

Université de Paris-8, St-Denis

Habilitation à diriger des recherches

présentée par

Pierre-Henri Cubaud

**Du texte au volume : contributions
aux bibliothèques numériques**

13 novembre 2006

M. Saleh, Imad	Tuteur
M. Beaudoin-Lafon, Michel	Rapporteur
M. Le Coadic, Yves	Rapporteur
M. Quint, Vincent	Rapporteur
M. Natkin, Stéphane	Examineur
Mme Pélachaud, Catherine	Examinatrice
M. Salaün, Jean-Michel	Examineur
M. Scholl, Michel	Examineur

Remerciements

J'ai parfois douté de pouvoir écrire un jour ces lignes, aussi c'est avec grand plaisir que j'y procède. Je voudrais en premier lieu remercier les membres du jury :

Le professeur Imad Saleh, qui a bien voulu être le tuteur - toujours disponible et encourageant - de cette habilitation à l'Université Paris 8. J'en profite également pour saluer ses collègues du département Hypermédia et les remercier collectivement de la chaleur de leur accueil.

Le professeur Stéphane Natkin, qui a patiemment réussi à créer au sein du CNAM un lieu original d'enseignement et de recherche axé sur le multimédia, et qui m'a toujours témoigné sa confiance.

Les rapporteurs : MM. les professeurs Michel Beaudoin-Lafon et Yves Le Coadic, ainsi que M. le directeur de recherche Vincent Quint. Je leur suis très reconnaissant du temps qu'ils ont bien voulu consacrer à l'analyse de mon travail ainsi que de leurs observations.

Les examinateurs : Mme la professeure Catherine Pélachaud, qui m'a très rapidement donné son accord pour examiner ce travail. M. le professeur Jean-Michel Salaün, pour son accueil à l'ENSSIB lors de mon congé sabbatique - une étape importante pour moi. M. le professeur Michel Scholl, pour ses encouragements depuis le début de mon travail sur le thème des bibliothèques numériques.

Les travaux que je présente dans ce mémoire n'auraient pas pu se concrétiser sans des collaborations : les co-auteurs de mes publications sont bien identifiables, mais je voudrais les remercier ici : Didier Girard, Alexandre Topol (multi-récidiviste), Claire Thiria, Pascal Stokowsky, Geneviève Deblock, Yvan Leclerc, Jérôme Dupire, Jean-François Haas, Rodrigo Almeida. Merci aussi à : Eric Dubreucq et aux autres membres de l'ABU. Brigitte Rozet, Françoise Chevalier, André Guillaume et les participants du projet CNUM. Stéphane Bortzmeyer, Gersan Moguerou et Joël Berthelin, les meilleurs ingénieurs système que je connaisse.

Merci encore à mes collègues du CEDRIC et du département d'informatique du CNAM, avec une spéciale dédicace pour le 55 rue de Turbigo où il fait bon (voire chaud) travailler.

Merci enfin à mes parents, Anne-Marie et Jean-Claude, qui le voulaient, et à Véronique et Lucie, qui ont dû faire avec.

Sommaire

Introduction	7
Les bibliothèques numériques.....	7
Numérisation, échantillonnage et moiré	12
Plan du mémoire et publications associées	15
1. Corpus textuels : constitution, diffusion et usages.....	18
Les bibliothèques numériques au CNAM.....	18
Une incursion en stylistique quantitative.....	27
Prolongements : les bibliothèques numériques et leur public.....	31
2. Interaction 3D pour les bibliothèques numériques.....	37
Le livre et les interfaces de lecture	38
La collection et les interfaces de recherche	44
Prolongements : interfaces pour le document et son contexte	50
3. Le livre comme volume.....	55
Numérisation 3D par photogrammétrie	55
Le livre objet : exemple des livres à systèmes	61
Prolongements	66
4. Eléments technologiques pour une machine à lire.....	70
Bureau 3D	70
Dispositif immersif de visualisation.....	74
Prolongements : dispositifs d'interaction pour le document 3D	79
Conclusion.....	88
Références bibliographiques	89
Résumé.....	100

Introduction

Les bibliothèques numériques

La recherche pour les bibliothèques numériques est actuellement en plein essor : une dizaine de revues et de conférences internationales spécialisées en témoignent*. Depuis la *Digital Library Initiative*** américaine [Fox et al., 1995], de nombreuses actions incitatives ont été mises en place par les organismes de financement de la recherche. Cette vitalité s'explique par la nature nécessairement pluridisciplinaire du thème et l'importance de ses enjeux sociaux. La diversité des propos a cependant conduit à rendre le terme « bibliothèque numérique » ambigu. On peut en effet considérer la bibliothèque numérique comme un lieu d'accueil (encore à inventer) pour les documents numériques que notre société produit maintenant en très grand nombre, ou bien, considérer la numérisation comme le prolongement naturel des missions de conservation et de diffusion communes à toutes les bibliothèques.

[Borgman, 1999] a très bien analysé les origines de cette dualité : « in general, researchers focus on digital libraries as content collected on behalf of user communities, while librarians focus on digital libraries as institutions or services (...) One reason for the confusion of terminology is that research and practice in digital libraries are being conducted concurrently at each stage of the continuum from basic research to implementation (...) Neither community is likely to surrender the term in favor of another » (p.229). Six ans plus tard, on retrouve encore un écho de cette difficulté dans la définition proposée dans l'appel à participation de la conférence JCDL'2005 : « Digital libraries are distinguished from information retrieval systems because they include more types of media, provide additional functionality and services, and include other stages of the information life cycle, from creation through use. Digital libraries also can be viewed as a new form of information institution or as an extension of the services libraries currently provide. ».

La figure 1 reproduit le résultat d'une requête sur le terme « bibliothèque numérique » effectuée sur le moteur Kartoo (www.kartoo.fr). La carte thématique

* Le thème a donné lieu à deux numéros spéciaux "fondateurs" de *Comm. ACM* (Avril 95 et Avril 98). Sont apparues depuis plusieurs revues à comité de sélection : *Intern. Journal on Digital Libraries* (IJDL, Springer), *D-Lib* (ARPA), *Journal of Digital Information* et en France *Document numérique* (Hermès). La conférence *ACM Digital Libraries* a fusionnée avec son homologue de l'IEEE-CS en 2001 (JCDL) La *European conf. on research and advanced technology for digital libraries* (ECDL) tiendra sa Xème édition en 2006. Une conférence similaire a lieu en Asie.

** Financée par la NSF, l'ARPA et la NASA auxquelles se sont joints la bibliothèque du congrès et d'autres grandes institutions fédérales. 24M\$ pour la phase 1994-98, programme reconduit en 1998-2003.

Les bibliothèques numériques ont aussi un passé, une histoire et un héritage (pour paraphraser [Buckland, 2002]), encore à retracer. La connaissance objective de cet héritage est pourtant fondamentale pour une compréhension des habitudes de travail, des motivations et des attentes des acteurs du domaine. Par exemple, l'attention un peu systématique portée actuellement en France à la numérisation des manuscrits enluminés* et aux Etats-Unis pour celle de la presse quotidienne locale du XIXème** a pour origine des campagnes de microfilmage effectuées dans ces deux pays dans l'entre-deux-guerres [Rouyer, 1999]. Les arguments avancés en faveur ou en défaveur du numérique par rapport au livre sont aussi souvent le décalque inconscient d'utopies (je pense ici au *Mondaneum* de Paul Otlet, au *Giant Brain* d'H.G. Wells, mais aussi à la pittoresque *Agonie du papier* d'A. Allais) ou de polémiques produites lors de la diffusion du microfilm il y a plus de 70 ans. Les échecs récents des livres électroniques, en tant qu'ordinateurs de poche spécialisés, ne sont pas non plus sans rappeler l'échec du télécopieur de journal à domicile de RCA, confronté dans les années 1940 à la télévision naissante.

On a bien sûr de nombreuses histoires de l'écrit, du livre, de la documentation, des bibliothèques (voir par ex. [Martin, 1990] pour une synthèse), ainsi que sur l'histoire des médias considérés dans leur ensemble [Barbier et Bertho Lavenir, 1996 ; Fayet-Scribe, 1997]. Le multimédia a aussi son anthologie, de « Wagner à la réalité virtuelle » [Packer et Jordan, 2001]. Il reste cependant à écrire l'histoire plus modeste des *machines à lire* (fig. 2), qu'il s'agisse des nombreux dispositifs « auxiliaires du livre » [Cim, 1900] : pupitres, lutrins, roues à livres, étagères, catalogues, fichiers, etc... ou des procédés substitutifs : reprographie, microfilm [Relier, 1966 ; Rouyer, 1999] et télécopie [Brethes, 1995]. Le MEMEX de V.Bush, inévitable icône du domaine, fait exception [Buckland, 1992]. La mécanisation des bibliothèques publiques (catalogues mécanographiés, puis mise en place des premiers systèmes en ligne) a récemment fait l'objet d'études historiques [Graham, Rayward et al. 2002]. Malgré cela, nous n'avons qu'une vue très parcellaire des premiers usages de l'ordinateur pour la manipulation de corpus textuels (travaux autour du *Trésor de la langue française*, par ex., dès 1957). Bien peu aussi de statistiques sur l'ampleur des campagnes de microfilmage, publiques ou privées, et surtout sur leurs usages ultérieurs. La chronologie des premières recherches véritablement informatiques sur le document numérique est décrite dans [André, 2003]. Il faudrait les mettre en regard avec celles relevant de l'*information retrieval* (travaux de G. Salton et le *Library of the future* de Licklider), durant la même période 1960-70. L'historique des développements

* Par ex. Base « Enluminures » du ministère de la Culture et travaux de la BM Lyon.

** Site « Memory of America » de la bibliothèque du Congrès

autour du concept d'hypertexte est bien résumé par [Nyce et Kahn, 1991 ; Laufer et Scavetta, 1995].

Le projet de numérisation de l'écrit le plus important de "l'ère pré-Web" est certainement celui initié lors de la conception de la nouvelle Bibliothèque de France fin des années 1980. Des moyens très importants étaient souhaités : 1 GF pour le numérique, 10 GF pour la bibliothèque elle-même*. Un plan de numérisation d'un million de titres était envisagé**, ainsi que le développement d'outils informatiques de consultation. On trouvera dans [Virbel, 1993] une description du *Poste de lecture associé par ordinateur* (PLAO) et dans [Maignien, 1998] un rappel des objectifs éditoriaux. [Pelfrene, 2002] donne un regard extérieur utile sur l'ensemble du projet.

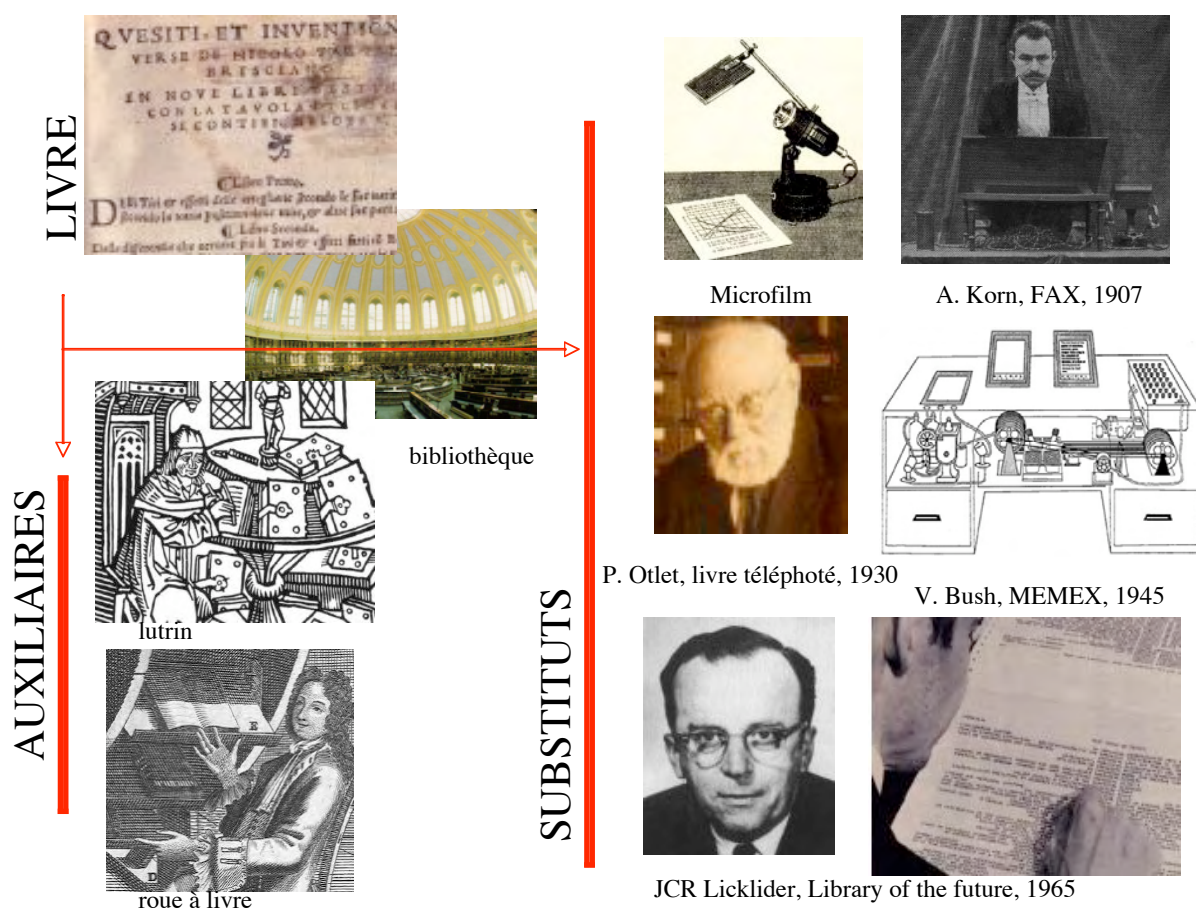


Figure 2. Machines à lire : le livre, ses auxiliaires, ses substituts

* En nov. 1997, J. Attali, conseiller du Président de la république et chargé de ce dossier à l'époque, déclarait regretter rétrospectivement ce cadrage : "on aurait dû faire l'inverse" (émission *Internet et compagnie*. Didier Sanz et Jérôme Paoli. Canal +, <http://didier.sanz.free.fr/I&Co/Web9711/I&Co1197.htm>)

** ramené à 300K titres en 1993 [Pelfrene, 2002]. Le seuil des 100K titres sera en fait atteint dans le courant de l'année 2006.

Passées progressivement du stade (totalement) expérimental au (relativement) opérationnel, les bibliothèques numériques ont depuis bénéficié du progrès constant des technologies de captation, de stockage et de transmission numérique ainsi que de la chute de leurs coûts [Lesk, 1997]. Le développement du World-Wide-Web [Gillies et Caillau, 2000] a permis d'atteindre un public international considérable par le biais d'une interface standard et commode. Les projets de numérisation se sont développés au fil des années : pour la France, le catalogue du ministère de la Culture* en recense mi-2006 près de 350 (mais il n'existe pas à ce jour de coordination de ces projets). Le nombre d'entreprises capables d'assumer les campagnes de numérisation a cru proportionnellement. Une dimension nouvelle vient d'être atteinte fin 2005 avec l'annonce de projets très ambitieux : l'accord de Google avec plusieurs très grandes bibliothèques américaines (Harvard, NYPL, Stanford, et surtout Université du Michigan), la mise en place d'une *Bibliothèque Numérique Européenne* (6 millions de titres d'ici 5 ans) et l'accord particulier de Microsoft avec la British Library (100K volumes en un an).

Ces annonces ne doivent pas faire oublier que plusieurs difficultés demeurent. Les plus médiatisées sont celles d'ordre juridique, en particulier la question du droit d'auteur. Du point de vue de la sélection des textes en vue d'une numérisation, on peut grouper la production imprimée en trois classes :

- les œuvres qui relèvent aujourd'hui du domaine public. En droit français, il s'agit de celles dont les auteurs sont décédés depuis 70 ans. En 2006, on peut grossièrement considérer que toute la production imprimée antérieure à 1850 relève de cette catégorie.

- les œuvres dont les ayants droit disposent d'une représentation numérique complète (sous la forme de fichiers de traitement de texte, SGML ou PostScript) permettant l'inclusion quasi automatique dans une bibliothèque numérique. Ce n'est en pratique le cas que depuis au mieux deux décennies, les grands éditeurs scientifiques étant précurseurs en la matière**.

- une classe intermédiaire, dont la période de production s'étale donc entre 1850 et 1990. Cette classe doit être subdivisée en deux groupes, selon que les ayants droit de ces œuvres sont identifiables, ou non : on parle alors d'œuvres « orphelines ».

Dans la pratique des projets de numérisation, ce dernier groupe revêt une importance considérable, car c'est lui qui regroupe la plupart des textes référencés dans la production intellectuelle d'aujourd'hui. On comprend donc que les projets

* <http://www.numerique.culture.fr/>

** En tant qu'expert auprès du Centre National du Livre depuis 2003, j'ai souvent été amené à constater le très faible niveau d'informatisation des éditeurs français.

de numérisation massive doivent passer par une négociation tout aussi massive des droits d'exploitation.

Mais il existe aussi des difficultés d'ordre plus technique, que nous allons examiner.

Numérisation, échantillonnage et moiré

Le point de départ de la construction d'une bibliothèque numérisée réside dans la numérisation du document sur support papier, pour en extraire, via un capteur, un nombre fini d'informations considérées comme significatives : ces informations (des nombres) sont ensuite susceptibles de traitements informatiques. La remédiation ainsi réalisée (du support papier au support mémoire magnétique, par ex.) ne peut pas se faire sans problème : les signaux acquis par le capteur coexistent avec du bruit, le choix des informations significatives lors de la numérisation doit souvent se faire *a posteriori*, mais, surtout, message et support entretiennent des relations trop complexes pour qu'on puisse les supposer toujours dissociables (fig. 3).

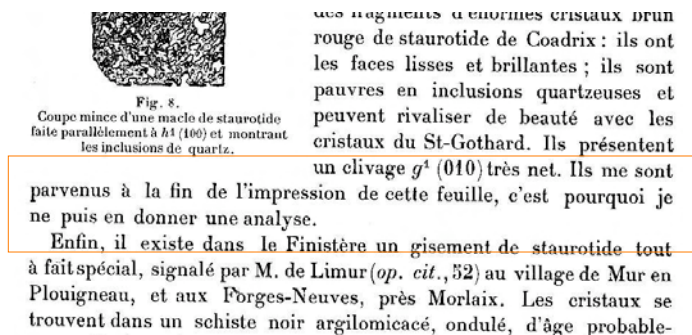


Figure 3. Séparation message/support difficile

(A. LACROIX. *Minéralogie de la France et de ses colonies*, T.1-1, p. 10, 1893, coll. part.)

L'expérience fondatrice en matière de numérisation de document est peut-être le photo-télégraphe d'Arthur Korn (1904). Dans cet appareil, une cellule au sélénium balaie une photographie en un nombre fini de lignes. Le signal électrique résultant est transmis sur une ligne téléphonique au récepteur, qui à son tour déplace une source lumineuse au-dessus d'un papier photosensible (fig. 4). Lorsqu'il juge de la qualité de reproduction de ses fac-similés, Korn montre bien les relations entre le nombre de lignes de balayage, le temps de transmission et les limitations de la technologie des lignes [Brethes, 1995]. Il a fallu attendre cependant une vingtaine d'années pour qu'Harry Nyquist propose une explication formelle du processus d'échantillonnage. Ce qu'on appelle maintenant le théorème de Nyquist donne une borne inférieure à la fréquence d'échantillonnage liée à la

bande passante du signal d'origine. Quand l'échantillonnage se fait à une fréquence inférieure à la limite de Nyquist, le signal résultant contient des formes absentes de l'original : du moiré.*



Figure 4. La téléphotographie à L'illustration. Conférence de M. le prof. Korn et expérience de transmission des photographies sur le circuit Paris-Lyon-Paris (supplément de l'illustration, 9 février 1907, coll. part.)

Les systèmes de numérisation actuels sont bien différents du photo-télégraphe de Korn, mais la nature même de la numérisation comme processus d'échantillonnage reste inchangée. Il est en pratique difficile de décider de la pertinence des signaux « émis » par un document et les exemples de sous-estimation sont fréquents, du fait des contraintes techniques ou économiques (sans oublier les approches naïves). Par exemple, la numérisation en noir et blanc est encore très courante dans les projets de numérisation du patrimoine imprimé, au lieu d'une captation couleur (fig. 5). Le traitement des planches hors-texte dépliantes est un exemple moins trivial. Cette technique de mise en page a été extensivement utilisée dans l'édition scientifique et technique jusqu'au milieu du XIXème. Elle permet en effet un commentaire (pour l'auteur) ou un examen (pour le lecteur) en parallèle, évitant les allers-retours entre page dus aux limites physiques imposées par la typographie. A ma connaissance, aucune interface de consultation de bibliothèque numérique actuellement en service ne transcrit ce type de fonctionnalité. Bien souvent, c'est à cause de la limitation des métadonnées associées au fac-similé. Dans notre projet de *Conservatoire Numérique* (chap. 1), nous avons bien distingué les hors-texte des autres fac-similés, mais pas le fait qu'une planche soit dépliant ou non. On doit dès lors considérer la version numérique du livre à planches comme un alias. D'une

* ou *alias*. On peut obtenir facilement des figures de moiré en superposant deux feuilles de celluloïd sur lesquelles on trace des réseaux de lignes de dimension et d'orientation légèrement différentes.

manière plus générale, un livre n'est pas seulement un ensemble de pages (planes), mais aussi un volume, qui se feuillette et peut être modélisé comme tel.

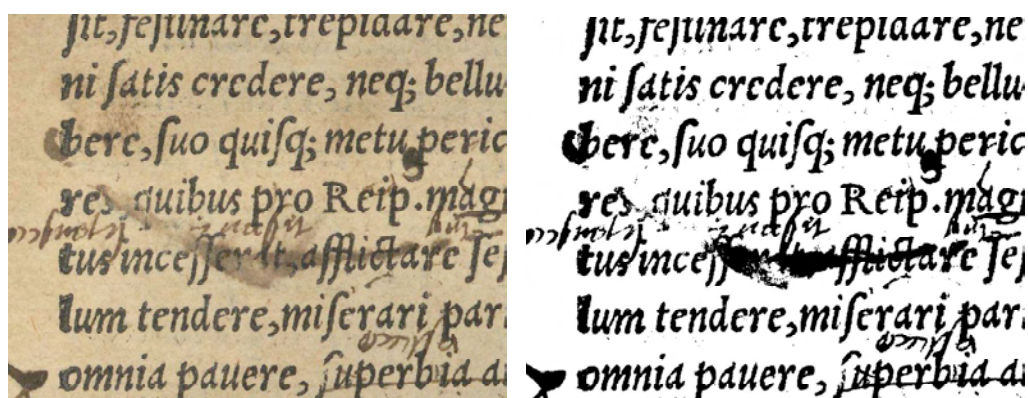


Figure 5. Gauche : numérisation couleur 600 DPI. Droite : image bitonale résultante, les tâches, les notes et l'encre se fondent en une forme unique (SALUSTE *De Catilinæ conjuratione* [...] trad. Melanchton, Cologne, 1539, p. 24. Coll. part.)

On retrouve cette problématique de l'échantillonnage et du moiré à d'autres niveaux : par exemple lorsqu'on réduit un périodique à la somme des articles qu'il contient, écartant au passage les couvertures, tables, errata (ex. de la bibliothèque numérique de l'ACM). Il me semble aussi que cette image de l'échantillonnage au niveau du document peut être généralisée au support et à l'environnement physique du texte (la bibliothèque), aux gestes et aux démarches des lecteurs et même dans la pratique de sélection de la « politique éditoriale » des bibliothèques numériques.

Les bibliothèques numériques souffrent par ailleurs des technologies qui sont à la base de leurs interfaces. Dans un chapitre intitulé « Les "oubliés" de la composition informatisée », [Laufer et Scavetta, 1992] faisaient déjà un constat similaire :

Le livre contemporain et ancien offre d'autres modèles qui n'ont pas tous été pris en compte. Songeons aux livres animés. Une tirette y fait sortir l'oiseau de la cage [...] Il devient possible aujourd'hui de simuler de tels livres. L'amélioration de la qualité des écrans était un préalable à toute lecture soutenue à l'écran. Le confort moyen demeure insuffisant. La petite dimension des écrans grand public freine l'hypertexte et le dévoie passagèrement. La question des formats présente pourtant une analogie évidente avec l'imprimé. Dans les rayons d'une bibliothèque, on ne place pas côte à côte les folios éléphants et les octavos nains. On n'y verse pas non plus les mêmes matières. [...] La structure physique de chaque page définit un ordre strictement linéaire. Mais la séquence des surfaces encrées sur les pages brise la ligne du discours en arborescences diverses, par le jeu des divisions et subdivisions. [...] Ces relations ne sont pas analysées mais perçues au feuilletage, autrement dit à une lecture graphique de survol toujours associée à la lecture verbale. [...] La physique du livre enfin ne se trouve pas seulement dans la succession de surfaces planes mais dans son assemblage tridimensionnel, son *volume*. C'est ce qui autorise les lectures transversales à l'aide de tables. [...] les mots d'onglets et d'index disent bien comment le livre se tient et se consulte. Le poids et

l'encombrement rendent mal commodes les plus beaux ouvrages de très grand format. Nous feuilletons avec le pouce et l'index et souvent de l'une ou l'autre main. Nous glissons même les autres doigts entre les pages en consultant plusieurs références. Toute simulation doit tenir compte du corps et de l'espace. [...] Pas de simulation interactive sans stimulation sensorielle. L'hypertexte a dû découvrir le rôle de la main dans la lecture.» (pp. 17 - 20)

La consultation de plusieurs documents en parallèle se traduit par des superpositions de fenêtres et des actions très nombreuses de redimensionnement, rendues d'autant plus nécessaires que les fac-similés imposent leurs dimensions au dispositif de visualisation, et non l'inverse. De fait, la plupart des utilisateurs préfèrent télécharger et imprimer leurs documents de travail. Avec la multiplication des bibliothèques en ligne d'accès international, les utilisateurs vont de plus être confrontés à un nouveau problème : évaluer l'adéquation de ces collections avec leur besoin de documentation. Pour estimer la valeur d'un corpus, il faut avoir recours à un moteur de recherche, ce qui suppose la capacité de l'utilisateur d'énoncer une telle requête (fig. 6). Comment dès lors déambuler dans une bibliothèque numérique ? Comment encourager la découverte accidentelle ? Quelles techniques de visualisation utiliser quand la taille de corpus peut atteindre le million d'items ?



Figure 6. Une bibliothèque difficile d'accès (<http://books.google.com>)

Il paraît donc nécessaire d'augmenter le débit d'informations entre les bibliothèques numériques et leurs usagers. Ceci implique l'obtention de données plus riches et la mise en œuvre d'une interaction plus fluide pour l'utilisateur. Alors, les bibliothèques numériques cesseront d'être des alias.

Plan du mémoire et publication associées

Ce mémoire résume mes contributions à ces questionnements sur les bibliothèques numériques et l'interaction humain-machine.* A quelques petits

* Pour des raisons de cohérence, je n'ai pas évoqué d'autres contributions consacrées à l'élaboration de modèles et d'outils pour l'évaluation de performance des systèmes parallèles, effectuées dans le prolongement direct de ma thèse de doctorat (soutenue en 1992). Je joins à ce mémoire un curriculum vitæ qui décrit ces travaux.

arrangements près, j'ai tenté de respecter au mieux la chronologie de mon parcours intellectuel.

Le chapitre 1. *Corpus textuels : constitution, diffusion et usages* décrit les projets de bibliothèques numériques dans lesquels j'ai été impliqué, et des exemples avancés d'utilisations de ces bibliothèques.

CUBAUD P. ABU : Un serveur de texte intégral sur Internet. Bulletin bibliographique de l'Institut nat. des Techniques Documentaires, Vol. 19, n°5, pp. 3-6. 1995.

CUBAUD P., GIRARD D. ABU : une bibliothèque numérique et son public. Document Numérique, vol. 2(3-4), 1998, pp. 13-30.

CUBAUD P., TOPOL A. A WWW-based digital library for antiquarian collection. Rapport de recherche du CEDRIC RR99-09. Mars 1999.

DEBLOCK G., ROZET B., CUBAUD P. Le Conservatoire numérique des arts et métiers : une création partenariale. Bulletin des bibliothèques de France, t. 46, n°4, 2001, pp. 43-49.

CUBAUD P., DEBLOCK G. Le conservatoire numérique des arts et métiers : historique du projet et organisation du site. Actes du colloque La numérisation des textes et des images : techniques et réalisations. Editions de l'Univ. Lille 3, pp. 13-22. Janvier 2003.

LECLERC Y., CUBAUD P. Bouvard : 618, Pécuchet : 598. Etude de statistique textuelle. Revue Flaubert n°3, 2003. www.univ-rouen.fr/flaubert

Je décris ensuite dans le chapitre 2. *Interaction 3D pour les bibliothèques numériques* un prototype d'atelier de lecture dont les briques de base sont un outil de parcours dans les collections par visualisation exhaustive et un outil de lecture de fac-similés.

CUBAUD P., THIRIA C., TOPOL A. Experimenting a 3D Interface for the access to a Digital Library. Proc. of ACM DL'98. Pittsburg, USA, June 1998.

CUBAUD P., TOPOL A., VODISLAV D. Les limites de VRML pour les comportements interactifs : étude de cas. Actes de la conférence ERGO-IHM'2000, Biarritz. Oct. 2000.

CUBAUD P., TOPOL, A. A VRML-based user interface for an online digitalized antiquarian collection. Proc. ACM SIGGRAPH Web3D Symp., Padderborn, April 2001, pp. 51-59.

CUBAUD P., STOKOWSKI P., TOPOL A. Mixing Browsing and Reading Activities in a 3D Digitalized Library. Proc. of ACM-IEEE JCDL'02, Portland, USA, June 2002.

CUBAUD P., DUPIRE J., TOPOL A. Textes, images, volumes : les bibliothèques numériques au CNAM. Actes Premières journées francophones de la toile (JFT). Tours, France, July 2003.

Le chapitre 3. *Le livre comme volume* présente une technique nouvelle de captation volumique par photogrammétrie, puis une modélisation de l'interaction pour les livres à systèmes.

CUBAUD P., HAAS J.F., TOPOL A. Numérisation 3D de documents par photogrammétrie. Proc. 8ème conf. francophone sur l'écrit et le document (CIFED'04). La Rochelle, France, June 2004.

CUBAUD P., DUPIRE J., TOPOL A. Digitization and 3D modeling of movable books. Proc. of ACM-IEEE JCDL'05, Denver, USA, pp. 244-5, June 2005.

Le chapitre 4. *Eléments technologiques pour une machine à lire* est plus particulièrement consacré à des expérimentations sur les dispositifs matériels et les services logiciels à offrir pour faciliter l'interaction avec les documents 3D.

TOPOL A. Immersion of XWindow applications into a 3D workbench. Proc. ACM CHI'2000, The Hague, Netherlands, 2000.

TOPOL, A. Interaction 3D pour les paysages informationnels. Thèse de doctorat en informatique. Conservatoire national des arts et métiers. Décembre 2002.

CUBAUD P., DUPIRE J., TOPOL A. Using game engines for non 3D gaming applications. Proc. of CGAMES'05. Angoulême, France, nov. 2005.

ALMEIDA R., CUBAUD P., DUPIRE J., TOPOL A. En hommage à Ramelli : un dispositif immersif de consultation de bibliothèques numériques. Actes de la conférence H2PTM'05. Université Paris VIII, Saint-Denis, Déc. 2005.

Chacune de ces parties se termine par l'évocation de thèmes de recherche qui vont ou qui pourraient prolonger les travaux effectués. Il s'agit dans le premier cas de thèses en cours dont je présente l'état, sans publication validée. Pour les pistes plus hypothétiques, j'ai renoncé à proposer un calendrier ou une évaluation des ressources nécessaires, mais je pense que ces travaux pourraient s'effectuer dans le cadre de collaborations entre laboratoires, par exemple en tant que réponses à des appels à projets nationaux ou européens.

