

Module 9 - Section 2 : Numériser les documents

Patrick PERROT

version 2 : 08/11/2014 ©AIAF - PIAF

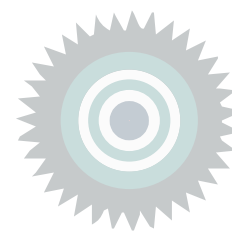
Table des matières

Objectifs	5
Introduction	7
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.	12
1.2.3. L'héritage actuel	12
1.3. La bonne numérisation	12
2. Notions fondamentales en numérisation	14
2.1. Image en pixels	14
2.2. Résolution	15
2.3. Modes colorimétriques	17
2.4. Analyse et typologie des documents à numériser	21
2.4.1. Analyse du document	22
2.4.2. Nature des documents déterminant la colorimétrie	22
3. Objectifs de la numérisation	24
3.1. Objectifs	24
3.2. Usage du fichier final	25
3.3. Portabilité	25
3.4. Répartition des documents numérisés selon l'usage	25
3.5. Usages mixtes	26
4. Formats d'enregistrement des images numérisées	28
4.0 Définitions	28
4.1. Les qualités des formats pour les besoins archivistiques	30
4.2. Formats sans perte de données	32
4.3. Formats avec pertes de données	35
4.4. Choix des formats d'enregistrement selon l'usage	41
4.4.1. Création d'une photothèque numérique	42
4.4.2. Création d'une bibliothèque de documents à pagination multiple	42
4.5. Formats libres et formats propriétaires	46

4.6. Microfilm et numérisation : des champs d'application connexes	47
5. Numériseurs	49
5.1. Format des originaux déterminant le numériseur	49
5.2. Le choix du numériseur	50
5.3. La disparition de certains types de numériseurs	51
5.4. Analyse des caractéristiques des numériseurs	51
5.5. Inventaire des fonctionnalités	52
5.6. Numériseurs de petit format	53
5.7. Numériseurs de grand format	55
5.8. Numériseurs de films et microfilms	59
5.9. La mise en place des éléments du poste de numérisation	67
5.10. Ségrégation des documents	67
5.11. Post-traitement des images	68
5.12. Informations textuelles dans le fichier image	69
5.13. Trucs et astuces	69
5.13.1. Documents plus grands que la vitre d'exposition.	69
5.13.2. Numérisation d'un ouvrage «à la baguette»	70
6. Numérisation des manuscrits	71
6.1. La nature des manuscrits	71
6.2. Choix d'un procédé physique pour les reliés	72
6.3. Choix d'un procédé physique pour les feuilles volantes	72
6.4. Choix du mode colorimétrique	73
6.4.1. Le mode bitonal	73
6.4.2. Le mode niveaux de gris	74
6.4.3. Le mode couleurs RVB	75
6.5. La numérisation de microformes	75
6.6. Importance de la taille des fichiers	76
7. Numérisation des photographies	79
7.1. Première approche en numérisation de photographies	79
7.2. Plaques de verre	79
7.3. Diapositives et négatifs de petit format	80
7.4. Photographies sur papier	82
7.4.1. Photographies monochromes	82
7.4.2. Photographies en couleur sur papier	83

8. Numérisation des gravures	84
8.1. Numérisation au bon mode	84
8.2. Cas spéciaux	87
9. Numérisation des imprimés	88
9.1. Contenu des imprimés, procédés d'obtention	88
9.2. Principe de la similigravure	88
9.3. Principe de l'héliogravure	90
9.4. Procédés de numérisation selon les objectifs	91
10. Numérisation des plans	94
10.1. Plans tracés	94
10.2. Tirages de plans	96
11. La post-production	98
La post-production	98
99. Solutions de numérisation en tableaux	99
11.1. Inventaire des typologies de documents	99
11.1.1. Documents figurés sans tramage photomécanique	99
11.1.2. Plans et dessins techniques	100
11.1.3. Photographies, plaques, films	100
11.1.4. Documents figurés en similigravure	100
11.2. Procédés de numérisation à partir de la typologie des documents	101
11.2.1. Manuscrits, dessins, gravures, lithos	101
11.2.2. Peintures et petits objets plats 11.2.3.....	102
Photographies, négatifs, diapositifs 11.2.4.	103
Plans d'architecture et techniques	104
11.3. Solutions de numérisation à partir de l'usage du fichier	105
12. Bibliographie en ligne	107
Numérisation	107
Bibliothèques en ligne	108
13. Evaluation des connaissances	109
Galleries associées à ce module	112
Ressources annexes	113

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. La numérisation est à présent prépondérante dans ce processus, prenant le pas sur la consultation des microfiches in situ. Un document numérisé est disponible au domicile du chercheur, quel que soit son éloignement du centre d'archives.

Le microfilmage reste le seul moyen pérenne de conservation des documents, mais il est à présent considéré comme un moyen de sauvegarde et non plus de diffusion

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 facteur 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
- introduire des métadonnées dans l'image pour déclencher le fonctionnement des moteurs de recherche à partir de l'objet.

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 fonds iconographiques

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.

Introduction



1. Définitions

Avant d'entrer dans le 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. Elles font partie des archives électroniques

Archives électroniques

Ensemble des archives nécessitant un moyen électronique de reproduction. Ces archives englobent les enregistrements par gravure en sillon, magnétisation, déformation permanente ou temporaire d'un support par un rayon laser. Les archives électroniques peuvent être analogiques (disques vinyl, cassettes audio...) ou numériques (CD-ROM, disquette...)

Archives numériques

Ensemble des archives numérisées (images des archives physiques) et des archives informatiques, conservées sur un support numérique. 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 obtenus par l'intermédiaire d'un dispositif appelé numériseur (ou scanner pour les anglophones), qui traduit des signaux analogiques (la lumière provenant de l'original et reconnue par des capteurs) en signaux numériques 0 et 1.

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.

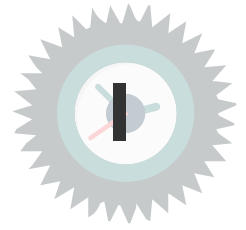
Bien sûr si quelques clips vidéo de quelques secondes ou quelques prises de vues documentaires se retrouvent dans un fonds d'archives raisonné, on les y laissera, sans faire de distinguo, pour ne pas attenter à la cohésion du fonds.

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



introduction binaire

1. Introduction à la numérisation



Introduction



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, dessinée, 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...

1.1. Définitions



Définitions

Avant d'entrer dans le 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 (images 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 signaux analogiques (la lumière provenant de l'original) en signaux numériques.

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 raisonné, on les y laissera, sans faire de distinguo, pour ne pas attenter à la cohésion du fonds.

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

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

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 19^e 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.

1.2.1. La transmission des documents écrits ou dessinés

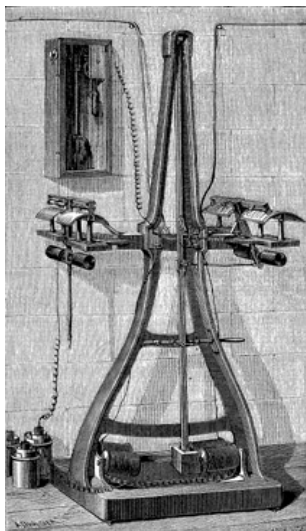
Le premier appareil capable de transmettre des images en deux valeurs, le noir et le blanc, est le télégraphe autographique (1862) inventé par l'abbé Caselli et perfectionné par Meyer. On l'appelle aujourd'hui télécopieur (fax), il servait à transmettre des textes et images que l'on traçait sur un papier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



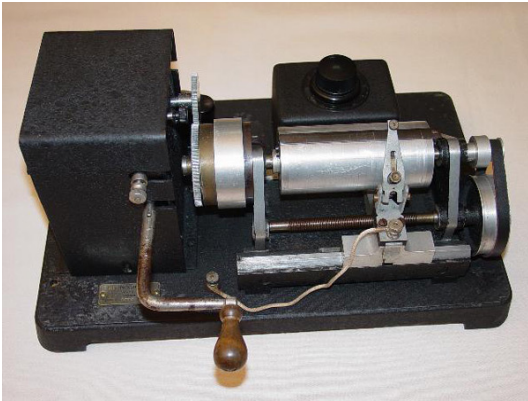
*Pantélégraphe
autographique de Caselli*

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.

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.



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

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 20e pour voir sa démocratisation!

(cf. Agrandir l'image) ^(cf. p.113)

2.3. 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 "discretisation" 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 terme 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.

1.3. La bonne numérisation



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



Introduction



Notions fondamentales en numérisation

Un certain nombre de connaissances sont indispensables pour comprendre les mécanismes de numérisation de l'image. La connaissance de ces principes 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 l'archiviste.

2.1. Image en pixels



Image en pixels

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.

Les limites de l'image sont invariablement celles d'un carré ou d'un rectangle, permettant le placement de l'image dans des compositions numériques (pages de magazines, de livres, pages web, etc.)

Cependant, l'image peut apparaître avec des contours non géométriques. Cela s'appelle le détourage.



Image 1 Image agrandie des pixels noirs et blancs, niveaux de gris et couleur.

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".



Image 2 Personnage détourné, et converti au format Gif pour conserver cette propriété dans un navigateur web

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:



Image 3 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 conservation et la restitution à l'échelle originale. Les mêmes résolutions s'appliquent aux documents en niveaux de gris.

Résolution des images bitonales (au trait)



Image 2 Aspects des résolutions 150, 300, 600 dpi pour un document bitonal (agrandissement 4x).

L'image bitonale de par sa nature brutale (noir ou blanc) va mettre en valeur les contours des contrastes et en basse résolution va produire des effets très crénelés nuisibles à la reconnaissance des caractères et désagréables pour le lecteur. On pratiquera donc la bitonalisation dans des résolutions plus élevées que pour les tons continus.

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 à 200dpi pour diffusion en ligne et documents commerciaux, 300 dpi pour l'impression domestique et 600dpi pour l'impression laser des dessins, des gravures... L'imprimerie utilise systématiquement des résolutions plus élevées: 1200 ou 2400dpi, pour s'aligner sur la finesse des presses offset.

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. On verra qu'il y a une exception : les négatifs photographiques qui devront être reproduits avec un agrandissement.

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 aux plus exigeant des usages que l'on en fera (par exemple celle de l'impression).

Par ailleurs, pour les fonds d'archives publiques, il faudra se conformer aux cahiers des charges standardisés, comme ceux préconisés par la Direction des Archives de France. Des liens contenus à la fin de ce cours au chapitre « Solutions de numérisation en tableaux ».

Génération de fichiers en résolution réduite

Pour la diffusion vers des réseaux publics et supports éducatifs, il suffira ensuite de générer des fichiers en plus basse résolution en utilisant les commandes de ré-échantillonnage des logiciels de traitement d'image.

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».

Il existe des moyens de calcul des résolutions optimales en fonction de la finesse du trait pour la reproduction des gravures sur bois (estampes).

Résolutions et droits



La diffusion en ligne de fichiers image à haute résolution pose des problèmes de droits d'auteur des oeuvres. Si la résolution et la qualité sont élevées, des éditeurs peu scrupuleux pourraient utiliser le document pour l'éditer avec des profits commerciaux, et ceci complètement à l'insu du centre d'archives. C'est pourquoi pour les iconographies modernes, comme les cartes postales, les affiches (surtout les affiches protégées par l'ADAGP), les photographies de personnalités...il faudra choisir des résolutions basses et surtout un nombre de pixels limité à 1000 sur le plus grand côté de l'image pour prévenir de ce piratage.

Cependant, l'archiviste n'appliquera pas forcément cette méthode aux documents très anciens appartenant à des collections publiques, qui ne présentent pas forcément autant d'enjeux médiatiques.

2.3. Modes colorimétriques

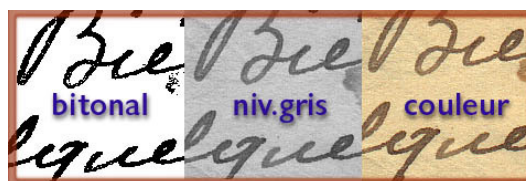


Modes colorimétriques

Pour la monochromie 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 essentiellement pour la reproduction des archives photographiques dites en « noir et blanc »,
3. Le mode RVB (rouge vert bleu) pour la reproduction de tous documents dont on veut restituer la couleur

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



De gauche à droite un mode bitonal, un mode niveau de gris, un mode couleur appliqués à un même document.

La notion de couches

L'image bitonale et l'image en niveaux de gris sont formées d'une seule couche, c'est à dire une seule série d'informations disposées en cartographie sur un plan pour créer une série de pixels juxtaposés. Ces pixels forment une image

Dans l'image bitonale comme dans l'image en niveaux de gris, la présence d'une seule série de nuances permet de leur donner une couleur de référence. Par défaut, le fichier bitonal et le fichier en niveaux de gris sont des variations du noir.

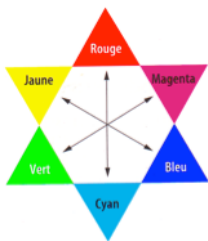
Dans les fichiers en couleurs, il a fallu faire la décomposition de l'arc-en ciel, mais en se limitant strictement aux couleurs de base rouge-vert-bleu. Ce système est appelé additif, parce que le mélange avec 100% pour chacune des couleurs donne le blanc.

S'il y a trois couleurs, il y aura forcément trois couches, qui seront le rouge, le vert, le bleu. (RVB en Français, RGB en anglais). Chacune des couches sera nuancée en un certain nombre de niveaux, et c'est la combinaison de ces niveaux qui va donner les couleurs intermédiaires (le jaune, le vert, les pastels...) Le nombre de couleurs intermédiaires que l'on peut obtenir est facteur de la profondeur d'analyse allouée lors de la discrétisation du signal analogique.

Un autre système d'encodage de couleurs existe qui ne fait pas appel à une notion de couches, mais à un modèle mathématique élaboré : le LAB (Luminosité, couleur A et couleur B.) Ce modèle offre le spectre colorimétrique le plus étendu, mais il n'est pas usité comme mode colorimétrique en archives, son usage est plutôt scientifique, médical ou astronomique.

La notion de couleur

Les couleurs perceptibles sont celles issues du spectre de la lumière solaire, le fameux « arc en ciel » que l'on peut caractériser par trois couleurs de base, le rouge, le vert, le bleu. L'addition de ces trois couleurs en quantité égale donne le blanc pur. En faisant varier la proportion entre ces couleurs, on obtiendra une infinité de nuances, qui font toute la subtilité des couleurs de la vie. Ce mélange est appelé mélange additif. Plus on ajoute des couleurs, plus on va vers le blanc. Dans le cas de l'informatique, ce sera la lumière incidente provenant de l'écran.



*roue
chromatique*

Lorsque l'on veut reproduire sur un support non lumineux (le papier blanc) une illusion de l'arc-en-ciel, il faudra faire appel à un mélange soustractif, c'est à dire que plus on ajoute des couleurs, plus on va vers le noir. On utilise pour cela des couleurs complémentaires. Ces couleurs complémentaires sont le Cyan, le magenta, le jaune. Ce sont les tubes de couleurs que tout peintre a dans sa trousse. Il a aussi un tube de noir pour renforcer les ombres, et un tube de blanc pour obtenir les nuances pastel. Dans le monde de l'imprimerie, on utilisera les encres Cyan magenta Jaune et du noir pour renforcer les contrastes, à cause de la transparence des encres.

La roue chromatique ci-dessous montre ces couleurs de base et les couleurs complémentaires. Alors que le rouge est une couleur de base, pour obtenir du rouge avec des couleurs complémentaires, il faudra mélanger du magenta et du jaune en proportions égales, du cyan et du jaune pour faire le vert, du cyan et du magenta pour faire le bleu.

Les spectres de restitution des couleurs

Comme on vient de voir précédemment, il y a deux principes de restitution des couleurs : par les couleurs de base, ou par les couleurs complémentaires. Dans l'image numérisée, ces deux familles de modes existent, mais ont des usages dédiés :

- par les couleurs de base, on fera tout ce qui est l'acquisition, le post-traitement, et le stockage de l'image,
- Par les couleurs complémentaires, on ne fera que la reproduction de ces images sur le papier, selon le modèle CMJN (Cyan Magenta Jaune Noir).

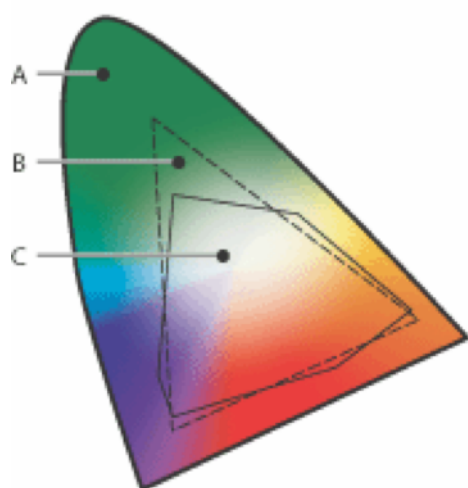
Ce dernier mode en « quadrichromie » sera donc ignoré des archivistes, parce qu'on va le voir, son spectre colorimétrique est restrictif, il engendre une perte de données, et prend plus de place en enregistrement à cause des 4 couches au lieu de 3.

