



HAL
open science

Images et contenus : classification et sémantisation de sources documentaires visuelles.

Yasmine Azzi

► **To cite this version:**

Yasmine Azzi. Images et contenus : classification et sémantisation de sources documentaires visuelles.. MIAjournal, 2006, 2, pp.1-6. halshs-00260889

HAL Id: halshs-00260889

<https://shs.hal.science/halshs-00260889>

Submitted on 31 Mar 2008

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

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

Images et contenus : classification et sémantisation de sources documentaires visuelles.

Yasmine AZZI

Etudiante Master Recherche Sciences de l'Information et des Systèmes
Laboratoire UMR 694 MAP – Gamsau
Ecole d'architecture de Marseille 184, avenue de Luminy
13288 MARSEILLE Cedex 09
yasmine.azzi@gamsau.map.archi.fr

Pour faire une étude sur l'évolution d'un bâtiment, dans le cadre de l'architecture patrimoniale, on est souvent confrontés à une quantité de sources documentaires hétérogènes et réparties.

Pour arriver à construire des représentations qui simulent l'évolution d'un édifice, on est généralement amené à interpréter ces différentes sources et faire en sorte qu'elles soient capables de nous renseigner sur leur contenu.

L'objectif de notre travail vise justement à ajouter une couche informative à un certain type de sources. En résumé, nous cherchons à exploiter un contenant informatif bitmap par des calques vectoriels en essayant de traduire concrètement cette relation, entre une source et son contenu « architectural », par un formalisme informatique non dépendant de solutions commerciales.

Mots-clés :

Architecture patrimoniale, concepts, dictionnaire, images bitmap, masque vectoriel, SVG

1 - Introduction

Dans le contexte de l'architecture patrimoniale, L'élaboration d'un modèle architectural suppose la définition préalable d'un ensemble de concepts représentatifs de l'édifice étudié. En effet la modélisation du corpus architectural est conditionnée par la détermination d'un ensemble de concepts univoques sur lesquels vont reposer les comparaisons et les analyses du bâti. Dans ce domaine, le travail d'étude sur l'évolution d'un bâtiment présente des difficultés d'approche car on est souvent confronté à une quantité de sources documentaires hétérogènes et réparties.

Pour parvenir à construire des représentations qui simulent l'évolution d'un édifice, on est généralement amené à interpréter ces différentes sources et faire en sorte qu'elles soient capables de nous renseigner sur leur contenu.

L'objectif de notre travail vise justement à ajouter une couche informative à un certain type de sources, en l'occurrence des images bitmap. En d'autres termes, nous cherchons à exploiter un contenant informatif bitmap par des calques vectoriels.

La démarche que nous avons retenue pour la mise en oeuvre de ce travail comporte deux grandes étapes:

La première concerne la construction d'une application qui va permettre la capture des contours des concepts. Nous aborderons la question des rapports entre termes et concepts et de la localisation de ces derniers sur l'image, et nous décrirons l'implémentation d'une application SVG/JavaScript qui permet de définir les contours des concepts sur l'image traitée, et ainsi d'appliquer la couche vectorielle sur cette dernière.

La seconde étape concerne la construction d'une seconde application, développée sur la base de la première, et via laquelle on pourra faire des requêtes sur les zones vectorielles définies lors de la première étape.

On précisera ici que le travail que nous menons n'est pas rattaché à un terrain d'expérimentation particulier. Il pourrait être appliqué à n'importe quelle image bitmap représentant un ou une partie d'édifice patrimonial.

Cependant, ce travail repose sur un support théorique qui prend la forme d'un dictionnaire fixant un jeu de termes, et qui est issu d'une expérience menée par l'UMR CNRS/MCC MAP dans le but d'organiser un jeu de ressources documentaires dites à entrées terminologiques [1].

2 – construction de l'application de capture des contours

2.1. Les concepts et leur localisation

Les sources documentaires relatives à un édifice architectural sont de natures diverses et variées. Elles ont cependant en commun des termes qui permettent d'identifier les entités élémentaires qui composent un bâtiment. Le problème que posent ces termes est qu'ils sont souvent à l'image des édifices qu'ils identifient : ambigus, incomplets, ... [2].

La description des entités élémentaires doit, dans ce cas, être rattachée à des jeux de « concepts fondamentaux reconnus par un jeu de termes » [2]. Le concept devient alors la référence qui permet de fédérer les diverses interprétations et déclinaisons du modèle théorique qu'il identifie.