Un dernier point : les publications que je viens de mentionner sont collectives. J'espère que les "je" et "nous" employés dans la suite de ce mémoire permettent de distinguer mon apport personnel à ces travaux.

1. Corpus textuels : constitution, diffusion et usages

Les bibliothèques numériques au CNAM

[Cubaud, 1995 ; Cubaud et Girard, 1998 ; Deblock, Rozet et Cubaud, 2001 ; Cubaud et Deblock, 2003]

L'ASSOCIATION DES BIBLIOPHILES UNIVERSELS (ABU)

L'ABU est un groupement informel de professionnels et de particuliers assez typique des actions coopératives d'internautes, dans le cadre par exemple du développement de logiciels libres ou des forums de discussion. La motivation de ses membres est double : accroître le corpus de textes diffusés librement sur Internet et étudier les modalités techniques de sa consultation via l'Internet. J'ai créé le service Web en Octobre 1993, quelques mois après la diffusion par le NCSA du logiciel client Mosaic (fig. 1.1 gauche). Le service a fait l'objet de trois versions successives, rendues nécessaires par l'accroissement du corpus et des possibilités offertes par le Web [Cubaud et Girard, 1998]. La version actuelle (<http://abu.cnam.fr>) a été mise en place en octobre 1996 : il s'agit donc d'un fossile vivant du Web... Je la présente brièvement ici, essentiellement du point de vue de son utilisation. Les aspects informatiques de son organisation interne n'ont plus - dix ans après - que peu d'intérêt : s'ils étaient innovants à l'époque (le serveur Web du CNAM, développé au centre de ressources informatiques de l'établissement, fut le premier à disposer des fonctions de HTML dynamique), ils sont devenus banals aujourd'hui. Les derniers textes ont été ajoutés au site en janvier 2002.

Interface utilisateur

La fig. 1.1 (droite) reproduit un exemple d'utilisation du service dans un contexte de lecture savante : l'utilisateur s'intéresse ici au « commodisme » d'Henri Poincaré. Il a sélectionné dans le catalogue (A) les textes qui lui paraissent adaptés. Une recherche d'occurrence dans ces textes (B) l'a guidé sur quelques passages qu'il a ensuite sélectionnés (C et D). Il prend un paragraphe important dans (C) pour l'incorporer dans son propre texte (E). Pour des raisons de confort visuel, l'utilisateur a modifié le corps des polices de base du client Web en 18 points. Pour ouvrir toutes les fenêtres, il a systématiquement utilisé un raccourci souris (« open link in new window » dans cet exemple réalisé en 1998 avec Netscape Macintosh).

littéraires de l'ABU sont d'inspiration philosophique ou en tout cas polémique. On peut considérer également que l'ensemble du corpus est par nature susceptible d'intérêt historique. La pauvreté du fonds en matière scientifique est en revanche regrettable. Elle s'explique en partie par la complexité structurelle de ce type de textes, qui nécessite des techniques de balisage sophistiquées (tableaux, formules, figures).

Le corpus diffusé à ce jour est composé de 288 titres, représentant plus de 11 millions de mots-occurrences et 70 millions de caractères. Il s'agit d'un corpus relativement modeste, si on le compare par exemple à la base FRANTEXT de l'INaLF (180 millions de mots-occurrences, 3500 « unités textuelles » d'après [INA]). Notons que ce dernier ayant été conçu à des fins d'études linguistiques, il n'intègre pas de description structurelle des textes et ne les transcrit pas toujours dans leur intégralité. Les conventions de balisage du corpus sont minimales. La description structurelle d'un texte peut être résumée par la grammaire formelle suivante (en notation EBNF) :

```
<texteABU> ::= <cartouche> { <partie> }
<partie > ::= [<titre_partie>] { <chapitre> }
<titre_partie> ::= <phrase> === {=} saut_ligne <phrase>
<chapitre> ::= [<titre_chapitre>] { <section> }
<titre_chapitre> ::= <phrase> --- {-} saut_ligne <phrase>
<section> ::= [<titre_section>] { <paragraphe> }
<titre_section> ::= --- {-} <phrase> --- {-}
<paragraphe> ::= saut_ligne { <phrase> } saut_ligne
```

Le rôle du cartouche est de contenir les métadonnées associées au texte diffusé. Mise en place dès 1995, la technique de description préfigure XML :

```
<IDENT methode>
<IDENT_AUTEURS descartesr>
<IDENT_COPISTES cubaudp>
<ARCHIVE http://www.abu.org/>
<VERSION 3>
<DROITS 0>
<TITRE Discours de la méthode>
<GENRE prose>
<AUTEUR Descartes>
<COPISTE Pierre Cubaud (cubaud@cnam.fr)>
<NOTESPROD>
  OEUVRES DE DESCARTES,
  PUBLIEES PAR VICTOR COUSIN,
  Tome Premier,
  Paris, 1824
  La pagination originale est indiquée entre crochets []
</NOTESPROD>
```

Les balises de mise en valeur du texte permettent la prise en compte :

- Des italiques, sans distinction de l'intention de l'auteur :
`_ Lacrymabiliter ! _`
- Des expressions mathématiques, avec la grammaire de TeX :
L'égalité $2 + 2 = 4$ n'a été ainsi susceptible d'une vérification que parce qu'elle est particulière.
- De la pagination de l'ouvrage original :
car pour la raison, ou le sens, d'autant qu'elle est [123] la seule chose qui nous rend hommes et nous distingue des bêtes
- Des notes de bas de page :
Un espion léger, infidèle et moqueur (* Causeries du Lundi, IV, 354.*), mais un observateur et un peintre
- Des ruptures forcées, pour des exergues par exemple :
Pè far la to vendetta \\
Sta sigur', vasta anche ella. \\
-- Vocero du Niolo.

Les copistes ont parfois eu recours à d'autres balises pour faire face à des cas particuliers (transcriptions de manuscrits, citations de langues non latines par exemple). Dans ce cas, la règle employée est explicitée dans le cartouche du fichier. Afin de simplifier le fonctionnement du service Web, nous avons utilisé le jeu de caractères ISO 8859-1 (ISO-LATIN-1) pour l'encodage des textes. La ligature Œ n'est donc pas prise en compte. Le tiret long est transcrit par deux tirets simples consécutifs.

Une faiblesse majeure du corpus d'ABU est de ne pas utiliser les acquis plus récents en matière de balisage structurel que constituent la *Text Encoding Initiative* (TEI) et le *Corpus Encoding Standard* (CES) [Ide et Veronis, 1995 ; Role, 2000]. Une conversion du format ABU en TEI-Lite est certainement automatisable. Le problème réside plus dans une rétro-conversion complète du fond pour tirer partie au maximum de la TEI. Les copistes de l'ABU ne pouvaient l'entreprendre sans un minimum de connaissances informatiques, aussi des ouvrages introductifs comme [Habert et al., 1998] ont-ils été particulièrement bienvenus. La rétro-conversion du fond ABU à la norme TEI a par ailleurs été partiellement effectuée par le *Electronic Text Center* de l'université de Virginie, mais les textes étaient rediffusés en HTML, ce qui nuisait à l'évaluation du procédé.

Le corpus ABU est totalement libre de copie pour un usage personnel, à des fins d'illustration pour l'enseignement, ou de recherche scientifique. Sa rediffusion ou son inclusion dans une autre œuvre est autorisée à la condition que le rediffuseur s'engage à signaler clairement que le texte provient de l'ABU qui l'a diffusé bénévolement et à indiquer les coordonnées du copiste. Nous retrouvons ici les conditions usuelles de diffusion des logiciels libres, transposées au texte

électronique. On peut d'ailleurs trouver les textes issus d'ABU dans de nombreux cédéroms de compilation de logiciels libres. La licence d'utilisation du corpus est accessible sur le service Web. Elle est automatiquement ajoutée en en-tête des fichiers à chaque téléchargement.

Rétrospectivement, il me semble que l'expérience ABU a été positive sur plusieurs points :

- Elle a contribué à démontrer, dès 1993-4, qu'au prix de quelques extensions, le Web pouvait être le vecteur idéal pour l'accès à des fonds d'ouvrages numérisés.
- L'appel à participation a permis de constituer en un temps assez bref un corpus significatif et de mettre en place une coopération durable entre des copistes motivés. La diversité des motivations des copistes a eu pour conséquence positive une grande variété dans le choix des textes, qui s'est traduit par une audience potentielle la plus vaste possible. L'appel au bénévolat nous a permis également de nous affranchir de toute contrainte commerciale ou légale pour la diffusion du corpus.
- La politique de « petits pas » suivie pour le développement du service, accompagnée par l'observation constante de l'usage qui en est fait, a été une bonne voie pour sortir les technologies associées au texte électronique du domaine des seuls spécialistes.
- L'utilisation intensive des outils Internet (messagerie, listes de diffusion, Web) a permis à l'association de fonctionner parfaitement sans que ses membres ne se soient jamais rencontrés (à quelques exceptions près). Les logiciels libres employés pour le développement ont permis de réduire les coûts à leur plus simple expression.

LE CONSERVATOIRE NUMERIQUE DES ARTS ET METIERS (CNUM)

En France, le projet de numérisation de la bibliothèque nationale (Gallica) a démontré vers 1996-7 que la diffusion sur Internet d'un vaste corpus d'ouvrages de caractère patrimonial numérisés en mode image pouvait rencontrer un lectorat très important. Le temps était donc venu pour les institutions publiques plus modestes d'entreprendre à leur tour un programme de numérisation et de diffusion des fonds dont elles étaient dépositaires. C'est le cas du CNAM qui dispose d'un fonds patrimonial de 60000 ouvrages scientifiques et techniques.

L'idée première de la Bibliothèque du CNAM était de préserver et valoriser un fonds patrimonial très consulté, majoritairement composé d'ouvrages et de périodiques du XIX^{ème} siècle et menacé par l'acidité des papiers qui le

composent. Le microfilmage pratiqué jusqu'alors pouvait être avantageusement complété, puis remplacé par la numérisation. Le *Conservatoire numérique des arts et métiers* (CNUM, <http://cnum.cnam.fr>) a démarré en janvier 1998 comme projet interne et autofinancé de trois services du CNAM : la bibliothèque centrale (qui a la responsabilité d'animation du projet), le Centre D'Histoire des Techniques (CDHT) et notre groupe au sein de l'équipe *Multimédia et interaction* du CEDRIC. Des experts d'autres institutions européennes y ont collaboré ponctuellement, essentiellement dans la phase de démarrage du projet [Blondel, 2000].

Constitution du fonds numérisé

Du point de vue éditorial, l'intention du projet est d'extraire du fonds patrimonial une collection francophone représentative permettant d'aborder l'histoire des techniques dans toutes ses dimensions [Deblock et al., 2001] :

- la constitution des savoirs, par le biais de monographies et d'articles de spécialistes en direction de leurs pairs,
- l'histoire des institutions : rapports d'activités, de commissions, de jurys, travaux des enseignants du CNAM,
- la médiation vers le public : ouvrages et revues d'enseignement populaire, de vulgarisation, récréations et romans scientifiques.

La diversité des titres retenus nous a contraints d'aborder de front toute la complexité de l'édition scientifique ancienne : lourd appareil éditorial, nombreuses illustrations, souvent en planches hors texte, paginations multiples et parfois défectueuses, ouvrages en plusieurs volumes, recueils factices, etc. Il a donc fallu élaborer un modèle descriptif valide pour tous les cas de figure rencontrés, qu'il s'agisse de monographies ou de périodiques [Cubaud et Deblock, 2003]. Ce modèle a ensuite permis de définir la structure des bases de données utilisées pour la production des fac-similés et leur mise en ligne (ces bases de données contiennent les fac-similés, les notices bibliographiques, les tables des matières et les tables des illustrations). On se sera peut-être étonné d'apprendre qu'il n'existe pas à ce jour d'ontologie apte à décrire l'organisation matérielle d'un livre ancien. Le modèle de la TEI est, par exemple, insuffisant pour décrire de manière standard la numérotation des pages correspondant aux fac-similés, ce qui rend impossible la possibilité de décrire ensuite des tables. Plus récent, le modèle *Metadata Encoding and Transmission Standard* (METS), développé par la bibliothèque du congrès américain,* est certainement utilisable dans ce contexte,

* <http://www.loc.gov/standards/mets/>

en s'appuyant également sur les résultats du projet européen METAe** (moteur docWORKS/METAe).

La volumétrie actuelle (février 2006) du site est la suivante : 426 volumes numérisés, correspondant à 178778 images (soit 420 pages/vol., ce qui est très au-dessus des estimations usuelles de la bibliothéconomie). Les tables saisies représentent 76589 entrées de tables de matière et 31976 illustrations. Au total, c'est donc 60% du corpus d'images qui est décrit sous la forme de métadonnées textuelles. Cette volumétrie devra doubler courant 2006, avec le traitement de deux importants périodiques français du XIXème (*La lumière électrique* et le *Bulletin de la Société d'Encouragement à l'Industrie Nationale*).

Diffusion des fac-similés

Un site Web de bibliothèque numérisée a une structure assez spécifique, avec peu de « profondeur » et une hyper-textualisation relativement faible. L'immense majorité de ses pages HTML sont des fac-similés de pages d'ouvrages papier et, du point de vue ergonomique, la navigation dans ces dernières doit être étudiée très finement. Contrairement aux procédures employées à la Bibliothèque Nationale, nous avons pris le parti de dissocier totalement le format des fac-similés diffusés de celui des numérisations. On sait en effet qu'aucun standard actuel de format de fichier image n'est totalement satisfaisant pour la diffusion de fac-similés d'ouvrages [Trin et LeBourgeois, 2001 ; Salomon, 1997]. Si les algorithmes de compression type Huffman (CCITT Groupe 3 et 4) ont été conçus pour les fac-similés de pages de texte, ils sont totalement inadaptés pour les gravures (fig. 1.3 gauche). La limitation et le caractère aléatoire du débit de l'internet « au large » imposent pourtant des contraintes assez strictes sur la taille des fichiers diffusés. Nous avons donc dégradé les images diffusées de 100 à 133 DPI (à peu près la résolution d'un écran) et 8 niveaux de gris. Les navigateurs Web les plus répandus du marché ne gèrent par défaut que trois formats de fichier : GIF, PNG et JFIF (JPEG). L'adoption de formats plus adaptés (TIFF-FX ou DjVU) au sein de ces navigateurs reste à ce jour problématique. La compression JPEG, conçue pour les tons continus et la couleur, n'a pas été retenue (voir [Mintzer et al., 1996] pour son utilisation dans le cas de documents couleurs). Le format GIF a finalement été préféré au PNG car il est le seul à permettre l'encodage des 8 niveaux de gris sur 3 bits. Les images résultantes ont ainsi une taille moyenne de 100 Ko, soit 10 à 20 fois moins que les originaux (400 DPI, 256 niveaux de gris), mais sont parfois aux limites de l'acceptable en terme de lisibilité [Cubaud et Topol, 1999]. La présence combinée de texte et

** <http://meta-e.aib.uni-linz.ac.at/>

d'illustrations au sein d'un même fac-similé pose problème dans la phase de sous-échantillonnage pour éviter qu'apparaissent des figures de moiré. Nous avons recours à une image intermédiaire bitonale où les illustrations sont tramées et c'est cette image qui est ensuite dégradée (fig. 1.2). Ce traitement améliore considérablement la lisibilité des illustrations, par comparaison à une image bitonale diffusée telle que, sans traitement au niveau du poste client (procédure suivie par Gallica, par ex.)

Il est évident que ces choix techniques devront être remis en cause dans le futur, selon l'ampleur des progrès en matière de débit de transmission chez les utilisateurs finaux. Au moment de la mise en place du site, ce point était considéré comme crucial. La figure 1.3 (droite) récapitule les résultats d'un test effectué en 1999 sur un premier lot de 11087 images de doubles pages. On y compare les tailles des fichiers d'origine (TIFF groupe 4 en 400 DPI) avec celles de fichiers dégradés en 100 DPI aux formats GIF et PNG en 16 niveaux de gris. On remarquera la forte variabilité des tailles de fichiers TIFF-g4, qui traduit l'inadéquation de ce type de compression pour les fac-similés de gravures. Les deux pics d'effectifs pour les formats GIF et PNG correspondent en revanche clairement aux deux types d'images traitées : fac-similés de texte et planches gravées.

L'utilisateur accède aux ouvrages par une page descriptive, incluant la notice et une table abrégée. Cette page est construite une fois pour toutes lors de la mise à jour du serveur et peut être référencée par des sites externes (moteurs de recherche, en particulier). Il est également possible de localiser un ouvrage par interrogation du contenu de sa notice et/ou de ses tables. Un logiciel d'interrogation élémentaire a été développé à cette fin. La navigation entre les pages elles-mêmes se fait par un « feuilletage » séquentiel. L'utilisateur peut aussi accéder au numéro de page de son choix. Il peut également naviguer exclusivement dans les planches hors-texte de l'ouvrage. Ce « typage » de la navigation nous a paru utile du fait de la spécificité des documents diffusés. Il pourrait être généralisé. La mise en regard de plusieurs pages est une fonction standard des navigateurs Web : elle trouve son plein intérêt ici (fig. 1.4). Le téléchargement intégral d'un volume (au format PDF) est actuellement en cours d'étude.

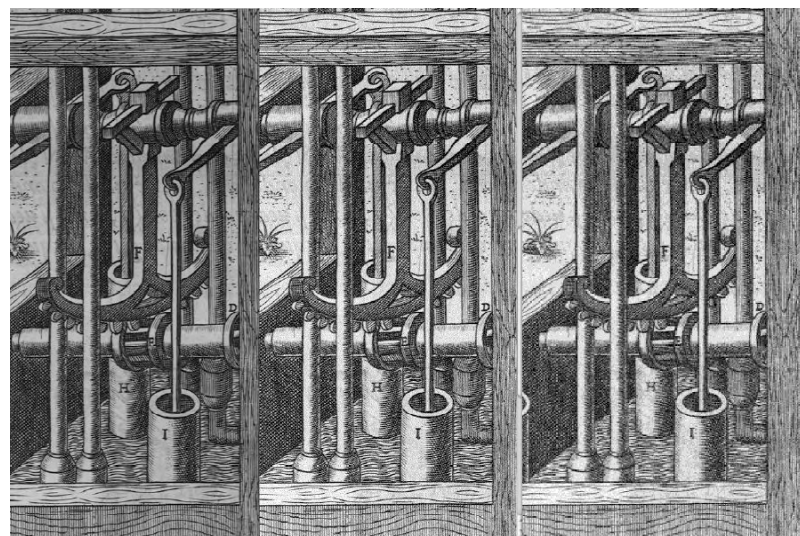
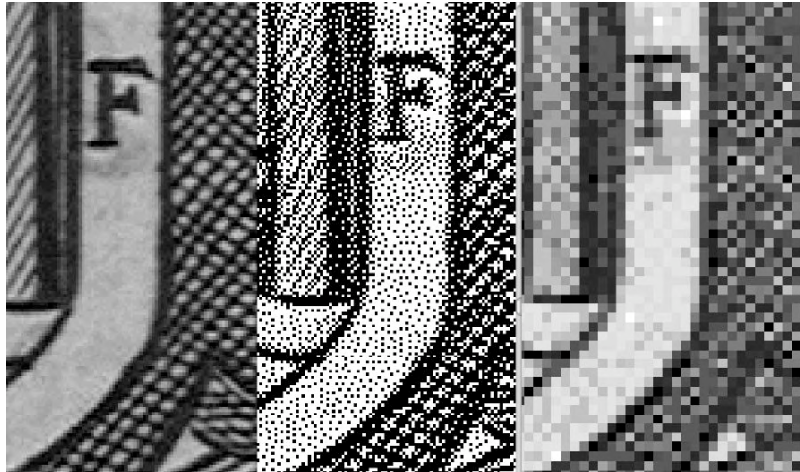


Figure 1.2. En haut : impact de la dégradation au niveau pixel. Gauche : image 400 DPI 256 niveaux de gris (png, 5.5 Mo). Milieu : idem bitonal (tiff, 1.1 Mo). Droite : image 133 DPI, 8 niveaux de gris (gif, 0.2Mo). En bas : résultat sur l'ensemble de l'image.

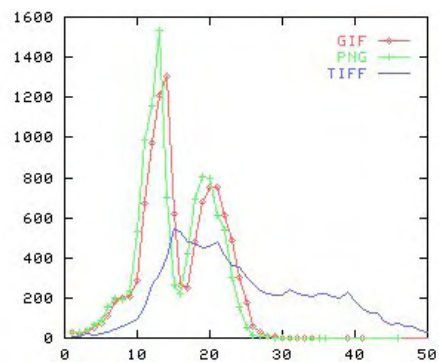
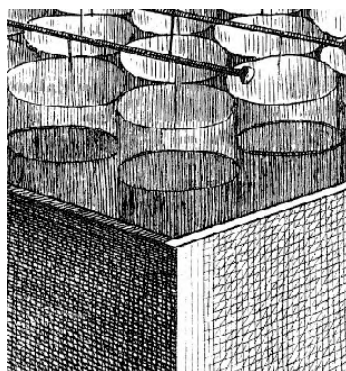


Figure 1.3. Gauche : une image difficilement compressible (Priestley, *Histoire de l'électricité*, 1771, CNAM 12°SAR6.3). Droite : histogramme des tailles de fichiers du lot d'images tests (unité = 10 Ko).

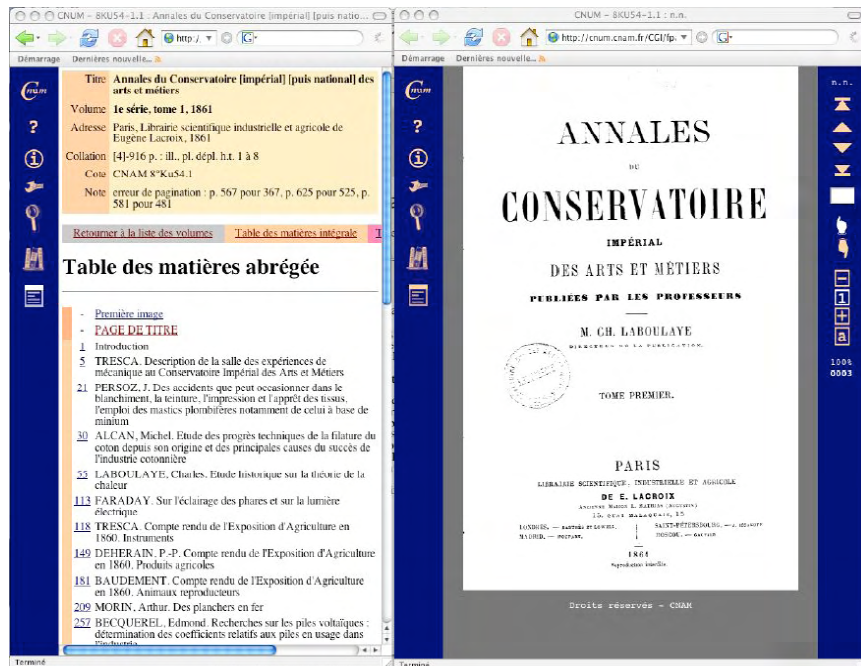


Figure 1.4. Session de travail sur le site CNUM.

Une incursion en stylistique quantitative

[Leclerc et Cubaud, 2003]

La mise à disposition d'un corpus numérisé trouve son plein intérêt en statistique textuelle [Lebart et Salem, 1994]. C'est d'ailleurs dans ce champ disciplinaire qu'ont été effectués les travaux pionniers en matière de numérisation (*Trésor de la langue française*, par ex.). L'application de ces techniques statistiques à des fins d'études littéraires, comme la stylistique ou la recherche d'attribution est elle aussi ancienne : on peut par exemple citer les travaux de Ch. Muller sur Corneille [Muller, 1967]. Pour A. Vuillemin, « il n'est guère de grands auteurs de W. Shakespeare à V. Hugo, de W. Goethe à P. Calderon, en passant par Homère, Virgile, Dante, Rabelais et bien d'autres, qui n'aient été ainsi étudiés » [Vuillemin, 2001]. Cependant, à l'exception de [Lebrave et al., 1991 ; Lecolinet et al., 1998], les travaux de stylométrie concernant Flaubert sont à ce jour rares, alors que les études sont facilitées par l'existence d'un riche corpus numérisé.* Il était donc utile que la *Revue Flaubert* consacre un numéro aux « apports de l'informatique et de l'Internet aux nouvelles formes d'édition, de lecture et d'interprétation des textes flaubertiens ». Y. Leclerc (dir. du Centre Flaubert à l'université de Rouen) et moi-même avons étudié la succession des occurrences

* Grâce aux efforts d'Eric Dubreucq, Vincent Maret et Joël Surcouf, l'ABU dispose de la presque totalité de l'œuvre de Flaubert. Depuis, se prépare à la bibliothèque municipale de Rouen la numérisation des manuscrits de ces œuvres (<http://www.univ-rouen.fr/psi/BOVARY/>).

Peut-on alors considérer la séquence {BP} comme le résultat d'une répétition de tirs à pile ou face équilibrés ? Sans doute non, du fait de l'auto-corrélation apparente de la séquence, qui interdit d'avoir recours à l'estimateur précédent de p. Il existe de nombreux tests de l'indépendance des termes d'une série chronologique (voir par ex. [Vaté, 1993], chapitres 1 et 3). Nous pouvons par exemple utiliser le rapport η de von Neuman :

$$\eta = \frac{d^2}{s^2} = \frac{1}{N-1} \sum_{i=1}^{N-1} (x_{i+1} - x_i)^2 \bigg/ \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \text{ avec } E(\eta) = \frac{2N}{N-1} \text{ et } V(\eta) = \frac{4(N-2)}{(N-1)^2} \approx \frac{4}{N}$$

On constate ici que $\eta = 0,6782 / 0,2499$ et $|\eta - E(\eta)| / \sqrt{V(\eta)} = 12,41 > 1,96$. On peut donc rejeter l'hypothèse d'indépendance des termes de la séquence (avec un risque d'erreur de 5%).

Un autre argument vient renforcer cette réserve : le nombre de répétitions **B**, **P**, **BB** et **PP** dans la séquence semble dépasser ce qu'on devrait obtenir dans le cas de tirs à pile ou face. On sait que dans ce cas, les séquences de $k = 1, 2, \dots$ symboles identiques apparaissent avec une probabilité $P[X = k] = p^k (1 - p)$. Nous indiquons ci-dessous les effectifs observés, avec en regard l'effectif théorique (avec $p = 0,5082$). Un test de chi-2 montre que les répartitions ne coïncident pas (pour les séquences B : avec 5 cellules, $D = 147,2 \gg \chi^2_{[95\%,3]} = 7,8$).

Séquence	obs.	théo.
B	275	152
BB	99	51
BBB	21	20
B*4	9	8
B*5	4	3
B*6	3	1
B*7	1	0.7

Séquence	obs.	théo.
P	287	152
PP	95	50
PPP	18	18
P*4	7	7
P*5	2	3
P*6	0	?
P*7	2	0.5
P*15	1	0.001

Après analyse de l'auto-corrélation de la séquence, nous avons pu modéliser la séquence binaire {BP} par une chaîne de Markov d'ordre 2, un peu plus complexe que la forme adoptée par Markov lui-même dans son analyse des successions voyelles-consonnes de textes russes [Petruszewycz, 1981]. Par ailleurs, on peut modéliser les alternances « **B** et **P** ensemble » / « **B** ou **P** seul » par un tirage au sort – et cela permet d'en déduire correctement le nombre de répétitions. En fait, ce résultat se retrouve dans tout phénomène « sans mémoire » : c'est le cas par exemple de la durée des conversations téléphoniques. Ici, l'auteur déterminerait son choix de poursuivre (par ex. à traiter « **B** et **P** ensemble ») à chaque étape, indépendamment des choix pris dans le passé. Il n'est bien sûr pas dans notre intention de conclure formellement que Flaubert procède de cette manière,

puisque des processus plus complexes pourraient conduire aux mêmes résultats. Le principe d'économie – fondement de la modélisation de systèmes physiques – ne s'utilise pas sans danger en matière de stylométrie [Bernard, 1999].

Les traitements statistiques et leur interprétation ont posé par ailleurs des questions de visualisation d'information et d'interaction avec le texte d'étude qu'il faudrait approfondir. La courbe fig. 1.5 représente l'évolution chronologique de la séquence {BP}. On indique en abscisse le numéro d'ordre de l'occurrence B ou P (en bas) et les divisions de chapitres (en haut). En ordonnée, on trouve la différence $nB - nP$ – très brièvement négative au tout début du roman. On peut être frappé par les brusques descentes (longues séquences **PPP...**) vers les occurrences 300, 1000 et 1100. Ces trois ruptures correspondent bien sûr aux trois séquences où Pécuchet se retrouve seul : au chapitre 3 consacré aux sciences, au chapitre 9 (religion) et, enfin, au dernier chapitre (pédagogie). La courbe suggère également que Flaubert rééquilibre son texte en **P** de manière massive alors qu'il introduit **B** plus « naturellement ».

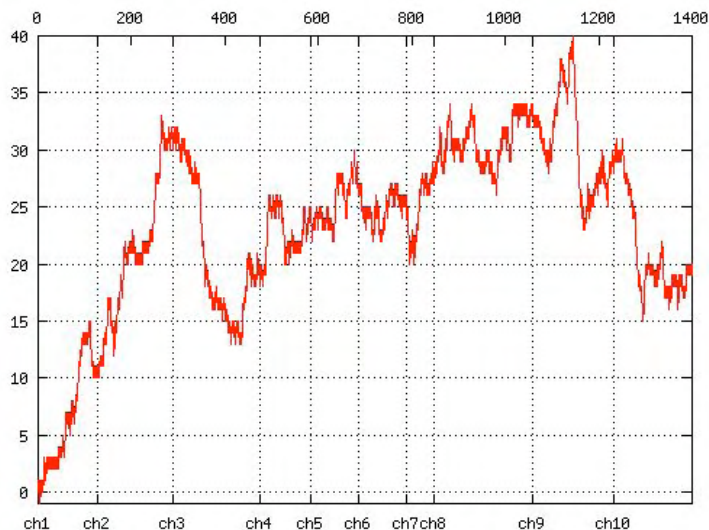


Figure 1.5. Evolution des occurrences **B** et **P**

La *Revue Flaubert* est une revue en ligne et la présentation que nous y avons faite tire parti de cette particularité : il nous a été en effet possible d'y associer notre analyse avec l'intégralité du texte de Flaubert, avec une remise en forme spécifique (fig. 1.6). On aimerait voir ce principe généralisé dans les revues de recherche et poussé plus loin dans son principe. Même si les outils d'édition sont encore insuffisants, le document numérique donne en effet à l'auteur la possibilité d'offrir simultanément au lecteur sa thèse, ses données, ses expériences et les moyens d'en envisager d'autres.