La figure ci-dessous montre les spectres de couleurs reproductibles par les modèles normalisés qui sont utilisés par les applications informatiques.

Le spectre le plus étendu identifié par A est celui du modèle LAB, (luminosité, couleurs A et B). Ce modèle est utilisé essentiellement en applications scientifiques, les numériseurs d'archives ne donnent pas de fichiers discrétisés dans ce mode. C'est aussi l'espace de travail de l'application Photoshop®, pour reculer les limites d'altération lors des transformations colorimétriques de tous les formats de fichiers.

Le spectre le plus répandu identifié par B est celui du modèle RVB (Rouge vert bleu). Il couvre la plupart des couleurs perceptibles par l'œil humain. C'est un spectre suffisant pour l'archivistique, bien que parfois certaines couleurs de papier ou de peintures très vives se trouvent à la frange de ce spectre.

Le spectre de l'imprimeur identifié par C est celui du modèle CMJN (Cyan Magenta Jaune Noir). Il apparaît nettement réduit par rapport au RVB, ceci apparaît nettement lors de la transformation d'une photo numérique contenant un ciel d'un bleu soutenu : il s'affadit brusquement. Le modèle CMJN est bridé de fait par la nature des colorants des encres d'imprimerie.



Les spectres LAB zone A, RVB zone B, CMJN zone C

Profils colorimétriques des modes RVB



Il n'existe pas qu'une sorte de RVB, différents profils existent sur le « marché » Ils ont été développés soit par des scientifiques, soit par des développeurs d'applications. Des traducteurs assurent la conversion de profils à l'entrée dans l'application. Hormis en reproduction des photothèques en couleurs, ces profils ont peu d'influence visible sur les documents anciens écrits comme les registres, les livres, les manuscrits.

Diverses recommandations existent dans les ministères chargés de la normalisation des formats d'archivages. Le profil RVB plus rencontré est le sRGB IEC6 1966 –2.1 pour la diffusion, mais la direction des archives de France recommande aussi le profil Adobe RGB 1998 pour les fichiers de conservation, parce que son spectre inclut les profondeurs d'échantillonnage plus larges.

Écrire un cahier des charges de numérisation¹

La profondeur d'analyse des couleurs ou échantillonnage

La profondeur d'analyse de la couleur, parfois appelée « échantillonnage » va permettre d'obtenir des nuances de plus en plus subtiles plus on augmentera le nombre d'informations de descriptions de la nuance sur la couche.

La profondeur d'analyse s'exprime en « bits », autrement dit le nombre d'informations 1 ou 0 qui définissent la couleur d'un pixel dans une couche.

L'échantillonnage le plus simple est celui du mode bitonal : 1 bit, autrement dit une information 0 pour l'absence de couleur et 1 pour la présence de couleur. (blanc=0, noir=1)

L'échantillonnage standard du niveau de gris se fait avec 8 bits, résultant en 256 niveaux de gris. Cela couvre les possibilités de perception de l'oeil humain dans cette gamme.

Pour faire de la couleur, on affecte 8 bits à chaque couche R, V, B, ce qui en les combinant donne environ 16 millions de couleurs. L'œil humain peut en distinguer 60 000 environ

Mais il existe des échantillonnages plus poussés, de 14, 16, 32 bits par couche, ce qui donne 65536 nuances de gris en 16 bits ou des milliards de couleurs en RVB.

A quoi cela sert-il ? Dans le cadre de l'archivistique, l'acquisition des négatifs photo en niveaux de gris utilise le 16 bits pour pouvoir restituer une gamme complète de niveaux en récupérant toutes les nuances subtiles des images à bas contraste.

Pour la couleur, un échantillonnage de 14 à 16 bits trouve son usage pour le « repêchage » des diapositives décolorées par le temps, cela permet de distinguer les nuances affaiblies par la dégradation des colorants.

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 dont la combinaison reconstituera la couleur de l'original.

1. <http://www.archivesdefrance.culture.gouv.fr/static/1309>

L'on verra par la pratique que la restitution des couleurs n'est pas si simple qu'il paraît : l'original est éclairé avec une lumière qui n'est pas parfaite, les capteurs ont des déficiences colorimétriques... le pilote du numériseur va introduire des corrections de gamma, de densité, et élaborer de façon automatique ou fixée par l'opérateur des corrections colorimétriques et densitométriques. C'est bien de ce pilote que dépendra la qualité finale du document.

Les modes colorimétriques spécifiques



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 **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 variantes 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'encrage 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'encrage permet d'améliorer le rendu des détails. Mais ce n'est pas un artifice utilisé par les archivistes, il intervient parfois lors de l'édition d'ouvrages patrimoniaux.



Une image originale en niveaux de gris et son traitement en bichromie

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'encrage 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)

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.

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.

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, cartes postales en phototypie
- lithographies et chromolithographies, affiches
- photographies, cartes postales photographiques
- 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 ou chromolithographies.

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).

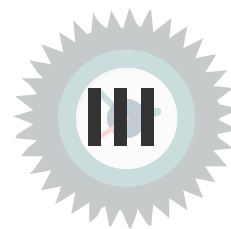
On s'apercevra par la pratique que les fonds iconographiques (photos, négatifs, diapositives, plaques de verre, peintures, chromolithographies, héliographies, cartes postales, affiches, plans en couleurs,...) sont aussi relativement aisés à numériser, car ils font appel à des modes en demi-tons, faciles à retoucher et à adapter à l'usage à l'aide des logiciels de retouche photo. Les difficultés apparaissent dès lors que l'on a affaire à des imprimés. Les pages des ouvrages et catalogues contiennent souvent des typologies mélangées (texte, gravures au trait, photos en couleur ou niveau de gris); l'idéal consisterait à numériser chaque page dans le mode qui lui convient le mieux, puis d'assembler ces sous-fichiers en un seul pour reconstituer le document intégral. Mais ceci est fastidieux pour la reproduction d'une reproduction ! Les progrès des numériseurs, et surtout des moyens de transmission des fichiers lourds ont permis de procéder systématiquement dans le mode le plus favorable à la reproduction des images. Les artifices de compression et détramage font le reste pour obtenir un fichier portable et agréable à lire à l'écran

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. Mais attention, les beaux jours sont comptés, avec l'apparition du dessin assisté par ordinateur, l'utilité de numériser un plan disparaît et les numériseurs à plans disparaîtront aussi du marché dans un délai plus ou moins court...

3. Objectifs de la numérisation



Introduction

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.

3.1. Objectifs

Les objectifs de la numérisation peuvent être:

1. un moyen de sauvegarde d'originaux menacés de dégradation physique (manipulations d'originaux fragiles et 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) ou si l'on choisit de reproduire des documents comportant des couleurs en mode bitonal.

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-après (voir 3.4),
- la haute résolution et une grande fidélité des couleurs pour l'objectif de la catégorie 4.

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 !

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.



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).

3.4. Répartition des documents numérisés selon l'usage

Dans ce qui suit, nous parlons de l'usage des fichiers, transformés ou non par un travail de post-production, c'est à dire tout ce qui est relatif à la préparation des fichiers maîtres et de leurs copies pour la distribution ou l'usage final. (voir chapitre 5-10)

Les catégories ci-après ne sont pas des standards internationaux, mais des familles rencontrées dans la pratique et nommées pour la compréhension de ce cours

Selon nos besoins et l'usage que nous voulons en faire, nous pouvons répartir les documents à numériser en différentes catégories, sachant que certains fichiers maîtres peuvent servir à générer des documents 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**. Idéalement, ces documents ont fait l'objet d'une indexation, c'est à dire une reconnaissance des caractères.

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. C'est par essence, de la documentation dans son acception la plus générale.

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 reproduction des négatifs et diapositives devrait permettre au moins l'impression sur papier en 18x24 cm à 300dpi.

La Direction des Archives de France recommande la conservation dans une taille qui permet la restitution dans une taille 30x40 cm en 300dpi. La performance des systèmes de stockage actuels permet cette taille.

Catégorie W : documents iconographiques pour la toile (Web)

Cette catégorie englobe les images à basse résolution que l'on rend disponibles en ligne pour les internautes, l'enseignement etc...et les images utilisées comme icônes d'accès aux fichiers haute résolution d'une photothèque 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 une page internet. La résolution est basse et l'on accepte des dégradations importantes par la compression. Ces fichiers sont constitués à partir d'une copie en résolution réduite réalisée à partir du fichier maître

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 générera des fichiers de moindre qualité pour les diffusions en ligne, par l'intermédiaire d'un logiciel de traitement des images. Cette opération couramment le sous-échantillonnage de l'image.



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



Introduction

Les formats d'enregistrement et les compressions choisies 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 (avec perte de données).

Les pertes de données par compression résultent uniquement en une réduction du nombre de couleurs du fichier, mais jamais du nombre de pixels.

Les fichiers destinée à l'archivage permanent

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. Il faut la conserver parce que reprendre un original est plus coûteux que le prix de l'espace disque.

Lorsque l'on acquiert une image il faut tout d'abord enregistrer le fichier d'acquisition dans un format n'entraînant pas de pertes de données. Ce fichier sera ensuite travaillé dans une phase dite de post-production comprenant le nommage, les éventuelles retouches colorimétrique et l'entrée des métadonnées. Le résultat sera de nouveau enregistré sans pertes de données pour donner le fichier maître destiné à l'archivage permanent.

Les copies dites de diffusion, ou d'exploitation

Du fichier maître, des **copies** de taille réduite seront générées de façon mécanique, pour le placement en ligne, afin de limiter les temps de téléchargement. Les métadonnées sont migrées elles aussi automatiquement depuis le fichier maître.

On pourra par la suite renouveler les copies pour la mise en ligne, en augmentant leur qualité au fur et à mesure de l'apparition de réseaux à haut débit.

A propos de ce chapitre des formats

Ce chapitre peut sembler long et indigeste pour l'apprenant. C'est pourtant toute la réussite d'un plan de numérisation qui en dépend. Face à l'archiviste, des commerciaux peu scrupuleux peuvent proposer des prestations au résultat dégradant, et au final l'argent est dépensé pour un résultat médiocre. Revenez sur ce chapitre par la suite si dans un premier temps il vous paraît trop ardu, mais faites d'abord un petit passage dans le chapitre qui suit, avec ses définitions.

4.0 Définitions

Dans les pages qui vont suivre, et déjà au début de ce chapitre, des termes nouveaux apparaissent et pourraient mettre le lecteur en difficulté. Mettons donc à plat ces termes pour bien comprendre le contenu des différents formats d'enregistrement.

Le format d'enregistrement

Le format d'enregistrement est l'ensemble des caractéristiques du fichier qui lui permettent d'être reconnu par les applications. La manière d'écrire le début et la fin de fichier selon certains codes déclenche le fonctionnement adéquat de l'application qui lit ou transforme ce fichier. Il existe des formats « propriétaires » dont les codes sont typiques à un éditeur, et des formats « universels » ou « libres » dont les codes sont divulgués publiquement. Nous les choisirons pour cette raison.

Le fichier maître

C'est le premier fichier enregistré sans perte de données et qui contient tout ce qui est nécessaire à la conservation de l'archive numérique et à la génération des sous-fichiers de diffusion: image, corrections et métadonnées.

L'image est l'objet informatique résultant de l'acquisition brute par le numériseur.

Les corrections sont des réglages de densité de l'image, de colorimétrie par exemple dans le cas de la restauration de diapos aux couleurs dégradées par le temps. Les corrections sont temporaires pendant la durée de post-production, puis appliquées de façon permanente après contrôle par un expert.

Les métadonnées sont les textes au format XML introduits dans le fichier image. Elles n'apparaissent pas sur l'image, mais dans les éditeurs de métadonnées ou les légendes automatiques de certains environnements. Dans le cas d'une image restaurée, ces métadonnées indiquent la fait de la restauration.

Le fichier de diffusion ou d'exploitation

La caractéristique principale de ce fichier est la portabilité (taille réduite pour diffusion en réseau). L'on cherchera toujours le format le plus efficace pour cette fonction. Le PDF (portable document format) est très avantageux, il permet d'assembler plusieurs pages d'images avec un descriptif ne nécessitant pas d'application spécifique à la lecture des métadonnées

Le calque

Le calque est une image ou une correction se superposant à d'autres calques dans un même fichier image. Lorsque l'on veut obtenir une image « plate » définitive, tous les calques sont fusionnés (confondus en un seul). Certains formats enregistrent les calques, d'autres pas.

Calque de réglage

Un logiciel de retouche photo peut générer des calques de réglage d'image pour changer les caractéristiques de densité, de contraste, de colorimétrie... de l'image placée en-dessous. Un fichier enregistré avec ces calques permet le changement de ces valeurs à l'infini, sans altérer l'image originale qui est la à base du fichier.

La couche Alpha

La couche alpha est une série de pixels transparents juxtaposés à des pixels colorés d'opacité 100%. Dans le format GIF, cela permet de détacher des objets sur un fond coloré dans une page internet. La limite est brutale, et le contour présente un aspect crénelé.

La transparence

La transparence permet d'enregistrer des parties d'images dont l'opacité n'est pas 100%. Il y a peu d'application de la transparence en archives, mais lors du détournage d'une carte postale par exemple, cela permet de ne pas avoir des bords crénelés comme pour la couche alpha, le bord est progressivement transparent sur 2 ou 3 pixels.

La compression

C'est l'action qui consiste à réduire la taille du fichier enregistré par un codage spécifique à chaque algorithme de compression. La compression peut se faire avec ou sans pertes de données. Une compression sans pertes de données réduit la taille de 20 à 30% et jusqu'à 50% pour la plus performante. La compression avec perte de données peut réduire la taille de 95%.

La perte de données

Un terme un peu choquant prend place dans notre propos : la compression avec perte de données. Ce procédé consiste à réduire l'échantillonnage des couleurs, c'est à dire le nombre de couleurs contenues dans l'image, mais pas le nombre de pixels. Il en résulte une taille réduite du fichier enregistré et des artefacts visibles lors d'une compression importante. C'est un procédé lié à un format développé par le Joint Photographs Expert Group (JPEG). Les métadonnées ne sont pas perdues lors de la compression du fichier.



Attention

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

Les métadonnées

Les métadonnées sont des textes qui sont insérés au fichier image, mais qui n'apparaissent pas visuellement sur celle-ci. Les métadonnées sont enregistrées au format d'échange XML et prennent les aspects que le lecteur de métadonnées leur donne. Les métadonnées sont distribuées dans des champs reconnus par des normes internationales dont le standard le plus connu est IPTC, créé à l'origine pour les photographes de presse internationale. Les appareils de numérisation complètent automatiquement les champs de métadonnées EXIF qui décrivent les conditions de prises de vues

La production

C'est la phase où l'on capture l'image de l'original et qu'on l'enregistre une première fois en vue de la suite des opérations.

La post-production

Cette phase comprend toutes les opérations de traitement d'images et d'insertion de métadonnées. A l'issue de ces opérations, le **fichier maître** est constitué avec un enregistrement sans pertes de données.

Le stockage

Le fichier maître constituera une **archive permanente**. Mais au cours de sa vie, les formats et supports physiques vont évoluer. Pour pallier leur obsolescence, il faudra régénérer les fichiers en les copiant sur de nouveaux supports (Disques optiques, serveurs sécurisés, disques durs, bandes magnétiques...)

4.1. Les qualités des formats pour les besoins archivistiques

Il existe une grande variété de formats informatiques pour l'enregistrement des images numérisées. Le but de l'archiviste est de choisir un format :

- pérenne
- libre, ouvert, et décrit par une norme
- universel, lisible sur toutes les plateformes
- pouvant être lu par des logiciels anciens en fichier de diffusion (compatibilité descendante des copies)

- pouvant contenir des métadonnées

A ces préceptes, on pourrait presque ajouter : le format de diffusion devrait être différent de celui du fichier maître d'archivage, permettant de le reconnaître d'un seul coup d'œil dans les répertoires.

Le format pérenne

C'est presque une hérésie de parler de format pérenne dans un monde changeant comme celui de l'informatique. Cependant, il existe des « stabilisateurs » de format, qui sont le nombre des usagers ayant utilisé tel ou tel format sur des quantités importantes de fichiers.

Au nombre de ceux-ci, on reconnaît les formats Tiff, Jfif (improprement appelé jpeg), PDF et png

Le format du domaine public

Le format choisi ne doit pas être propriétaire, c'est à dire que ses clés de développement doivent être publiques et non soumises à droits d'exploitation limités. C'est le cas des formats dont nous allons parler.

Le format universel

Les formats cités dans ce cours ont tous « élu domicile » sur les plateformes Windows-Mac-Unix-linux, et sont lisibles parce que leurs formats ont peu changé avec le temps. Dans le cas du PDF, un format spécial « archives » a été créé et répond à la norme ISO 19005. Cette démarche vise à garantir une lecture dans le futur par une stabilisation du standard, qui cependant en est déjà à sa troisième version depuis sa création...

