

Quelques techniques de visualisations de contrôle pour la numérisation massive

Rodrigo Almeida, Pierre Cubaud

*Centre d'études et de recherche en informatique (CEDRIC).
Conservatoire national des arts et métiers (CNAM)*

Mots-clés : numérisation massive, visualisation d'images, contrôle qualité, bibliothèques numériques

Keywords: massive digitization programs, visual quality control, digital libraries, image browsing and visualization

Résumé : Les programmes de numérisation de masse ont besoin de nouvelles techniques de visualisation adaptées pour le contrôle qualité. Nous décrivons quelques prototypes fonctionnels d'interfaces fluides pour un logiciel permettant l'inspection rapide de la conformité de grands lots de numérisations d'ouvrages.

Abstract : Massive digitization programs need massive visualization techniques for quality control. We describe some functional prototypes of a fluid interactive environment enabling a rapid inspection of pages conformity for large batches of digitalized books.

1 Introduction

En matière de bibliothèques numérisées, une dimension nouvelle a été atteinte fin 2005 avec l'annonce de projets très ambitieux : l'accord de Google avec plusieurs très grandes bibliothèques américaines (en particulier avec l'Université du Michigan qui atteint le million de titres en ligne début 2008) et la mise en place d'Europeana, une fédération de bibliothèques européennes (6 millions de titres prévus en 5 ans). Chez les sous-traitants qui effectuent la numérisation, un opérateur peut atteindre la vitesse de 700 p./h. Ainsi, pour un centre de numérisation d'une dizaine de numériseurs, la productivité est de l'ordre d'un millier de volumes traités par semaine. Avec de telles charges, il est clair que le processus de numérisation doit être soigneusement mis au point. Malheureusement, la productivité chute pour les collections patrimoniales du fait de leur hétérogénéité (reliures pincées, existence d'hors-textes, pagination erratique, tables défectueuses, etc.). De fait, la dissimilitude entre les

volumes et les particularités de chacun imposent un plus grand nombre d'intervention humaine pour le paramétrage du scanner. Le projet DEMAT-FACTORY, une collaboration regroupant pour trois ans des entreprises de numérisation (A2IA, Banctech, SAFIG, Themis) et des laboratoires (CNAM, ESIEE, Univ. Paris 6), s'est donné pour but d'étudier ces problèmes.

Une des pistes explorées consiste à évaluer l'apport de techniques récentes de visualisation d'information pour les interfaces de contrôle qualité. Ce domaine n'a, à notre connaissance, fait l'objet d'aucune étude systématique. Les logiciels actuellement en usage au sein du consortium sont basés sur les techniques usuelles d'association de vignettes et de vues de page en gros plan (fig. 1). Ces applications, à la fois conçues selon la logique de la visualisation de photos et non destinées à la visualisation de milliers d'images, n'offrent que peu de modalités d'affichage prenant en compte les particularités graphiques des pages des documents.

Nous allons d'abord présenter la problématique du contrôle de qualité visuelle des images. Ensuite, nous indiquerons quelques aspects des techniques de visualisation traditionnelle qui les rendent peu adaptées à ce type de tâche. Puis, nous présenterons quatre techniques de visualisation conçues pour ce contexte de travail que nous avons explorées.

2 Contrôle qualité d'un lot de pages numérisées

Lorsque les images sont livrées par les prestataires de numérisation, on peut identifier différents types d'erreur, quelques-uns assez imprévus. De fait, lorsqu'on numérise des ouvrages rares et difficilement accessibles ailleurs, le contrôle qualité devrait mériter encore plus d'attention afin que les documents numérisés soient suffisamment fiables pour être consultés à la place des originaux. Certains lecteurs trouveraient peu utilisables des documents numérisés dont quelques pages ou planches sont absentes. De plus, la ré-insertion de ce volume dans un nouveau train peut paraître moins intéressante que la numérisation d'un titre pas encore numérisé. La phase de contrôle qualité est donc essentielle afin de découvrir des images manquantes, les pages absentes dans le volume original (et non détectées avant l'envoi du volume), des prises de vue inexploitable, etc. Cette phase de contrôle occupe de fait une place importante dans les chaînes de traitement des projets de numérisation [1].

Certains aspects du contrôle qualité peuvent être validés de façon automatique ou quasi-automatique. Par exemple, les métadonnées de chaque volume, les noms donnés aux fichiers image, l'intégrité de ces fichiers, la résolution des images, etc. [2]. Il est cependant recommandable de procéder à un contrôle visuel des images livrées [2].

Lors du contrôle visuel, on cherche des erreurs qui ne sont pas identifiables par les procédures automatiques. Ces erreurs peuvent être classées selon le moment où ils sont survenues : avant, durant ou après la prise de vue. Le tableau 1 présente une liste non-exhaustive des problèmes que l'on peut trouver lors d'un contrôle visuel des pages numérisées.

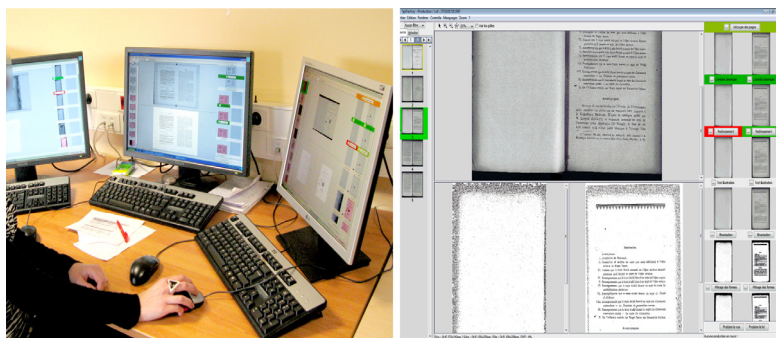


Figure 1. Contrôle qualité courant dans Demat-Factory : espace de travail et interface.