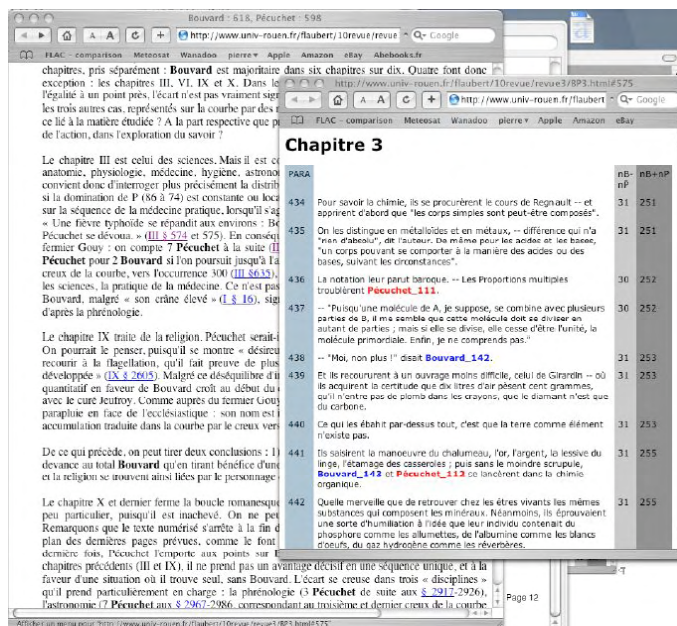


Figure 1.6. Texte colorisé et analyse/glose sur le site de la revue *Flaubert*

Prolongements : le public des bibliothèques numériques

Rassemblons les faits pour nous donner des idées
(Buffon, Histoire naturelle)

Etudier l'usage du service Web de l'ABU a toujours fait partie intégrante du projet. Nous pensons en effet, avec [Le Coadic, 1997, p. 22], que « le but ultime d'un système d'information doit être pensé en fonction des usages qui sont faits de l'information et des effets résultants de cet usage sur les activités des usagers. ». Il nous a donc paru important de publier et d'analyser les statistiques que nous avons de l'usage réel du site [Cubaud, 1995 ; Cubaud et Girard, 1998 ; Cubaud et Topol, 1999]. Une étude d'usage beaucoup plus poussée a ensuite été menée sur le site ABU pour les personnes non-voyantes [Savary, 2002]. La montée en puissance du site CNUM permet de donner à ces études une tonalité nouvelle, car nous avons maintenant la possibilité de procéder à des analyses *comparatives* des usages de bibliothèques numériques. Ce qui, à ma connaissance, n'a pas encore été mené en France.

Outre le texte général [Le Coadic, 1997], on trouvera un état de l'art et un guide méthodologique spécifique aux bibliothèques numériques dans [Bryan-Kinns et Blandford, 2000]. Les études concernant les bibliothèques numériques dont les fonds correspondent aux thématiques d'ABU et du CNUM sont en fait assez

rares^{*}. On citera particulièrement : [Assadi et al., 2003] pour l'étude *Bibusage* menée conjointement par FT R&D et la BNF, complété par [Lupovici et al., 2003] pour les aspects spécifiques à *Gallica* ; les études sur le site *Early Canadiana Online* [Duff et Cherry, 2000 ; Cherry et Duff, 2002] et sur la *Bibliothèque électronique de Lisieux* [Bogros, 2003 ; Bogros, 2004]. Enfin, un aperçu utile des usages généraux de l'internet dans le grand public se trouve dans [Rouet, 2005].

Du point de vue méthodologique, nos analyses se sont basées sur une analyse quantitative, avec comme matériaux les journaux de bord du service. Cette technique est ancienne [Peters, 1993], et s'est popularisée avec le développement de l'Internet et du Web [Cooley, 2003 ; Dhyani et al., 2002]. Même pour des sites modestes, comme ceux de l'ABU et du CNUM, l'analyse des journaux de bord est un processus coûteux en ressource machine. Leur étude à long terme est souvent rendue délicate par les modifications de la structure du site web au fil des années. Mais surtout, ce sont les caractéristiques mêmes du système client-serveur Web qui rendent l'analyse délicate ou, à tout le moins, parcellaire :

- Tout d'abord, il n'est pas possible *a priori* de faire l'amalgame entre site client et utilisateur humain : certains sites sont multi-utilisateurs et même dans le cas de petits ordinateurs, rien ne permet de supposer que c'est toujours la même personne qui se connecte (sauf bien sûr si il y a une procédure d'identification sur le site serveur). Identifier la machine elle-même n'est pas toujours possible car une même machine peut se voir attribuer plusieurs adresses (cas des liaisons intermittentes « dialup »).
- Par ailleurs, le protocole HTTP étant en mode non connecté, il est difficile de reconstituer la session de travail (la « visite ») d'un utilisateur. La solution communément employée consiste à faire des hypothèses sur les durées écoulées entre les requêtes (voir pour une revue [Tanasa et al., 2004]).
- Enfin, le mécanisme de cache de données des logiciels clients et l'utilisation des serveurs de caches « proxies » rendent bien sûr invisibles au serveur certaines requêtes^{*}.

^{*} Beaucoup font en effet une acception assez large du terme "digital library", ce qui limite leur impact sur nos propres travaux. Par exemple, [Jones et al., 1998] est en fait une étude de l'usage fait des requêtes booléennes sur le moteur de recherche d'un (riche) serveur de pre-prints en informatique de l'université de Waikato.

^{*} sauf à employer le mécanisme des « cookies ». Dans une ou plusieurs pages HTML diffusées par le serveur, un script demande au site client d'enregistrer un fichier texte (le « cookie »). Les informations contenues dans le cookie sont ensuite retournées par le logiciel client lors des visites ultérieures. Pour des raisons évidentes de protection de la vie privée, il est possible de configurer son navigateur pour refuser le stockage de cookies. Ce n'est toutefois pas le mode par défaut.

Le tableau ci-dessous récapitule l'état actuel des journaux de bord archivés pour les deux sites à partir de janvier 2000. Au total, c'est près de six millions d'actes de consultations (visites) qui ont été archivés. Leur dépouillement reste à ce jour à accomplir et il est clair que de nombreuses études pourraient être menées à partir de ce matériau.

ABU	Visiteurs différents	Visites	Pages	Hits	Transferts	
Total 2000	241551	343362	2375888	2522555	48.76 Go	
Total 2001	579623	841585	5761795	6107897	116.21 Go	
Total 2002	725592	1029444	7607935	8001653	137.41 Go	
Total 2003	729037	1012981	6614392	6977031	124.03 Go	incident 8/04
Total 2004	430921	565691	3324337	3506377	59.22 Go	incident 4/04 à 10/04
Total 2005	963563	1219462	7843516	8241910	134.84 Go	incident 8/05

CNUM	Visiteurs différents	Visites	Pages	Hits	Transferts
Total 2000	9668	12612	243739	478825	2.68 Go
Total 2001	14286	19425	381364	732900	4.63 Go
Total 2002	30284	42665	1189566	1968043	15.49 Go
Total 2003	101813	141457	3667180	6122392	60.79 Go
Total 2004	200641	268946	6125842	10864681	124.47 Go
Total 2005	211564	272612	8819747	14056966	166.08 Go

Une difficulté pour l'analyse statistique apparaît immédiatement à la lecture de ces tables : la fréquentation des sites n'est pas stable au fil du temps. On considère généralement que le nombre d'ordinateurs (donc sans doute aussi d'utilisateurs) connectés à l'internet a crû de manière exponentielle depuis sa création (avec une loi de la forme $T(n) = 2^n$). Cette croissance générale impacte forcément sur les usages d'un service particulier, mais pas forcément au point de les faire croître eux aussi de manière exponentielle, puisque le service est confronté à une « concurrence » d'offre elle aussi en croissance. Lorsqu'on examine, sur la décennie 1996-2006, le nombre total de textes téléchargés sur le site ABU, on distingue (fig. 7) deux périodes : une période de croissance linéaire jusque vers nov. 2001 (mois n° 68) à laquelle succède une période stable, ou de faible croissance. Il est frappant de constater que la date du pivot correspond aux dernières modifications du site. En quelque sorte comme si se normalisaient depuis cette date les rapports entre le public et la bibliothèque achevée. Pour le site CNUM et beaucoup de sites similaires, nous ne sommes sans doute pas entrés dans cette phase stationnaire.

On peut ensuite chercher quelle part est due dans cette croissance à l'augmentation de l'usage par chacun des usagers. Cette question me paraît très importante dans le cas des bibliothèques numériques, car elle témoigne de

l'appropriation par les usagers de l'outil proposé. Les courbes fig. 1.8 traduisent l'évolution mensuelle du ratio des pages vues par visite pour les deux sites. Les valeurs absolues ne sont pas immédiatement comparables, car l'organisation des pages des sites diffère. Il est cependant clair que, pendant les cinq ans écoulés, les deux sites ont été utilisés de manière très différente. L'interprétation des résultats est délicate. Pour le CNUM, où chaque consultation de page correspond au téléchargement d'un fichier image assez lourd, la hausse progressive peut par exemple s'interpréter par le progrès croissant des équipements des usagers et des infrastructures réseau. On peut aussi évoquer l'accroissement du corpus numérisé, qui rend le site plus attractif. Des études menées sur un panel plus large de bibliothèques numériques permettraient peut-être de répondre à ces interrogations.

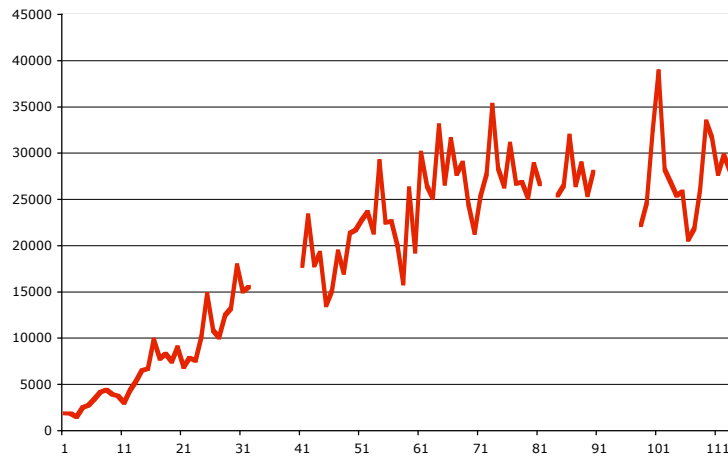


Figure 1.7. Evolution mensuelle des téléchargement sur le site ABU (10/1996 à 2/2006)

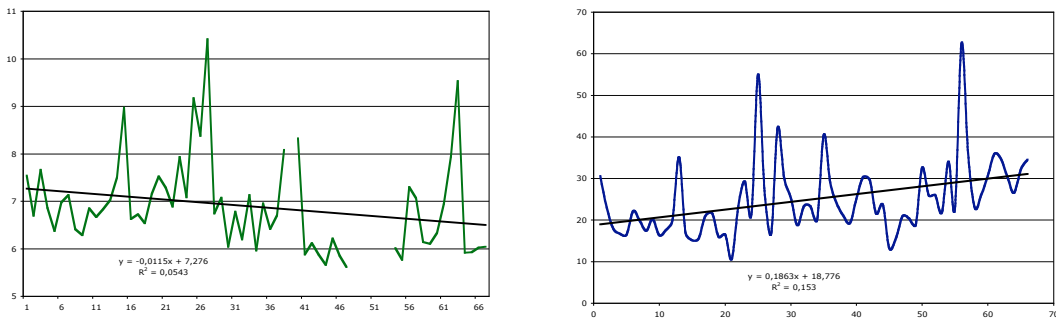


Figure 1.8. Evolution mensuelle du nombre de pages/visite (6/2000 à 12/2005).
Gauche : site ABU. Droite : site CNUM.

La grande déviation autour des moyennes pose une deuxième difficulté d'analyse. Les pics constatés fig. 1.8 correspondent en fait à un très petit nombre d'actes de consultation extrêmement lourds. Lorsque des usagers font un usage répété d'un même service, on observe fréquemment de leur part un comportement s'assimilant à une répétition de tirages au sort indépendants. En notant p , la

probabilité d'arrêter d'utiliser le service, le nombre total d'usages N obéit à la loi géométrique : $P[N = k] = p(1-p)^{(k-1)}$ avec $k = 1, 2, \dots$ (moyenne $EN = 1/p$). La probabilité de petites valeurs de N est donc beaucoup plus grande que celle des grandes. Il en est ainsi du nombre de tâches successives effectuées par un lecteur, ou du nombre de fois qu'un livre est emprunté dans une bibliothèque [Le Coadic, 1997, pp. 38-42 ; Lafouge et al., 2003].

On pourrait penser que de telles distributions se retrouveraient par exemple, pour le nombre de pages vues sur un site par un internaute, ou pour la durée des sessions. En fait, ce n'est pas le cas. On a constaté en général des distributions beaucoup plus variantes (dites « heavy tail » en anglais), c'est à dire dont la fréquence des très grandes valeurs n'est plus complètement négligeable devant celle des petites. Aux grandes valeurs, ces distributions ont une allure hyperbolique, comme la loi de Pareto : $P[Y > t] = a / (t^b)$ avec $0 \leq a < t$ et $b > 1$. Une littérature considérable a été produite à ce sujet depuis le début des années 1990 [Adler et al., 1998], car ce comportement impacte sur les performances des infrastructures de l'Internet [Pastor-Satorras et Vespignani, 2004]. On retrouve ce type de phénomène sur les sites ABU et CNUM. La figure 1.9 reproduit la répartition du nombre total de téléchargements (ABU) et de consultations de pages de texte (ABU et CNUM) par visiteur, sur la période 7/2000-6/2001. Les trois distributions sont hyperboliques (linéaires sur l'échelle logarithmique utilisée pour les graphiques). La similarité des distributions du téléchargement d'ABU et du « feuilletage » de fac-similés sur le CNUM peut s'interpréter par la similarité des tailles de fichiers associés.

Un phénomène identique de grande variation se retrouve au niveau de l'utilisation des ressources offertes (règle dite des "80-20", où 20% des services concentrent 80% des usages). On trouvera dans [Cubaud et Girard, 1998] un dépouillement des usages par rubriques du site ABU. Ces statistiques sont stables au fil des années et un phénomène similaire est décrit dans [Bogros, 2003 ; Lupovoci et al., 2003]. Il en est de même avec la répartition des usages pour les textes téléchargés (fig. 1.10).

Croissance et grande déviation font que la plupart des études quantitatives d'usage sur Internet n'ont souvent qu'un intérêt circonstanciel, voire douteux dès lors que les analyses se portent sur un usage « moyen » sans réalité statistique. Interroger les pratiques de lectures dans les bibliothèques numériques ne pourra se faire que sur une échelle de temps assez longue, en impliquant le plus grand nombre possible de sites.

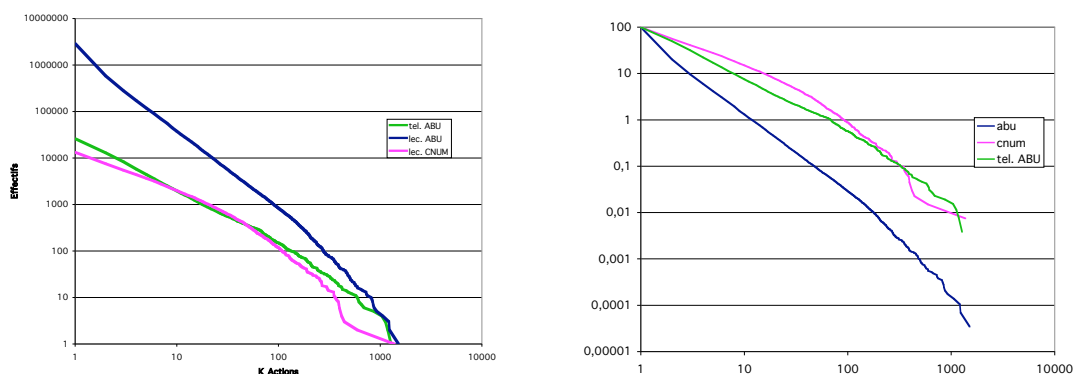


Figure 1.9. Répartition du nombre total de téléchargement intégral (ABU) ou de consultation de page de texte (ABU et CNUM) par visiteur. A gauche : effectifs cumulés, à droite : idem, origine 100.

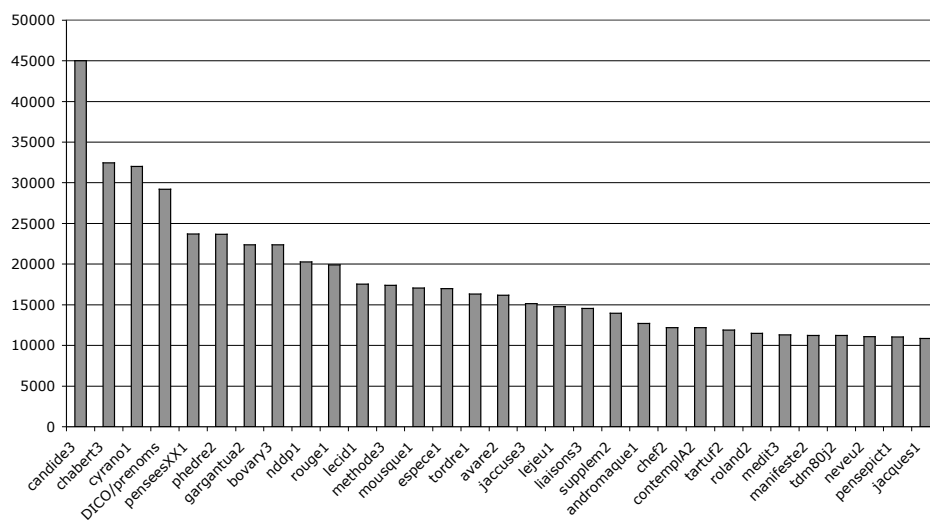


Figure 1.10. Trente premiers textes les plus téléchargés sur le site ABU (2/2000-2/2006)

2. Interaction 3D pour les bibliothèques numériques

Au chapitre des « inventions à faire » de son traité fondateur sur la documentation paru en 1934, Paul Otlet avait décrit à grands traits un procédé qu'il baptise le « livre téléphoté » [Otlet, 1934, p. 238]. On ne peut s'empêcher d'y voir quelques analogies avec le programme des recherches actuelles en matière de bibliothèques numériques :

« 8 - Télélecture - comme application particulière de la télévision. 1° Donner des textes en lecture à distance. 2° Permettre à chacun par un dispositif approprié de prendre connaissance à distance de textes publiquement exposés à cet effet. 3° Permettre la vision à distance des textes de livres disposés sur les rayons d'une bibliothèque ou des feuilles réunies en dossier dans les classeurs. (Service : accroître ainsi la diffusion des collections des bibliothèques et des offres de documentation.) » [Otlet, 1934, p. 289 et suiv.]

Ce n'est finalement pas la télévision qui a pu servir de médium au « livre téléphoté », mais le Web et l'internet avec le succès que l'on sait. Toutefois, si le Web et ses technologies associées (formats de description de métadonnées, systèmes d'échange et de propagation de requêtes) permettent certainement de « prendre connaissance » des centaines de milliers d'ouvrages numérisés actuellement en ligne sur Internet, il faut bien reconnaître qu'il est difficile d'en avoir « la vision ».

Héritière de la cartographie et du graphisme technique [Bertin, 1967 ; Tufte, 1992], la visualisation d'informations non quantitatives est un domaine récent de la recherche en informatique. Elle est encouragée par les progrès des matériels d'affichage graphique et motivée par l'explosion des masses d'informations disponibles sous forme numérique. [Gershon et Eick, 1995] la définit comme « a process of transforming data and information that are not inherently spatial, into a visual form allowing the user to observe and understand the information ». Son but est d'utiliser nos capacités visuelles extrêmement performantes pour percevoir, analyser et mémoriser plus facilement des données abstraites complexes. De nombreux dispositifs et logiciels d'interaction humain-machine ont été proposés pour aider l'utilisateur dans son appréhension de grandes organisations textuelles en réseaux (hypertextes, résultats de recherches) [Benford et al., 1999 ; Mukherjea, 1999], mais il reste encore beaucoup à faire pour évaluer leur potentiel dans le contexte des bibliothèques numériques [Chen et Börner, 2002 ; Bertini et al., 2005]. Comme son titre le suggère, mon exposé se limitera à l'étude du potentiel des techniques d'interaction 3D, dans la lignée du texte

fondateur [Robertson et al., 1993]. D'autres procédés existent, décrits dans les ouvrages de synthèse [Card et al., 1999 ; Chen, 1999]. On trouvera également un état de l'art dans le mémoire de thèse [Topol, 2002].

Le livre et les interfaces de lecture

[Cubaud, Topol et Vodislav, 2000 ; Cubaud et Topol, 2001 ; Topol, 2002]

Les interfaces de bibliothèques numériques reposent actuellement sur HTML et quelques technologies associées. Prenons comme illustration celles réalisées au CNAM : ce ne sont sans doute pas les plus importantes, ni les plus sophistiquées, mais elles me paraissent assez représentatives.

La littérature scientifique et technique a plusieurs caractéristiques intéressantes du point de vue éditorial. Les illustrations y jouent par exemple un très grand rôle. Celles-ci sont généralement groupées en planches séparées (jusque vers la moitié du XIXe siècle) et ces planches ont une très grande variété de format, parfois au sein d'un même volume. La structure éditoriale est en général très sophistiquée (par opposition au roman, par ex.), mais aussi parfois incohérente quand la publication s'étale sur de longues années. Les citations et les notes jouent également un rôle essentiel dans les publications à caractère académique. Une lecture séquentielle sera donc relativement rare, et ces livres seront rarement lus isolément.

Nous avons essayé de prendre en compte cette situation lors de la définition de l'interface du CNUM :

- l'utilisateur peut facilement ouvrir plusieurs fenêtres pour naviguer parmi différents fac-similés de textes, ou parcourir leurs tables.
- des fonctions de zoom permettent d'adapter la page courante à la taille de la fenêtre qui la contient.
- des outils de navigation spécifiques ont été réalisés pour passer de planche en planche, ou pour aller directement d'une page quelconque aux planches mentionnées dans cette page.

Ces fonctions ont été réalisées en HTML dynamiques, sans recours à un plug-in extérieur (une contrainte fixée dans notre cahier des charges). La fig. 2.1 est une copie d'écran d'une session de travail hypothétique, consacrée à l'arithmétique de Pascal. On y trouve une page du catalogue du CNUM (a), ainsi que deux fenêtres de consultation de document (b), dont l'une réduite par l'utilisateur. Une autre fenêtre correspond à un autre texte de Pascal sur le site de l'ABU (c).

L'utilisateur prend également des notes ou des extraits de texte avec son logiciel favori (d).



Figure 2.1. Copie d'écran d'une session de travail avec les sites CNUM et ABU.

Malgré ces efforts, il semble bien que la plupart des usagers de bibliothèques numériques téléchargent leurs documents pour les imprimer. Les courriers d'usagers du CNUM mentionnent pour l'essentiel des problèmes relatifs à l'impression des documents et à leur téléchargement intégral. J'ai présenté au chapitre 1 les résultats d'une étude comparative pour le site de l'ABU de l'usage des fonctions de téléchargement comparé à celles de feuilletage. On y a constaté un rapport de l'ordre de 8 actes de feuilletage pour un téléchargement en moyenne. L'étude *BibUsages* effectuée sur le site Gallica montre un rapport encore plus décevant pour le feuilletage, de l'ordre de 1,7 (548/314) [Assadi et al., 2003, p.32]. De ce point de vue, les interfaces Web n'offrent finalement qu'un progrès limité par rapport à celles du FTP utilisées auparavant.

De nombreuses autres interfaces que le Web sont actuellement utilisées pour la consultation de fac-similés d'ouvrages. La plus répandue est certainement celle d'Acrobat (Adobe), mais son utilisation pour de gros ouvrages numérisés en mode image reste délicate (la navigation dans les vignettes est par exemple peu utile dans ce contexte, la recherche d'occurrences limitée et lente). Des systèmes expérimentaux ont été développés dans plusieurs laboratoires : [André et al., 1995], BAMBI [Calabretto et al., 1998], Philectre [Lecolinet et al., 1998 ; Robert et Lecolinet, 2000] et DEBORA [Debora, 2002] par exemple (pour nous limiter à des travaux français récents). Comme le Web, toutes ces interfaces ont en commun de reposer sur les briques de bases du paradigme WIMP (Window, Icon,

Menu, Pointer). Les limites de ce procédé apparaissent clairement dans l'encombrement de l'écran de la fig. 2.1. Il serait bien délicat d'y ajouter une fenêtre pour un texte supplémentaire. De plus, l'organisation même de ces fenêtres est en soi une activité coûteuse, en terme de clics, de déplacements, d'iconifications.

L'effort imposé à l'utilisateur d'un système de lecture à l'écran a été analysé dans [O'Hara et Selen, 1997], en comparaison à la lecture conventionnelle. Ce travail a contribué à mettre en évidence le besoin de disposer pour les interfaces de lecture, de techniques de navigation plus fluides, ainsi que d'une plus grande flexibilité dans le contrôle de l'organisation spatiale des documents. Des travaux plus récents ont confirmé ce constat et [Marshall et Bly, 2005] insiste sur l'écart existant entre le livre physique et son feuilletage et la navigation au sein d'une interface de bibliothèque numérique, en termes de performance et de confort de lecture. La majorité des lecteurs sont par ailleurs gênés par la luminosité des écrans informatiques et par la nécessité de devoir faire défiler verticalement le document. [Gould et al., 1984] a mis en évidence le rapport entre la représentation et la vitesse d'acquisition des informations. Pour les documents numériques, on a constaté que la résolution du dispositif de sortie est un élément déterminant dans la vitesse de lecture [Gould et al., 1987]. Or, la résolution d'un écran excède rarement 100 points par pouce alors qu'un confort optimal de lecture sans distinction du moiré nécessite une résolution supérieure à 300 points par pouce. Malgré ces éléments négatifs, fortement dépendants des progrès technologiques, la lecture sur écran est devenue une pratique courante et acceptée depuis l'avènement de l'Internet. Il reste toutefois d'autres aspects d'ordre logiciel à améliorer pour, par exemple, faciliter la lecture croisée ou l'organisation d'un espace de travail. En cela, une solution 3D offre des possibilités de représentation beaucoup plus riches que la 2D sans pour autant pénaliser de manière rédhibitoire le rendement des utilisateurs comme le montrent plusieurs études [Sebrechts et al., 1999 ; Cockburn et al., 2001 ; Westermen et al., 2005].

Le *Web Forager* [Card et al., 1996] est une interface 3D conçue pour organiser une collection de pages HTML. Bien qu'il ne traite pas directement des bibliothèques numériques, les métaphores proposées en sont issues. Cet outil de data mining propose une interface très conviviale et intuitive dans laquelle l'organisation des informations se fait sur deux niveaux différents. Le premier autorise l'utilisateur à regrouper des pages HTML dans une sorte de classeur, les *WebBooks* qui ont l'apparence et la structure séquentielle des livres. Plusieurs outils sont offerts pour pouvoir manipuler ces classeurs, parmi lesquels un outil de feuilletage et une loupe à document pour pouvoir examiner toutes les pages en un

coup d'œil [Card et al., 1991]. Le deuxième niveau d'organisation concerne le rangement des WebBooks. L'utilisateur dispose pour cela de différents « meubles » pour y déposer les livres. Il pourra, par exemple, ranger des livres qu'il n'est pas certain de réutiliser dans un avenir proche dans les étagères alors qu'une collection de pages qu'il utilise souvent, pourra rester sur le bureau au premier plan. L'utilisation de telles métaphores (étagères, bureau) permet, d'après les auteurs, de profiter de la mémoire spatiale de l'utilisateur et de la rendre encore plus efficace en se reposant sur les habitudes de rangement que l'on peut avoir dans un bureau réel.

Les détails d'implémentation du Web Forager ne sont pas explicités et aucune démonstration n'est plus disponible. Beaucoup de questions se posent quant à son fonctionnement interne. Le point de vue est-il fixe ? Quelles sont les contraintes de positionnement des livres ? L'interaction est-elle fluide ? Néanmoins, les principes décrits par cette interface (la spatialisation des livres, le feuilletage interactif, l'organisation de l'espace de travail) sont une grande source d'inspiration même si nous ne partageons pas cette approche orientée réalité virtuelle très contrainte qui induit une surcharge inutile de l'interface avec des objets réels. Certes, il est important de limiter les mouvements et les actions de l'utilisateur dans une interface 3D sous peine de la rendre inutilisable. Mais certaines métaphores ne nous paraissent pas judicieuses pour permettre un travail efficace sur les livres. Le bureau, par exemple, oblige l'utilisateur à poser les livres à plat, horizontalement par rapport au point de vue. Nous doutons que la lecture soit possible à cause de cette vue en perspective. Un livre positionné perpendiculairement à la direction du point de vue permet une lecture nettement plus confortable. Le Web Forager n'interdit pas de lire verticalement un livre en arrière plan mais la place de choix au premier plan est tenue par le bureau et son plateau horizontal.

Dans [Cubaud et Topol, 2001], nous avons décrit une première expérience d'un tel environnement de lecture basé sur des fenêtres de consultations 2D translucides suspendues dans un espace 3D (fig. 2.2 haut). L'objectif de l'étude était d'étudier le potentiel d'un dispositif de type WebBook pour la manipulation de fac-similés. Réalisée en VRML2 scripté, l'interface implémente les fonctions classiques de manipulation de fenêtres (création, maximisation, iconification, fermeture) et surtout la spécification du point de vue de l'utilisateur. Un mécanisme de punaisage des fenêtres rend possible d'associer/dissocier la position des fenêtres au point de vue. Le zoom résultant du déplacement de point de vue peut ainsi n'être affecté qu'aux fenêtres choisies. Les possibilités offertes par VRML pour programmer de telles scènes interactives se sont avérées à l'usage

limitées [Cubaud et al., 2000 ; Topol, 2002] et nous nous sommes tournés vers d'autres technologies à partir de 2002 (décrites dans la section suivante).

Dans l'interface reproduite fig. 2.2 (bas), la scène 3D est limitée verticalement par un sol. La caméra est fixée en permanence au-dessus de ce sol et son champ est fixe. Les fac-similés sont présentés sur un outil spécifique en forme de trépied (sorte de lutrin simplifié). Le livre peut être positionné de manière arbitraire sur le sol, repoussé, tiré ou pivoté par une action de l'utilisateur sur la barre verte au bas du trépied. Cette manipulation s'effectue en temps réel et de manière continue avec un pointeur ordinaire du type souris : deux degrés de liberté suffisent en effet pour déterminer la position du trépied sur le sol. Plusieurs trépieds peuvent ainsi être positionnés sur le plan de travail. Le déplacement d'un trépied peut être contraint par la présence des autres si on choisit de détecter les collisions éventuelles de ces objets 3D. La destruction du trépied s'effectue par un clic de bouton droit sur la barre verte. Il est possible d'imaginer d'autres raccourcis pour l'interaction pour, par exemple, permettre le passage rapide du trépied en premier plan (mode plein écran). Ces raccourcis devraient trouver leur place sur des équivalents de boutons ou de menus sur la base du trépied.

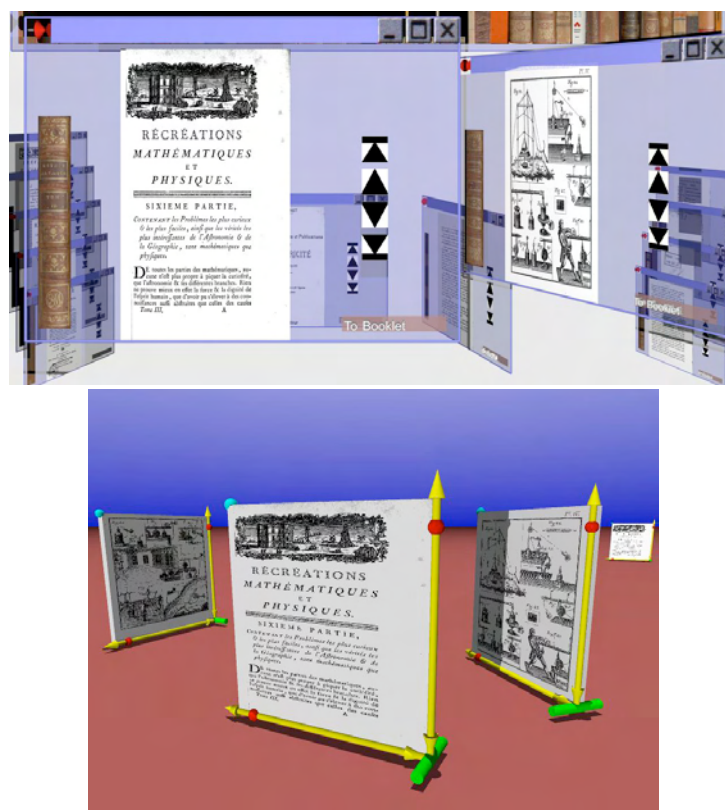


Figure 2.2. haut : fenêtres flottantes. bas : trépied de lecture.

