

La maquette numérique comme support pour la recherche visuelle d'informations patrimoniales : définition d'une approche pour la sémantisation de sources iconographiques par référencement spatial.

Chawee Busayarat

▶ **To cite this version:**

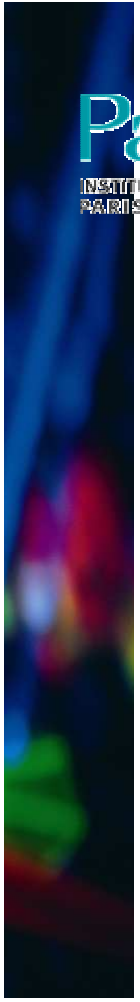
Chawee Busayarat. La maquette numérique comme support pour la recherche visuelle d'informations patrimoniales : définition d'une approche pour la sémantisation de sources iconographiques par référencement spatial.. Multimédia [cs.MM]. Arts et Métiers ParisTech, 2010. Français. NNT : 2010ENAM0060 . pastel-00557514

HAL Id: pastel-00557514
https://pastel.hal.science/pastel-00557514

Submitted on 19 Jan 2011

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.



2010-ENAM-0060

École doctorale n°432 : SMI

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

l'École Nationale Supérieure d'Arts et Métiers

Spécialité Conception

présentée et soutenue publiquement par

Chawee BUSAYARAT

le 16 décembre 2010

La maquette numérique comme support pour la recherche visuelle d'informations patrimoniales

Définition d'une approche pour la sémantisation

de sources iconographiques par référencement spatial.

Directeur de thèse : **Michel FLORENZANO**
Co-Directeur de thèse : **Philippe VERON**

Co-Encadrement de thèse : **Livio DE LUCA**

Jury

M. Jean-Claude LEON, Professeur des universités, INPG	Président du jury
Mme. Michela SPAGNUOLO, CNR Senior Research, IMATI, CNR Italie	Rapporteur
M. Marc PIERROT-DESEILLIGNY, Directeur de recherche, IGN	Rapporteur
M. Michel FLORENZANO, Directeur de recherche CNRS, MAP	Directeur
M. Philippe VERON, Professeur des universités, LSIS, Arts et Métiers ParisTech	Co-directeur
M. Livio DE LUCA, Ingénieur de recherche CNRS , MAP	Encadrant

Arts et Métiers ParisTech - Centre de Aix-en-Provence
FRE 3315 MAP / UMR CNRS 6168 LSIS

T H È S E

Remerciements

En préambule à cette thèse, je souhaitais adresser mes remerciements les plus sincères aux personnes qui m'ont apportées leur aide et qui ont contribué à l'élaboration de ce mémoire ainsi qu'à la réussite de ces formidables années universitaire.

Je tiens à remercier sincèrement:

Michel Florenzano, en tant que directeur de thèse, pour son accueil bienveillant et ses conseils avisés, et cela malgré son emploi du temps chargé. Philippe Véron, co-directeur de thèse, pour son aide constante et minutieuse tout au long du développement et de la rédaction de la thèse. Michel Berthelot, pour m'avoir offert la possibilité de participer à de nombreuses expériences au cours de ces années.

Mes remerciements s'adressent également à Livio de Luca, l'encadrant de thèse, qui s'est toujours montré à l'écoute et très disponible tout au long de la réalisation de ce travail, que ce soit pour l'inspiration, l'aide et le temps qu'il a bien voulu me consacrer et sans qui ce mémoire n'aurait jamais vu le jour. Jean-Yves Blaise et Iwona Dudek, véritables sources d'inspiration de mon travail. Menehould Chataignier, secrétaire général du laboratoire MAP-GAMSAU, pour son amitié, générosité et la grande patience.

J'exprime ma gratitude à tous les consultants et chercheurs rencontrés lors des recherches effectuées et qui ont accepté de répondre à mes questions avec gentillesse.

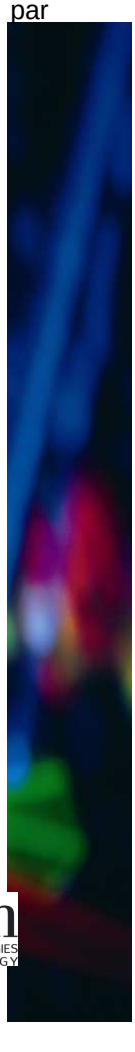
Je n'oublie pas mes parents pour leur contribution, leur soutien et leur patience. Enfin, j'adresse mes plus sincères remerciements à tous mes proches et amis, qui m'ont toujours soutenu et encouragé au cours de la réalisation de ce travail.

Merci à tous et à toutes.

La maquette numérique comme support pour la recherche visuelle d'informations patrimoniales : définition d'une approche pour la sémantisation de sources iconographiques par référencement spatial

RESUME : La collecte, l'organisation et la recherche de sources iconographiques est certainement un problème majeur au sein des communautés qui s'intéressent à l'étude, la conservation et la valorisation du patrimoine architectural. L'organisation et la structuration de ces ressources sont alors un problème essentiel si l'on s'inscrit dans l'idée de concevoir les archives de demain. Aujourd'hui, la méthode la plus couramment utilisée, pour classer et rechercher des contenus iconographiques au sein d'une base de données, est la recherche par mots-clés. En particulier, dans le domaine de la documentation architecturale, au vu du degré de complexité (en termes de richesse d'information) et d'hétérogénéité (en termes de variété de supports et de techniques d'exécution) des sources iconographiques (photos, dessins, peintures, etc.), les solutions existantes d'annotation sémantique d'images s'avèrent inefficaces. Cette thèse considère le modèle 3D interactif comme un moyen d'accès privilégié à l'information patrimoniale. Dans ce sens, ce travail se concentre sur l'exploitation de la représentation de la morphologie d'un édifice comme source principale pour distribuer / propager les attributs sémantiques sur l'ensemble des sources iconographiques (le représentant en deux dimensions) qui se trouvent dans une condition de cohérence géométrique/spatiale avec son modèle tridimensionnel. Un modèle de description sémantique (structurant l'ensemble de termes pouvant décrire la morphologie d'un édifice) est alors utilisé comme dénominateur commun permettant d'établir des relations entre la représentation 3D complète des formes architecturales qui composent l'édifice et l'ensemble des sources iconographiques (collections d'images 2D segmentées) correspondantes. Ce travail abouti à la définition et au développement d'un système d'informations permettant de rechercher visuellement, au sein d'un modèle 3D, les sources iconographiques corrélées en fonction de critères spatiaux, morphologiques et sémantiques. Il s'agit d'une application Web utilisable par des spécialistes du domaine de la documentation architecturale comme par le grand public.

Mots clés : Patrimoine architectural, iconographie, recherche d'informations, annotation sémantique d'images, référencement spatial



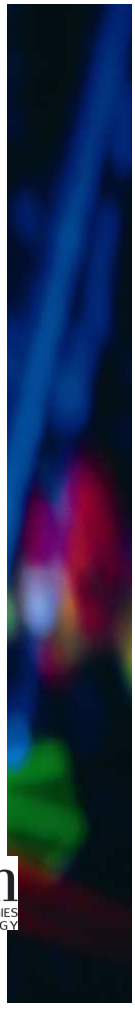
ARTS ET MÉTIERS ParisTech

ParisTech INSTITUT DES SCIENCES ET TECHNOLOGIES PARIS INSTITUTE OF TECHNOLOGY

The digital model as a support for the visual search for heritage information: an approach for the semantic enrichment of iconographic sources by spatial referencing

ABSTRACT: The collection, the organization and the search for iconographic sources are certainly a main issue within the communities that are interested in the study, the conservation and the valorization of the architectural heritage. The organization and the structuring of these resources are an important problem if for researcher involved into the design of tomorrow's digital archives. Today, the most used method to classify and search iconographic contents within a database is the key words-based search. Nevertheless, the association between the research topics and the contents of images is not entirely developed. In particular, in the field of architectural documentation, according to the degree of complexity (in terms of wealth of information) and the heterogeneity (in terms of variety of supports and techniques of execution) of iconographic sources (photographs, drawings, paintings, etc), existing solutions for semantic annotation of images are ineffective. This thesis starts from the assumption that the interactive 3D model interactive can be an important support to access at the heritage information. In this direction, this work focuses on the use of the building's morphology representation as primary source to distribute/propagate semantic attributes on the whole of the iconographic sources (the 2D representations), which are under a condition of geometrical/spatial coherence with its three-dimensional model. A semantic description model (structuring the whole of terms which can describe the morphology of a building) is then used as common denominator in order to establish relations between the 3D representation of the architectural shapes that compose the building and the corresponding iconographic sources (collection of segmented 2D images). This work led to the definition and the development of an information system allowing visually search, within a 3D model, iconographic sources according to spatial, morphological and semantic criteria. This system is a Web application is usable by specialists in the field of architectural documentation as by the general public.

Keywords: Architectural heritage, iconography, information retrieval, semantic annotation, spatial referencing



ARTS ET MÉTIERS ParisTech

ParisTech INSTITUT DES SCIENCES ET TECHNOLOGIES PARIS INSTITUTE OF TECHNOLOGY

Table de matières

1. INTRODUCTION 11

1.1. CONTEXTE 11
1.2. OBJECTIF 12
1.3. PORTEE, LIMITES, APPORTS PRINCIPAUX 13
1.4. STRUCTURE DU DOCUMENT 14

2. LES SOURCES ICONOGRAPHIQUES POUR LA DOCUMENTATION ET L'ETUDE DU PATRIMOINE BATI 15

2.1. CLASSIFICATION PAR L'ETAT DE L'EDIFICE 15
2.1.1. L'état actuel 16
2.1.2. L'état historique 17
2.1.3. L'image hypothétique 18
2.2. LE DEGRE DE CERTITUDE 20
2.3. LES TECHNIQUES D'ELABORATION 21
2.3.1. Photographies 21
2.3.2. Dessin 23
2.3.3. Document métrique 23
2.4. LES TYPES DE PROJECTION 24
2.4.1. Projection perspective 24
2.4.2. projection parallèle 29
2.4.3. Projection panoramique 32

3. METHODES D'ANNOTATION SEMANTIQUE D'IMAGES 35

3.1. METHODES MANUELLES 35
3.2. METHODES AUTOMATIQUES 39
3.3. METHODES SEMI-AUTOMATIQUES 42

4. UNE APPROCHE POUR LA CARACTERISATION SEMANTIQUE 2D/3D D'EDIFICES PATRIMONIAUX 45

4.1. APPROCHE PROPOSE 47
4.2. PROBLEMES A TRAITER 49
4.2.1. Acquisition de données 3D et reconstruction géométrique basés sur la réalité (chapitre 5) 49
4.2.2. Caractérisation sémantique de la représentation 3D (chapitre 5) 50
4.2.3. Référencement spatial de sources iconographiques (chapitre 6) 50
4.2.4. Indexation du contenu des sources iconographiques (chapitre 6) 50
4.2.5. Recherche visuelle de sources iconographiques basée sur critères spatiaux, morphologiques et sémantiques (chapitre 7) 51
4.3. DEVELOPPEMENT INFORMATIQUE 52
4.3.1. Présentation du projet NUBES 52
4.3.2. Présentation de l'environnement de développement 52
4.3.3. L'architecture du système 53

5. RECONSTRUCTION 3D ET STRUCTURATION SEMANTIQUE D'EDIFICES PATRIMONIAUX 55

5.1. ACQUISITION INFORMATIONS SPATIALES 55
5.1.1. La photogrammétrie 56
5.1.2. L'acquisition par scanner laser 3D 57
5.2. LA RECONSTRUCTION GEOMETRIQUE DE LA MORPHOLOGIE 60
5.2.1. Reconstruction 3D à partir de nuage de points 60

5.2.2. Reconstruction 3D basée sur image 63
5.3. CARACTERISATION SEMANTIQUE DU MODELE 3D 64
5.3.1. Travaux relatifs 65
5.3.2. Solution retenue 67

6. REFERENCEMENT SPATIAL ET SEMANTISATION DE SOURCES ICONOGRAPHIQUES 71

6.1. METHODES DE REFERENCEMENT SPATIAL 71
6.1.1. Méthodes manuelles 72
6.1.2. Méthodes semi-automatiques 73
6.1.3. Méthodes automatiques 74
6.2. ANALYSE DES SOLUTIONS EXISTANTES 79
6.3. SOLUTION RETENUE 80
6.3.1. Téléchargement sur serveur des sources iconographiques 81
6.3.1. Superposition manuelle de sources à projection orthographique et pseudo-perspective 82
6.3.2. Résection spatiale de sources à projection perspective 83
6.3.3. Récupération directe des paramètres de photographies orientées 86
6.4. ANNOTATION SEMANTIQUE DES IMAGES PAR PROJECTION 86
6.4.1. Présentation de l'environnement du développement 87
6.4.2. Développement de la procédure de sémantisation d'images 88

7. UN SYSTEME POUR LA RECHERCHE VISUELLE DE SOURCES ICONOGRAPHIQUES 95

7.1. ARCHITECTURE DU SYSTEME 96
7.1.1. Base de données 97
7.1.2. L'interface WEB du système de la consultation 99
7.1.3. Fonctions pour la recherche d'images 100
7.2. RECHERCHE D'IMAGES BASEE SUR CRITERES SPATIAUX 101
7.3. RECHERCHE D'IMAGES BASEE SUR CRITERES MORPHOLOGIQUES 104
7.4. RECHERCHE D'IMAGES BASEE SUR CRITERES SEMANTIQUES 106
7.5. RECUPERATION DE POINT DE VUE IMAGE 107

8. CONCLUSIONS ET PERSPECTIVES 109

8.1. LIMITES 110
8.2. PERSPECTIVES DE RECHERCHE 112

BIBLIOGRAPHIE 113

LISTE DE PUBLICATIONS 125

1. Introduction

1.1. Contexte

Dans le domaine de la documentation architecturale, le développement de descriptions qualitatives d'édifices est un domaine de recherche vaste. Au vu du progrès des technologies de numérisation 3D, il apparaît essentiel aujourd'hui d'étudier les conditions au travers desquelles les informations quantitatives, extraites du relevé de la morphologie d'un édifice, et les informations qualitatives, produites par l'interprétation des données acquises ou par l'analyse de sources iconographie puissent être consultés, analysées et structurées au sein d'une plateforme unique.

Parmi les nombreuses sources documentaires utilisées dans ce domaine, il est certain que l'iconographie (photo, dessin, peinture, croquis et aussi l'image de synthèse créée par l'ordinateur, etc..) occupe une place centrale. Il est assez fréquent que sur un édifice historique, on trouve de l'iconographie en grande abondance, quelque soit la période historique de sa construction ou de sa restauration. Des images des édifices historiques sont parfois utilisées comme arrière-plan des scènes picturales ; les édifices sont également systématiquement photographiés par des touristes lors des visites ; des dossiers beaucoup plus approfondis (photos, dessins, relevés, etc..) ont été produits (et continuent à être produits) par les chercheurs qui s'intéressent à leur étude.

L'iconographie a une valeur énorme pour les études architecturales et archéologiques car ses différentes sources sont un véritable témoignage de l'état d'un édifice à un moment de son histoire. Qu'il s'agisse de son état actuel ou d'un état antérieur. Des chercheurs collectent les sources pour comprendre l'histoire d'un édifice, comprendre les transformations physiques ou fonctionnelles, ou encore la composition des matériaux et son état de conservation.

L'évolution technologique dans le domaine de la photographie permet aujourd'hui de produire des images d'une manière plus simple et plus rapide qu'auparavant. Le réseaux et Internet permettent également de partager et de gérer ces sources plus facilement. Ceci permet de retrouver une énorme quantité d'images directement sur le Web. Pour récupérer l'information spécifique d'une grande source telle que l'Internet, une manière pratique de classifier, documenter et rechercher l'information architecturale est nécessaire. L'efficacité de l'accès, la rapidité de la recherche et la capacité d'assortir la plupart des informations importantes sont aussi essentielles.

De nos jours, la méthode la plus pratique pour la recherche d'images architecturales se base sur la recherche par mots-clefs de l'objet présent dans l'image (par exemple chapiteau, porte, base, etc.). En d'autres termes, ce type de recherche

est basé sur les termes utilisés pour décrire les objets qui sont présent dans l'image. Ce type de recherche nécessite de conduire une annotation sémantique des images. Ce domaine de recherche suscite aujourd'hui de plus en plus d'attention.

Dans les outils communs pour la recherche d'images, les recherches se basent sur la détection du texte pré-associé aux images (sous forme de nom, description et mots-clés) par un étiquetage du contenu sémantique (également connu sous le nom d'annotation sémantique d'image). L'objectif d'un procédé d'annotation est de créer des correspondances entre :

- les objets (ou des segments d'objet) ;
- des étiquettes de mots-clés, des textes ou des concepts appropriées qui reflètent son contenu visuel.

Des images annotées sémantiquement peuvent habituellement être trouvées en utilisant une recherche par mots-clés, alors qu'il est extrêmement difficile de retrouver des images non annotées au sein d'une grande base de données. L'utilisation d'images annotées sémantiquement d'une manière correcte et détaillée permettra d'exploiter plus facilement les moteurs d'indexation et de récupération d'informations visuelles sur le Web.

Néanmoins, l'association d'attributs sémantiques aux images n'est pas encore suffisamment développée aujourd'hui, surtout dans le domaine de la documentation architecturale. Les méthodes conventionnelles pour l'annotation sémantique d'images telles que l'annotation manuelle ou la reconnaissance automatique d'objet ne peuvent pas être suffisamment efficaces dans le domaine de l'architecture historique (voir chapitre 3). D'ailleurs, même si l'on peut conduire une annotation sémantique d'une image concernant un artefact architecturale ou archéologique, rien ne nous permet d'associer à cette image des informations importantes comme sa position dans l'espace et/ou sa morphologie. En manque de ces types d'informations, les outils de recherche de sources documentaires reliées à un édifice historique ou à un site archéologique s'avèrent donc inefficaces.

1.2. Objectif

L'objectif de ce travail de thèse est de développer une méthodologie et un système permettant la collecte, la structuration et la recherche des sources iconographiques servant à l'étude des bâtiments historiques, basé sur le référencement spatialisé et sur l'annotation sémantique d'images.

Ce système doit permettre de fournir des informations importantes sur les objets ou les parties d'objets qui apparaissent sur les images :

- il doit être capable de visualiser leur position dans l'ensemble architectural, leur morphologie et leur composition en parties. Cette information est essentielle pour l'étude et la compréhension du bâti patrimonial ;

- il doit permettre d'associer les attributs sémantiques (avec un niveau de précision important) permettant de décrire les formes et d'observer dans les images en relation à un univers de connaissances formalisées, d'où la nécessité de baser l'annotation sur des ontologies qui permettent une capitalisation des termes de description ainsi que la relation entre ces termes ;

- il doit être capable de gérer des représentations complexes d'édifices et d'assurer la gestion de sources iconographiques relatives aux différentes échelles de sa morphologie ;

- il doit permettre des mises à jour faciles (ajout, modification d'iconographie) et il doit permettre également de changer la structure de description de la morphologie de l'édifice en fonction du type d'analyse que les spécialistes de la conservation veulent conduire. Ceci est essentielle car dans le domaine de la documentation architecturale, l'état de connaissance relatif à l'histoire d'un édifice peut varier facilement à chaque découverte (ou nouvelle interprétation) d'une source iconographique.

Au-delà de la sémantisation de sources iconographiques l'objectif de ce travail est également le développement d'un système de recherche d'images d'édifices historiques basé sur toute la richesse informationnelle produite par la méthodologie ci-dessous décrite.

1.3. Portée, limites, apports principaux

Ce sujet se situe dans le domaine de la sémantisation de sources iconographiques relatives à l'étude et à la documentation du patrimoine bâti. Compte tenu de l'ampleur de la problématique, nous avons concentré notre attention principalement sur les images photographiques. Tout particulièrement, notre intérêt porte sur les images représentant l'état actuel du bâtiment.

Notre problématique comprend différentes préoccupations car elle s'intéresse aux relations entre des sources bidimensionnelles, des représentations tridimensionnelles et des modèles de description sémantiques. Dans ce sens, le premier apport de cette thèse est d'avoir identifié un processus innovant pour la sémantisation de sources iconographiques mettant en relation ces trois dimensions. Notre approche se base sur un processus structuré en phases. Chaque phase relève de problèmes spécifiques qui seront traités à travers la définition de méthodes et le développement d'outils informatiques. Par conséquence, ce document est organisé à l'image du processus identifié et présente pour chaque phase de ce processus, l'étude des l'état de l'art relié, la définition de méthodes et les apports produits.

1.4. Structure du document

Ce document est structuré de la façon suivante. Dans les deux premiers chapitres on s'intéressera à l'étude bibliographique relative aux sources visuelles d'architectures historiques et aux méthodes actuelles d'annotation sémantique. Le deuxième chapitre se focalisera plus particulièrement sur les sources iconographiques bidimensionnelles d'architectures historiques en donnant une attention particulière aux différents formats et types de sources à prendre en compte. Le chapitre 3 décrira les approches actuelles de la sémantisation d'image en s'appuyant sur les travaux de recherche s'intéressant plus particulièrement au domaine de la documentation d'artefact patrimoniaux, mais tout en étendant l'étude à d'autres domaines.

A partir d'analyse des travaux décrits au sein du chapitre 3, le chapitre 4 présentera l'approche de sémantisation d'images développée. Cette approche est particulièrement adaptée aux problèmes de la documentation architecturale. Ce chapitre se terminera avec la mise en évidence des problèmes relatifs aux différentes phases de l'approche et qui seront abordés de façon spécifique dans les chapitres suivants.

Les chapitres 5 et 6 approfondiront les problèmes identifiés dans la définition de l'approche générale. Chaque chapitre est divisé en deux parties. La première partie décrit les technologies et les solutions existantes à ce jour. La deuxième partie analyse les travaux présentés et filtre les aspects provenant de ces contributions afin de les intégrer à l'intérieur d'une ou des solutions retenues. Le chapitre 5 traite du problème de traitement des données à référence spatiales. En abordant les méthodes de relevé de l'état actuel des édifices, ce chapitre se concentre ensuite sur les techniques de génération d'un modèle 3D sémantiquement enrichi. Le chapitre 6 est centré sur le problème du référencement spatial d'images en fonction de plusieurs échelles de représentation ainsi qu'en relation à plusieurs technologies. Ce chapitre se termine sur la définition d'une méthode permettant d'exploiter les résultats du référencement d'images sur la représentation tridimensionnelle d'un édifice afin d'exploiter les relations projectives 2D/3D pour la sémantisation d'images.

Le chapitre 7 décrira le développement d'un outil pour la recherche visuelle de sources iconographiques développé à partir de notre approche. Ce chapitre est divisé en deux parties. La première partie présente un outil pour la segmentation et l'annotation sémantique d'images orientées sur une représentation 3D d'un édifice ; la deuxième partie présente un système de consultation sur Internet permettant la recherche visuelle de sources iconographiques basé sur une navigation 3D temps réel au sein de la représentation de la morphologie de l'édifice.

Le chapitre 8 présentera la conclusion, les limites du travail réalisé et les perspectives de recherche

2. Les sources iconographiques pour la documentation et l'étude du patrimoine bâti

L'iconographie est l'étude de l'identification, la description et l'interprétation du contenu des images relatives à l'histoire de l'art. Dans le domaine spécifique du patrimoine architectural, ce terme est utilisé pour indiquer l'ensemble des représentations relatives au même artefact [Kleinbauer W.E. et al., 1982].

Dans le domaine du patrimoine architectural, il est essentiel d'étudier les conditions du bâtiment (présentes, antérieures et futures) au travers d'informations produites par l'interprétation et l'analyse de sources iconographiques. Une des sources iconographiques plus importantes pour l'étude du bâti est certainement la source visuelle bidimensionnelle. Ces informations ont beaucoup d'importance pour l'étude architecturale et archéologique du patrimoine parce qu'elles peuvent facilement décrire (et d'une manière particulièrement efficace) l'état d'un bâtiment à un instant T de son histoire (que ce soit l'état actuel ou un état antérieur). Les sources iconographiques disponibles à ce jour pour les communautés de chercheurs ou de professionnels s'intéressant à un édifice, ont été collectées pour la compréhension de l'histoire, des changements physiques ou fonctionnels du bâtiment, ou encore des techniques employées pour ses restaurations. Ces informations peuvent être également utilisées et référencées comme données de base pour la conservation et/ou les rénovations futures. Les sources iconographiques relatives au patrimoine architectural sont de différent format, élaborées à partir de différentes techniques et provenant de différentes procédures de conservation et de duplication.

Dans ce chapitre, nous nous concentrons sur une classification des sources iconographiques. Cette étude est essentielle pour une exploitation correcte et appropriée des informations visuelles que l'on cherche à enrichir afin de les organiser au sein d'une base de données.

2.1. Classification par l'état de l'édifice

L'étude des états historiques d'un édifice est généralement basée sur la collecte et l'analyse des sources documentaires. L'iconographie est utilisée pour des buts différents. C'est une source d'information géométrique. C'est aussi une preuve de l'état d'un édifice à un moment précis de son histoire.

2.1.1. L'état actuel

L'image (ou les images) de l'état actuel d'un édifice est l'information la plus facile à retrouver. Ce type de source est présente (en grande quantité) dans les bases de données d'images sur internet. L'état actuel d'un édifice peut être représenté dans des photographies, des peintures, ou des dessins, etc. C'est une représentation de l'édifice dont l'état ne diffère pas (sauf en cas de transformations récente) au condition actuelle d'édifice réel.

La définition d'état actuel diffère par rapport de chaque bâtiment. Il s'agit d'un bâtiment qui n'a pas changé de condition (interne et externe) pendant une certaine période. Toutes les images de ce bâtiment prises ou élaborées lors de cette période seront considérées comme images de son état actuel. Mais si il y a une modification permanente d'état du bâtiment, à partir du moment de changement, toutes les images prises ou élaborées auparavant sont par conséquent classées comme des sources historiques. Pour donner un exemple, l'état actuel du palais du Louvre (aujourd'hui musée du Louvre) peut se compter depuis la construction de la pyramide en verre en 1983 (voir figure 1).



Figure 1 L'état actuel du palais du Louvre contient des bâtiments du 17e siècle et une construction moderne

Aujourd'hui la diffusion d'appareils photo numériques et Internet contribuent grandement à la collecte et à l'archivage de ce genre d'images. Ces images sont les informations qui permettent l'interactivité majeure entre le grand public et les bases de données d'images. Qu'il s'agisse de l'augmentation de la richesse des informations associées aux images ou de la recherche et de l'accès à l'information, il y a beaucoup de travaux de recherche qui s'intéressent à ce domaine.

La règle 3x3 [Waldhäusl P., Ogleby C., 1994] a été proposée dans le cadre du développement du système APIS (Architectural Photogrammetry Information System) [Herbig U. et al.,1998.] pour la préservation d'informations sur le patrimoine culturel. L'objectif principal est de créer un pont entre les amateurs et les professionnels de la documentation photogrammétrique (voir figure 2). Ce système a besoin d'une

administration des données pour savoir quels bâtiments sont documentés et où les photographies ont été prises.

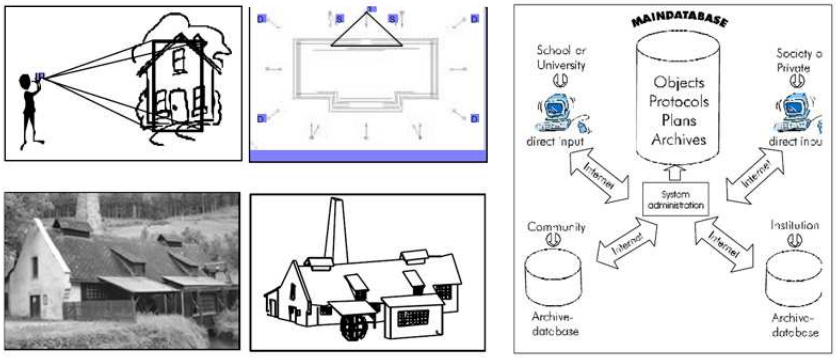


Figure 2 Documentation photogrammétrique du système APIS [Herbig U. et al.,1998.]

Le *CIPA - Test Karlsplatz, Viena* [Almagro A. et al., 1996] a montré que la restitution à partir d'appareils photo de petit format et non-métriques (typiquement les appareils photos des touristes) apporte des résultats avec une précision suffisante pour la plupart des besoins de la documentation architecturale, ou au moins pour conserver une information permettant la reconstruction physique d'un édifice (suite à un événement brusque, comme un incendie par exemple). Les instituts participant à ce projet ont augmenté leurs connaissances sur le traitement des photos prises par des caméras d'amateurs. Leur savoir-faire peut être utilisé pour la création d'une collection d'images de documentation architecturale servant de support pour la préservation du patrimoine.

2.1.2. L'état historique

Les états passés du bâtiment sont caractérisés par des propriétés spatiales et temporelles incertaines parce qu'ils sont connus grâce à des sources iconographiques et documentaires très variées, parfois imprécises et de fiabilité incertaine. On parle de sources iconographiques des états historiques des bâtiments quand ces sources représentent un état de l'édifice différent de son état actuel. Ces sources peuvent être basées sur quasiment toutes les techniques d'élaboration (depuis les plus anciennes aux plus actuelles).

Le niveau de précision de la représentation contenue dans ces types de sources diffère logiquement en fonction du type de support et/ou de technique de génération ou d'exécution. Par exemple des photos anciennes sont parmi les représentations les plus précises que l'on peut avoir sur l'état historique d'un édifice. Tandis que le dessin architectural (plan, coup, élévation) est particulièrement tributaire de la méthode de relevé qui est à la base de la restitution graphique (manuel, topographique, etc.). Les dessins à main levée peuvent logiquement contenir plus d'erreurs, d'approximations et d'incohérences avec la réalité. L'époque dans laquelle l'image a été créée peut également conditionner l'exactitude et la cohérence de l'image qui représente un état du bâtiment. Pour donner un exemple, les techniques de représentation utilisées au moyen âge diffèrent bien, à la fois en termes de projection et en termes de rigueur d'exécution, des techniques qui se sont développées à la

renaissance : l'introduction de la perspective a en effet permis de dessiner (des fois même à l'aide de machines) des représentations des espaces architecturaux relativement cohérentes avec la réalité.

La pratique de représention de l'architecture passée n'est pas une démarche contemporaine. Ses origines remontent à l'époque romaine. Vitruve au Ier siècle av. J.-C. ou Palladio au XVIe siècle ont décrit l'architecture dans deux fameux traités [Palladio A., 1992; Perrault C., 1684]. Entre le 1541 et le 1549, Palladio se déplaçait des nombreuses fois à Rome pour effectuer des relevés des bâtiments romains. Aujourd'hui on peut lire des traces de ses raisonnements dans ses esquisses (figure 3). Il y a 500 ans Palladio commençait à interpréter les fouilles et à représenter l'architecture d'une époque précédente à la sienne dans le but de la comprendre et de la réutiliser en clé moderne dans ses bâtiments. Ses esquisses sont une première opération de représentation de l'architecture.

