



HAL
open science

La numérisation des images. Un petit guide "maison"

André Pelle

► **To cite this version:**

André Pelle. La numérisation des images. Un petit guide "maison". Cahier des thèmes transversaux ArScAn, 2001, Outils et méthodes de la recherche : Les systèmes d'information, I, pp.236-241. hal-02093899

HAL Id: hal-02093899

<https://hal.science/hal-02093899>

Submitted on 9 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La numérisation des images

Un petit guide « maison »

L'idée d'une séance d'échange sur la numérisation des images a fait son chemin au début de l'année 1999, moment où les équipes et les services communs de la Maison René Ginouvès ont été nombreux à acquérir l'équipement nécessaire et à se lancer dans des numérisations de lots importants d'images. Nous avons lors de cette séance abordé les différents principes à mettre en œuvre aux étapes successives de ce travail, de l'acquisition à la diffusion, en passant par le stockage et l'archivage. Il s'agit ici de diffuser le fruit de ces échanges, sous la forme d'un petit guide maison qui fera gagner du temps et conserver quelque énergie face à ces nouveaux outils.

De l'image argentique à l'image numérique

André Pelle (UMS 844, Service photographique de la Maison René Ginouvès)

La photographie avant d'être numérique fut argentique. Si l'image numérique peut théoriquement fournir par la copie des générations d'images, identiques, proches du « clonage », il n'en est rien de la photographie argentique. Elle perd et modifie à chaque génération les informations contenues dans la génération précédente : une perte de netteté ou « piqué » est principalement due à la qualité de l'objectif utilisé ; une perte de nuance de gris ou de couleurs, vient de l'utilisation d'un film ou d'un papier de tirage qui a ses propres caractéristiques de contraste et de dominante chromatique. Un tirage photographique argentique résulte toujours de l'interprétation d'un tireur.

Il est donc préférable, dans la mesure du possible, de toujours numériser le document support, soit le négatif ou la diapositive (qu'il ne convient pas de confondre avec l'internégatif ou le duplicata).

Les divers matériels disponibles sont : les scanners à plat (pouvant être équipés d'un dos transparent), les scanners pour film de 35 millimètres (négatif ou diapositives 24 x 36) ou les appareils photographiques et caméras numériques.

Ce tableau propose le type de scanner à utiliser et la configuration « couleurs » envisageable pour un document donné.

négatif N. & B.	24 x 36	demi-teinte	scan de dia et inversion	niveau de gris
	supérieur à 24 x 36		scan à plat avec système pour document transparent	
	24 x 36	au trait	scan de dia	trait ou niveau de gris et inversion
	Supérieur	au trait	Scan à plat avec système pour document transparent	Idem
négatifs couleurs	24 x 36	demi-teinte	scan de dia.	sRGB ou RGB ou CMJN
négatifs couleurs	supérieur		scan à plat	idem

Dia.	24 X 36		Scan de dia.	sRGB ou RGB ou CMJN
	Supérieur		Scan à plat avec système pour document transparent	Idem

tirages	au trait	scan à plat	trait
	demi-teinte	idem	niveau de gris
	couleurs	idem	sRGB ou RGB ou CMJN

Acquisition des images numériques

Roland Bierry (UMR ArScAn – Proche-Orient hellénistique et romain)

Dans le cas des appareils photographiques et des caméras numériques, la résolution de l'image est fixée par le constructeur (deux à trois millions de pixels au mieux). Le problème du choix de la résolution des images se pose lorsque l'on numérise une image.

Résolution d'entrée de l'image

La résolution d'entrée de l'image désigne le nombre de points (pixels) numérisés sur une longueur d'un pouce (inch en anglais). La résolution maximale permise par les meilleurs scanners pour films est de 2700 ppi (cas du scanner NIKONSCAN 2000).

Un petit calcul très simple permet de comprendre ce que cela implique pour une diapositive ou un film 24x36 :

- taille d'une diapositive : 2,4cm x 3,6cm.
- sachant que 1 pouce = 2,54cm, 2700 ppi donnent $2700/2,54 = 1063$ ppcm (pixels par centimètre) comme résolution maximale du scanner exprimée en pixels par centimètre ;
- une diapo numérisée à la résolution maximale présentera donc les caractéristiques suivantes : nombre de pixels pour la longueur ($3,6 \times 1063 = 3826,8$ pixels) ; nombre de pixels pour la largeur ($2,4 \times 1063 = 2551,2$ pixels) ; nombre total de pixels de l'image ($3826,8 \times 2551,2 = 9762932$ pixels) ; soit près de 10 millions de pixels (à la résolution maximale du scanner).