Le défilement dans les pages du livre se fait par un clic sur la page courante. Comme pour une fenêtre 2D conventionnelle, le trépied dispose sur son axe vertical d'un « ascenseur » pour accélérer la progression dans le document (représenté fig. 2.2 par une boule rouge). L'axe horizontal permet de contrôler le rapport entre la dimension (en pixels) du fac-similé et celle choisie pour le trépied. On réalise ainsi une fonction de zoom indépendante de la position du trépied sur le plan de travail. La taille du trépied lui-même est négociable par le biais d'une poignée spécifique (boule bleue en haut à gauche du trépied).

L'épaisseur du livre pourrait également être prise en compte. La mise en perspective des tranches du livre permettrait de mieux localiser la page courante dans la globalité du texte. Les tranches pourraient être augmentées de signets correspondants aux chapitres ou à des passages sélectionnés par l'utilisateur, comme dans le système *3Book* [Card et al., 2004]. Il est également possible de modifier l'apparence du trépied pour permettre l'affichage du livre ouvert. L'action de feuilletage peut alors être simulée de manière assez convaincante par une animation de la circulation de la page courante d'un bord vers l'autre (fig. 2.3). Un procédé similaire (non 3D) avait été proposé dans l'expérience *Turning the page* de la British Library [Carpenter et al., 1998]. Il est maintenant régulièrement utilisé - avec grand succès - dans des expositions grand public de textes précieux (exposition des manuscrits de Léonard de Vinci, au Louvre en 2004, par ex.). Une animation plus sophistiquée, basée sur le comportement physique d'une feuille de papier fléchie, a été décrite dans [Witten et al., 2003 ; Chu et al., 2004]. La fonction de feuilletage est surtout intéressante si elle permet une lecture de survol. Sa mise en œuvre est à ce jour problématique si on souhaite atteindre la vitesse du pouce qui parcourt toute l'épaisseur des pages [Marshall et Bly, 2005].

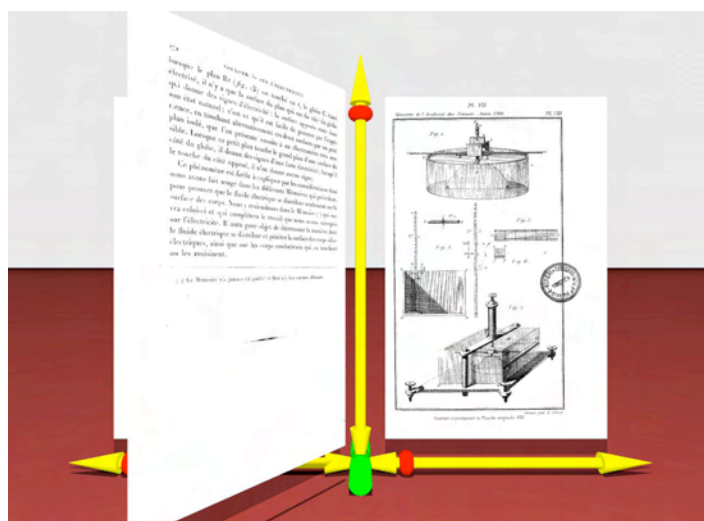


Figure 2.3. Le trépied en mode feuilletage.

Le design proposé pour le trépied/lutrin de lecture est évidemment arbitraire. Pour des raisons de performance du rendu 3D, nous l'avons limité à des primitives géométriques élémentaires, et l'analogie avec une fenêtre conventionnelle est voulue. L'intérêt principal du dispositif, en comparaison aux propositions plus récentes de [Card et al., 2004] et [Chu et al., 2004], réside dans la possibilité de faire coexister un assez grand nombre de trépieds sur le même plan de travail. Il est également possible de le compléter par des outils de recherche et de navigation dans des collections.

La collection et les interfaces de recherche

[Cubaud, Thiria et Topol, 1998 ; Thiria 1999 ; Topol, 2002 ; Cubaud, Stokowsky, Topol, 2002 ; Cubaud, Dupire, Topol, 2003]

La figure 2.4 reproduit l'interface en mode « parcours » de trois bibliothèques numériques très représentatives : le CNUM (CNAM), Gallica (BNF) et *Memory of America* (lib. of Congress). Aucune d'entre elles ne pourrait satisfaire un lecteur contemporain comme Alberto Manguel [Manguel, 1998, p. 155] :

« Mes mains, lorsque je choisis un livre à emmener au lit ou à poser sur un pupitre, à offrir ou à lire dans le train, tiennent compte de la forme autant que du contenu. En fonction de l'occasion, en fonction de l'endroit où j'ai décidé de lire, je préférerai quelque chose de petit et intime, ou d'ample et de substantiel. Les livres s'affirment grâce à leurs titres, leurs auteurs, leurs places dans un catalogue ou une bibliothèque, leurs illustrations de couvertures. Par leur taille aussi. En fonction des époques et des lieux, j'ai appris à attendre des livres des apparences diverses et, comme dans toutes les modes, ces traits changeants attachent un caractère précis à la définition d'un livre. Je juge un livre à sa couverture ; je juge un livre à sa forme. »

En pratique, le mode « parcours » est toujours associé à un moteur de recherche : une fonction jugée indispensable dès lors que le corpus atteint une certaine taille. Aussi sophistiqué soit-il, ce type de fonctionnement n'a d'intérêt que pour l'utilisateur muni d'un projet préconçu : il n'est donc pas possible de flâner dans une bibliothèque numérique...

Il est assez simple de compléter les interfaces strictement textuelles des outils de recherche par un fonctionnement beaucoup plus graphique, par exemple en numérisant les dos d'ouvrages à leurs dimensions respectives (fig. 2.5). Cependant, pour les trois sites mentionnés, ces informations n'ont pas été acquises lors de leur conception. On peut également penser que dans bien des cas, la couverture ne sera pas assez discriminante. A partir des quelques informations disponibles sur la constitution physique des ouvrages (pagination, date de publication, ...) il serait peut-être possible de bâtir une interface plus abstraite, reposant sur une « grammaire » de formes préétablies dont il faudrait dresser

l'inventaire. Une approche similaire a été suivie dans le projet *LibViewer* [Rauber et Bina, 2000]. Dans ce système, les métadonnées sont interrogeables et le résultat des requêtes est affiché sous la forme d'étagères de livres en pseudo 3D. Ces interfaces ont toutefois l'inconvénient de ne pas permettre un affichage simultané d'un grand nombre de volumes. Une fonction de zoom est bien sûr réalisable, mais son utilisation dans un contexte graphique 2D ne permet pas une navigation fluide. Nous retrouvons là les limites du procédé WIMP évoquées dans la section précédente.

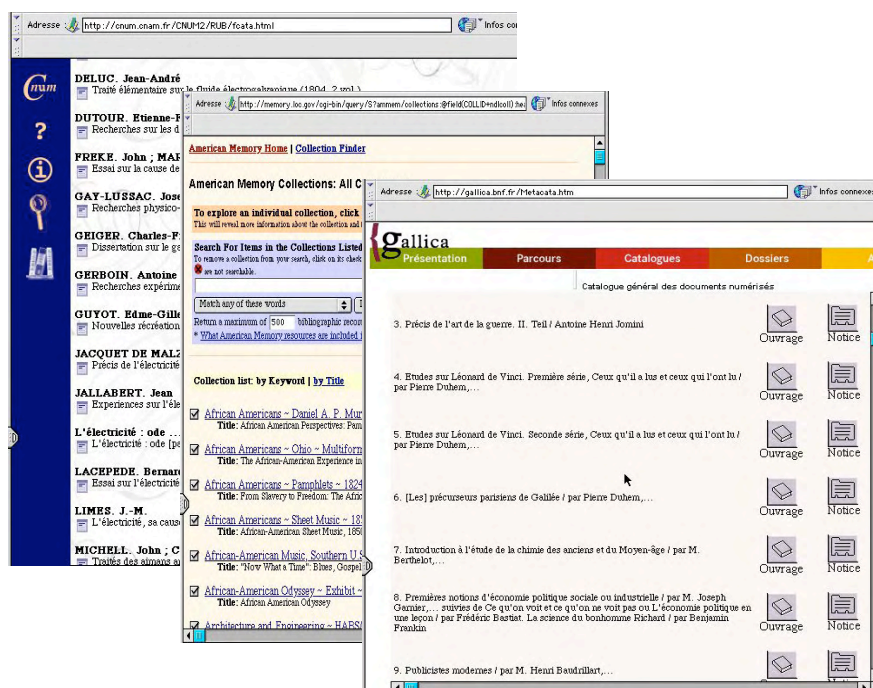


Figure 2.4. Extrait de trois bibliothèques numérisées (pages de catalogues)

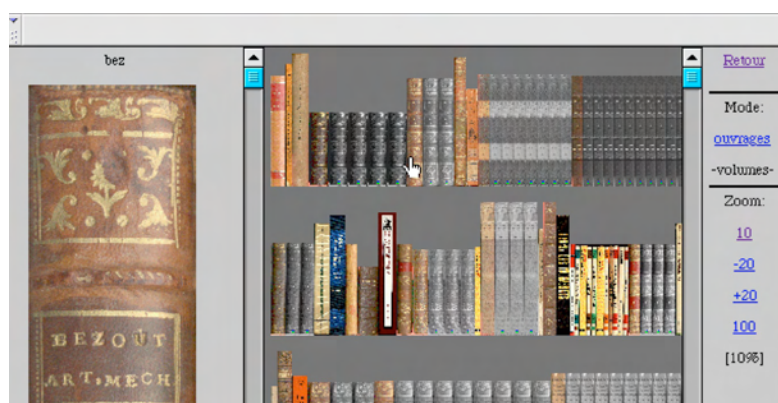


Figure 2.5. Interface visuelle alternative (<http://jasmin.cnam.fr:8081/REL/>)

Le parcours dans une bibliothèque par des techniques de réalité virtuelle a été étudié par plusieurs équipes. Une première représentation 3D d'une bibliothèque existante (avec VRML) a été réalisée à l'IRCAM [Fingerhut, 2000]. Des expériences similaires ont eu lieu à l'école des mines de Nantes [Plénacoste et al., 2001] et à l'université de Karlsruhe [Christoffel et Schmitt, 2002]. Dans ces interfaces, l'organisation de la bibliothèque réelle est respectée (salles de travail, zones de rayonnage, voire couloirs et escaliers). Au delà de leur intérêt indéniable en tant qu'outil de communication, de tels procédés ont évidemment des limites. La navigation d'une salle virtuelle à l'autre est forcément coûteuse en temps (on n'ose pas imaginer une telle interface représentant la bibliothèque François Mitterrand...). De plus, toute collection numérisable n'est pas forcément accessible au public, ni même associée à un établissement précis. On trouvera ainsi dans [Fox et Kipp, 1997] le compte-rendu d'expériences menées pour la navigation en 3D dans un grand corpus national de thèses numérisées et dans [Jacquemin et Jardino, 2002] la présentation d'une interface 3D générique pour la visualisation dans de grands documents XML.

Dans [Cubaud et al., 1998], nous avons présenté une interface du même type pour les utilisateurs du CNUM. Ceux-ci peuvent naviguer à l'intérieur d'une scène VRML représentant de grands rayonnages virtuels. Plusieurs architectures ont été étudiées : linéaires ou structures hélicoïdales (fig. 6), en premier lieu [Thiria, 1999]. Pour faciliter la navigation dans la scène, une structure cylindrique paraît satisfaisante. Dans une telle scène (fig. 2.7), l'utilisateur peut pivoter le point de vue et se déplacer pour zoomer vers les étagères. Le rangement de la collection peut également être recomposé selon des requêtes de l'utilisateur. Dans ce cas, le point de vue est orienté vers les ouvrages correspondant à la requête, tandis que le reste de la collection est placé dans le dos de l'utilisateur. Pour des requêtes à deux critères (par ex., « date de publication = XVIIIe ET sujet = électricité »), une rotation du point de vue montre les livres obéissant au premier critère, tandis que l'axe vertical permet de naviguer dans la collection selon le deuxième. Ces deux mouvements s'effectuent facilement avec une souris ordinaire. Des recherches multi-critères plus générales peuvent également être abordées avec des géométries comme les hyper-sphères [Topol, 2002], au prix d'une complexification de la navigation dans l'espace de réponses (fig. 2.8).

La standardisation en cours des techniques d'inter-connexion de bibliothèques numériques pour l'échange de métadonnées (notamment par le biais du protocole de l'*Open Archive Initiative**) offre de nouvelles perspectives d'expérimentation,

* <http://www.openarchive.org>. On trouvera aussi des tutoriels très complets sur l'OAI sur les sites de la BNF et du ministère de la culture.

avec des corpus beaucoup plus riches. Depuis janvier 2006, le CNUM et le site Gallica sont connectés par OAI, il en sera bientôt de même avec le portail du *Système Universitaire de Documentation* (SUDOC, ABES), actuellement riche de près de cinq millions de références. Le couplage de l'interface de visualisation avec l'entrepôt de métadonnée a été étudié dans [Shen et al., 2002 ; Ong et Leggett, 2005], mais il n'existe pas à ce jour de résultats d'expérience de visualisation sur de telles masses de données.

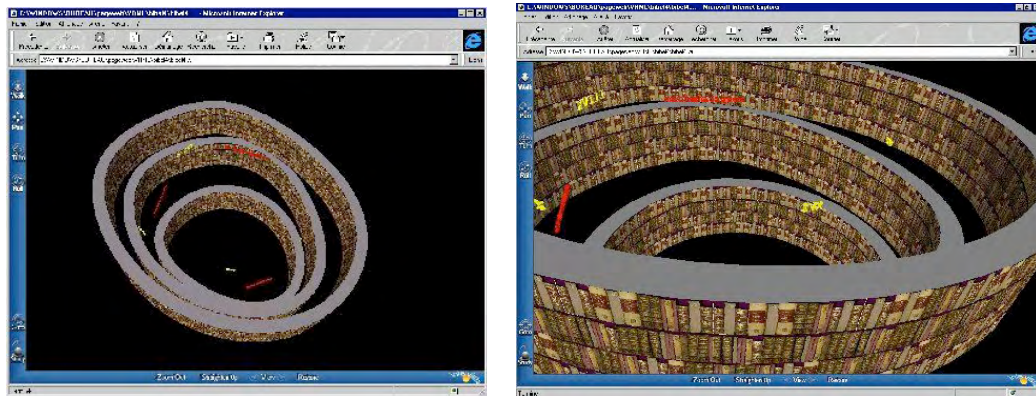


Figure 2.6. Scène 3D pour la navigation dans une collection

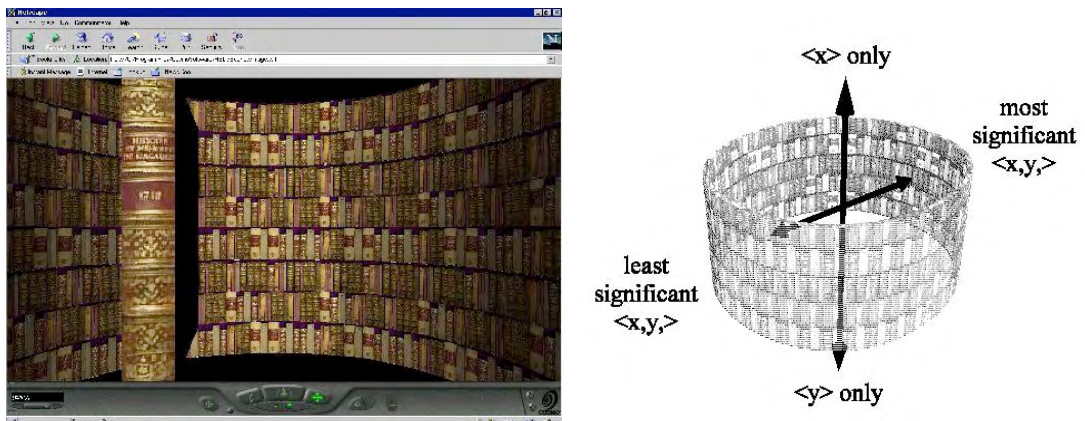


Figure 2.7. Scène VRML pour les rayonnages virtuels. L'ouvrage sélectionné apparaît au premier plan (gauche). Critères de classement des étagères (droite).

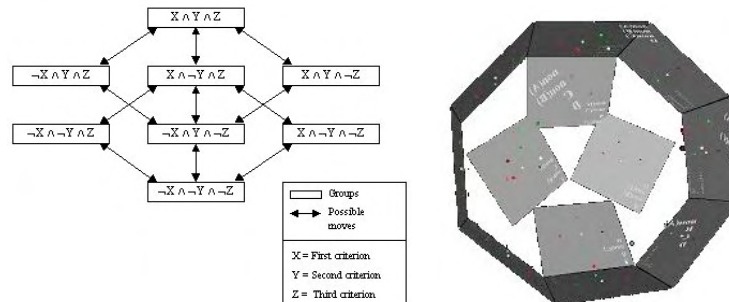


Figure 2.8. Représentation d'un résultat de recherche multi-critères sous la forme d'hyper-sphère [Topol, 2002].

COMBINER RECHERCHE ET LECTURE

L'étagère virtuelle peut être insérée dans l'interface générale de lecture de plusieurs manières. Un compromis doit être trouvé pour éviter l'encombrement de la surface de l'écran (ce qui tend à disperser les objets de la scène), et réduire le temps de navigation dans la scène (qui s'allonge si les objets sont distants). La fig. 2.9 (gauche) reproduit une organisation « en cockpit » de la scène 3D, assez semblable à celle du *Web Forager* [Card et al., 1996]. La collection est toujours visible en arrière-plan et le zoom sur les rayonnages n'affecte pas le positionnement des trépieds. On comprend qu'avec cette disposition, l'écran se surcharge dès que quelques trépieds sont activés, et une grande partie des rayonnages devient invisible. Une scène plus large doit donc être définie. Sur la fig. 2.9 (droite), la zone d'évolution des trépieds est organisée en disque autour de l'utilisateur. Par une rotation à 180° sur l'axe Y, il passe de manière fluide de la zone des trépieds à celle des rayonnages. Il garde également la possibilité de positionner des trépieds devant les rayonnages. D'autres organisations sont bien sûr possibles. On trouvera dans [Cubaud et Topol, 2001] et [Cubaud et al., 2002] des propositions de designs alternatifs.

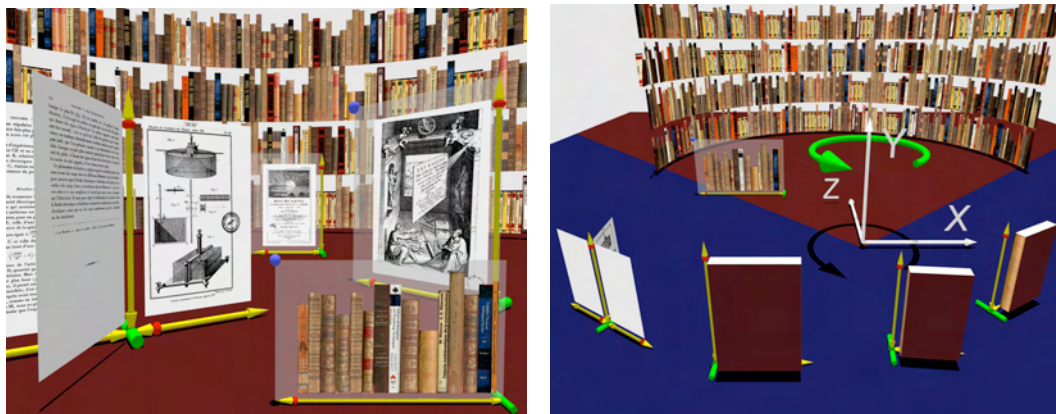


Figure 2.9. Gauche : toute la collection est affichée et accessible en fond. Les ouvrages favoris (panier) sont accessibles en avant-plan. Droite : organisation de l'espace de travail pour la lecture et la recherche dans la collection.

Un premier démonstrateur (basé sur OpenGL) a été présenté à des professionnels des NTIC lors des conférences JFT'2003 [Cubaud et al., 2003b] et ECDL'03 [Cubaud et al., 2003]. Une audience plus large a été atteinte durant l'exposition grand public *Image par Image* (Montreuil, mars 2003) et les journées nationales *Sciences en fête* auxquelles le CNAM participe (oct. 2003). A chacune de ces expositions, nous avons installé un poste dédié à l'atelier 3D et un autre pour l'accès au site web du CNUM. La session était limitée par une horloge à quelques minutes et les actions des utilisateurs étaient enregistrées. Nous avons pu

conclure de ces démonstrations que les utilisateurs atteignent vite une certaine aisance dans la manipulation des trépieds de lecture. La fonction de pagination a été bien accueillie du grand public et des bibliothécaires (mais pas toujours par les informaticiens professionnels). Quelques modifications du démonstrateur ont paru nécessaires. L'amplitude de déplacement des trépieds a été bornée : un trépied ne peut pas être plus grand que l'écran ni réduit à une taille limite raisonnable. La détection de collision entre les objets a été améliorée, ainsi que l'éclairage et le rendu des ombres.

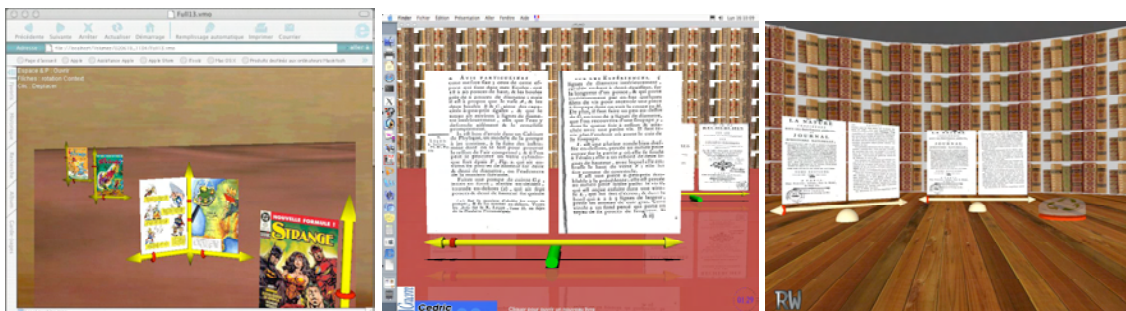


Figure 2.10. Démonstrateurs de l'atelier de visualisation 3D avec diverses technologies de programmation 3D. De gauche à droite : Virtools, OpenGL, Criterion renderware .

Plusieurs fonctions de notre atelier de visualisation 3D de documents restent à mettre en œuvre. Le démonstrateur actuel n'inclut en effet qu'une collection simplifiée, la création/destruction des trépieds et le feuilletage. Nous entreprenons à cette fin une réécriture complète du logiciel en utilisant *Criterion Renderware* [Cubaud et al., 2005b]. Plusieurs logiciels très intéressants ont en effet été introduits récemment dans l'industrie du jeu vidéo pour s'affranchir des limitations de standard de programmation 3D tels que OpenGL et Java3D (fig. 2.10).

La comparaison des pratiques de lectures des utilisateurs du CNUM avec celles résultant de l'utilisation de l'atelier 3D sera notre prochaine tâche. Nous souhaitons associer des méthodes quantitatives, comme celles présentées au chapitre précédent, avec des entretiens de « grands lecteurs » du CNUM. La banalisation progressive du matériel 3D et des connexions internet haut débit permettent d'espérer que l'accès aux bibliothèques numériques deviendra « as easy as a game » [Christoffel et Schmitt, 2002] dans un futur proche. Il reste auparavant à comparer, dans un cadre expérimental commun, les procédés proposés par les différentes équipes qui contribuent à ce thème de recherche.

Prolongement : le document et son contexte

Parallèlement aux travaux sur l'interaction pour les documents textuels cités au chapitre précédent, de nombreuses expériences sont menées en matière de numérisation 3D pour les musées. Après la numérisation d'objets isolés (projet *Digital MichelAngelo* [Levoy et al., 2000], par ex.), on s'est intéressé à la constitution et l'interrogation de grandes bases d'objets numérisés puis d'interfaces d'accès permettant la visite d'expositions virtuelles [Hemminger et al., 2005 ; Alisi et al. 2005]. Ce dernier type d'application rend particulièrement nécessaire l'emploi de métadonnées textuelles, dont la présentation reste totalement isolée des interfaces de navigation des objets 3D.

Il nous semble dès lors intéressant d'étudier les modalités de la coexistence, au sein d'un unique environnement de visualisation 3D, de documents textuels et d'objets 3D. Une application de tels dispositifs se trouve en histoire des techniques, où l'on voudrait pouvoir associer des reproductions d'appareils scientifiques, par exemple, et les ouvrages qui les décrivent. Nous nous intéressons en particulier aux situations où ce type d'association contextuelle n'est pas conçue *a priori*, par l'auteur d'un hypermédia, mais *a posteriori* par un lecteur qui a recours à plusieurs sources d'informations qu'il confronte. Cette mise en contexte de documents est également utile quand le document étudié est le résultat d'un processus éditorial complexe. C'est par exemple le cas pour la bande dessinée (maquettes, photos, différents états des planches originales, bleus, épreuves, etc.) qui n'a fait l'objet, à notre connaissance, d'aucune étude spécifique en matière d'interfaces. Notons aussi qu'un nombre considérable d'utilisateurs du Web accomplissent au fond la même tâche en confrontant toutes sortes de tarifs et de brochures descriptives textuelles et 3D pour acheter mieux ou plus.

Prenons comme exemple une situation que nous jugeons assez typique des usagers du site *Conservatoire numérique*. Il s'agit ici d'étudier le fonctionnement d'un appareil complexe ancien (ici, un polarimètre*) à partir de documents historiques. La polarimétrie est une technique de chimie analytique qui permet le dosage de la teneur d'une solution en substances optiquement actives. On y a recours pour évaluer par exemple la teneur en sucre du jus de betterave. Le

* Dans cet appareil, un faisceau lumineux, engendré par une source monochromatique, traverse un polariseur (lame de quartz). Une moitié de la lumière qui sort polarisée linéairement va directement vers la solution à analyser, tandis que l'autre moitié traverse une lame demi onde. Les deux faisceaux, avec plans de polarisation différents, sont analysés par un analyseur. Sur l'oculaire se trouve un vernier circulaire. En tournant l'analyseur, l'intensité des deux faisceaux change. On obtient deux angles dans lesquels les deux moitiés sont faiblement ou fortement illuminées mais avec la même intensité. La différence d'angles relevée est proportionnelle à la concentration de la solution étudiée (loi de Biot).

polarimètre de Laurent (1878) a été particulièrement employé à cette fin (fig. 2.10). Cet appareil a été décrit dans de nombreuses publications de l'époque, dont certaines sont d'ores et déjà numérisées et accessibles sur le Web (au CNAM et à la BNF, par ex.)

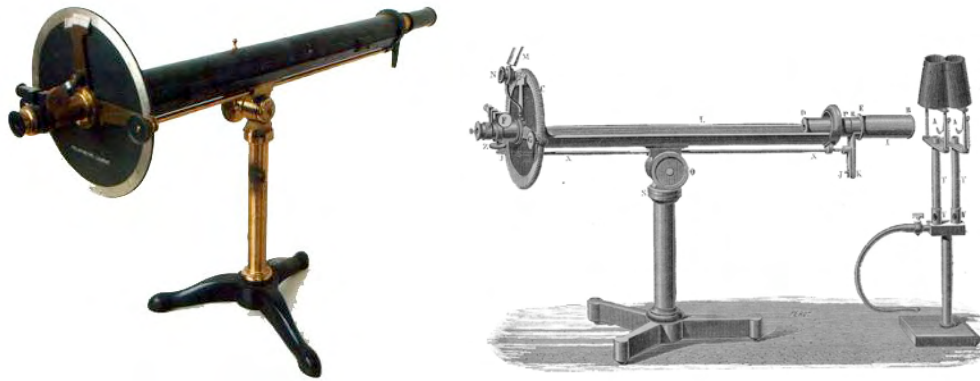


Figure 2.10. Polarimètre Laurent et sa reproduction in D. Sidersky *Polarisation et saccharimétrie*. Gauthier-Villars, 1900, p. 61 (coll. part.)

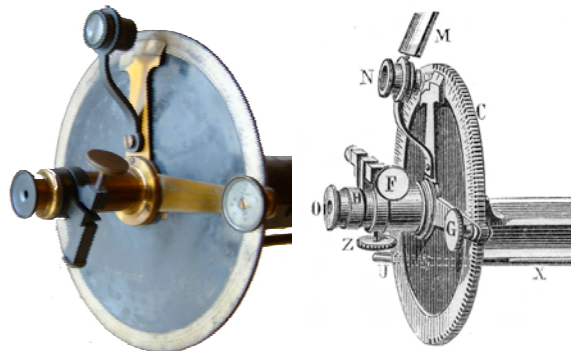


Figure 2.11. Détail du cadran gradué et du porte-oculaire

La confrontation entre la gravure et l'objet réel est intéressante (fig. 2.11). D'abord, leur similarité rassure. Un examen de détail révèle ensuite des différences. Le capot du polarimètre est absent de la gravure, sans doute pour exposer les parties plus essentielles qu'il cache. On voit aussi que l'exemplaire conservé est incomplet : il manque le miroir M. Le montage du bouton de réglage F diffère. Une rotation de 120° de celle-ci, possible avec les trois vis de fixation, permettrait de reproduire l'installation décrite sur l'illustration. L'objet a-t-il été mal (re)monté ? S'agit-il de versions différentes de l'appareil, produit pendant plusieurs décennies, ou bien s'agit-il d'un modeste exemple d'adaptation par l'utilisateur, pour une meilleure ergonomie au quotidien, par exemple ? Il n'est pas dans notre propos de répondre à ce type de question, mais plutôt de s'interroger sur les outils informatiques à fournir pour assister ces

questionnements lorsqu'ils s'effectuent uniquement à l'aide de documents numérisés.

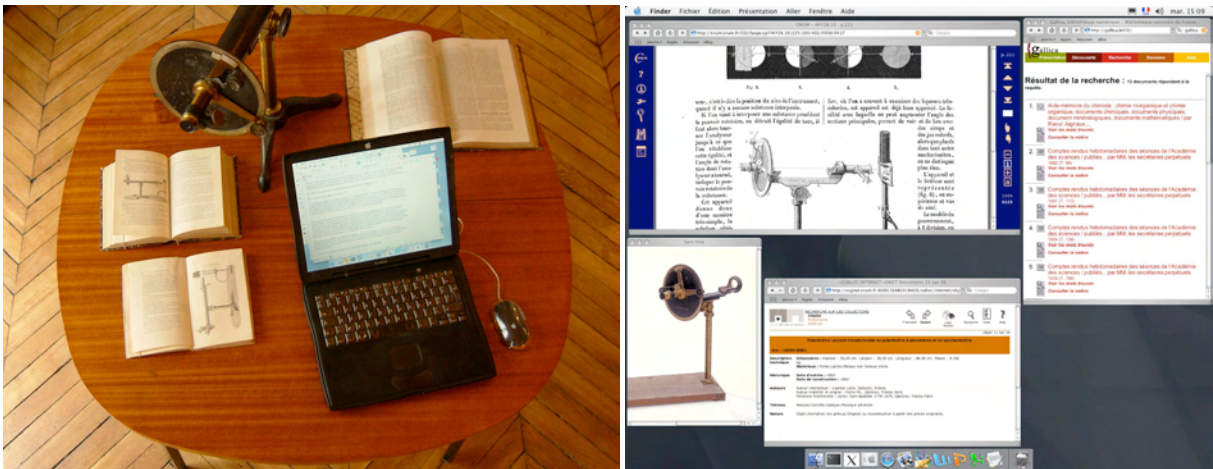


Figure 2.12. Gauche : Une table de travail avec les documents originaux. Droite : Une session de travail avec les documents numérisés sur le Web (sites Gallica/BNF, Conservatoire numérique/CNAM et musée des Arts et Métiers).