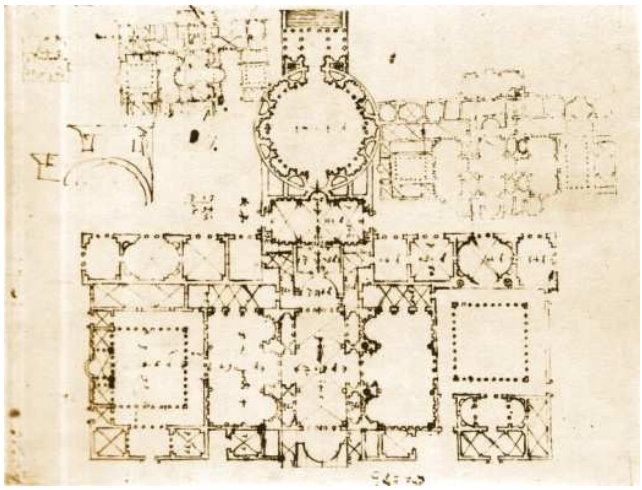


Figure 3 Andrea Palladio, Thermes d' Agrippa et Panthéon, RIBA, Londres, Vol. VII, fol. 2.

2.1.3. L'image hypothétique

L'image d'un état hypothétique d'un édifice est une représentation normalement élaborée par un chercheur (archéologue, architecte, historiens, etc.) s'appuyant sur des preuves, des traces et d'autres sources documentaires et se basant sur un rayonnement construit sur des méthodes scientifiques. Les représentations hypothétiques se bases le plus fréquemment sur des dessins conventionnels (plans, élévations et coupes) ou sur des vues d'ensemble (perspectives, axonométries). Même si ces sources représentent un état du bâtiment, cela ne signifie pas qu'il s'agisse du véritable état passé du bâtiment. La fiabilité d'une représentation hypothétique est contrainte à la dimension subjective de l'auteur de la source ; par ailleurs différentes sources iconographiques hypothétiques (parfois contradictoires) sont formulées par différents chercheurs à différents périodes.

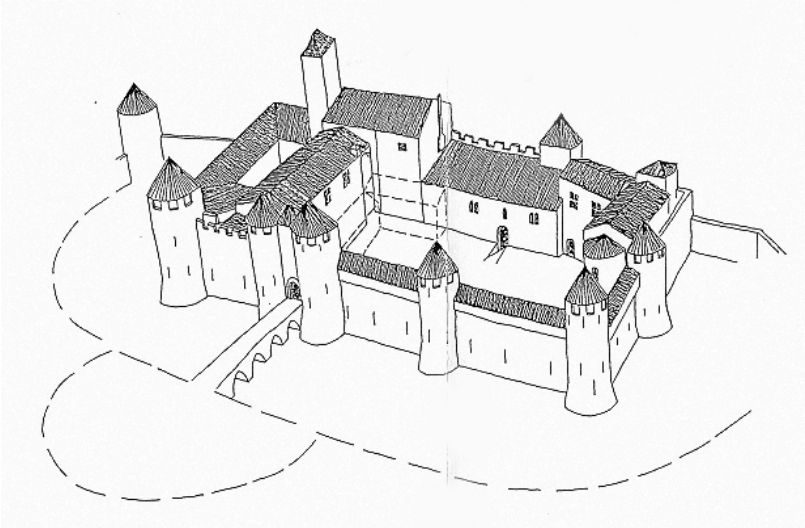


Figure 4 L'état hypothétique du château Comtal de Carcassonne au milieu du XIV^{e} : reconstitution et représentation de [Guyonnet F., 2000].

Le passé étant inaccessible, si les sources offrent des informations incomplètes sur la morphologie d'un artefact, les experts sont amenés à interpréter la forme et à formuler des hypothèses parfois très variées sur les états passés. De plus, la comparaison des sources peut induire à des raisonnements contradictoires et à des hypothèses qui précisent le sens d'incertitude sur la *forme*, la *position* et la *datation* des artefacts. Les hypothèses sont un cas particulier d'interprétation des incertitudes appliquées à l'architecture (voir section 2.2). Plus précisément, les historiens et les archéologues s'appuient sur différents raisonnements pour formuler leurs hypothèses [Strothotte T. et al., 1999]:

- *déduction*. Ce type d'interprétation se base sur la détermination des formes architecturales à partir d'éléments géométriques incomplets mais pour lesquelles la géométrie est reconnaissable (fouilles archéologiques ou sources iconographiques).

- *analogie*. L'historique d'un lieu peut être supposé à partir de la documentation sur les techniques de l'époque. Le processus d'analogie s'appuie sur les connaissances architecturales telles que les traités classiques d'architecture, les relations de composition, la connaissance de constructions similaires pendant la même époque, ou encore de constructions similaires en proximité du site.

Chaque perte d'information ouvre plusieurs possibilités d'interprétation. Les restitutions hypothétiques sont donc des *versions* possibles, proches aux versions des produits d'un processus de conception. Les hypothèses peuvent enfin être validées ou démenties selon leur niveau de confiance et le processus d'interprétation des experts. La plupart des restitutions des états passés s'appuient conjointement sur ces types de raisonnements per la définition de la morphologie des éléments architecturaux.

2.2. Le degré de certitude

Dans des domaines variés, diverses études sur la perception de l'incertitude ont été amenées afin d'évaluer les paramètres qui en conditionnent la présence, afin d'interagir avec elle et la rendre un support pour la prise de décisions [MacEachren A.M. et al., 2005; Thomson J. et al., 2005]. [Pang A.T. et al., 1997] parle d'incertitude à trois niveaux : dans le processus d'acquisition, d'analyse et de visualisation des données. En patrimoine historique, les facteurs d'incertitude d'une source sont liées à trois aspects : la qualité, la cohérence et l'objectivité. [Thomson J. et al., 2005] propose une classification plus détaillée de ces aspects basée sur 9 sous-catégories (Fidélité, Précision, Intégralité , Cohérence, Filiation, Pertinence, Crédibilité, Interdépendance) pour décrire l'incertitude en patrimoine historique.

De façon générale, il est impossible de définir un ordre absolu de fiabilité des sources. Trop de paramètres rentrent en jeu pour des sources ayant des objectifs trop variés. On peut affirmer que le niveau de certitude augmente en fonction de la richesse d'informations métriques et visuelles de chaque source ; mais l'évaluation réelle du niveau de certitude / incertitude doit être établie cas par cas, selon la spécificité de chaque source. De ce fait, le niveau de précision, de cohérence et l'objectivité est plus élevé pour les sources les plus riches en détails métriques, visuels ou descriptifs. Les sources sont donc ordonnées selon leur apport métrique et visuel croissant (des sources les plus imprécises aux sources les plus détaillées).

Quand l'on étudie des bâtiments historiques, les informations qui concernent ses états antérieurs sont souvent contradictoires, hétérogènes, incertaines et incomplètes. Chaque source est caractérisée par une certaine *qualité*, une *cohérence* et un niveau d'*objectivité*. Ces trois aspects dépendent d'une composante humaine et des moyennes physiques qui ont permis de représenter l'objet.

- *La qualité de la source* dépend de son état de conservation, des outils de représentation employés et de l'objectif de la description. Premièrement, la qualité graphique des sources n'est pas souvent très élevée pour tirer des informations sur les proportions. Deuxièmement, dès que l'on a des informations métriques sur un objet, le niveau de fiabilité augmente immédiatement. Troisièmement, le niveau de fiabilité augmente encore plus si l'on analyse un relevé fait à l'aide d'un dispositif d'acquisition particulièrement précis et automatisé (c'est le cas du scanner laser par exemple).

- *Le niveau de cohérence* dépend de la concordance des informations véhiculées par les sources. Il s'appuie sur la comparaison avec diverses sources.

- *Le niveau d'objectivité* est lié à l'auteur de la source, car souvent il transfère graphiquement son propre point de vue par rapport à son contexte de vie, ou parfois il représente des architectures selon un projet pas encore réalisé.

2.3. Les techniques d'élaboration

A différence d'autres domaines où une collecte systématique d'informations permet d'effectuer des bilans statistiques, de créer des catégorisations ou de prendre des décisions, en patrimoine historique les informations sur les artefacts sont transmises de façon fortuite, grâce aux sources conservées ou découvertes. Les sources historiques et documentaires d'un site, par leur nature, peuvent être très hétérogènes par typologie et qualité de l'information communiquée. Leurs typologies sont détaillés comme suive.

2.3.1. Photographies

Cette catégorie inclue les photos historiques et actuelles, les photos stéréoscopiques, les mosaïques de photos et les orthophotos. Les photos peuvent être exploitées en diverses façons : la combinaison d'ensembles de photos à axes parallèles permet la génération de photos orthogonales (orthophotos terrestres et aériennes) ; les ensembles de photos à axes convergents et de photos stéréoscopiques (vues légèrement décalées d'une scène) permettent l'extraction des coordonnées spatiales et la restitution géométrique en utilisent des procédés de photogrammétrie.

2.3.1.1. Photo historique

En l5l5, Léonard de Vinci expliquait : << En laissant les images des objets éclairés pénétrer par un petit trou dans une chambre très obscure, tu intercepteras alors ces images sur une feuille blanche placée dans cette chambre [...] mais ils seront plus petits et renversés >> . Cette << camera obscura›> simple, dite « sténopé », est progressivement perfectionnée par l'ajout d'une lentille, puis d'un diaphragme et d'un miroir incliné à 45°, l'ancêtre du reflex. C'est ce dispositif qui, en remplaçant l'écran par une surface photosensible fixant la lumière, sera à la base de l'invention de la photographie en I839 par Nicéphore Niépce et Louis Daguerre. La photo historique peut être daté jusqu'à cette époque [Gernsheim H., 1997; Hirsch R., 2000].



Figure 5 Une vieille photo sur une carte postale présente une partie du château Comtal du Carcassonne.

Généralement les photos historiques peuvent être abîmées au cours du temps ou manquent parfois de netteté. Toutefois, elles peuvent témoigner des états passés d'un artefact disparus. Une vieille photo pourrait permettre la restauration des données dimensionnelles et géométriques relatives à la forme des éléments disparus. En outre, il peut parfois fournir des informations précises sur l'aspect visuel des matériaux qui composent le bâtiment.

2.3.1.2. Photo numérique

La photographie numérique est une forme de la photographie qui utilise un de capteurs sensibles à lumière pour capturer l'image focalisée par la lentille, par opposition à une exposition sur le film argentique. Elle ne requière pas la phase de développement. L'image capturée est ensuite stockée dans un fichier numérique prêt pour le traitement numérique (correction des couleurs, re-dimensionnement, recadrage, etc), l'affichage ou l'impression.

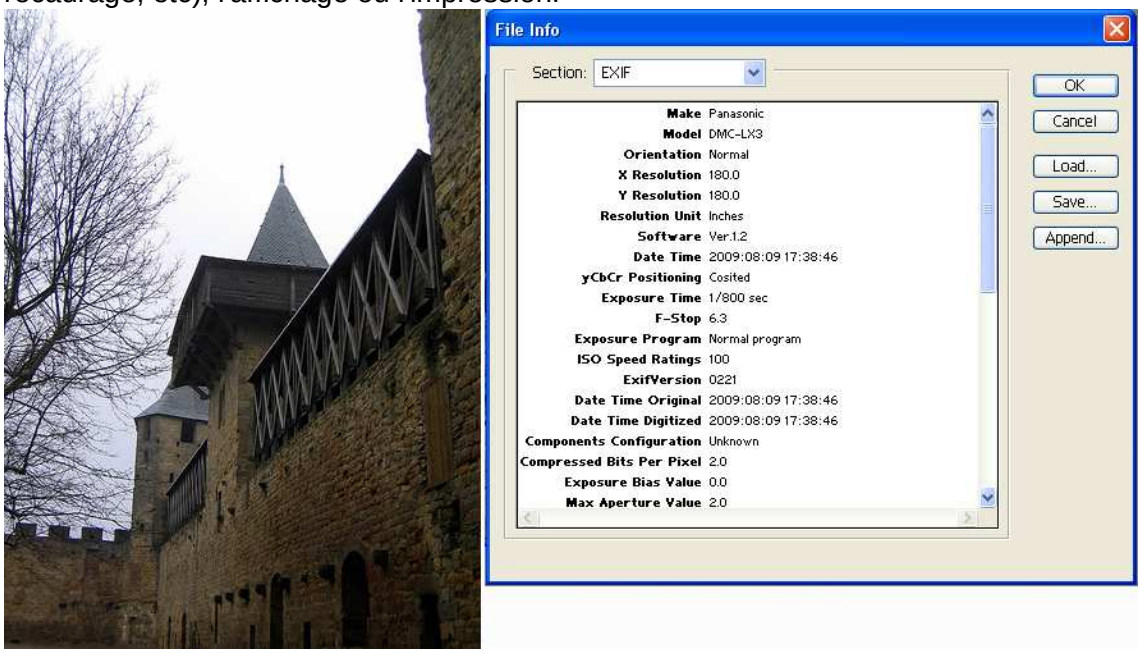


Figure 6 Métadonnées de format EXIF d'une photo numérique

Aujourd'hui, des images numériques peuvent être acquises directement par une caméra à matrice DTC (en anglais CCD) qui acquiert directement des images numériques. Elle offre un contrôle direct de la qualité de l'image et a de bonnes caractéristiques géométriques [Galer M. et al., 2005; Rand G. et al., 2005 ; Rosen M.J. et al., 2002].

Un des avantages majeur de la photographie numérique est la possibilité d'enregistrer, au moment de la prise de vue, les métadonnées EXIF relatives au paramétrage de l'appareil photo (Figure 6) [Clark R.N., 2002], comme par exemple la date et l'heure de la photo, le modèle de la caméra, la vitesse d'obturation, l'utilisation du flash, la distance focale, etc. De ce point de vue, les caméras argentiques ont logiquement une capacité bien limitée dans l'enregistrement de métadonnées, bien que de nombreux appareils photo peuvent graver une date sur une image en exposant le film à un matrice DEL (diode électroluminescente, en anglais LED) ou d'autres périphériques qui affichent une date.

2.3.2. Dessin

Le terme « dessin » inclut diverses représentations graphiques [Gerald F et al., 1998]. Les *croquis* sont des dessins faits à la main très rapidement, sans recherche de détails. Leur objectif est de donner un aperçu d'un élément afin de *noter* des aspects, des caractères ou d'être un support explicatif à la communication. Les *esquisses* sont des dessins rapides qui anticipent une œuvre plus sophistiquée, complexe et enrichie en détails, comme par exemple des *peintures*. Les *gravures et les estampes,* obtenues respectivement en incisant un matériau ou en faisant une impression ou un dessin sur un support, communiquent une richesse de détail plus élevée.

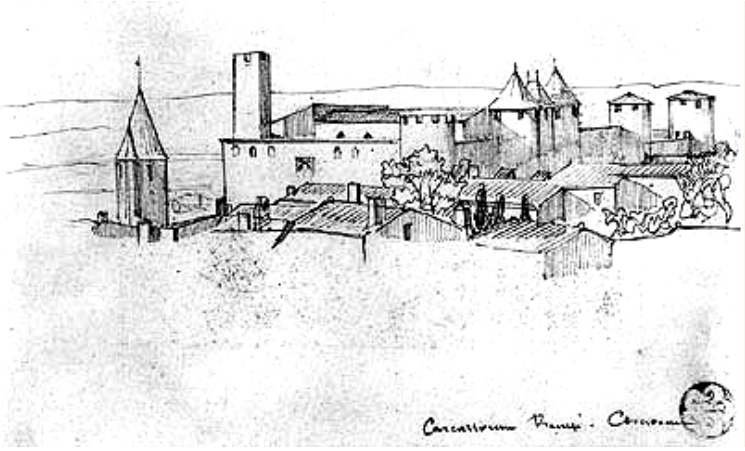


Figure 7 Dessin de vue générale sur le château comtal depuis le sud réalisée par Viollet-le-Duc.

2.3.3. Document métrique

La documentation métrique comprend généralement des représentations orthogonales de l'espace, exécutées à la main ou à l'aide d'outil de dessin technique, à des échelles variées, dont l'objectif est de fournir les mesures et les proportions entre les diverses parties. A cette catégorie appartiennent plans historiques, cadastres, plans urbains, coupes et élévation.

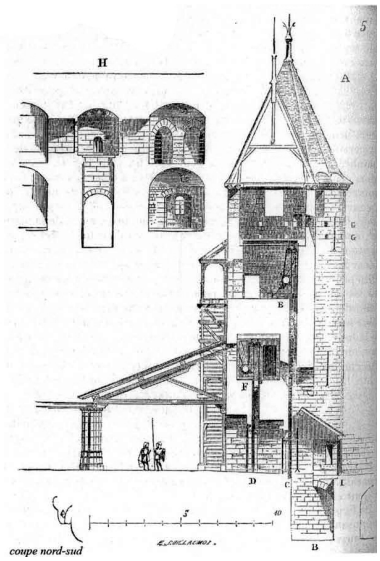


Figure 8 Coupe de la porte d'entrée du château Comtal de Carcassonne

2.4. Les types de projection

Le dessin de l'architecture a évolué au cours de l'histoire des méthodes de représentation en servant de différentes méthodes de construction et de contrôle géométrique de l'espace et des formes. Diffèrent types de projections ont été utilisée, au cours des siècles, comme principes sous-jacents aux pratiques de représentation de l'architecture.

2.4.1. Projection perspective

La projection perspective est la représentation d'une image telle qu'elle est perçue par l'œil. Le principe général de construction de ce type de représentations est l'utilisation de la distance entre l'objet et le point d'observation. Si l'objet est loin, il apparaît petit, et si l'objet est proche, il apparaît grand. La projection perspective est créé par des règles géométriques codifiées dans l'époque de la Renaissance.

Il existe une forte relation entre la perspective et les mesures. En effet, si les techniques de relevé utilisées au Moyen Age sont impliquées dans les origines de la perspective, après son invention, cette dernière a rapidement été employée comme instrument de mesure.

Au XV^{e} siècle, Filippo Brunelleschi introduit un procédé géométrique permettant de peindre avec exactitude une portion observée de l'espace réel. Dans son livre *De pittura*, il proposait une représentation de la perspective à un point de fuite principal pour simuler la vision humaine. Il s'agissait de la formalisation du passage entre la perspective «naturalis» (optique) et « artificialis » (géométrie) [Damisch H., 1994 ; Hyman I., 1974].

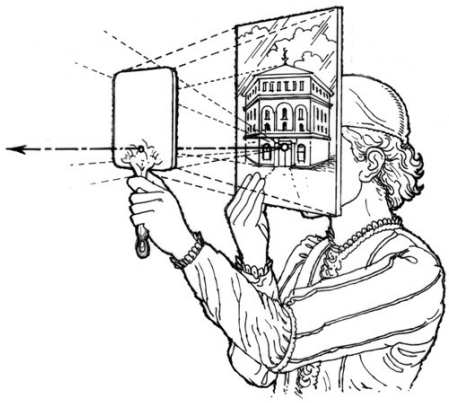


Figure 9 Illustration du dispositif de la perspective de Brunelleschi.

La définition de perspective donnée par Léonard de Vinci est : « la perspective n'est rien d'autre que voir un objet derrière une vitre plane et bien transparente, dont la surface contient toutes les choses qui sont derrière : elles peuvent être ramenées au point de l'œil par des pyramides et ces pyramides coupent la vitre mentionnée ».

D'autres principes, comme le rayon visuel ou la ligne de fuite, déjà présents dans la première formulation du procédé, furent ensuite repris par Jean-Henri Lambert pour développer les règles inverses de la géométrie projective consistant à déterminer la position de l'observateur, à trouver la ligne d'horizon et à estimer l'orientation des objets contenus dans un tableau. Lambert comprend que les points de fuite pouvaient être exploités comme support géométrique pour l'inversion de la perspective [Kemp M., 1992].

Ce type de projection est parmi les plus diffuses dans les collections de sources iconographiques d'architectures historiques. En faisant abstraction de détails d'ordre géométrique ou technique, on peut retrouver la projection perspective dans des dessins et de peintures de bâtiments (même si certains des documents graphiques, tels que les dessins, ne se basent pas forcement sur une construction géométrique rigoureuse de l'espace), jusqu'aux images élaborées par les nouvelles technologies d'infographie (comme par exemple les rendus en trois dimensions).

La photographie se base aussi sur une représentation en perspective de l'espace. De ce point de vue, la photographie permet une projection particulièrement fine de la réalité photographiée, à l'exception des erreurs introduites par la distorsion dépendent de plusieurs facteurs (lentilles, position du film, etc..).

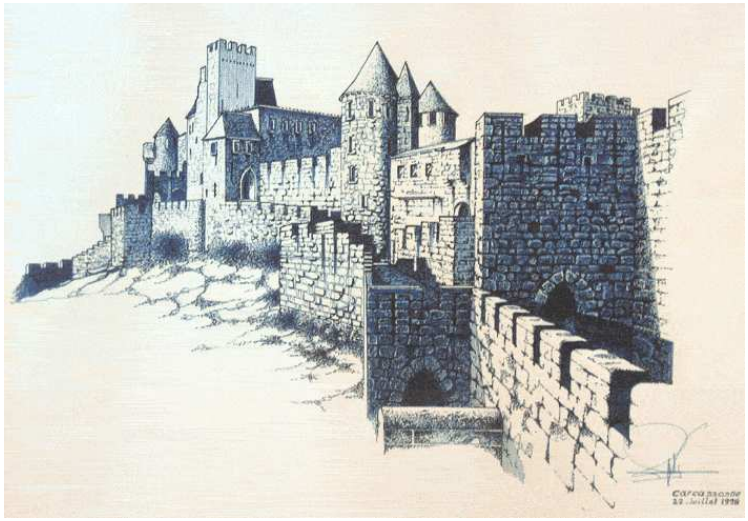


Figure 10 Un exemple de dessin en perspective du château Comtal de Carcassonne

Comme mentionné auparavant, la codification de la représentation en perspective a été introduite dans l'époque de la Renaissance. Par conséquent, les images qui ont été élaborées au cours des périodes précédentes (par exemple, l'époque médiévale), tout en se basant sur la perception des espaces et des formes par l'œil humain, ne se basent pas sur une règle formalisable (d'un point de vue géométrique au moins) de construction géométrique. C'est la raison pour laquelle, sauf dans certains cas rares, on ne peut pas considérer ce type de sources comme une information exploitable à des fins d'extraction et d'analyse métrique de la scène ou de la forme représentée. Dans certains cas, ces types d'images se rapprochent plus d'une projection axonométrique (voir section 2.4.2.2).

En dehors de la représentation d'un point de vue unique et statique, la projection perspective peut être retrouvée dans des procédés techniques plus complexes comme par exemple l'image stéréo et la séquence vidéo.

2.4.1.1. Image stéréo

L'image stéréo est une technique capable d'enregistrer en trois dimensions de l'information visuelle ou de créer l'illusion de profondeur dans une image à projection perspective. L'illusion de profondeur dans une image en deux dimensions est créée en présentant une image légèrement différente à chaque œil. L'image stéréoscopique a été inventée par Sir Charles Wheatstone en 1840 [Welling W., 1978]. Des images stéréoscopiques peuvent se trouver dans des collections de photographies anciennes et/ou de photographies actuelles.



Figure 11 Un exemple de photographie stéréo (porte de Lanhydrock House, Cornwall)

La stéréoscopie est utilisée en photogrammétrie pour la visualisation et la restitution géométrique d'images en calculant la différence de position de points homologues présents sur les deux images à la fois.

La photographie moderne peut acquérir et enregistrer l'information en 3 dimensions [Wiley C., 2008]. Les informations sur la profondeur de la scène est reconstruite à partir d'une corrélation automatique de chaque pixel d'un couple d'images.

Même si aujourd'hui on dispose de technologies de balayage laser 3D pour le relevé et la numérisation tridimensionnelle de l'architecture. Les solutions basées sur la photogrammétrie, stéréo et multi-vues restent très utilisées et constituent un moyen particulièrement flexible et démocratique pour l'enregistrement d'informations métriques et géométriques sur le patrimoine bâti.

2.4.1.2. Séquence vidéo et Multimédia

La vidéo regroupe l'ensemble des techniques permettant l'enregistrement ainsi que la restitution d'images animées, accompagnées ou non de son, sur un support électronique. Comme la photo numérique, et toujours en faisant abstraction des phénomènes de distorsion, une séquence d'images produite par caméra vidéo se base sur une projection perspective. Cette catégorie inclue également les animations créées par le rendu de modèles 3D en utilisant des solutions infographiques.

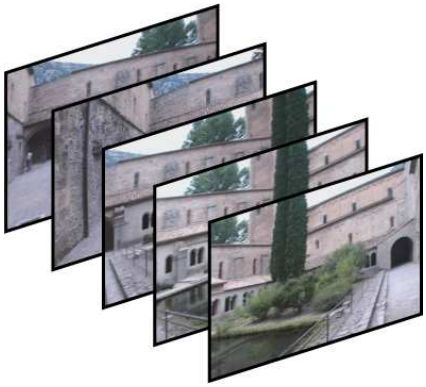


Figure 12 Un séquence d'images extraite d'une séquence vidéo

Des techniques de *camera tracking* (reconstruction du parcours d'une caméra) et de *shape from video* (reconstruction de la forme à partir d'une séquence vidéo) permettent de récupérer à la fois la structure de la scène filmée et le mouvement que la caméra a effectué à son intérieur. Ces techniques se basent sur des solutions d'analyse et de traitement d'image automatiques ou semi-automatiques [Pollefeys M. et al., 2000]. La flexibilité de l'approche permet d'appliquer cette technique dans le domaine de l'étude et de la représentation spatiale de sites patrimoniaux. Ces techniques peuvent en effet être utilisées pour enregistrer et reconstruire des scènes, des immeubles, des éléments d'architecturaux, des statues et de la géométrie de terrain.

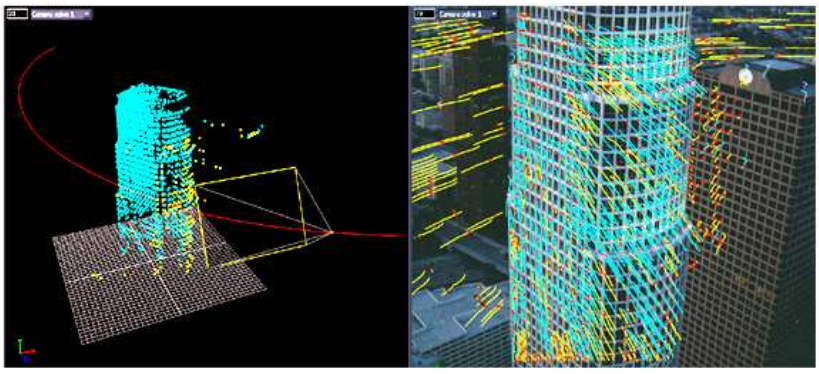


Figure 13 Un exemple d'application de la technique de *Camera tracking*

[Cosmas J. et al.,2001] utilisent ces techniques dans le cadre du projet *3D MURALE*. Ce projet se concentre sur la mise en relation d'un ensemble d'outils à l'intérieur d'un système multimédia pour l'enregistrement, la reconstruction, le codage, la visualisation et l'interrogation de base de données pour la documentation de sites archéologiques.

2.4.2. projection parallèle

La projection parallèle correspond à une projection perspective avec un point de vue hypothétique, se situant à une distance infinie de l'objet [McReynolds T. et al., 2005].

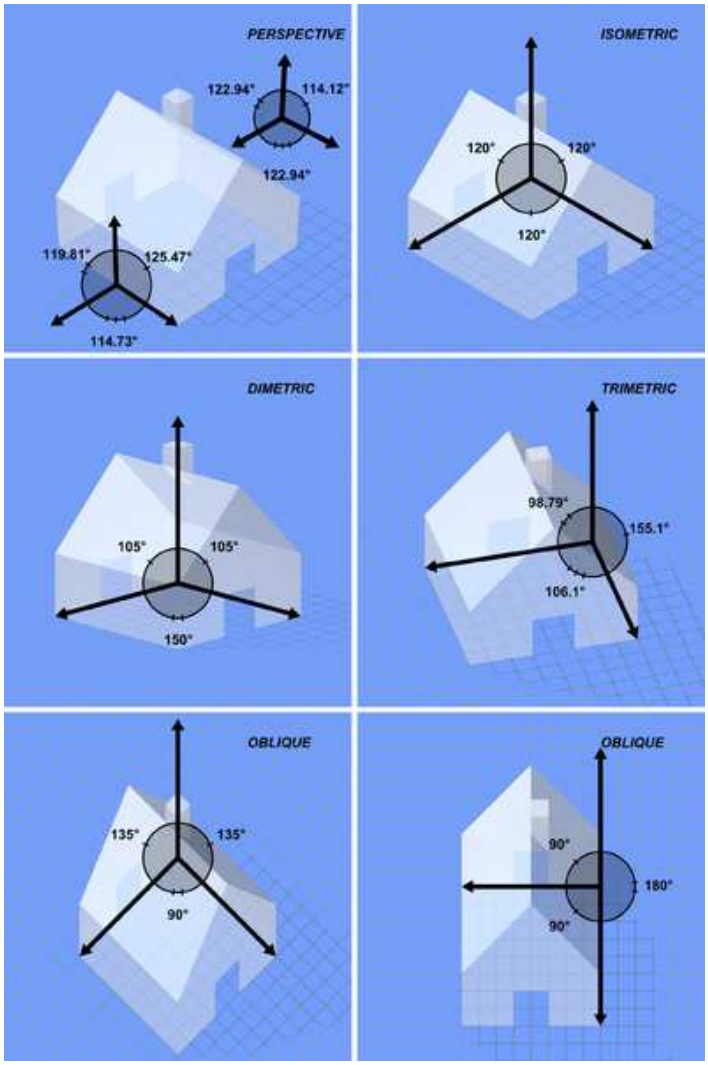


Figure 14 Tableau résumant les différents types de projection parallèle

2.4.2.1. Projection orthogonale

La projection orthogonale est une représentation en deux dimensions d'un objet en trois dimensions. Cela correspond à une projection à l'infini [Maynard P., 2005]. C'est le type de projection le plus utilisé pour l'élaboration de dessins techniques. Dans les sources iconographiques relatives au patrimoine, ce type de projection est assez utilisé dans les dessins d'édifices (comme les plans, élévations et les coupes).

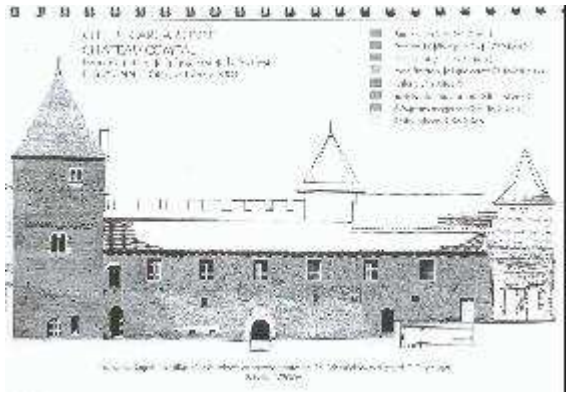


Figure 15 Élévation de l'aile ouest du château Comtal de Carcassonne. Mélange entre une orthophotographie et un dessin au trait.

