?

Si vous voulez le vérifier, faites les exercices proposés ci-dessous.

Si vous ne savez pas répondre, ne regardez pas trop vite le corrigé, travaillez à nouveau la (les) section(s) précédente(s) où vous découvrirez les solutions.

Bien sûr, si vous n'y arrivez vraiment pas, vous pouvez consulter les réponses. Ne les lisez pas avec précipitation mais avec une grande attention et surtout essayez de comprendre.

A vous de jouer...

Exercice 1

[Solution n°1 p 109]

Parmi les énoncés suivants concernant les archives virtuelles, lequel est faux?

- a. Les archives informatiques sont uniquement des documents créés originalement au moyen d'équipement informatique
- b. Archives numériques et archives électroniques sont des synonymes
- c. Les photographies numériques sont des archives électroniques
- d. Les dessins assistés par ordinateur sont des documents numériques
- e. Aucune de ces réponses

Exercice 2

[Solution n°2 p 109]

Associer les concepts et les énoncés suivants.

- I - c. Le nombre pixels
- II - a. Le mode d'encodage bitonal
- III - b. Le format

- | | | |
|----------------------------------------------------------|----------------------------------------------------------------------------------------|--------------------------------------------------|
| 1. Il détermine la résolution (ou définition) de l'image | 2. Il s'utilise pour la reproduction des documents et des textes illustrés de gravures | 3. Il détermine le type de numériseur à utiliser |
|----------------------------------------------------------|----------------------------------------------------------------------------------------|--------------------------------------------------|

Exercice 3

[Solution n°3 p 110]

Parmi les énoncés suivants, lesquels sont-ils vrais ?

- | | |
|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/> | 1. Il est recommandé de numériser le document physique dans la plus petite résolution possible puisqu'une grande résolution exige la création de fichiers énormes, ce qui est plus dispendieux |
| <input type="checkbox"/> | 2. Il existe une pratique uniforme pour la numérisation de tous les types de documents |
| <input type="checkbox"/> | 3. On peut transformer un document de haute résolution en basse résolution, mais pas l'inverse |
| <input type="checkbox"/> | 4. Les formats d'enregistrement sans pertes de données doivent être utilisés pour constituer une banque de fichiers source |
| <input type="checkbox"/> | 5. Le choix des formats d'enregistrement doit seulement se faire en fonction du type de documents numérisés et des besoins de l'organisme |

Exercice 4

[Solution n°4 p 110]

Parmi les énoncés suivants, lesquels sont-ils vrais ? (suite)

- | | |
|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/> | 6. La mise de photographies en ligne requiert que l'on offre aux internautes des documents de la plus haute qualité possible, donc comprenant une excellente définition |
| <input type="checkbox"/> | 7. Le format PDF, bien que populaire, est à déconseiller puisqu'il s'agit d'un format propriétaire |
| <input type="checkbox"/> | 8. Les formats les plus usuels sont les formats libres |
| <input type="checkbox"/> | 9. De tous les supports de stockage, ce sont les supports informatiques qui ont la plus longue longévité |

Exercice 5

[Solution n°5 p 111]

Associer les besoins de la situation au choix de la qualité de la résolution.

- I - c. Vous optez pour une basse définition et la compression des documents. Vous offrez des documents lisibles, mais dont l'impression est de piètre qualité.
- II - b. Vous optez pour la plus haute résolution possible.
- III - a. Vous visez une portabilité (petite taille) permettant un flux rapide dans les réseaux et la diffusion par Internet et vous optez pour une résolution bitonale.
- IV - d. Vous optez pour la lisibilité tout en acceptant une certaine perte de qualité au niveau de la définition.

- | | | | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|
| 1. Vous mettez en ligne la copie numérique de lettres du XIII ^e siècle que les internautes pourront consulter et faire imprimer pour leurs besoins personnels | 2. Vous mettez en ligne des aperçus des publications de votre organisme. Ils comportent à la fois du texte et des images | 3. Vous numérisez un testament du XVI ^e siècle en très mauvais état | 4. Vous mettez en ligne des documents administratifs textuels en noir et blanc qui vont servir temporairement dans votre organisme |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|

Exercice 6

[Solution n°6 p 111]

Les formats d'enregistrement avec perte de données permettent de créer des fichiers de petite taille qui facilitent le téléchargement. Associer les différents formats aux utilisations que l'on peut en faire.

I - a. JPEG

II - c. PNG

III - d. PDF

IV - b. GIF

1. Documents multipages fait de plusieurs formats

2. Diffusion sur le Web mais sans perte de données

3. Diffusion sur le Web

4. Diffusion d'images

Exercice 7

[Solution n°7 p 112]

Parmi les énoncés suivants, lesquels sont-ils vrais ?