La fig. 2.12 oppose une session de travail effectuée avec les documents et objets originaux, avec celle que peut effectuer aujourd'hui un lecteur du Web. Tout espace de lecture nécessite une organisation, une ergonomie voire une esthétique adaptées pour favoriser un travail efficace. Plus généralement, un espace de travail tel qu'un bureau, réel ou métaphore informatique, est organisé en fonction des aptitudes et préférences de son utilisateur. Ainsi, chacun étale à sa guise, selon ses propres critères, les différents documents sur lesquels il travaille. La mémoire spatiale du lecteur et son sens de l'organisation sont mis à profit et laissés libres sur une surface réduite limitée à la taille de son bureau. Les bibliothèques publiques tentent également de recréer un tel environnement favorable en offrant de larges tables et une lumière adaptée. La métaphore du bureau des interfaces WIMP reprend ce principe et permet également à l'utilisateur d'organiser les différentes fenêtres. Les documents peuvent y être regroupés, dimensionnés très librement. Cependant, la surface de ces bureaux virtuels est très réduite et ne permet pas d'atteindre une productivité comparable à un bureau classique. Les fenêtres recouvrantes habillant les documents conduisent inévitablement et rapidement à un enchevêtrement de données qui s'oppose clairement à une bonne organisation.

La première fonctionnalité importante concernant les objets 3D est de pouvoir, au même titre que les livres, les déplacer dans l'environnement de travail. La deuxième est de permettre leur étude sous tous les angles. Ce mode de travail, rencontré en particulier dans les logiciels de modelage, reprend la métaphore de la

pomme que l'on peut regarder sous tous les angles en orientant la main. Afin de pouvoir étudier un objet, on doit pouvoir le déplacer librement. Le contraindre à rester sur le sol gênerait cette tâche. En effet, lors de la manipulation de l'objet, sa géométrie associée à la contrainte de gravité pourrait nécessiter des mouvements compliqués pour arriver au résultat souhaité. D'autre part, la consistance et l'organisation de l'environnement seraient mises à mal si tous les éléments de l'interface ne subissaient pas les mêmes lois.

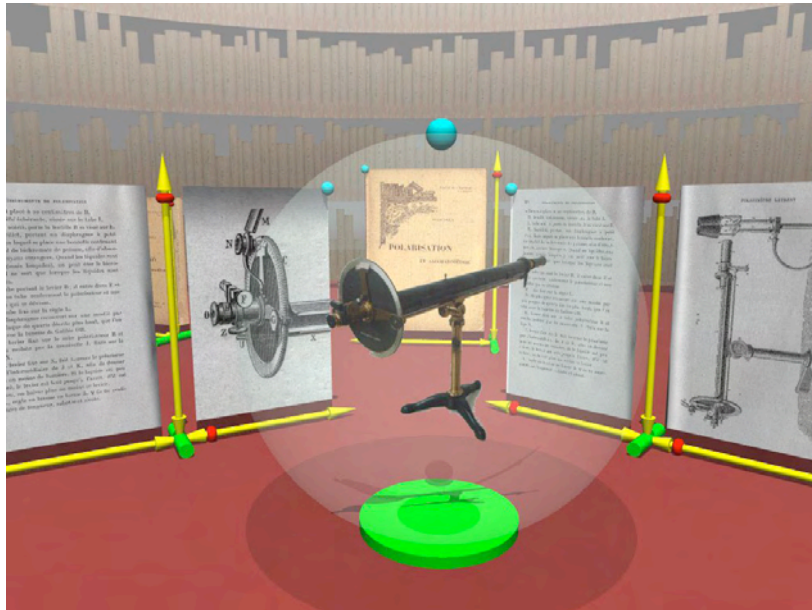


Figure 2.13. Simulation d'une session de travail dans l'environnement 3D

Pour contourner le problème, nous pouvons attribuer des fonctions différentes aux éléments du trépied. Celui-ci servirait toujours de widget de manipulation pour l'objet 3D qui serait en lévitation entre ses axes. Sachant qu'un objet 3D n'est pas feuilletable et que zoomer sur sa structure n'est pas nécessaire, nous pouvons utiliser ces deux axes pour d'autres types d'interaction. Par exemple, chaque axe prendrait en charge chacune des deux rotations selon les axes X et Z non utiles pour les livres. Cependant, là encore, une telle solution serait inconsistante puisque des fonctionnalités différentes seraient attribuées à des mêmes éléments.

Une autre solution est de séparer les actions de positionnement et de manipulation d'un objet 3D. Pour cela, une simple structure englobante sphérique (et non un trépied mal adapté aux objets 3D) non sensible aux rotations, associée à la même poignée de positionnement (un cylindre à sa base), conviendrait. Les déplacements de souris sur la structure englobante permettrait de l'orienter et d'orienter par la même occasion l'objet qu'elle renferme. L'ensemble de ces deux

actions laissent ainsi toutes les possibilités pour étudier et positionner les objets 3D. Dans son micro-environnement limité à sa sphère englobante, l'objet 3D ne subit la gravité, on peut donc le laisser reposer dans la position souhaitée qui répond à nos besoins d'étude. La sphère et les artefacts de positionnement subissent la gravité et les lois de l'environnement comme les trépieds.

3. Le livre comme volume

Numérisation 3D par photogrammétrie

[Haas, 2003 ; Cubaud et Haas, 2004]

Pour les ouvrages dont on peut sacrifier un exemplaire, la numérisation est pratiquée à une échelle quasi industrielle avec des machines de grande capacité, après massicotage et placement à plat dans des chargeurs automatiques. Sans massicotage apparaît un problème de déformation dû à la perspective. En effet, la surface à numériser est rarement plane, ne serait-ce qu'à cause de la présence de la reliure (fig. 3.1, haut). C'est pourquoi les ouvrages sont en général placés sous une vitre qui force la planéité. On conçoit cependant que cette technique ne soit pas appropriée pour des ouvrages précieux qu'il est hors de question de manipuler de la sorte. Par ailleurs, la nécessité de plaquer le livre contre la vitre à chaque page ralentit considérablement le processus et le rend difficilement automatisable. On utilise alors d'autres machines, qui numérisent les ouvrages tels quels, sans vitre. Certaines sont depuis peu équipées d'un tourne-page automatique (4DigitalBooks en Suisse, Kirtas aux Etats-Unis). Les déformations des images sont alors traitées par logiciel, selon diverses heuristiques : rectification des bords [Brown et al., 2003] , des lignes de texte (logiciel I2S Book Restorer), inclinométrie [Wada et al., 1995 ; Zang et al. 2004]. Bien que tout à fait acceptables pour des ouvrages reliés en bon état, aucune de ces techniques ne peut s'appliquer à corriger les artefacts apparaissant* sur le fac-similé reproduit fig. 3.1 (bas).

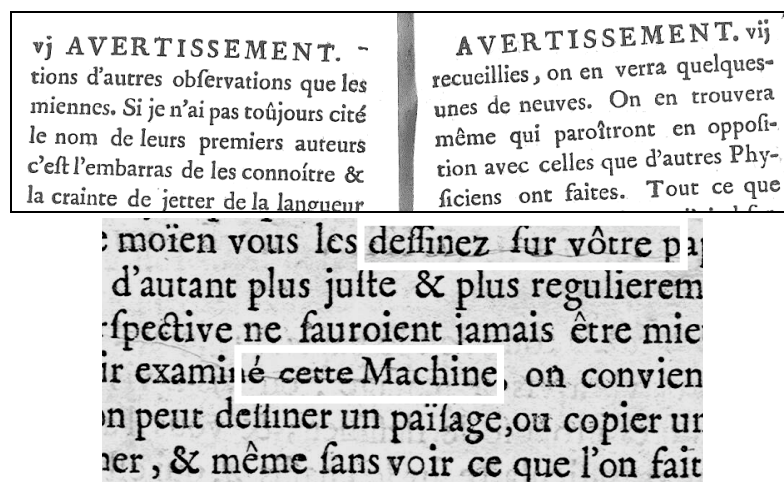


Figure 3.1. Pliures de papier et ombres résultantes. (A. Ramelli. *Le diverse et artificiose machine* (...). Parigi : in casa del'autore, 1588. Accessible en ligne au CNAM <http://cnum.cnam.fr/fSYN/fDY3.html>)

* L'impact de ces artefacts sur le confort de lecture du fac-similé n'a fait l'objet à ma connaissance d'aucune étude.

Il semble dès lors naturel de chercher à en capturer la géométrie par une numérisation 3D. Celle-ci pourrait également trouver des débouchés dans le commerce électronique de livres rares mais plus conventionnels, actuellement en plein essor. Enfin, elle trouve une application évidente pour la numérisation de documents très abîmés. Nous reproduisons comme autre exemple fig. 3.2 un volume extrait du fonds ancien du CNAM, particulièrement déformé par l'humidité.



Figure 3.2. Exemple très abîmé par l'humidité : Privat, Deschanel *Précis de physique*, Paris, 1855, in-8°.CNAM 8°C44

Malgré le recours de plus en plus systématique à la numérisation 3D dans le domaine culturel et le développement d'un marché spécifique, il n'existe pas à l'heure actuelle de matériel dédié à la numérisation 3D de documents. Deux techniques ont fait cependant l'objet d'études :

LUMIERE STRUCTUREE

Un motif lumineux connu, par exemple une grille, est projeté sur la surface, et une prise de vue effectuée sous un angle différent. La déformation du motif sur l'image permet de déduire la surface. La précision obtenue est liée à la finesse du motif. Les différentes lignes de la grille doivent pouvoir être identifiées sans ambiguïté. On peut par exemple leur donner des couleurs différentes. C'est le principe de l'appareil 3D1500 de Minolta, qui a été utilisé dans le projet *Digital Atheneum* [Brown et Seale, 2000] pour des manuscrits du Moyen-âge très gondolés. Il a également fait l'objet d'un test au CEDRIC en juin 2000 [Dequier, 2000]. Un résultat de ce test appliqué à une reliure in-12 assez typique du XVIIIe français (Nollet, *Leçons de physique*) est montré fig. 3.3. Les ondulations du papier (de chiffon à l'époque, donc épais), ainsi que la gorge de reliure, sont bien rendues. On notera cependant les artefacts d'arrondi des angles et les ondulations au niveau des tranches dues à l'interférence entre les bandes et la liasse des feuilles. Le 3D1500 ne semble plus commercialisé.

Il existe également des appareils « tout en un », appelés « scanner 3D », constitués d'un laser à balayage associé à un appareil photographique numérique. Il s'agit d'un cas particulier de lumière structurée, où la ligne projetée est générée par le balayage d'un rayon laser (dans un plan perpendiculaire à celui de la figure), elle-même balayant l'objet (dans le plan de la figure). Nous avons pu tester le scanner Minolta VIVID 700, qui fonctionne selon ce principe, sur un exemple de livre ancien du même type que le précédent (Lalande, Almanach de physique, in-12 broché). La texture des caractères imprimés est interprétée comme un léger relief (fig. 3.4) : l'appareil ne fonctionne en effet qu'avec une radiométrie de l'objet uniforme. Cependant, la précision de restitution est très bonne dans les autres zones : inférieure au 1/10 mm. Un dispositif plus économique a été décrit dans [Brown et Seales, 2001]. Il consiste en un balayage du document par une ligne lumineuse, les images étant successivement photographiées. La précision semble comparable à celle obtenue par laser.

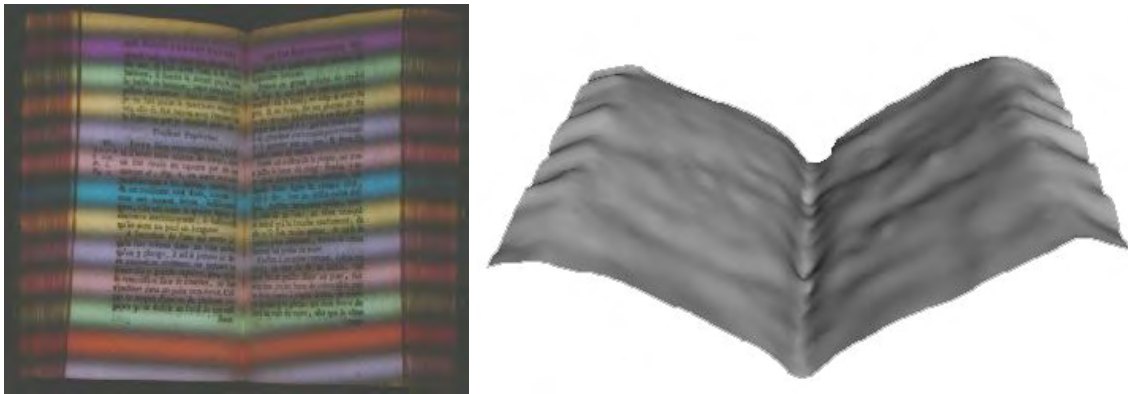


Figure 3.3. Test du 3D1500 de Minolta: projection de bandes de couleurs (gauche) surface obtenue (droite).

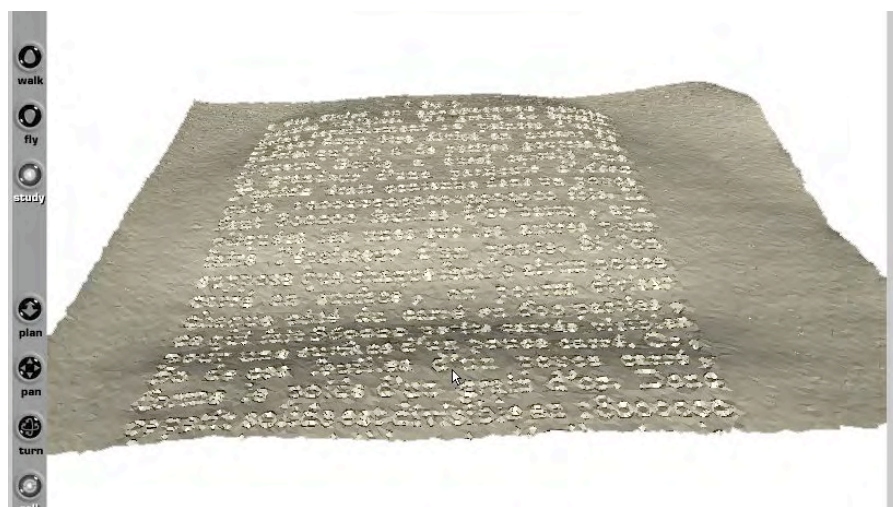


Figure 3.4. Surface obtenue avec le VIVID 700 (20 000 triangles).

PHOTOGRAMMETRIE

La photogrammétrie permet la reconstruction de la surface à partir d'un couple de photographies stéréoscopiques [Egels et Kasser, 2001]. Chaque couple de points identifié par corrélation entre les deux images donne les coordonnées du point objet (3D) correspondant. Par rapport aux techniques précédentes, la photogrammétrie offre plusieurs avantages : couplage entre acquisition de la texture et acquisition de la surface, séparation totale entre la phase d'acquisition (rapide) et la phase de calcul (lente), rusticité de l'installation : un matériel photographique standard, orienté approximativement et non étalonné donne déjà des résultats. Si l'on procède à un étalonnage de l'équipement, il est possible d'obtenir directement des informations métriques absolues sur la surface du document (il devient ainsi possible d'exploiter un plan côté, par ex.). En revanche, la photogrammétrie présente deux inconvénients : la nécessité de la présence d'une texture sur le document, qui exclut son utilisation pour des zones uniformes (marges, interlignes, etc.) et le risque de fausse corrélation, qui reste le problème majeur de la photogrammétrie si on l'envisage sans aucun contrôle humain.

La prise de vue de document a quelques caractéristiques qui la distinguent des applications plus courantes (photogrammétrie aérienne ou architecturale par exemple). Certains aspects facilitent le processus de reconstruction 3D : l'orientation externe des prises de vue est fixe et connue ; l'objet est relativement plat, sans discontinuités ni occlusions dans le cas général, ce qui facilite la corrélation entre les images. Cependant, la mise au point est rapprochée et variable, ce qui rend difficile un étalonnage unique préalable. La forte convergence entraîne un problème de profondeur de champ. Les livres présentent une texture avec une grande répétition de motifs (les lettres dans du texte, traits de gravure dans les illustrations), ce qui présente un danger de fausses corrélations. Nous décrivons dans [Cubaud et Haas, 2004] l'organisation d'une chaîne photogrammétrique qui tient compte de ces spécificités. Le détail des développements informatiques et des essais se trouvent dans [Haas, 2003].

Dans un logiciel de photogrammétrie, la partie délicate est l'identification des couples de points homologues sur les deux vues. Pour chaque point d'intérêt de la vue gauche par exemple, on cherche à identifier son homologue sur la vue droite par corrélation automatique. La description d'une surface nécessitant des dizaines de milliers de points, le temps de calcul peut devenir rédhibitoire, et ceci d'autant plus qu'on cherche une fiabilité (absence de fausses corrélations) élevée. Nous utilisons dans ce logiciel une approche itérative destinée à minimiser ce temps de calcul. En effet, la surface d'une page de livre a des propriétés de continuité qu'il

convient d'exploiter : la hauteur d'un point n'est jamais très éloignée de celle de ses voisins.

Un premier semis de points (coins de Harris) est sélectionné sur l'image de gauche. Le calcul de leurs homologues à droite n'est pas très long puisqu'ils sont peu nombreux (environ 60 points). Une approximation de la surface est ensuite obtenue par triangulation des points-objets alors calculés. Un second semis de points, quatre fois plus dense, est sélectionné à gauche. Pour chacun d'eux, la surface approximative obtenue précédemment permet de prédire la plage de hauteurs possibles du point objet correspondant. La recherche de l'homologue à droite se fait alors sur un segment « prédicteur », ce qui réduit le temps de calcul de corrélation pour chacun des points. Une surface plus proche de la réalité est obtenue par triangulation entre ces nouveaux points. On recommence ensuite le processus : à chaque boucle, le nombre de points à corrélérer est plus grand, mais la surface triangulée étant plus précise, l'approximation de la hauteur des points-objets est meilleure, et les segments de recherche sont plus courts. L'algorithme général du logiciel est le suivant :

```
Chargement image et paramètres d'orientation approximatifs
répéter
  Production des points d'intérêt à gauche (coins de Harris)
  Sélection d'un jeu de points d'intérêt
  Recherche des points homologues à droite
  Affinage de l'orientation et calcul des points objets
  Ajout de points à la triangulation
jusqu'à nombre de points suffisants
Production du fichier VRML, de la texture et de l'ortho-image
```

Sa mise en œuvre représente 5600 lignes de code Java, sous Windows. Il est complété par une feuille de calcul Excel où sont entrés les paramètres géométriques du banc de prise de vue et les paramètres extrinsèques des caméras, dont on déduit la résolution à attendre du modèle 3D.

Un banc expérimental a été monté sur un bâti de prises de vues professionnel. Sur le socle, deux appareils photographiques grand public (Konika Digital Revio KD500Z et Canon PowerShot S45, 4Mpix) ont été fixés à l'aide de deux « bras magiques » et visent le document à 30° par rapport à la verticale. Le banc est équipé de quatre projecteurs au tungstène, deux de chaque côté, orientés à 45°. La résolution attendue est de l'ordre de 0,1 mm. Nous avons fait l'impasse sur l'étalonnage des appareils. Celui-ci aurait nécessité une étude à part entière, en particulier la réalisation d'un objet étalon aux dimensions parfaitement connues. Cet aspect n'a pas été jugé prioritaire, dans la mesure où l'on peut faire de la restitution 3D non métrique (c'est-à-dire à un facteur d'échelle près) en se passant

des paramètres d'orientation interne (focale et coordonnées du point principal d'autocollimation). Dans les faits, une échelle approximative est déterminée par l'entrée manuelle dans le logiciel de la longueur de la base stéréoscopique, mesurée au double-décimètre. En ce qui concerne la distorsion, sa non-prise en compte (même si celle-ci est prévue dans le logiciel) conduit à rechercher chaque point homologue non plus strictement sur un segment, mais sur une plage autour de ce segment, ainsi qu'à une légère erreur dans le positionnement des points-objets.

Pour l'ouvrage de test de la fig. 1, nous obtenons par exemple (fig. 3.5) une surface de 30K triangles en 5 min. env. (pentium 4, 2 GHz). Les fausses corrélations représentent 0.6% du total des sommets (un exemple est donné fig. 3.6). Celles-ci sont cependant très gênantes, car elles provoquent des aberrations de surface inacceptables (fig. 3.7). Le temps de calcul devrait pouvoir être réduit par parallélisation des traitements afin d'envisager des applications industrielles.

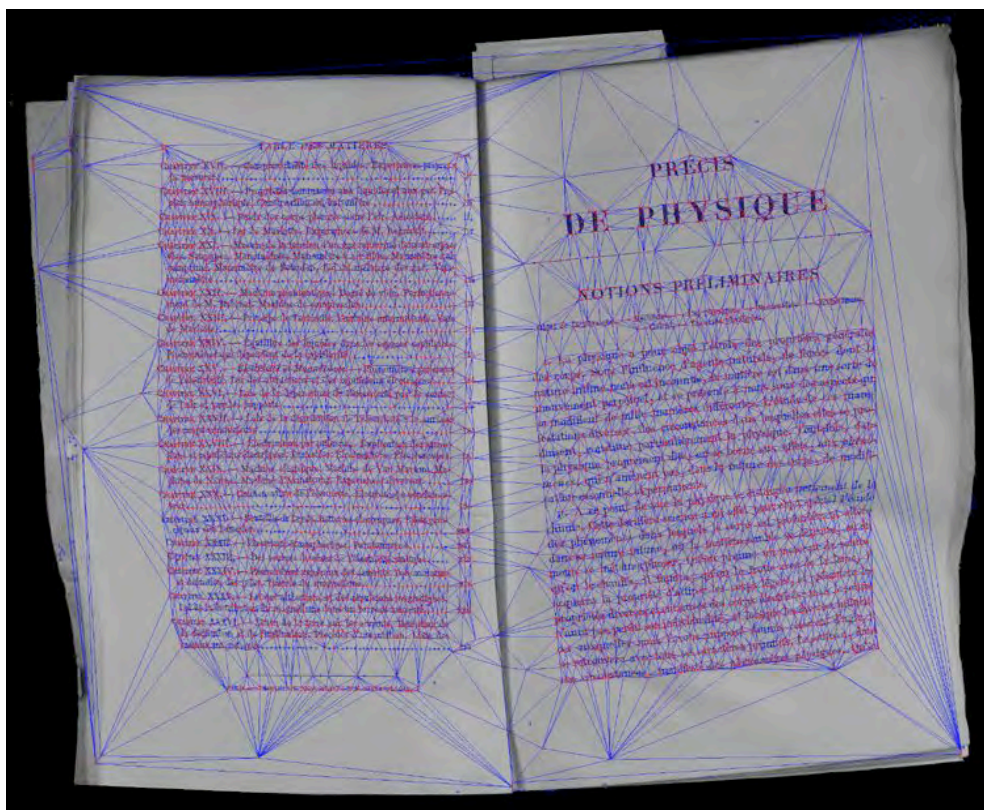


Figure 3.5. Vue gauche après triangulation (27 075 triangles).

Exemple de fausse
corrélation (point 230)
sur
un motif répétitif parallèle à
la base.

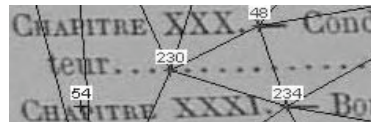


Image gauche

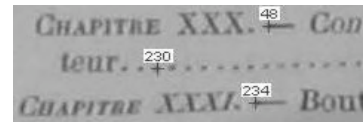


Image droite

Figure 3.6. Motif répétitif horizontal (points et tirets dans une page de table de matière).

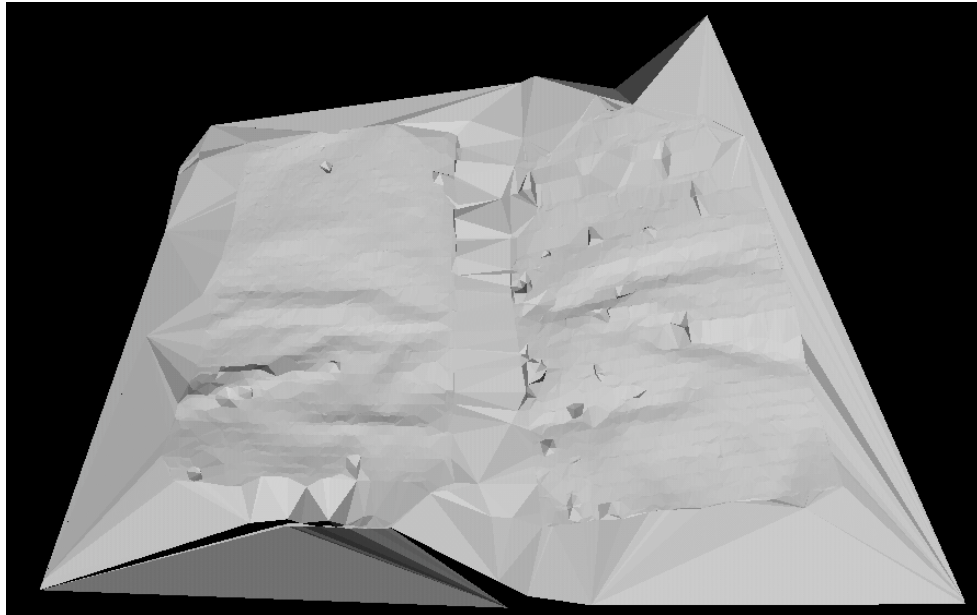


Figure 3.7. Surface 3D traduite en VRML, sans plaqué de texture pour faire ressortir les fausses corrélations.

Le livre-objet : exemple des livres à systèmes

[Cubaud, Dupire et Topol, 2005]

Dès lors que l'objet livre devient lui-même centre d'intérêt, la mise à plat devient de toute manière insuffisante. On peut citer ainsi les livres d'artistes, avec souvent de nombreux collages superposés, les herbiers, et bien sûr les livres « à systèmes », c'est à dire avec parties mobiles. Les livres à systèmes ont une longue et fascinante histoire, qui remonte aux manuscrits du Moyen-âge. Ils ont d'abord été conçus à des fins didactiques, ou pour réaliser des outils d'aide au calcul. Depuis le XIXème siècle, les livres à systèmes sont principalement réalisés pour les enfants et utilisent une très grande variété de techniques d'animation : le tableau ci-dessous en récapitule les principaux. A partir des années 1960, les procédés s'industrialisent et une écriture spécifique à ce type de média apparaît [Haining, 1979].

Dépliant	Partie d'une page, repliée sur celle-ci, permettant, lorsqu'elle est dépliée, d'étendre la surface de la page en dévoilant du texte et/ou des images.
Volet	Pièce de papier fixée à la page en un point et qui dévoile, lorsqu'elle est soulevée, le texte et/ou les images cachés dessous.
Tirette	Languette de papier qui, lorsqu'elle est tirée/poussée, entraîne l'animation d'autres pièces de papier de la page.
Roue	Disque de papier, souvent inséré dans l'espace de deux pages consécutives collées, actionné directement par une ouverture sur le bord des pages, permettant de modifier le contenu d'une fenêtre découpée sur la page courante.
Métamorphose	Système fonctionnant sur le même principe que les stores vénitiens. En déplaçant une languette de papier, on fait apparaître 2 images différentes.
Volvelle	Disques de papiers, concentriques, pouvant être manipulés indépendamment les uns des autres, permettant d'obtenir une information ou une image différente en fonction de leurs positions relatives.
Harlequinade	Illustrations repliées les unes sur les autres, permettant de modifier la scène de départ lorsqu'elles sont dépliées, faisant ainsi avancer l'histoire.
Pop-up	Livre en relief, dont les scènes ou éléments se déploient lors de l'ouverture des pages.
Carrousel	Livre s'ouvrant à 360°, et présentant une histoire en relief à la manière d'un manège.
Peep-show (tunnel book)	Livre se déployant en profondeur, permettant de voir une scène en perspective, grâce aux différents plans illustrés qui le composent, donnant cette impression de « tunnel ».

De tels ouvrages procurent une expérience de lecture allant beaucoup plus loin que le simple parcours du texte et des illustrations. Nous prendrons comme exemple un texte de l'ingénieur de la Renaissance Salomon de Caus, consacré à la gnomonique. La page reproduite fig. 3.8 inclut un modèle en papier de gnomon. Il est composé de deux plats posés initialement l'un sur l'autre. Le plus petit plat, qui représente le stylet du gnomon, doit être levé en premier. Le grand plat se lève ensuite en glissant le stylet dans une échancrure située sur le bord du grand plat. Une fois le modèle construit, le lecteur doit placer le livre au soleil, l'orienter dans la direction Est-Ouest et, grâce à l'ombre du stylet, lire l'heure solaire sur les graduations latérales.



Figure 3.9. Salomon de Caus. *La pratique et démonstration des horloges solaires*. Paris, Drouart, 1624. (Bib. du CNAM)

La numérisation de tels textes serait extrêmement utile, car leur manipulation répétée cause des dommages irréversibles. On conçoit toutefois qu'elle soit délicate, dans la mesure où elle impose souvent un démontage des parties mobiles du livre. Une numérisation 3D par le dessus, du type de celle décrite dans la section précédente, ne suffit en effet pas pour des montages de papiers perpendiculaires au plan du livre, ni pour résoudre les occlusions de parties cachées. L'exemple évoqué ici est toutefois suffisamment simple pour se prêter à des prises de vues successives des différentes parties du système.

Une fois photographiées les différentes parties du système, il est possible d'en décrire l'organisation sous la forme d'un arbre [Cubaud et al., 2005]. Chaque partie mobile est nommée et typée (volet, tirette, disque mobile, etc.) et possède une vue avant et arrière. A chaque type sont associés des attributs géométriques spécifiques (position et orientation de l'axe de rotation, direction de rotation, position initiale). Un attribut *Parent* permet de situer la partie mobile dans la hiérarchie du système, la racine de l'arbre étant la page du livre sur laquelle le système est posé. Un certain nombre d'actions sont associées à la partie mobile, qui sont effectuées par le lecteur et qui peuvent affecter le positionnement d'autres parties mobiles.

Deux options sont envisageables pour la reconstruction : nous pouvons, d'une part, choisir de reconstruire virtuellement toutes les parties du système (même celles qui sont invisibles, souvent cachées dans l'épaisseur de la page) afin d'obtenir une copie stricte de l'original : tous les mécanismes, leurs articulations et leurs effets seront préservés. Leur intégration se fera grâce à l'utilisation d'un moteur gérant la physique de l'animation 3D (*Open Dynamic Engine*, par

exemple^{*}) dans l'environnement de consultation. Celui-ci déterminera les positions de chaque partie du système, en temps réel, en fonction des actions de l'utilisateur.

Nous pouvons, d'autre part, décider que ces mêmes mécanismes sont d'un intérêt négligeable et que leur numérisation n'apporterait rien de fondamental. Seules les parties terminales (actionneur et actionné) seront donc prises en considération. Cette seconde option permet de s'abstenir de la reconstruction des structures cachées de la chaîne cinématique. D'un point de vue fonctionnel, la perte d'information (liens entre les parties du système) devra être compensée. Ceci pourra être réalisé par la création d'animations, mettant en scène les mouvements simultanés des systèmes liés entre eux. Idéalement, ces séquences permettront de visualiser toutes les configurations relatives du système, en partant de la position initiale jusqu'à la position finale des parties mobiles. A chaque action de l'utilisateur pour modifier la configuration d'un système donné, le fichier de l'animation sera joué, en avant ou en arrière à partir de la position courante, afin de faire correspondre à l'action effectuée (interaction utilisateur) les mouvements des différentes parties mobiles.

Pour ce faire, une description de la structure de la page reste indispensable pour stocker les informations d'interdépendance entre les parties et des fonctionnalités associées. Nous avons choisi de spécifier ces informations au travers d'une hiérarchie de classes (fig. 3.9). Chaque type de système est représenté par une classe dérivée introduisant une sémantique spécifique. Ces différentes classes comportent des méthodes permettant d'opérer les actions spécifiques au type de système représenté. Chaque système d'une page est une instance d'une de ces classes. Tout système ainsi modélisé dérive d'une classe introduisant les transformations géométriques les décrivant. A ce niveau d'abstraction, on peut distinguer deux types de systèmes mettant en œuvre respectivement :

- une translation caractérisée par une contrainte de longueur maximale possible de la translation et par l'état courant de la translation,
- une rotation caractérisée par un angle maximum et par l'angle courant.