La compatibilité descendante

Les formats choisis doivent aussi pouvoir être lus par des logiciels datant de quelques années, ou des logiciels grégaires comme les lecteurs d'images intégrés à Windows. Cette compatibilité sert aussi à faire fonctionner des banques de données propriétaires ou libres qui se basent sur des standards anciens pour éviter le paiement de brevets.

Le contenu de métadonnées

Les formats devront pouvoir contenir des métadonnées au langage XML pour que lorsque l'image est détachée de son site de diffusion, l'utilisateur puisse toujours l'identifier complètement sans risque d'intervention lors de la copie (fastidieuse) de légendes en ligne. Les métadonnées sont une part essentielle du contenu du fichier maître. Le choix des champs complétés doit être judicieux pour conserver une compatibilité descendante.

Particularités des formats « photo »



L'avantage des formats d'images courants est qu'ils sont destinés à l'échange, donc en principe stables, compatibles et ouverts.

Une particularité existe cependant pour le format « Adobe Photoshop » qui est un format propriétaire de la société Adobe®. Ce format mondialement reconnu offre une option de compatibilité descendante à l'enregistrement dans le logiciel Adobe Photoshop® (maximiser la compatibilité). Il est par ailleurs reconnu par nombre de logiciels libres ou d'autres éditeurs à cause de son omniprésence dans le milieu de l'image. Il reste cependant la propriété de son créateur qui le fait évoluer en permanence.

On notera aussi une exception avec les nouveaux formats « raw » des appareils photo et chambres numériques, qui peuvent nécessiter un filtre de rendu propriétaire à chaque constructeur, livré avec l'appareil de prise de vues. Le format « raw » est un peu comme les anciens négatifs argentiques, il est nécessaire d'ajuster les valeurs lors du « tirage » de la photo. Les clichés réalisés dans ces formats devront donc être convertis en formats universels pour l'archivage permanent.

4.2. Formats sans perte de données

2.1 Aperçu des formats sans pertes de données

Les formats sans pertes de données sont à **utiliser pour constituer la banque de fichiers maîtres destinés à la conservation et/ou à la sauvegarde.**

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.

Les formats sans pertes de données peuvent être utilisés sans activer leurs moteurs de compression, mais c'est stupide car cela amène à une occupation importante d'espace disque et ralentit les échanges avec le serveur distant. Les formats modernes ont une activation automatique de la compression (cas du png).

La caractéristique typique de ces formats est bien celle des algorithmes de compression. Nous allons en faire le tour avant d'entrer dans les formats, car dans certains cas ils sont communs à plusieurs d'entre eux.

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.

2.2 Les algorithmes de compression

Il existe plusieurs algorithmes de compression qui peuvent coexister au sein d'un même format, dans ce cas, le choix du procédé sera à faire par l'utilisateur lors de l'enregistrement. De même, on retrouve les mêmes « compresseurs » dans plusieurs formats.

LZW Lemple-Zif-Welch du nom de ses inventeurs est utilisé dans les formats TIFF, PDF, GIF et est performant essentiellement lorsque les images contiennent de grandes zones monochromes. Mais comble de l'ironie, il augmente la taille du fichier lorsque l'échantillonnage est supérieur à 8 bits par couche, parce qu'il n'est pas prévu pour cet usage. Il appartient maintenant au domaine public.

ZIP est utilisé par les formats de fichier PDF et TIFF. Comme la technique LZW, la compression ZIP est plus efficace pour des images présentant de grandes zones monochromes, mais il est efficace sur les fichiers à plus large échantillonnage de nuances (16 et 32 bits).

CCITT groupe 4 (Comité Consultatif International Télégraphique et Téléphonique) est utilisé pour les images bitonales, pris en charge par les formats de fichier PDF et de langage PostScript.

Deflate-inflate est un algorithme inventé par Phil Katz pour le format de compression PKZIP en y associant le codage Huffman. Il est utilisé librement dans le format PNG, lui aussi développé par Phil Katz et dans le format Zip actuel (PKZIP pour Phil Katz ZIP). Les codes sources sont accessibles.

2.3 Les formats d'enregistrement sans pertes de données

Le format TIFF

Le format TIFF (*Tagged Image File Format*) est plus particulièrement utilisé pour les documents mono-pages, les photothèques. C'est le plus connu et répandu avec son extension .tif.

Il s'avérera aussi idéal lors des phases de post-production pour l'enregistrement transitoire avant archivage définitif, par son aptitude à conserver les calques image et calques de réglage, avec leurs gradations. C'est le format utilisé en majorité par le monde de l'édition imprimée.

Une option multi-pages du format TIFF existe et est très usitée en Allemagne, mais ne peut pas être lue par tous les logiciels, aussi faudra-t-il en faire un usage mesuré.

PNG

Le format PNG (Portable Network Graphics) -prononcer peng- est un format libre et qui propose la plus grande universalité et un des meilleurs potentiels en termes de gain d'espace disque grâce à sa compression automatique « deflate ». Il reproduit les transparences de l'image et la transparence totale par couche alpha (pixels transparents), mais n'enregistre pas plusieurs calques. Il peut contenir des métadonnées qui ne sont pas toujours affichées par les éditeurs de logiciels propriétaires. Il n'est pas encore très répandu dans le monde des archives, mais des centres précurseurs l'introduisent dans leurs photothèques. Il est reconnu par un très grand nombre d'applications sur toutes les plateformes, et les systèmes libres (Linux).

JPF, JP2

Le format Jpeg 2000 étendu JPF est une version plus complète que l'ancien JP2 auquel il emprunte le moteur de compression avec ou sans perte de données, au choix de l'utilisateur. Le JPF montre les performances les plus intéressantes en matière d'épargne d'espace disque avec son option « sans pertes de donnée » activée. Le principal inconvénient est que son usage n'est pas encore très répandu bien que sa pérennité soit assurée, vu qu'il est normalisé. Le .jpf est encore illisible dans de nombreuses applications, vu que c'est un format récent qui n'est pas encore inclus par défaut aux filtres import-export des applications. Il est cependant très attractif pour l'archivage permanent des fichiers maîtres aplatis (en un seul calque) avec une taille pouvant atteindre la moitié de celle du png ou du Tiff Zip.

Ce format est étudié plus en détail dans la section « formats avec pertes de données » parce que c'était sa destination initiale.

PDF

Le format PDF (Portable Document Format) est plus particulièrement pertinent pour les documents multi-pages et les documents informatiques générés par les applications. L'enregistrement sans perte de données est à choisir par l'utilisateur dans les préférences avant de commencer les enregistrements. Avec le format PDF-A (PDF Archives) il offre une option de pérennité certaine. Il permet aussi d'enregistrer des pages textes et des pages images dans le même fichier.

Nota :Le PDF peut aussi être utilisé avec une compression engendrant des pertes de données, ce qui est intéressant pour la mise en ligne. Ce format est étudié plus en détail dans la section « formats avec pertes de données » parce que c'est sa destination première.

PSD

Le format PSD de Adobe Photoshop n'est pas un format d'archivage à proprement parler, mais on peut l'utiliser dans les copies de travail intermédiaires de post-production, lorsque des calques de réglage sont utilisés temporairement pour le rendu de l'image, ou pour d'autres opérations demandant des fonctions avancées (détourage, redressements, retouches non irréversibles...) Les formats photoshop les plus récents sont enregistrés avec une compression sans perte de données.

GIF

Le format GIF (Graphic Interchange Format) ne sera pas étudié ici, il a été avantageusement remplacé par le format PNG avec des fonctionnalités supplémentaires.

2.4 propriétés intrinsèques des formats sans pertes de données

Potentiels de contenu des fichiers en phase d'enregistrement de la capture du numériseur					
Propriétés	TIFF	PNG	JPF	PDF	PSD
Modes Couleur	RVB, LAB	RVB	RVB	RVB	RVB, LAB
Mode Niveaux de gris	oui	oui	oui	oui	oui
Mode bitonal	oui	oui	non	oui	oui
Profondeur d'échantillonnage des couleurs par couche	1 à 32 bits	1 à 16 bits	8 à 16 bits	1 à 16 bits	1 à 32 bits
Type de compression sans perte de données et bits par couche	LZW 1 à 8 bits				
Zip 1 à 32 bits	« deflate » tous modes	Jpeg 2000 option « sans pertes »	Zip (gris-couleur) CCITT ou Zip bitonal	Compression genre ZIP	
Choix de la compression	utilisateur	automatique	utilisateur	utilisateur	automatique
Forme d'enregistrement classique	Enregistrement « plat », sans calques, à la sortie du numériseur				

Potentiels de contenu et d'usage des fichiers en phase de post-production					
Propriétés	TIFF	PNG	JPF	PDF	PSD
Enregistre les calques de réglage	oui	non	non	non	oui
Enregistre plusieurs calques image	oui	non	non	Sur option	oui
aplatit en calque image unique avec transparence		oui	Oui, transparence sur option	Sur option	
Enregistre les métadonnées IPTC, EXIF, XMP...	oui	oui	Oui sur option	Oui	oui
Options spéciales		Entrelacé pour affichage web	Compression avec perte de données possible dans ces formats	Conserve tracés vectoriels et caractères	
Forme d'enregistrement classique	Enregistrement avec les calques de réglage tant que dure la phase de traitement, puis aplatissage à la fin.				

4.3. Formats avec pertes de données

4.3.1 aperçu des formats

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 **PDF** avec perte de données est un format de diffusion acceptant des documents multi-pages pouvant contenir des images, des textes, des dessins. Lorsqu'il est utilisé à l'enregistrement d'images avec pertes de données, il contient les formats de la famille JPEG étudiés ci-après. Le choix du mode et du taux de compression est fait par l'utilisateur, et le programme prévoit des cas d'usage prédéfinis.

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, ou à une image bitonale. La réduction du nombre de couleurs intervient lors de l'enregistrement, c'est pour cela qu'on le considère comme un format avec pertes de données bien que sa compression LZW n'en génère pas intrinsèquement. Il est progressivement abandonné au profit du jpeg.

Le format **JBIG** est destiné au codage des fichiers au mode bitonal émis par les télécopieurs. Son usage a pu être proposé par certains prestataires, mais il n'apporte rien par rapport aux TIFF et PDF avec compression CCITT groupe 4.

4.3.2 les formats avec pertes de données

JPEG -JFIF

Le format JPEG désigne en réalité plusieurs formats développés par le Joint Photographs Expert Group. Le plus répandu, que l'on peut trouver avec les extensions .jpeg, .jpg, .jpe, .jfif, .jfi. est un format jfif pour JPEG File Interchange Format qui a été créé pour permettre la diffusion sur toutes les plates-formes et en ligne des images numériques en couleur. Il répond à la norme ISO/CEI 10918-1 (JPEG baseline) ou UIT-T recommandation T81, et son code est ouvert et non propriétaire. Son algorithme permet une compression progressive en 10 à 12 niveaux selon les logiciels, avec une dégradation invisible (niveau 12) à importante (niveau 0).

Le format JPEG supporte les modes niveau de gris, couleurs RVB et CMJN. On peut y introduire des métadonnées, et un profil de couleur. Ce format ne supporte pas le mode bitonal (bitmap) ni les transparences.

En archivistique, son usage sera limité à la génération des sous-fichiers à placer en ligne, et à la numérisation de cartes de grand format pour limiter la taille du fichier, bien que dans ce cas, le fichier JPEG 2000 étendu soit préférable par sa performance.

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.



Niveaux de dégradation des images JPEG



Le premier enregistrement en jpeg même en qualité maximale réduit déjà imperceptiblement la gamme des couleurs de l'image.

Le réenregistrement avec le même niveau de qualité peut aussi réduire la qualité de l'image selon les applications, car seul le format du fichier compressé est normalisé, mais chaque développeur peut développer une « moulinette » différente pour y arriver.

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 post-production ou interprétation artistique (travail des couleurs, des densités 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 subtils (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 (effet de carreau) en accentuant les défauts cités précédemment



Cet ultime 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

Définition complète du format Jpeg/Jfif²

JPEG 2000 (JP2, JPF)

Le format Jpeg 2000

JPEG 2000 est un procédé de compression d'images développé aussi par le groupe de travail Joint Photographic Experts Group. C'est devenu une norme commune à l'ISO, la CEI et l'UIT-T sous la dénomination ISO/CEI 15444-1. Le fichier JPEG 2000 peut être enregistré avec ou sans perte de données, en utilisant une transformée en ondelettes (méthode mathématique de codage du fichier natif).

Les performances de JPEG 2000 en compression avec perte de données sont supérieures à celle de la méthode de compression JPEG. On obtient donc des fichiers d'un poids inférieur pour une qualité d'image égale. De plus, les contours nets et contrastés sont mieux rendus, le halo et l'effet de « sable » n'apparaissent plus. Alors qu'un taux de compression de fichier de 60% en JPEG traditionnel (rapport de 0.4 : 1) est presque inévitablement générateur d'artefacts, détruisant de manière visible certains détails chromatiques et la continuité des lignes obliques par effet de crénelage « aliasing », un taux de compression supérieur à 80% en JPEG-2000 demeure visuellement excellent.

Le JEG 2000 contient plusieurs résolutions d'affichage possibles : on peut faire apparaître l'image dans un navigateur internet à différentes tailles en pixels, ce qui est un grand avantage pour moduler l'affichage en réseau en fonction du débit. Le webmestre n'installera qu'un fichier source et l'éditeur internet choisira la taille d'affichage nécessaire.

² <http://www.w3.org/Graphics/JPEG/jfif3.pdf>

Les formats avec compression JPEG2000 supportent les modes niveau de gris, couleurs RVB et CMJN, et en plus par rapport au JPEG, le bitonal. On peut y introduire des métadonnées, et un profil de couleur.

Le format Jpeg 2000 est caractérisé par les extensions .JP2, .J2K, .JPC, .JPX

Le format Jpeg 2000 étendu

Le format JPEG 2000 étendu est une amélioration du précédent, en intégrant les outils indispensables à l'archiviste. Il répond à la norme ISO/CEI 15444-2:2004 et UIT-T Recommendation T.801 il permet en plus :

- l'intégration des métadonnées d'image XMP, EXIF, IPTC, sur option par l'utilisateur,
- la conservation d'un calque unique avec sa transparence (sans crénelage des bords francs) sur option de l'utilisateur (à défaut le calque est aplati sur fond blanc)
- l'insertion d'un profil de couleur ICC sur option par l'utilisateur
- l'utilisateur peut régler la taille du carreau sur lequel se fait la compression, pour optimiser la taille.

Le format Jpeg 2000 étendu est caractérisé par l'extension .JPF



Le premier enregistrement en JP2 lorsque l'option « sans pertes » n'est pas cochée réduit déjà imperceptiblement la qualité de l'image.

Le réenregistrement avec le même niveau de qualité peut dans certaines applications chercher à appliquer de nouveau le même rapport de compression à partir du fichier qui avait déjà été compressé. Le résultat est ravageur. (par exemple si on avait voulu gagner 30% la première fois, en le faisant une deuxième fois, le fichier sera alors réduit de 51% de sa taille originale et la qualité s'en ressentira)



PDF en usage de diffusion

Il est ici nécessaire de parler du format PDF avec des pertes de données, car c'est sous cet aspect qu'il est majoritairement utilisé dans la diffusion des documents d'archives.

Le format PDF (portable document format) est un langage de description de page c'est à dire qu'il peut positionner divers éléments sur une page qui a une échelle de restitution connue en centimètres. Un des principaux avantages de ce format est que les fichiers au format PDF sont fidèles aux documents originaux : les polices, les images, les objets graphiques et la mise en forme du fichier source sont préservés, quelles que soient l'application et la plate-forme utilisées pour le créer. Il a été créé à l'origine pour le monde de l'édition imprimée, afin d'envoyer à l'imprimeur un fichier à restituer sans surprise. Bien sûr, dans ce cas d'usage, toutes les images sont enregistrées sans pertes de données.

Ce format peut contenir ces éléments divers :

- des images bitonales
- des images en tons continus (gris ou couleur)
- des dessins vectoriels
- du texte

- mais aussi des signets, hyperliens, des sons, des séquences vidéo.

A l'origine, la page est décrite selon le langage PostScript® Apple, puis les éléments sont pris un à un et compressés à l'aide des algorithmes les plus performants pour chacun d'entre eux.

Ainsi, les images peuvent-elles être encodées en JPEG-JFIF ou en JPEG 2000. Dans les applications générant un fichier PDF sur un ordinateur à partir du fichier source d'une application, il existe des fenêtres permettant de régler le taux de compression de l'image, mais aussi de réduire la résolution (taille en pixels) des images.

La taille en pixels des images se règle par un filtre passe-bas, avec une réduction de la résolution des images se trouvant au-dessus d'un nombre prédéfini de pixels par pouce à la taille de restitution. Une imprimante virtuelle PDF permet de produire cet artefact à partir d'un fichier PDF à haute résolution/sans perte de données et de générer des copies légères à basse résolution/avec pertes de données pour la mise en ligne.