L'opération de localisation des concepts constitue précisément la première étape de notre travail. Il s'agit ici d'analyser les images « tests » dont on dispose afin, d'une part, de déterminer les concepts qui y sont présents, et d'autre part, de les localiser sur l'image en question, c'est-à-dire d'en définir les contours. Il faut entendre ici, contours « sémantiques » du concept, au sens de limites physiques de l'objet. Il est important de souligner ici que cette opération ne peut pas se suffire de méthodes de détection automatique car elles ne peuvent prendre en charge toutes les dimensions du problème posé. Elles sont, d'une part, inefficaces sur des ressources dont le contenu est complexe, et d'autre part, ces méthodes nous renvoient des informations déjà présentes dans la ressource, alors qu'en règle générale nous cherchons ce qui est caché ou sous-entendu derrière ce que l'on voit [3].

Pour effectuer cette étape de localisation, nous nous sommes référés à l'ouvrage de l'architecte français Jean-Marie Pérouse de Montclos « *Architecture voculaire - Principe d'analyse scientifique* ».

Dans son ouvrage, l'auteur présente un dictionnaire théorique d'un grand nombre d'objets du corpus architectural [4] et propose, pour chacun de ces objets, des illustrations sur plusieurs édifices différents. Il pose également les règles de distinction dont il se sert dans la détermination du vocabulaire qu'il recense.

On peut trouver, dans cet ouvrage, plusieurs propositions qui vont dans le sens de notre démarche. La plus importante d'entre elles est la description des édifices comme un ensemble de pièces architecturales élémentaires et qui sont combinées entre elles à travers des règles de composition que l'auteur a pris soin de préciser et de nommer.

L'ouvrage de Pérouse de Montclos s'avère par conséquent une référence précieuse pour l'étude approfondie des rapports entre les objets et leurs désignations. Ce travail est à la base de la construction à laquelle nous nous proposons de contribuer. Cette l'étape servira de « préalable » à

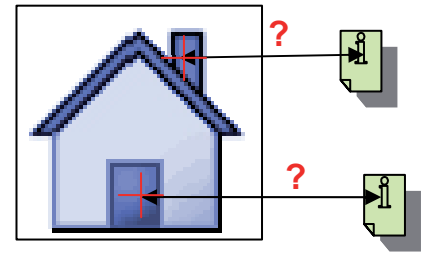


Fig. 1. Figure schématisant l'existence de plusieurs informations localisées sur une image [aut. Y. Azzi, 2006]

l'application d'un masque vectoriel sur les images ; En effet, les contours que l'on aura définis recevront une couche vectorielle qui permettra de faire des requêtes et ainsi accéder aux fiches descriptives relatives aux concepts en question.

2.2. L'application SVG/JavaScript

Cette opération a constitué la seconde étape de notre travail. Après avoir déterminé et localisé les concepts présents dans nos images, il s'agissait de construire une application qui permette d'appliquer la couche vectorielle sur les zones définies lors de l'étape précédente.

Nous avons choisi de développer cette application dans le langage SVG pour certaines raisons dont, la volonté de nous affranchir de toute solution commerciale et les avantages que ce langage présente : Instance XML, Affiché dans un navigateur, Gratuit, Format ASCII.

Une des premières contraintes rencontrées fut la gestion du zoom. Il allait faire en sorte de pouvoir récupérer des données exactes en x, y pixels quel que soit le facteur de zoom utilisé, afin de pouvoir adapter l'application à différentes tailles d'images. Pour cela, nous avons procédé à l'écriture d'un script en JavaScript et en jouant sur l'attribut « ViewBox » du noeud racine du fichier SVG.

Un autre problème s'est posé au moment de la construction des modes de capture. Pour proposer un jeu de modes de capture adaptés aux différentes possibilités, il a fallu "balayer" un certain nombre de cas. On a été amené ainsi à réfléchir, à travers des exemples, et des cas de figure, au problème des limites « physiques » des objets et à leurs limites « dessin ».

Pour construire ces modes de capture, on a d'abord listé les formes géométriques existantes dans le langage SVG, et pour chacune d'entre elles, on a indiqué les informations nécessaires pour l'obtenir [5]. Par exemple :

Un rectangle est défini par 2 points

Une courbe est définie par une liste de points

Ces modes de capture sont accessibles sous la forme d'une barre d'outils dans l'application de saisie d'informations.

Le procédé à rechercher était de parvenir, par saisie sur l'image, à obtenir les informations nécessaires et suffisantes pour repérer de façon non ambiguë chaque objet.

En raison de l'importante variation morphologique existant dans les éléments d'architecture que l'on cherche à extraire, il était donc nécessaire d'avoir des modes de capture différents pour arriver à superposer le calque vectoriel sur l'ensemble de la zone extraite.

3 – construction de l'application d'interrogation des concepts

3.1. Construction des fiches descriptives des concepts en XML

Avant toute chose, on précisera ici qu'un concept est d'abord un élément du dictionnaire évoqué en introduction et qui, à défaut d'un terrain d'expérimentation, est la base théorique sur laquelle repose notre travail. Et c'est aux éléments de ce dictionnaire que l'on cherche à rattacher un certain nombre d'informations.