Une classe de base commune à tous les systèmes rassemble les informations géométriques et hiérarchiques :

- la position du système dans le repère local du système père auquel il est rattaché,
- l'axe de rotation ou de translation du système,
- les informations pour savoir si un système peut être manipulé ou pas (*isMovable*) et s'il se trouve dans sa position finale ou pas (*isActive*),

* <http://ode.org>

- les références des systèmes attachés (les systèmes fils).

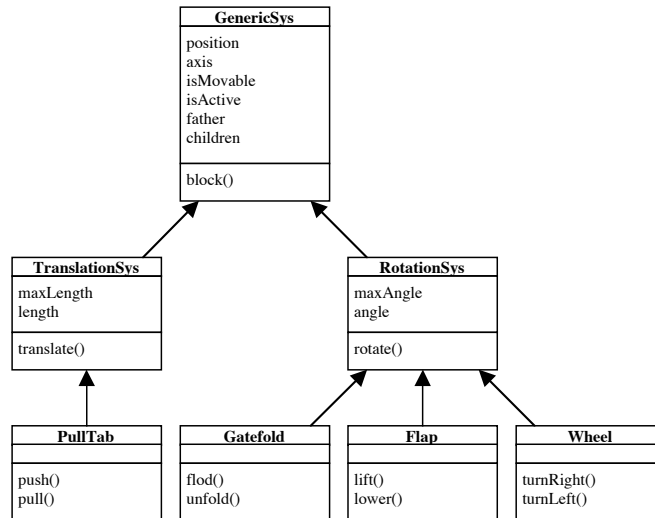


Figure 3.9. Diagramme UML des différentes classes de base.

Cette classe de base permet de construire un graphe de dépendance des différents systèmes. Tout système fils ne pourra être activé que si le système père est activé. De même, tout système ne pourra être activé que si ses frères (les systèmes se trouvant dans la liste des fils de leur système père commun) sont désactivés. A chaque page d'un livre à système sera également associé une liste de systèmes présents sur cette page. Les méthodes de cette classe de base permettent d'effectuer les actions réalisables sur un type particulier de système. Des appellations communes sont envisageables pour chacune des classes dérivées pour mettre, via des méthodes virtuelles, une gestion polymorphe de la classe de base.

Par ces mécanismes simples, il est possible de modéliser le gnomon de Salomon de Caus. Le graphe de dépendance est le suivant : un premier volet *flap1* est rattaché à la *page* et le deuxième volet *flap2* est rattaché à *flap1*. Cette manière de connecter les trois éléments (*page*, *flap1* et *flap2*) traduit simplement, d'après notre mécanisme décrit précédemment, les contraintes suivantes :

- *flap1* ne peut pas être activé (relevé) tant que la page n'est pas active (ouverte),

- *flap2* ne peut pas être actionné tant que *flap1* n'est pas en place (formant un angle de 90° avec la page – attribut *activate* à *true*).

Dans le cas particulier où *flap1* est en position relevée et *flap2* abaissé, nous aurons les valeurs d'attributs suivantes : *flap1.activate* = vrai, *flap1.movable* = vrai, *flap2.activate* = faux, *flap2.movable* = vrai. Il sera donc possible soit d'abaisser *flap1*, soit de lever *flap2*. Si l'on choisit de relever le second volet, alors *flap1.movable* devient faux et seul ce volet reste manipulable.

Une fois ces informations acquises, une reconstruction en image de synthèse du système devrait pouvoir être obtenue automatiquement, et nous souhaitons pouvoir poursuivre nos travaux en ce sens. La fig. 3.10 reproduit quelques copies d'écrans obtenues avec le logiciel 3DStudioMax par modelage manuel de la scène 3D. Cette première expérience est encourageante dans la mesure où le système est correctement reproduit et très lisible.

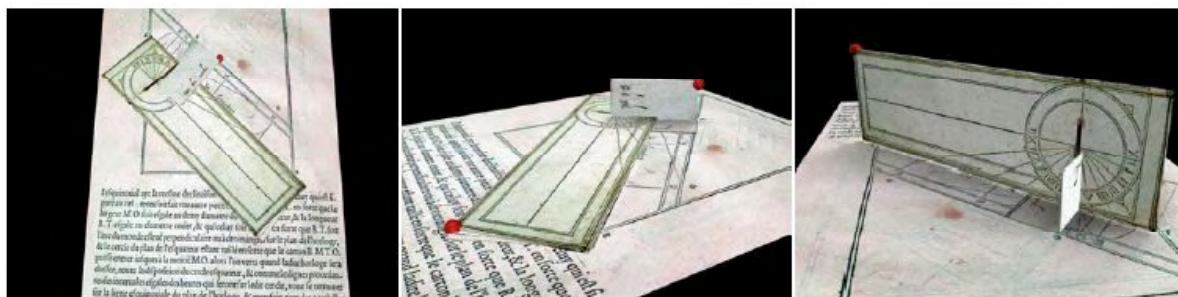


Figure 3.10. Différentes vues d'une reconstruction synthétique de la page à système

Prolongements

PERFORMANCE DE LA NUMERISATION 3D

La robustesse de la numérisation peut être considérablement améliorée par un certain nombre de techniques, qu'il s'agisse d'éviter les fausses corrélations ou bien de les supprimer à posteriori. Ceci constitue un sujet en soi pour une étape ultérieure de ce travail. Parmi ces techniques, citons l'introduction :

- d'une troisième image,
- d'une corrélation croisée : l'homologue de l'homologue d'un point doit être ce point lui-même [Ulges et al., 2004].
- d'une fenêtre de corrélation adaptative, c'est-à-dire tenant compte de la surface approximative locale déjà calculée,
- de contraintes sur les vecteurs normaux de la triangulation, afin de tenir compte des éventuelles propriétés de continuité et de dérivabilité de la surface.

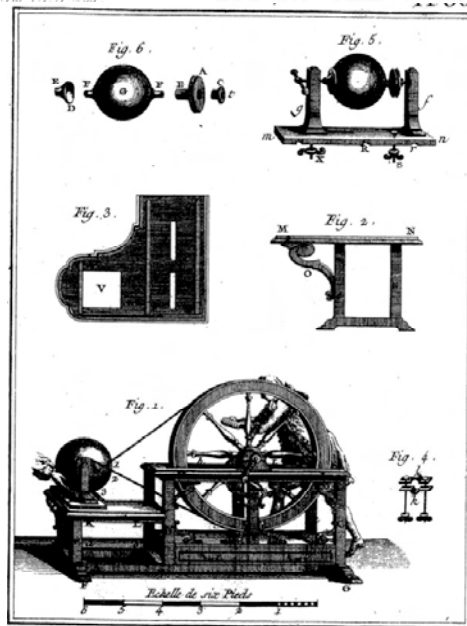
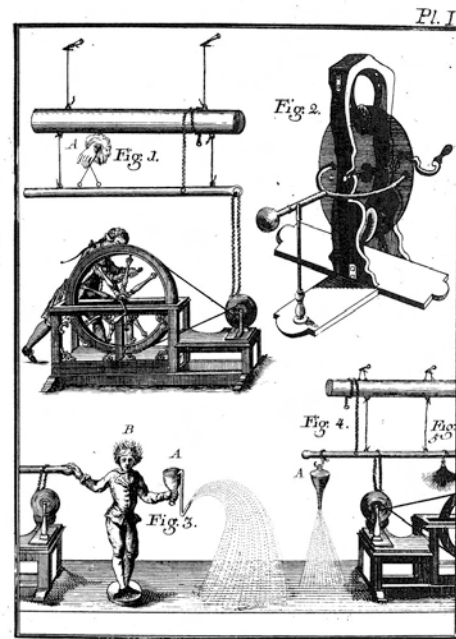
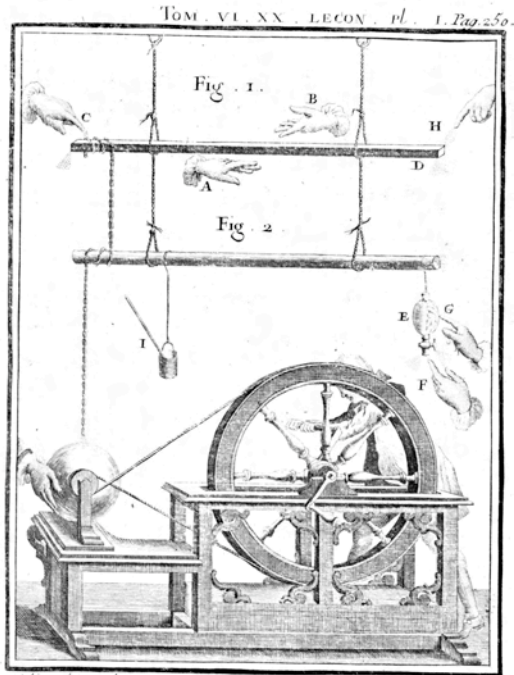
INTERFACE DE VISUALISATION ET DE MANIPULATION

A partir de la connaissance de la surface 3D du document numérisé, on peut tenter de le déplier ou de le mettre à plat. C'est utile pour les manuscrits reliés (registres paroissiaux, ou manuscrits d'auteurs réassemblés), pour lesquels les heuristiques de rectification de lignes de textes sont inopérantes. C'est

historiquement la première application envisagée pour la numérisation 3D [Brown et Seales, 2000 ; Pilu, 2001 ; Haas, 2003]. Les progrès effectués en matière de visualisation 3D temps-réel permettent toutefois de s'affranchir de ce procédé et d'offrir à l'utilisateur une interface de consultation purement 3D. L'utilisation d'un navigateur 3D généraliste (au format VRML, Metastream ou QTVR) s'avère cependant vite fastidieuse pour la manipulation du document. En effet, seules les actions de zoom et la rotation autour de l'axe défini par la gorge de reliure sont vraiment utiles. Pour la consultation de l'ensemble d'un même texte, on peut également vouloir offrir une possibilité de défilement des pages. L'objet 3D pourrait ensuite être plongé au sein d'une interface 3D de bibliothèque numérique plus générale. Surtout, cette interface devrait permettre à l'utilisateur de mesurer les dimensions de parties du fac-similé pour tirer pleinement partie du procédé photogrammétrique.

RECHERCHE DE SIMILARITES DANS DES CORPUS D'IMAGES

Les techniques de recherche de similarité s'appuient sur des algorithmes assez proches de ceux développés pour la reconstruction 3D (mise en correspondance de points d'intérêt). Une application potentielle serait la recherche d'extraits semblables dans les fac-similés d'ouvrages scientifiques pour étudier comment les planches gravées sont copiées d'un ouvrage à l'autre. Ce phénomène est en effet très fréquent jusque vers 1850 (fig. 3.11). Ce type d'application est à ma connaissance nouveau (il existe cependant plusieurs travaux sur la recherche de similarité d'éléments typographiques complexes, comme les lettrines [Debora]). Il n'est d'ailleurs pas garanti que les méthodes actuelles soient adaptées à ces images (gravure au trait, pas de tons continus) car elles conduisent à une surabondance des points d'intérêt (fig. 3.11 en bas à droite).



Physique.



Figure 3.11. De gauche à droite : Haut : NOLLET. *Leçons de physique expérimentale*, t.6. 8e ed. Durand, 1780 (1ère ed. 1743-48) ; JACQUET DE MALZET. *Précis de l'électricité (...)*. Trattner, 1775 ; Bas : BRISSON. *Dictionnaire raisonné de physique. Seconde édition. Librairie économique*, 1800 (CNUM et coll. part.) • Semis de points d'intérêt sur une image de planche gravée

LIVRES A SYSTEME

La réalisation d'une interface utilisateur spécifique sera la prochaine étape de notre travail. Cette interface devra désigner clairement les parties mobiles pouvant être manipulées par l'utilisateur. Ce processus doit être réalisé de manière à ne pas surcharger la page qui contient elle aussi des informations (textes, images). Par exemple, l'utilisation de widgets 3D inadaptés peut avoir pour conséquence de masquer des informations situées sur la page du livre (en lecture simple). De plus, la nature de ces nouvelles interactions varie selon les systèmes rencontrés (soulever un volet, tirer une languette, etc.). Le lecteur doit pouvoir identifier efficacement quelle(s) action(s) sont réalisables sur le système qu'il considère. Ces interactions peuvent, par ailleurs, être combinées entre elles dans le cas de systèmes hiérarchiques, dans lesquels une partie du système en contraint une autre. Là encore, les choix existent lors de la conception du livre à système virtuel, de laisser ou non la possibilité au lecteur de commander finement les systèmes. En effet, dans un souci de simplification de l'interaction, nous pouvons envisager le déploiement automatique et complet d'un système sur une page. Cette solution permet de réduire l'interaction à une seule commande, et de réduire le nombre d'informations à présenter à l'écran. Cette approche est tout à fait envisageable pour la plupart des systèmes et permet de guider plus efficacement le lecteur dans sa découverte du livre. Inversement, on doit pouvoir proposer aussi au lecteur un mode exploratoire libre, dans l'esprit des navigateurs de scènes 3D (comme les navigateur VRML), afin de pouvoir gérer les cas de déplacements de point de vue plus complexes (fig. 3.12).



Figure 3.12. Système de type tunnel nécessitant un placement de caméra libre (<http://www.populady.com>)

4. Eléments technologiques pour une machine à lire

Bureau 3D

[Topol 2000 ; Topol 2002 ; Dupire, Topol, Cubaud, 2005]

Un environnement mêlant 2D et 3D serait idéal pour mettre en commun les différents outils de consultation de bibliothèques numériques présentés aux chapitres précédents. La figure 4.1 illustre le potentiel d'une telle interface. Dans celle-ci sont représentés non seulement les éléments de la bibliothèque numérique, mais aussi des applications 2D standards pour annoter les livres ou rechercher des informations complémentaires. La collection est schématisée dans cette figure sous la forme d'un arbre conique [Robertson et al., 1993]. Cela illustre un autre aspect que nous voudrions étudier : la souplesse de représentation de la collection. En fonction des préférences de l'utilisateur nous pourrions représenter la collection dans un cylindre, une sphère, dans le lieu réel reconstruit ou encore un arbre conique, etc.

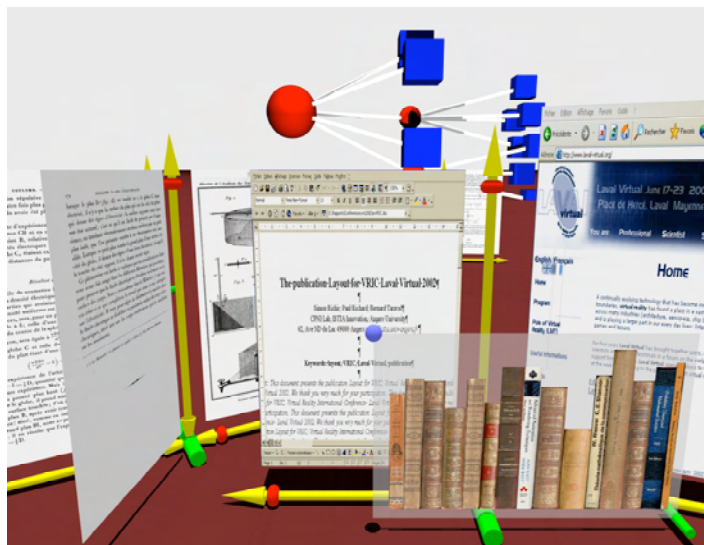


Figure 4.1. Utilisation du trépied comme métaphore d'interaction pour un gestionnaire de fenêtres.

Le remplaçant du bureau 2D actuel n'a pas encore été trouvé. Cependant, nous sommes persuadés qu'il devra intégrer, à la fois, les applications avec une interface 2D et celles avec une interface 3D, qui sont de plus en plus nombreuses. Il n'est pas envisageable d'abandonner les applications 2D dont la plupart n'ont aucune raison d'être transformées en application 3D. De la même manière, les

applications en mode texte ont été intégrées dans toutes les interfaces graphiques 2D grâce à l'utilisation de consoles texte. Un sur-ensemble 3D permettrait donc, non seulement d'englober les applications 2D, mais également les applications en mode texte (elles-mêmes englobées par les interfaces 2D). De plus, la situation actuelle concernant la gestion des applications 3D n'est pas satisfaisante. Les applications 3D sont en effet exécutées dans une fenêtre 2D et forcent l'utilisateur à changer de métaphores de navigation, lorsque qu'il passe d'une application avec une interface 2D à une application en 3D. Ce problème est constamment rencontré lorsque l'on visualise des scènes VRML, par exemple.

Le travail que je décris dans cette section a été principalement réalisé par A. Topol dans le cadre de sa thèse [Topol, 2000 ; Topol, 2002]. Nos expériences sur le bureau 3D ne se sont à ce jour pas poursuivies, essentiellement parce qu'elles relèvent maintenant plus du développement pré-compétitif industriel (projet *Looking Glass* de Sun, ou *Vista* de Microsoft, par ex.) que de la recherche proprement dite. De plus, le portage récent des API 2D industrielles au-dessus des standards graphiques 3D (OpenGL, Direct 3D) pour bénéficier de l'accélération matérielle du rendu 3D, a modifié totalement la technique de programmation mise en œuvre.

TRAVAUX SIMILAIRES

D'autres projets ont également eu pour tâche d'inclure des applications 2D dans un environnement 3D. Le gestionnaire de fenêtre *MaW3D* place les fenêtres dans un « tunnel » [Leach, 1997]. Microsoft a également présenté un tel projet à la conférence CHI'2001, la Task Gallery [Robertson et al., 2000]. La métaphore utilisée est celle de la galerie d'art. La ressemblance avec *MaW3D* est donc assez frappante puisque l'utilisateur est contraint de naviguer dans le couloir. Les fenêtres, dans ces deux tentatives, sont plaquées sur des murs. Nous pensons qu'une interface 3D ne doit pas être aussi contrainte. Ces limitations fortes n'autorisent pas l'inclusion d'objets 3D et rien (ni dans les articles cités, ni dans les copies d'écran) ne nous indique qu'il est possible de le faire dans ces deux environnements. L'apport des rotations est très important pour l'organisation d'un espace 3D, aussi bien pour déplacer le point de vue sur la scène que pour la manipulation des objets. Cependant, cela nécessite d'étudier les politiques de placement et les techniques de manipulation.

La deuxième tentative, plus originale, des chercheurs de Microsoft se propose de présenter les fenêtres sur un plan incliné. Dans la *Data Mountain*, les informations peuvent, en effet, être déplacées et positionnées où que ce soit sur ce

plan incliné, par rapport au point de vue de l'utilisateur. Les seules contraintes imposées sont que les documents reposent sur le plan et qu'ils ne sortent pas de ses limites. Dans cette interface, les fenêtres peuvent être positionnées en fonction de l'intérêt que leur porte l'utilisateur et plus encore, en fonction de sa propre notion d'importance du placement. Un utilisateur pourra placer les documents pertinents au premier plan parce qu'il peut ainsi les voir complètement à tout moment, alors qu'un autre les placera tout en haut du plan pour signifier leur plus grande importance. La mémoire spatiale sera donc complètement stimulée par cette notion personnelle d'importance attribuée aux informations. Cependant, pour ce projet également, l'utilisation d'objets 3D n'a pas été intégrée, en prévision des futures applications véritablement 3D.

Les projets les plus proches du nôtre sont sans conteste *3Dwm* (<http://www.3dwm.org>) et celui de [Roussel, 2001] car ils ont été également développés sous Linux. Cependant, ces gestionnaires de fenêtres ont besoin, pour leur fonctionnement, de l'émulateur d'écran inter-plateformes *Virtual Network Computing* (VNC) d'AT&T, [Richardson, 1998]. *3Dwm* réorganise le *frame buffer* de XWindow sur une seule et unique face texturée avec l'image complète du bureau 2D de XWindow. Contrairement à notre prototype, chaque fenêtre n'est donc pas orientable et/ou positionnable indépendamment. Elles peuvent être déplacées mais doivent rester dans le même plan. Cette solution, malgré une esthétique très travaillée, n'apporte donc rien de plus que la 2D pour la gestion des fenêtres. Elle offre néanmoins la possibilité de manipuler dans un même environnement aussi bien des objets 2D que 3D.

ORGANISATION DU PROTOTYPE

Notre gestionnaire 3D a été développé sur XWindow, car son architecture client/serveur permet de récupérer et de détourner l'usage des structures de données stockées par le serveur X. La même opération sous Windows nécessiterait de modifier le code de l'interface graphique (ce qui a d'ailleurs été fait par Microsoft pour implémenter la *Task Gallery*). Or, nous ne disposons pas de ces codes sources. La première étape fut de récupérer les images des différentes fenêtres et de les inclure sur des faces OpenGL que nous avons ajoutées dans notre visualisateur de fichiers VRML. La récupération des images n'est pas un problème majeur. Nous avons implémenté exactement la même méthode que celle des programmes de capture d'écran sous XWindow (comme *xwd*). Cette méthode est également utilisée par VNC.

Le schéma ci-dessous les différentes opérations effectuées lors de l'initialisation de la scène et lors de l'interaction sur une fenêtre. Pour l'initialisation, l'arbre des fenêtres est récupéré par un appel à la fonction `XQueryTree`. Pour chacune des fenêtres filles de la fenêtre principale (root), nous récupérons ensuite sa taille. La fenêtre est ensuite redimensionnée de telle manière qu'elle ait une taille de $2n * 2m$ pixels, la taille de texture imposée sous OpenGL. L'image est ensuite récupérée et plaquée sur une face dont la taille est proportionnelle à l'image. Le sous-arbre des fenêtres de chacune des fenêtres de premier niveau, est ensuite sauvegardé pour déterminer, durant la phase d'interaction, sur quelle fenêtre l'utilisateur a agi. Il faut en effet savoir que sous XWindow, tous les éléments graphiques (widgets) sont des fenêtres. Lorsque l'utilisateur effectue un clic de souris, on détermine d'abord sur quel objet il a eu lieu. Si c'est une fenêtre, les coordonnées 3D du point cliqué sont calculées, pour pouvoir déterminer les coordonnées 2D de l'endroit cliqué dans la fenêtre, par rapport à son coin supérieur gauche. Le widget sur lequel a eu lieu ce clic de souris est ensuite extrait grâce à la sauvegarde de l'arbre des sous-fenêtres. Les événements "bouton appuyé" et "bouton relâché" sont envoyés à ce widget. Les nouveaux widgets que cette action a pu créer sont stockés dans l'arbre des sous-fenêtres. Les modifications d'apparence de la fenêtre à la suite à cette interaction, sont reflétées automatiquement par le moteur OpenGL lorsqu'il est inoccupé. Toutes les fenêtres sont donc périodiquement remises à jour par le moteur. Ainsi, les animations automatiques sont également visibles. La figure 4.2 illustre la différence visuelle entre une application 3D exécutée dans un système de fenêtrage 2D et des applications 2D accessibles depuis le prototype de bureau 3D.

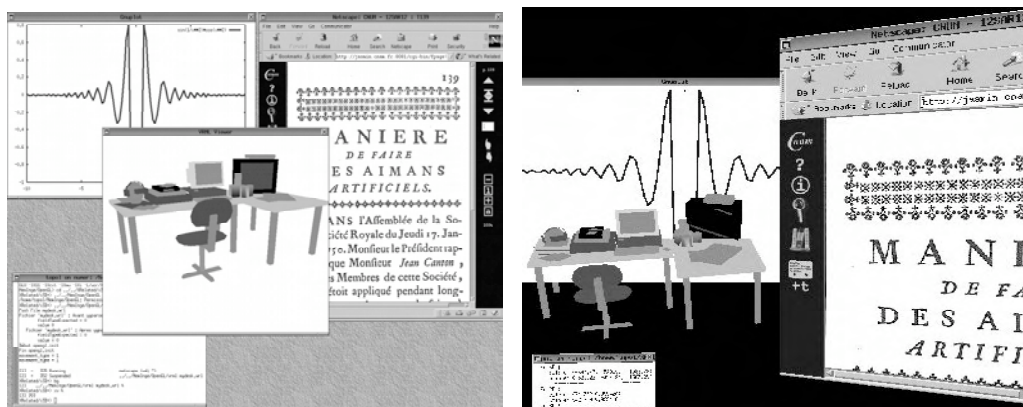
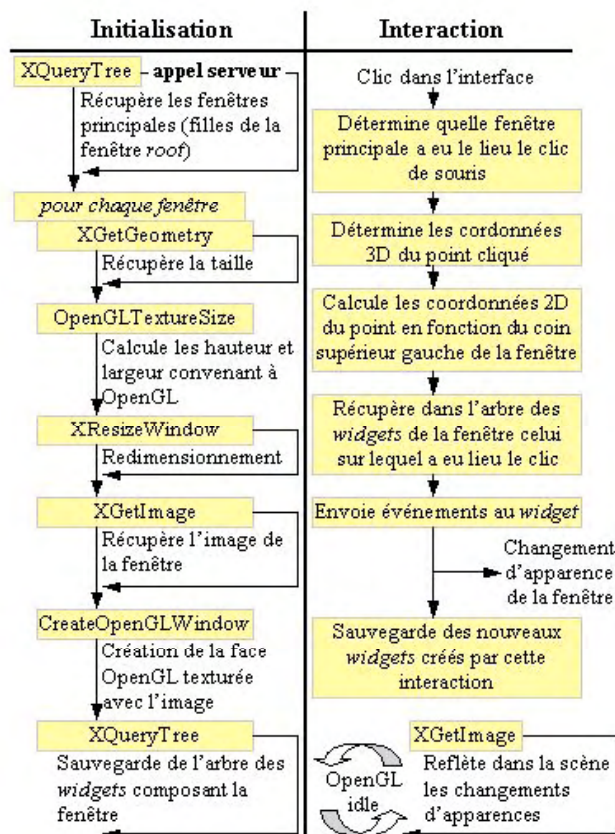


Figure 4.2. gauche : Fenêtre 3D insérée dans un bureau 2D (gnuplot, netscape, xterm et un navigateur VRML). Droite : Les mêmes documents dans l'environnement 3D prototypal.



Initialisation et gestion des interactions avec le contenu d'une fenêtre.

Dispositif immersif de visualisation

[Almeida, 2004 ; Almeida et Dupire, 2004 ; Almeida, Cubaud, Dupire et Topol, 2005]

En 1588, l'ingénieur italien Agostino Ramelli publiait en France ses *Diverse et artificieuse machine*. Cet ouvrage traite principalement de machines de pompage ou de guerre, mais l'une d'entre elles est particulièrement importante pour l'histoire de la visualisation d'informations : il s'agit de la roue à livre (fig. 4.3, gauche). Les 195 machines hypothétiques de Ramelli sont toutes décrites en une page en italien puis en français, avec une ou deux planches en regard. Le texte concernant la roue à livres est donc assez court pour être cité intégralement :

C'este cy est une belle & artificieuse machine, laquelle est fort utile & commode à toute personne qui se delecte à l'étude, principalement à ceux qui sont mal dispos & subjects aux gouttes ; car avec ceste sorte de machine un homme peut voir & lire une grande quantité de livres, sans se mouvoir d'un lieu : outre, elle porte avec soy une belle commodité, qui est de tenir & occuper peu de place, au lieu où on la met, comme tout homme d'entendement peut bien comprendre par son dessein. Cette rouë est faicte avec l'artifice que on voit, à sçavoir, elle est construitc de telle maniere, qu'en mettant les livres sur les tablettes, combien qu'on tourne la dicte rouë tout autour, jamais lesdits livres ne tomberont, ni se remueront du lieu où ils sont posés, ains demeureront tousjours en un mesme estat, & se représenteront devant

le lecteur en la mesme maniere qu'ils ont esté mis sur les tablettes. Ceste rouë se peut faire grande & petite, selon la volonté de celuy qui la faict faire, observant toutesfois les proportions de chascune partie des artifices de la dicte rouë, comme il pourra fort bien faire, considerant diligemment toutes les parties de ceste petite rouë, & les autres artifices qui se voyent en icelle machine : lesquelles parties sont faictes par mesures & proportions. Et pour donner plus grande intelligence & cognoissance à un chascun qui desirera faire mettre en œuvre la dicte machine, j'ay mis icy à part & découvert tous les artifices qui sont requis en telle machine, afin qu'un chascun les puisse mieux comprendre, & s'en servir à son besoin. (fol. 316 verso)

Il est tentant de reconnaître la roue de Ramelli comme le premier dispositif de lecture associant contexte et focus. Une autre caractéristique importante de la roue de Ramelli est qu'elle s'appuie sur la mémoire spatiale et les gestes du lecteur. En effet, celui-ci doit se souvenir de l'ordre dans lequel ses ouvrages sont placés et son geste va dépendre de la rotation à effectuer sur la roue. Mais le mouvement de l'utilisateur est directement associé à la navigation parmi les ouvrages stockés dans le dispositif. Les technologies de documents et de bibliothèques numériques d'aujourd'hui dépassent très largement les capacités de stockage de la roue à livre, et ne sont pas limitées au seul média textuel. Il faut toutefois reconnaître que dans la routine quotidienne de manipulation de documents numériques, les lecteurs sont souvent frustrés par leur équipement (fig 4.3, droite) et pourraient regretter le confort de visualisation et de manipulation offert par la roue de Ramelli...

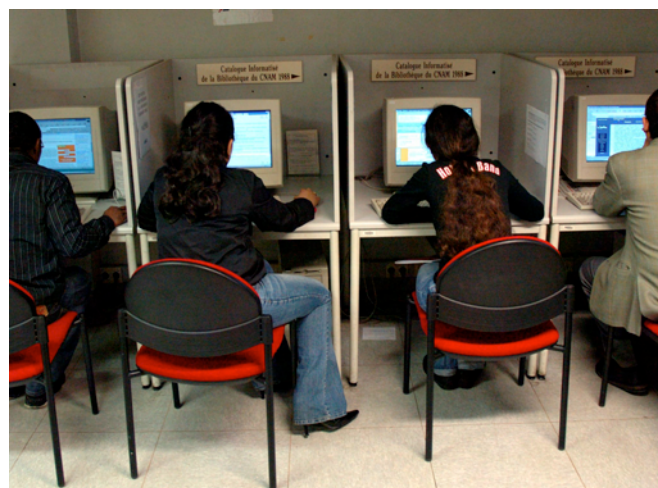
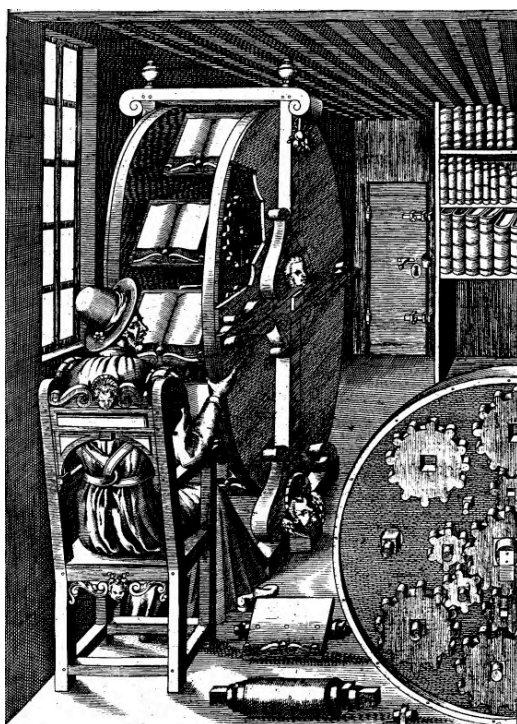


Figure 4.3. Roue de Ramelli (gauche) Postes de consultation en bibliothèque publique (droite).