Phase	Type de problème
avant	<ul style="list-style-type: none"> - pages manquantes - pages abîmées (pliées, déchirées, tachées) - mauvais volume (mauvaises informations sur la reliure)
durant	<ul style="list-style-type: none"> - volume peu ouvert - volume mal orienté - pages mal tournées - planches non dépliées - doigts ou d'autres objets couvrant la page - images floues - bordures du <i>setup</i> trop larges - artefacts numériques - variation de la couleur parmi les pages d'un même volume - illisibilité des caractères/dessins les plus petits
après	<ul style="list-style-type: none"> - images « binarisées » trop foncées ou trop claires - mélange d'images issues de volumes différents - images répétées - images manquantes - incohérence entre le n° de la page imprimé et celui déclaré dans le fichier - incohérence entre le n° du vol. imprimé et celui déclaré dans le répertoire

Tableau 1 : Type de problème de numérisation en fonction du moment où ils apparaissent.

Quelques-uns des erreurs décrits dans le tableau 1 (pages répétées, incohérence entre le numéro imprimé sur la page et le numéro dans le nom du fichier, page mal orientée) pourraient éventuellement être détectés avec l'aide des techniques de vision par ordinateur. Cette approche automatique est plutôt complémentaire à l'approche visuelle. Tandis que la première permet de détecter vite des erreurs « modélisables », un avantage clair de la seconde est de permettre que l'on décèle des erreurs qui n'étaient pas « prévues » ou qui sont plus facilement détectées par l'œil humain.

Dans un contrôle qualité traditionnel, un petit échantillon de pages numérisées est examiné et on cherche des problèmes récurrents. Les vignettes sont rapidement balayées et quelques-unes sont chargées en taille normale. Par ailleurs, dans une chaîne non de numérisation non-massive, le volume physique original peut être facilement accessible par le personnel de la bibliothèque numérique, qui connaît ses particularités, et qui peut être confronté à son équivalent numérique. En revanche, dans un programme de masse, des connaissances sur le document numérisé, ou en cours de numérisation, sont réduites, voire inexistantes, à l'exception de celles explicitement déclarées dans le cahier des charges. De même la comparaison des images numérisées avec l'original physique est infaisable.

Nous considérons ainsi que, pour compenser les manques d'indices plus riches sur la pertinence de la numérisation, des chaînes de numérisation massive devraient adopter une stratégie de *visualisation exhaustive*, et non pas *par échantillon*. On devrait pouvoir rapidement balayer *toutes les images* et pouvoir passer sans effort à une observation détaillée de toutes celles qui attirent son attention.

3 Vue globale et vue détaillée

Les applications traditionnelles d'affichage d'images sont principalement conçues pour la gestion et la visualisation de photos personnelles, plutôt que pour la visualisation de pages numérisées. L'utilisation des pixels de l'écran (« *screen real estate* ») n'est pas optimale. De plus, passer d'une photo à l'autre peut prendre du temps lorsqu'on souhaite en voir un grand nombre. On accorde plus d'importance à la notion d'élément individuel qu'à la notion d'un ensemble d'éléments disposant de traits en commun (comme c'est le cas des pages numérisées issues d'un même volume). Or, dans la visualisation de pages, l'important est de voir chaque image dans son « contexte », c'est-à-dire entourée par les pages « voisines » et à côté des pages issues des volumes d'un même lot, et de pouvoir en visualiser un grand nombre avec un minimum d'encombrement. Dans un bon nombre d'applications, une interface pour la visualisation et la navigation

des vignettes existe, bien qu'elle soit souvent présentée comme un mode de consultation auxiliaire. Dans ce type d'interface, la priorité est donnée pour la présentation de la page dans sa totalité.

L'interface *Space-filling Thumbnails* (SFT) propose d'exploiter davantage la présentation de vignettes qui miniaturisent les pages [3]. Cette interface présente toutes les pages d'un livre distribuées de façon matricielle où chaque page a une position fixe par rapport aux pages voisines. La matrice a toujours la même configuration (la réduction de la fenêtre réduit la taille de toutes les cellules sans changer les nombres de lignes et de colonnes), ce qui permet à l'utilisateur de mémoriser la position relative d'une vignette et de la revisiter facilement. Par ailleurs, cette interface ne dispose d'aucun mécanisme de défilement. De ce fait, toutes les pages (d'un même document) sont toujours simultanément visibles sur la fenêtre. Pour cela, la taille de chaque vignette décroît à proportion que le nombre total de vignettes s'accroît. Le désavantage de ce changement d'échelle automatique est le fait que les vignettes deviennent microscopiques lorsque le volume est doté d'un grand nombre de pages. Outre la visualisation globale via un affichage « miniaturisé » du volume, le passage à un affichage de la page en haute résolution est une fonctionnalité également importante. Certains problèmes (comme la lisibilité des caractères du texte ou la netteté des illustrations) ne sont identifiables que lorsqu'on visualise l'image en « taille réelle ». Les mécanismes de zoom sont variés : le zoom peut agrandir toute la page qui occupera toute fenêtre (comme dans *Adobe Photoshop*), il peut agrandir toutes les photos présentées dans la fenêtre en même temps (comme dans *Picasa*) ou il peut agrandir temporairement qu'une photo en conservant au fond les autres images (comme le zoom contextuel de SFT). La transition entre les différents niveaux d'échelle peut aussi être soit discrète (comme dans *Adobe Acrobat Reader*), soit continue (comme dans *Photomesa* [4]). Les applications, dites « zoomables » comme *Picasa* et *Photomesa* présentent un bon compromis entre les vues globale et détaillée.