Il est important de comprendre qu'à partir du moment où une image a été numérisée, le nombre de pixels est fixé et n'est pas modifié par les dimensions que l'on donne ensuite à l'image. Le nombre de pixels par unité de longueur est inversement proportionnel à la dimension linéaire donnée à l'image. Plus une image est agrandie, moins elle est nette (puisque le nombre de pixels est identique).

Reprenons le cas d'une diapositive 24mm x 36 mm (soit 2,4cm x 3,6cm) numérisée à 2700 ppi et donnons lui la dimension 12cm x 18cm (par exemple sous Photoshop). Le rapport de taille est de 5 (en effet $12 = 5 \times 2,4$ et $18 = 5 \times 3,6$). La résolution de l'image dans la dimension 12cm x 18cm est alors $2700/5 = 540$ ppi (résolution de scannage de la diapo d'origine divisée par le rapport de taille). Vérification :

- 540 ppi donne $540/2,54 = 212,6$ ppcm (pouce par centimètre) ;
- nombre de pixels pour la longueur : $18 \times 212,6 = 3826,8$ pixels ;
- nombre de pixels pour la largeur : $12 \times 212,6 = 2551,2$ pixels ;
- nombre total de pixels de l'image : $3826,8 \times 2551,2 = 9762932$ pixels.

On retombe sur le même nombre de pixels pour cette image 18cm x 12cm que pour la diapo d'origine numérisée à 2700 ppi.

En résumé :

Pour avoir la résolution à une dimension d'image désirée, il faut diviser la résolution de scannage de l'image d'origine par le rapport de taille entre l'image désirée et l'image d'origine. Dans le cas d'une diapo numérisée à 2700 ppi, on obtient le tableau suivant :

Taille de l'image (cm)	Résolution (ppi)	Rapport de taille
2,4 x 3,6	2700	1
8 x 12	810	3,33
12 x 18	540	5
18 x 27	360	7,5
20 x 30	324	8,33
30 x 45	216	12,5

De même, pour toute résolution de scannage d'une diapo 24 x 36, on peut trouver la résolution pour une taille d'image donnée.

Résolution de sortie

La résolution de sortie désigne le nombre de points par pouce qu'un périphérique de sortie (écran ou imprimante) est capable de reproduire. Intéressons-nous aux imprimantes. Pour obtenir de bons résultats, il faut utiliser une résolution d'image proportionnelle — mais non identique — à la résolution de l'imprimante.

Type d'imprimante

IMPRIMANTES LASER

Dans la pratique, on obtient de bons résultats pour une imprimante laser 600 ppi avec une résolution d'image de 150 à 200 ppi.

IMPRIMANTES À JET D'ENCRE

Dans la pratique, on obtient de bons résultats pour une imprimante à jet d'encre 720 ppi avec une résolution d'image de 240 ppi ; pour une imprimante-jet d'encre 1440 ppi, on obtient de bons résultats avec une résolution d'image de 360 ppi.

IMPRIMANTE LASER 600 PPI

Prenons l'exemple d'une image imprimée de format A4 (21cm x 29,7cm) ; en tenant compte des marges, la dimension d'image, homothétique d'une diapositive, la plus proche est 18cm x 27cm (rapport 7,5) ; comme on obtient une excellente sortie avec 200 ppi de résolution d'image, cela entraîne $200 \times 7,5 = 1500$ ppi comme résolution d'entrée pour la diapo.

Donc, pour obtenir une bonne image A4 sur une imprimante laser 600 ppi, on peut se contenter de scanner la diapo (ou le négatif) à environ 1500 ppi.

IMPRIMANTE JET D'ENCRE COULEUR 720 PPI

La dimension maximale d'une image de qualité est obtenue en divisant la résolution maximale de la diapo d'entrée (2700 ppi) par 240 ppi (résolution nécessaire pour obtenir de bons résultats avec ce type d'imprimante), soit $2700/240 = 11,25$ (rapport de taille), puis en multipliant la dimension de l'image d'entrée par le rapport de taille, soit $2,4 \times 11,25 = 27$ et $3,6 \times 11,25 = 40,5$. Cette dimension de 27cm x 40,5cm correspond en gros à un format A3 (29,7cm x 42cm) en tenant compte des marges. On peut donc imprimer une image A3 (avec marges) de bonnes qualités à partir d'une diapo numérisée à la résolution maximale du scanner.

IMPRIMANTE JET D'ENCRE COULEUR 1440 PPI

La dimension maximale d'une image de qualité est obtenue en divisant la résolution maximale de la diapo d'entrée (2700 ppi) par 360 ppi (résolution nécessaire pour obtenir de bons résultats avec ce type d'imprimante), soit $2700/360 = 7,5$ (rapport de taille), puis en multipliant la dimension de l'image d'entrée par le

rapport de taille, soit $2,4 \times 7,5 = 18$ et $3,6 \times 7,5 = 27$. Cette dimension de 18cm x 27cm correspond en gros à un format A4 (21cm x 29,7cm) en tenant compte des marges.