Pour les acquisitions de séries de documents en mode bitonal, c'est la machine de numérisation qui donne accès aux réglages de résolution en clavier de commande ; la bitonalisation et la compression se font directement dans la machine et les pages sont livrées assemblées en un seul fichier PDF reproduisant le lot de chargement de l'appareil.

4.3.3 Le format PDF-A

Le format PDF-A est un format spécialement dédié à l'archivage issu du format Adobe PDF et qui a été normalisé pour assurer la pérennité à long terme. Etant normalisé, ce format échappe à la propriété de son éditeur, et les codes sources deviennent accessibles et publics.

La spécification PDF/A a été publiée par l'ISO et est utilisée par les organismes d'archivage du monde entier pour garantir la sécurité et la fiabilité de la diffusion et des échanges de documents électroniques.

Le contenu de la norme ISO 19005 optimise l'indépendance matérielle et logicielle ainsi que l'auto-documentation. Il comprend la définition du format PDF/A, mais aussi la façon de développer un outil de visualisation de fichier conforme à ce format. Cela garantit la possibilité future de disposer d'un outil de visualisation. La norme ISO 19005 contient également le document « *PDF reference manual* » d'Adobe® systems

Par ailleurs, l'intégralité du format PDF dans la version 1.7 a été normalisée par l'ISO en juillet 2008 sous la référence ISO 32000. Cette version sert de base au PDF-A le plus récent

Évolutions du format PDF-A

On croyait lors de la création de la norme avoir fixé à jamais le PDF dans le marbre, mais en fait, le format PDF archives suit les évolutions du format PDF créé par Adobe, pour profiter des dernières ressources dynamiques, ceci afin de satisfaire aux exigences du monde des affaires.

La norme ISO 19005 évolue par adjonction de parties, qui définissent à chaque fois des compléments au format de base.

ISO 19005 - Gestion de documents – Format de fichier des documents électronique

Partie	Nom	Contenu	Date d'édition	Norme ISO	Version PDF de base
Part 1	PDF/A-1	Utilisation du PDF 1.4 (PDF/A-1)	2005	ISO 19005-1	PDF 1.4 (Adobe Systems, référence PDF 3e édition, 2001)
Part 2	PDF/A-2	Utilisation du PDF selon ISO 32000-1 (PDF/A-2)	2011	ISO 19005-2	PDF 1.7 (ISO 32000-1:2008)
Part 3	PDF/A-3	Utilisation de l'ISO 32000-1 avec support pour fichiers encapsulés (PDF/A-3)	2012	ISO 19005-3	PDF 1.7 (ISO 32000-1:2008)

On peut juste regretter que le normalisateur n'ait pas créé des champs de métadonnées avec des libellés adaptés au monde des archives : cote, dates extrêmes... mais l'évolution est toujours possible.

Devenu normalisé, le format PDF est omniprésent, et a un bel avenir devant lui. Chaque internaute y a accès avec le graticiel Adobe Reader® distribué en ligne, et disponible pour toutes les plateformes et versions de systèmes d'exploitation.

Au sujet de la protection

Le format PDF-A offre automatiquement à l'enregistrement une protection du document contre les modifications.

Pour le PDF standard, il existe des options permettant de protéger le fichier PDF contre la copie du contenu, voire l'impression, mais certains éditeurs de graticiels contournent ces protections. Le seul procédé résistant bien est l'encryptage et le mot de passe à l'ouverture, mais c'est peu utilisable dans un flux intense de documents en ligne.

On préférera réduire la qualité graphique du contenu si on veut empêcher l'utilisation des images en ouvrages imprimés.

4.3.4 propriétés intrinsèques des formats avec pertes de données

Le tableau ci-dessous dresse un inventaire des formats avec pertes de données, et de leurs propriétés. Le format PDF peut contenir les formats JFIF et/ou JPEG 2000, générés par les machines ou les applications.

Dans tous les cas, si des fichiers à mode bitonal sont enregistrés et compressés sans pertes de données.

Potentiels de contenu des fichiers avec compression à perte de données					
Propriétés	JPEG-jfif	JPEG 2000 – JP2	jpeg 2000 étendu JPF	PDF , PDF/A	GIF (pour mémoire)
Modes Couleur	RVB, CMJN	RVB, LAB, CMJN	RVB, LAB, CMJN	Celui du jpeg choisi	
	RVB				
Mode Niveaux de gris	oui	oui	oui		oui
Mode bitonal (mode enregistré sans pertes)	non	oui	oui	oui	
	oui				
Profondeur d'échantillonnage des couleurs par couche	8 bits	1 à 16 bits	1 à 16 bits, capacité jusqu'à 38 bits	1 à 16 bits	4 bits
Type de compression avec perte de données	Jpeg 1991 (IBM)				
6 à 12 niveaux selon application	Transformée en ondelettes taux continu de compression de 0 à 99%	Celui du jpeg choisi			
CCITT groupe 4 pour le bitonal	Compression ZIP				
Choix de la compression	utilisateur	utilisateur	utilisateur	Utilisateur via fenêtre compression image	automatique
Forme d'enregistrement classique	Enregistrement « plat »,	plat	Plat ou un calque avec transparence	Selon le jpeg choisi	plat
Compatibilité en décompression (lecture)	Totale, avec petites différences imperceptibles selon l'encodage et les marqueurs du fichier			Universelle ISO 32000	Totale (libre)

Utilité d'un calque avec transparence ?



Il peut sembler incongru de citer un mode permettant d'enregistrer et de présenter des images avec des transparences, alors que les archivistes travaillent sur des documents bien opaques. En fait, c'est une utilité essentielle pour détourer les cartes postales par exemple, qui seront présentées avec leur contour physique, mais pas le dos du numériseur. L'élimination du fond se fait lors du post-traitement et parfois automatiquement à la prise de vues.

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 pour une même image. En revanche, il présente la possibilité de définir des zones transparentes, ce qui est utile pour le détournage 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.



Aspects des options de format Gif en 256 couleurs

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.

Le GIF permet de créer de petites animations où différentes images contenues dans le fichier s'affichent successivement selon le rythme défini par leur créateur.

4.4. 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 mono-page pour chacun des documents et individuellement.

- dans certains cas, les séries de cartes postales (dont l'original est monté en accordéon) seront restitués dans un format multi-pages, mais avec une qualité photographique.
- **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,

La numérisation se fera dans un format permettant l'assemblage des pages comme le PDF.

4.4.1. Création d'une photothèque numérique

La photothèque a une double identité : il existe la photothèque de fichiers maîtres, dont la préservation doit suivre les règles de l'archivage permanent, et une photothèque pour la mise en ligne, constituée généralement de copies à basse résolution. Par mesure de prudence, seule cette dernière doit être accessible depuis un réseau connecté à internet.

Dans une photothèque en ligne, il y aura un fichier par image, enregistré dans un format reconnu par tous les logiciels "de base" du monde Apple et du monde PC.

L'accent sera mis sur la portabilité

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**.

Cette photothèque n'est pas obligatoirement en ligne; elle peut être stockée sur un serveur, un disque dur, des DVD; elle peut être mise à disposition seulement **à titre payant** pour les éditeurs qui en tirent un profit pécuniaire. Les fichiers maîtres doivent contenir des indications sur les droits de reproduction pour l'édition imprimée (introduire ces informations dans les métadonnées de l'image).

L'entretien d'une photothèque source, comme tous les stockages d'archives, nécessite une surveillance et a un coût.

- les supports optiques (DVD) doivent être recopiés périodiquement
- le contenu du serveur doit être sauvegardé régulièrement sur un jeu alternatif de bandes, ou sur un serveur distant en réseau sécurisé (*cloud*).

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, , voire le PDF si l'on veut insérer des légendes visibles ou diffuser des séries assemblées. 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, l'auteur, le nom du centre d'archives, la cote de l'original 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. Le format JPEG 2000 peut être utilisé pour contenir en un même fichier diverses résolutions.

4.4.2. 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 résolution, la numérisation des documents d'archives sériés sert généralement simplement à une consultation à distance.

La numérisation peut se faire dans ce cas dans des formats avec pertes de données puisque l'aspect n'est pas l'objectif premier et que le but est d'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. (cela n'a pas empêché certains prestataires de le faire au temps des réseaux indigents, mais c'est à présent stupide)

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.

- Diffuser directement un fichier ou des groupes de pages au format *PDF-A*^{PDF/A}

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. (exemple Gallica³) La diffusion en PDF est donc la plus adéquate pour les documents à pages multiples.

Les trois paragraphes qui suivent décrivent la méthodologie pour constituer un document multi-page à diffuser 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.*

4.4.2.1. Numérisation des éléments à assembler

Lorsque les pages d'un document sont à assembler en un fichier PDF, il est pratique de les numériser directement dans ce format et d'en régler les propriétés sur la machine si aucun post-traitement n'est nécessaire, cela permet de les assembler éventuellement avec des graticiels.

Pour les documents numérisés en mode bitonal, il ne sera pas fait de réduction de résolution ultérieure, les documents seront numérisés directement dans la résolution utile.

Pour les documents numérisés en couleurs (les manuscrits, les chartes...), dont on envisage une diffusion sur plusieurs années, un fichier maître avec la meilleure qualité envisagée pour le projet sera constitué à la capture, dans le but d'anticiper les possibilités des futurs réseaux. Ceux-ci vont permettre dans moins de 5 ans de diffuser des fichiers de 100 MO comme on le fait d'une page internet de 100kO de nos jours. En attendant, des copies de plus basse qualité seront à mettre en ligne.

Lors de la numérisation en mode « livre ouvert » les machines actuelles permettent de redresser l'image, effacer l'ombre de l'espace de reliure, d'effacer les bords...

3. Gallica bibliothèque numérique - <http://gallica.bnf.fr>

4.4.2.2. Assemblage des fichiers au format PDF

L'assemblage constitue la post-production. Alors que la numérisation est machinale, l'assemblage nécessitera des formes de discernement propres à l'archiviste. Si ce travail n'est pas effectué par le centre d'archives, il aura fallu au préalable installer dans le document physique des balises matérialisées par des pages qui contiennent toutes les informations nécessaires à un travail effectué de façon mécanique et aveugle :

- le titre du bloc
- les métadonnées à entrer
- éventuellement la page de garde à assembler...
- le mode colorimétrique de numérisation
- l'emplacement des signets

Le format PDF va permettre d'assembler des pages qui auront été numérisées dans de différents modes, des pages de tailles différentes, des pages de sources différentes : pages numérisées, pages de description au mode texte ajoutées par le centre d'archives...

Par exemple les pages de texte pur peuvent avoir été traitées en mode bitonal et celles imagées en mode couleur ou niveaux de gris. Une ségrégation peut être pratiquée à la numérisation, mais selon le coût de prestation, il peut être moins coûteux de tout numériser en couleurs et de bitonaliser le texte par la suite.

Beaucoup de prestataires proposent des numérisations qui sont sensées reconnaître les plages de texte et d'images (tuilage de l'image) mais cela ne donne pas toujours de bon résultats. Par exemple, les gravures sont reconnues comme des images en niveaux de gris à cause d'une résolution insuffisante alors que ce sont des modèles parfaits de « bitonalité ». On se gardera donc d'utiliser ce procédé pour les ouvrages anciens, ou un test pilote sera mené auparavant.

L'assemblage se fera selon les circonstances de diffusion, en reliant un nombre de pages qui correspond à la capacité de diffusion et/ou au centre d'intérêt de création d'un bloc, par exemple :

- une année d'un registre paroissial en bitonal,
- Un chapitre d'un livre très imagé,
- Tout un ouvrage s'il n'y a que du texte (bitonal),
- Toutes les parties intéressant un sujet et extraites de divers livres, magazines, archives, publications électroniques...

L'assemblage est la partie où l'intelligence de l'archiviste se fait sentir : il est facteur du contenu, et non de l'aspect trivial d'un emplacement ou d'une date de dépôt...

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

Livre blanc Adobe® PDF, norme d'archivage (cf. pdfarchiving1.pdf)

4.4.2.3 L'indexation ou reconnaissance de caractères

Le format PDF est typiquement celui dédié aux représentations imagées des anciens imprimés, qui ne sont pas des documents texte mais des images.

Qu'est-ce qu'une indexation ?

C'est un moyen informatique de reconnaître le texte représenté sur l'image numérisée.

Le procédé s'appelle ROC en Français pour Reconnaissance Optique des Caractères et OCR en anglais (Optical Recognition of Characters)

Lorsque ce texte est reconnu dans la page, dans le document, il est possible d'utiliser un moteur de recherche de texte et trouver directement l'objet de sa recherche.

Les lecteurs de documents PDF ont un moteur de recherche intégré, et les disques de stockage indexés permettent l'identification directe de tout fichier contenant les mots recherchés.

On voit là directement l'intérêt d'indexer les pages de texte numérisées et aussi d'intégrer des métadonnées en clair dans les images.

L'indexation de l'image d'un texte

On peut indexer un fichier numérisé en conservant son aspect original, en couleurs, niveaux de gris ou bitonal, autrement dit en ayant le texte sous-jacent à l'image réelle du texte, exactement à sa place physique. C'est le mode préférable pour l'archiviste, parce que :

- les caractères mal formés sont toujours reconnaissables à l'œil, alors qu'ils feront chuter le logiciel de reconnaissance optique des caractères
- l'aspect du document original sera préservé, ce qui est la prétention d'un ouvrage d'archiviste.
- Le logiciel ROC reconnaît tout ce qui lui semble un caractère, y compris les taches, traits, petits dessins, mais il n'interprète pas les aplats, les images.

Ce procédé est promis à un grand avenir et est appliqué à des bibliothèques en ligne comme Gallica. Mais il a aussi ses limites, pour les raisons suivantes :

- les textes manuscrits ne sont pas reconnaissables, sauf pour les calligraphies et écritures de dessin industriel,
- La limite du ROC apparaît vite lorsque les caractères sont mal formés, peu contrastés, trop petits, d'une police trop ornée...
- Le fonctionnement du ROC nécessite une résolution de contraste assez élevée, que l'on peut estimer à 300 dpi pour les documents modernes très contrastés et grands caractères de 10-12 points, et de 400 à 600 dpi pour les caractères plus petits ou peu contrastés.

Certaines machines intègrent dès la capture un post-traitement d'indexation du texte, généralement associé à un mode de fonctionnement bitonal.

Diffusion des fichiers PDF indexés

Lors d'une demande de prestation, l'attention est à porter sur la mise en ligne des documents indexés. Certains prestataires indexent le texte, mais les textes indexés restent dans le site internet et ne migrent pas dans le fichier PDF téléchargé par l'internaute. Il doit donc faire un téléchargement du texte seul (quand c'est possible). Le texte se trouve alors dissocié de son image dans le document original. C'est parfois malcommode... à part pour les plagiaires ! Un bon document PDF doit contenir son texte indexé sous-jacent.

Dans certains lecteurs de fichiers PDF, il est possible d'extraire en une seule fois tout le texte contenu dans le document par une commande d'export ou « d'enregistrer sous ». Il sort avec ses imperfections de reconnaissance, s'il n'a pas été vérifié manuellement.

L'enregistrement final

Par simplification, et pour assurer la plus grande pérennité possible, l'archiviste devra utiliser le format PDF-A pour la diffusion hors des espaces de travail du centre et pour le stockage.

Liens dynamiques et protection individuelle



Liens dynamiques du document PDF

Le lien dynamique le plus usité dans un document PDF est le signet. Dans une fenêtre juxtaposant le document, l'archiviste peut créer des signets avec le titre de la partie ainsi repérée.

Un éditeur PDF permet aussi d'ajouter des liens dynamiques (hypertexte) vers les pages du document lui-même, à partir de son sommaire.

Des liens peuvent aussi être créés vers un site internet, ou vers un autre document placé sur un serveur local ou distant. Mais dans ce cas, il faut bien être sûr de la stabilité de l'emplacement des sources.

Protection du document PDF

Le format PDF 1.X 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, tant qu'on utilise Adobe Reader® ou un logiciel approuvé. Malheureusement, certains logiciels de lecture ne respectent pas la protection ou transforment les pages en simples images, sans fonctionnalités ni de protection, ni d'indexage.

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, mais il n'est pleinement efficace que s'il conditionne l'ouverture. Cette disposition est intéressante pour **protéger les fonds d'archives à un niveau juste nécessaire**, sans pénaliser les utilisateurs, ni favoriser les plagiaires.

4.4.2.4. Recopie des documents en basse résolution

C'est avec imprimante virtuelle 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 maître 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 résolution**, il suffira de les copier du document source et de les réintégrer dans le document allégé par la commande de remplacement de pages.

Vérifier si la réimpression en PDF ne détruit pas les liens hypertexte que l'on a pu placer dans l'original, si on utilise des gratuits pour la conversion.

4.4.2.5. 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 imprimante pdf 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 les codes source qui sont publics..

4.5. Formats libres et formats propriétaires

Les formats propriétaires proposés par des distributeurs de matériel de numérisation existent encore, mais tendent à se raréfier. La plupart des machines enregistrent de nos jours en Jpeg, tiff et jpeg 2000, et en PDF contenant l'un de ces formats.