4 Visualisation par réduction extrême

Nous avons commencé à explorer les apports de la visualisation par vignettes en produisant des miniatures pour les périodiques de « La Nature »¹. Ce périodique est le plus visité de la bibliothèque numérique du CNAM (Conervatoire Numérique, CNUM) et l'un des plus volumineux. Il compte 32.500 pages, avec environ une gravure toutes les deux pages ; actuellement 32 années (soit 1696 fascicules reliés dans 65 volumes) de cette publication sont disponibles en ligne. À l'aide de la

¹ <http://cnum.cnam.fr/redirect?4KY28>

bibliothèque de programmes *NetPBM*.², nous avons produit des miniatures, chacune fait 10 par 16 pixels, pour chaque page de tous les volumes de cette collection.



Figure 2 : Les 32.500 pages de La Nature affichées (chaque page fait 10 x 16 pixels) sur une même image (8000 x 600 pixels). Chaque ligne correspond à un volume de fascicules reliés. Détail sur la fin des volumes dans les années 1890.

Nous avons ensuite produit un script Processing³, pour que les pages d'un même volume soient affichées sur une même ligne (fig. 2). Cette configuration fournit des pistes visuelles très intéressantes. On peut, par exemple, observer : la longueur relative des volumes les uns par rapport aux autres, les pages typiques qui apparaissent en fin de volume et celles qui apparaissent en début de volume, la régularité des pages illustrées, si un volume est plus illustré que les autres. Nous pouvons aussi voir que dans le volume 34, les suppléments qui apparaissent en fin du volume sont visiblement manquants (ce que nous ignorions avant cette expérience !). Toutes ces indications sont utiles pour repérer des lacunes dans les volumes, qu'elles soient dues à la numérisation ou aux volumes physiques fournis à la numérisation.

Cette visualisation statique met en évidence les points forts d'avoir une vue globale sur une collection numérisée, principalement pour vérifier son intégrité et comprendre quels sont les types de page qui la composent. La visualisation interactive d'un ensemble de telle grandeur, proche du giga-pixel, est actuellement hors de portée de la puissance de calcul des ordinateurs standard. Nous avons commencé à explorer des techniques interactives, sur des ensembles de pages plus réduits.

5 Interface Grille de détails

Nous avons développé l'interface « Grille de détails » pour que les pages numérisées puissent être vérifiées de façon *exhaustive* et non pas par échantillons. Il s'agit d'une interface Web, utilisée au sein du projet CNUM, qui affiche les vignettes des pages numérisées dans une

² <http://netpbm.sourceforge.net/>

³ <http://processing.org/>

disposition matricielle (fig. 3). Dans cette interface, hormis l'espace d'un pixel entre les vignettes, tous les pixels de la fenêtre sont utilisés pour afficher des zones des pages numérisées. Quelques aspects visuels de toutes les pages d'un volume de mille pages peuvent ainsi être vérifiés en quelques minutes. En outre, les vignettes sont partiellement recoupées afin de prioriser la zone de la page contenant du texte ou de l'illustration. L'affichage résultant est donc dense en informations visuelles. Enfin, le fait que les pages soient serrées les unes contre les autres, sans beaucoup d'espace entre elles, facilite la détection d'anormalités. Par exemple, une page avec du texte plus foncé que le texte des pages voisines saute aux yeux.

Modalités d'affichage. Les vignettes des pages utilisées par la Grille de détails sont générées lorsqu'on charge les images dans le serveur du CNUM. Deux types de vignette sont fabriqués à partir des images master : une miniature de la page et un détail du centre de la page. Ce détail a la même résolution de l'image qui sera diffusée dans le site.⁴ Pour la miniature, on recoupe un carré dont le côté fait la dimension la plus courte de la page et on l'échantillonne ensuite à 128 x 128 pixels. Pour le détail, on échantillonne d'abord l'image à la résolution du format de diffusion ; ensuite on recoupe, au centre cette image échantillonnée, un carré de 128 x 128 pixels. Pour la miniature des images brutes, les vignettes font 256 x 128 pixels (ces images sont recadrées le moins possible). Ces vignettes sont aussi fabriquées par lot à l'aide des programmes *NetPBM*. La grille de détails affiche les vignettes dans une page générée par du code *PHP* dont des fonctions *JavaScript* offrent quelques micro interactions.

Inspection des images de diffusion. Ce mode de visualisation est actuellement utilisé dans la vérification des images qui seront mises en ligne. Les vignettes de chaque page sont affichées dans une matrice qui se sert de tout l'espace de la fenêtre. L'utilisateur décide si, par défaut, toutes les vignettes affichent la vue « détaillée » ou la vue « miniature » de la page. Lorsqu'on passe le curseur sur une vignette, celle-ci bascule vers le mode de présentation qui n'est pas celui par défaut. En cliquant sur la vignette, on charge la page dans le format GIF de diffusion. La vue de détail permet de vérifier, par exemple, si le texte est lisible (si les caractères ne sont ni trop rongés ni trop petits) ou si le résultat du tramage des gravures est satisfaisant. Ce changement rapide et sans clic entre la vue miniature et la vue détail fonctionne comme un zoom contextuel et discret (d'un seul pas). Il permet de vérifier simultanément des indices visuels de tout un volume avant de charger les pages en taille réelle.