Certains procédés de photogrammétrie permettent l'élaboration d'orthophotographie, c'est à dire de projection orthographique de photographies à partir d'une transformation géométrique appelée homographie (Figure 15). Les orthophotographies peuvent fournir des informations métriques avec une résolution et un niveau de précision important en intégrant des éléments de détail des surfaces.

2.4.2.2. Projection axonométrique

Axonométrie est un type de projection orthographique où le plan ou l'axe de l'objet représenté n'est pas parallèle au plan de projection [Godse A.P., 1980; Gary R. B. et al., 2002], de telle sorte que plusieurs côtés d'un objet sont visibles dans la même image [Maynard P., 2005]. La projection axonométrique peut être de trois différents types : isométrique, dimétrique et trimétrique, en fonction de l'angle formé entre l'axe et le point de vue orthogonal [McReynolds T. et al., 2005]. Une caractéristique typique des images axonométriques consiste en le fait que un des axes est généralement représenté en vertical.

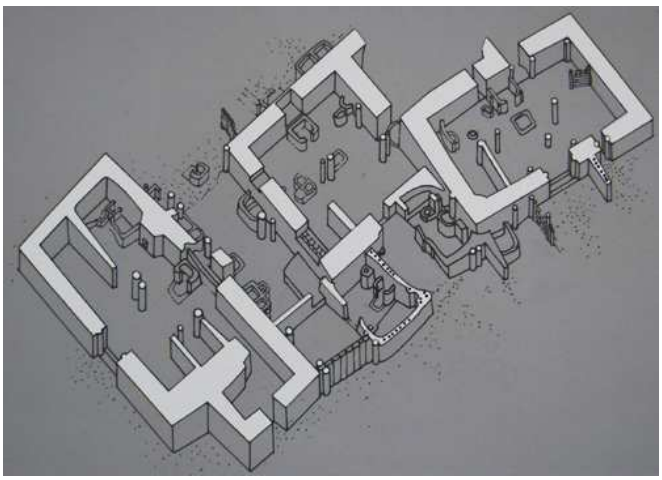


Figure 16 Un exemple de l'image isométrique de la fouille archéologique (maisons néolithiques, Hacilar, Turquie)

Il n'existe pas beaucoup d'images anciennes construites selon une projection axonométrique correcte. Normalement, c'est un type de construction géométrique dont la pratique s'est diffusée dans l'époque moderne. Si des images dont la construction se rapproche d'une axonométrie existent avant l'époque moderne, c'est d'habitude des images crées dans le but d'arriver à suggérer un point de vue perspectif. Ce type de représentation a une grande diffusion dans les dessins et les peintures de monuments ou de villes.

2.4.2.3. Projection oblique

Dans les projections obliques les rayons de projection parallèle ne sont pas perpendiculaires au plan de visualisation comme dans le cas de la projection orthographique, mais orientés d'un angle qui diffère de 90 degrés [Maynard P., 2005]. Dans les projections orthographiques et obliques les lignes parallèles dans l'espace semblent parallèles sur l'image. Grâce à la simplicité de sa construction graphique, cette technique de représentation s'est particulièrement diffusée dans la période antécédente à l'introduction de la perspective. C'était d'ailleurs la volonté de figurer l'espace et sa perception, avant la codification de la perspective, qui a déterminé une utilisation des projections obliques. Ce n'est pas par hasard que les représentations d'époque médiévale basées sur des projections obliques sont aussi appelées (images en perspective gauche).

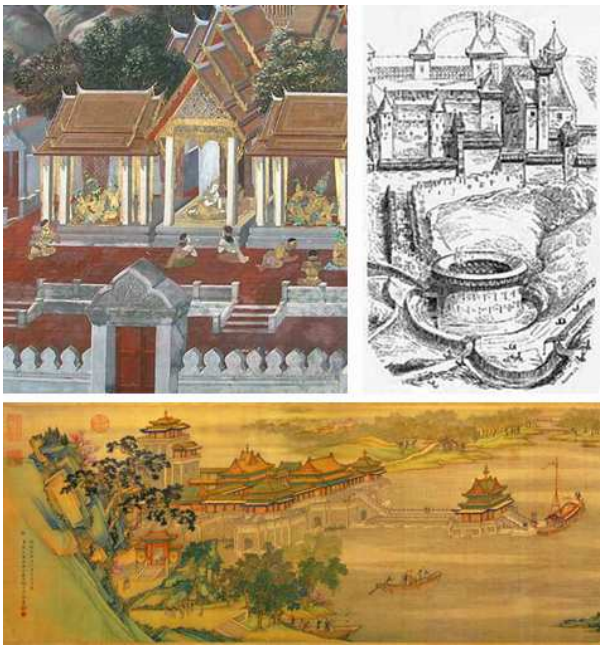


Figure 17 Des exemples d'images à projection oblique
(dessin de palace thaïlandais du 17e siècle, perspective gauche du château comtal de Carcassonne, paysage chinois)

2.4.3. Projection panoramique

L'image panoramique est basée sur l'assemblage d'images prises à partir de la même position dans l'espace mais avec des orientations différentes. Ces image sont très utiles pour la documentation d'espace architecturaux : la seule phase d'assemblage est suffisante pour élaborer des image interactives permettant d'explorer l'espace photographié à 360° [Ward G., 2006; Little field R. ,2006].

Les assemblages panoramiques peuvent aussi être utilisés pour reconstruire de façon sommaire la géométrie de l'espace à partir d'une orientation approximative s'appuyant sur l'identification d'un repère orthonormé dans la scène. Finalement, il est possible d'exploiter une ou plusieurs images panoramiques ou de les intégrer à des images prises avec une configuration à axes convergents ou parallèles pour reconstruire de manière détaillée des espaces architecturaux via des solutions de photomodélisation.

L'acquisition d'une image panoramique s'appuie sur la notion de recouvrement. Dans ce sens, la mise en correspondance des différentes images est l'aspect fondamental de l'assemblage d'une image panoramique. L'assemblage panoramique consiste a détecter des points caractéristiques communs a des couples d'images. Cette mise en correspondance d'images peut être automatisée compte tenu des conditions particulières de prise de vue (toutes les images partagent le même centre optique) [Szeliski R., 2005 ; d'Angelo P., 2007]. Désormais, il existe des méthodes entièrement automatiques pour analyser et faire correspondre des portions d'images ayant des zones de recouvrement. Le résultat d'un assemblage panoramique peut être généralement rendu selon trois projections différentes (cylindrique, sphérique et cubique).

2.4.3.1. Projection cylindrique.

Les photographies assemblées sont projetées à partir du centre optique sur un cylindre dont les dimensions verticale et horizontale représentent respectivement la largeur et la hauteur de l'image panoramique.

Ce type de projection est particulièrement efficace pour l'enregistrement d'informations spatiales distribuées sur une ligne d'horizon ne présentant pas d'élévations importantes.

2.4.3.2. Projection sphérique.

Les photographies assemblées sont projetées à partir du centre optique sur une sphère.

L'espace photographié est donc distribué sur 360° s elon l'axe horizontal et sur l80° selon I' axe vertical. La projection sphérique permet donc de représenter la totalité de l'espace environnant de façon homogène où la valeur angulaire relative à un pixel horizontal est liée a celle d'un pixel vertical selon un rapport de 360° / l80°.

2.4.3.3. Projection cubique.

Les photographies acquises et assemblées sont projetées (selon une projection perspective classique) a partir du centre optique sur les six faces d'un cube. Comme pour le cas d'une projection sphérique, une projection cubique permet la représentation complète de l'espace environnant.

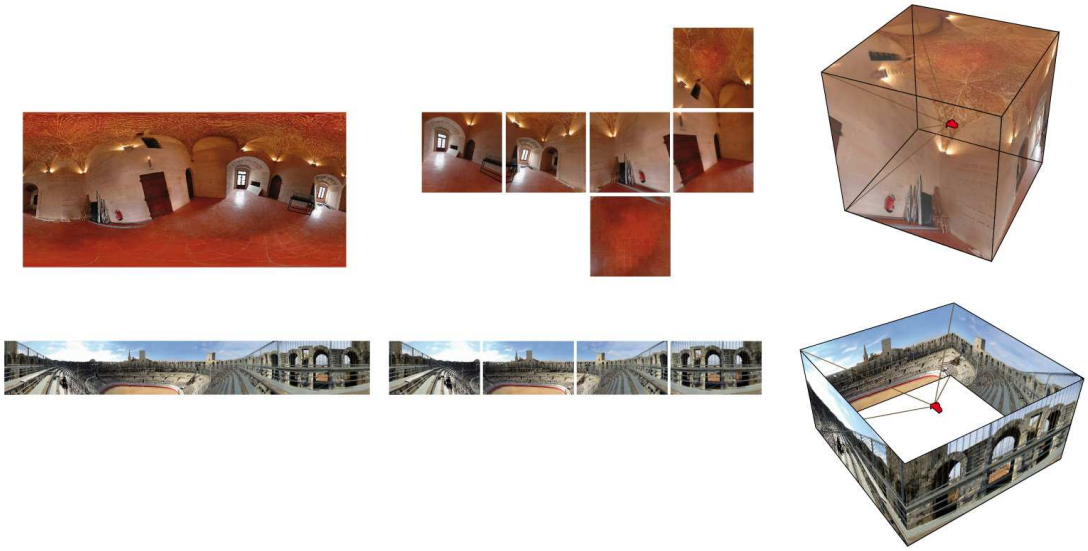


Figure 18 Des exemples d'images panoramiques

D'un point de vue géométrique, l'aspect intéressant des images panoramiques est qu'en étant générées à partir d`un assemblage d'images, tous les types de projections (cylindrique, sphérique et cubique) partagent le même principe de génération : on projette les pixels contenus dans les photographies assemblées à partir d'un centre optique central vers trois types de surfaces. Compte tenu des cohérences géométriques existantes entre les trois projections, il est alors possible de généraliser le point de vue permettant de représenter une image panoramique dans l'espace quelque soit sa projection.

3. Méthodes d'annotation sémantique d'images

Dans le chapitre précédent, nous nous sommes intéressés à une classification, basées sur différents critères, des différents types d'images que l'on peut utiliser pour documenter un édifice patrimonial. Dans ce chapitre, nous nous intéressons aux méthodes et aux solutions existantes pour la sémantisation d'images.

L'étiquetage du contenu sémantique des images au travers d'un ensemble de mots clés est une problématique connue sous le nom d'« annotation d'image ». L'annotation est principalement utilisée pour la gestion de base de données d'images, notamment afin d'améliorer les possibilités de la recherche. Des images annotées peuvent habituellement être retrouvées en utilisant une recherche par mots-clés, alors que des images non-annotées peuvent être extrêmement difficiles à trouver dans une base de données. L'annotation d'images par mots clés permet également d'améliorer les processus d'indexation d'informations utilisés par les moteurs de recherche Web [Rodden K.,1999].

Aujourd'hui, Il y a approximativement trois manières pour enrichir d'un point de vue sémantique des images [Wenyin L. et al., 2001] :

- l'annotation manuelle ;
- annotation semi-automatique ;
- annotation automatique.

L'annotation sémantique n'est pas encore utilisé dans le domaine spécifique de l'architecture patrimoniale, par conséquent, nous présentons ici les travaux de recherche et les développements relatifs à un domaine d'application beaucoup plus large.

3.1. Méthodes manuelles

La plupart des systèmes de base de données d'images utilisent aujourd'hui l'annotation manuelle [Gong Y. et al., 1994]. Il s'agit d'une méthode d'annotation qui permette aux utilisateurs d'associer des mots-clés descriptifs aux images conservée au sein d'un base de données. Deux types d'annotation différentes existent : celle basée sur les mots-clés et celle basée sur les ontologies. Dans le premier cas les utilisateurs sont libres d'étiqueter les ressources considérées avec n'importe quel mot-clé qu'ils retiennent, adéquat à la description du contenu de l'image ; dans le deuxième cas les

mots à utiliser pour la description du contenu des images sont reliés à une structure de concept représentant un domaine de connaissances précis.

Dans le cas de l'annotation libre par mots-clés les utilisateurs ne sont pas forcés de suivre un arrangement formalisé. Puisque aucune conceptualisation partagée n'est prise en considération, l'association de l'étiquette à une interprétation sémantique précise peut seulement être accidentellement réalisée. Les exemples bien connus de ce genre d'annotation pour les images 2D sont [Flickr] et [Riya]. Ces systèmes se basent sur des interfaces permettant aux utilisateurs de tracer un rectangle autour d'un objet dans une photographie puis d'introduire quelques termes descriptifs.

Différents travaux de recherche ont exploré des techniques pour améliorer le processus d'annotation sémantique manuelle. [Shneiderman B. et al., 2000] ont développé une méthode d'étiquetage des noms des personnes dans un ensemble de photos dans le cadre du projet «*PhotoFinder* » (voir figure 19). Par cette méthode, l'utilisateur peut simplement choisir manuellement un mot clé à partir d'une liste de noms et se servir d'une manipulation de type « Glisser - déposer » (*drag & drop*) sur les images afin d'en indiquer la personne correspondante. Bien qu'elle évite une grosse partie de travail de dactylographie, cette méthode reste toujours manuelle et requiert beaucoup de manipulations.



Figure 19 L'interface de PhotoFinder

Une technique permettant d'accélérer cette tache lourde est l'annotation par lots : un group d'images sont sélectionnées, puis le même mot-clé est appliqué à chaque photo avec une seule action [Kang H. et al., 2000; Kuchinsky A. et al., 2002].

Du coté de l'annotation basée sur l'ontologie, les étiquettes sont définies par des spécifications formelles et explicites d'une conceptualisation partagée d'un domaine de connaissances [Gruber T.R. 1993], [Guarino N., 1998].

Puisque la conceptualisation est partagée, il n'y a aucune liberté dans le choix des noms d'étiquettes, mais cette contrainte est par contre récompensée par un arrangement commun des étiquettes. D'ailleurs, la conceptualisation partagée peut

également être traitée par des applications informatiques, fournissant des possibilités particulièrement probantes en terme de recherche de contenu [Saathoff C., 2006].

Photostuff [Halaschek-Wiener C. et al., 2005] donne aux utilisateurs la possibilité d'annoter des régions d'images en utilisant une ontologie ainsi que d'éditer les métadonnées de l'image dans le Web. Dans M-OntoMat-Annotizer [Petridis K. et al., 2006], l'utilisateur peut dessiner et accentuer des segments (i.e. régions) d'une image, visualiser (browse) des ontologies spécifiques à différents domaines, puis annoter les régions identifiées (voir figure 20).

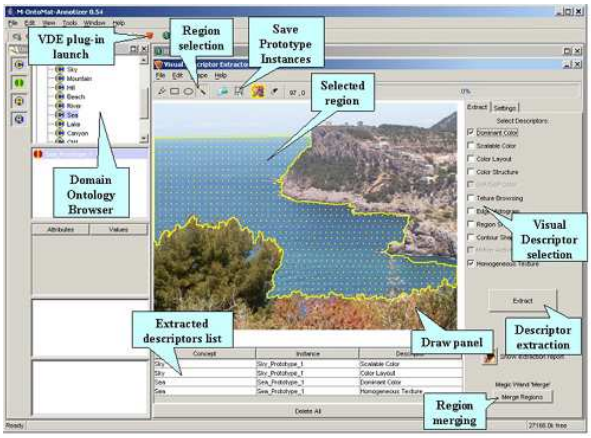


Figure 20 L'interface de M-OntoMat-Annotizer

Dans le domaine de l'architecture archéologique [Blaise J.Y. et al., 2007] et [Azzi Y., 2008] proposent un système qui permet à l'utilisateur d'identifier des régions des images en traçant la silhouette des objets à décrire puis en associant à chaque région crée un termes du vocabulaire architectural faisant partie d'une base de données ontologique (voir figure 21).

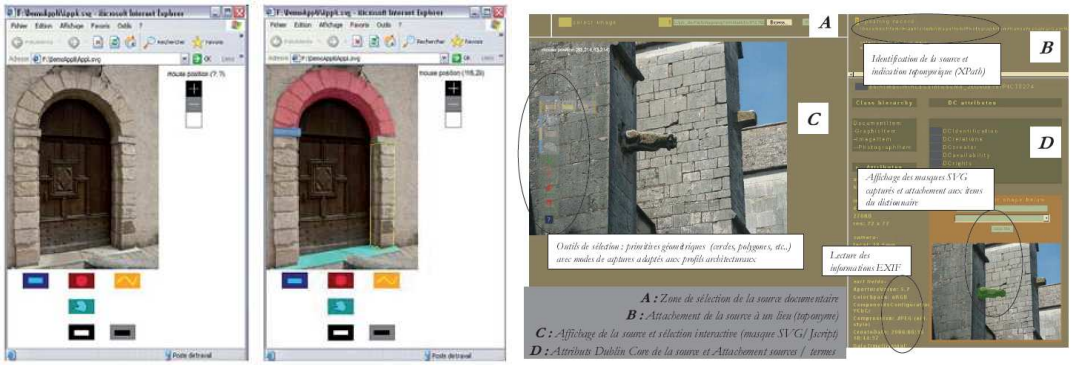


Figure 21 Annotation sémantique d'images d'architecture patrimoniale
[Azzi Y., 2008; Blaise J.Y. et al., 2007]

[Cross J. et al.,] proposent un dispositif de collecte et d'annotation d'images pour les applications de terrain dans les chantiers de fouilles archéologiques. Les images sont recueillies par webcam et géo-rectifiées. L'utilisateur peut annoter les images en utilisant un système de couches d'annotations. Plusieurs objets peuvent

être attachés à une annotation, et les annotations peut être présentées individuellement (en une seule couche), ou comme un composite.

Analyse : Bien que l'annotation manuelle du contenu d'image soit considérée un meilleur cas en termes d'exactitude du contenu sémantique des images, puisque des mots-clés choisis sont basé sur la détermination humaine, c'est un processus qui a besoin de beaucoup de travail. En outre, l'annotation manuelle peut également présenter des erreurs de récupération à cause de oubli d'utilisateurs ou descripteurs après une longue période de travail.

Les critères pour saisir un mot-clé au travers de cette méthode sont complètement subjectifs. Les termes ne sont pas référencés par le même standard. Même si certains systèmes sont basés sur des mots-clés structurés dans une ontologie, le nombre et l'association des termes est toujours dépendante d'une décision personnelle. Une personne peut entrer un seul mot-clé par image, alors que une autre personne pourrait en utiliser plusieurs. Certaines annotations jugées inutiles pour certains pourraient être importantes pour d'autres.

En outre, dans les solutions basées sur cette méthode il est extrêmement difficile de conduire de mise à jour de la structure de la base de données. Car, comme les mots-clés sont entrés manuellement, lorsqu'on modifie un mot-clef ou l'organisation des mots-clés, la seule manière de mettre à jour la base de données est de modifier les termes associés image par image.

Aujourd'hui, de nombreuses expériences du domaine de la documentation du patrimoine ont adoptée cette méthode de sémantisation des images grâce à sa simplicité d'utilisation. Si on prend en considération cette méthode dans une utilisation basique, elle peut aujourd'hui répondre aux besoins exprimés par les communautés scientifiques et professionnelles concernées. Mais si l'on s'intéresse à l'apport potentiel de ces technologies dans le domaine de la documentation patrimoniale, il est certain qu'il y aura à terme un besoin plus important de manipuler des collections d'images beaucoup plus importantes et un besoin complémentaire d'augmenter la complexité des descriptions (plusieurs niveaux de détails). Au vu des limites actuelles de ces méthodes manuelles, il est d'hors et déjà évident que ces méthodes ne pourrons pas répondre aux besoins potentiels des utilisateurs de cette communauté.

3.2. Méthodes automatiques

Les méthodes automatiques d'annotation sémantique d'images se basent sur des techniques d'analyse d'image basées sur le contenu (*content-based image analysis techniques*) [Flickner M. et al., 1995]. Normalement, ces processus sont composés de deux tâches, la segmentations d'image (détection automatique les forme d'éléments dans l'image) et l'identifications d'objets (la comparaison entre les formes détectées et des objets cohérents dans une bibliothèque).

L'identification des classes d'objets peut être réalisée en utilisant une combinaison des modèles particuliers [Winn J. et al., 2005 ; Fergus R. et al., 2003] (voir figure 22). Plusieurs chercheurs ont considéré ces deux tâches séparément. Par exemple [Kumar S. et al., 2004] se sont concentrés seulement sur le problème de la segmentation.

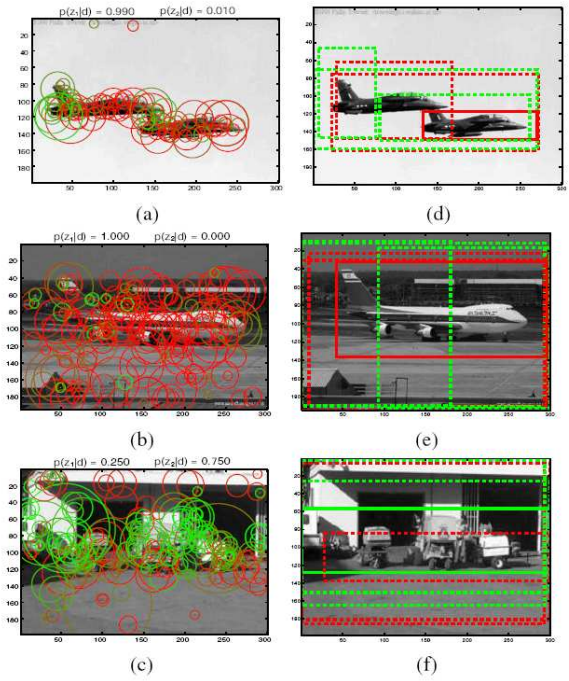


Figure 22 Les résultats d'une détection sémantique proposée par le modèle TSI-pLSA de [Fergus R. et al., 2005]

L'outil TextonBoost, proposé par [Shotton J. et al., 2006] combine la reconnaissance de formes et la segmentation d'images. Ils utilisent une combinaison de caractéristiques spécifiques appelé *Texton* pour encoder la forme, la texture et l'apparence de différentes classes d'objets. Bien que leur méthode a produit des résultats significatifs, les formes bruitées et les textures causent son échec sur les frontières des objets (voir figure 23).

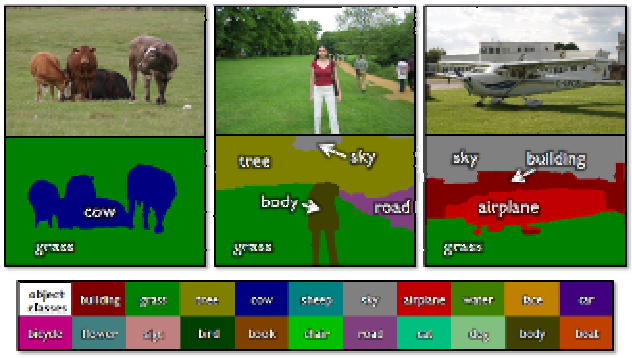


Figure 23 Les résultats de la détection sémantique de TextonBoost [Shotton J. et al., 2006]

Dans le domaine de l'architecture patrimoniale, il n'y a pas beaucoup de recherches s'appuyant sur l'analyse automatique de la sémantique d'image. Par contre, dans un domaine proche, certains chercheurs ont développé des méthodes et des outils pour analyser l'architecture urbaine. [Johansson B. et al., 2002a] proposent une méthode pour détecter un élément architectural simple dans la scène d'une ville, comme par exemple la détection automatique des fenêtres des bâtiments.

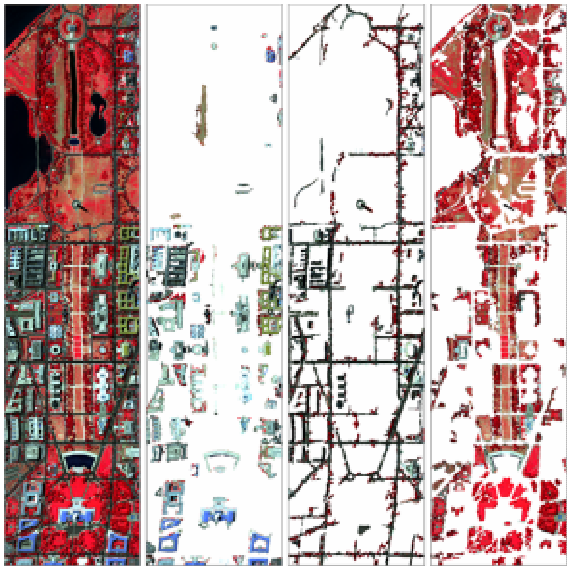


Figure 24 Un exemple de détection sémantique à partir d'images aériennes [Akcay H.G. et al., 2007]

À l'échelle urbaine, la sémantisation automatique d'images aériennes peut produire beaucoup plus de résultats utilisables. Des techniques spécifiquement développées pour la détection de bâtiments peuvent être trouvées dans la littérature de [Mayer H., 1999]. La détection automatique d'objets géographiques comme les structures urbaines se fondent sur la position, l'échelle et l'orientation d'éléments urbains dans l'image [Akcay H.G. et al., 2007]. (voir figure 24).

La plupart des recherches sur l'analyse de l'image urbaine se concentrer sur la classification basée sur pixel. Les textures peuvent également être utilisés comme méthode alternative [Karathanassi V. et al., 2000 ; Bhagavathy S. et al., 2007]. Des éléments architecturaux ont également été utilisés pour classifier les régions d'images

de zones rurales, résidentielles ou urbaines [Unsalan C. et al., 2005 ; Aksoy S., 2006] (voir figure 25).

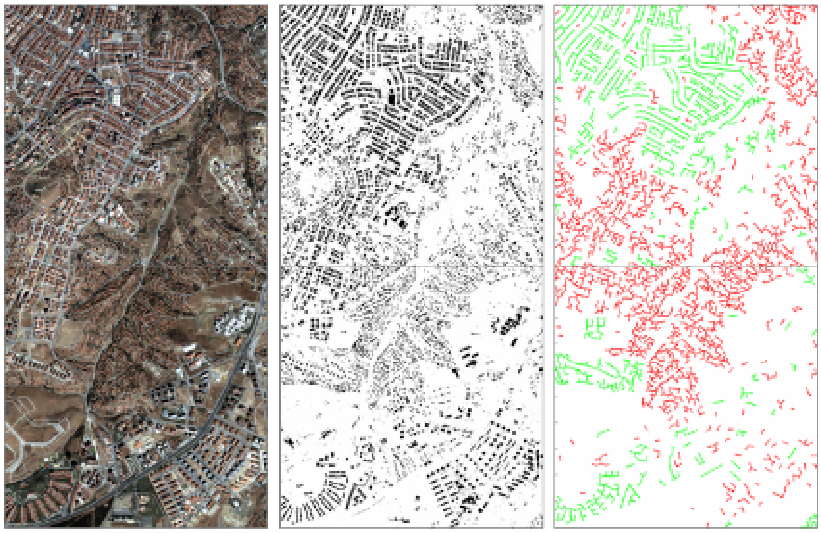


Figure 25 Les résultats de la détection sémantique d'images aériennes à l'échelle urbaine utilisant des patterns de bâtiment [Aksoy S., 2006]

Analyse La sémantisation automatique par analyse d'image est en général très rapide et nécessite de moins de travail humain (par rapport à l'annotation manuelle). Malgré cela, les solutions actuelles, ne sont pas encore suffisamment précises et fiables. Au-delà du problème de la précision, les limites de cette méthode sont que le champ d'application possible est très restrictif. En effet, les deux étapes d'analyse d'image (segmentation et identification d'objets) doivent avoir les données de base appropriées.

En ce qui concerne la segmentation d'image, le problème de la détection des formes est particulièrement difficile à traiter en particulier dans le cas d'images qui présentent beaucoup d'éléments hétérogènes à leur intérieur. D'autres problèmes sont du à la présence de plusieurs objets similaires ou quand la couleur et le contraste de fond est proche de celui du sujet principal de l'image.

En ce qui concerne l'identification d'objets, les solutions actuelles nécessitent d'une très grande base de données afin de permettre l'identification de plusieurs typologies d'objets. Si la base de données ne contient suffisamment de modèles pouvant être superposés aux formes détectées dans l'image, ces systèmes ne peuvent pas reconnaître leur sémantique. Et même dans le cas où le modèle de l'objet détecté dans l'image est présent dans la base de données, il est aussi nécessaire que le point de vue (et donc la projection) de cette objet coïncide avec le ou les points de vue associés au modèle. C'est la raison pour la quelle l'utilisation de cette méthode dans les images aériennes est beaucoup plus fiable, car elle bénéficie d'un point de vue zénithale.

En revenant à notre domaine d'intérêt, l'architecture patrimoniale, il est sur que les limitations actuelles des méthodes automatiques rendent ces méthodes particulièrement inadaptée aux besoins des scientifiques et des professionnels de ce

domaine. En effet, en termes de complexité d'image, les photographies d'édifices architecturaux sont souvent assez complexes d'un point de vue de la quantité d'éléments présents dans la scène (souvent de différentes tailles) et rarement bien identifiables. En ce qui concerne la constitution de bases de données d'images correspondantes aux formes architecturales, la richesse et la diversité du patrimoine architecturale empêchent d'imaginer d'identifier des classes d'éléments pouvant être superposés correctement aux éléments photographiés. Enfin, en ce qui concerne la variété de point de vue (ou de points de prise photographique), les photos acquises pour des besoins d'étude patrimonial ou de documentation sont souvent assez hétérogènes en termes de paramétrage e de configuration de prise de vue.

3.3. Méthodes semi-automatiques

Les méthodes semi-automatiques d'annotation sémantique d'images se basent sur des techniques combinant les solutions présentées précédemment (automatiques et manuelles). Des algorithmes sont utilisés pour le filtrage de la sémantique appropriée en utilisant les information textuelles liées à l'image ou par l'identification des contenus d'image (comme dans les méthodes complètement automatiques), puis l'utilisateur peut choisir manuellement les mots-clefs qui se rapportent à l'objet dans l'image. En employant le contrôle des utilisateurs, la sémantique obtenue est mises à jour de façon incrémentale. Intégrées à des interfaces appropriées, les méthodes d'annotation semi-automatique ont un grand potentiel en terme d'application à plusieurs domaines [Suh B. et al., 2004].

[Shen H.T. et al. 2000] ont employé une méthode dans laquelle les images sont utilisées pour indexer automatiquement d'autres images dans la même page Web. Ils emploient l'information textuelle riche de la page pour extraire des éléments de description potentielle des images. [Srihari R. et al., 2000] extraient du texte des entités textuelles (par exemple, les nom des gens, les nom des endroits et des choses, etc.) pour indexer automatiquement les images qui accompagnent le texte des pages web.