Il existe cependant une exception à cette règle, c'est le PIX2 PDF proposé par un distributeur français, dont le but principal est de reproduire les documents imprimés et écrits modernes, en couleurs, et dans une taille très petite correspondant à un fichier bitonal. L'astuce est simple (en apparence).

Toutes les parties en fort contraste de l'image sont bitonalisées puis coloriées dans leur ton d'origine, et placées en premier plan alors que le fond de page est entièrement flouté pour constituer une image très légère à l'enregistrement. Le résultat est bluffant, mais il y a plusieurs inconvénients majeurs :

- A chaque image numérisée, il faut payer une redevance à l'éditeur, quand le crédit est fini, il faut racheter des unités...
- Il est impropre au traitement des images, l'objectif est le formulaire imprimé moderne
- Les fichiers ne sont pas toujours imprimables sur les imprimantes à langage PCL, l'aplatissement de l'image ne se fait pas toujours, et seul le fond flou apparaît.
- Il s'agit d'un artefact, plutôt destiné à l'usage commercial qu'aux archives durables.
- Enfin, étant propriétaire, il n'a pas de pérennité certaine et est hors normes internationales de facto.

4.6. 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. Les nouvelles générations de films argentiques mis au point et identifiés LE 500 par les normes ANSI et ISO, préparent à une durée de vie minimum de 500 ans

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.



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



Introduction

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 le système d'exploitation Mac OSX est spécialement conçu pour le traitement d'images, c'est pourquoi il sera préféré pour les opérations lourdes de post-production, alors qu'un PC est suffisant pour le traitement trivial de l'image juste acquise.

5.1. Format des originaux déterminant le numériseur

5.1.1. Aspects d'une numérisation de masse externalisée

Au fil des contraintes budgétaires et des objectifs, un plan de numérisation sera établi à partir des utilités de numérisation des documents, soit :

- Le nombre de chercheurs qui s'y intéresse ou
- Le péril encouru par les originaux en cours de dégradation

C'est le choix des objets à numériser qui va fixer les machines à utiliser. Une numérisation de masse fera appel à des prestataires qui disposent de moyens automatisés de production, et d'une chaîne de production travaillant avec un manuel d'assurance de la qualité, suivant ISO 9000.

En quelques mots, il faudra :

- Sélectionner et faire la ségrégation des originaux par formats, par types, etc ;
- assurer la sécurité physique des originaux pendant le transport et le stockage intermédiaire chez le prestataire
- La numérisation dans le format de fichier maître,
- le contrôle qualité et le récolement des fichiers obtenus
- la restitution et le reclassement des originaux

A l'issue de ces opérations, le nommage et la post-production sont à effectuer sur les fichiers fournis.

Pour un tel traitement, il s'agira d'une gestion de projet, et l'archiviste du centre sera plus un préparateur et un donneur d'ordre qu'un praticien. Cet aspect de projet de numérisation est bien développé dans le chapitre « cahier des charges de numérisation » du site internet de la Direction des Archives de France.

L'annexe 1 de ce cahier des charges fixe les résolutions et formats approuvés par la DAF, et il conviendra de s'y conformer, même lorsqu'on numérise avec les propres moyens du centre d'archives.

5.1.2. Aspects d'une numérisation avec les propres moyens du centre

Actuellement, les grands centres d'archivages régionaux et nationaux ont leurs propres moyens de production, parfois aussi performants que ceux des prestataires, mais dont ils réservent l'usage à des originaux de grande valeur, des petites séries, ou des demandes ponctuelles sur une partie d'un fonds.

Pour faire un inventaire « convivial » des moyens de numérisation que l'on peut posséder en propre, nous allons faire un petit inventaire des numériseurs les plus répandus, abordables et utilisables pour de petites séries.

La systématisation n'étant pas l'objet de ce type de machines, c'est plutôt la variété des travaux qui va amener à se constituer un petit « laboratoire » d'acquisition.

Mais la numérisation étant devenue un objectif commercial et industriel, il se trouve à présent sur le marché une grande variété de machines de petit format dont la productivité est assez élevée.

? 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.

5.2. Le choix du numériseur

La gamme des appareils d'utilisation courante existant sur le marché est adaptée à la numérisation des originaux en feuilles détachées actuels. Ces appareils permettent de numériser jusqu'au format A3, dans tous les modes colorimétriques.

Pour les registres, les livres, les plans de moyen format, les archivistes font appel à des numériseurs spéciaux, qui sont bâtis comme les bancs de microfilmage. La caméra statique ou dynamique selon les appareils, scrute l'original disposé en-dessous, dans un arrangement que l'on nomme « à livre ouvert ». L'original est éclairé par le haut et la numérisation se fait par l'image réfléchie.

Enfin, pour la numérisation des plaques de verre, des négatifs, diapositifs et ektachromes, des fabricants de matériel photo proposent des numériseurs à plat avec un couvercle lumineux et des porte-films. D'autres appareils sont dédiés à la numérisation des films en bandes. Dans ces deux cas, La numérisation se fait par **transparence**

5.3. La disparition de certains types de numériseurs

Aussi incroyable que cela paraisse, les numériseurs photographiques à plat et les numériseurs à films sont condamnés à court terme. Ils ont été l'objet d'un engouement passager lorsque le grand public voulait transférer ses images familiales, diapos, négatifs et tirages sur un support numérique. Depuis 2010, le marché se réduit à peau de chagrin parce que l'image est directement acquise en numérique avec les appareils photo et les téléphones multifonctions. Les fabricants abandonnent ces fabrications non rentables, et les appareils encore abordables actuellement, vont devenir dans l'avenir des machines rares et chères réservées exclusivement à l'usage des professionnels du patrimoine et des archivistes.

Les numériseurs de bureautique échapperont à cette disparition, parce qu'ils sont seulement une partie d'un appareil combiné photocopieur – numériseur – imprimante – Téléfax. Mais leurs qualités de numérisation sont insuffisantes au traitement des images de qualité.

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 devrait être 3 fois l'épaisseur des traits les plus fins des gravures (estampes) à analyser, soit en pratique au minimum 1200 dpi par 1200 dpi pour les documents opaques.
- 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 pour agrandir 8 fois. Les appareils du marché actuels proposent 9600 x 9600 dpi, c'est amplement suffisant, car au-delà, on va juste analyser les grains de l'émulsion photographique.
- La capacité à **distinguer et restituer les couleurs**. On trouve des numériseurs de 24 à 48 bits, soit de 16 millions à des milliards de couleurs possibles pour chaque pixel. Une profondeur de 24 bits, convient à la reproduction des documents d'archives opaques. Il est nécessaire de numériser à 48 bits uniquement pour les diapositives ou les tableaux qui contiennent des couleurs subtiles ou à la limite du modèle RVB.

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 qui peuvent présenter d'importants contrastes ou inversement une faiblesse de tons nécessitant une analyse très sensible.

Le choix raisonné



Pour un usage en centre d'archives, ce n'est pas la vitesse de numérisation qui prime, mais plutôt la qualité. Le pilote de l'appareil peut être programmé par un tiers ou par le fabricant de l'appareil. Si l'on exécute simultanément des numérisations sur plusieurs appareils de marques différentes, l'usage d'un pilote commun peut donner des résultats plus constants.

5.5. Inventaire des fonctionnalités

Les fonctionnalités des numériseurs destinés à la photographie sont communes à tous ces appareils, on y retrouve :

- un aperçu rapide et des fenêtres de cadrage que l'on peut mémoriser
- des contrôles de densité, point blanc, point noir, gamma
- des contrôles de rapport de restitution (réglage programmé du rapport entre la taille d'image et la taille dans le fichier enregistré, en archives on choisira 100% pour les documents opaques)
- un contrôle de détramage pour la reproduction des imprimés
- un choix du fichier cible, et une numérotation automatique des clichés.
- Un choix de profil colorimétrique et de format d'enregistrement, souvent sans compression pour les fichiers sans perte de données ; une post-production est nécessaire.

Fonctionnement technique d'un numériseur

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 discrétisée par une conversion des signaux analogiques en valeurs numériques. Et c'est dans cette opération que se décide la qualité du résultat.

C'est le pilote du numériseur qui réalise cette conversion

Lorsque le numériseur travaille par transparence, la lumière traverse l'original. Il apparaît à ce moment les mêmes sujétions que pour le tirage photo en agrandisseur : les poussières apparaissent sur le tirage numérique.

Les numériseurs à livre ouvert, ou caméra dynamique, éclairent l'original à distance, l'ambiance du laboratoire de prises de vue a alors une importance, la lumière doit être diffuse.

Effets de la lumière du numériseur

En règle générale, les éclairagements produits par les numériseurs modernes sont peu agressifs pour les originaux, il n'y a pas d'échauffement, et peu d'ultra-violets et surtout pendant une durée très courte. Les lampes à filaments utilisées auparavant pour la photo « au statif » étaient bien plus agressives, surtout par la chaleur qui séchait inconsidérément une seule face du document. Cependant, certains conservateurs restent frileux quant à l'exposition d'un original à une lumière électrique, préférant le laisser en dépôt ce qui a pour effet de manquer d'en faire une sauvegarde numérique qui lui serait plus salvatrice que nuisible.

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.
- 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 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).

5.6. Numériseurs de petit format

Cette famille de numériseurs sera essentiellement compétente pour ce que nous avons appelé la "photothèque" constituée de documents plats et mono-page. Ils sont dédiés à des traitements fins et circonstanciés, généralement à la demande. La rapidité n'est pas le premier critère, qui est plutôt la qualité. Ces appareils peuvent aussi traiter des originaux transparents (plaques, films).

Numériseurs à plat à vocation photographique

Format	jusqu'à A3, A2 dans certains cas
Originaux	opaques ou transparents avec le dos lumineux, exposition sur vitre
Vocation	photographies sur papier, films en châssis, diapositives, plaques de verre.
Vocation supplémentaire	petites estampes, lithographies, cartes postales, lettres, actes...mono-page
Profondeur de champ	1 cm environ pour les appareils à miroir et optique
Temps de numérisation	20-30 secondes à plusieurs minutes selon les tailles d'originaux
Automatisation du passage	non
Automatisation des prises de vues de diapos et films	oui la plupart du temps
Réglage du gamma et des couleurs	Par l'opérateur, préréglages enregistrés possibles
Restauration automatique des couleurs	sur option
Résultat	très bon fichier pouvant servir de sauvegarde des originaux en voie de dégradation

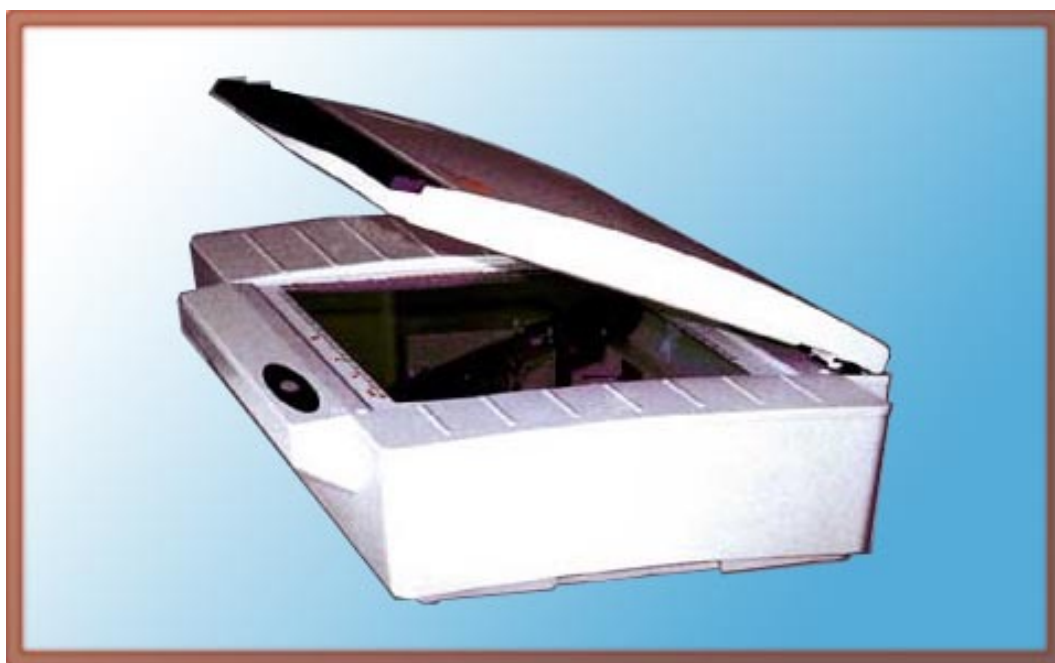


Image 4 Numériseur A3 à plat avec zoom optique interne

Le numériseur à plat à vocation documentaire

C'est par excellence le copieur multi-fonction utilisé couramment dans l'industrie et le commerce. Il ne traite que les documents opaques. Il est connectable à un réseau par lequel le document numérisé va rejoindre son lieu de post-traitement, voire de stockage si on le choisit ainsi.

Cet appareil est universel et très rapide. Les plus récents offrent une qualité de numérisation telle que l'on peut les utiliser pour traiter soi-même des séries à la demande. Un opérateur peut numériser plusieurs centaines de photographies ou documents dans une journée. Les formats sans pertes de données sont disponibles, dont le tiff et le jpeg 2000.

Format	Jusqu'à A3
Originaux	Opacques, exposition sur vitre
Vocation de base	Documents manuscrits ou imprimés, affichettes, petites estampes, lithographies, lettres, actes...
Vocation supplémentaire	photographies sur papier, pages d'albums photo, cartes postales...
Profondeur de champ	1 cm environ pour la plupart des appareils, 0,2mm pour les petits appareils à éclairage par diodes
Temps de numérisation	2 à 5 secondes selon les tailles d'originaux et la résolution
Automatisation du passage des originaux	Oui avec le chargeur automatique feuille à feuille
Possibilité de faire des fichiers PDF multi-pages	Oui, c'est la vocation première. PDF sans pertes ou haute qualité possible selon les appareils, Tiff multi-pages
Réglage des couleurs et du gamma	Non, seulement le mode et la densité
Résultat	très bon fichier pouvant servir à la diffusion, et à la copie numérique de sauvegarde simple



Copieur numérique connecté

Le numériseur à plat à vocation documentaire pour la série

Dans cette catégorie des numériseurs à plat de bureau, il existe actuellement des numériseurs dédiés à la dématérialisation des documents commerciaux et industriels, intégrés dans le flux de gestion des GED (gestion Electronique des Documents). Ces numériseurs ont les caractéristiques informatiques des

copieurs connectés, mais ils sont limités à la capture en série de documents de format A4, généralement au format bitonal où ils excellent. Le pilote peut comprendre un système automatique de reconnaissance des caractères (ROC) pour indexer les documents. Leur usage s'étend à la numérisation en couleurs et niveaux de gris pour la capture des documents imagés comme les rapports d'expertise, magazines...

Format	A4 (210x297) et/ou légal américain (215 x355 mm)
Originaux	Opaques, exposition automatique au défilement
Vocation de base	Documents feuille à feuille, couleur ou monochromes
Vocation supplémentaire	aucune
Temps de numérisation	1 seconde par page ou moins
Automatisation du passage des originaux	toujours avec le chargeur automatique feuille à feuille
Possibilité de faire des fichiers PDF multi-pages	Oui, c'est la vocation première.
Réglage des couleurs et du gamma	Non, seulement le mode
Résultat	Fichier pour la documentation ou l'archivage non historique



Numériseur multi-pages pour les documents d'affaires.

5.7. Numériseurs de grand format

Numériseur à caméra dynamique

Les numériseurs de grand format sont initialement dévolus à la numérisation à livre ouvert. Leur forme de construction est directement issue des machines de microfilmage des registres d'état civil, avec plateaux compensés pour placer les pages dans un plan. Tout comme dans le précédent appareil, la caméra surplombe le livre à numériser, et est dynamique (la caméra se déplace) pour couvrir les très grands formats. Ces appareils conçus pour les grands registres sont aussi très utilisés pour les affiches, les cartes, mais ne se trouvent que dans des centres d'archives nationales ou chez les prestataires.

Les appareils actuels numérisent en couleurs. Ce sont des chambres de reproduction améliorées par le mouvement de la caméra, permettant de compenser la petite taille du capteur par des prises de vues en carreaux complémentaires. Ces appareils n'ont pas la finesse d'exploration d'un numériseur à plat, du fait même de leur conception qui est plutôt celle d'un appareil photo.