Les éléments de ce dictionnaire méthodologique multilingue, développement dans le cadre duquel s'inscrit notre travail, ont vocation à identifier des types architecturaux que l'on retrouvera dans les sources bitmap sous la forme d'instances de ces types.

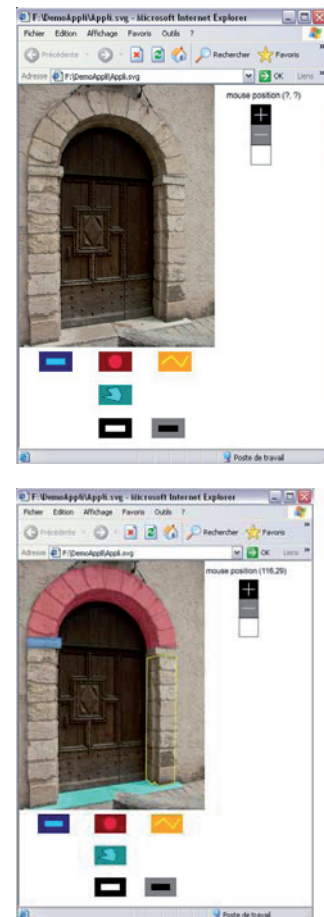


Fig. 2. Exemple de capture de quelques zones sémantiques [aut. Y. Azzi, 2006]

Le rôle de ce dictionnaire est donc bien sûr celui de tout dictionnaire, c'est-à-dire donner un sens au terme (et ici traduire aussi), mais aussi de servir plus généralement dans l'étude des types que les termes désignent. C'est pourquoi il est important de rassembler autour du terme un jeu « d'instances » qui doivent contribuer à sa définition. Dans ce cadre, il est important de disposer d'une méthode efficace pour attacher des « instances » au terme, et ces instances nous sont fournies par la réalité des objets d'architecture. Les sources bitmap à analyser sont donc des images figurant les types, images choisies pour en cerner les contours. A partir de là on comprend que la priorité pour notre travail a été de développer une méthode adaptée, c'est à dire s'inscrivant dans une démarche plus générale sur l'usage de la terminologie.

Les termes utilisés sont ceux repérant les objets physiques qui, à différentes échelles, permettent de décrire le bâti. Les termes identifiés dans ce dictionnaire sont caractérisés par un jeu de critères descriptifs.

Dans un premier temps le travail qui nous a été confié consistait à étudier de façon générique comment lier deux éléments :

- une localisation x, y associée à une forme 2D, dans une image bitmap, sous la forme d'un calque vectoriel.
- un terme du dictionnaire cité ci-dessus, formalisé par une fiche XML.

La démarche a consisté à simuler la fiche dictionnaire par un jeu d'informations type qui sont des éléments à ajouter à la fiche du dictionnaire existant pour la lier aux images illustrant le concept que la fiche définit.

Ces informations, que l'on appellera fiches descriptives, seront stockées dans des fichiers XML.

Le choix du format XML peut être justifié par deux raisons :

- c'est dans la nature même du XML de pouvoir supporter des données structurées, et de les interpréter simplement par une feuille de style XSLT, par exemple (Contrairement aux métadonnées EXIF qui ne possèdent pas la capacité de stocker ce type d'informations, ou du moins la capacité de les interpréter sans décodeur spécifique).
- Comme l'application de capture a été développée en SVG, il nous apparaissait plus « naturel » de nous diriger vers le format dont le SVG est une instance.

Ces fiches contiennent les informations suivantes :

- image source, et les caractéristiques de l'image au moment de la saisie pour détecter par exemple des re-dimensionnements intempestifs.
- le nom du concept capturé
- Sa définition
- le mode de capture

Et selon le mode de capture :

- les données quantitatives adéquates

3.2. La construction de l'application PHP/SVG/JavaScript

Cette application est celle qui doit nous permettre de faire correspondre, à chaque zone détectée lors de la première étape, sa fiche descriptive au format XML. Pour cela, il faudra intégrer des scripts PHP à la première application développée, afin d'établir le lien entre le masque vectoriel

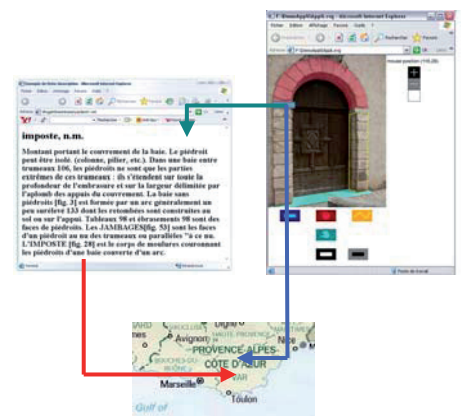


Fig. 3. Illustration du principe du triangle concept/instance/lieu [aut. Y. Azzi, 2006]

SVG et le fichier XML. Ce dernier, on le rappelle, est interprété au moyen de feuilles de style XSLT.