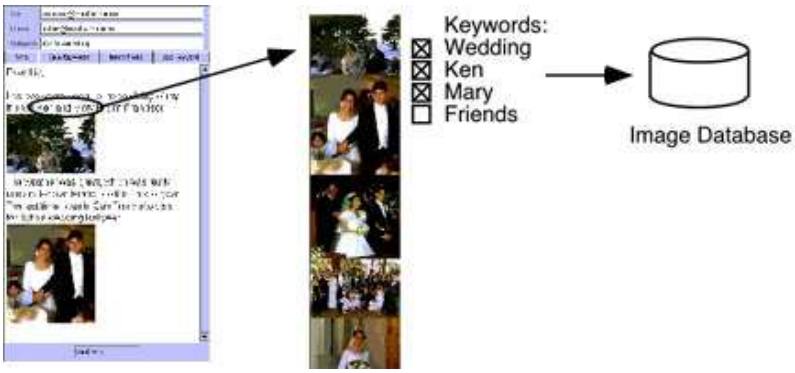


Figure 26 Le processus de la détection sémantique du ARIA

[Lieberman H., 2000] décrit le système ARIA (Annotation and Retrieval Integration Agent) qui intègre la récupération et l'utilisation d'image. Le système emploie l'extraction des mots présents dans les messages électroniques (dans lesquels des images sont incluses) pour indexer ces images (voir figure 26). Cette méthode semble fournir des résultats raisonnables, mais la précision du contexte textuel n'est pas aussi haute que l'indexation manuelle. [Wenyin L. et al., 2001] développent un système capable de générer et d'affiner progressivement des annotations sémantiques d'images de façon automatique en demandant à l'utilisateur la validation et l'organisation (ou la réorganisation) du résultat final (voir figure 27).

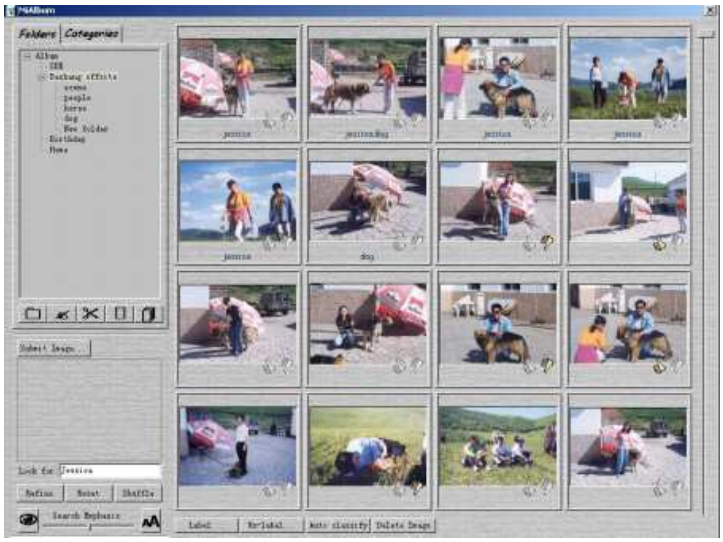


Figure 27 L'interface de *MiAlbum* pour la recherche basée sur texte, pertinence de résultat et d'annotation semi-automatique

[Ono et al. 1996] ont essayé d'employer des techniques d'identification d'objets pour choisir automatiquement des mots-clés descriptifs appropriés (dans un ensemble prédéfini) pour chaque image. Cependant, ils ont seulement examiné leur système avec des mots-clés et des images limités, ainsi la possibilité d'employer ce système dans un contexte plus général est peu claire. D'ailleurs, puisque les techniques d'identification d'image ne sont pas complètement fiables, les utilisateurs doivent en autre confirmer ou vérifier des mots-clés produits automatiquement par le système.

Analyse Il est certes qu'en combinant les techniques automatiques et manuelles, les méthodes semi-automatiques d'annotation sémantique d'images permettent d'obtenir des résultats meilleurs en termes de précision et de productivité effective.

Tout en restant probablement les méthodes les plus performantes, dans le domaine spécifique de la documentation architectural le problème de l'incohérence des formes persistent. Car les édifices historiques sont généralement très complexes et composés de nombreux éléments hétérogènes difficiles à classifier de façon préalable dans le dessein de constituer une base de modèles à utiliser pour la reconnaissance.

4. Une approche pour la caractérisation sémantique 2D/3D d'édifices patrimoniaux

Suite à l'analyse des méthodes actuelles d'annotation sémantique d'images, on peut retenir que les solutions actuelles, tout en donnant des résultats applicables à certains contextes, ne sont certainement pas encore au point pour aborder le problème plus complexe de la segmentation et annotation précise d'artefacts patrimoniaux. En effet, même si les méthodes semi-automatiques semblent offrir des résultats convenables, le domaine de la documentation du patrimoine reste particulièrement difficile à traiter au vu de l'élevée complexité et hétérogénéité des formes architecturales.

Les vrais problèmes dérivent probablement du fait que cette complexité et cette hétérogénéité (typique du contexte de l'architecture patrimoniale) oblige à traiter les problèmes de classification des formes non pas à partir d'un simple aperçu bidimensionnelle de leur aspect (sur les images), mais à partir d'une véritable analyse de leur forme (sur des modèles 3D).

Pour donner un exemple, la documentation photographique des artefacts se base, le plus souvent, sur la prise de multiples photographies, à partir de différents points de vue, avec différents niveaux de zoom des mêmes formes. Or, si l'on imagine d'utiliser des solutions d'annotation sémantique d'images telles que l'on vient d'analyser dans le chapitre précédent, il est évident qu'il serait particulièrement problématique de garder (sauf dans le cas de l'utilisation d'une solutions complètement manuelle) sur toutes les images un niveau cohérent de précision à la fois en termes de segmentation et en termes d'annotation sémantique. De plus, si l'on s'intéresse à une annotation particulièrement fine d'édifices patrimoniaux (c'est le cas dans les travaux menés pour des finalités de conservation et/ou de restauration), il sera probablement nécessaire de conduire l'annotation sémantique non pas seulement au niveau d'une segmentation générale du bâti (colonne, porte, mur, etc..), mais au niveau des éléments architecturaux (base, fut, chapiteau, etc.) ainsi que des sous éléments (tore, baquette, scotie, etc.).

Ceci indique bien, que dans ce domaine, l'étude et la description des éléments, doit nécessairement s'appuyer sur un support de représentation plus sophistiqué, capable d'expliciter la complexité morphologique d'un édifice (relation entre ses parties et l'ensemble) ainsi que les caractères formels spécifiques aux éléments qui le composent (aspects spécifiques reliés à l'apparence visuelle des surface).

Or, dans le domaine du relevé et de la documentation graphique du patrimoine architectural, des techniques et des outils (scanner laser 3D, photogrammétrie,

photomodélisation, etc.) permettent aujourd'hui de collecter, de structurer et de traiter des informations acquises (réseaux de points, nuages de points, photographies, etc.) afin de construire des représentation tridimensionnelles d'édifices patrimoniaux. Ceci correspond à créer un modèle 3D des artefacts étudiés contenant le niveau d'information (morphologie, apparence visuelle) nécessaire à une analyse complète.

Une hypothèse, introduite par [Blaise J.Y. et al., 2004] considère que le modèle 3D interactif peut être considéré comme un moyen d'accès privilégié à l'information patrimoniale. En s'inscrivant au sein de ce concept, nous avons alors la possibilité de traiter notre problème dans un contexte bien plus large : **on peut utiliser la représentation de la morphologie d'un édifice comme source principale pour distribuer / propager les attributs sémantiques sur l'ensemble des sources iconographiques (le représentant en deux dimensions) qui se trouvent dans une condition de cohérence géométrique/spatiale avec son modèle tridimensionnel.**

En lisant le problème dans cette nouvelle dimension, notre approche peut être ainsi décrite : au lieu de créer une relation directe entre chaque région de chaque image 2D et une annotation sémantique, nous pouvons utiliser un modèle de description sémantique (structurant l'ensemble de termes pouvant décrire la morphologie d'un édifice) comme dénominateur commun permettant d'établir des relations entre la représentation 3D complète des formes architecturales qui composent l'édifice et l'ensemble des sources iconographiques (collections d'images 2D segmentées) correspondantes (voir figure 28).

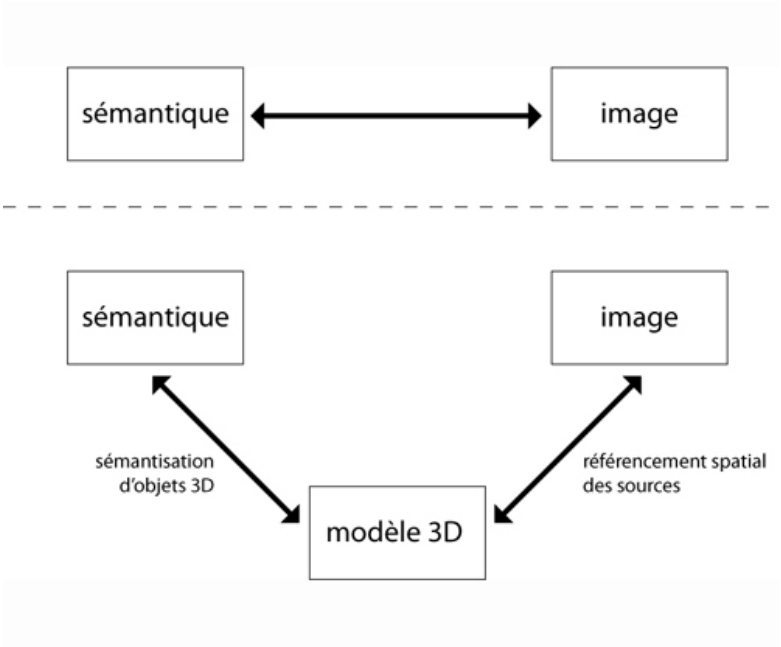


Figure 28 Différence entre les approches actuelles de sémantisation d'image (en haut) et l'approche proposée (bas)

Ceci correspond à mettre en place une méthodologie complète de caractérisation sémantique de représentations 2D et 3D d'édifices en s'appuyant sur la notion de référencement spatial : la détermination des relations projectives existantes entre la morphologie d'un artefact (modèle 3D) et l'ensemble des ses représentations

iconographiques (images 2D). La relation projective établie sera ensuite exploitée pour véhiculer des annotations sémantiques de la représentation 3D à l'ensemble de représentations 2D.

4.1. Approche proposé

Notre approche est composée de cinq étapes.

La première étape se concentre sur l'obtention d'une représentation tridimensionnelle complète de l'édifice à partir d'un relevé de terrain. Ce processus peut être divisé en sous étapes : l'acquisition des données à référence spatiale sur l'édifice réel et la construction de son modèle 3D à partir de ces données.

La deuxième étape consiste en la caractérisation sémantique du modèle 3D. Cette phase consiste à une décomposition de la représentation tridimensionnelle de l'édifice afin d'en isoler les éléments caractéristiques et de les qualifier par des attributs sémantiques. Dans cette phase on décrit la morphologie du bâti en fonction du vocabulaire de l'architecte (chapiteau, base, fenêtre, colonne, porte, etc.) en utilisant des thesaurus spécialisés.

La troisième étape se concentre sur l'établissement de la relation entre le modèle 3D et les images 2D, en s'appuyant sur le référencement spatial. Cette étape permet de retrouver le point de vue exacte (ou approximatif) de l'image dans l'espace 3D (voir figure 29). Différentes méthodes seront étudiées pour cette étape.

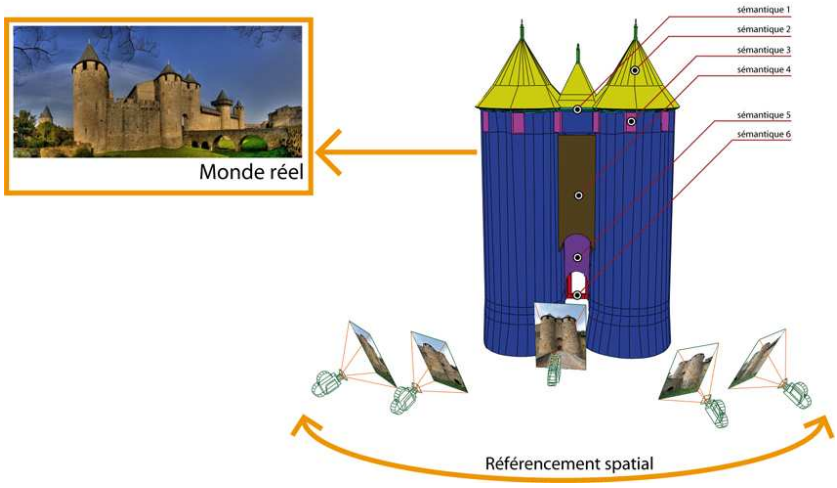


Figure 29 Les trois premières étapes de l'approches proposée : un modèle 3D est réalisé à partir d'une acquisition de l'édifice réel ; le modèle 3D est caractérisé du point de vue sémantique ; les images sont référencées spatialement par rapport au modèle 3D.

La quatrième étape consiste en la projection de la description sémantique du modèle 3D (décomposition) vers les images référencées spatialement (segmentation). Le résultat de cette étape se concentre sur en l'ajout d'une couche sémantique (projections des contours des formes 3D) sur les images (voir figure 30). Cette couché est réalisée par une représentation vectorielle (coïncidente avec l'image originelle) composée de polygones reliés à des attributs sémantiques.

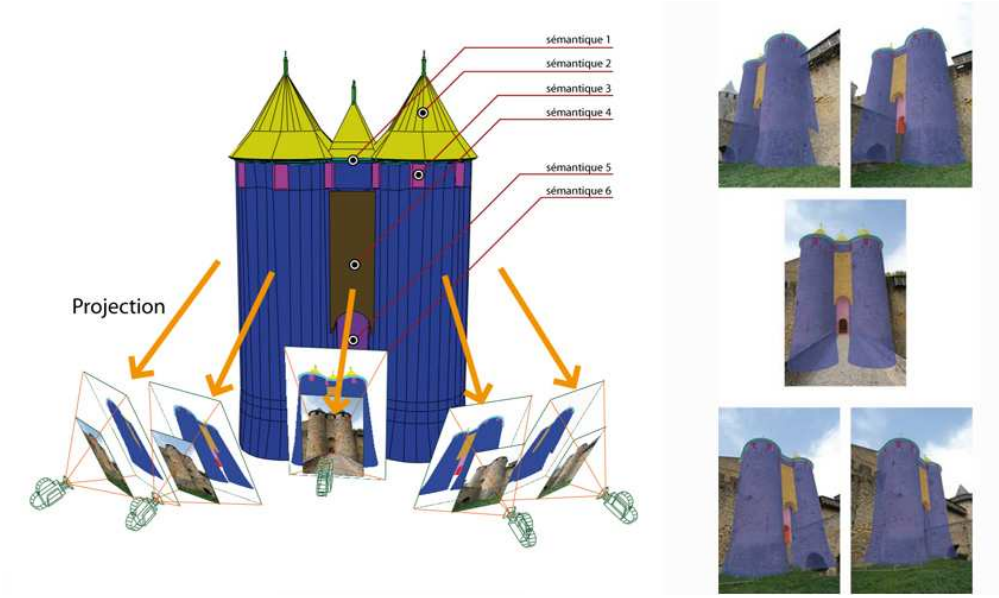


Figure 30 La quatrième étape de l'approche proposée : projection (du modèle 3D vers les images 2D) d'une couche sémantique.

Le grand avantage de notre approche consiste en la capacité à gérer des descriptions très complexes (c'est le cas de l'architecture patrimoniale) et surtout en fonction de multiples stratégies de description. L'annotation sémantique est en effet conduite directement sur la représentation tridimensionnelle, puis elle est systématiquement et automatiquement transférée dans les images référencées spatialement. La cohérence établie entre l'édifice réel (relevé 3D), sa morphologie (reconstruction 3D et structuration sémantique) et les images (référencement spatial sur le modèle 3D) permet la génération de contours particulièrement précis de formes (parfois très complexes) pouvant être projetées sur les images afin de les segmenter en fonction de la structuration du modèle 3D. De plus, la structure de description sémantique de chaque image serait toujours basée sur la même norme.

Un autre grand avantage de notre approche consiste en la facilité à modifier la structure de la description sémantique. Compte tenu du fait que chaque image est référencée spatialement par rapport au modèle 3D, a chaque changement au niveau de la décomposition de la morphologie de l'édifice ou au niveau de l'annotation sémantique des parties isolées, correspondra une mise à jour automatique de la segmentation et de l'annotation sémantique des images corrélées. Cette possibilité est particulièrement importante dans le domaine de la documentation architecturale, en effet, si d'un coté il y a souvent l'exigence de faire évoluer la description de l'édifice (soit en termes de granularité spatiale, soit en termes de choix des termes de vocabulaire), de l'autre coté plusieurs description des mêmes formes (décompositions et descriptions par termes) peuvent être conduites sur le même édifice en fonction de différentes exigences d'analyse (conservation, restauration, aménagement, valorisation ,etc.). Dans ce cas, il sera alors possible d'envisager des descriptions sémantiques multicouches (voir figure 31).

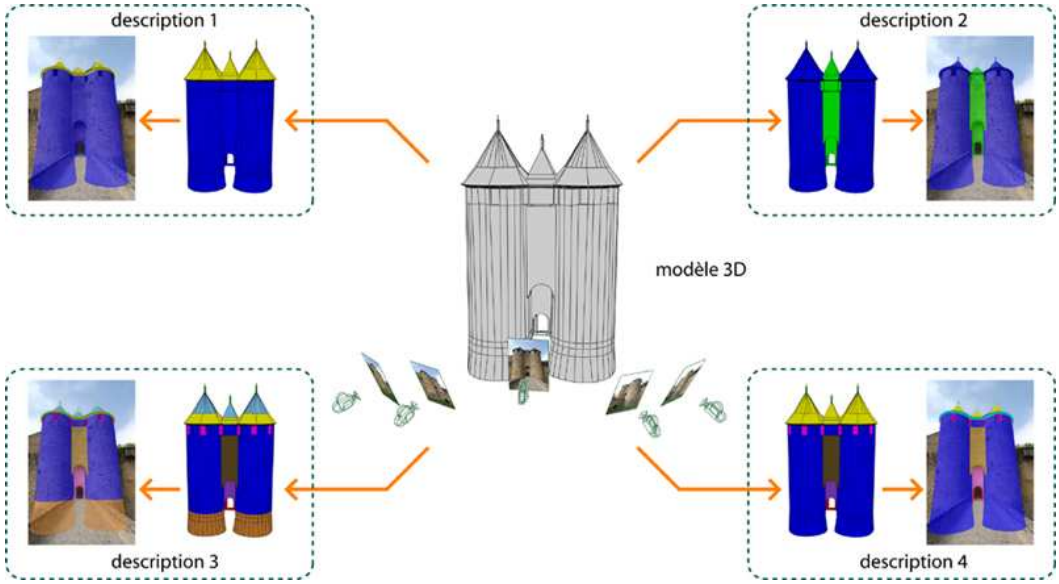


Figure 31 Un exemple de description sémantique multicouche : à chaque décomposition du modèle 3D corresponde une projection de couche sémantique différente sur les images.

La dernière étape de notre approche, se concentrera sur la définition et le développement informatique d'un système d'informations capable de rechercher visuellement les résultats du processus entier ci-dessus décrit en fonction de critères spatiaux, morphologiques et sémantiques. Il s'agit d'une application Web utilisable par des spécialistes du domaine de la documentation architecturale comme par le grand public.

4.2. Problèmes à traiter

A partir de l'approche proposée qui contient cinq étapes de démarche, le travail effectué prend en compte également cinq préoccupations distinctes.

4.2.1. Acquisition de données 3D et reconstruction géométrique basés sur la réalité (chapitre 5)

Notre approche utilise une acquisition de données 3D comme base de départ. Ces données doivent nécessairement être produite à partir du relevé d'un édifice réel afin de garantir la cohérence spatiale (avec différents niveaux d'approximation) avec les sources iconographiques (photographies, dessin, peintures, etc..) le représentant. La reconstruction géométrique à partir de ces données doit elle aussi être le plus cohérent possible avec la réalité afin de garantir (une fois structurée d'un point de vue sémantique) une projection correcte des couches sémantiques sur les images. Plus le modèle 3D sera approximé, plus la segmentation des images corrélées sera imprécise et incohérente.

Pour l'étape de relevé, l'approche devra alors prendre en considération l'utilisation (individuelle et ou complémentaire) des moyens actuels d'acquisition de données 3D (balayage laser et photogrammétrie) afin de définir leur niveau d'adaptation à notre objective.

Même pour l'étape de reconstruction géométrique, il sera nécessaire d'étudier les moyens existants le plus appropriés à notre problème spécifique. L'objectif de cette étape consiste en effet à identifier les moyens permettant d'élaborer une représentation 3D la plus cohérente possible avec la réalité et en même temps permettant une structuration et une annotation sémantique adaptée aux besoins spécifiques du domaine. Un dernier aspect concerne également l'évaluation des possibilités effectives d'exploitation de cette représentation, que ce soit en termes de manipulation, stockage et/ou visualisation.

4.2.2. Caractérisation sémantique de la représentation 3D (chapitre 5)

La logique des décompositions sémantiques d'un édifice est toujours contrainte à la description de modèles théoriques. Ces approches sont en effet fondées sur la structuration de formes symboliques. L'objectif de cette partie de notre travail consiste à ajouter une couche sémantique à la description de la morphologie d'édifices existants.

Notre approche devra alors prendre en compte l'étude des méthodologies existantes de caractérisation sémantique des représentations 3D en fonction des véritables exigences de description que la morphologie d'un édifice requiert.

4.2.3. Référencement spatial de sources iconographiques (chapitre 6)

L'approche devra prendre en compte l'utilisation d'espaces 3D comme dénominateur commun pouvant établir des relations projectives entre la représentation 3D et les représentation 2D de l'édifice étudié. Ceci permettra d'associer à chaque image 2D un modèle géométrique exprimant la position et l'orientation de son point de vue dans le référentiel du modèle 3D.

A ce fin, nous avons besoin d'étudier les méthodes existantes afin de choisir la ou les plus indiquées en fonction de différentes contextes d'utilisation (professionnel, amateur, etc.). Par ailleurs, il sera nécessaire de s'intéresser à la définition d'un modèle générique de représentation des points de vue des images en fonction de la diversité des sources iconographiques prises en considération (types de projection, niveaux de cohérence avec le modèle, etc.).

Enfin, on évaluera également le niveau d'approximation intrinsèque aux différentes techniques et/ou outils de référencement spatial d'images.

4.2.4. Indexation du contenu des sources iconographiques (chapitre 6)

L'approche devra définir en détail le processus de création de couches sémantiques à projeter sur les sources iconographiques à partir de la caractérisation sémantique du modèle 3D et en s'appuyant sur le référencement spatial. La projection de la silhouette bidimensionnelle de la structuration sémantique du modèle 3D se base en effet sur le point de vue (déterminé par la phase de référencement spatial) des sources iconographiques.

On pourra traiter ce problème de deux étapes. Dans la première étape on s'intéressera à la création d'une image vectorielle à partir de la décomposition de la représentation 3D et d'une projection des contours de chaque élément isolée. Dans ce contexte, nous étudierons une technique de rendu 2D permettant le transfert et la conservation d'attributs relevant de la structure de description morphologique de la représentation 3D.

Dans la deuxième étape, on se concentrera sur l'association entre la sémantique associée à la représentation 3D (reliée à des termes structurés dans des thesaurus spécialisés) et les régions polygonales projetés sur les images.

4.2.5. Recherche visuelle de sources iconographiques basée sur critères spatiaux, morphologiques et sémantiques (chapitre 7)

En ce qui concerne la conception et le développement d'un système d'informations capable d'intégrer, de gérer et d'exploiter l'enrichissement sémantique des sources iconographiques qu'il est possible de conduire au travers de l'approche que nous venons de décrire, nous nous intéresserons à l'identification des moyens qui permettent de profiter pleinement des apports produits par la sémantisation 2D/3D d'images.

Dans ce sens on se concentrera sur la création d'une interface permettant la recherche de sources iconographiques à partir de trois critères différents.

- Critère spatial : l'interface devra permettre de retrouver des sources iconographiques en fonction du point d'observation de l'utilisateur à l'intérieur d'une scène 3D.
- Critère morphologique : l'interface devra permettre de retrouver des sources iconographiques en fonction de la sélection d'un ou plusieurs éléments de la représentation 3D de l'édifice.
- Critère sémantique : l'interface devra permettre de retrouver des sources iconographiques contenant des éléments reliés à un ou plusieurs attributs sémantiques spécifiques.

L'interface devra également être d'utilisation facile afin de permettre un enrichissement effectif de la base de données iconographiques relative à l'édifice (par les spécialistes de la documentation comme par le grand public).

4.3. Développement informatique

Les apports méthodologiques de ce travail de thèse sont expérimentés et validés au sein de différents développements informatiques intégrés au projet NUBES [Nubes] mené au sein du laboratoire MAP-Gamsau.

4.3.1. Présentation du projet NUBES

Le projet NUBES concerne le développement d'un système d'informations à l'échelle architecturale qui prend en compte les relations que l'on peut établir entre la représentation de l'édifice (forme, dimensions, état de conservation, restitution hypothétique, etc.) et des informations hétérogènes concernant différents domaines (technique, documentaire, historique, etc.). La démarche proposée vise à organiser autour d'un modèle de description sémantique de multiples représentations (et informations associées) dans le but de définir un système pour l'observation pluridisciplinaire de l'édifice. La représentation numérique de la morphologie des édifices étudiés a pour vocation d'être exploitée pour la constitution de banques de données patrimoniales et devient alors le support privilégié de navigation et d'interactions avec les sources documentaires qui lui sont associées. Ces informations et les représentations géométriques multiples des objets constituant l'édifice peuvent être restitués de manière dynamique en fonction de différents points de vue utilisateurs.

Dans le cadre du projet NUBES, les développements informatiques abordés dans le cadre de notre thèse concernent spécifiquement le module « Nubes Imago », un système pour la recherche visuelle de sources iconographiques en fonction de critères spatiales, morphologiques et sémantiques. Les fonctionnalités implémentées dans cette application concernent :

- le téléchargement sur serveur (upload) de sources iconographiques ;
- le référencement spatial de sources iconographiques ;
- la manipulation dans l'espace de représentations tridimensionnelles d'édifices et des sources iconographiques superposées ;
- la recherche de sources iconographiques en fonction de critères spatiaux, morphologiques et sémantiques.

4.3.2. Présentation de l'environnement de développement

4.3.2.1. PHP/MySQL

PHP est un langage interprété exécuté du côté serveur. La syntaxe du langage provient de celles du langage C, du Perl et de Java. Ses caractéristiques principaux concernent la possibilité de d'inclure le script PHP au sein d'une page HTML et l'interfaçage avec MySQL, un SGBD (Système de Gestion de Bases de Données) disponible sur de nombreuses plateformes : Unix, Linux, Windows, MacOs X, Solaris, etc.

4.3.2.2. Virtools DEV

Le paradigme de développement spécifique de Virtools considère les objets 3D comme des composants individuels, indépendants des données associées. L'architecture résultante permet d'attribuer des comportements modulaires aux objets qui composent la scène. Les applications crées avec ce système sont directement exploitables comme exécutables ou comme plugin Internet.

Virtools DEV répond à tous les besoins de notre système car il peut afficher un rendu basé sur les points pour les nuages de points, un rendu filaire ou ombragé pour la maquette d'un édifice, mais aussi un rendu texturé pour l'affichage des photos dans l'espace. Virtools DEV est composé de 5 composants clefs : l'interface Graphique, le moteur à comportement, le SDK (un ensemble d'outils de développement) et le langages de script (VLS).

4.3.3. L'architecture du système

En relation à la structure générale de la plateforme NUBES, notre système se base sur une architecture en trois parties :

- une base de données développées en MySQL contenant les représentations 3D des édifices, les nomenclatures utilisées pour l'annotation sémantique et les sources iconographiques ;
- une scène 3D interactive pour l'affichage et la manipulation des représentations 3D et des superpositions image/modèle, développée en Virtools DEV ;
- une application web en PHP qui constitue l'accès utilisateur aux données stockées dans la base et fournit les passerelles nécessaires pour le dialogue entre la scène 3D et la base de donnée.

La scène 3D est intégrée à l'application Web sous forme d'un plug-in qui communique avec l'application en PHP/MySQL par des fonctions javascript.

L'application Web est basée sur le dialogue entre une base relationnelle SQL et la scène 3D par le biais de script php et java. En ce qui concerne l'envoie de commande de la partie Web vers la scène 3D nous utilisons des fonctions javascript spécifiquement programmées. Dans le sens inverse un bloc DEV (GetWebData) permet de formuler des requêtes (URL avec des variables) dans la base de données en passant par une page PHP. Pour gérer la synchronisation entre les représentations dans l'espace et les données stockées dans la base, nous dupliquons de façon dynamique les tableaux relatifs aux entités manipulées dans une base de données interne à la scène 3D.

5. Reconstruction 3D et structuration sémantique d'édifices patrimoniaux

Comme on l'a vu dans le chapitre précédent, notre approche se base sur une représentation tridimensionnelle de la morphologie de l'édifice basé sur le réel, c'est à dire sur l'acquisition et le traitement de données métriques de terrain. Il ne s'agit pas d'une prérogative absolue, mais il facile de comprendre que cet aspect conditionne forcement le degré de précision en phase de référencement spatial des sources iconographiques et de segmentation. Pour donner un exemple, prenons le cas le plus simples, celui d'une photographie d'un édifice. Plus d'approximation existera entre le modèle 3D et l'édifice réel, moins précise sera la superposition entre la photo et le modèle et l'annotation sémantique résultante.

L'élaboration de la maquette 3D d'un édifice (tel que construit) passe par l'expression géométrique des objets qui le composent puis par la détermination de l'aspect de surface de ses différentes parties [De Luca L., 2006]. Deux phases d'élaboration interdépendantes sont nécessaires :

L'acquisition des données spatiales. Il s'agit du relevé de la morphologie, des dimensions et des aspects de surface de l'édifice étudié. Cette phase peut utiliser différents dispositifs basés sur le principe de mesure par scanner laser 3D, ou sur la photogrammétrie. Le résultat de cette phase consiste en un nuage de points qui peut présenter des caractères différents de morphologie du bâtiment par rapport de la fonction de la modalité d'acquisition employée.

La reconstruction géométrique de la morphologie. C'est l'étape de modélisation qui consiste à construire le modèle géométrique de l'édifice en s'appuyant sur les mesures issues de la phase de relevé. Plusieurs techniques permettent une reconstruction automatique, semi-automatique ou manuelle des surfaces à partir des nuages de points. Ces techniques diffèrent en fonction des données d'entrée qu'elles peuvent traiter et du type de représentation géométrique qu'elles peuvent générer.

5.1. Acquisition informations spatiales

Cette phase du relevé peut utiliser des dispositifs d'acquisition directe (scanners laser 3D), qui restituent les points acquis sous forme de coordonnées dans l'espace 3D, ou des méthodes d'acquisition indirectes (photogrammétrie), qui se basent sur l'exploitation du support photographique pour extraire les coordonnées spatiales à l'aide de différentes techniques de restitution.

5.1.1. La photogrammétrie

« La photogrammétrie est la science ou la technique permettant d'obtenir des informations fiables sur l'espace naturel ou sur des objets physiques par l'enregistrement, la mesure et l'interprétation d'images photographiques …. » [AmericanSociety of Photogrammetry, 1980].

L'enregistrement photogrammétrique a utilisé pendant longtemps uniquement des chambres métriques et semi-métriques [Carbonnel M., 1968]. Aujourd'hui, des images numériques peuvent être acquises directement par une caméra à matrice DTC (en anglais CCD). Qu'elles soient employées en configuration mono ou stéréo, en photogrammétrie, les chambres appartiennent à trois catégories principales [Grussenmeyer P. et al., 2001] : chambres métriques, chambres semi-métriques, chambres à DTC.