Ce que l'on peut donc imprimer

On peut donc imprimer une image A4 (avec marges) de bonnes qualités à partir d'une diapo numérisée à la résolution maximale du scanner. Certaines imprimantes à colorants nécessitent jusqu'à 400 ppi de résolution d'image.

Si les diapositives numérisées sont destinées à être imprimées (ou données à des éditeurs), il serait bon de prendre une résolution de l'ordre de 360 à 400 ppi pour l'image finale, ce qui correspond, après scannage à la résolution maximale d'une diapo, à une dimension d'image de 18cm x 27cm à 16,2cm x 24,3cm (pour 400 ppi). Si les diapositives sont uniquement destinées à être regardées sur écran, une résolution d'entrée bien moindre est suffisante (voir doc. Photoshop p. 44)

Etant donné les progrès constants des imprimantes, et à moins d'être certain de ne jamais vouloir imprimer (ou faire éditer dans un ouvrage) une image, il semble intéressant de scanner une diapositive à la résolution maximale (2700 ppi pour l'image originale 2,4cm x 3,6cm). Cela permet des sorties sur diverses imprimantes à des dimensions allant jusqu'au format A4 (ou même A3 selon l'imprimante) avec une bonne définition.

Résolution de trame, notion de linéature

Lorsqu'une image est destinée à être imprimée (insertion dans une revue ou un ouvrage), il faut tenir compte de la résolution de trame utilisée par l'imprimeur. La résolution de trame, ou linéature, désigne le nombre de cellules par pouce d'une trame en demi-teinte et se mesure en lignes par pouce (lpp). Pour obtenir une image en demi-teinte de qualité optimale, il faut utiliser une résolution d'image une et demi à deux fois la résolution de trame.

La plage de résolution acceptable dépend de la linéature du périphérique de sortie. En connaissant la linéature de celui-ci et la dimension de l'image finale, on peut déterminer directement la résolution de numérisation de l'image d'origine à l'aide du logiciel Photoshop. (Voir p. 50, 52-53 doc. Photoshop).

Le tableau suivant donne quelques valeurs de linéature habituellement utilisées par les imprimeurs.

Linéature (lpp)	Utilisation
65	Impression de bulletins
85	Impression de journaux
133	Impression en quadrichromie
177	Impression de livres d'art

Si l'on prend la linéature la plus élevée (livres d'art) on obtient une résolution d'image de $177 \times 2 = 356$ ppi, soit environ 360 ppi, ce qui correspond, pour une diapo numérisée à la résolution maximale de 2700 ppi (en 2,4cm x 3,6cm), à une dimension d'image finale de 18cm x 27cm.

Conclusion

Toute diapositive ou négatif numérisé et destiné à être, un jour ou l'autre, imprimé ou édité dans un ouvrage, doit être numérisé au maximum de résolution du scanner (2700 ppi pour 2,4cm x 3,6cm) afin d'obtenir les meilleures qualités d'impression pour des images de taille raisonnable (entre 160 dpi et 360 dpi).

Pour un simple examen sur écran ou envoi sur le Web (et pour ces usages uniquement) on peut se contenter de résolutions d'entrée bien moindres (<72 dpi).

Le tableau suivant indique la résolution de scannage de la diapositive (ou du négatif) d'origine en ppi pour obtenir de bons résultats en fonction de la dimension de l'image finale désirée et du type d'imprimante utilisée.

Type d'imprimante	Dimension de l'image finale désirée				
	2,4 x 3,6	8 x 12	12 x 18	18 x 27	20 x 30
Laser 600 ppi	200	666	1000	1500	1666
Jet d'encre 720 ppi	240	800	1200	1800	2000
Jet d'encre 1440 ppi ou impression d'art	360	1200	1800	2700	-

Le cas des scanners à plat

Tous les raisonnements précédents sont valables pour les scanners à plat, mais habituellement, l'image de départ est beaucoup plus grande qu'une diapositive. On peut donc se permettre de scanner avec une résolution d'entrée beaucoup plus faible que 2700 ppi (les scanners à plat ayant d'ailleurs souvent un maximum de résolution de 1200 ppi).