⁴ Il faudrait à terme ajouter une étape de reconnaissance du contenu de la page pour que le détail recoupé encadre forcément du texte ou de l'illustration et non pas du vide (comme il arrive dans certaines pages).

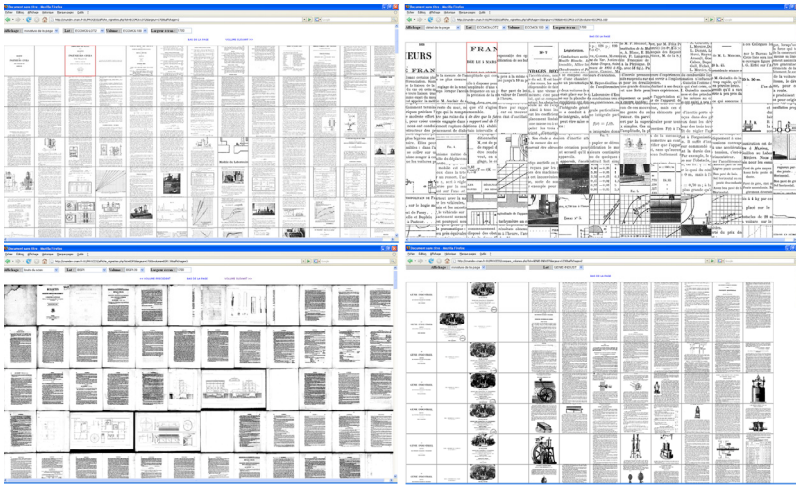


Figure 3. Copies d'écran de l'interface par Grille de détails. En haut à gauche, les vignettes affichent les pages en miniature ; à droite on voit le détail du centre des pages. En bas à gauche, les miniatures des bruts de scan ; à droite les dix premières pages des volumes d'une même collection.

Inspection des images de préservation. Cette deuxième interface présente les miniatures des images brutes en double page. Le fait d'afficher les pages impaires à côté des pages paires permet de voir si une page est manquante dans le volume original ou si le fichier image a été « perdu » lors des transferts ou des traitements des images brutes. Par ailleurs, les pages apparaissent dans l'ordre dans lequel elles ont été numérisées, ce qui facilite la vérification de l'emplacement des planches ou d'autres éléments sans pagination. Enfin, les tonalités de gris et d'autres éléments visuels, comme des traits ou des taches, communiquent une certaine homogénéité des pages et peuvent éclairer l'utilisateur sur des pages « étrangères » ayant été placées par erreur dans le répertoire du volume.

Visualisation comparative de plusieurs volumes. Cette troisième méthode n'est pas faite pour vérifier les images elles-mêmes mais pour vérifier la cohérence entre les images des volumes appartenant à une même collection. On se sert des mêmes vignettes (miniature et détail) utilisées dans la première interface et les images de chaque volume sont affichées dans une ligne.⁵ Ainsi, en balayant les lignes de haut en bas, on

⁵ Comme il serait trop lourd d'afficher toutes les images de chaque volume (certains lots font 24.000 pages), seules les 15 premières vignettes de chaque volume sont affichées.

peut voir si des différences importantes entre les volumes (ou des similarités indiquant des volumes répétés) sautent aux yeux. Cette présentation est spécialement intéressante dans la visualisation des revues qui se servent d'une même structure au fil de plusieurs volumes (comme, par exemple, La Nature). On peut identifier des volumes répétés ou des pages manquantes.

Comme le montre le tableau 2, on peut considérer ces différentes options d'affichage comme des modalités d'échantillonnage auquel l'ensemble d'un volume est soumis. Vu la difficulté pratique de visualiser *tous les pixels de toutes les pages du volume*, on échantillonne soit le nombre de pages, soit la zone visualisée de chaque page, soit la résolution de la page.

Type d'affichage	Nombre de pages	Zone observée	Résolution
Page entière	2%	100%	100%
Miniature	100%	90%	10%
Détail du centre	100%	10%	100%

Tableau 2 : Résumé des caractéristiques de l'« échantillonnage visuel » procuré par chaque méthode.

6 Mur de pages

La *grille de détails* augmente le débit de pages que nous pouvons visualiser mais elle ne fournit qu'un niveau de zoom pour chaque page. On sacrifie le paramétrage de la visualisation au profit de la simplicité d'interaction (ainsi qu'au profit d'une technologie Web facile à déployer). Lorsque l'utilisateur souhaite plus de détail, il va devoir charger l'image entière. Il s'agit donc d'une *interaction discrète* à trois pas : (1) miniature, (2) zoom serré et (3) chargement de l'image en taille réelle.

Pensant à ces limitations, nous avons développé un prototype qui pourrait restituer les détails de l'image de façon fluide et progressive. Comme dans la grille de détails, les pages sont disposées de façon serrée dans une matrice. Grâce à cette disposition, l'ensemble de pages est perçu comme une grande image. Utilisant la technologie 3D, les pages sont disposées dans un environnement où le zoom et la navigation deviennent des opérations très naturelles. Le zoom en arrière réduit la grille progressivement jusqu'à ce que l'intégralité des pages soit visible en même temps. Par le zoom en avant, l'utilisateur peut rapidement trouver le niveau de détail qui lui convient. Il peut alors défiler les pages à une

très haute vitesse grâce au dispositif isométrique utilisé pour contrôler la caméra.