Le principe fondamental sur lequel la mesure à partir des photographies se base est la déduction des paramètres d'orientation externes et internes de l'appareil photo [Kraus K. et al., 1997]. L'orientation interne consiste à fixer la position du centre de projection par rapport au support de l'image (pellicule ou capteur DTC) et à chiffrer la distorsion d'image. L'orientation externe détermine la position et l'orientation de la caméra dans l'espace. On détermine six paramètres, trois pour la position et trois pour l'orientation dans l'espace (voir détaille en 6.1 et 6.2).

La restitution photogrammétrique permet d'extraire des images un ensemble de coordonnées exprimées dans l'espace. Ces méthodes diffèrent en fonction de la configuration de l'acquisition : mono, stéréo ou multi images [Grussenmeyer P. et al.,2001] (voir figure 32).

Restitution à partir d'une seule image. La restitution à partir d'images isolées de parties planes d'un objet s'appuie sur la mesure d'un certain nombre de points. Pour développer un redressement (ou une rectification), on considère, par approximation, que la surface est définie uniquement par des plans identifiés par un minimum de quatre points d'appui. L'homographie est l'équation projective qui établie la relation entre le plan de l'objet et le plan de l'image.

Restitution à partir d'un couple d'images. La restitution à partir d'un couple d'images utilise une configuration stéréo. En phase d'acquisition, les chambres (avec une orientation interne connue et étalonnée et une distorsion négligeable de l'objectif) sont presque parallèles entre elles. Dans les solutions de restitution numérique, l'observation stéréoscopique consiste en un système d'asservissement en temps réel qui agit sur l'un des clichées et qui prend en compte les paramètres d'orientation internes et externes des perspectives. Ce dispositif commande les déplacements en translation du deuxième cliché en fonction de la position observée sur le premier de façon à ce que les deux points observés soient toujours homologues. Ce mécanisme permet à l'opérateur de réaliser des représentations bidimensionnelles de façades, ou des modèles en trois dimensions

Restitution à partir d'un bloc multi-image. La restitution à partir d'un bloc multi-images peut utiliser des photographies convergentes, horizontales, verticales ou obliques. La restitution se base sur l'ajustement des blocs d'images par la méthode

des faisceaux (conditions de colinéarité) à partir d'un ensemble de correspondances (points, lignes) sélectionnées sur les photographies. Un des avantages principaux des méthodes basées sur ce principe est la possibilité d'exécuter un auto-étalonnage de la chambre directement à partir de photos de l'objet. En effet, une première solution de calibration déterminée, l'ajout de correspondances permet d'affiner le calcul des paramètres (externes et internes) par minimisation d'erreur. Cette approche n'est donc pas limitée à des chambres métriques ou étalonnées ce qui la rend beaucoup plus flexible.

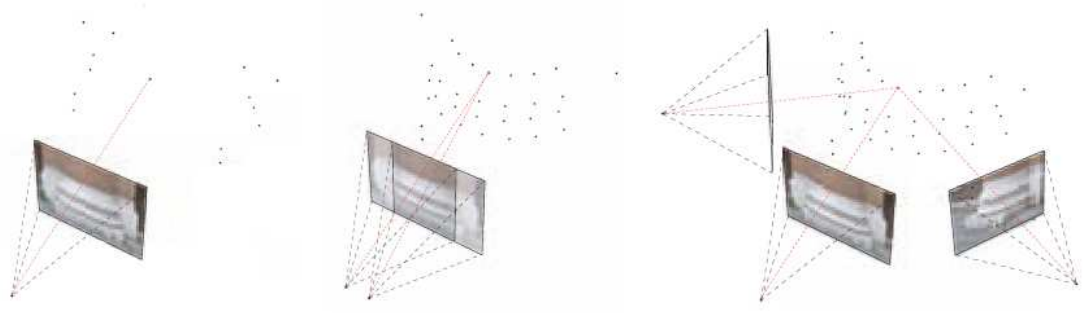


Figure 32 Schéma représentant les trois techniques de restitution photogrammétrique. De gauche à droite : la restitution à partir d'une seule image, d'un couple d'images et d'un bloc multi-image.

La restitution photogrammétrique à partir d'une seule image utilise des points d'appui préalablement mesurés qui sont nécessaires aux opérations de redressement. L'extraction de mesures supplémentaires est donc toujours contrainte à l'orientation du plan choisi pour la rectification. Cela rend cette technique une solution valable exclusivement pour la restitution des éléments sur un support bidimensionnel même s'il est possible (en exploitant une collection de points d'appui supplémentaire) d'assembler différentes restitutions bidimensionnelles dans l'espace pour restituer une composition tridimensionnelle.

Comme la restitution à partir d'une seule photo, la restitution d'un bloc d'images (dans les logiciels du marché) est une opération relativement manuelle (même si aujourd'hui des solutions automatiques commencent à se développer [Pierrot-Deseilligny M., 2007 ; Vu H. et al. 2009]) qui permet de déterminer la position d'un ensemble de coordonnées dans l'espace une fois sélectionnées par l'utilisateur sur les photos. Le choix de ces coordonnées est conditionné par la morphologie d'objet de la restitution. Il s'agit donc de sélectionner des informations pertinentes aux objectifs de la représentation (production d'un plan, ou d'un modèle 3D).

5.1.2. L'acquisition par scanner laser 3D

Le balayage laser terrestre est « l'utilisation d'un dispositif basé au sol, qui utilise un laser pour mesurer les coordonnées tridimensionnelles d'une région donnée de la surface d'un objet de façon automatique, dans un ordre systématique et à un taux élevé de vitesse, près du temps réel » [Boehler et al., 2002a]. Les modules de balayage laser permettent une acquisition directe de milliers de coordonnées par seconde. Au-delà des coordonnées spatiales (x,y,z), certains dispositifs de balayage

permettent aussi d'acquérir les valeurs d'intensité (niveaux de gris) et les valeurs colorimétriques (valeurs RVB) (voir figure 33).

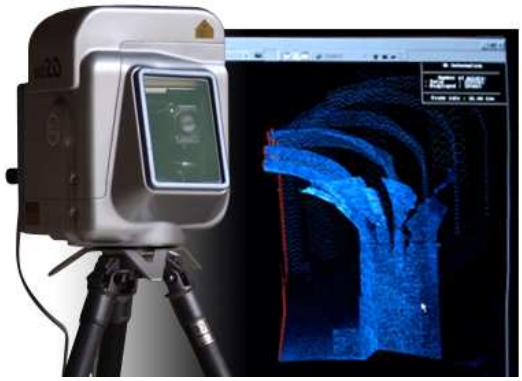


Figure 33 Un scanner laser « temps de vol » et résultat produit (nuage de point.)

Les modules de balayage laser 3D applicables au relevé d'édifices se divisent en deux catégories principales. Les scanners à longue portée mesurent des angles horizontaux et verticaux et mesurent la distance en calculant le temps de vol ou en comparant le changement de phase de l'onde transmise et reçue d'un signal modulé, et les scanners par triangulation incluent une base et calculent le point d'impact du faisceau laser en utilisant une ou deux caméras DTC (voir figure 34). Les deux principes sont différents soit en termes d'exactitude, soit en terme de portée. [Marbs A. et al., 2001] fournissent une description des principes de fonctionnement des scanners appartenant aux deux catégories.

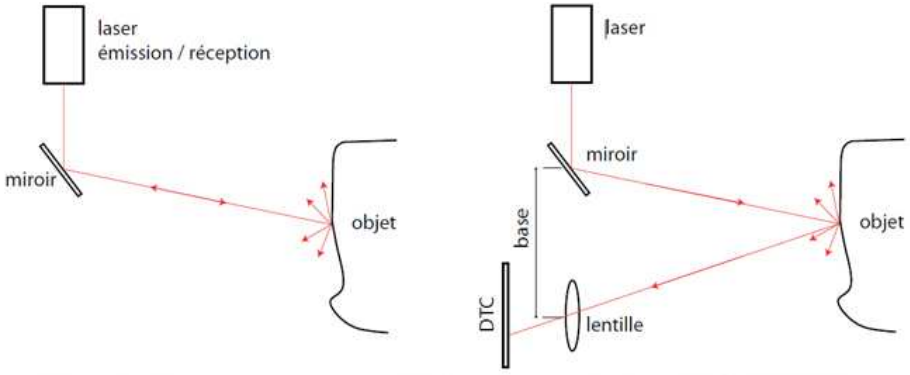


Figure 34 Illustration schématique des deux principes de fonctionnement des dispositifs de balayage laser 3D. À gauche le principe du temps de vol, à droite celui de la triangulation. [Marbs A. et al., 2001].

[Bohler W. et al., 2003] présentent une analyse comparative de la précision des scanner 3D. Dans la documentation du patrimoine architectural, un écart de quelques millimètres pour un seul point balayé n'est pas trop important si ce point fait partie d'un élément à géométrie régulière (plane, cylindrique ou analogues) et si le relevé sert juste à récupérer les paramètres qui décrivent cet élément dans une représentation bidimensionnelle. Par contre, si des surfaces irrégulières doivent être modélisées, les nuages de points bruités peuvent causer différents problèmes en phase de traitement.

Par conséquent, le procédé de balayage devrait être réalisé avec le dispositif le plus précis disponible pour la taille de l'objet à numériser et en fonction de la distance d'acquisition.

Les procédures de restitution peuvent produire les nuages de points trop denses, des disparités, des points non pertinents et des parties absentes sont souvent constatés dans les résultats. Des procédés du traitement des nuages sont alors nécessaires

Des traitements d'échantillonnage des nuages selon différents critères permettent de réduire la densité d'une zone du nuage (par critère spatial ou séquentiel) ou d'extraire des entités caractéristiques (analyse des discontinuités ou des valeurs d'intensités). On peut extraire aussi des informations pertinentes de façon semi-automatique ou manuelle.

Une simple acquisition (à partir d'un seul point de vue) n'est jamais suffisante pour enregistrer un ensemble patrimonial. Différentes acquisitions sont alors nécessaires, surtout quand le but est d'enregistrer un bâtiment entier, ou un site, à plusieurs niveaux de détail. Il est donc nécessaire de transformer les nuages de points collectés séparément dans un système de coordonnées unique. Trois stratégies sont possibles [Boehler W. et al., 2002b] :

- La collecte des données peut être accompagnée par une saisie de points d'appui au moyen d'un autre système de relevé (par exemple une station totale avec des coordonnées et une orientation connues) ;
- Des cibles (naturelles ou artificielles) peuvent être utilisées pour transformer les données acquises localement en fonction d'un système de coordonnées commun ;
- Des procédures automatiques de superposition de surfaces obtenues par maillage automatique (uniquement pour les objets de petite taille comme les sculptures ou les décors) peuvent être utilisées.

Pour les objets à l'échelle architecturale, la technique utilisée le plus fréquemment, est celle qui utilise des cibles. Ces points d'appui peuvent juste être des parties reconnaissables de l'objet (par exemple des coins) ou des cibles spéciales (sphères, cibles planes avec une réflectivité élevée) placées à des endroits judicieusement choisis. En ce qui concerne la première méthode, si la région acquise est assez grande, il y aura beaucoup de points disponibles pour assurer la robustesse géométrique de l'opération de consolidation. Malgré cela, il faut considérer que la propagation d'erreur peut être très importante quand il s'agit de considérer plus de deux stations. Dans la deuxième méthode, désormais supportée de façon complètement automatique, trois ou quatre cibles artificielles sont suffisantes pour obtenir une consolidation précise. Ces cibles pourront après être acquises aussi par des méthodes topographiques (théodolite, station totale) dans le cas d'un référencement local ou par des méthodes géodésiques (GPS - Global Positionning System) pour un référencement global.

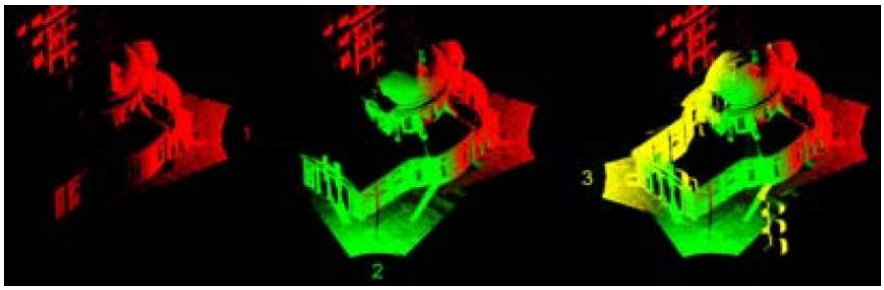


Figure 35 Consolidation d'un nuage de points issu d'un balayage laser 3D. Dans [De Luca L., 2006].

5.2. La reconstruction géométrique de la morphologie

Les procédures d'acquisition d'information ne prennent pas en compte habituellement les conditions géométriques de l'objet, il est donc nécessaire de transformer les nuages de points (aléatoirement) produits en structures polygonales ou en surfaces de haute qualité et sans perte d'informations importantes.

Aujourd'hui il y a plusieurs manières de créer un modèle 3D basé sur une architecture de monde réel. Par exemple un maillage automatique de nuage de point rassemblée par un outil de balayage laser ou modélisation photométrique créée des photos numériques ou de photo stéréo etc.

5.2.1. Reconstruction 3D à partir de nuage de points

Habituellement, les nuages de points ne sont pas suffisants comme résultat final de la documentation. Ils pourraient être considérés comme des moyens d'archiver la géométrie d'un objet en prévision d'exploitations futures. Dans ce but, beaucoup de techniques (automatiques, semi-automatiques et manuelles) existent aujourd'hui. Nous allons les présenter dans cette section en les regroupant en deux catégories principales :

La reconstruction automatique de surface: il s'agit de procédures d'interpolation ou d'instanciation automatique ou semi-automatique qui utilisent comme données d'entrée un nuage de points issu d'un balayage laser;

La modélisation interactive: il s'agit de techniques d'approximation manuelles qui décrivent les surfaces au travers de la manipulation d'un certain nombre de primitives géométriques et de fonctions de modélisation.

5.2.1.1. Les procédures de reconstruction automatiques

Le maillage polygonal est habituellement la méthode la plus adaptée pour représenter correctement les résultats des mesures, fournissant une description cohérente avec les données d'entrée. Beaucoup de méthodes ont été développées [Mencl R. et al., 1998] pour créer une représentation (triangulaire) régulière et continue de maille à partir d'un nuage de points (voir figure 36). Une fois la surface polygonale déterminée, diverses techniques de post-traitement peuvent être employées pour l'optimisation du résultat ; lissage, remplissage de trous, etc. La conversion des données mesurées en une surface polygonale cohérente est généralement basée sur quatre étapes:



Figure 36 Un scanner laser 3D à triangulation optique et le résultat obtenu (maillage polygonal d'un élément sculpté).

Prétraitement. Dans le cas des données acquises par un module de balayage, cette étape est obligatoire afin de réduire le volume de données à traiter et donc le temps de calcul (cas d'une surface acquise à partir de points de vue multiples) et d'éliminer une certaine quantité d'erreurs produites par le module de balayage [Floater M. et al.1998]. Pour la réduction du bruit de mesure dans un nuage issu de balayage laser, des méthodes statistiques sont appliquées prenant en compte la courbure extérieure du nuage et essayant de préserver les points mesurés. Pour les nuages obtenus à partir d'une corrélation d'images, les correspondances fausses peuvent être enlevées automatiquement [Borghese A. et al., 2000] ou manuellement par une inspection visuelle.

Détermination de la topologie globale de la surface de l'objet. les relations de voisinage entre les parties adjacentes de la surface doivent être dérivées. Cette opération a besoin typiquement d'informations additionnelles (normales) pour la prise en compte de contraintes de discontinuité, dans le but de préserver principalement la définition des bords

Maillage automatique. Des mailles triangulaires sont créées répondant à certaines exigences de qualité. Une triangulation convertit l'ensemble des données d'entrée en modèle polygonal cohérent (maille). Cette opération divise les données d'entrée et produit des sommets, des arêtes et des faces (représentant la surface) qui

se relient entre elles sur les bords partagés. Des méthodes de calcul par éléments finis sont employées pour discrétiser le domaine d'entrée en la divisant en ensemble de petits éléments, typiquement des triangles ou des quadrilatères en deux dimensions ou des tétraèdres en trois dimensions. Une triangulation optimale est définie en mesurant les angles, la longueur des arêtes et la hauteur de l'aire des éléments où l'erreur de l'approximation par éléments finis est normalement reliée à l'angle minimal des éléments [Edelsbrunner H., 2001]. Les sommets de la triangulation peuvent être exactement les points d'entrée ou les points supplémentaires, appelés les points de Steiner, qui sont insérés pour créer une maille plus homogène. Le maillage polygonal peut être exécuté en 2D, 2D et demi ou en 3D, selon le type des données d'entrée [Remondino F., 2003].

Post-traitement. Une fois le modèle créé, des opérations d'édition sont généralement appliquées pour raffiner et perfectionner la surface polygonale résultante. Les polygones créés ont besoin habituellement de quelques améliorations pour corriger les imperfections ou les erreurs engendrées par les procédures de maillage automatique. Ces opérations (principalement manuelles) se basent sur la manipulation de simples triangles

- Correction de bords en présence de triangles divisés ou renversés ;
- Insertion de triangles en présence de trous.

Le modèle polygonal peut également être amélioré en ajoutant de nouveaux sommets et/ou en ajustant la position des sommets.

5.2.1.2. La modélisation interactive

Si l'objet balayé est composé d'un ensemble de primitives simples (tels que des plans, cylindres, cônes, sphères...) les primitives géométriques correspondantes peuvent être adaptées aux différentes parties du nuage de points. La reconstruction géométrique résultante sera par contre réduite à une représentation plus approximative. Les solutions commerciales disponibles pour le traitement du nuage de points diffèrent considérablement en ce qui concerne les outils qui exploitent cette technique.

D'autres solutions peuvent s'appliquer à la reconstruction de structures beaucoup plus compliquées. Des procédés manuels permettent enfin l'extraction de profils en utilisant des plans en intersection avec le nuage de points. Ces plans isolent les points compris dans une épaisseur (réglable) et offrent un support pour la reconstruction du profil décrit par des lignes ou des courbes.

Les fonctionnalités communes aux logiciels de CAO permettent aujourd'hui de manipuler ces lignes et courbes dans l'espace. À partir des courbes et/ou de simples coordonnées, plusieurs fonctionnalités ont été implémentées dans les logiciels de modélisation géométrique pour générer des surfaces. Certains d'entre eux permettent de manipuler des surfaces NURBS (Non-Uniform Rational B-Splines). Les procédures seront décrites de façon générale indépendamment du type de génératrice utilisée (poly-ligne ou courbe) et du type de représentation géométrique résultante (filaire, par frontières ou paramétrique). La figure 37 montre une restitution tridimensionnelle

réalisée à partir d'un ensemble de points pertinents et conduite par des procédures de modélisation manuelles.

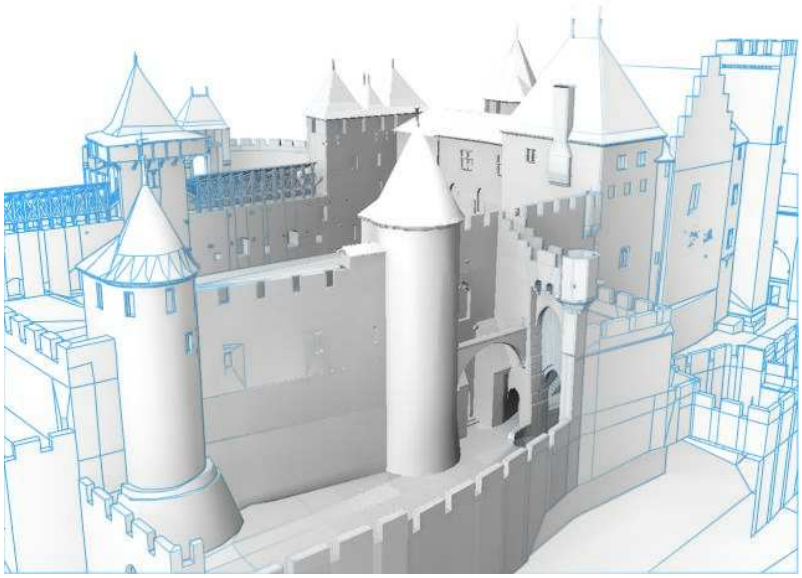


Figure 37 Modèle 3D réalisé par des procédures de modélisation interactive. Château Comtal de Carcassonne.

5.2.2. Reconstruction 3D basée sur image

La photographie est un support bien adapté à l'interprétation des formes. Les approches et les techniques basées sur ce principe peuvent potentiellement éliminer la phase laborieuse normalement requise par la modélisation de structures géométriques détaillées. En effet, un des bénéfices majeur de ces techniques réside dans leurs capacités à restituer de manière fine les effets de textures et les détails liés aux imperfections des objets réels, qui sont très difficiles à obtenir à l'aide des techniques conventionnelles de synthèse d'image [Foley J. et al., 2000]. La solution offerte par ces systèmes consiste à remplacer les modèles conventionnels avec des représentations géométriques approximatives enrichies par l'image [Shum H.Y. et al.,2000]. En fonction de la quantité d'informations géométriques utilisées pour la reconstruction et du nombre d'images nécessaires pour la reconstruction.

Ces techniques basent la procédure de reconstruction 3D à partir d'un bloc d'images non calibrées et en deux phases:

Dans une première phase, on établi les correspondances entre différentes prises de vue. Ces correspondances peuvent être des points (ImageModeler, Photomodeler) ou des lignes présumées perpendiculaires ou parallèles [Cipolla R. et al., 1999] sélectionnées sur les images. Certains systèmes utilisent des techniques éprouvées de vision par ordinateur afin d'automatiser le plus possible cette tache [Faugeras O., 1993]. De là, les matrices de projection des images peuvent êtres calculées à une transformation projective près grâce aux propriétés de la géométrie. Ces procédés permettent de déterminer les paramètres externes et internes des appareils photo (voir section 6.2).

La deuxième phase consiste en une véritable reconstruction des éléments de la scène en utilisant des primitives géométriques et des fonctions de déformation. Quelque solution commerciale exploite les propriétés des représentations par frontières (Brep) : les transformations simples (translation, rotation, changement d'échelle) peuvent alors être affectées à la primitive entière aussi bien qu'à une ou plusieurs de ses composantes (faces, arêtes, sommets). La figure 38 montre la restitution tridimensionnelle d'un édifice obtenue par une démarche de photo-modélisation.

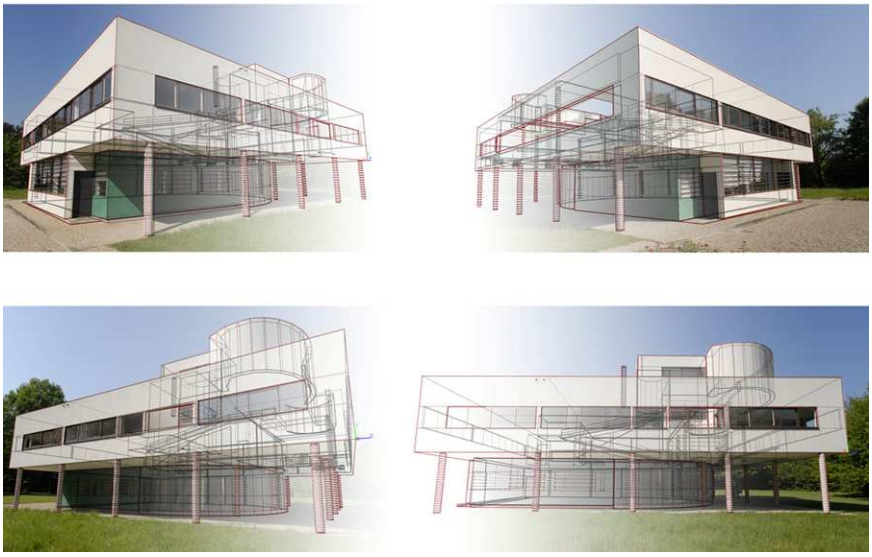


Figure 38 Exemples de modélisation à partir d'images.

5.3. Caractérisation sémantique du modèle 3D

La phase de relevé est juste le commencement du long processus qui caractérise la rédaction des documents montrant l'état actuel d'un édifice ou l'analyse des transformations qu'il a subies pendant son histoire. Différents traitements visent en effet, non seulement à collecter et à organiser les informations mais surtout à produire de la documentation pour différentes analyses spécifiques. Il apparaît donc essentiel d'étudier les conditions au travers desquelles des représentations produites à partir du relevé, pourraient être organisées autour d'un modèle de description de l'édifice. Cette problématique revient à considérer le modèle 3D comme moyen d'accès privilégié à l'information. Ceci impose, à partir d'une analyse sémantique de l'objet, d'isoler une structure d'entités dont les représentations peuvent être organisées en fonction de plusieurs points de vue.

La logique des décompositions sémantiques d'un édifice est toujours contrainte à la description de modèles théoriques. Ces approches sont en effet fondées sur la structuration de formes symboliques. L'objectif de cette partie de notre travail consiste à ajouter une couche sémantique à la description de la morphologie d'édifices existants. [Blaise J.Y., 2003] a montré que pour l'analyse architecturale, les approches de classification par intension (association d'un modèle théorique à un élément réel) sont incompatibles avec les exigences de l'étude du patrimoine construit, alors que des

approches par extension (extraction et groupage de descripteurs communs à plusieurs éléments) semblent plus pertinentes.

L'élément d'architecture, en tant que forme symbolique, appartient en effet principalement au domaine de la représentation mentale [Goulette J.P., 1999] : c'est un objet abstrait pour penser et concevoir un projet d'architecture. Le vocabulaire de l'architecte organise en effet le bâti selon une structure signifiante qui n'est pas seulement caractérisée par une apparence visuelle ou une nature géométrique. De plus, il faut remarquer que les connaissances qui interviennent dans un découpage sémantique de la morphologie d'un édifice sont toujours relatives à un objectif d'analyse.

5.3.1. Travaux relatifs

La sémantisation des représentations d'édifice suit la notion de grammaire des formes proposée par [Stiny G. et al., 1979] et [Stiny G., 1975].

Le premier exemple de la modélisation 3D basé sur une classification sémantique a été introduit par [Quintrand P. et al., 1985]. Les auteurs indiquent que : pour associer de la sémantique à la représentation d'une forme architecturale il est nécessaire de considérer le bâtiment comme une collection d'objets élémentaires.

Plusieurs recherches se sont concentrées sur le développement des classifications des éléments architecturaux dans les cadres théoriques [Tzonis A. et al., 1986] ou dans des applications de la modélisation géométrique [Gaiani M., 1999]. La modélisation sémantique a été également utilisée dans des applications de modélisation automatique de tissus urbains [EmGard K.L. et al., 2008] ou d'éléments architecturaux répétitifs [Mueller P. et al., 2006].

[Attene M. et al. 2007] ont développé ShapeAnnotator, un système modulaire pour caractériser d'un point de vue sémantique des représentations d'objets numérisés (sous forme de maillage polygonaux) au travers d'un processus d'annotation sémantique basé sur une ontologie. Pour la caractérisation sémantique, les auteurs utilisent des solutions de segmentation (automatiques et manuelles) des surfaces puis des comparaisons avec une base de données correspondantes.

Dans le domaine spécifique de l'architecture patrimoniale, la méthode définie par [De Luca L., 2006] permet d'associer des attributs sémantiques à un modèle de description conceptuel de la morphologie d'un édifice (représenté sous forme de graphe de relation) qui structure des multiples représentations des éléments architecturaux. Telle approche considère l'édifice comme un système de connaissances architecturales à partir duquel il est possible d'extraire un modèle sémantique qui permet, en fonction des points de vue spécifiques, de définir diverses représentations de la forme (voir figure 39).

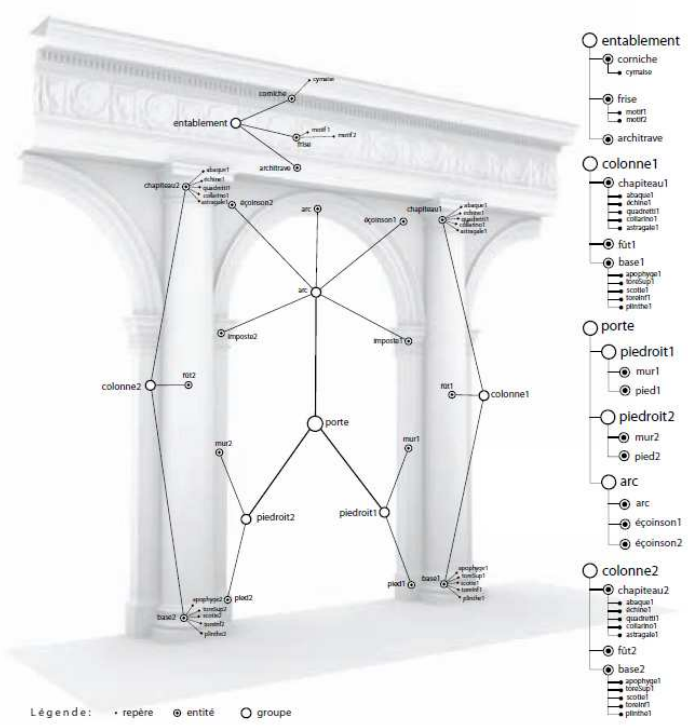


Figure 39 Graphe de description sémantique associé au modèle 3D [De Luca L., 2006]

[Manferdini A.M. et al., 2008] proposent un système de classification sémantique pour l'étude de sites archéologiques. Le modèle géométrique est segmenté de façon semi-automatique et classées selon les règles archéologiques et architecturales. La classification sémantique des découvertes est ensuite lié à une base de données archéologiques (voir figure 40) indiquant entre autres la provenance parties qui composent les objets architecturaux (originales ou appartenant à d'autres éléments).



Figure 40 Classification sémantique d'objets archéologiques. Dans [Manferdini A.M. et al.,2008]

A l'échelle urbaine, [Grussenmeyer, P. et al., 1999] proposent un principe dans lequel la structure sémantique associée à la géométrie est la base de la modélisation. L'espace est décomposé en différents concepts sémantiques, chaque concept comprenant une structure sémantique inhérente. Cette structure est basée sur

connaissances des règles de composition qui peuvent être programmées pour aider la reconstruction semi-automatique de différents objets géométriques. La restitution à partir de photos aériennes permet de procéder à la modélisation géométrique et sémantique des objets urbains, en incluant les parties visibles du monument vu de dessus. Ensuite, la photogrammétrie terrestre peut restituer les détails géométriques du monument en considérant ses aspects sémantiques (voir figure 41).