Mais au-delà du dispositif technique nécessaire à la mise en place de cette application, il y a aussi un aspect méthodologique que l'on ne doit pas négliger.

En effet, l'objectif recherché consiste à réaliser une forme d'intégration apte à saisir les ressources documentaires sous des angles différents mais reliés entre eux.

De façon schématique, les ressources documentaires seraient intégrées dans un triangle liant « le concept », qui est un terme du dictionnaire, « le lieu », qui est la localisation toponymique de l'objet représentant le concept, et « l'instance », qui serait un jeu de ressources documentaires. La ressource bitmap peut localiser un des trois sommets du triangle, celui du « concept », (une ressource particulière utilisée comme définition visuelle du concept), et permettrait alors d'interroger les deux autres sommets. Mais la ressource bitmap est aussi ressource documentaire, simple illustration du concept, permettant d'établir pour chaque concept un catalogue d'instances.

4 - Conclusion

L'hypothèse avancée, sur laquelle nous avons travaillé, consiste à dire que si l'on s'appuie sur les connaissances architecturales il doit être possible d'aider à la capture de zones sémantiques en fournissant à l'utilisateur l'équivalent graphique de masques de saisie. Ces masques de saisie graphiques seraient une aide à la détermination de contours qui utiliserait des formes prototypiques simples.

Le travail mené a démontré que cette voie est possible, et en a souligné les difficultés techniques (notamment compte tenu du choix de plateforme excluant toute solution commerciale) mais aussi l'apport potentiel. Le travail mené n'est que la première partie du processus proposé, mais elle doit permettre de continuer en étant assurés de la crédibilité des solutions qui ont été choisies.

A partir de là un couplage de cette approche avec celles de travaux sur la détection automatique de contours est peut-être possible.

L'application que nous avons commencé à développer est certes, modeste, mais ce premier travail a permis d'affiner les approches et d'expérimenter des solutions dont on perçoit mieux la portée, la pertinence ou les limites.

L'application est, bien sûr, encore à un stade expérimental. Pour qu'elle devienne opérationnelle, plusieurs problèmes doivent être préalablement résolus.

En dehors des difficultés techniques rencontrées (lire un fichier XML à partir d'une zone SVG au travers d'un script PHP), et des problèmes d'ergonomie qui sont présents mais que nous ne pouvons traiter dans le cadre du master, il reste aussi tous les aspects de navigation auxquels il faudra réfléchir, comme :

- l'aspect de catégorisation des jeux d'images : Il s'agira dans ce cas d'organiser les images dans des catégories en faisant appel aux *a priori* de l'utilisateur (Par exemple, si l'utilisateur cherche un pilastre, une colonne, il pourra le faire dans la catégorie « organes de support »...)
- le côté utilisateur de l'application : Réfléchir à la façon de mettre des points d'accroche qui signifient que l'objet sélectionné est actif et qui

laissent également apparaître les limites de cet objet. Il faudrait peut-être aussi penser à faire varier le symbole selon la catégorie de l'objet.

- L'aspect de changement d'échelle : Réfléchir au procédé qui permettrait de passer de la plus petite entité architecturale à l'échelle de l'édifice et, pourquoi pas, à l'échelle urbaine voire territoriale ?

Mais ce travail a montré aussi la difficulté d'adaptation de méthodes et d'outils aux contraintes particulières du patrimoine bâti. La réponse apporté a cependant montré le gain potentiel de cette adaptation : servir réellement l'étude du bâti.

Références bibliographiques

- [1.] UMR CNRS/MCC 694 MAP, « *Programme STRABON WorkPackage 6 Initiative EUMEDIS, Interfaces de navigation 2D/3D dans les contenus* », compte rendu de travaux, Mai 2006
- [2.] Blaise (J.Y) et Dudek (I.), « *des termes au concept, le cas du vocabulaire architectural* », communication présentée aux journées scientifiques LTT, Bruxelles, 2005
- [3.] Zlatoff (N) et al., « *Image Understanding Using Domain Knowledge* », in *Proceedings of Recherche d'information Assistée par Ordinateur (RIAO)*, Avignon, 2004, pp. 277-290
- [4.] Pérouse De Montclos (J.M.), « *Architecture vocabulaire - Principe d'analyse scientifique* », Imprimerie Nationale, 1972-88.
- [5.] J. Frost, S. Goessner, M. Hirtzler, « *Apprivoiser SVG* » (E-book), <http://www.learnsvg.com/>, 2002 (consulté le 08 mars 2006)