6.1 Stratégies de navigation

Deux stratégies de navigation sont importantes dans le contexte d'une tâche de contrôle de qualité : nous les dénommons la « navigation global-détail » et le « balayage séquentiel ».

Dans la *navigation global-détail*, l'utilisateur peut visualiser la totalité des pages d'un coup d'œil. Cette vue globale est importante pour qu'il puisse estimer la proportion entre des pages de texte, d'illustration et des planches. Cet affichage permet aussi de voir s'il y a des pages dont les aspects visuels (niveaux de gris, taille des caractères, bordure) sont différents de la majorité des pages. Puis, une fois identifiée des « zones du mur » qui méritent d'être observées avec plus d'attention, l'utilisateur doit pouvoir rapidement s'en approcher allant, si nécessaire, jusqu'à la trame du papier.

Pour la stratégie *balayage séquentiel*, l'utilisateur peut « balayer » toutes les pages à une résolution et à une vitesse qui lui conviennent. En voyant les pages dans l'ordre dans laquelle elles ont été numérisées, il pourra plus aisément avoir des indications sur des pages manquantes (par exemple, un chapitre qui finit dans une page et la page suivante commence par du texte). De même, si on souhaite vérifier la qualité des illustrations dans un volume qui en possède beaucoup, il sera plus facile de faire défiler toutes les pages plutôt que de faire des allers-retours de zoom sur chaque illustration.

C'est pour cette raison aussi que ce défilement doit pouvoir se faire à une grande vitesse : pour que l'utilisateur puisse « sauter » les zones visuelles méritant peu d'attention. Enfin, il est important que l'utilisateur puisse positionner la caméra à une résolution donnée (affichant ou bien la totalité de chaque page, ou bien uniquement le détail de chaque page) et qu'il puisse en suite avancer en gardant fixe cette distance. C'est pensant à ces deux stratégies que nous avons créé et adapté la scène et la navigation du *Mur de pages*.

6.2 Scène et contrôle de la caméra

Les pages sont plaquées sur la face interne d'un hémicylindre (fig. 4, gauche). Lorsque la caméra se trouve au centre de la scène, toutes les pages sont visibles : le rayon du cylindre est calculé en fonction du numéro de pages à afficher et de l'amplitude du champ de vision de la caméra virtuelle. Les limites du mouvement de la caméra : en zoomant, elle peut aller jusqu'au détail de la page ; en reculant, elle doit s'arrêter au centre du cylindre. Cette disposition permet que les trajectoires rectilignes du centre du cylindre vers n'importe quelle page aient la même

longueur. En revanche, si le mur était plat, il serait plus coûteux de déplacer la caméra jusqu'aux pages qui se trouvent dans un point extrême de cette surface.

Navigation cylindrique. Dans un premier prototype du Mur de pages, il était possible de contrôler la caméra dans quatre degrés-de-liberté (DDL), à savoir, les déplacements le long des trois axes et la rotation autour de l'axe vertical. Cependant, nous avons constaté que les utilisateurs n'arrivaient pas à maîtriser facilement ces quatre DDL et cela les retardait pour atteindre les positions qui leur convenaient. La rotation de la caméra autour de l'axe Y et son déplacement le long de l'axe X jouent, dans ce type de tâche, des rôles similaires. Nous avons ainsi décidé de restreindre la navigation horizontale à la rotation autour de l'axe vertical. Alors, bien que le déplacement le long de X et de Z soit interdit, toutes les zones du mur restent accessibles et visibles (dans une topologie convexe comme celle-ci). Comme le déplacement de la caméra suit des coordonnées cylindriques, elle reste toujours parallèle au mur. L'utilisateur voit ainsi les images sans aucune distorsion de perspective. Contrairement à d'autres interfaces qui se servent des lignes de fuite pour montrer la « périphérie » d'un document [5], la non-distorsion du *Mur de pages* réduit les chances que des erreurs de numérisation soient confondues avec des effets de perspective.

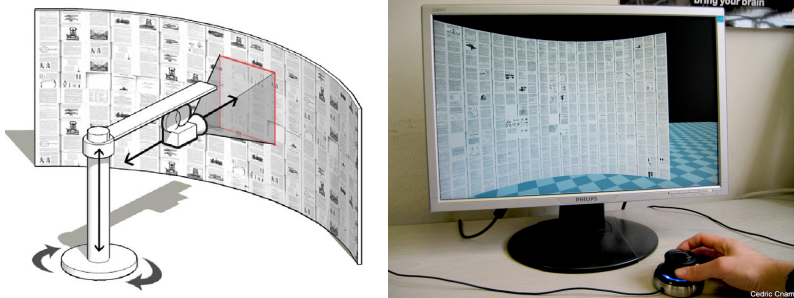


Figure 4. Interface Mur de pages. Gauche : Principe de navigation. Droite : utilisation.

Navigation sur l'axe vertical. Dans le *Mur de pages*, les pages sont disposées sur deux dimensions pour tirer un maximum de profit de l'espace de la fenêtre. Puisque les pages sont ordonnées le long des rangées, parcourir l'axe vertical procure une « visualisation aléatoire » des pages, ce qui est intéressant dans le contexte du contrôle qualité. Par exemple, si l'utilisateur zoome sur la page 7 et monte la caméra, il verra passer la page 36, puis 65, 104 et ainsi de suite. Cette navigation lui permet de visualiser rapidement un échantillon de pages du volume.