- | | |
|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/> | 1. Une fois la numérisation d'un microforme terminé, on peut le jeter et utiliser la version numérique car elle est plus facilement accessible |
| <input type="checkbox"/> | 2. Afin d'améliorer la visibilité lors de l'opération de numérisation, il est préférable d'installer le poste de travail près d'une fenêtre afin de bénéficier de la lumière du jour |
| <input type="checkbox"/> | 3. Pour le traitement des images et leur enregistrement sur des supports physiques, il est préférable d'opter pour des unités de travail Mac |
| <input type="checkbox"/> | 4. Dans certains cas, il est utile de traiter l'image après sa numérisation |
| <input type="checkbox"/> | 5. La mise en ligne de photographies, documents fort recherchés, nécessite des copies à haute définition |

Exercice 8

[Solution n°8 p 112]

Parmi les énoncés suivants, lesquels sont-ils vrais ? (suite)

- | | |
|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/> | 6. Lors de la numérisation d'un négatif sur plaque de verre, il faut s'assurer que celui-ci soit bien plaqué contre la vitre de l'appareil afin d'éviter un bris |
| <input type="checkbox"/> | 7. Lors de la numérisation de photographies couleur, il y a parfois lieu d'effectuer des retouches |
| <input type="checkbox"/> | 8. La numérisation des gravures se fait prioritairement en mode bitonal |
| <input type="checkbox"/> | 9. Le format TIFF est recommandé pour la sauvegarde des archives |
| <input type="checkbox"/> | 10. Le format TIFF est recommandé pour la mise en ligne d'images, le téléchargement d'images et leur transmission par courriel |

Galleries associées à ce module

Ressources annexes

- Bélinographe Tpe B.E.L.1, (1928) collection F.B.



Solution des Quiz

> Solution n°1 (quiz p. 101)

- a. Les archives informatiques sont uniquement des documents créés originalement au moyen d'équipement informatique
Bravo, cet énoncé est faux. Les documents numérisés (de la correspondance manuscrite ou une vieille photographie, par exemple) sont considérés comme des archives informatiques si le centre d'archivers ne possède pas les originaux. (chapitre 1.1)
- b. Archives numériques et archives électroniques sont des synonymes
Cet énoncé est vrai. Les archives numériques ou électroniques sont disponibles seulement par le biais de l'utilisation d'un ordinateur et sont en cela synonymes, mais le terme "archives électroniques" paraît désormais démodé (voir glossaire).
- c. Les photographies numériques sont des archives électroniques
Cet énoncé est vrai. Le modèle original ayant servi à les générer n'existe plus et les tirages sur papier en sont les copies et non les sources.
- d. Les dessins assistés par ordinateur sont des documents numériques
Cet énoncé est vrai. Les dessins assistés par ordinateur sont créés à l'origine sur un ordinateur et n'ont pas d'originaux sous forme papier. La copie papier est en fait une reproduction ou une impression du document électronique.
- e. Aucune de ces réponses
Voir les autres possibilités.

> Solution n°2 (quiz p. 101)

1. Il détermine la résolution (ou définition) de l'image	c. Le nombre pixels
2. Il s'utilise pour la reproduction des documents et des textes illustrés de gravures	a. Le mode d'encodage bitonal
3. Il détermine le type de numériseur à utiliser	b. Le format

> Solution n°3 (quiz p. 101)

- | | |
|-------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/> | 1. Il est recommandé de numériser le document physique dans la plus petite résolution possible puisqu'une grande résolution exige la création de fichiers énormes, ce qui est plus dispendieux
<i>FAUX. La plus petite résolution (économie d'espace sur le disque dur) est souvent insuffisante pour les besoins archivistiques. (chapitre 2.2)</i> |
| <input type="checkbox"/> | 2. Il existe une pratique uniforme pour la numérisation de tous les types de documents
<i>FAUX. Une pratique uniforme ne répondrait pas aux besoins archivistiques de respecter la valeur intrinsèque du document. (chapitres 2.4.2 et 3)</i> |
| <input checked="" type="checkbox"/> | 3. On peut transformer un document de haute résolution en basse résolution, mais pas l'inverse
<i>VRAI. C'est pourquoi il est préférable de sélectionner la qualité de la numérisation en prévoyant les besoins et les utilisations futures. (chapitre 3.2)</i> |
| <input checked="" type="checkbox"/> | 4. Les formats d'enregistrement sans pertes de données doivent être utilisés pour constituer une banque de fichiers source
<i>VRAI. A partir de ces fichiers maîtres, des copies avec pertes de données, plus petites, peuvent être effectuées afin, par exemple, de mettre les documents en ligne. (chapitre 4)</i> |
| <input type="checkbox"/> | 5. Le choix des formats d'enregistrement doit seulement se faire en fonction du type de documents numérisés et des besoins de l'organisme
<i>FAUX. Si le format choisi est incompatible avec les applications courantes la numérisation est un échec. (chapitre 4.3.1)</i> |