Procurer à un utilisateur une meilleure immersion dans l'espace virtuel 3D nécessite de modifier les conditions de restitution visuelle de celui-ci [Lantz, 1997 ; McKinlay et al., 2003]. [Patrick et al., 2000] a montré que la visualisation d'une scène qui occupe tout le champ de vision de l'utilisateur peut lui apporter une plus grande appropriation cognitive de l'espace virtuel. Dans cette perspective, seuls de rares périphériques, comme le casque de réalité virtuelle et les écrans larges, rempliraient les conditions préconisées [Bryson et al., 1997]. Nous avons choisi comme système de visualisation la *VisionStation* de la société Elumens. La *VisionStation* est un écran hémisphérique de diamètre 1,5 m. couplé à un vidéo-projecteur à lentille modifiée pour fournir une image visible sur un champ de 160 par 160 degrés. L'utilisateur est assis devant une petite tablette qui contient le vidéo-projecteur et sur laquelle peut être placé un périphérique d'interaction. Selon la position de la tête de l'utilisateur, celui-ci peut avoir presque tout son champ de vision rempli par la projection (fig. 4.4). On peut donc qualifier la *VisionStation* de dispositif semi-immersif, en ce sens qu'il s'intègre à un dispositif matériel plus traditionnel au lieu de le remplacer.

La *VisionStation* n'est pas seulement un grand écran : un logiciel spécifique est utilisé pour adapter les images de la scène 3D projetées sur l'écran hémisphérique. Quatre rendus de la scène doivent être produits simultanément pour que le champ de vision atteigne 180° car une projection conventionnelle produirait des distorsions de perspective. Les quatre rendus doivent être assemblés en temps-réel dans une image unique, qui est ensuite déformée sphériquement pour apparaître plane sur l'écran (fig. 4.5).

Le principal inconvénient rencontré dans l'utilisation de la *VisionStation* comme outil de consultation de documents numériques est la résolution de l'image projetée. Elle est au maximum de 1024 x 768 pixels (limitation du vidéo-projecteur). Cette résolution reste très satisfaisante pour des applications comme les simulateurs de vol, les jeux vidéo et de manière plus générale, pour l'affichage d'images animées. En revanche, lors d'un travail plus précis sur des documents textuels ou graphiques, la qualité de l'image perçue devient un critère essentiel pour le confort et la bonne compréhension des informations par l'utilisateur [Kasik et al., 2002]. On sait aussi que la vitesse de lecture sur un écran d'ordinateur est proportionnelle à la résolution de l'image affichée [Mutter et al., 1991]. C'est dans ce contexte particulier que la *VisionStation* montre ses limites. On peut en effet discerner à l'affichage le détail des pixels projetés. Ce phénomène constitue un obstacle important à l'utilisation de ce dispositif, dans cette configuration, pour ce type d'application. [Baudish et al., 2001] a proposé un dispositif composé de matériel courant (un vidéo-projecteur, un écran de

projection et un moniteur) permettant de projeter simultanément une vue globale de la scène et une partie de celle-ci à une résolution plus élevée. Ce système a été testé pour la visualisation et le travail informationnel en 2D tel que l'analyse de cartes ou d'images satellitaires. Ce dispositif ne semble pas approprié pour gérer simultanément des contextes différents, la vue détaillée étant inévitablement une partie de l'image globale. De plus, le problème du bruit visuel lié aux écrans traditionnels se retrouve encore un peu dans ce système. Nous avons donc ajouté à notre dispositif un écran plat tactile (TFT 17 pouces). Celui-ci est capable d'afficher des résolutions plus fines que la VisionStation (jusqu'à 1280x1024). Nous avons envisagé l'utilisation de ce moniteur pour différentes tâches. Il pourrait permettre d'afficher une partie de la scène 3D, qui nécessiterait des conditions compatibles avec un travail sur les documents (confort visuel), ou d'isoler une partie de la scène, permettant une vision précise et plus détaillée de celle-ci. Nous pourrions ainsi permettre à l'utilisateur d'accéder, par cet affichage complémentaire, à des éléments d'interaction non disponibles dans la vue globale.

Plusieurs essais ont été nécessaires pour obtenir une visualisation satisfaisante des fac-similés d'ouvrages sur la VisionStation. Du fait du pitch important du projecteur, les images de fac-similés sont fortement moirées. Comme l'affichage de la scène nécessite quatre calculs de rendu par trame, un sur-échantillonnage s'avère trop consommateur d'espace mémoire. Une alternative satisfaisante consiste à flouter légèrement les fac-similés et à recourir à du « mip-mapping » pour gérer les transformations dynamiques de textures. Au final, le dispositif fonctionne correctement et est simple d'utilisation (fig. 4.6). Le déplacement dans la scène 3D pour la sélection d'ouvrages fonctionne de manière fluide, malgré la charge de calcul imposée au sous-système de rendu 3D (nous avons eu recours à une carte QuadroFX 3000 sur PC Pentium4 3,2 GHz). La VisionStation, à la différence d'autres dispositifs immersifs comme les casques de visualisation, ne fatigue pas l'utilisateur dans une utilisation longue (on regrette toutefois le bruit de ses ventilateurs !). La visualisation sur le deuxième écran offre un plus grand confort quand des informations plus détaillées sur le document désiré sont nécessaires.



Figure 4.4. Comparaison entre champs de vision à 60° (gauche) et 160° (droite). La quantité d'images affichées à droite est doublée

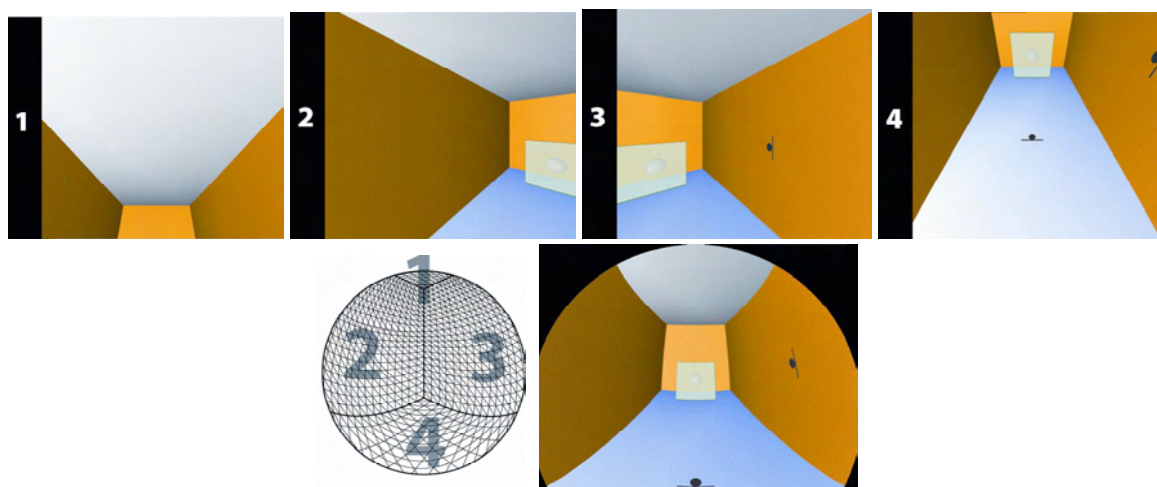


Figure 4.5. Exemple de scène avec les 4 rendus initiaux et (en bas) l'image résultante de la fusion des précédentes, après déformation.



Figure 4.6. Session de travail avec la VisionStation

Prolongements : dispositifs d'interaction pour le document 3D

La manipulation du trépied de lecture décrit au chapitre 2 et repris plus haut peut s'effectuer avec une souris ordinaire, par désignation de la base de celui-ci. Pour effectuer une rotation du trépied sur son axe vertical, l'utilisateur doit sélectionner un coin spécifique de cette base. Même si cette disposition peut être revue, il reste qu'il est impossible avec une souris à deux degrés de liberté de fournir simultanément les informations nécessaires au déplacement du trépied sur le plan de travail et son orientation angulaire sur l'axe vertical. Nous explorons actuellement deux directions : un dispositif généralisant la souris standard en lui ajoutant un degré de liberté supplémentaire, et un dispositif basé sur une désignation manuelle plus directe. Ces deux dispositifs sont en cours d'évaluation et ils devront ensuite être comparés entre eux.*

UNE SOURIS A 3 DEGRES DE LIBERTE

Nous proposons un dispositif d'entrée qui puisse contrôler, simultanément et à travers un « mapping » direct, trois degrés de liberté. A l'instar de la souris traditionnelle, ce dispositif doit être posé sur une surface plate, sur laquelle il peut être déplacé. Par ailleurs, outre le repérage des déplacements sur l'axe X (gauche, droite) et sur l'axe Y (proche, éloignée), ce dispositif doit aussi repérer la rotation autour de son axe vertical. Puisqu'il s'agit d'un dispositif généraliste, il sera capable d'opérer une rotation sur plusieurs objets virtuels différents. De cette façon, il accumulera une orientation qui sera la somme de toutes les rotations faites depuis la première opération. Pour que cette accumulation ne soit pas une contrainte pour l'utilisateur, il est nécessaire que ce dispositif ne soit pas attaché à un fil.

Travaux Similaires

Les recherches concernant les dispositifs d'entrée à degrés de libertés multiples sont très actives. Un premier état de l'art sur leur utilisation en interaction 3D a été fait dans [Topol, 2002]. S'agissant de notre projet spécifique, une expérience proche est décrite dans [McKenzie et al., 1997] avec l'utilisation de 2 boules dans une souris pour capter son orientation relative. Cependant, seule l'étape initiale de cette recherche a été publiée, avec des données très incomplètes. [Rekimoto et Sciammarella, 2000] décrit un dispositif qui capte six degrés de liberté. Il s'agit

* Ce travail s'effectue dans le cadre des thèses de R. Almeida et J. Dupire.

d'une boîte contenant 3 capteurs magnétiques de stylets Wacom. Lorsque cette boîte est manipulée (déplacée, tournée, inclinée) au-dessus d'une tablette, ces mouvements peuvent être repérés et correctement interprétés. [Balakrishnan et al., 1997 ; Hinckley et al., 1999] ont proposé la *VideoMouse*, une souris qui utilise une « webcam » comme capteur optique. Cette caméra est orientée vers le plan de travail (qui est éclairé par une Led). La *VideoMouse* ne peut pas être utilisée sur une surface quelconque car l'analyse d'images en temps réel qui est faite, permettant d'interpréter ses mouvements en termes de translations et de rotations, fonctionne uniquement sur une surface avec une grille imprimée. Cette grille permet de rendre plus rapides les calculs nécessaires à l'interprétation des déplacements. Les principaux désavantages de ce dispositif résident dans le besoin de cette surface spéciale, le temps de calcul des mouvements effectués (qui restent élevés) et le fait que ce dispositif soit relié à l'ordinateur au moyen d'un fil.

Notre prototype

Le dispositif que nous proposons est similaire à celui décrit par [McKenzie et al., 1997] puisque nous sommes aussi partis des composants des souris existantes (fig. 4.7). Deux souris, collées l'une à l'autre latéralement, nous fournissent quatre valeurs à chaque mouvement : la position relative de la souris gauche, notée $\langle X1, Y1 \rangle$ et la position relative de la souris droite, notée $\langle X2, Y2 \rangle$. Nous extrayons donc les moyennes de ces valeurs sur chaque axe pour avoir la moyenne relative de la position de notre dispositif. Si le dispositif n'est pas tourné depuis son initialisation, $Y1$ et $Y2$ restent identiques. En revanche, si l'on tourne le dispositif dans le sens horaire, par exemple, $Y1$ va croître, tandis que $Y2$ va décroître de façon symétrique (fig. 4.8). L'écart δ entre la valeur Y d'une des souris et la moyenne $(Y1+Y2)/2$ nous permet d'établir alors une relation linéaire entre δ et l'angle τ de la rotation opérée. Le coefficient de proportionnalité K sera lié, parmi d'autres facteurs, à la distance entre le centre du capteur optique de chaque souris. Ainsi, les mouvements de translation de la souris seront transmis en termes de moyennes de X et de Y , tant que son orientation relative sera transmise comme δ/K .

Même si le dispositif fonctionne correctement pour la détection des rotations, il reste des problèmes lors de mouvements de translation et de rotation intercalés et non corrigés. Pour corriger les translations il est toujours nécessaire de pré-calculer, à partir d'un mouvement sur X ou Y , l'impact final de ce mouvement sur les axes virtuels X et Y . Si, par exemple, le dispositif a été initialement tourné de 90 degrés, et qu'il a été ensuite déplacé vers le fond, c'est un déplacement sur l'axe X qui sera notifié. Pourtant, nous devons l'interpréter comme un déplacement sur

l'axe Y, puisque nous savons que le dispositif était tourné. Pour réaliser ce calcul de correction de l'axe X, par exemple, nous utiliserons le sinus de l'angle du dispositif pour extraire son déplacement virtuel en Y et son cosinus pour son déplacement virtuel en X. Ce seront uniquement ces valeurs finales qui seront envoyées vers l'application contrôlée par le dispositif.

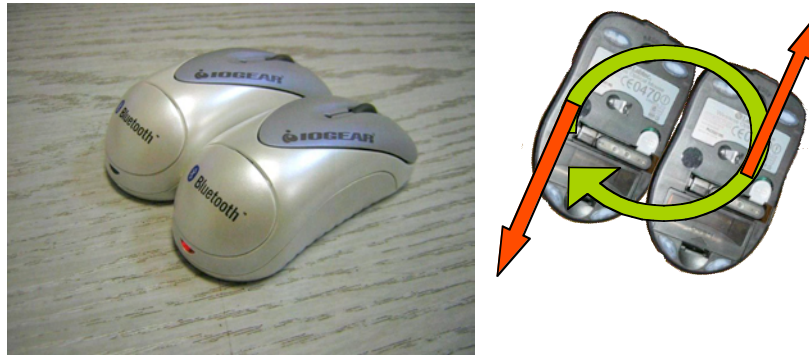


Figure 4.7. Captation de l'angle de lacet par couplage mécanique de souris standard

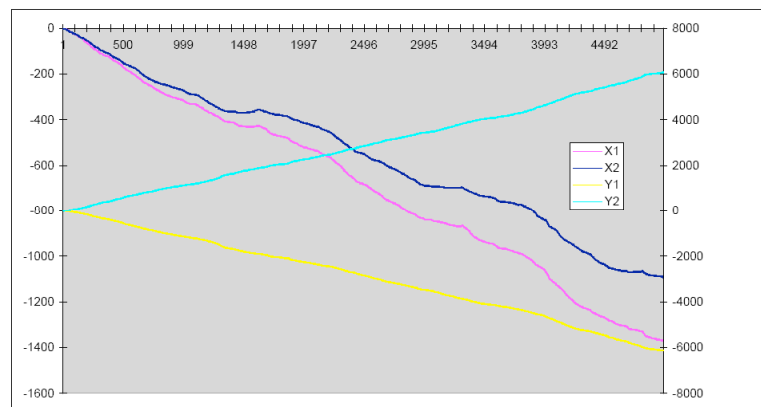


Figure 4.8. Evolution des valeurs sur les 4 axes pour une rotation horaire de 3 x 360°

L'intérêt de ce dispositif expérimental réside essentiellement dans sa mise en œuvre très simple, qui devrait favoriser la réplication d'expériences. Il existe certainement d'autres méthodes, basées par exemple sur l'utilisation d'accéléromètres intégrés*. Nous pensons que le dispositif actuel est de qualité suffisante pour procéder à une évaluation.

Acceptation

* Un tel dispositif semble avoir été développé dans le cadre d'un projet d'étudiants en Australie : <http://www.designawards.com.au/ADA/02-03/STUDENT%20DESIGN/668/668.HTM> (fev. 2006)

Dans le cadre d'une première évaluation pour la validation du dispositif, nous n'avons utilisé qu'un trépied 3D simplifié. De cette manière, nous ne testons qu'une variable : la performance du dispositif d'entrée pour le positionnement d'un document 3D sur le plan de travail. Pour cela, nous comparons le temps passé par un échantillon d'utilisateurs, à réaliser cette tâche avec différents dispositifs :

- la souris stéréoscopique « nue » (noté dans ce qui suit **3D**), réalisée avec deux souris de taille réduite sans fil (Bluetooth)
- le même dispositif, emballé dans une demi-sphère de polystyrène (**3DR**) (fig. 4.9),
- une souris traditionnelle, qui avec une touche peut alterner entre les modes translation et rotation (**2D**),
- une souris équipée d'une molette (**2DW**) servant à effectuer la rotation du document 3D.

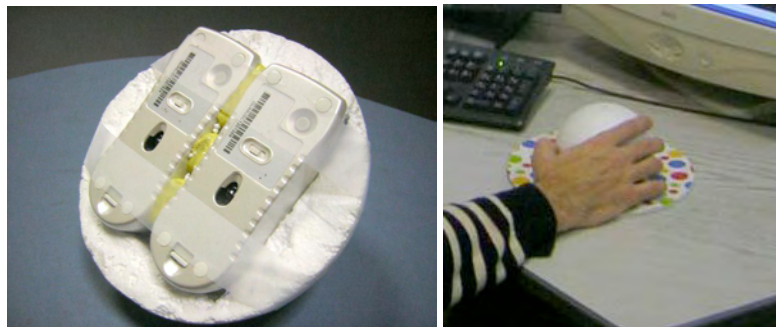


Figure 4.9. Habillage de la souris 3D dans une demi-sphère de polystyrène.

La tâche de test consiste à placer et orienter un document dans une boîte transparente qui est positionnée aléatoirement dans un espace 3D, ces deux objets reposant toujours sur le « sol » (fig. 4.10). Le temps pour compléter cette tâche est mesuré. En outre, nous demandons à l'utilisateur de classer les dispositifs entre eux et de répondre à un questionnaire.

Nous avons effectué en déc. 2005 une série de tests avec 16 participants volontaires non rémunérés (11 hommes, 5 femmes, droitiers, âge moyen 27,6 ans). Tous les participants ont une expérience d'utilisation de la souris avec ou sans molette. Treize personnes se sont déclarées des "utilisateurs confirmés d'ordinateurs" mais seulement quatre se considèrent des "joueurs sur ordinateur confirmés". Enfin, 12 participants ont déclaré avoir déjà utilisé d'autres types de périphériques (joystick, SpaceMouse, joypad, retour d'effort). Chaque participant a utilisé les 4 dispositifs dans une session unique d'une durée de 50 min. environ. Toutes les évaluations ont été faites la même semaine. Chaque test consiste en 20 mesures successives de temps de positionnement du document 3D. Les cinq

premières mesures sont en fait écartées de l'analyse. Cette procédure est semblable à celle décrite dans [Hinckley et al., 1997].

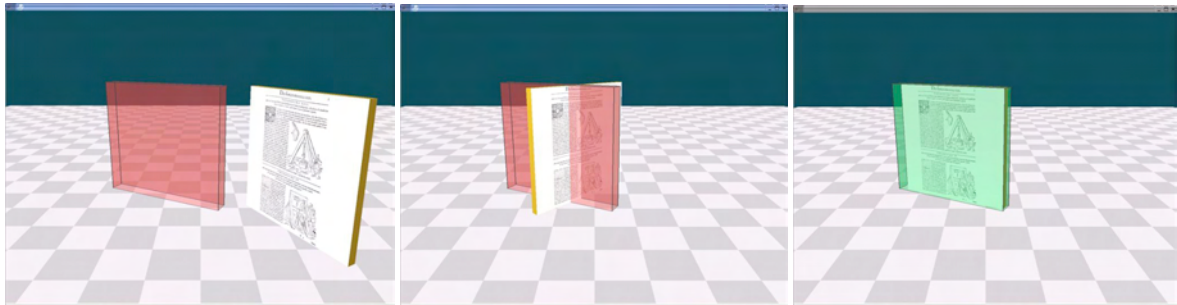


Figure 4.10. Les trois étapes du test de positionnement de trépied dans l'environnement 3D.

La fig. 4.11 reproduit l'histogramme des mesures obtenues. On voit que les deux dispositifs 3D ont conduit majoritairement à des temps de réaction plus courts. Une analyse factorielle ANOVA révèle que les écarts entre les dispositifs 2D et 3D sont significatifs. En revanche, on ne peut montrer de différence significative entre **3D** et **3DR**. Le dépouillement des questionnaires est actuellement en cours (février 2006).

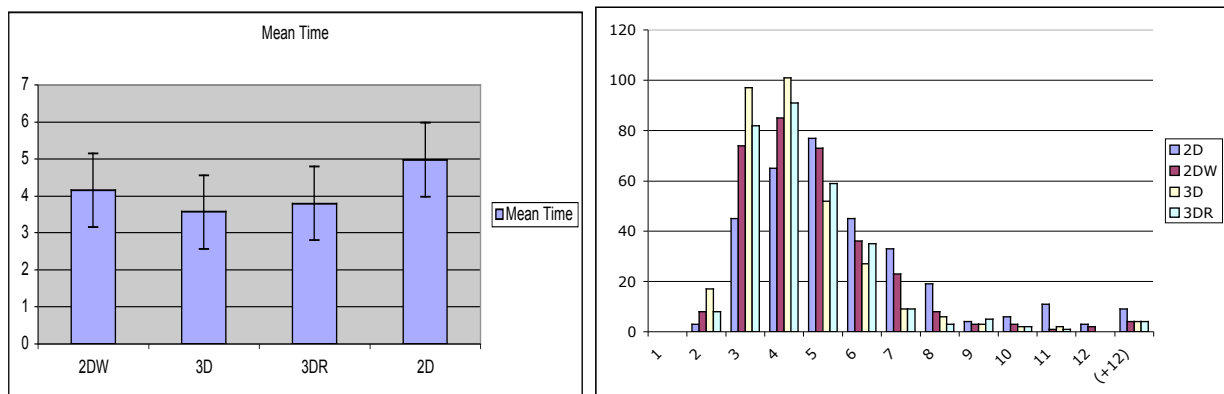


Figure 4.11. Droite : temps moyen de réaction (s) pour les 4 dispositifs et valeurs extrêmes. Gauche : Histogramme des temps de réaction (s) pour les 4 dispositifs.

Ces résultats, encore partiels, sont encourageants et nous envisageons d'étendre le champ d'application de notre souris stéréo, en particulier dans le contexte du jeu vidéo et des environnements collaboratifs [Kruger et al., 2005].

SELECTION DIRECTE

La tâche de sélection est de loin l'action la plus fréquente dans le contexte actuel des interfaces graphiques. Toujours associée à celle-ci, au point parfois de

confondre les deux concepts, la tâche de pointage est devenue l'incontournable préalable dans cette démarche globale de choix dans toutes les interfaces basées sur le modèle WIMP. En effet, quelle que soit la déclinaison du type d'environnement graphique considérée, on retrouve toujours un pointeur (ou l'une de ses métaphores, comme une main virtuelle par exemple) grâce auquel un utilisateur sélectionnera une cible, c'est-à-dire « [...] to identify a particular object among a whole of objects available » [Bowman et al., 2004].

En réduisant l'espace de sélection aux seuls objets existants dans la scène, l'utilisateur est capable d'en désigner un simplement, avec une amplitude gestuelle minime et sans manipulation complexe, par le déplacement de la sélection courante d'un objet à un autre. Cette technique permet de réduire la quantité de mouvement réalisée par un utilisateur, ainsi que la bande passante nécessaire pour véhiculer les informations utiles à la réalisation de cette action.

Etat de l'art

On retrouvera dans [Bowman et al., 2004] un état de l'art des techniques de pointage dans les environnements en trois dimensions (3D). Parmi elles, plusieurs s'appuient sur le principe du lancer de rayon, permettant de désigner un objet par l'intersection d'un rayon virtuel avec celui-ci. L'origine du rayon est souvent le point de vue de l'utilisateur (caméra virtuelle), la direction étant fixée par la position du pointeur dans l'environnement virtuel [Poupyrev, 1998]. Cette technique atteint rapidement ses limites lorsqu'elle est utilisée dans des scènes complexes et/ou denses, dans lesquelles des cas d'occlusion apparaissent fréquemment (objets alignés le long de l'axe de vision de l'utilisateur). Une amélioration de cette technique consiste à laisser la possibilité de spécifier librement l'origine du rayon virtuel, celle-ci pouvant donc être différente de la position de la caméra virtuelle. Cette variante constitue un début de solution au problème des occlusions, mais demeure assez lourde à manipuler : en plus des problèmes de précision de la trajectoire planifiée, deux étapes sont nécessaires pour désigner un objet, une première pour définir l'origine du rayon et la seconde pour en fixer la direction [Mine et al., 1997]. Une autre adaptation du lancer de rayon se retrouve dans la métaphore de la lampe torche [Liang, 1994]. L'objet est désigné en orientant le cône lumineux d'une torche virtuelle vers lui. Cette approche, si elle réduit favorablement le problème de précision évoqué auparavant, est encore moins performante face à celui des occlusions et de la densité de la scène. Enfin, une autre déclinaison du lancer de rayon, dans laquelle il est possible d'en fixer la longueur, donne une première solution aux cas d'occlusion. Cependant, cette approche est pénalisée, d'une part, par la mauvaise

perception de la profondeur dans les environnements en 3D et, d'autre part, par le niveau de précision important requis par son utilisation.

Des techniques de manipulation plus « directe » existent. L'intégration de pointeurs virtuels par exemple permet à un utilisateur d'interagir avec les objets de la scène. Là encore, de nombreuses déclinaisons existent. Parmi celles-ci, nous pouvons citer celle dans laquelle une main virtuelle est la reproduction exacte de la main physique de l'utilisateur (position, orientation et configuration). Des améliorations à cette configuration de base existent, visant principalement à réduire les déplacements de la main de l'utilisateur et la fatigue inhérente (*GoGo Technique* [Poupyrev et al., 1996], *world in miniature* [Stoakley, 1995], etc...).

Chacune des techniques présentées offre des avantages liés à un contexte d'usage particulier, mais aussi un ensemble d'inconvénients rarement négligeables. Nous proposons une technique de désignation d'objet, insensible aux phénomènes d'occlusion et possédant une latitude confortable (et personnalisable) dans l'évaluation de la précision des gestes de l'utilisateur, sans pour autant être plus fatigante pour ce dernier.

Principe du dispositif

A chaque déplacement significatif du périphérique de pointage est associé un saut d'un objet (le dernier sélectionné) vers son successeur (celui que l'on souhaite sélectionner), sans position intermédiaire. Le principe de base est de réduire l'espace de sélection initial (l'espace 3D continu) à un ensemble fini contenant les objets « sélectionnables » de la scène. Par la discrétisation de l'espace de sélection, la précision de la sélection de la cible en tant qu'objet de la scène est maximale, cette même sélection ne pouvant pas glisser ou manquer son objectif. Le temps de passage d'un objet à un autre est aussi fortement réduit, celui-ci étant uniquement dépendant de la période d'évaluation de l'atteinte du seuil critique de déplacement et du temps de traitement du saut par le PC (négligeable). Nous exploitons dans notre algorithme les caractéristiques neurophysiologiques de l'organisation du mouvement chez l'humain. En effet, la réalisation d'un geste volontaire suit une courbe en forme de cloche et se décompose en trois phases : la première, lente, est le temps de réaction. La seconde, rapide, est un déplacement grossier de l'effecteur vers la cible. La troisième, enfin, est lente et correspond à l'ajustement fin de l'effecteur sur la cible. Ce modèle s'applique bien sûr aux déplacements de la souris par un utilisateur. Notre méthode permet d'optimiser la commande motrice en réduisant le temps et l'énergie dépensés pour la sélection d'un objet, puisque seule la prise

en compte de la première phase du mouvement est nécessaire à la désignation d'un nouvel objet. En effet, la phase d'initiation du mouvement contient à elle seule toutes les caractéristiques géométriques du mouvement. La trajectoire (direction) est donc détectée très tôt. L'analyse des deux phases suivantes peut être évitée puisqu'elles n'apportent pas d'information supplémentaire. Par la discrétisation de l'espace de sélection, la distance inter-objets n'a plus d'influence sur les caractéristiques du mouvement et la sélection d'un objet n'est plus la conséquence d'un déplacement global et complexe d'un pointeur, mais simplement de la détection précoce d'une direction particulière, entraînant le passage de la sélection courante d'un objet vers un autre. Le gain est alors double, d'une part sur temps de réalisation et d'autre part sur la distance parcourue, sans parler de l'économie « physiologique ».

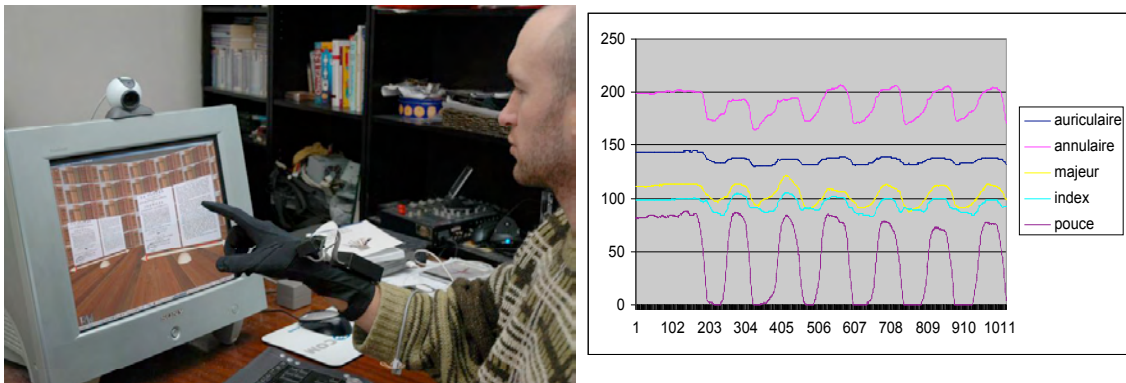


Figure 4.12. Gauche : Implémentation en cours du dispositif de sélection. Droite : Flexion des doigts mesurée par le capteur PATRIOT, évolution temporelle [Dupire et al., 2005].

Implémentation

Nous utilisons comme périphérique de pointage un *tracker* à 6 degré de liberté de type *Polhemus PATRIOT* pour la détection des mouvements et un gant de réalité virtuelle de type *Data Glove 5 (5DT)* pour la sélection des différents modes : déplacement du trépied ou sélection (fig. 4.12, gauche). La détermination des différents réglages de sensibilité (gain, seuils, etc.) fut la partie la plus longue de l'implémentation. En effet, il existe un certain nombre de mouvements parasites (les tremblements de la main par exemple) qui sont détectés par le PATRIOT mais qui sont du bruit pour notre système. On voit en effet sur l'enregistrement reproduit fig. 4.12 (droite) que la captation d'un mouvement répété est délicate. Nous souhaitons donc qu'un filtrage des mouvements de l'utilisateur soit effectué afin de ne prendre en compte que les mouvements significatifs, c'est-à-dire ceux dont les caractéristiques d'amplitude et de vitesse remplissent certaines conditions minimales. Pour cela, nous avons utilisé les réglages proposés dans [Zelevnik et Forsberg, 1999]. Ceux-ci correspondaient à la

détermination de deux seuils, l'un définissant une distance minimale de déplacement du capteur, l'autre une durée minimale de déplacement. Les valeurs respectives de distance parcourue par le pointeur (15 pixels) et de durée (100 ms) ont donc été implémentées puis adaptées à notre contexte particulier en 3D. Après quelques ajustements, il s'est avéré que les valeurs de 150 ms et 5 cm (de déplacement réel du capteur) constituaient une base moyenne correcte. Ces valeurs restent cependant à confirmer par la mise en place d'une évaluation de notre système par un panel plus large d'utilisateurs, qui nous permettra en outre de récupérer des données quantitatives sur les mouvements effectués et des retours d'expérience subjectifs.

Conclusion

When engine-driven vehicles were first introduced, they were built in the shape of a horse-drawn carriage and indeed were called "horseless carriages"; it took some time to take full advantage of the new technology and engineer the modern automobile. Much of DL [digital libraries] practice is still at the stage of the "horseless carriage"; we must move on to the modern automobile. [Soergel, 2002]

Le document textuel est actuellement la seule marchandise à circuler de manière économiquement significative sous une forme entièrement numérique, de son producteur/auteur à son consommateur/lecteur. L'explosion de la presse et des bibliothèques numériques en témoigne. Pourtant, l'écrit est le média dont la numérisation est la plus difficile (par comparaison au son ou à l'image animée). De plus, bien qu'étant une activité quotidienne pour des millions de gens, la lecture « à l'écran » (et/ou « en ligne ») est unanimement considérée comme pénible : le dispositif idéal de lecture de documents numériques reste à inventer.