Balayage séquentiel sans rupture. L'axe horizontal, privilégié par la disposition des pages, est consacré au balayage séquentiel. Son effet est similaire à l'action de défiler rapidement un document dans une application comme *Adobe Acrobat Reader*. Cependant, l'utilisateur doit passer à la ligne du dessus lorsqu'il atteint la fin de la rangée de pages où il se trouve. Ce mouvement ponctuel de monter la caméra d'une ligne introduit une interruption dans le ballayage. C'est pourquoi nous avons considéré d'autres topologies de scène qui éviteraient cette opération. L'idéal serait que l'utilisateur puisse parcourir toutes les pages avec un même geste de contrôle pour que le contrôle porte uniquement sur la vitesse du balayage et l'attention sur l'identification des problèmes.

Dans cette perspective, un cylindre complet, où les pages sont disposées de façon hélicoïdale, produirait une bande d'images continue de la première à la dernière page [6]. Cependant, dans cette configuration il n'y a aucune point où la caméra puisse voir toute le scène (il y aurait toujours une partie du mur derrière la caméra). Par ailleurs, l'environnement serait trop homogène, sans points de repère, ce qui pourrait rendre plus difficile à l'utilisateur de retourner à des pages déjà visitées.

Comportements pour passer d'une rangée de pages à l'autre. Nous étudions deux techniques qui ont pour but de combler cette lacune de notre scène 3D. Un test d'ergonomie pourra montrer si elles sont efficaces et naturelles. Le comportement *monte-page* serait déclenché dans des conditions précises : en réalisant un mouvement panoramique à la proximité du mur de pages (une rangée de pages occupe plus de la moitié de la hauteur du *viewport*), on atteint la fin de cette rangé de pages. Le *monte-page* transporte alors la caméra à la rangée de pages d'au-dessus. Bien que cette solution impose un changement dans le sens du mouvement appliqué au dispositif d'entrée qui contrôla la caméra, elle a l'avantage de permettre que les pages soient disposées dans un ordre facile à suivre visuellement. Ainsi, il est inutile de tourner le regard vers le début de la ligne lorsqu'on atteint la fin d'une rangée. Le deuxième comportement, le *vide-de-pages*, fait que la caméra soit transportée d'un côté à l'autre du hémicylindre lorsqu'elle atteint la fin d'une ligne. L'avantage de ce deuxième comportement est de permettre que les pages soient disposées toutes dans le même ordre (croissant de gauche à droite, quelle que soit la ligne).

6.3 Implémentation du prototype

Un *SpaceNavigator* a été utilisé comme périphérique d'entrée pour le contrôle de la caméra. Il s'agit d'un sort de joystick isométrique à six DDL fabriqué par la société *3D Connexion*. Il est formé par une base lourde sur laquelle un petit cylindre (la manette) est monté. On saisit cette manette par les bouts des doigts. Les dispositifs isométriques procurent un contrôle plus efficace pour une tâche de navigation car, par une simple

modulation de la pression appliquée, on peut faire défiler toutes les pages [7]. En relâchant la manette, la caméra s'arrête.

L'application prototype a été développée en *C* et *OpenGL*. Les images master en niveaux de gris ont été légèrement recadrées pour que leurs dimensions aient un ratio de 2:1. Elles ont été ensuite six fois échantillonnées en avance dans des versions plus petites. Toutes ces versions avaient des dimensions qui étaient des puissances de deux (de 2048x1024 pixels jusqu'à 32x16 pixels). Des images de ces dimensions tirent profit de la mémoire de la carte graphique lorsqu'elles sont chargées en tant que textures. Le jeu de versions réduites de chaque image alimente le filtrage par *MIP mapping*, ce qui procure un zoom plus fluide et plus net.

Les GIFs sont chargées dans la mémoire de la carte graphique utilisant une fonctionnalité de compression de textures lors de l'initialisation de l'application. Comme les textures sont toutes résidentes, il n'y a pas de délai dû à la pagination de la mémoire virtuelle et l'animation n'est pas saccadée. En effet, dans un ordinateur doté d'une carte GeForce 9400M, nous avons chargé 500 pages et l'utilisateur peut zoomer sur n'importe quelle page en moins de deux secondes (fig. 4, droite).

6.4 Retour des utilisateurs

Plusieurs utilisateurs (20 personnes parmi bibliothécaires, professionnels de la numérisation, utilisateurs débutants et avancés) ont déjà essayé notre prototype. Il a été présenté dans un salon des techniques de la numérisation (FAN'2008, Paris) où il a été très apprécié par les personnes qui l'ont utilisé. Ceux qui travaillent dans la numérisation (ou dans la visualisation d'images en général) ont trouvé que cette interface apporte une facilité qu'ils ne connaissent pas dans d'autres applications. Certains utilisateurs, surtout ceux qui ne sont pas habitués aux dispositifs 3D, ont eu des difficultés pour contrôler les mouvements de zoom de la caméra. La forme du dispositif et son faible feed-back élastique paraissent troubler les utilisateurs qui finissent pour réaliser des mouvements imprécis. Le mouvement vertical, par exemple, est assez difficile puisque la prise pour ce type de mouvement n'est pas confortable.

Ajustement de la vitesse du panorama. Igarashi et Hinckley évoquent le problème du « débordement visuel » généré par le défilement accéléré d'un document (par exemple, une longue page Web) [8]. Cet effet est bien similaire à celui expérimenté par les utilisateurs qui ont essayé le *Mur de pages* lorsque la caméra se trouve proche des pages. Dans notre application, nous avons adopté une « navigation cylindrique ». De ce fait, la vitesse angulaire parcourue par la caméra, qu'elle soit proche du centre ou dans l'extrémité du cylindre, est toujours la même. En revanche, la longueur parcourue sera toutefois beaucoup plus grande lorsque la caméra se trouve proche au mur (c'est-à-dire loin du centre du cylindre).