> Solution n°4 (quiz p. 102)

- | | |
|-------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/> | 6. La mise de photographies en ligne requiert que l'on offre aux internautes des documents de la plus haute qualité possible, donc comprenant une excellente définition
<i>FAUX. La mise en ligne sans contrôle des documents de haute qualité n'est pas recommandée, car il est pénalement très risqué de faire des reproductions en faisant fi des droits d'auteur. (chapitre 4.3.1)</i> |
| <input type="checkbox"/> | 7. Le format PDF, bien que populaire, est à déconseiller puisqu'il s'agit d'un format propriétaire
<i>FAUX. Il est vrai que l'assemblage des fichiers requiert l'achat d'un logiciel spécifique. Cependant, le logiciel de lecture (Acrobat Reader) est distribué gratuitement et demeure aisément disponible. (chapitre 4.3.2.2)</i> |
| <input checked="" type="checkbox"/> | 8. Les formats les plus usuels sont les formats libres
<i>VRAI. Ces formats ont l'avantage d'être lus au moyen de toutes les applications. (chapitre 4.4.1)</i> |
| <input type="checkbox"/> | 9. De tous les supports de stockage, ce sont les supports informatiques qui ont la plus longue longévité
<i>FAUX. En fait, la micrographie et la matérialisation de ces informations les mettent à l'abri de l'effacement par erreur, de l'incompatibilité informatique et de la fragilité des supports d'enregistrement informatiques. (chapitre 4.5)</i> |

> Solution n°5 (quiz p. 102)

1. Vous mettez en ligne la copie numérique de lettres du XIII ^e siècle que les internautes pourront consulter et faire imprimer pour leurs besoins personnels	d. Vous optez pour la lisibilité tout en acceptant une certaine perte de qualité au niveau de la définition.
2. Vous mettez en ligne des aperçus des publications de votre organisme. Ils comportent à la fois du texte et des images	c. Vous optez pour une basse définition et la compression des documents. Vous offrez des documents lisibles, mais dont l'impression est de piètre qualité.
3. Vous numérisez un testament du XVI ^e siècle en très mauvais état	b. Vous optez pour la plus haute résolution possible.
4. Vous mettez en ligne des documents administratifs textuels en noir et blanc qui vont servir temporairement dans votre organisme	a. Vous visez une portabilité (petite taille) permettant un flux rapide dans les réseaux et la diffusion par Internet et vous optez pour une résolution bitonale.

Réponse a. C'est un document qui se retrouve sous la catégorie 1. On retrouve dans cette catégorie les documents administratifs utilisés dans les réseaux internes ou pour diffusion sur le Web. (chapitre 3.4)

Réponse b. C'est un document qui se retrouve sous la catégorie 4. Ce sont des documents sévèrement endommagés ou dédiés à être imprimés. (chapitre 3.4)

Réponse c. C'est un document qui se retrouve sous la catégorie 2. C'est souvent le choix fait par les services intranet. Les documents sont de piètre qualité pour l'impression mais tout à fait acceptable pour consultation à l'écran. (chapitre 3.4)

Réponse d. C'est un document qui se retrouve sous la catégorie 3. Les centres d'archives ou les bibliothèques peuvent rendre ainsi des documents disponibles en ligne. (chapitre 3.4)

> Solution n°6 (quiz p. 102)

1. Documents multipages fait de plusieurs formats	d. PDF
2. Diffusion sur le Web mais sans perte de données	c. PNG
3. Diffusion sur le Web	b. GIF
4. Diffusion d'images	a. JPEG

Réponse a. Créé pour permettre la diffusion sur toutes les plates-formes et en ligne des images numériques en couleur. (chapitre 4.2)

Réponse b. Réservé à l'usage sur le Web et n'est pas utile pour les activités de numérisation des archives mais seulement pour leur diffusion sur le Web. (chapitre 4.2)

Réponse c. Moins volumineux que le format GIF et de meilleure qualité que le format PNG ; ce format sera capable bientôt de rentrer en compétition avec le format TIFF utilisé aux archives. (chapitre 4.2)

Réponse d. Le format PDF se prête à plusieurs utilisations, entre autres la compilation avec d'autres fichiers images PDF. (chapitre 4.2)

> Solution n°7 (quiz p. 103)