Format	Jusqu'à A0 environ
Originaux	Opaques, posés sur table ou plateaux compensés
Vocation de base	Registres, livres
Vocation supplémentaire	Cartes, plans, affiches
Profondeur de champ	Plusieurs centimètres
Temps de numérisation	30 secondes à une ou 2 minutes selon le format et la résolution
Automatisation du passage des originaux	Non, sauf cas particulier d'utilisation d'un tourne-page automatique pour la numérisation des livres
Réglage des couleurs et du gamma	Sur certains pilotes
Résultat	très bon fichier pouvant servir à la diffusion, et à la copie numérique de sauvegarde simple

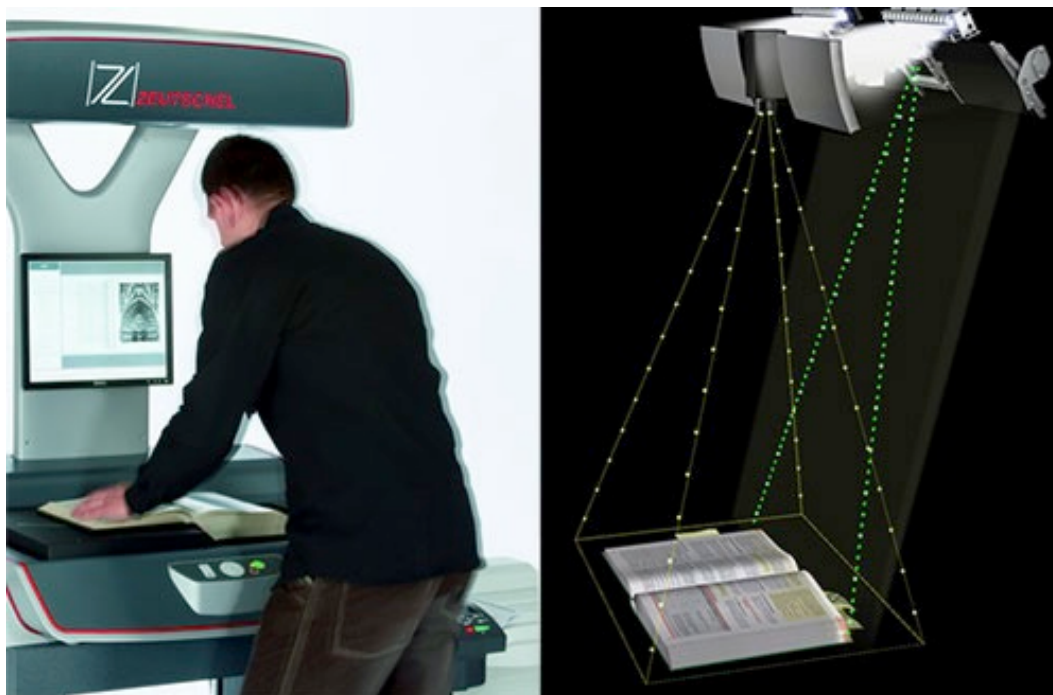


Numériseur à caméra dynamique et la station d'acquisition

Numériseur de livres ouverts

Le numériseur à livre ouvert, spécifiquement dédié et conçu pour cela, fonctionne selon un procédé spécifique. L'acquisition de l'image se fait par une rangée de cellules disposées de façon linéaire, et une barrette de lumière scrute progressivement l'original. Ce mouvement étant fait à partir d'un axe de rotation, il y a des compensations de mise au point et géométriques.

Le logiciel d'acquisition comporte un effacement des bords, un redressement et une séparation des pages, une linéarisation des parties se trouvant dans la courbure due à la reliure, l'effacement de l'ombre de la reliure, et des algorithmes de bitonalisation agissant par carreaux pour tenir compte des différences d'éclairement. Certains ont une vitre que l'on peut appliquer sur l'original pour corriger sa planéité, mais cela peut être nuisible aux reliures fragiles.



plateaux compensés à barrette lumineuse

Cette illustration montre l'opérateur disposant un livre sur les plateaux compensés et le principe d'éclairage – Éclairage par une barrette lumineuse.

Numériseur pour petits livres

Il existe des versions allégées de ces appareils pour l'acquisition des livres in-quarto, destinés aux bibliothèques et centres de documentation. Leur prix est plus abordable, et ils permettent de numériser les petits livres sans les presser trivialement sur un numériseur à plat !



Le numériseur pour petits livres

Le numériseur pour petits livres



Plus jamais cela ! la reliure crie de douleur !

Plus jamais cela ! la reliure crie de douleur !

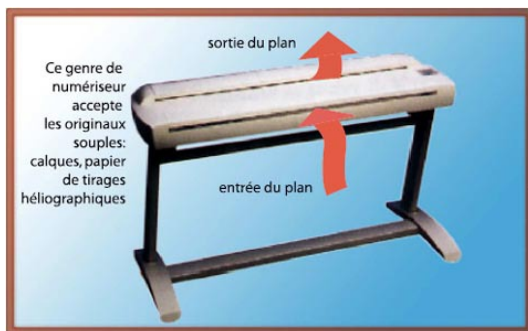
Numériseurs à plans



Complément

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

Numériseur à plans et documents mono-page de grand format.

Format de sortie des numériseurs de grand format



Complément

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.

5.8. Numériseurs de films et microfilms

Ces appareils sont conçus pour la numérisation des films photographiques en bandes ou des microfilms. Mais même s'ils sont conçus sur les mêmes principes, un seul appareil ne peut pas remplir les deux fonctions. Alors que les numériseurs photo sont conçus pour la couleur avec des pilotes spéciaux qui éliminent la couche orangée des négatifs, les numériseurs de microfilms sont dédiés à l'image monochrome et sont automatisés pour l'avance vue par vue sur toute la longueur d'une bobine.

Numériseurs de films photo

Le numériseur de film photo comporte d'une source lumineuse diffuse qui traverse le film. L'image ainsi produite est analysée à travers un système optique, puis discrétisée avec application de filtres, corrections de gamma et de niveaux, qui sont ajustés sur une prévisualisation. Ces appareils acceptent le film 35 mm et parfois les diapositives 5X5. Quelques appareils acceptent le rollfilm 6cm (formats 6x6, 6x9, 6x4,5). Les planfilms ne peuvent pas être numérisés avec ces appareils, il faut utiliser un numériseur à plat adapté



numériseur à films professionnel

Les numériseurs à plat pour documents opaques et transparents



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.



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

Numériseur d'imprimerie au format A2.

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**.

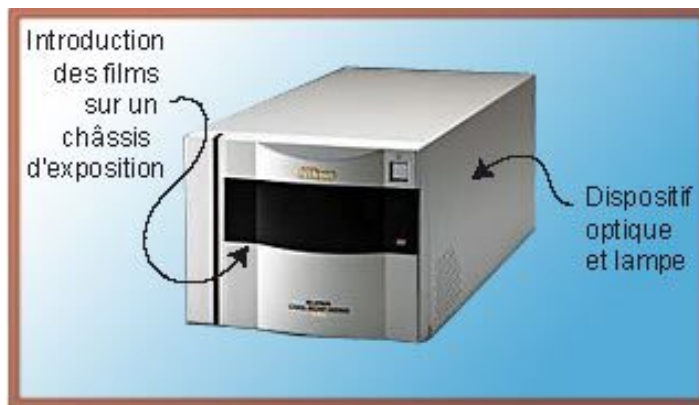
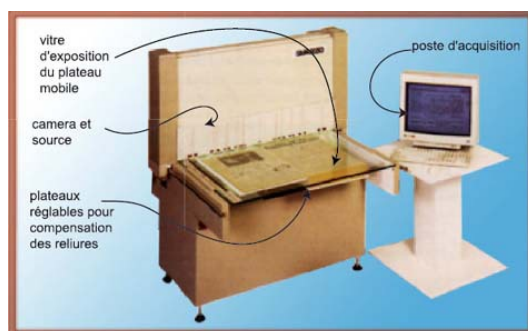


Image 5 Numériseur à diapos 5x5 et négatifs de petit format

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

Un numériseur pour livres et grands documents au format A1.

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**

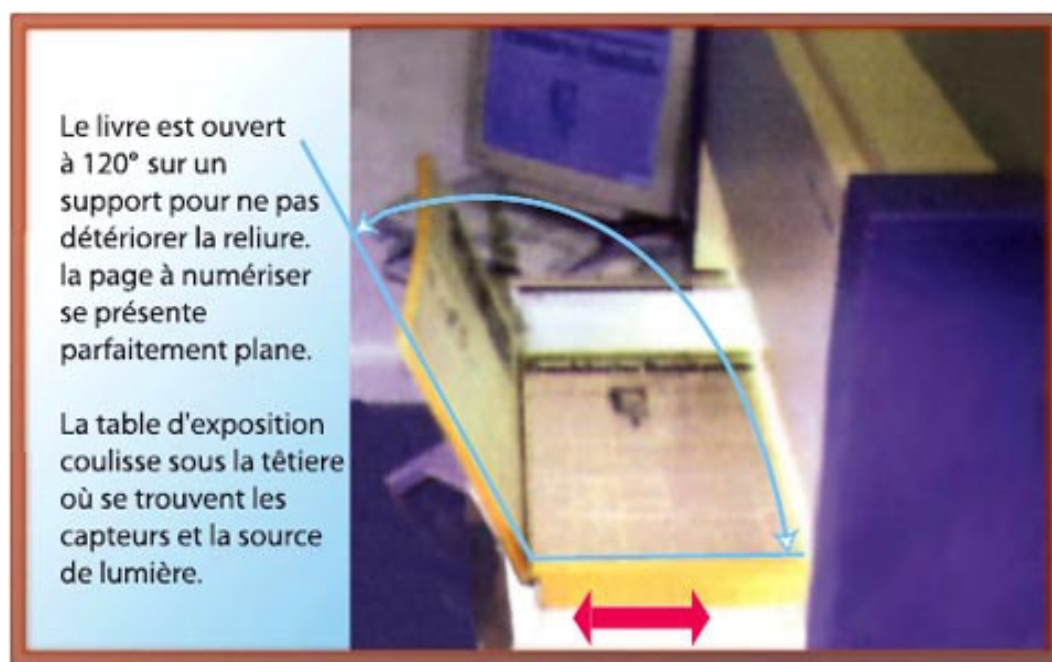


Image 6 Numérisation d'un livre sans détériorer le dos de la reliure

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.

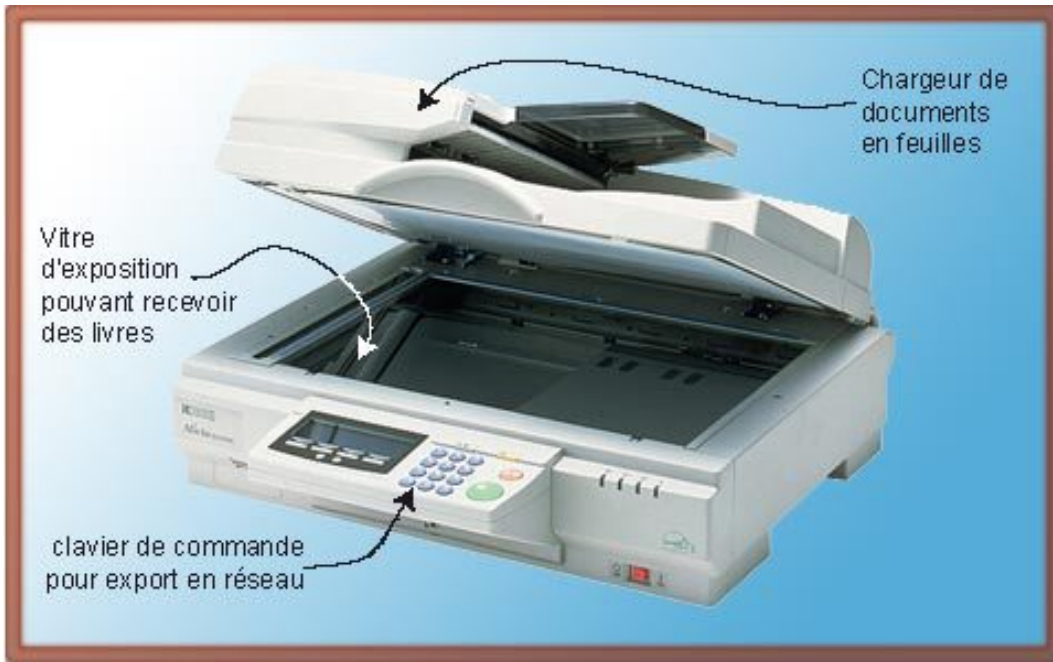


Image 7 Numériseur doté d'un chargeur automatique de documents. (appareil de bureautique)

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.

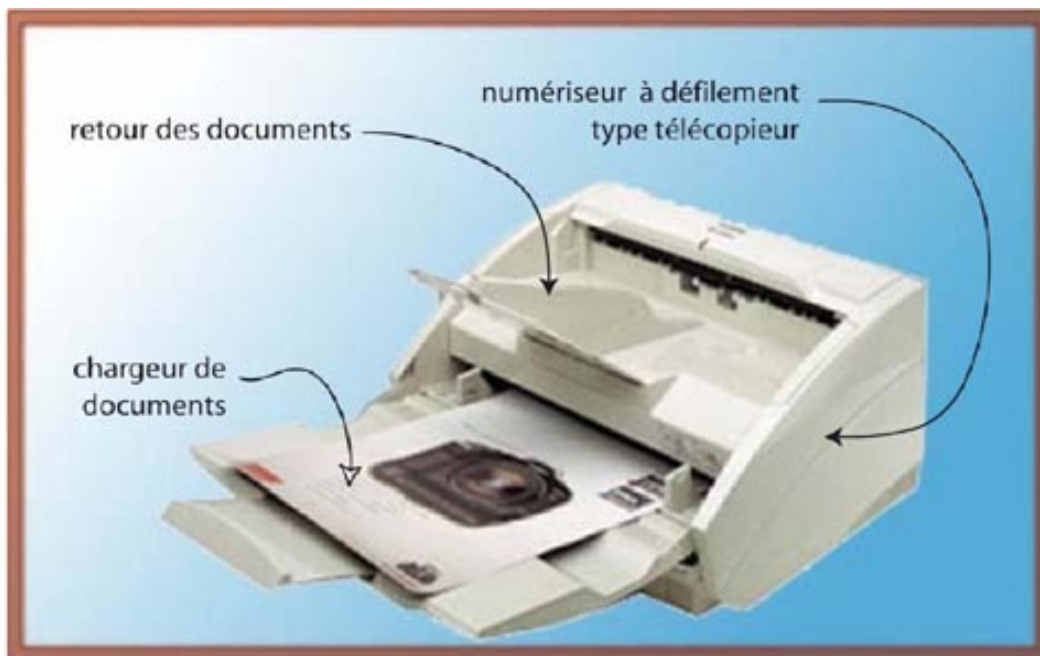


Image 8 Numériseur multi-pages pour les documents d'affaires.

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 :

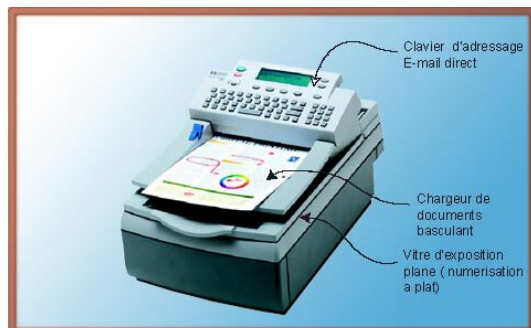
- mode «texte» générant un fichier bitonal à basse-moyenne résolution (catégorie 1- texte lisible qualité télécopie en mode fin)
- mode «photo» générant un fichier couleurs RVB à basse-moyenne résolution (catégorie 2)
- mode «document en couleurs» qui active un module d'analyse et mélange les modes bitonal 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**.



Numériseur autonome à chargement automatique.

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

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.



Image 9 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.

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.

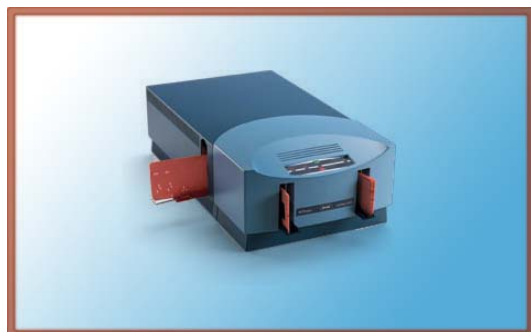
La définition indiquée par le logiciel d'acquisition, par exemple 300dpi pour un document A4, est celle du document restitué à sa taille d'impression, c'est pourquoi il faut bien prendre garde au taux de réduction appliqué lors du microfilmage pour ne pas sous-échantillonner ni sur-échantillonner le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

Numériseur de cartes à fenêtre



Complément

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.



Numériseur de cartes à fenêtre

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

Numériseur de microfiches



Complément

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.



Image 10 Numériseur de microfiches

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.