Nous étudions à cette fin un *mapping* variable pour réduire la vitesse de rotation de la camera en fonction de sa distance par rapport au mur. Cela ferait en sorte que l'effet visuel généré par la rotation soit toujours le même. Un deuxième *mapping* que nous étudions est une adaptation directe de la technique de Igarashi et Hinckley, c'est-à-dire reculer la caméra (vers le centre du cylindre) lorsqu'on la tourne rapidement la caméra et la rapprocher lorsque que le mouvement se ralentit ou s'arrête.

7 Panorama de détails

Comme nous avons vu dans la *Grille de détails*, il peut être avantageux visualiser les détails de plusieurs pages en même temps. On se sert des indices disponibles dans le détail affiché pour vérifier la qualité de l'image. Dans l'interface utilisée actuellement, on pré-fabrique ces détails en recoupant le centre de la page. Il s'agit d'un « cache graphique ». Ce cache graphique est nécessaire pour que l'interface soit utilisable pour les lots de toutes tailles et aussi pour qu'on la consulte via des navigateurs différents et sans *plugin*. Cependant, la grille de détails ne nous permet pas de changer dynamiquement la position de la fenêtre de détail. On peut vouloir balayer les détails du haut de la page, pour vérifier si des notes ont été prises. Ou bien, on peut vouloir aussi voir une zone de détail un peu plus grande de chaque page. Ces changements en temps réel ne pourraient pas être pris en compte par la plate-forme de travail actuelle.

Nous avons donc exploré une première maquette d'une « grille de détail paramétrable », que nous appelons « Panorama de détails ». Il s'agit d'une fenêtre similaire à la Grille de détails, mais où chaque carré correspond à une « sous-fenêtre ». On peut imaginer que l'image-détail de chaque page est fournie par une caméra qui plane au-dessus et parallèle à la page correspondante. Dans cette maquette, l'utilisateur contrôle la position et le zoom d'une « camera-abstraite », celle-ci est associée à toutes les caméras qui planent sur chaque page de la grille. Il est ainsi possible de faire un panorama multiple qui s'affiche dans chaque sous-fenêtre de l'interface.

7.1 Paramètres de navigation

Outre le mouvement de panorama sur les pages, notre maquette permet aussi de changer la taille de la fenêtre de visualisation de chaque page. L'utilisateur peut ainsi, par exemple, voir une partie plus grande des pages à l'échelle 100%. Cependant, l'augmentation de la taille des sous-fenêtres réduit le nombre de sous-fenêtres qui peuvent être simultanément affichées sur la fenêtre principale. La solution est de considérer la fenêtre principale elle aussi comme une « super-caméra » qui plane sur une grille. On peut approcher ou écarter la super-caméra des sous-fenêtres.

Comme le montre la figure 5, on peut contrôler les mouvements de la camera-abstraite et les mouvements de la super-camera.

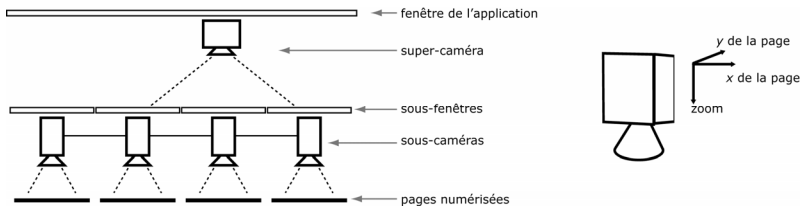


Figure 5. A gauche, schéma avec les caméras virtuelles de la maquette du Panorama de détails. Les sous-caméras sont reliées les unes aux autres et exécutent toujours les mêmes mouvements. A droite, les DDL le long desquels les caméras virtuelles se déplacent. Les sous-caméras restent toujours parallèles aux images des pages numérisées. La super-caméra reste toujours parallèle au plan des sous-fenêtres.

7.2 Interface de navigation

L'obstacle que nous avons rencontré dans cette première maquette est le grand nombre de paramètres de visualisation à contrôler, à savoir le zoom et le déplacement panoramique à la fois de la caméra-abstraite et de la super-caméra (trois DDL pour chaque caméra). Il s'agit ainsi d'une interface à six DDL. Dans cette maquette, nous avons choisi de représenter les viewports des deux caméras par des icônes qui fonctionnent comme interface de navigation. Ces représentations, à l'instar du *widget Navigator* dans *Photoshop*, peuvent être glissées et redimensionnées directement avec la souris ou par des touches du clavier. Elle a l'avantage d'indiquer quels sont la position et le niveau de zoom de la caméra par rapport à la page.

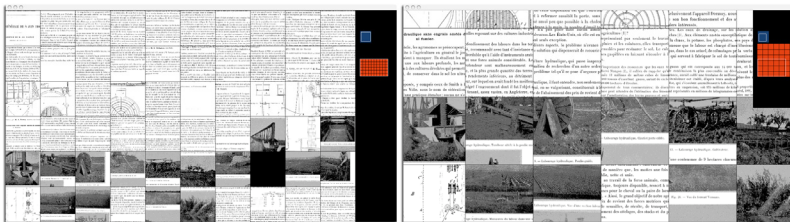


Figure 6. A gauche, schéma avec les caméras virtuelles du Panorama de détails. Les sous-caméras sont reliées les unes aux autres et exécutent toujours les mêmes mouvements. A droite, les DDL le long desquels les caméras virtuelles se déplacent. Les sous-caméras restent toujours parallèles aux images des pages numérisées. La super-caméra reste toujours parallèle au plan des sous-fenêtres.