- 1. Une fois la numérisation d'un microforme terminé, on peut le jeter et utiliser la version numérique car elle est plus facilement accessible
FAUX. En fait, la copie numérisée du microforme reste une copie de travail, le microforme en lui-même est un meilleur médium de conservation. (chapitre 5.7)
- 2. Afin d'améliorer la visibilité lors de l'opération de numérisation, il est préférable d'installer le poste de travail près d'une fenêtre afin de bénéficier de la lumière du jour
FAUX. En fait, peu importe le type de numériseur car la lumière du jour peut causer des reflets ou des taches sur le film. (chapitre 5.8)
- 3. Pour le traitement des images et leur enregistrement sur des supports physiques, il est préférable d'opter pour des unités de travail Mac
VRAI. Les Mac peuvent lire et enregistrer sur tous les supports de tous les systèmes d'exploitation, MAC, PC et UNIX, ce qui permet de les relire universellement. (chapitre 5.8)
- 4. Dans certains cas, il est utile de traiter l'image après sa numérisation
VRAI. Un tel traitement permet notamment d'enlever des bordures noires, ainsi que de redresser l'image et d'améliorer son rendu. (chapitre 5.10)
- 5. La mise en ligne de photographies, documents fort recherchés, nécessite des copies à haute définition
FAUX. En fait, les copies disponibles gratuitement pour les utilisateurs sont des copies à basse résolution. La numérisation en très haute qualité reste dans les serveurs du centre d'archives ou sur des supports de sauvegarde comme le CD-ROM.

> Solution n°8 (quiz p. 103)

- 6. Lors de la numérisation d'un négatif sur plaque de verre, il faut s'assurer que celui-ci soit bien plaqué contre la vitre de l'appareil afin d'éviter un bris
FAUX. Cela pourrait provoquer un effet de ventouse qui risquerait d'entraîner le bris du document au moment de le retirer de la vitre. (chapitre 7.2)
- 7. Lors de la numérisation de photographies couleur, il y a parfois lieu d'effectuer des retouches
VRAI. Les caractéristiques du support entraînent, avec le temps, l'altération des couleurs. Les retouches permettent, dans une certaine mesure, de restituer l'aspect original. (chapitre 7.4.2)
- 8. La numérisation des gravures se fait prioritairement en mode bitonal
VRAI. Ce mode évite les altérations dues au mode en niveaux de gris, comme la transformation des zones franches entre le noir et le blanc en flous gris. (chapitre 8.1 et 8.2)
- 9. Le format TIFF est recommandé pour la sauvegarde des archives
VRAI. Il possède notamment l'avantage de conserver la gamme complète des couleurs ou de niveaux de gris.
- 10. Le format TIFF est recommandé pour la mise en ligne d'images, le téléchargement d'images et leur transmission par courriel
FAUX. Le format TIFF ne permet pas de réduire le nombre de couleurs de pixels ; le fichier demeure volumineux, ce qui ralentit le téléchargement.

Glossaire

Fichier

Ensemble de fiches ou d'enregistrements informatiques contenant des informations de même nature.

Manuscrit

Texte écrit à la main ou, par extension, élaboré par son auteur à l'aide d'une machine à écrire ou d'un ordinateur.

Numérisation

Procédé électronique de reproduction d'un document d'archives sous forme de document électronique.

Type de document

Catégorie de documents distingués en raison de critères communs, matériel (par exemple: aquarelle, dessin) ou fonctionnels (par exemple: journal, livre de compte, main-courante, registre de délibération...).

Bibliographie

[Pour la numérisation]

Conseils à voir dans le chapitre 12 de la deuxième section du présent module.

[Pour le microfilmage]

La conservation entre microfilmage et numérisation. Actes des journées patrimoniales organisées par la BNF tenues à Sablé les 8 et 9 novembre 1993.

[Pour le microfilmage]

LEISINGER, ALBERT H. *La microphotographie aux Archives*. Bruxelles, ICA, 1975 (traduit de l'anglais par Christian GUT).

[Pour le microfilmage]

Recueil de normes françaises. Supports d'informations, micrographie. AFNOR, 5ème édition, 1992.

[Pour le microfilmage]

ROPER, MICHAEL. *Directives pour la préservation des microformes*. Paris, édition française préparée par Georges WEILL, ICA, Etudes n° 2., 1990.

[Pour le microfilmage]

SCOM. *Méthodologie d'emploi de la micrographie*. Paris, ministère de l'Economie et des Finances, direction du Budget, 3ème édition, 1983.

[Pour le microfilmage]

FAVIER, JEAN, NEIRINCK, DANIELÈ. *La pratique archivistique française*. Paris, Archives nationales, 1993, p.507-512.

[Pour le microfilmage]

KORMENDY, LAJOS, KEENE JAMES A., POWELL TED F., WEILL, GEORGES. *Manual of Archival Reprography*. Munich, Londres, New-YORK, Paris, ICA Handbooks Series, volume 5. K.G., SAUR, 1989.