- Exemple : pour une image de dimension 18cm x 27cm (format A4 avec marges), il suffit de scanner à 360 ppi pour pouvoir éventuellement l'imprimer sur une imprimante-jet d'encre couleur 1440 ppi ou pour qu'elle puisse être éditée dans un livre d'art dans cette dimension ; pour une impression sur une imprimante laser 600 ppi, il suffit de numériser à 200 ppi ; si l'image d'origine est plus petite (par exemple 8cm x 12cm), le rapport de taille étant de 2,25 ($18/8 = 2,25$ et $27/12 = 2,25$), il suffit de numériser à $200 \times 2,25 = 450$ ppi pour l'imprimante laser (et pour un image finale de 18cm x 27cm) ou à $360 \times 2,25 = 810$ ppi pour une imprimante-jet d'encre couleur 1440 ppi ou une impression d'art (toujours pour une image finale de 18cm x 27cm).

Stockage et sauvegarde des fichiers d'images

Marie-Élisabeth Boutrou (UMR ArScAn - ESPRI)

Une fois numérisées, les images doivent être stockées et archivées de façon à rester disponibles dans des formats qui permettent à la fois de garder les informations les plus proches possibles du document original et, d'autre part, de conserver des archives d'un poids mémoire acceptable.

Il existe actuellement trois formats : *tif* ou *TIF* ou *TIFF*, format universel, reconnu par toutes les applications qui peut ou non autoriser une compression de l'image numérisée. Dès l'origine, ce format est lisible sous MAC et PC ; il semble autorisé par les plates-formes Linux et UNIX. L'ouverture et la fermeture des documents n'entraînent pas de pertes d'informations dans l'image. La compression n'est pas paramétrable.

- EPS** : format reconnu par la plupart des logiciels de dessin vectoriel (Illustrator par exemple). C'est le format qui permet le minimum de pertes d'informations lors du stockage. Cependant, il n'est reconnu que par un nombre limité d'applications. Photoshop et Illustrator le reconnaissent, dans la version Mac aussi bien que dans la version PC. Des logiciels moins complets risquent de ne pas pouvoir lire ce format.
- JPG ou JPEG**. Ce format d'image présente l'avantage d'être universellement reconnu (norme pour les images fixes), et de devenir peu à peu l'un des deux formats d'usage avec GIF pour les documents HTML (format des documents du Web). La sauvegarde au format JPEG implique la compression des fichiers, compression qui reste paramétrable et permet pour l'image de référence une réduction du poids final de l'image qui varie de 1 à. La compression de l'image provoque cependant une perte d'informations graphiques qui peut être plus ou moins acceptable selon l'origine du document, sa disponibilité, l'usage qu'on veut en faire, le type de sortie qu'on envisage. Il semble en outre que l'ouverture et la fermeture successives d'images jpeg entraînent une perte progressive d'informations graphiques. Il importe donc de garder une image jpeg sur un support non modifiable (cédérom non RW par exemple), pour archiver une version optimale de l'image et de la doubler avec un autre format, Photoshop ou TIFF.

- *GIF*. Le second format des images numérisées destinées à être proposées sur Internet. Son utilisation est surtout recommandée pour les images animées.
- Les formats dits « *propriétaires* ». Ce sont ceux qui sont proposés par les logiciels de retouche graphique. Leur inconvénient principal réside dans le fait qu'ils ne seront le plus souvent reconnus que par l'application qui les a créés. Leur utilisation ne peut donc être que locale. Les propriétés de ces logiciels varient évidemment d'un logiciel à l'autre et posent quelquefois le problème de la lecture des fichiers quand on change de plate-forme. Le format BMP, par exemple spécifique des applications graphiques de Windows, sert à créer les fonds d'écrans dans le monde PC. Il ne sera lu par un Mac qu'après traitement par un convertisseur graphique. Le stockage d'une image rare, ou dont on a perdu la disponibilité après numérisation, ne doit donc pas être fait dans ces formats.

Nous numérisons des images, mais nous savons que nous devons leur associer, sous cette forme comme lorsqu'elles étaient sur les supports traditionnels, des informations textuelles enregistrées dans des banques de données documentaires. Ces banques ont à la fois une visée de gestion et de recherche des informations relatives au document lui-même ou aux « objets archéologiques » qu'il représente. Les recherches précédentes ont montré qu'il fallait réfléchir dès le départ aux choix d'information en fonction des divers usages des images dans la discipline et des différents besoins des usagers. Elles ont également montré que ces banques sont lourdes à constituer, mais nécessaires pour améliorer la qualité du traitement d'information lors des recherches. Enfin, à l'heure de la multiplication des bornes interactives, des cédéroms et des pages multimédias sur Internet, on réalise la valeur de ces documents pour la valorisation des recherches et l'on sait qu'il est important de réfléchir à leur mise en consultation (cf. la bibliographie).

Bien sûr, il s'agit d'un état des savoir-faire qu'il faudra réactualiser.

Éléments bibliographiques

Un site web permet de calculer la taille des images numériques par rapport aux capacités de stockage : http://www.ccr.jussieu.fr/urfist/image-numerique/im_age.htm