Ma contribution à cette thématique de recherche pendant la décennie écoulée a été délibérément expérimentale. Elle peut se résumer de la manière suivante :

- J'ai conçu et participé à la réalisation de deux bibliothèques numériques importantes à l'échelle internationale : les sites ABU et CNUM.

- Par l'observation des usages effectifs de bibliothèques numériques, leur comparaison avec les « usages rêvés » [Bogros, 2003] de leurs concepteurs, j'ai contribué à mettre en évidence les limitations de ces dispositifs.

- J'ai ensuite étudié en quoi, en aval, des procédés d'interaction 3D pour la lecture, puis, en amont, de numérisation volumique, plus riches que ceux actuellement en usage, pouvaient contribuer à résoudre ces problèmes. Je tente d'évaluer ces procédés en utilisant les corpus et le public réel des sites que je gère.

C'est donc l'ensemble du processus numérisation/diffusion/usage du document que j'ai eu l'occasion d'étudier. Je n'ai clairement pas épuisé ce programme de travail, mais je voulais, par le biais d'un mémoire d'habilitation à diriger des recherches, en fixer l'état actuel et ma vision des pistes à venir.

Références bibliographiques

- ADLER R.J., FELDMAN R.E., TAQQU M.S. (eds.). A practical guide to heavy tails : statistical techniques and applications. Birkhäuser, 1998.
- ALISI T. M., BIMBO A. D., VALLI A. Natural Interfaces to Enhance Visitors' Experiences. *IEEE MultiMedia* 12(3), pp. 80-85. Jul. 2005.
- ANDRÉ J., FEKETE J.D., RICHY H. Traitement mixte image/texte de documents anciens. *Cahiers GUTenberg* n°21, pp. 75-85. Juin 1995
- ANDRÉ, J. Une vision historique du concept de « document numérique ». Document de travail, diffusé au sein du RTP Document Numérique (CNRS), mars 2003, 15 p.
- ASSADI H., BEAUVISAGE T. et al. Usages des bibliothèques en lignes. *Projet BibUsages. Rapport final*. France Télécom R&D et Bibliothèque nationale de France. Juillet 2003. 58 p. <http://www.bnf.fr>
- BALAKRISHNAN R., BAUDEL T., KURTENBACH G., FITZMAURICE G. The Rockin'Mouse : Integral 3D Manipulation on a Plane. *Proc. of ACM CHI'97*, pp. 311-318. 1997.
- BARBIER, F., BERTHO LAVENIR C. *Histoire des médias, de Diderot à l'Internet*. A. Colin, 1996.
- BAUDISH P., GOOD N., STEWART P. Focus plus context screens : combining display technology with visualization techniques. *Proc. of UIST'01*. Orlando, USA, Nov. 2001.
- BENFORD S., TAYLOR I., BRAILSFORD D., KOLEVA B., CRAVEN M., FRASER M., REYNARD G., GREENHALGH C. Three dimensional visualization of the World Wide Web. *ACM Comput. Surv.* vol. 31(4). Dec. 1999.
- BERNARD M. Mesurer le mètre ? Poésie et statistique. *Bulletin de l'Institut International de Statistique*. 52ème session (tome LVIII) ISI'99. Helsinki, Finlande, août 1999.
- BERTIN J., *La sémiologie graphique. Les diagrammes, les réseaux, les cartes*. Mouton, 1967.
- BERTINI E., CATARCI T., DI BELLO L., KIMANI S. Visualization in Digital Libraries in HEMMJE et al. (eds.) *Integrated Publication and Information Systems to Virtual Information and Knowledge Environments*, Springer-Verlag Berlin, 2005.

- BLONDEL C. L'électricité et le magnétisme au XVIIIe siècle à travers la bibliothèque virtuelle du CNAM. *Annales historiques de la révolution française*, 2000, n°2.
- BOGROS O. La bibliothèque électronique de Lisieux. Etats des lieux : mythes et réalité. *Bulletin des bibliothèques de France*, tome 48, n°4, 2003. pp. 45-48.
- BOGROS O. Une non-mesure de l'Internet : usages subjectifs et rêves d'un bibliothécaire, ses contes et mécomptes. in GUICHARD E. (dir) *Mesures de l'Internet*. Les canadiens en Europe. vol. 6, pp. 227-237, 2004.
- BOLT R. A. "Put-that-there" : Voice and gesture at the graphics interface. *Proc. of ACM SIGGRAPH'80*, Seattle, USA, pp. 262-270, 1980.
- BORGMAN, C.L. What are digital libraries? Competing visions. *Information Processing & Management*, 38(3), pp. 227-243, 1999.
- BOWMAN D., KRUIJFF E., LAVIOLA J., POUPYREV I. *3D user interfaces : theory and practice*. Addison-Wesley, 2004.
- BRETHES, J.C. *Histoire de la télécopie*. Que-sais-je n°3007, PUF, 1995.
- BROWN M. S., SEALES W.B. Document Restoration Using 3D Shape : A General Deskewing Algorithm for Arbitrarily Warped Documents. *Proc. Int'l Conference On Computer Vision (ICCV'01)*, Vancouver, 7-14 juillet 2001, vol.2, pp.367-375.
- BROWN M. S., TSOI D. Correcting Common Distorsions in Camera-Imaged Library Materials. *Proc. IEEE & ACM Joint Conference on Digital Libraries (JCDL'03)*, Houston, may 2003, p.367-368.
- BROWN M.S., SEALES W.B. Beyond 2D images : Effective 3D imaging for library materials. *Proc. of the 5th ACM conf. on Digital Libraries*, San Antonio, June 2000, p. 27-36.
- BRYAN-KINNS N., BLANDFORD A. A survey of user studies for digital libraries. *RIDL Working paper*, July 2000.
- BRYSON S., ZELTZER D., BOLAS, M.T. et al. The future of virtual reality : head mounted displays versus spatially immersive displays. *Proc. of ACM SIGGRAPH'97*, pp. 485-6.
- BUCKLAND, M. Emanuel Goldberg, electronic document retrieval, and Vannevar Bush's Memex. *JASIS* 43(4), pp. 284-294, may 1992.
- BUCKLAND, M. Histories, heritages and the past : the case of Emanuel Goldberg. *Proc. 2nd conf. on the History and Heritage of Sientific and Technical Information Systems*, Philadelphia, nov. 2002.

- CALABRETTO S., PINON J.M., BOZZI A. BAMBI : Système de Gestion de Manuscrits Anciens pour Historiens. *Document Numérique*. Vol. 2, n° 3-4, 1998. pp. 31-50.
- CARD S.K., HONG L., MACKINLAY J.D., CHI E.H.. 3Book: A Scalable 3D Virtual Book ». *Proc. of ACM CHI'04*. Vienna, Austria, April 2004.
- CARD S.K., MACKINLAY J.D., SHNEIDERMAN B. (eds.) *Readings in Information visualization. Using Vision to Think*. San Francisco, Morgan Kaufmann, 1999.
- CARD S.K., ROBERTSON G., MACKINLAY J.D., The Information Visualizer, an Information Workspace. *Proc. of ACM CHI'91*, New Orleans, USA, 1991.
- CARD S.K., ROBERTSON G., YORK W. The WebBook and the Web Forager : An Information Workspace for the World-Wide-Web. *Proc. of ACM CHI'96*. Vancouver, Canada, April 1996.
- CARPENTER L., SHAW S., PRESCOTT A. (eds) *Towards the digital library : The british library initiatives for access programme*. London, British lib., 1998 (p. 61 et suiv.)
- CHEN C. *Information visualisation and virtual environments*. Springer Verlag, 1999.
- CHEN C., BÖRNER, K. Top ten problems in visual interfaces to digital libraries. *in* K. Börner, C. Chen (eds) *Visual interfaces to digital libraries*, LNCS 2539, 2002.
- CHERRY J.M., DUFF W.M. Studying digital library users over time: a follow-up survey of Early Canadiana Online. *Information Research*, vol. 7 (2), 2002.
- CHRISTOFFEL M., SCHMITT B. Accessing libraries as easy as a game. *in* K. Börner, C. Chen (eds) *Visual interfaces to digital libraries*, LNCS 2539, 2002.
- CHU Y.C., BAINBRIDGE JONES D., WITTEN I. Realistic books : a bizarre homage to an obsolete medium ? *Proc. of ACM-IEEE JCDL'04*, Tucson, June 2004.
- CIM, A. *Le livre*. Flammarion, 1900 (4 vol.).
- COCKBURN A., MCKENZIE B. 3D or not 3d? Evaluating the effecte of the third dimension in a document management system. *Proc. of the ACM Conference on Human Factors in Computing Systems (CHI'2001)*, Seattle, USA, March 2001.
- COOLEY R. The use of web structure and content to identify subjectively interesting web usage patterns. *ACM Trans. Internet Technology* 3(2), pp. 93-116, May 2003.

- CUBAUD P. ABU : Un serveur de texte intégral sur Internet. *Bulletin bibliographique de l'Institut nat. des Techniques Documentaires*, Vol. 19, n°5, pp. 3-6. 1995.
- CUBAUD P., DEBLOCK G. Le conservatoire numérique des arts et métiers : historique du projet et organisation du site. *Actes du colloque La numérisation des textes et des image : techniques et réalisations*. Editions de l'Univ. Lille 3, pp. 13-22. Janvier 2003.
- CUBAUD P., DUPIRE J, TOPOL A. Digital libraries at CNAM : 1993-2003. *Proc. of ECDL'03*. Trondheim, Norway, Aug. 2003.
- CUBAUD P., DUPIRE J., TOPOL A. Digitization and 3D modeling of movable books. *Proc. of ACM-IEEE JCDL'05*, Denver, USA, pp. 244-5, June 2005.
- CUBAUD P., DUPIRE J., TOPOL A. Textes, images, volumes : les bibliothèques numériques au CNAM. *Actes Premières journées francophones de la toile (JFT)*. Tours, France, july 2003.
- CUBAUD P., DUPIRE J., TOPOL A. Using game engines for non 3D gaming applications. *Proc. of CGAMES'05*. Angoulême, France, nov. 2005.
- CUBAUD P., GIRARD D. ABU : une bibliothèque numérique et son public. *Document Numérique*, vol. 2(3-4), 1998, pp. 13-30.
- CUBAUD P., HAAS J.F., TOPOL A. Numérisation 3D de documents par photogrammétrie. *Proc. 8ème conf. francophone sur l'écrit et le document (CIFED'04)*. La Rochelle, France, June 2004.
- CUBAUD P., STOKOWSKI P., TOPOL A. Mixing Browsing and Reading Activities in a 3D Digitalized Library. *Proc. of ACM-IEEE JCDL'02*, Portland, USA, June 2002.
- CUBAUD P., THIRIA C., TOPOL A. Experimenting a 3D Interface for the access to a Digital Library. *Proc. of ACM DL '98*. Pittsburg, USA, June 1998.
- CUBAUD P., TOPOL A. A WWW-based digital library for antiquarian collection. *Rapport de recherche du CEDRIC RR99-09*. Mars 1999.
- CUBAUD P., TOPOL A., VODISLAV D. Les limites de VRML pour les comportements interactifs : étude de cas. *Actes de la conférence ERGO-IHM'2000*, Biarritz. Oct. 2000.
- CUBAUD P., TOPOL, A. A VRML-based user interface for an online digitalized antiquarian collection. *Proc. ACM SIGGRAPH Web3D Symp.*, Padderborn, April 2001, pp. 51-59.
- DEBLOCK G., ROZET B., CUBAUD P. Le Conservatoire numérique des arts et métiers : une création partenariale. *Bulletin des bibliothèques de France*, t. 46, n°4, 2001, pp. 43-49.

- DEBORA : Digital Access to Books of the Renaissance. *EC Telematics app. program #5608*. <http://debor.enssib.fr>
- DEQUIER A., *La numérisation 3D à des fins culturelles*. Mémoire du DEA ESTC sous la direction de P. Cubaud, CNAM, Paris, 2000.
- DHYANI D., NG W.K., BHOWMICK, S.S. A survey of Web metrics. *ACM Comput. Surveys* 34(4), pp. 469-503. Dec. 2002.
- DUFF W.M., CHERRY J.M. Use of historical documents in a digital world : comparisons with original materials and microfiche. *Information Research*, Vol. 6 No. 1. October 2000.
- DUPIRE, J., FARINONE, J.M., CUBAUD, P. Un nouveau périphérique : la flûte virtuelle. Actes de l'Atelier Formations transdisciplinaires en interaction sensorielle pour la création artistique, IHM'2005, Toulouse, 2 p., 2005.
- EGELS Y., KASSER M. *Photogrammétrie numérique*. Hermès Science Publications, 2001, 661 p.
- FAYET-SCRIBE, S. Chronologie des supports, des dispositifs spatiaux, des outils de repérage de l'information. *Solaris*, sept. 1997.
- FINGERHUT M. Le site Web de la bibliothèque considéré comme un espace. *Bulletin des bibliothèques de France*. t. 45, n°3, mai 2000.
- FOX, E. A., AKSCYN, R. M., FURUTA, R. K., and LEGGETT, J. J. Digital libraries. *Comm. ACM* 38(4) pp. 22-28. Apr. 1995.
- FOX E., KIPP N. et al. Networked Digital Library of Theses and Dissertations : An International Effort Unlocking University Resources. *D-lib magazine*, Sept. 1997.
- GERSHON N., EICK S. Foreword. *Proc. IEEE Symp. on Information Visualization, (InfoVis'95)*, IEEE CS Press, 1995.
- GILLIES, J., CAILLIAU, R. *How the Web was born*. Oxford Univ. Press, 2000.
- GOULD, ALFARO, FONN, HAUPT, MINUTO, SALAUN. Why Reading was slower from CRT displays than from paper. *Proc. of ACM CHI+GI'87*, Toronto, Canada, 1987.
- GOULD, GRISCHOKOWSKY. Doing the same work with hard copy and with Cathode-Ray tube (CRT) computers terminals. *Human Factors Journal*, vol. 26(3), pp. 323-337, 1984.
- GRAHAM, R, RAYWARD, W.B. (eds). Computer applications in libraries. Num. spéciaux de *IEEE Annals of the history of computing*. vol. 24, n°2 et 3, Avril à nov. 2002.
- GUICHARD E. (dir.) *Mesures de l'internet*. Les canadiens en europe, vol. 6. 2004

- HABERT, B., FABRE C., ISSAC F. *De l'écrit au numérique. Constituer, normaliser et exploiter les corpus numériques*. InterEditions, 1998.
- HAINING P. *Movable books, an illustrated history*. New English Library, 1979.
- HEMMINGER B., BOLAS G., SCHIFF D. Capturing content for virtual museums : from pieces to exhibits. *J. of Digital Information*, vol. 6(1), march 2005.
- HINCKLEY K., SINCLAIR M., HANSON E., SZELISKI R., CONWAY M. The VideoMouse: a camera-based multidegree-of-freedom input device. *Proc. of ACM UIST'99*, pp. 103-112, 1999.
- HINCKLEY K., TULLIO J., PAUSCH R., PROFFITT D., KASSELL N. Usability Analysis of 3D Rotation Techniques. *Proc. of UIST'97*, pp. 1-10. 1997.
- IDE N., VERONIS J. *The Text Encoding Initiative : Background and Context*. Kluwer Academic Publishers, Dordrecht, 1995.
- JACQUEMIN C., JARDINO M. Multi-dimensional and Multi-scale visualizer of large XML documents. *Proc. of Eurographics'02*, Saarbrücken, Germany. 2002.
- JONES S., CUNNINGHAM S. J., McNAB R. Usage analysis of a digital library. *Proc. Third ACM Conference on Digital Libraries*, Pittsburgh, USA, pp. 293-294, June 1998.
- KASHIWAZAKI K., KIKUTA Y., OHIGASHI M., TANAKA Y. Interactive 3D learning contents based on 3D meme media architecture. *Proc. of the 4th int. Symp. on information and Communication Technologies (Cape Town, South Africa, Jan. 2005)*. *ACM International Conference Proceeding Series*, vol. 92, pp. 14-19. 2005.
- KASIK D., TROY J., AMOROSI S., MURRAY M., SWAMY S. Evaluating Graphics Displays for Complex 3D Models. *IEEE Computer Graphics & Applications*, 22(3):56-64. 2002.
- KRUGER R., CARPENDALE S., SCOTT S. D., TANG A. Fluid integration of rotation and translation. *Proc. ACM CHI'05*, Portland, USA, pp. 601-610, april 2005.
- LAFOUGE, T., LE COADIC, Y.-F., MICHEL, C. *Eléments de statistique et de mathématique de l'information*. Cahiers de l'ENSSIB n°1. Presses de l'ENSSIB, 2003.
- LANTZ E. Future Directions in Visual Display Systems. *Computer Graphics*, 31(2), pp. 38-45, 1997.

- LAUFER R., SCAVETTA D. *Texte, hypertexte, hypermédia*. Que-sais-je n°2629, PUF, 1995 (2ème ed.)
- LE COADIC Y. *Usages et usagers de l'information*. Nathan, 1997.
- LEACH G., AL-QAIMARI G., GRIEVE M., JINKS M., MACKAY C. Elements of the Three-Dimensional Graphical User Interface, *Proc. of INTERACT'97*, Sydney, Australia, 1997.
- LEBART L., SALEM A. *Statistique Textuelle*. Dunod. 1994.
- LEBRAVE J.L., GRÉSILLON A., FUCHS C. ; Flaubert : "Ruminer Hérodiad" Du cognitif-visuel au verbal-textuel, in *L'Écriture et ses doubles : genèse et variation textuelle*. Editions du CNRS, p. 27-109. Coll. Textes et Manuscrits. 1991
- LECLERC Y. *La spirale et le monument. Essai sur Bouvard et Pécuchet de Gustave Flaubert*. CDU-SEDES, 1988.
- LECLERC Y., CUBAUD P. Bouvard : 618, Pécuchet : 598. Etude de statistique textuelle. *Revue Flaubert* n°3, 2003. www.univ-rouen.fr/flaubert
- LECOLINET E., LIKFORMAN-SULEM L., ROBERT L., ROLE, F, LEBRAVE J.L.. An integrated reading and editing environment for scholarly research on literary works and their handwritten sources. *Proc. ACM conf. on Digital Libraries*, Pittsburgh, 1998, pp. 144-151.
- LESK, M. *Practical Digital Libraries : Books, Bytes, and Bucks*. Morgan Kaufmann, 1997.
- LEVOY M. et al. The Digital Michelangelo Project: 3D Scanning of Large Statues. *Proc. ACM SIGGRAPH'00*, New Orleans, USA, 2000.
- LIANG, J., GREEN M. JDCAD: A Highly Interactive 3D Modeling System. *Computers and Graphics* 18(4), pp. 499-506, 1994.
- LUPOVICI C., CLOAREC T., DE CHARENTENAY F. Les usages de Gallica. *Bulletin des bibliothèques de France*, tome 48, n°4, pp. 40-44, 2003.
- MAIGNIEN, Y. Chronique partielle d'une bibliothèque virtuelle. *Document numérique* vol. 2(3-4), pp. 281-290, 1998.
- MANGUEL A. *Une histoire de la lecture*. Actes Sud, 1998.
- MARSHALL C.C., BLY S. Turning the page on navigation. *Proc. ACM-IEEE JCDL'05*, June 2005, pp. 225-234.
- MARTIN, H.J. (avec la collab. de Bruno Delmas) *Histoire et pouvoirs de l'écrit*. Perrin, 1990, 518 p.

- McKENZIE I.S., SOUKOREFF R.W., PAL C. A Two-Ball Mouse Affords Three Degrees of Freedom. *Proc. of ACM CHI'97 Conference Companion*, pp. 303-304, 1997.
- McKINLAY J.D., HEER J., ROYER C. Wideband Visual Interfaces : Sensemaking on Multiple Monitors. *PARC Technical Report*, 2003.
- MINE M., BROOKS F., SEQUIN C. Moving Objects in Space: Exploiting Proprioception in Virtual Environment Interaction. *Proc. ACM SIGGRAPH*. 1997.
- MINTZER F. C., BOYLE L. E., CAZES A. N., CHRISTIAN B. S., et al. Toward online, worldwide access to Vatican Library materials. *IBM J. Res. Dev.* 40(2), pp. 139-162. Mar. 1996.
- MUKHERJEA S. Information visualization for hypermedia systems. *ACM Comput. Surv.* 31(4) Dec. 1999.
- MULLER Ch. *Étude de statistique lexicale. Le vocabulaire du théâtre de Pierre Corneille*. Larousse, 1967.
- MUTER P., MAURUTTO P. Reading and skimming from computer screens : the paperless office revisited. *Behavior and Information Technology*, 10(4), 257-266, 1991.
- NYCE, J., KAHN, P. *From Memex to hypertext : Vannevar Bush and the Mind's machine*. Academic Press, 1991.
- O'HARA K., SELLEN A. A comparison of reading paper and on-line documents. *Proc. ACM conf. on Human factors in Comp. Sys. (CHI'97)*, Atlanta, 1997.
- OTLET P. *Traité de documentation. Le livre sur le livre, théorie et pratique*. Bruxelles, van Keenberghen, 1934. Réimpr. Liège, C.L.P.C.F., 1989.
- PACKER, R. JORDAN, K. (eds.) *Multimedia : from Wagner to Virtual reality*. Norton, 2001.
- PASTOR-SATORRAS R., VESPIGNANI A. *Internet, structure et évolution*. Belin, 2004.
- PATRICK E., COSGROVE D. et al. Using a Large Projection Screen as an Alternative to Head-Mounted Displays for Virtual Environments. *Proc. ACM conf. on Human factors in Comp. Sys. (CHI'00)*, Avril 2000, pp. 478-485.
- PEDAQUE, R. T. (collectif) *Document : forme, signe et médium, les reformulations du numérique* (08/07/2003) Diffusé sur le site du RTP DOC Document et contenu : création, indexation et navigation <http://rtp-doc.enssib.fr/>

- PELFRENE, A. Les postes de lecture assistée par ordinateur (PLAO) in C. Paganelli (dir.) *Interaction homme-machine et recherche d'information*. traité des sciences et techniques de l'information. Hermès. 2002.
- PETERS T.A. The history and development of transaction log analysis. *Library Hi Tech* 42 (11:2), pp. 41-66, 1993.
- PETRUSZEWYCZ M. *Les chaînes de Markov dans le domaine linguistique*, Slatkine, 1981.
- PILU M. Undoing Paper Curl Distorsion Using Applicable Surfaces. *Proc. Computer Vision and Pattern Reconition (CVPR 01)*, Hawaii, dec 2001.
- PLÉNACOSTE P., LECOLINET E. et al. Zoomable and 3D Representations for Digital Libraries. *Proc of IHM-HCI 2001*, Lille, France, Sept. 2001
- POUPYREV I., WEGHORST S., et al. Egocentric Object Manipulation in Virtual Environments : Empirical Evaluation of Interaction Techniques. *Computer Graphics Forum*, 17(3), 1998.
- POUPYREV, I. et al. The Go-Go Interaction Technique: Non-linear Mapping for Direct Manipulation in VR. *Proc. ACM UIST*, 1996.
- RAUBER A., BINA H. Visualizing electronic document repositories: drawing books and papers in a digital library. *Proc. of the 5th IFIP 2.6 work. conf. on Visual Databases Systems (VDB5)*, Fukuoka, Japan, May 2000.
- REKIMOTO J., SCIAMMARELLA E. ToolStone : effective use of the physical manipulation vocabularies of input devices. *Proc. of the 13th annual ACM symp. on User interface software and technology*, San Diego, USA. p.109-117, Nov. 2000.
- RELIER, Y. *Le microfilm*. Que-sais-je n°1221, PUF, 1966.
- RICHARDSON T., STAFFONT-FRASER Q., WOOD K.R., HOPPER A. Virtual Network Computing. *IEEE Internet Computing*, vol. 2(1), pp. 33-38, jan. 1998.
- ROBERT L., LECOLINET E. Techniques d'interaction et de visualisation pour l'accès à des documents numérisés. *Actes de la conférence ERGO-IHM*, Biarritz, pp. 178-185. 2000.
- ROBERTSON G. , VAN DANTZICH M., ROBBINS D., CZERWINSKI M., HINCKLEY K., RISDEN K., THIEL D., GOROKHOVSKY V. The task gallery : a 3D window manager. *Proc. ACM CHI'2000*, ,The Hague, Netherlands, pp. 494-501, 2000.
- ROBERTSON G., CARD S., MACKINLAY J. Information visualization using 3D interactive visualization. *Comm. of the ACM*, 36(4), pp. 56-71, April 1993.
- ROBERTSON G., VAN DANTZICH M., ROBBINS D., CZERWINSKI M., HINCKLEY K., RISDEN K., THIEL D., GOROKHOVSKY, V. The task

- gallery: A 3D window Manager. *Proc. of ACM CHI'2000*, The Hague, Netherlands, pp. 494-501, 2000.
- ROLE F. *TEI (Text Encoding Initiative)*. in *Techniques de l'ingénieur; série H : Informatique (H7158)*. 2000.
- ROUET J.F. *Cent fenêtres sur Internet*. Atlantique, éditions de l'Actualité Poitou-Charentes, 2005.
- ROUSSEL, N. Exploring new uses of video with videoSpace. *Proc. of EHCI'01, the 8th IFIP International Conference on Engineering for Human-Computer Interaction*. Springer, pp. 73-90, 2001
- ROUYER, P. *L'âge d'or du microfilm*. Thèse de doctorat de l'université Paris III, 1999.
- SALOMON D. *Data compression, the complete reference*. Springer Verlag, 1997.
- SAVARY J.P. Proposition méthodologique d'évaluation. Le cas de la conception d'une bibliothèque virtuelle pour les personnes non-voyantes. *Mémoire pour le diplôme d'ergonome du CNAM* (P. Beguin, dir.), 122 p., 2002.
- SEBRECHTS M.M., CUGINI J., LASKOWSKI S.J., VASILAKIS J., MILLER M.S. Visualization of Search Results: A Comparative Evaluation of Text, 2D, and 3D Interfaces. *Proc. of ACM SIGIR'99*, 1999.
- SOERGEL D. A Framework for Digital Library Research. Broadening the Vision. *D-Lib Magazine*. Vol. 8(12). Dec. 2002.
- STERN D. (ed.) Digital libraries. Philosophies, technical design considerations, and example scenarios. *Science and tech. libraries* 17 (3-4), 1999.
- STOAKLEY R., CONWAT M., PAUSCH R. Virtual Reality on a WIM : Interactive World in Miniature. *Proc. ACM CHI*, 1995.
- SUGAYA N. Bouvard et Pécuchet, l'exposition critique d'un paradigme médical. *Revue Flaubert* n°4, 2004. <http://www.univ-rouen.fr/flaubert/>
- TANASA D., TROUSSE B., MASSEGLIA F., Fouille de donnée appliquée aux logs web : état de l'art sur le Web usage mining. in GUICHARD E. (dir.) *Mesures de l'Internet*. Les canadiens en Europe. vol. 6, pp. 126-143, 2004.
- THIRIA, C. *Bibliothèques numériques : du monumental au virtuel*. Mémoire du DEA ESTC sous la direction de P. Cubaud, CNAM, Paris, 1999.
- TOPOL A. Immersion of XWindow applications into a 3D workbench. *Proc. ACM CHI'2000*, The Hague, Netherlands, pp. 355-6, 2000.
- TOPOL, A. *Interaction 3D pour les paysages informationnels*. Thèse de doctorat en informatique. Conservatoire national des arts et métiers. Décembre 2002.

- TRINH E., LEBOURGEOIS F. Compression et formats de documents textuels. *Actes de la journée d'information et d'étude sur la numérisation des collections*, INSA-RFV et ENSSIB, Villeurbanne, Juin 2001.
- TUFTE E.R. *The Visual Display of Quantitative Information*, Graphics Press, 1992.
- ULGES A., LAMPERT C. H., BREUEL T. Document capture using stereo vision. *Proc. of the 2004 ACM Symposium on Document Engineering*. Milwaukee, USA, Oct. 2004. pp. 198-200.
- VAN DAM, A. Escaping flatland in user interface design. *Proc. Symposium on Interactive 3D graphics*, Cambridge, MA, Etats-Unis, 1992, p. 9
- VATÉ M. *Statistique chronologique et prévision*. Economica, 1993.
- VIRBEL, J. Reading and managing texts on the Bibliothèque de France station. In G. P. Landow et P. Delany (eds.) *The Digital Word: Text-Based Computing in the Humanities*, Mit Press On Technical Communication And Information Systems Series, vol. 5. MIT Press, pp. 31-51, 1993.
- VUILLEMIN A. L'informatique et les études littéraires. *ASYMETRIA, Revista de cultura, critica si imaginatie*. I(3). Janvier 2001. <http://www.asymetria.org/vuillemininformatique.html>
- WADA T., UKIDA H., MATSUYAMA T. Shape from shading with interreflections under proximal light source : 3D shape reconstruction of unfolded book surface from a scanner image . *Proc. Int. Conf. of Computer Vision (ICCV'95)*, 1995, pp 66-71.
- WESTERMAN S.J., COLLINS J., CRIBBIN, T. Browsing a document collection represented in two- and three-dimensional virtual information space. *Int. J. of Human-Computer studies*, vol. 62, pp. 713-736.
- WITTEN, I. et al. How to turn the page. *Proc. ACM-IEEE JCDL'03*, Houston, may 2003. pp. 186-188.
- ZELEZNIK R., FORSBERG A. UniCam—2D gestural camera controls for 3D environments. *Proc. ACM Symposium on Interactive 3D Graphics*, 1999.



Résumé :

Depuis les premières expériences au début de la décennie 1990, les bibliothèques numériques ont bénéficié du progrès constant des technologies de captation, de stockage et de transmission numérique ainsi que de la chute de leurs coûts. Le développement du World-Wide-Web a également permis d'atteindre un public international considérable par le biais d'une interface standard et commode. Une dimension nouvelle vient d'être atteinte fin 2005 avec l'annonce de projets très ambitieux : Google-Print et Bibliothèque Numérique Européenne, entre autres. Ces annonces ne doivent pas faire oublier que plusieurs difficultés demeurent. Du point de vue technique (le seul abordé dans ce mémoire), la principale réside à mon sens dans l'effet de moiré que produit la numérisation, comprise non seulement comme processus d'échantillonnage d'un signal (ici, celui provenant d'une caméra CCD visant une page de livre), mais en un sens plus large, qui englobe le support et l'environnement physique du texte, les gestes et les démarches des lecteurs. Il paraît donc nécessaire d'accroître le volume des informations extraites lors du processus de numérisation et d'étendre, par le biais d'interfaces plus fluides, le fonctionnement du dispositif offert au public. Les deux aspects sont évidemment liés.

Ce mémoire résume mes contributions à ces questionnements sur les bibliothèques numériques et l'interaction humain-machine. Le premier chapitre décrit les deux projets de bibliothèques numériques dont j'ai participé à la création et au suivi : ABU et CNUM, ainsi que des exemples avancés d'utilisations de ces bibliothèques. Je décris ensuite dans le chapitre 2 un prototype d'atelier 3D de lecture dont les briques de base sont un outil de parcours dans les collections par visualisation exhaustive et un outil de lecture de fac-similés. Le chapitre 3 présente une technique de captation volumique par photogrammétrie, puis une modélisation de l'interaction pour les livres à systèmes. Le chapitre 4 est plus particulièrement consacré à des expérimentations sur les dispositifs matériels et les services logiciels à offrir pour faciliter l'interaction avec les documents 3D. J'indique en conclusion de chaque chapitre des perspectives de recherche que je souhaite pouvoir donner à ces travaux.

Pierre Cubaud <cubaud@cnam.fr>

Maitre de conférences

Centre d'études et de recherche en informatique (CEDRIC)

Conservatoire national des arts et métiers (CNAM)

292 rue saint-Martin 75003 Paris