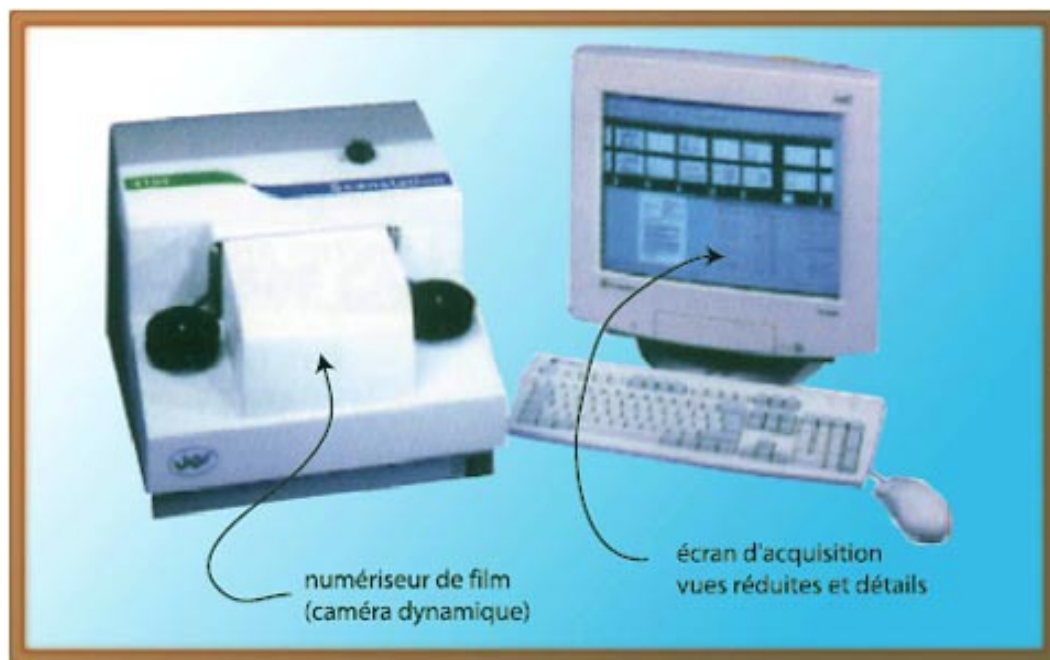


Image 11 Numériseur de microfilms

5.9. 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 webmasters.

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.

5.10. Ségrégation des documents

Ségrégation des documents mono-page

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

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 sérieuse, 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 procédera 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).

5.11. 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 détourage,
- 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).

5.12. 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.

5.13. 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.

5.13.1. 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.

5.13.2. 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 une 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



Introduction

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.

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.

6.2. Choix d'un procédé physique pour les reliés



Un numériseur de livres à caméra, conçu comme un banc de microfilmage.

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).

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.



Image 12 Exemple de numériseur multi-pages à utiliser seulement si les originaux sont en bon état.

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.

Les modes d'acquisition du document manuscrit

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.



Image 13 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.

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



Image 14 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.

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



Image 15 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.

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.

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).

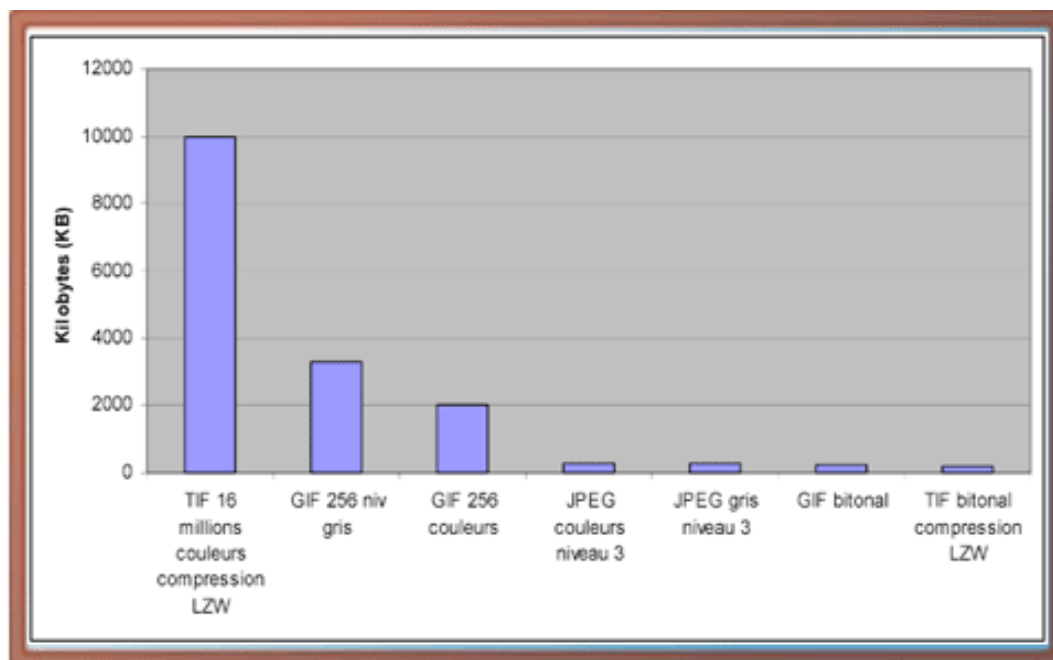


Image 16 Analyse comparée des tailles de fichiers

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.

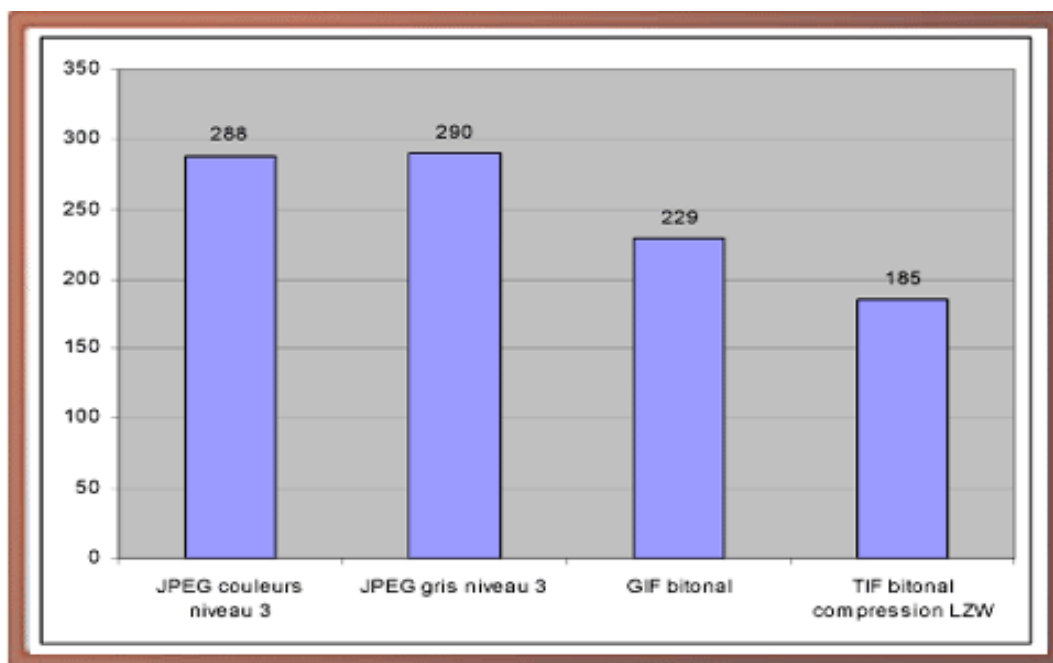


Image 2 Les modes d'enregistrement

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.

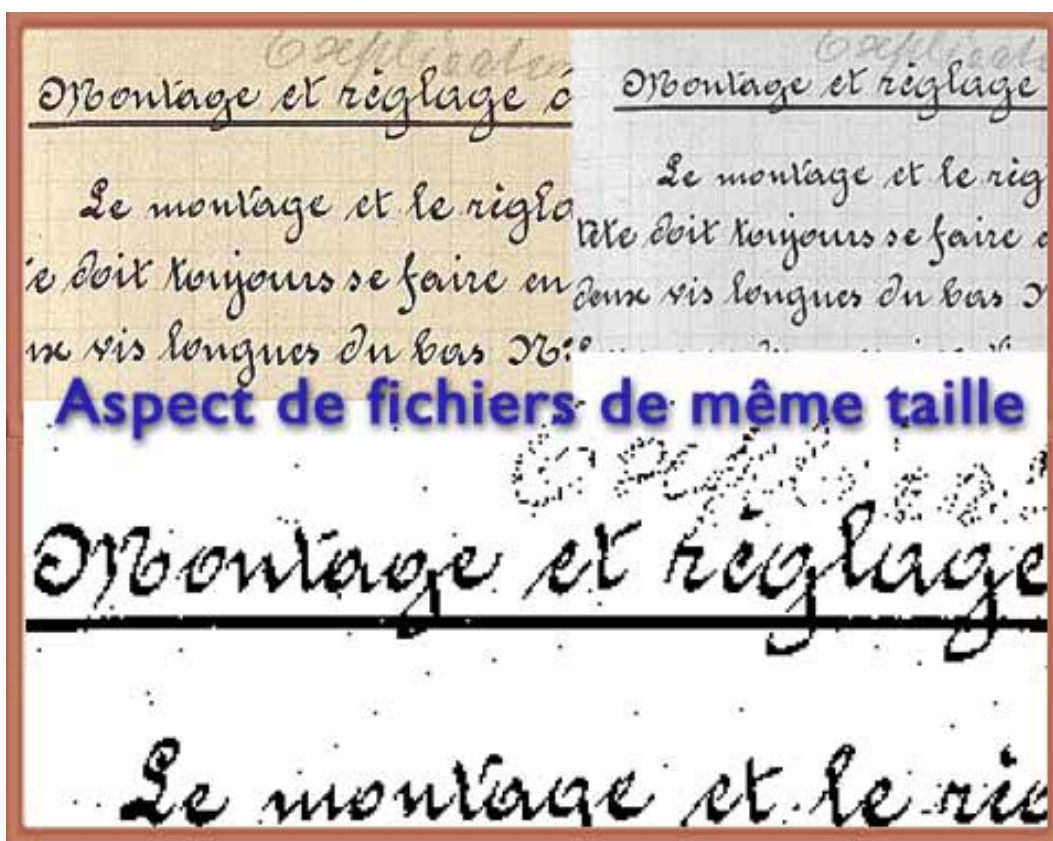


Image 3 Aspect de fichiers de même taille

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

Introduction

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.

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.

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



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.

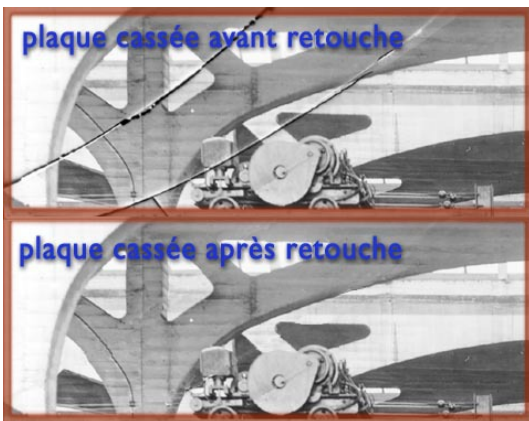
Il peut se former en outre des spires de Newton dans la zone de contact des verres (taches irisées).

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.

Retouche d'une plaque cassée et altérée à la cassure

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 les 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.



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 180mm / 36 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».

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.

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

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

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.

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.



Princesse Potochka

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.

7.4.2. Photographies en couleur sur papier

Les photos en couleur contemporaines présentent pour certaines des dégradations dues au vieillissement des copulants. Une dominante rouge-violet apparaît sur l'ensemble de l'image, les verts et les bleus s'affaiblissent. Cette dégradation est inexorable et se produit même à l'abri de l'air. Il est donc urgent de numériser ces images pour sauver ce qui peut encore l'être.

Il existe pour ce faire deux chemins :

- la saisie brute avec tous les défauts existants en vue d'un post – traitement
- la « réparation » des couleurs avec un filtre expert du numériseur.

Dans ce second cas, le numériseur utilise des systèmes experts et parfois une exploration de l'image par les infra-rouges, ce qui donne un résultat beaucoup plus rapide et efficace qu'une correction par tâtonnements. L'image peut parfois avoir un aspect criard ou peu réaliste, mais il sera plus facile de repartir de cet aspect pour refaire quelque chose de vraisemblable.

Il faut garder à l'esprit que la photographie en couleurs sur papier est toujours un « mensonge arrangé » qui a suivi les modes et les goûts du public, on trouve même des dominantes différentes selon les contrées : brun -violet aux Etats-Unis, vert en Angleterre, couleurs chaudes en Allemagne...

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 couleur

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

8. Numérisation des gravures



Introduction

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.

8.1. Numérisation au bon mode



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.



Gravure acquise en mode bitonal et gravure en niveaux de gris

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.



Réimpression d'une gravure

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

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.



Réglage de la résolution

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.

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.



Image 17 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.

Gravure numérisée avec un seuil de conversion trop haut

- 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.



Gravure numérisée correctement

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).

8.2. Cas spéciaux

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

- lorsqu'elle doit être réduite en taille pour une impression domestique,
- si le rapport de taille d'impression (en ouvrage) est inférieur à 50%.
- lorsqu'elle est à diffuser sur la toile ou pour consultation à l'écran uniquement.

On peut :

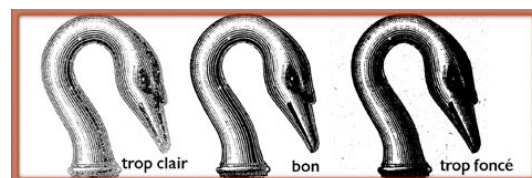
- soit acquérir directement la gravure en niveaux de gris,
- soit convertir un fichier bitonal en niveaux de gris, en réduisant la résolution pour l'adapter à l'usage final

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.



Gravures du cygne

9. Numérisation des imprimés



Introduction

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.

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

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 grosseurs 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.



Image 18 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

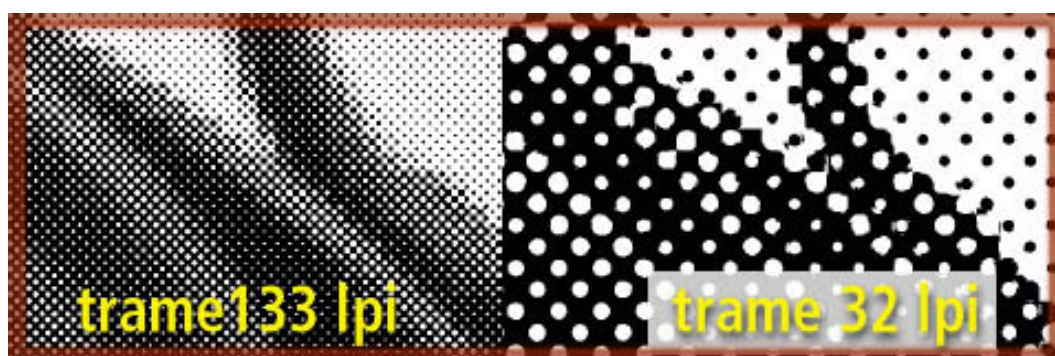


Image 19 Principe de la similigravure (2)

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.

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.

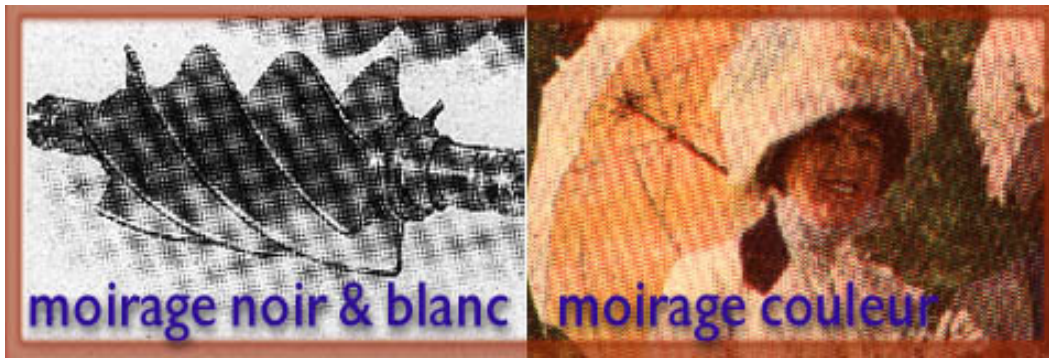


Image 20 Moirage

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.

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 similigravure 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

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 Hèlène Mégret et Robert Minder- impression Braun à Mulhouse)

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**).

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

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

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.

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

Page numérisée en niveaux de gris

Rappel !

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

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 technique (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

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*

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.

*Première couche*

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 (préresse). 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.

*Page finie*

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



Introduction

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 palisement 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.

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.