Le contrôle de six DDL est naturel s'il s'agit des six variables qui définissent la position et l'orientation d'un objet dans l'espace, c'est-à-dire (X, Y, Z) et (R_x , R_y , R_z). Un joystick isométrique pourrait offrir une solution bureautique pour cette manipulation. Nous ne voyons cependant pas comment associer un mouvement six DDL à la métaphore de caméra que nous avons décrite. Une piste à explorer serait de se servir d'une manipulation bi-manuelle. La main gauche pourrait contrôler la super-caméra tandis que la main droite manipule la caméra-abstraite.

8 Conclusion et perspectives

Les bibliothèques numériques massives ont (ou vont avoir) besoin de procédés de visualisation également massifs. Les premières expériences décrites ci-dessus sont encourageantes dans la mesure où elles montrent que les systèmes graphiques actuels ont la puissance nécessaire pour de telles visualisations interactives. Nous avons décrit quatre techniques de visualisation, en différents niveaux de développement, qui ont pour but de favoriser la compréhension d'un lot de pages numérisées. L'interface *Grille de détails* est maintenant utilisée de manière systématique pour la revue des lots du CNUM.

Le prototype de *Mur de pages* a, quant à lui, reçu un accueil positif de la part des personnes qui l'ont testé. Ce prototype, bien qu'il soit techniquement bien avancé, ne procure pas encore de fonctions qui nous permettraient de l'utiliser dans un cadre réel de contrôle qualité. Par exemple, il n'est pas encore possible de récupérer des informations additionnelles sur une page donnée (nom du fichier, date de numérisation, dimensions en pixel, colorimétrie) et nous ne disposons pas non plus d'une interface d'annotation intégrée. Outre le développement de ces fonctions, nous souhaitons conduire une étude ergonomique qui nous aiderait à choisir un *mapping* plus efficient et confortable pour la caméra. Deux idées concernant le *mappings* ont déjà été proposées et sont en cours d'implémentation : le passage semi-automatique d'une ligne à l'autre lors d'un balayage séquentiel et l'adaptation « écologique » de la vitesse de rotation (en fonction de la distance qui sépare la caméra des pages).

D'autre part, nous souhaitons rendre le prototype robuste pour la visualisation d'un ensemble de dimensions très importantes (plus de 2000 pages). Une architecture qui prenne en compte le chargement et le déchargement dynamique des textures doit être étudiée [9]. Ils nous semble aussi que les travaux liés à la visualisation des images giga-pixel peuvent nous fournir des pistes [10]. D'autant plus que ces techniques évoluent rapidement grâce à leur popularisation via les outils Web de consultation de cartes et photos satellitaires. Un questionnaire de

chargement adapté aux parcours caractéristiques de la tâche de contrôle de qualité pourra donc être développé.

Nous allons ensuite approfondir l'analyse du travail des opérateurs de contrôle de qualité des partenaires de DEMAT-FACTORY, et continuer le développement avec une démarche de design participatif. Nous procéderons ensuite à une évaluation quantitative. Au-delà, il conviendrait d'étudier en quoi ces techniques seraient utilisables pour faciliter la navigation des utilisateurs finaux des grands corpus numérisés.

Remerciements : Nous remercions M. Emile Prior, directeur technique de SAFIG, ainsi que le personnel du centre de La Châtre. Merci également à Alexandre Topol et Pedro Alessio pour leur aide dans l'expérience du mur de page.

9 Références bibliographiques

- [1] Site Web du projet de numérisation de la bibliothèque de l'Université de Michigan, "<http://www.lib.umich.edu/news/millionth.html>"
- [2] J. Riley et K. Whitsel. Practical Quality Control Procedures for Digital Imaging Projects. *OCLC Systems & Services* 21(1), pp. 40--48 (2005)
- [3] A. Cockburn, C. Gutwin et J. Alexander. Faster Document Navigation with Space-Filling Thumbnails. *Proc. of CHI '06*, pp. 1--10. ACM Press, New York (2006)
- [4] B. Bederson. PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps, *Proc. of UIST '01*, pp. 71--80. ACM Press, New York (2001)
- [5] Y. Guiard, O. Chapuis, Y. Du et M. Beaudouin-Lafon. Allowing Camera Tilts for Document Navigation in the Standard GUI: A Discussion and an Experiment. *Proc. of AVI '06*, pp. 241--244. ACM Press, New York (2008)
- [6] A. Topol. Interaction 3D pour les paysages informationnels. Thèse de doctorat en informatique, Conservatoire national des arts et métiers, Paris, France (2002)
- [7] S. Zhai, B. A. Smith et T. Selker. Improving Browsing Performance : A Study of Four Input Devices for Scrolling and Pointing Tasks. *Proc. of INTERACT '97*, pp. 286--293, Chapman & Hall, London, UK (1997)

- [8] T. Igarashi et K. Hinckley. Speed-Dependent Automatic Zooming for Browsing Large Documents. *Proc. of UIST '00*, pp. 139--148. ACM Press, New York (2000)
- [9] D. Cline et P. K. Egbert. Interactive Display of Very Large Textures. *Proc. of VIS '98*, pp. 343--350, IEEE Press, Los Alamitos, CA, USA (1998)
- [10] J. Kopf, M. Uyttendaele, O. Deussen et M. F. Cohen. Capturing and viewing gigapixel images. *ACM Trans. Graph.*, 26(3) : 93, ACM Press, New York (2007)