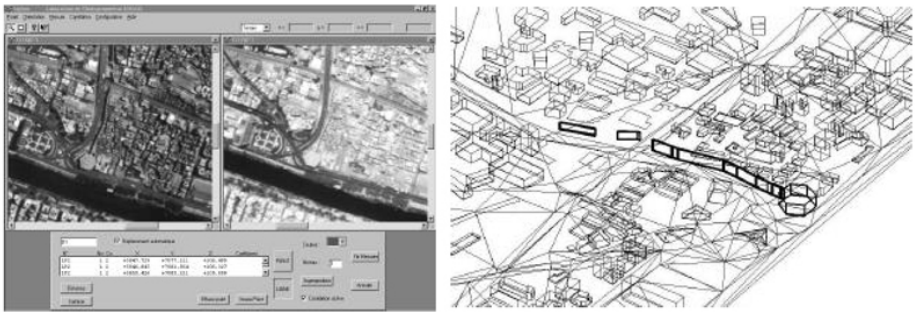


Figure 41 L'interface d'un système pour la modélisation géométrique et sémantique d'objets urbains [Grussenmeyer, P. et al., 1999]

Du coté de l'annotation sémantique automatique de formes, la solution optimale pour annoter des modèles 3D serait d'utiliser un algorithme de segmentation forme capable de détecter automatiquement toutes les caractéristiques conceptualisées par une ontologie. Cette approche est loin d'être opérationnel pour l'instant. En effet, les algorithmes de segmentation existants se basent principalement sur des critères géométriques. Une analyse du problème de la segmentation de maillages polygonaux est fournie par [Shamir A., 2006], une comparaison des solutions existantes est traitée par [Attene M. et al., 2006].

5.3.2. Solution retenue

Comme on l'a vu dans la section 5.3.1 de ce chapitre, il existe aujourd'hui beaucoup de recherches qui traitent la problématique de l'annotation sémantique de représentations 3D. Notre travail se situe en continuité de l'approche développé par [De Luca L., 2006] qui définie un procédé complet pour obtenir des représentations tridimensionnelles structurés du point de vue sémantique à partir du relevé d'un édifice réel. Cette solution implique trois phases : premièrement la classification des composantes d'un édifice selon un point de vue (c'est à dire en fonction d'un vocabulaire, ou d'une analyse spécifique), deuxièmement la décomposition morphologique de l'édifice en composantes élémentaires et troisièmement la création d'associations entre concepts et les formes à l'aide des graphes de description.

La notion de point de vue comprend trois aspects :

- L'utilisation d'un vocabulaire pour isoler des éléments ;
- L'identification d'une logique pour le regroupement de ces éléments dans un ensemble cohérent ;

- Le choix d'un type de représentation capable d'extraire de la forme les informations nécessaires à l'analyse.

En accord avec les principes introduits par [Falcidieno B. et al. 1998], un modèle de description (définit par sémantique et structure) est construit autour de la morphologie de l'édifice.

Une première phase du travail est dans le « niveau sémantique ». Elle consiste à isoler des concepts « termes de la description » et à lui associer des parties de la forme. La seconde phase « niveau structurel » permet d'établir un graphe de relations entre ces concepts pour organiser les éléments de la scène par rapport à une exigence de description. Le dernière « niveau de représentation » permet d'associer à chaque concept isolé une ou plusieurs représentations.

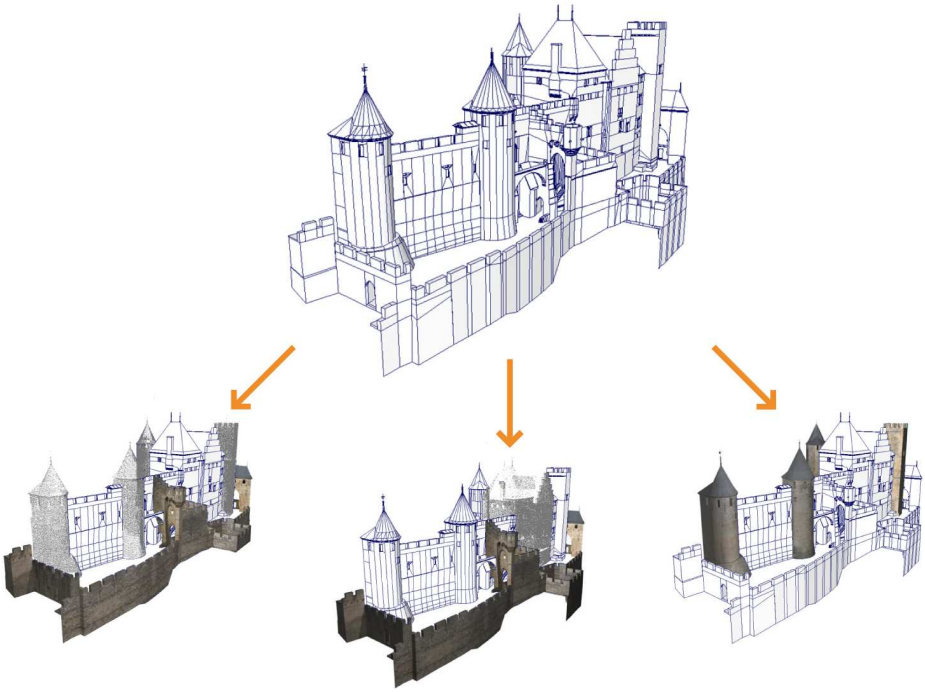


Figure 42 Des multiples représentations du modèle 3D élaborées à partir du même modèle de description sémantique.

Cette méthode offre une liberté importante dans l'exploitation des données issues du relevé de l'édifice à partir du même modèle géométrique de la morphologie, on peut élaborer différentes descriptions sémantiques qui exploitent un système multi-représentations (voir figure 42).

Comme on le verra dans la suite, dans le cadre spécifique de notre travail de thèse, ce modèle de description de l'édifice n'est pas seulement utilisé pour la structuration de la maquette tridimensionnelle, mais il devient le dénominateur commun pour l'établissement de relations bilatérales entre les régions des images (sources iconographiques) et les annotations sémantiques associées (voir section 6.4).

Afin de constituer un cadre d'expérimentation à notre thèse, nous avons structuré le modèle tridimensionnel du château Comtal de Carcassonne (reconstruit à

partir d'une utilisation conjointe de balayage laser 3D et de photogrammétrie) selon l'approche de [De Luca L., 2006] en nous s'insérant dans le développement de la plateforme NUBES. Pour supporter la phase de description sémantique de la morphologie de l'édifice, on utilise une représentation symbolique, dans l'espace 3D, des relations (partie/ensemble) établies entre les éléments isolés. Il s'agit d'un graphe 3D (arborescence dans l'espace) dont la configuration dépend de la manipulation (relations hiérarchiques) d'un ensemble de termes dans une liste.

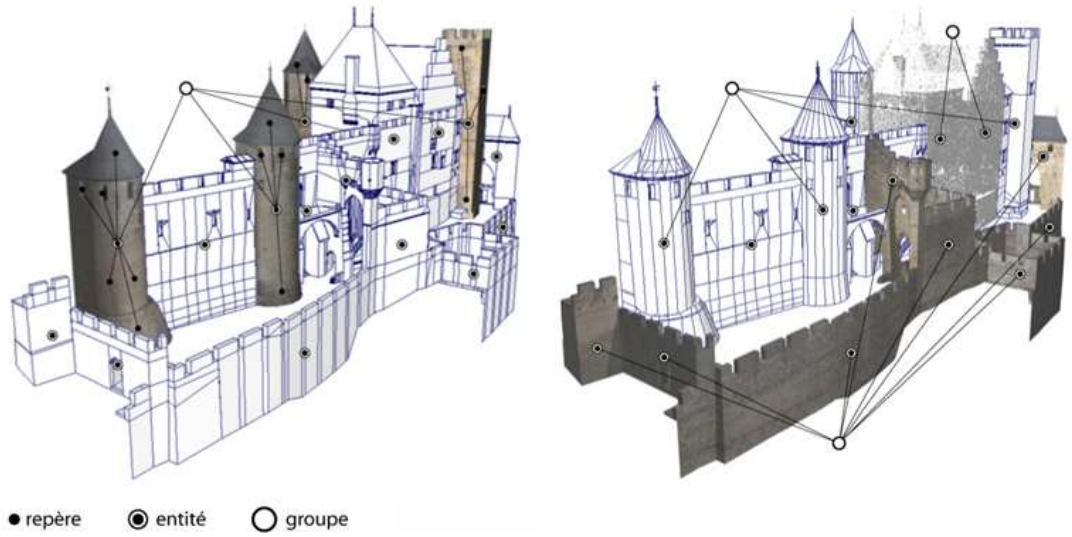


Figure 43 Des exemples de structuration sémantique du modèle élaborés en organisant les éléments architecturaux en fonctions des concepts de « entité morphologique », « groupe finalisé » et « repère ».

Ce graphe est construit à l'aide de trois concepts : entité morphologique, groupe finalisé, repère (voir figure 43).

Une *entité morphologique*, qui résulte d'une décomposition morphologique, constitue un concept identifié par l'utilisateur pour la description de l'édifice qui peut être associé à une ou plusieurs représentations géométriques.

Un *groupe finalisé* est un nœud qui regroupe des entités morphologiques et qui n'a pas de représentation géométrique propre. Son extension spatiale résulte en effet de l'union des enveloppes englobant les entités qui lui appartiennent.

Un *repère* indique un aspect particulier à l'intérieur d'une entité. Par exemple une moulure d'un élément ou un aspect particulier sur la surface de l'entité (personnage historique d'un bas-relief, dégradation du matériau, etc.).

A partir de la structure du modèle 3D (décomposition morphologique) on produit une première description en associant à chaque concept isolé un nœud de type « entité morphologique » dans une liste de termes et en générant le symbole correspondant dans l'espace 3D. Ce symbole est situé au barycentre de l'enveloppe englobant l'entité. L'utilisateur peut alors organiser hiérarchiquement les entités sur la liste de termes (voir figure 44) en introduisant des nœuds de type « groupe finalisé » ou « repère ».

La création d'un groupe correspond à la création d'un concept qui réunit un ensemble d'entités : dans le graphe un symbole 3D est alors créé au barycentre de l'enveloppe englobant toutes les entités morphologiques sélectionnées. La création d'un repère est au contraire conduite manuellement en sélectionnant dans l'espace 3D un point sur la surface d'une entité morphologique. La correspondance établie entre un éditeur (liste de termes) et la manipulation des entités dans l'espace permet de conduire la structuration hiérarchique dans la scène 3D et/ou dans la liste d'entités. La représentation des relations dans l'espace est en effet calculée en temps réel par une procédure qui génère une chaîne de liaisons 3D en fonction de la lecture des niveaux hiérarchiques exprimés dans la liste des termes.

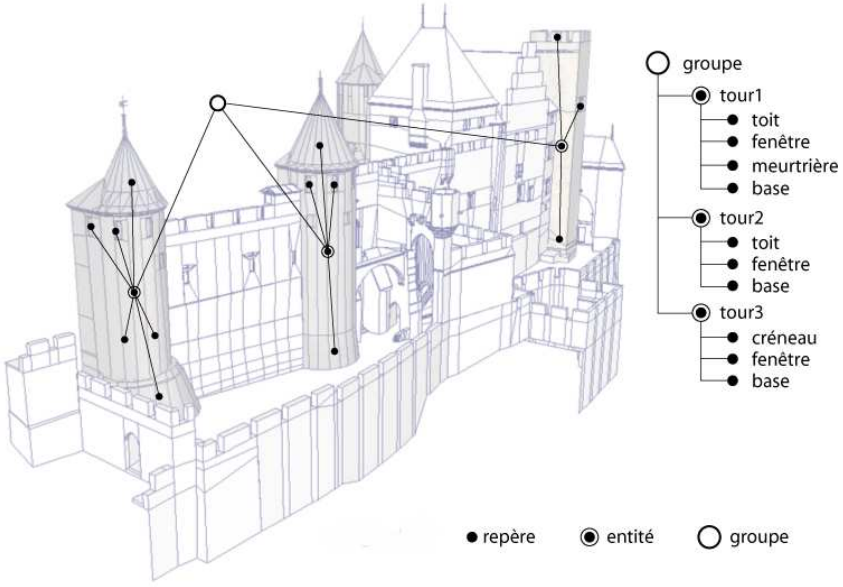


Figure 44 Termes de descriptions associés à la morphologie de l'édifice organisés selon les relations hiérarchiques établies dans le graphe de description.

6. Référencement spatial et sémantisation de sources iconographiques

Dans le chapitre précédent nous nous somme intéressé à la reconstruction de la morphologie de l'édifice afin de produire une représentation tridimensionnelle enrichie du point de vue sémantique. Dans ce chapitre nous nous concentrerons sur les méthodes permettant de mettre en relation spatiale des sources iconographiques (sous forme d'images 2D) par rapport à la morphologie de l'édifice reconstruite. Comme on le verra dans le chapitre suivant, cette relation spatiale servira aussi pour transférer les attributs sémantiques associés au modèle 3D sur les images 2D. C'est le point central de notre approche : l'utilisation d'une description sémantique de l'édifice (voir section 5.3.2) comme dénominateur commun qui relie la représentation 3D des formes architecturales aux sources documentaires bidimensionnelles.

Ce chapitre se divise en deux partie principales : dans la première partie nous nous intéressons aux méthodes actuelles pour le référencement spatial d'images 2D sur des ensembles de coordonnées 3D, c'est à dire de superposition spatiale entre l'image et le modèle 3D. Dans la deuxième partie de ce chapitre nous nous concentrons sur les modèles géométriques permettant d'exprimer les modèles géométriques des sources iconographiques dans l'espace 3D. Cette opération permet d'exprimer la position et l'orientation d'une image en fonction d'un référentiel, notamment celui du modèle 3D de l'édifice. Une analyse des différents modèles géométriques à utiliser ainsi que des différents niveaux d'approximation à prendre en compte (en fonction du type d'iconographie) permettra enfin d'identifier les méthodes le plus indiquées pour la spatialisation de différents types d'images (photos, vidéos, panoramiques, dessins, peintures, etc.).

6.1. Méthodes de référencement spatial

Afin de s'intéresser à la structuration des bases de données de sources iconographiques en relation à la morphologie d'un édifice, notre travail aborde le problème de l'organisation spatiale de photographies sur une représentation tridimensionnelle de l'édifice issue d'un relevé (tel qu'un balayage laser 3D ou une restitution photogrammétrie). La notion de « résection spatiale » [Hartley R. et al., 2000] (orientation d'une image sur un ensemble de coordonnées 3D connues) appliquée au relevé d'édifice, peut constituer le point de départ pour le développement d'un système de référencement spatial de sources documentaires (photographies, images panoramiques, séquences vidéo, etc.) à l'intérieur d'une scène architecturale.

Le référencement spatial est une opération qui créée une relation spatiale entre l'image 2D et la représentation 3D qui correspond à une superposition. Il s'agit en

quelque sorte de retrouver les paramètres du point de vue de la source afin d'observer le modèle 3D à partir de ce point de vue. Ce point de vue peut être estimé de façon fine (c'est le cas d'une photographie, où la cohérence entre la réalité et l'image est forte) ou de façon approximative (c'est le cas des dessins, ou la cohérence entre la réalité et l'image est faible).

Les méthodes de référencent spatial permettent d'obtenir différents types d'informations géométriques relatives aux images (position, orientation, paramètres optiques, etc.) et à différents niveaux de précision. Une fois analysée les travaux et les solutions existantes dans ce domaine, on s'intéressera à l'identification des plus adaptées à chaque type de source iconographique à prendre en compte en fonction de ses propriétés. Comme on le verra dans les chapitres suivants, le référencement spatial des sources iconographiques permet d'envisager des modalités d'interrogation des bases de données d'images en fonction de critères spatiaux (à différents niveaux de précision). Nous allons présenter ici un ensemble de méthodes de référencement spatial en les regroupant en trois catégories distinctes : méthodes manuelles, semi-automatiques et automatiques.

6.1.1. Méthodes manuelles

Le référencement manuel des photographies est certainement la solution la plus simple à mettre en place et la plus facile aussi du point de vue de l'utilisateur final. Dans le domaine de la documentation architecturale et archéologique différentes recherches vise aujourd'hui à créer une relation entre les sources 2D et une représentation de la morphologie d'un édifice au travers d'outils permettant d'indiquer de façon approximative la position de la prise de vue d'une photographie. D'habitudes, ces solutions utilisent des représentations cartographiques (les SIG par exemple) permettant d'indiquer sur le plan d'un édifice la position (et parfois la direction de vue) d'une photographie. La position et l'orientation du point de vue de l'image sont généralement présentées par un symbole graphique, comme un point ou une flèche.

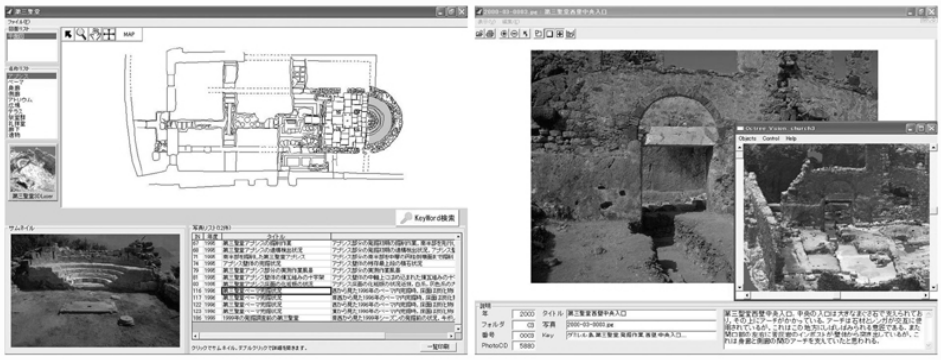


Figure 45 L'interface du système de [Kadobayashi R. et al.,2003], capable à référencer manuellement les information 2D et 3D

[Kadobayashi R. et al.,2003] propose un système qui permet d'effectuer des recherches d'images sur la base des informations stockées dans les légendes des photos. En même temps, ce système permet de manipuler les nuages de points et/ou le modèle 3D résultat du relevé et crée par maillage automatique (voir figure 45). Des

cibles de référencement (placées sur différentes parties de l'édifice) sont utilisées comme points de référence pour retrouver des photographies corrélées ; leurs coordonnées sont obtenues par des mesures topographiques avec une station totale. Chaque image a son propre système de coordonnées basé sur la position du scanner au moment de l'acquisition des informations. Ces images sont orientées selon un système de coordonnées géographique. L'interface du système de consultation comporte une vue planimétrique, une élévation et une vue 3D. L'utilisateur peut rechercher des images dans la base de données en cliquant sur les positions relatives indiquées sur les plans de l'édifice ou sur le plan de masse mais aussi en indiquant des mots-clés.

Le système APIS (Architectural Photogrammetry Information System) [Waldhäusl P. et al., 1994] dont certains aspects ont déjà été présenté dans le chapitre 2 se base sur une base de données qui répertoire les bâtiments documentés par des acquisitions photogramétriques. Le système est basé sur un outil permettant d'enregistrer des informations sur les édifices documentés et pour enregistrer les positions d'où les photographies ont été prises au moment du relevé photogrammétrique. Pour démarrer un projet avec ce système, les photographes doivent préparer leur propre esquisse sous forme de plans et élévations du bâtiment (comme indiqué dans la règle organisationnelle 3x3) pour ensuite y marquer les position et les direction de la caméra de façon approximative (voir figure 46).

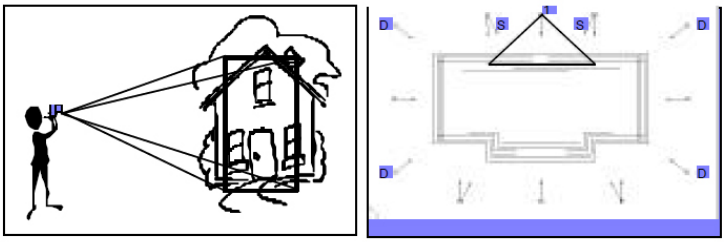


Figure 46 Planification des photographies du règle 3x3

Même si la méthode de référencement manuel des photographies est la méthode la plus simple et la plus compréhensible pour les utilisateurs du système, l'association d'images à leur point de vue nécessite une gestion manuelle de l'administrateur système. Les erreurs de détermination humaines sont toujours possibles, cela peut réduire la précision du système. De plus, c'est une méthode qui a besoin de beaucoup de temps de travail pour arriver à un niveau de précision suffisant pour qu'il soit utilisable dans notre méthodologie.

6.1.2. Méthodes semi-automatiques

Les méthodes semi-automatiques de référencement spatial d'images se basent sur des solutions géométriques appelées de calibration et orientation des caméras. Ceci se réfère au processus d'utilisation des valeurs numériques pour établir des paramètres géométriques et optiques de la caméra virtuelle qui correspondent à un point de vue de l'image par rapport à un référentiel 3D (ensemble e points, modèle 3D, etc.) [Cipolla R. et al, 1999]. Les méthodes de calibration et orientation des caméras peuvent prendre en compte les paramètres intrinsèques (distance focale, distorsion,

etc.) et extrinsèques (translation et rotation) de l'appareil photo (ou plus généralement de la camera) au moment de la prise de vue rapport à un repère tridimensionnel (voir figure 47). Plusieurs méthodes ont été développées.

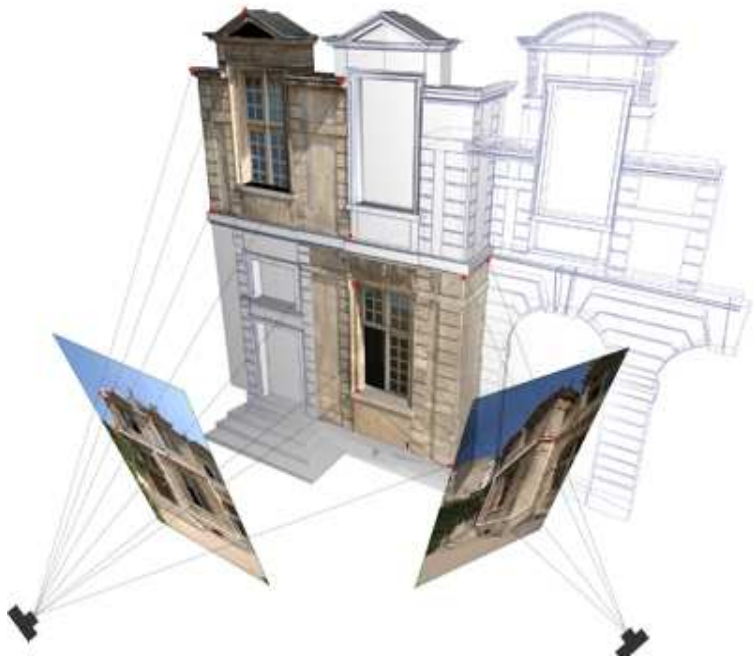


Figure 47 Calibration et orientation des caméras

La méthode DLT (Transformation linéaire directe) est une méthode rapportée par [AbdelAziz et al., 1971]. La méthode de DLT emploie un ensemble de points de commande dont les coordonnées d'espace-plan d'objet sont déjà connues. Les points de commande sont normalement fixés à un cadre rigide. Le problème principal de la méthode DLT est que les paramètres du calibrage ne sont pas mutuellement indépendants.

Parmi les recherches développées pour orienter une image sur un ensemble de coordonnées connues dans le domaine de la photogrammétrie [Hatze, 1988] et de la vision par ordinateur [Faugeras, 1993 ; Fitzgibbon et al., 1998 ; Heikkila et al., 1997 ; Kumar R. et al., 1994], une des méthodes les plus utilisées et celle proposée par [Tsai, 1986]. Son implémentation requiert des correspondances (points) sélectionnées respectivement dans l'espace 3D et sur l'image 2D. Cette méthode est basée sur un calcul en deux phases : premièrement on calcule la position et l'orientation, deuxièmement les paramètres internes de la caméra. Dans plusieurs applications qui nécessitent une bonne précision, on considère la caméra déjà calibrée. Cela veut dire que les paramètres internes sont déjà connus. Dans ce cas, seulement les six paramètres relatifs à la position et à l'orientation doivent être estimés.

6.1.3. Méthodes automatiques

Nous pouvons référencer les images au modèle géométrique tridimensionnel par une technique automatique en deux façons. La première consiste à utiliser un outil pour enregistrer l'information spatial du point de vue d'image au moment qu'elle été prise (hardware). Une autre option est d'utiliser un algorithme de l'ordinateur (software)

pour trouver le point de vue d'image qui a été prise sans enregistrée les données spatiales.

6.1.3.1. Solutions hardware

Dans cette section nous nous intéressons aux solutions hardwares existantes pour l'enregistrement automatique des paramètres spatiaux d'un appareil photo au moment de la de prise de vue.

[Hu W. et al. 2006] ont développée une solution d'estimation de la position et de l'orientation d'une camera basée sur un système hybride mêlant un GPS, un gyroscope et un algorithme de vision par ordinateur pour superposer des indicateurs 3D de navigation et de trafic routier dans un environnement de réalité augmentée.

Pour la réalisation de représentations tridimensionnelles photoréalistes d'objets architecturaux [Abmayr T. et al., 2004] proposent une solution pour enrichir les nuages de points avec les valeurs colorimétriques acquises par une caméra haute résolution intégrée au système de balayage produit par la société Zoller+Fröhlich. La caméra, développée en collaboration avec la German Aerospace Center (DLR), consiste en une unité pivotante qui intègre une caméra DTC capable d'acquérir des images à 10200 x 500000 pixels. L'acquisition laser et photographique est effectuée l'une après l'autre. Une fois la numérisation laser terminée, la caméra est montée sur le même pied à l'aide d'un adaptateur qui assure le positionnement précis du centre optique de la caméra sur celui du scanner. Cela permet (par la simple détermination du paramètre de rotation horizontale de la caméra) de superposer l'image sur le nuage de points. Cette procédure consiste à associer les attributs RVB extraits de l'image à chaque coordonnée du nuage de points. La figure 48 illustre un nuage de points enrichi par cette technique.



Figure 48 Le hall du château Neuschwanstein : l'image de luminance du scanner est mélangée avec les valeurs RVB correspondantes acquises par la caméra. [Abmayr T. et al., 2004].

[Salemi G. et al., 2005] présentent une expérimentation pour l'intégration d'images panoramiques dans un nuage de points issus d'un relevé laser. L'objectif est d'associer des informations métriques à un système pour l'exploration de panoramas cylindriques. L'intégration est basée sur la fusion des données issues de deux

systèmes d'acquisition. À partir d'une numérisation panoramique (360°) avec un module de balayage temps de vol et d'une acquisition photographique faite par la caméra embarquée dans le scanner (calibrée préalablement par les constructeurs du scanner), les auteurs se focalisent sur l'extraction d'entités caractéristiques sur les images issues de la phase d'acquisition. Le processus d'extraction d'entités se base sur trois phases : la première est l'application d'un filtre « Edge-Sobel » sur les valeurs d'intensité acquises par le scanner (nuances de gris). La seconde est l'application du même filtre en utilisant la modalité « Keep sign of result (dans une seule direction) » sur les valeurs RVB de l'image. De cette manière les arêtes sur l'image sont contournées. Une application de la méthode a été expérimentée à l'occasion du relevé du bâtiment « Suriano Bellotto » à Venise (voir figure 49). Trois images sources ont été utilisées pour le traitement : l'image panoramique en couleurs visible, l'échelle d'intensité en nuances de gris et la carte de profondeur.

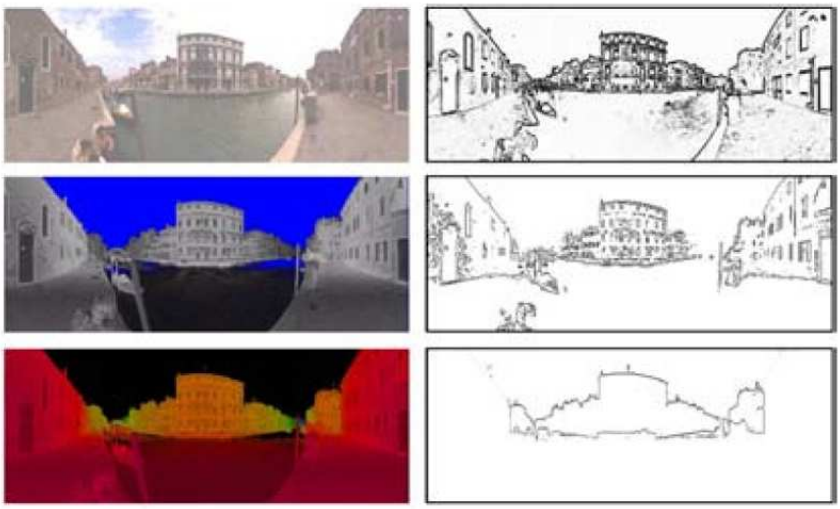


Figure 49 Acquisition laser panoramique à Venise. À gauche, images de l'acquisition laser. À droite, les images filtrées correspondantes. De haut en bas : l'image en vraies couleurs, l'image d'intensité en nuances de gris, la carte de profondeur. [Salemi G et al., 2005].

Le module [Omega] a été développé par le laboratoire MAP-GAMSAU en relation avec notre travail de thèse. Celui-ci permet d'enregistrer, de façon automatique, la position et l'orientation des photographies au moment de la prise de vue. Il utilise une station totale possédant une camera vidéo embarquée. La station total, en mode « tracking », possède un prisme, positionné sur le module OMEGA, ce qui lui permet d'atteindre de très longues portées (jusqu'à 500 mètres) en balayant sur 360° autour de son axe.

La station totale utilise ce qu'on appelle une « méthode télémètre» (en anglais ; range-finder) : un faisceau infra-rouge est intégrer dans le processus qui utilise un réflecteur à prisme tétraédrique comme une cible. La station totale est également capable de prendre des photos à partir d'une camera interne dont la position et l'orientation relative sont connues. Le prisme (dont la station totale contrôle la position dans l'espace) et les boules de repères montées sur le module OMEGA sont alors utilisés comme références pour déterminer l'orientation de l'appareil photo au travers d'un algorithme d'analyse d'image (voir figure 50).



Figure 50 Système de OMEGA, une méthode de référencement automatique d'image.

Comme les positions relatives entre le prisme, les boules et l'appareil photo montées sur le module sont connues a priori, la position et l'orientation du module peut être déterminé à partir de la détection de la position des boules (par un algorithme d'analyse d'image) enregistrée par la camera intégrée à la station totale au moment de la prise de vue.

6.1.3.1. Solutions software

Concernant les méthodes algorithmiques basées sur l'analyse et le traitement d'images, [Lee S.C. et al.,2002] proposent une méthode pour estimer la position et l'orientation des caméras à partir d'une extraction automatique des points de fuite dans l'image. Les points de fuite sont extraits à partir d'un ensemble de lignes parallèles dans l'image et sont alors utilisées pour l'estimation des paramètres des cameras.