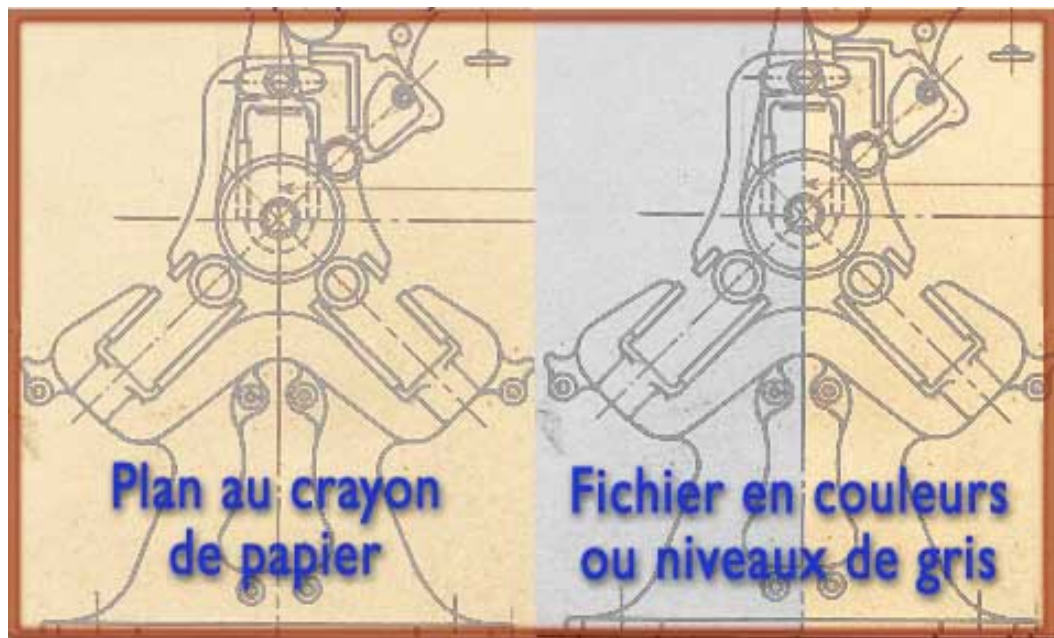


Image 21 Plans tracés à l'encre de chine.

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).

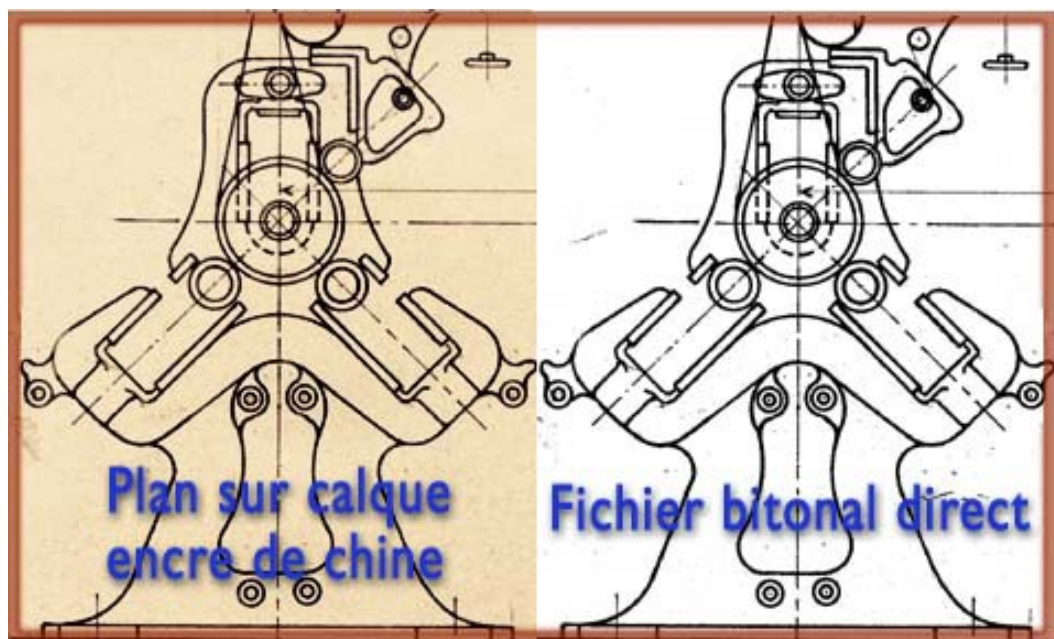


Image 22 Plans coloriés

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.

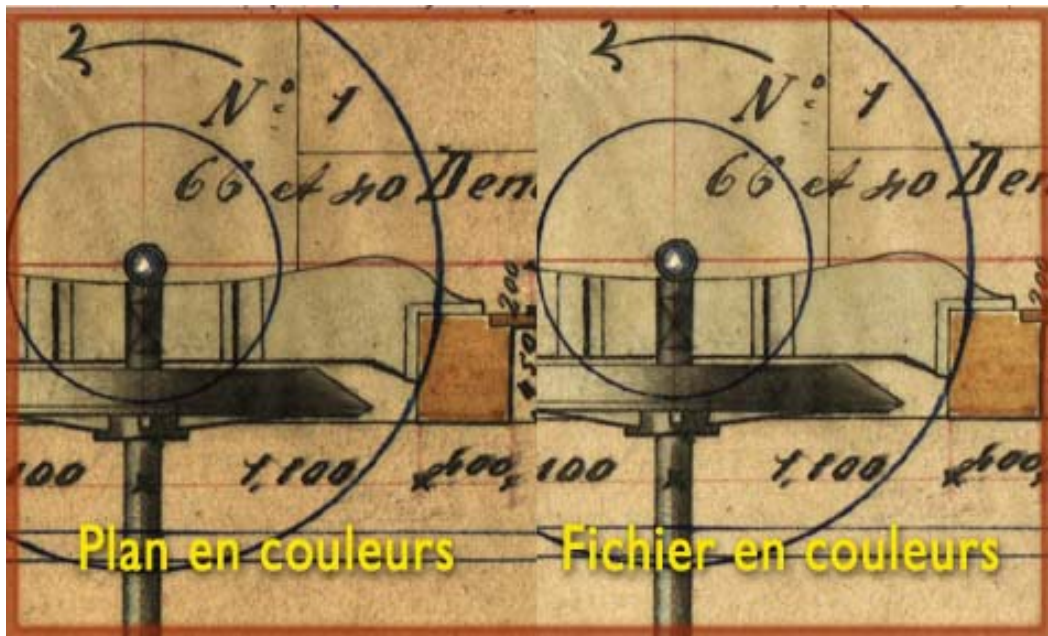


Image 3 Plan et fichier en couleurs

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.

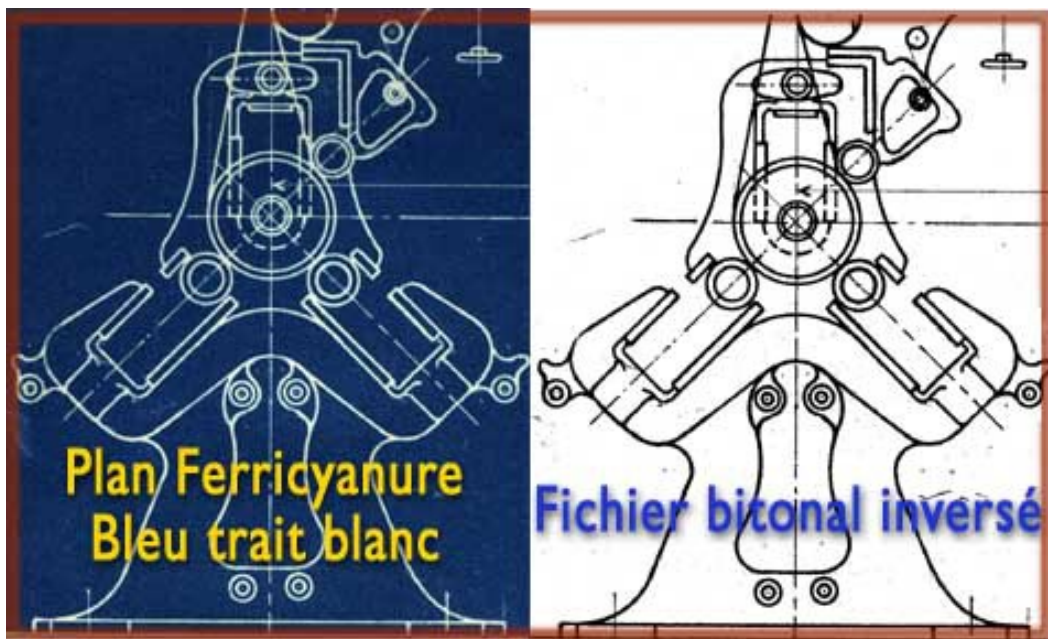


Image 23 Tirages au ferricyanure

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.

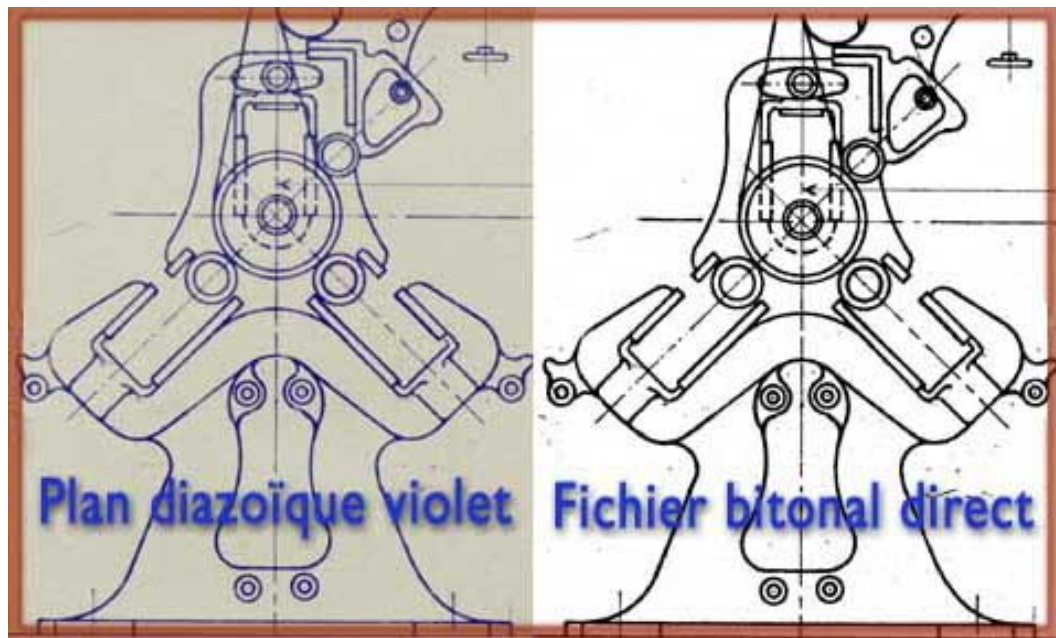


Image 2 Tirage diazo

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.

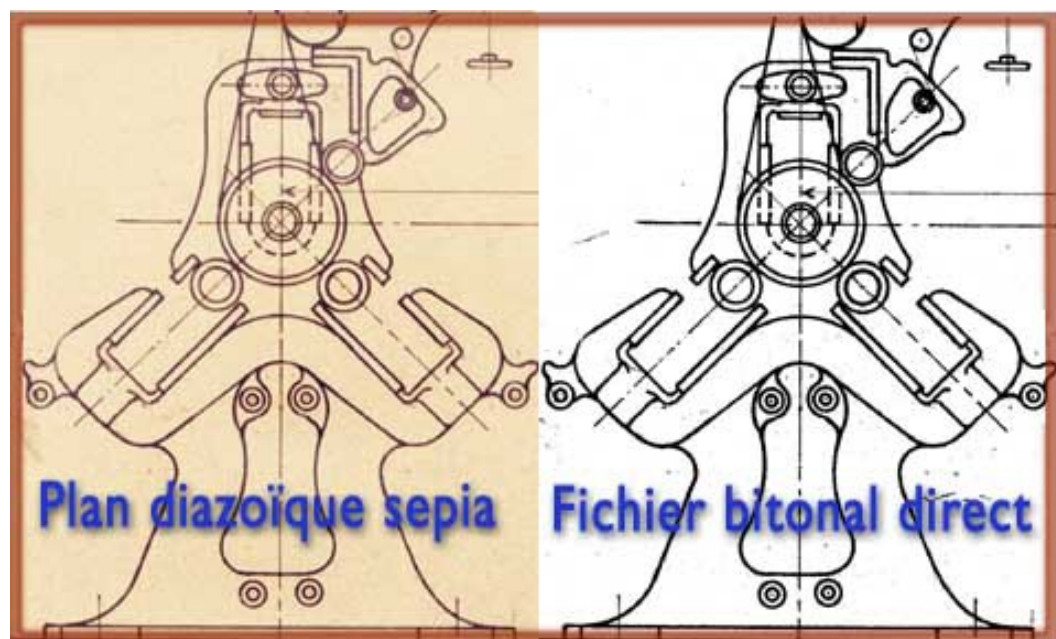


Image 3 Tirages diazo sepia



11. La post-production

Introduction

Les mécanismes d'acquisition en série
Le réglage et l'étalonnage des machines
La création d'un original brut
Le contrôle qualité et le récolement
La mise sur un support pérenne et protégé
La création de la copie de travail réversible
Les retouches de valeurs
La mise à la taille d'impression des négatifs mini format
Les règles de restitution des plaques photographiques
La confirmation des transformations par le personnel habilité
L'aplatissement des fichiers maîtres et l'enregistrement définitif
L'entrée des métadonnées en masse et individuellement

1. La post-production

99. Solutions de numérisation en tableaux



Introduction

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.

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 similigravure.

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.

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 similigravure 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.

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.

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.

11.1.4. Documents figurés en similitravure

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 similitravure en zinc.



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.

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».

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 ténus 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, gravures sur cuivre	conserver la gravure avec les défauts et le grain du papier, mais en altérant la finesse du trait à l'impression	couleurs RVB	numériseur à plat jusqu'au format A2 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

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.

11.2.3. Photographies, négatifs, diapositifs

Nature du document	Qualité principale à préserver	Mode conseillé	Moyen d'acquisition
Photographies en noir et blanc sur papier, simple ou retouchée à l'aérographe.	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...	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 en couleurs	Les photos en couleur dégradées sont parfois à retoucher pour diminuer les dominantes dues au vieillissement	Couleur RVB	Numériseur à plat.
Photographies négatives ou positives sur plaques de verre ou sur film 9x13 cm et au-delà, monochromes	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	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	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 toutes les nuances et contrastes	Couleur LAB	Numériseur à négatifs, ou numériseur à 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

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 coloriés au lavis technique, ou tirages coloriés	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)
--	--	-------------	---

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 et objectif	Résolution recommandée du fichier d'utilisation	Description et Mode de colorimétrie	Format d'enregistrement mono – page
Gravures ou similigravures à préserver, fichier source (catégorie_4)	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)	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)	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)	150 dpi	Mode bitonal (Bitmap)	.tif avec compression LZW dans tous les cas
Textes manuscrits à lire en ligne (catégorie_1)	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))	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) Reproduction de similigravures pour réimpression	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)	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

Introduction

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

1. 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>

- ⁴ <http://www.konicaminolta.fr/>

- Doc. de numériseur de livres couleur Digibook 6002

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

- Doc. de numériseur de livres couleur digibook 10 000

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

⁴. <http://www.konicaminolta.fr/>

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>

2. 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 Figuier", 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+Figuier&p=1&lang=fr>⁵

- 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 (catégorie2) 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.

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

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

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

Associer les concepts et les énoncés suivants.

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

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

Parmi les énoncés suivants, lesquels sont-ils vrais ?

- ☐ 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
- ☐ 2. Il existe une pratique uniforme pour la numérisation de tous les types de documents
- ☐ 3. On peut transformer un document de haute résolution en basse résolution, mais pas l'inverse
- ☐ 4. Les formats d'enregistrement sans pertes de données doivent être utilisés pour constituer une banque de fichiers source
- ☐ 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

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

- ☐ 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
- ☐ 7. Le format PDF, bien que populaire, est à déconseiller puisqu'il s'agit d'un format propriétaire
- ☐ 8. Les formats les plus usuels sont les formats libres
- ☐ 9. De tous les supports de stockage, ce sont les supports informatiques qui ont la plus longue longévité

Exercice 5

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

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.

d. Vous optez pour la lisibilité tout en acceptant une certaine perte de qualité au niveau de la définition.

b. Vous optez pour la plus haute résolution possible.

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.

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

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.

b. GIF	c. PNG	a. JPEG	d. PDF
--------	--------	---------	--------

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

Parmi les énoncés suivants, lesquels sont-ils vrais ?

- ☐ 1. Une fois la numérisation d'un microforme terminé, on peut le jeter et utiliser la version numérique car elle est plus facilement accessible
- ☐ 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
- ☐ 3. Pour le traitement des images et leur enregistrement sur des supports physiques, il est préférable d'opter pour des unités de travail Mac
- ☐ 4. Dans certains cas, il est utile de traiter l'image après sa numérisation
- ☐ 5. La mise en ligne de photographies, documents fort recherchés, nécessite des copies à haute définition

Exercice 8

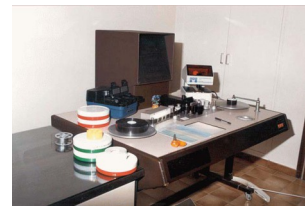
Parmi les énoncés suivants, lesquels sont-ils vrais ? (suite)

- ☐ 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
- ☐ 7. Lors de la numérisation de photographies couleur, il y a parfois lieu d'effectuer des retouches
- ☐ 8. La numérisation des gravures se fait prioritairement en mode bitonal
- ☐ 9. Le format TIFF est recommandé pour la sauvegarde des archives
- ☐ 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



Développement



Tables de montage

Ressources annexes



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



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