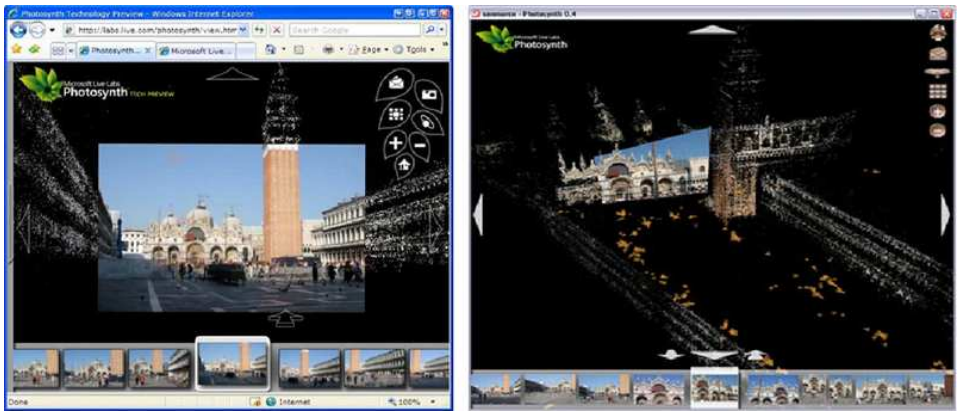


Figure 51 L'interface de Photosynth

Photosynth [Snavely N. et al.,2006 ; Snavely N. et al.,2007] est un logiciel développé par Microsoft Live Labs et l'Université de Washington. Ce logiciel analyse des photographies numériques et génère un nuage de points correspondants aux coordonnées extraites à partir des points homologues repères de façon automatique

dans une collection d'images (voir figure 51). Le système permet des mises à jour des collections d'images au travers d'une procédure de superposition manuelle de nouvelles images sur les images déjà orientées dans l'espace. Une fois une solution approchée est indiquée par l'utilisateur, l'algorithme d'extraction de points homologues [Lowe D., 1999] crée des correspondances entre la nouvelle image et une ou plusieurs images orientées, ces correspondances sont ensuite utilisées pour l'orientation de la nouvelle image par un calcul d'ajustement des faisceaux.

[Johansson B., 2002b] décrit un système automatique pour l'estimation de la position et de l'orientation d'une image simple représentant une scène urbaine, dans le cas ou les modèles des édifices présents dans l'image sont disponibles. Chaque édifice a un modèle qui consiste en un certain nombre de plans parallèles associés. Une homographie (un plan dans la scène superposé à une image) est estimée de façon automatique pour chaque édifice présent dans l'image. Le modèle ayant le meilleur score de superposition est choisi pour déterminer la position et l'orientation de la camera. Cette solution s'appuie sur l'hypothèse que les édifices sont composés principalement par des plans et que ces plans ont des arêtes verticales et horizontales bien visibles (fenêtres et/ou portes des édifices)(voir figure 52).

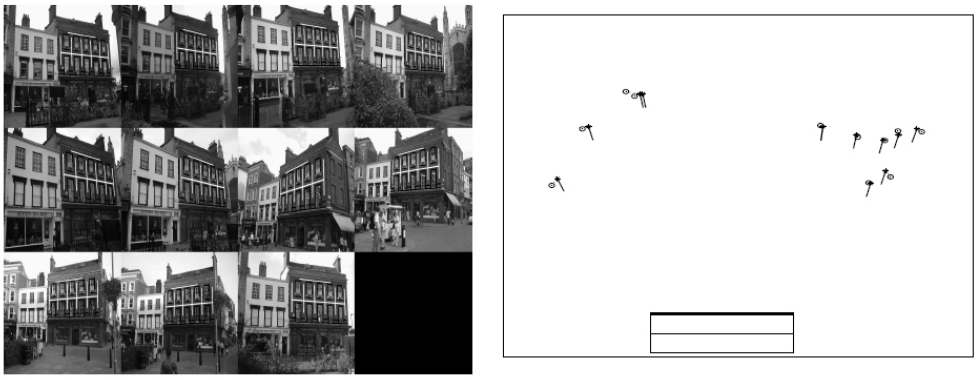


Figure 52 À gauche, images de différents points de vue avec le modèle adapté à un bâtiment. À droite, Le résultat de l'estimation de la position

Pour chaque édifice dans l'image, un modèle 3D est généré. Ce modèle consiste en un certain nombre de plans en 3D en fonction de la détection de la position de leurs arêtes. Pour chacun des édifices candidats, le modèle correspondant est superposé à l'image. Dans une deuxième phase, l'estimation de la position et de l'orientation de la camera associées à l'image sont calculées. Ceci est possible en ayant des informations métriques concernant le modèle. Dans ce cas la méthode utilisée est l'auto-calibration qui fonctionne avec un simple plan dans une seule image. Comme la camera a 11 paramètres, 6 externes et 5 interne, quand la géométrie du plan est connue, l'homographie donne deux contraintes dans la calibration.

6.2. Analyse des solutions existantes

A partir de l'analyse des solutions de référencement spatial d'images étudiés, un certain nombre de considération sont à prendre en compte afin d'établir leur degré d'adaptabilité au problème spécifique que nous souhaitons traiter.

GPS, tout en étant particulièrement flexibles et faciles à mettre en place dans leur utilisation, ne peuvent pas vraiment être prise en compte pour nos exigences à cause de leur faible niveau de précision. En effet, pour les utilisations que nous envisageons de mettre en place au sein de notre système d'informations, nous nécessitons d'un degré de cohérence géométrique très fin, de l'ordre du décimètre en termes de positionnement de l'appareil photo, ce niveau de précision ne pouvant pas être obtenu par la technologie GPS. Ces technologies ne sont pas prises en compte non plus à cause du fait qu'au delà de la position de l'appareil photo, nous devons également connaître son orientation (toujours de façon assez précise) dans l'espace. En revanche, les solutions basées sur une superposition fine des images panoramiques sur un nuage de points, ou encore la solution OMEGA permettent un référencement en termes de positionnement et d'orientation à la fois tout à fait compatible avec les niveaux de cohérences nécessaires à une superposition précise des images photographiques sur le modèle 3D résultant du relevé architectural. Toujours en prenant en considération la précision, il est certes que toutes les solutions softwares basées sur un calcul de calibration, puis orientation d'ensembles d'images (manuel ou automatique) restent les plus adaptées à l'exploitation des relations projectives entre les images photographies et le modèles 3D, car toute coordonnées utilisée pour la construction de ce modèle est directement extrait des images. Cela veut dire que le modèle 3D issus d'une reconstruction à partir d'image se trouvera toujours en cohérence géométrique avec les images photographiques.

En ce qui concerne la flexibilité d'application, il est évident que l'utilisation de dispositifs hardware contraint énormément le champ d'application aux seules photographies ou images vidéos, de plus acquises dans des conditions particulières de prise de vue (notamment dans des campagnes de relevé métrique). Dans ce sens, les solutions hardwares ne seront pas une solution effective pour toute les sources iconographiques collectées par étude documentaire et provenant d'archives historiques. Concernant la flexibilité d'acquisition, les solutions basées sur un calcul de résection spatiale (points 2D saisi sur l'image et points 3D correspondants saisi sur le modèle 3D) s'avèrent les plus adapté à la grande hétérogénéité de formats, de niveaux de résolutions, d'état de conservation caractéristique des fonds photographiques des monuments historiques. En effet, ces solutions déterminent la superposition image/modèle à partir d'une estimation à la fois des paramètres intrinsèques et extrinsèques en ayant ainsi la possibilité de s'adapter (avec des niveaux d'approximation conséquent) à différents types de photographies. De plus, ses solutions sont aussi adaptées au référencement spatial de représentations artistiques (dessins, peintures, gravures, etc.) dont la construction géométrique se base sur la perspective. En effet, malgré les incohérences (parfois particulièrement accentuées) existantes entre un dessin en perspective et la réalité, le calcul d'ajustement de faisceaux permet d'obtenir une estimation de la position et de l'orientation de l'image convenable, tout en restant dans la forte approximation.

6.3. Solution retenue

Afin d'implémenter dans notre système une solution pour le référencement spatial de sources iconographiques, nous avons pris en considération deux aspects principaux.

Le premier aspect concerne la grande hétérogénéité d'images qui constituent l'iconographie relative à un édifice historique. Comme expliqué dans le chapitre de cette thèse, plusieurs supports (dessins, photos, peintures), plusieurs formats et résolutions, plusieurs états de conservation et surtout différents degré de certitude peuvent être attribués aux sources iconographiques. Dans ce sens, une première préoccupation a été de concevoir une solution particulièrement facile à gérer par les utilisateurs finaux dont les exigences sont limitées à une spatialisation approximative des documents graphiques.

Le deuxième aspect concerne le fait que, comme nous l'avons expliqué dans le chapitre 4 présentant le schéma général de notre approche, nous nous appuyons sur une représentation 3D de l'édifice conformé à la réalité (toujours avec un niveau d'approximation) en termes métriques et géométriques. C'est la raison pour laquelle nous n'avons pas besoin de mettre en place de solutions qui créent des représentations tridimensionnelles directement à partir des images.

A partir de ces deux constants nous avons pris en compte l'implémentation de trois différentes solutions de référencement spatial pouvant représenter trois différents degrés de précision.

Au premier niveau nous trouvons une solution pour la superposition manuelle de sources à projection en perspective (ou pseudo-perspective) et orthogonale directement sur le modèle 3D au travers du control du déplacement et de l'orientation de la camera de navigation à l'intérieur de la scène 3D de l'édifice. Cette solution est particulièrement flexible et adaptée au référencement spatial de sources iconographiques comme les dessins, les peintures et les gravures à projection perspective (ou presque) et orthogonale (plans, élévations et coupes).

Au deuxième niveau nous trouvons une solution pour la superposition géométrique d'images photographiques basée sur un calcul de résection spatiale. Cette solutions est suffisamment flexible pour prendre en compte la grande variété (en termes de résolutions, de format et de précision) des fonds de photographies concernant les édifices patrimoniaux (photos d'archive, cartes postales, photos de tourisme et photos de relevé).

Au troisième niveau nous trouvons une solution pour la récupération directe des paramètres de calibration et d'orientation de photographies de relevé calculés dans des solutions de photogrammétrie pure ou dans le cadre de l'utilisation complémentaire et intégrée de balayage laser 3D et photogrammétrie.

L'aspect intéressant de notre solution est que toute source iconographique spatialisé par une solution approximative, peut toujours faire l'objet d'un nouveau référencement spatial conduit par une méthode plus précise. Il est par exemple possible de superposer rapidement un jeu de photographies avec la méthode de

référencement manuel pour ensuite affiner leur orientation via la méthode d'orientation géométrique.

Dans les paragraphes suivants nous présentons les solutions adoptées en fournissant aussi des détails sur leur implémentation informatique au sein de notre système.

6.3.1. Téléchargement sur serveur des sources iconographiques

La procédure de référencement spatial implémentée démarre par le téléchargement des images sur le serveur du système. Ce processus sa base sur un script PHP permettant de sélectionner une image parmi les répertoires de l'ordinateur du coté client. Les images sont enregistrées sur le serveur en passant par un formulaire permettant de leurs associer des attributs descriptifs tel que le nom de l'auteur, le type de sources, le format, la propriété, etc (voir figure 53). Ces attributs sont organisés conformément au modèle de description des sources iconographiques de la base de données officielle du ministère de la Culture et de la Communication français.

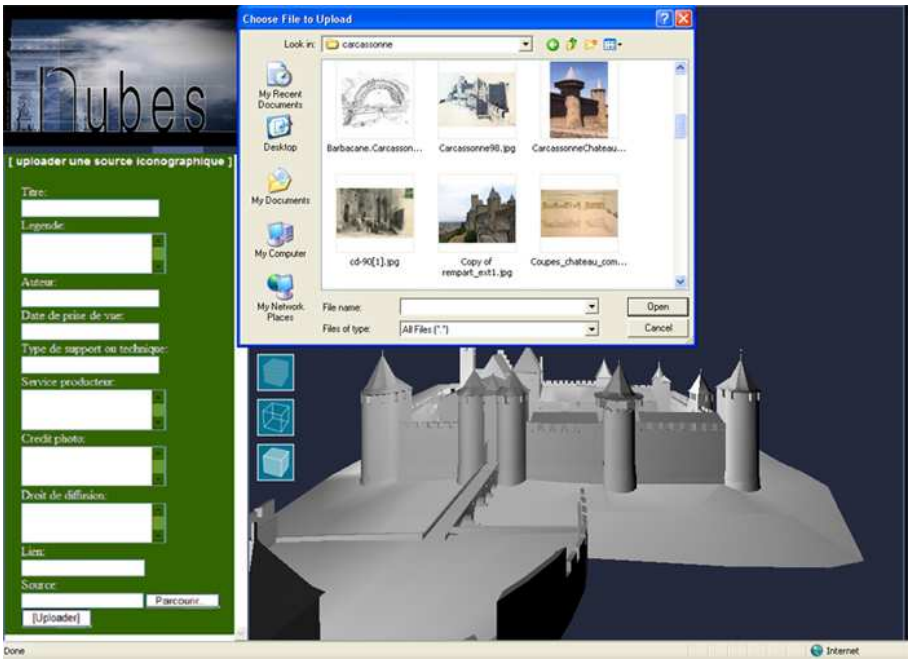


Figure 53 L'interface pour le téléchargement d'images sur serveur

Aux informations descriptives rentrées par l'utilisateur s'ajoutent des informations directement extraites de la lecture des paramètres de l'image (dimensions en pixels) et l'url du fichier image stocké dans la base de données. Un script PHP envoie ensuite une commande à la scène 3D temps réel afin que l'image soit téléchargée et visualisées dans la scène 3D. Un plan (d'un niveau de transparence configurable de façon interactive) est alors placé dans la scène 3D de manière orthogonale à la camera de navigation. La taille du plan est calculée en fonction des proportions de l'image (à partir des informations dimensionnelles stockées dans la base de données).

6.3.1. Superposition manuelle de sources à projection orthographique et pseudo-perspective

Pour la superposition manuelle de sources iconographiques à projection perspective et pseudo perspective nous avons développé une interface interactive 3D. L'objectif de cet outil est de fournir une voie très rapide et facile aux utilisateurs pour positionner de façon approximative des images comme des dessins, des peintures et des gravures dont la construction graphique se base sur la perspective aussi bien que des photographies. L'interactivité programmée permet aux utilisateurs de manipuler un plan 2D (sur lequel l'image est plaquée) en le contrôlant dans l'espace 3D en liaison directe avec la camera de navigation temps réel.

Au travers de contrôles de mouvement et de rotation, l'utilisateur peut déplacer la camera de navigation (toujours en ayant l'image à superposer comme plan d'image de la camera) afin de retrouver le point de vue correspondant dans la scène 3D (voir figure 54)



Figure 54 L'interface pour le référencement manuel.

Au-delà de la position et de l'orientation du point de vue, un contrôle permet de modifier (toujours en temps réel) l'angle de champ de la camera (distance focale) afin d'adapter la vision de la scène 3D à la perspective de l'image (voir figure 55).



Figure 55 L'interface pour le référencement manuel. Variation de l'angle de champ de la caméra.

Afin d'offrir une manipulation plus simple en phase de superposition, d'autres contrôles permettent de pivoter l'image (en sens horaire et antihoraire) autour de l'axe principal de la camera (voir figure 56).

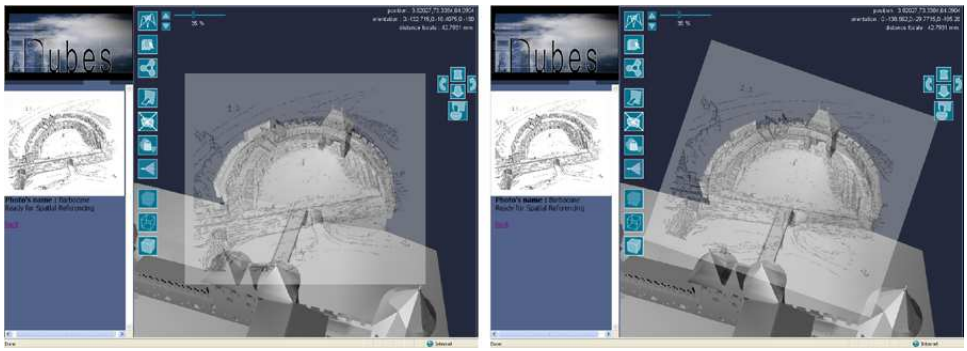


Figure 56 L'interface pour le référencement manuel. Rotation du plan image**.**

Afin de prendre en compte la superposition de plans, de coupes et d'élévations, un contrôle permet de basculer entre la vision de la scène 3D au travers d'une camera à projection perspective à une vision à projection orthogonale (voir figure 57). Dans ce cas, le contrôle de la distance focale est remplacé par un contrôle de zoom.

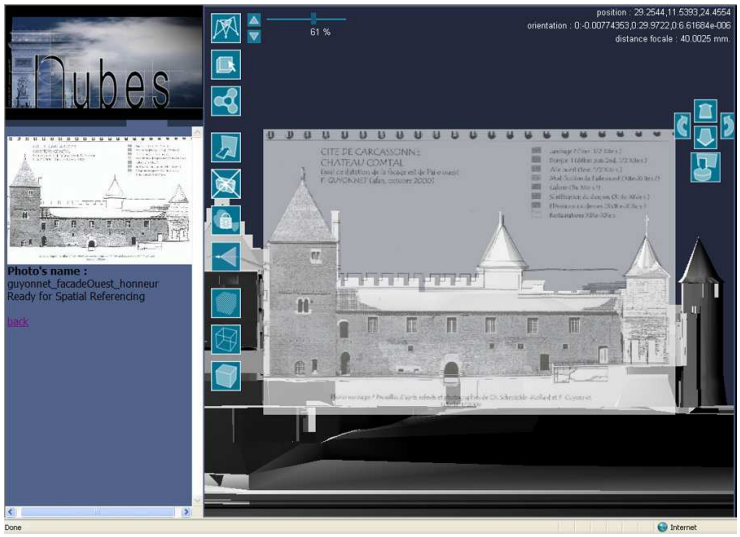


Figure 57 L'interface pour le référencement manuel. Projection orthographique.

Pour prendre en compte la superposition de sources présentant des sections (plans et coupes par exemple), une fonction supplémentaire permet de contrôler les limites de la pyramide visuelle de la camera (frustum) afin de limiter la vision aux seuls éléments représentés dans la source.

6.3.2. Résection spatiale de sources à projection perspective

Dans ce cas de figure nous prenons en compte un modèle géométrique de camera à projection en perspective défini par des paramètres intrinsèques et extrinsèques.

Les paramètres intrinsèques expriment les propriétés projectives de l'appareil photo (ou d'une perspective dessinée). Il s'agit de la distance principale (ou focale de l'objectif dans le cas d'une photographie), des coordonnées du point principal (centre perspectif projeté sur le plan image) et de la distorsion qui correspond aux erreurs de représentation introduites par la forme et les imperfections de l'objectif (toujours dans le cas d'une photographie). La distorsion est par un simple coefficient de distorsion radial. En effet, les corrections de la distorsion se caractérisent par un déplacement quasiment nul au point principal et graduellement accentué vers les extrémités de l'image. Ce modèle, beaucoup plus simple d'un modèle polynomial généralement utilisé en photogrammétrie, offre des avantages en termes de flexibilité, notamment en présence d'images avec une faible cohérence géométrique avec le réel (photos déformées, dessins en perspective, etc.)

Les paramètres extrinsèques représentent la position et l'orientation de l'appareil photographique au moment de la prise de vue. La position et l'orientation de la caméra par rapport à ce repère sont déterminées par une translation et une rotation au moyen de six composantes au total. Une translation est définie par les composantes (Tx, Ty, Tz) relatives aux axes. Une rotation est décrite par trois angles (Rx, Ry, Rz) autour de chacun des axes.

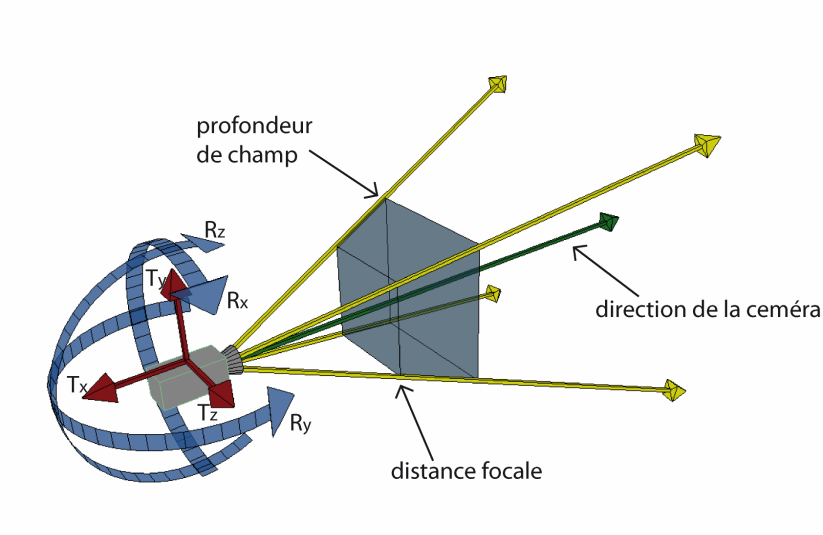


Figure 58 Le modèle géométrique d'une caméra à projection perspective.

Pour réaliser le calcul de résection spatiale nous utilisons l'algorithme de [Tsai R.Y., 1986] qui est basé sur un modèle de projection perspective et permet d'estimer les onze paramètres suivants : distance focale de la caméra ; coefficient de distorsion radiale de l'objectif ; coordonnées du centre de distorsion radiale ; facteur d'échelle pour prendre en considération l'incertitude due aux imperfections en phase de numérisation de l'image; angles de rotation pour la transformation entre les coordonnées de l'objet et de la caméra; composantes de translation pour la transformation entre les coordonnées de l'objet et de la caméra.

La calibration de Tsai consiste en la transformation des coordonnées 3D d'un point en coordonnées correspondantes dans l'image. Une calibration coplanaire requière que la composante Z des coordonnées 3D soit fixée à 0 et au moins 5 points

d'entrée. Tandis qu'une calibration basique non coplanaire requiert au moins 7 points d'entrée.

Malgré cela nous avons pu constater qu'une calibration optimisée requiert 11 points minimum pour les deux modalités. D'autres considérations autour de la qualité de la calibration concernent la stratégie de saisie des correspondances : les points choisis doivent en effet être distribués largement dans le champ visuel de la caméra et couvrir toute la gamme de profondeur présente sur l'image (ou dans la géométrie que l'on veut restituer).

Cette solution offre l'avantage d'estimer les paramètres internes et externes séparément. Il s'agit d'une option particulièrement utile car elle donne la possibilité de fixer les paramètres internes de la caméra calculés sur la première image d'un bloc puis de faire seulement l'estimation des paramètres externes des autres images de la même série (prises avec le même appareil photographique et avec le même niveau de zoom par exemple).

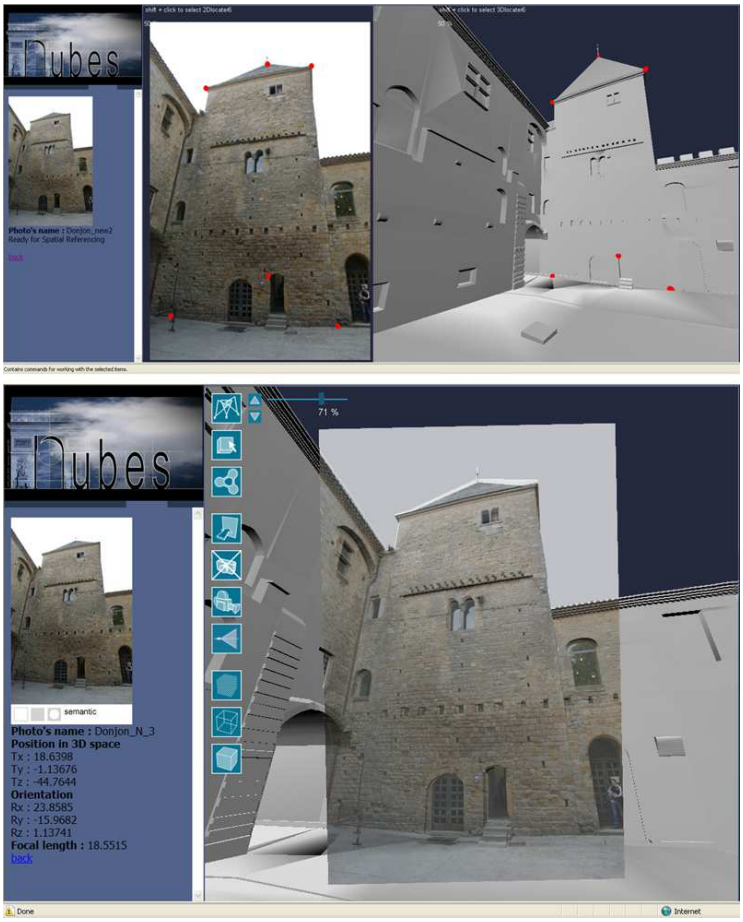


Figure 59 L'interface pour la résection spatiale d'images. En haut : le processus de sélection des points d'entrée sur le modèle 3D et sur l'image, en bas : le résultat du référencement.

A partir de nombreuses applications nous avons pu apprécier un niveau de précision acceptable pour notre usage (avec des erreurs de re-projection de l'ordre de 4 ou 5 pixels) sur des photographies issues d'appareils photo numériques conventionnels (voir figure 59). Elle est également une solution particulièrement flexible

car elle ne nécessite pas d'informations sur l'appareil photographique utilisé pour la prise de vue (c'est bien le cas des photographies historiques par exemple). Malgré cela, elle demeure une technique manuelle importante du côté de l'utilisateur (saisie d'au moins 7 points dans la scène 3D et dans l'image 2D).

6.3.3. Récupération directe des paramètres de photographies orientées

Afin d'intégrer des orientations de photographies provenant des solutions de photogrammétrie, notre système permet de rentrer directement les valeurs relatifs aux paramètres intrinsèques et extrinsèques des cameras au travers d'un formulaire. Ceci permet de bénéficier d'un degré de cohérence géométrique maximale, notamment dans le cas où la morphologie de l'édifice est directement reconstruite à partir d'images. C'est le cas des représentation 3D élaborées par la technique de la photo-modélisation [De Luca L., 2009] qui se sert de points extraits des photographies pour conduire la reconstruction géométrique des éléments architecturaux en utilisant des fonctions de modélisation interactive.

6.4. Annotation sémantique des images par projection

Comme décrit dans la présentation générale de notre approche de spatialisation et de sémantisation de sources iconographiques (chapitre 4), nous utilisons la relation spatiale entre la représentation 3D et les images 2D. L'utilisation de l'information 3D comme solution pour créer une relation entre une description sémantique et des images corrélées est originale dans le domaine de l'annotation sémantique d'images. Une fonction dans Photosynth [Snavely N. et al., 2006] utilise un concept similaire pour explorer une relation bidirectionnelle entre l'image et la sémantique en utilisant un nuage de points (extrait directement des images) comme véhicule de distribution/propagation. Cette fonction permet de transférer de façon automatique la sémantique d'une image à une autre que contient la(es) même(s) région(s) de la scène.

L'utilisateur de ce système peut sélectionner manuellement la région d'une image et entrer une annotation textuelle. L'annotation est alors enregistrée et reliée à la portion de nuage de points sélectionnée et apparaît comme une enveloppe (rectangulaire dans l'image). Une fois l'annotation crée elle est automatiquement transférée sur toutes les autres images concernées. Pour déterminer si une annotation est appropriée pour une camera donnée, le système teste sa visibilité dans l'image et adapte l'enveloppe. Au-delà de l'enrichissement sémantique des images, l'habilité de transférer les annotations de ce système offre une solution très portable pour le développement de systèmes de réalité augmentée basé sur la vision par ordinateur [Feiner S. et al. 1997].

Comme nous l'avons déjà dit, dans notre travail, nous nous concentrons sur la relations entre une représentation 3D détaillée (et non pas un nuage de points approximatif de la scène) et les images 2D. Cette relation est réalisée par la projection des contours des objets géométriques (en utilisant spécifiquement une représentation polygonale) sur les images 2D référencées spatialement dans le but de générer de

façon automatique l'annotation sémantique de ces images. Par rapport à la méthode développée au sein de Photosynth, la différence la plus importante consiste en le fait que nous utilisons un vrai modèle géométrique 3D comme distributeur / propagateur, à la place d'un nuage de points. Ceci nous permet de gérer une complexité, une précision et une lisibilité majeure des régions projetées.

En utilisation les résultats des méthodes de référencement spatial décrites dans les paragraphes précédents, le point de vue d'une camera virtuelle relative à chaque image corresponde au point de vue dans le monde réel. De cette manière les contours des objets de la scène sont projetés sous forme d'une couche vectorielle additionnelle s'associant à l'image originelle. Dans l'image vectorielle projetée, chaque segment de polygone corresponde parfaitement à la forme des objets 3D de la scène en fonction au point de vue de l'image et associe aussi tout les attributs sémantiques qui qualifient les éléments 3D (voir section 5.3.2). Ce processus est complètement automatique.

6.4.1. Présentation de l'environnement du développement

Pour réaliser la projection des couches sémantiques du modèle 3D aux images 2D spatialisées nous avons développée un processus en s'appuyant sur les langages PHP/MySQL (voir section 4.3.2.2), MEL et SVG.

Langage MEL (Maya Embedded Language). : est un langage de script de Maya, un logiciel de modélisation 3D. Le langage MEL permet d'avoir accès à toutes les fonctions du logiciel concernant la modélisation, le rendu, l'animation ainsi que la création d'interfaces customisées. Maya a une architecture basée sur les commandes. Il est possible de rentrer explicitement des commandes dans les langages de script. Chaque commande a plusieurs arguments (flags) qui établissent la façon dont la commande est exécutée. Les arguments et leurs valeurs implicites ou explicites sont requis pour exécuter une commande.

SVG est un format de données conçu pour décrire des représentations graphiques vectorielles basé sur XML. Ce format est spécifié par le World Wide Web Consortium. Les coordonnées, les dimensions et les structures des objets vectoriels sont indiquées sous forme numérique dans le document XML. Un système spécifique de style (CSS ou XSL) permet d'indiquer les couleurs et les polices de caractères à utiliser. Ce format gère les formes géométriques de base, les parcours (qui utilisent les courbes de Bézier), le remplissage (qui peuvent être créés à l'aide de gradients, des patterns ou de filtres), les markers (des motifs distribués le long des chemins) et le canal alpha pour la transparence. Comme dans tout document XML, les objets sont organisés sous forme d'arbre et de balises. Le format SVG, accessible depuis JavaScript, permet l'intégration d'animations et la manipulation des objets graphiques grâce à des scripts, à des événements et par l'intermédiaire du modèle Document Object Model (DOM). SVG peut être visualisé nativement dans la plupart des navigateurs web ou via un plug-in.

6.4.2. Développement de la procédure de sémantisation d'images

Du point de vue du développement, le système de projection sémantique est basé sur la communication entre la base de données MySQL et Maya. Cette communication permet de collecter les informations nécessaires depuis la base de données et générer les modèles géométriques des cameras associées aux sources iconographiques spatialisées dans Maya (dans lequel la représentation polygonale de la scène 3D est aussi chargée). Le moteur de rendu de Maya est ensuite utilisé pour calculer un rendu vectoriel de la scène de chaque image 2D (à partir de chaque point de vue). La dernière phase consiste à associer chaque région projetée (polygone en SVG) aux attributs sémantiques stockés dans la base de données.

Ce processus est donc composé de quatre étapes : collecte des données depuis la base de données, reconstruction des points de vue image, projection des couches sémantiques, ajout d'attributs supplémentaires et de comportements interactifs dans les images sémantisées. Le schéma présenté dans la figure 60 explique ce processus.

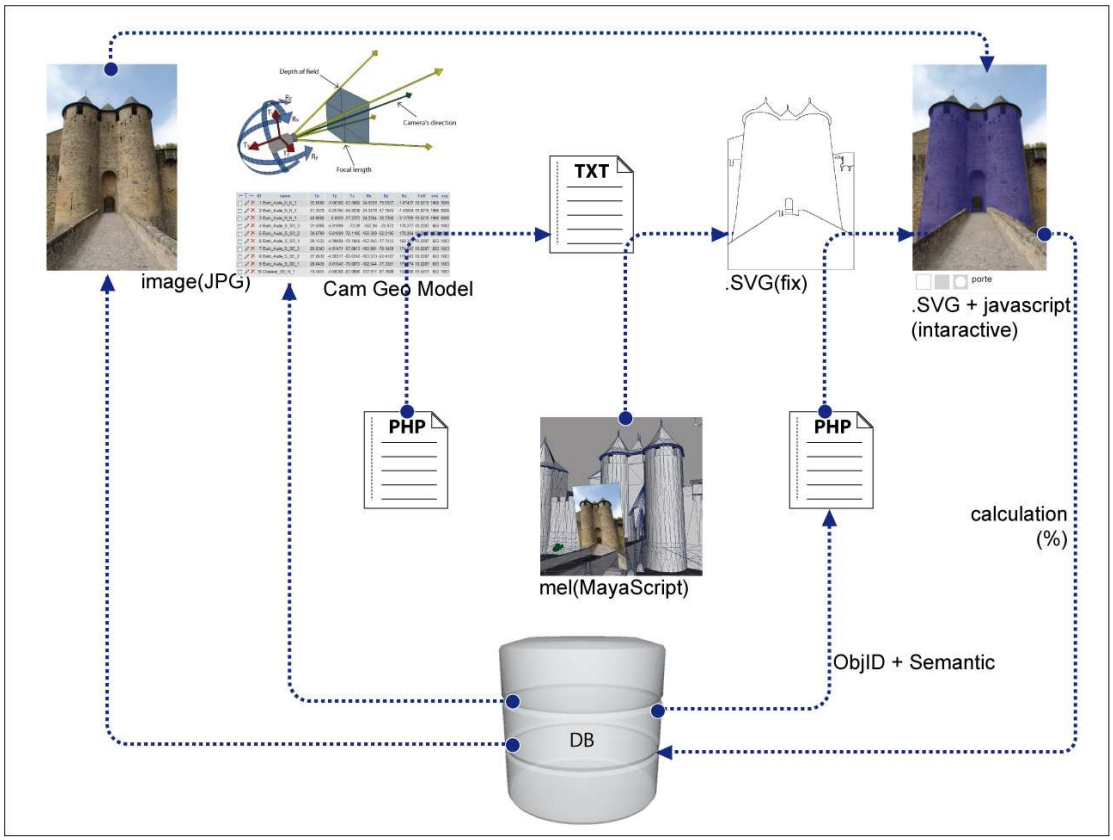


Figure 60 Schéma de la sémantisation d'images.

6.4.2.1. Collecte de données depuis la base de données

Trois types d'informations sont collectées pour ce processus provenant des phases antérieures de notre approche (annotation sémantique des formes 3D, sources iconographiques et leur point de vue) :

- l'information 2D des images originales (largeur et hauteur) à partir de la table des sources iconographiques ;
- l'information spatiale relative au modèle géométrique des caméras à partir de la table des référencements spatiaux;
- les identifiants des éléments 3D et leurs attributs sémantiques à partir de la table de la description sémantique de la morphologie de l'édifice.

Ces informations sont nécessaires pour recréer les points de vue des caméras dans Maya. Nous nécessitons de les transférer depuis la base de données MySQL dans une forme lisible et représentable par Maya. Les deux premières catégories de données sont exportées depuis la base de données en utilisant un script PHP et réécrite sous forme d'un tableau textuel configuré en fonction d'un parseur MEL.

6.4.2.2. Reconstruction des points de vues des sources iconographiques

Un script MEL lie le tableau généré et utilise les valeurs pour créer et paramétrer des caméras virtuelles dans la scène 3D afin de reproduire la position et l'orientation (ainsi que les paramètres intrinsèques) des images spatialisées.

Les valeurs nécessaires pour la reconstruction des points de vue des images sont :

- Le vecteur de position et d'orientation, pour déterminer le point de vue dans l'espace 3D ;
- Une valeur décimale pour la distance focale, pour déterminer l'angle de champ de la caméra ;
- Les proportions de l'image originale, pour déterminer les dimensions du plan d'image.

Chaque caméra est crée automatiquement, et nommée par l'identifiant de l'image originale, ensuite ses paramètres sont ajustés aux valeurs lus dans les tableau.

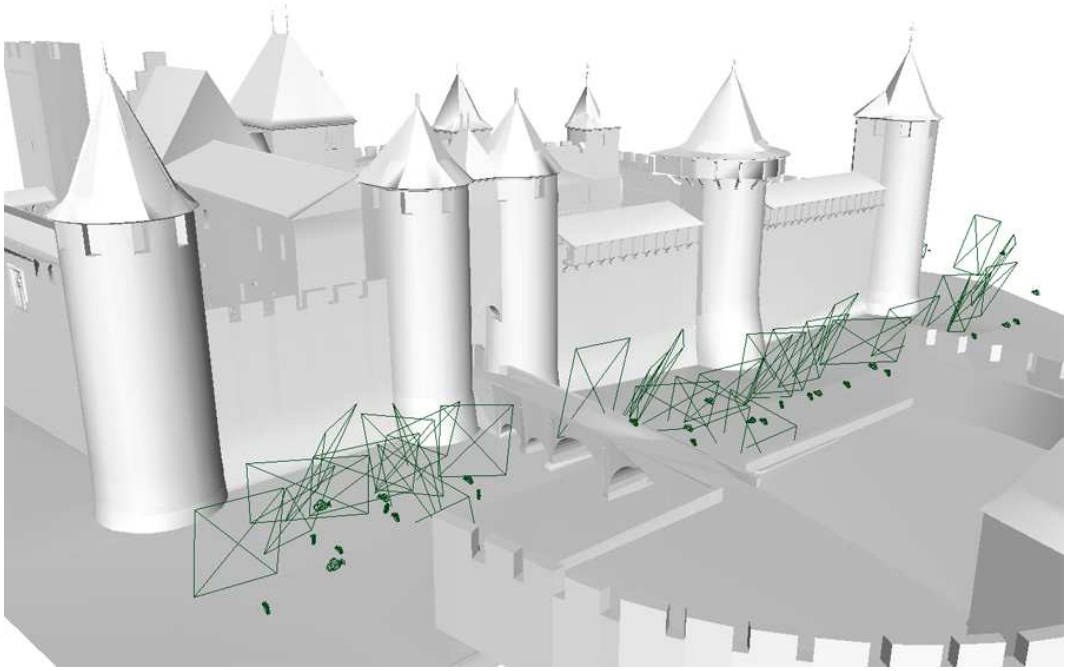


Figure 61 Reconstruction des points de vues de images dans une scène 3D (Maya).

Une fois les caméras virtuelles crées dans la scène 3D, Maya télécharge automatiquement les éléments 3D de la scène à partir de la liste des représentations polygonales présentes dans la base de données (voir figure 61). Cette liste de représentation fait référence au modèle de description sémantique de la morphologie de l'édifice et elle véhicule les identifiants des entités morphologiques reliées aux annotations sémantiques. Une fois ce téléchargement terminé, la scène 3D est prête au calcul de rendu vectoriel.

6.4.2.3. Rendu vectoriel des couches sémantiques

Afin de créer une image vectorielle de la scène 3D, nous utilisons une libraire de fonctions de rendu appelée MayaVector. La procédure MEL développée lance un rendu automatique de la scène 3D à partir de chaque caméra virtuelle reconstruite (relative aux points de vue des images spatialisées). Cette libraire nous permet de projeter facilement la silhouette des éléments 3D sur l'image entière dans la base de données. Cela produit une image vectorielle en format SVG ayant un point de vue correspondant à celui de l'image originale. La procédure de rendu est programmée de façon à calculer seulement les arêtes de contours des objets dans la scène (en incluant les arêtes d'intersection entre les différents objets) et d'ignorer les autres arêtes polygonales (voir figure 62).

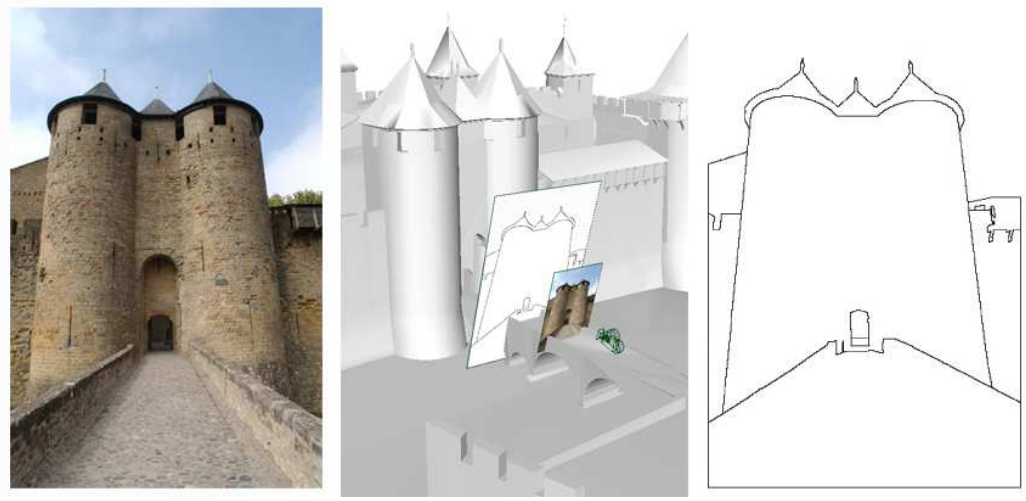


Figure 62 Projection des contours des éléments 3D sur l'image. A gauche, l'image originale, à droite l'image vectorielle SVG.

Au moment du calcul du rendu vectoriel nous associons l'identifiant de l'élément 3D (entité morphologique de la description sémantique) à la région polygonale 2D correspondante dans l'image vectorielle. L'image SVG est ainsi enregistrée dans la base de données (à coté de l'originale) et prête à être enrichie par des attributs supplémentaires.

6.4.2.4. Ajout d'attributs sémantiques et de comportements interactifs

La procédure pour l'ajout d'attributs sémantiques est réalisée par un script PHP. Ce script utilise des commandes pour déstructurer, restructurer et éditer l'image SCG en fonction des informations stockées dans la base de données et relatives à la description sémantique de la morphologie de l'édifice. Chaque région polygonale de l'image est associée à une annotation sémantique au travers de l'identifiant du polygone SVG provenant de la description de la morphologie de l'édifice et véhiculé de la scène 3D à l'image 2D par la procédure que l'on vient de décrire précédemment.

Le dernier processus se concentre sur l'ajout d'un comportement interactif de l'image en utilisant javascript. Cette interactivité est utilisée dans notre système de recherche visuelle (voir détails dans la section 7.1.3). Le résultat de ce processus est une image interactive (sensible au mouvement du curseur) affichant le nom et la sémantique associée aux régions de l'image. Cette image possède l'image d'original comme un fond. Quand l'utilisateur passe sur un objet avec le curseur, la région polygonale relative est colorisée et un texte dynamique affiche l'annotation sémantique correspondante (voir figure 63).

Le niveau de précision de la segmentation et de l'annotation sémantique des images dépend des résultats du référencement spatial.

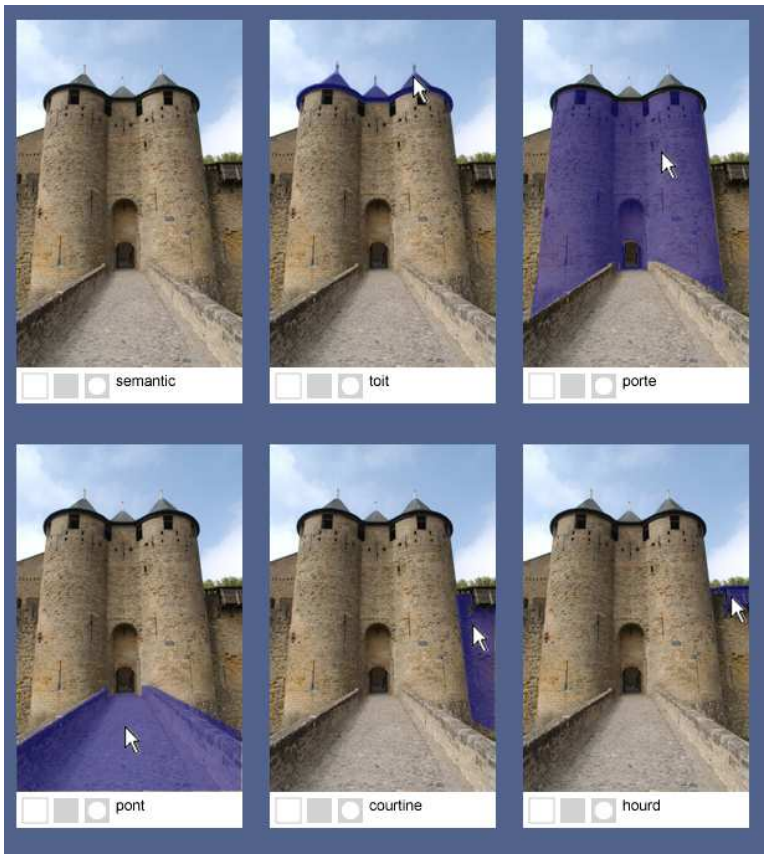


Figure 63 Image semantisée SVG associant des comportements interactifs au passage de la souris

Une valeur relative à la surface de chaque région de l'image est calculée en fonction de la position des sommets du polygone. Cette information est utilisée dans les procédures de recherche (voir section 7.3.1), par exemple en utilisant un pourcentage de surface comme critère pour rechercher des images contenant un élément spécifique ou encore pour trier les résultats d'une recherche.

La phase finale du processus de sémantisation des images est la mise à jour de la base de données. Un script PHP collecte l'information ajoutée à l'image SVG (l'identifiant des régions présentes dans l'image, la surface de chaque élément sur l'image) et les enregistre dans la base de données.

Ces images interactives SVG sont des représentations hybrides combinant trois types d'informations : l'image 2D, les éléments 3D représentés à son intérieure et leurs attributs sémantiques. Ces représentations permettent de voir les sources iconographiques d'une manière inédite. La visualisation de ces informations peut en effet assister les spécialistes de la documentation architecturale pour étudier les sources iconographiques, et peuvent aussi permettre au grand public de mieux comprendre l'architecture patrimoniale.

Le diagramme dans la figure 64 représente les dialogues entre les parties du processus de téléchargement, le référencement spatial et la projection de couche sémantique.

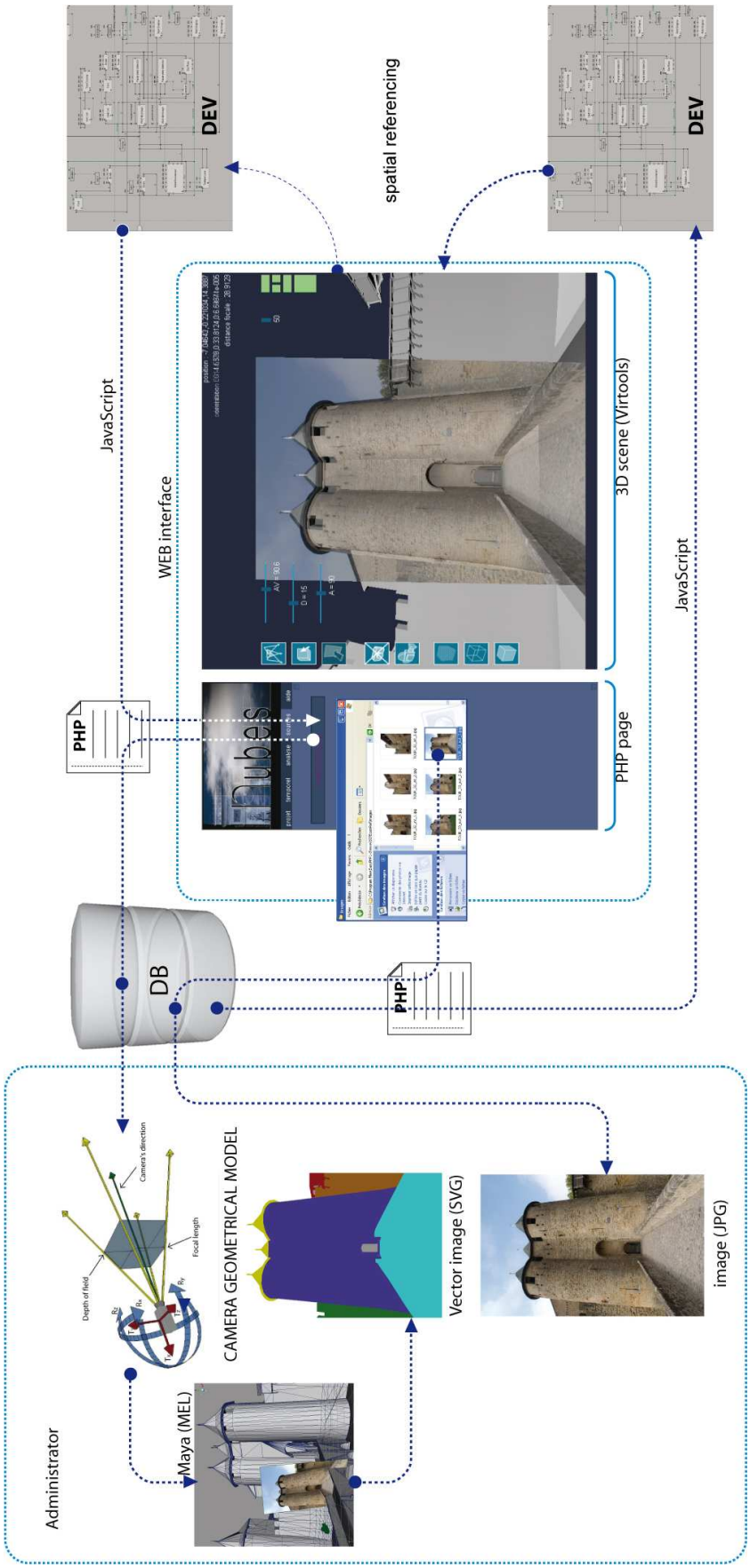


Figure 64 Dialogue entre les parties du système pour les processus de téléchargement, de référencement spatial et de projection des couches sémantiques

7. Un système pour la recherche visuelle de sources iconographiques

Dans ce chapitre nous allons présenter un système pour la recherche visuelle de sources iconographiques qui exploite à fond toute la richesse informationnelle que nous ajoutons aux images au travers de l'approche décrit dans les chapitre 4, 5 et 6.

Dans le domaine de la documentation architecturale et archéologique, différents travaux de recherche ont produit des systèmes pour la consultation d'archives d'images en utilisant de l'information 3D. Par exemple [Kodabayashi R. et al. 2005] proposent une méthode pour la recherche de photographies qu'utilise des points de vue tridimensionnels comme requêtes. La recherche basée sur les points de vue 3D est utilisée pour recherche différentes images relatives au même objet archéologique. Le système défini le point de vue d'une image comme « la position et l'orientation de l'appareil à partir desquels la photo a été prise ». Comme la réalité est tridimensionnelle, la position et orientation de la caméra servent à fournir des coordonnées 3D. Le point de vue est associé à l'image sous forme de métadonnées et enregistré dans une archive numérique. La recherche d'images est conduite alors en s'appuyant sur la similarité du point de vue.

Un type de recherche dans Photosynth [Snavely N., 2006] est "montre-moi les photographies de cet objet", où l'objet en question peut être sélectionné directement dans une photographie ou dans le nuage de points. Ce type de recherche est appliqué aux vidéos dans [Sivic J., 2003]. Avec ce système d'exploration de photos, l'utilisateur sélectionne un objet en dessinant un rectangle entourant une région de l'image courante ou dans un nuage de points. Tous les points se situant à l'intérieur de l'enveloppe sont sélectionnés. La fonction que classe les images trouvées est base sur trois critères : 1) les points sélectionnés sont visibles, 2) l'objet est vu à partir d'un bon angle, 3) l'objet apparaît dans ses détails.

L'exploration des images dans l'espace 3D a également été étudiée dans ce projet. L'utilisateur peut déplacer la caméra virtuelle en toutes les directions et peut aussi contrôler le zoom et la rotation. Ceci permet à l'utilisateur de se déplacer librement dans la scène et fournit une voie simple pour recherche des points de vue intéressant et trouver des photographies correspondantes. A tout moment, l'utilisateur peut cliquer sur une pyramide visuelle dans la vue principale et la camera virtuelle se déplace jusqu'au point de vue de l'image sélectionnée.

Dans notre travail, nous avons développé un système pour aider les architectes, les archéologues et les documentalistes dans la tache de classification, de gestion et de visualisation des fonds iconographiques relatifs aux monuments historiques en exploitant les résultats de notre méthodologie.

Les bases de données permettent aujourd'hui de gérer plusieurs types d'informations nécessaires à différents supports de documentation. De plus, des nouveaux moteurs d'affichage 3D permettent une interaction spatiale en temps réel avec les nuages de points et les maquettes résultantes des phases de traitement. Aujourd'hui, il apparaît donc essentiel de s'intéresser aux relations que l'on peut établir entre les informations métriques et photographiques à l'intérieur d'une scène 3D.

Ceci est le propos à la base du développement de "Nubes Imago", un système de consultation sur Internet pour la recherché visuelle de sources iconographiques relative à la morphologie des édifices patrimoniaux. Les fonctionnalités actuellement implémentées au sein de ce système sont l'accès et la manipulation de reconstructions tridimensionnelles d'édifices (modèles 3D, nuages de points et photographies) au sein d'une scène 3D interactive. L'objectif de ce projet est de fournir aux utilisateurs (architectes, archéologues, historiens, grand public, etc.) une manière effective pour accéder à l'information (en particulier à l'iconographie) relative d'un édifice historique en fonctions de critères spatiaux, morphologiques et sémantiques.

7.1. Architecture du système

Comme notre système est un module de la plateforme NUBES [De Luca et al. 2010], l'architecture du système s'inscrit dans la structure générale de ce projet. Il s'agit d'une base de données développée en MySQL, une scène 3D interactive développée en Virtools DEV, puis une application Web développée en PHP. La scène 3D est intégrée à l'intérieur de l'application Web sous forme d'un plug-in qui communiqué avec l'application PHP/MySQL au travers de commandes Javascript. La figure 65 explique le dialogue entre les parties de notre système. Les pages dynamiques PHP de l'application sont utilisées comme support pour la formulation des requêtes, ainsi que pour l'insertion et la modification des données iconographiques stockées dans la base de données MySQL.

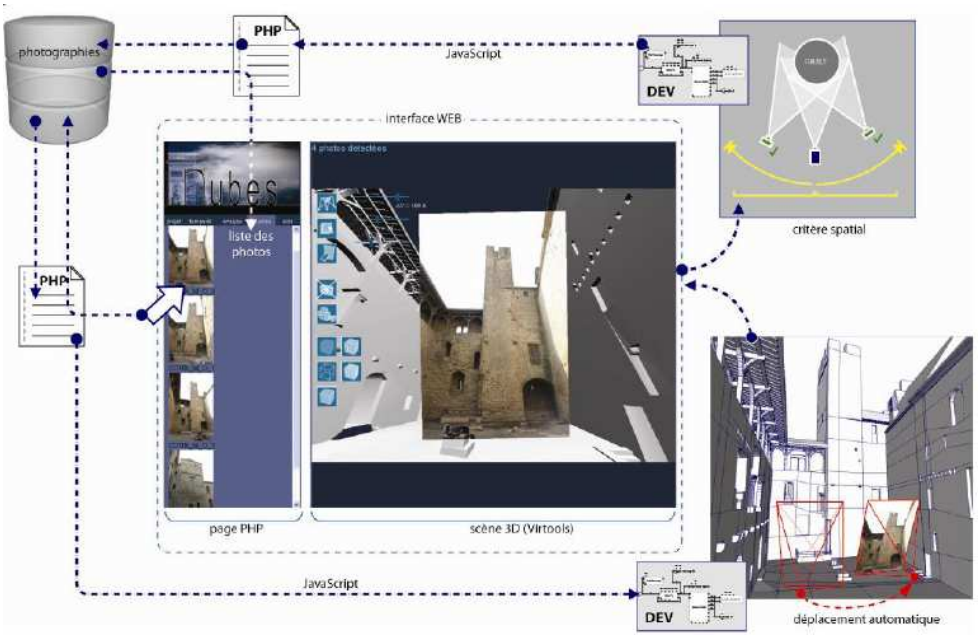


Figure 65 Dialogue entre les parties du système

7.1.1. Base de données

En relation avec la structure générale de la plateforme NUBES, notre système permet la gestion de multiples projets et de multiples utilisateurs. Ceci nous permet d'identifier chaque image et représentation 3D présents dans la base au travers d'un attribut correspondant à un projet spécifique (et dans certains cas à un utilisateur spécifique). Au lancement du système, les seules données relatives au projet sélectionné (et à l'utilisateur concerné) sont téléchargées dans la scène 3D.

Pour le stockage des données de référencement spatial des sources iconographiques, des tables spécifiques sont crées dans la base de données MySQL et concernent quatre catégories d'informations pour chaque source.

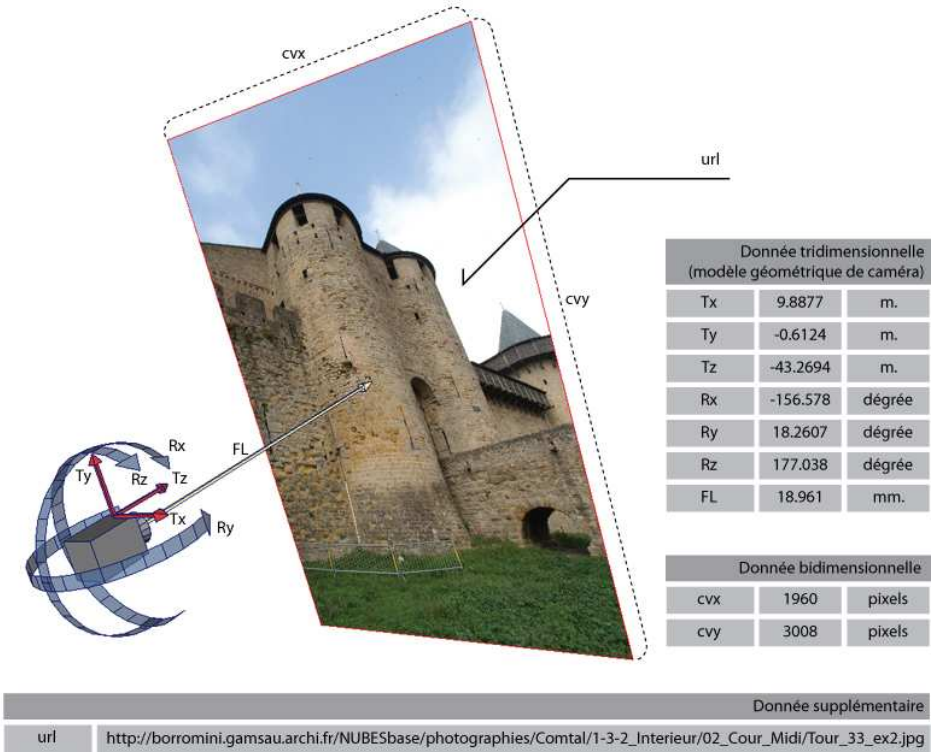
Données tridimensionnelles (modèles géométriques des caméras) : Ces valeurs sont celles obtenues par les différentes méthodes de référencement spatial (manuel, géométrique, automatique, etc.) que nous avons décrits dans le chapitre précèdent. La scène 3D utilise ces informations pour la rechercher d'images basées sur des critères spatiaux.

Les données tridimensionnelles contenues dans la base sont :

- la position du point de vue de chaque image (stockée dans des champs Tx,Ty,Tz). Ces valeurs sont utilisées pour le déplacement de la caméra de navigation jusqu'au point de vue de la photo sélectionnée ;
- l'orientation du point de vue de chaque image (stockée dans des champs Rx,Ry,Rz). Ces valeurs sont utilisées pour la rotation de la caméra de navigation jusqu'au point de vue de la photo sélectionnée ;
- type de projection. Cette valeur est utilisée pour modifier le système de projection de la caméra (parallèle, perspectif) en fonction du type de source iconographique sélectionnée ;
- la distance focale de chaque image (stockée dans un champ FL). Cette valeur est utilisée pour calculer la distance du plan image 2D de l'image par rapport à la position du point de vue de l'image sélectionnée.

Données bidimensionnelles. Ces données sont relatives aux dimensions des sources iconographiques. La largeur et la hauteur sont stockées dans deux champs différents. Cette information est utilisée pour calculer la proportion de l'image 2D associée à la caméra de navigation.

- largeur (stockée dans un champ cvx) et hauteur (stockée dans un champ cvy) en pixel de chaque photo



Donnée tridimensionnelle (modèle géométrique de caméra)		
Tx	9.8877	m.
Ty	-0.6124	m.
Tz	-43.2694	m.
Rx	-156.578	dégrée
Ry	18.2607	dégrée
Rz	177.038	dégrée
FL	18.961	mm.

Donnée bidimensionnelle		
cvx	1960	pixels
cvy	3008	pixels

Donnée supplémentaire	
url	http://borromini.gamsau.archi.fr/NUBESbase/photographies/Comtal/1-3-2_Interieur/02_Cour_Midi/Tour_33_ex2.jpg

Figure 66 Données relatives à un point de vue relatif à une source iconographique

Données additionnelles comme l'URL de l'image originale (utilisée pour le téléchargement de l'image dans la scène 3D) et la méthode de référencement spatial (utilisée si l'image a déjà été spatialisée ou non et le type de référencement).

Données descriptives. Ces données sont rentrées par l'utilisateur au moment du téléchargement sur serveur des sources iconographiques. Ces données concernent :

- le titre de l'image
- une légende descriptive
- l'auteur
- la date de prise de vue ou d'exécution
- le type de support
- le service de production
- les crédits de l'image
- les droits de diffusion
- les liens vers des informations supplémentaires
- les liens vers des références bibliographiques

Données partagées. Ces données sont des clés pour le dialogue entre les parties du système. La scène 3D temps réel envoie ces données comme critères pour les requêtés SQL aux pages PHP afin de sélectionner et de visualiser les images correspondantes. Quand l'utilisateur sélectionne une image, la page PHP renvoie l'identifiant de l'image sélectionnée à la scène 3D et les valeurs correspondantes (retrouvés à partir des clés) sont utilisées pour retrouver les points de vue de l'image.

Les données partagées sont :

- l'identifiant de l'image
- le nom de l'image
- l'identifiant de l'entité morphologique
- nom et description de l'entité morphologique
- attributs sémantiques de l'entité morphologique.

7.1.2. L'interface WEB du système de la consultation

Notre interface est développée en Virtools DEV (scène 3D interactive) et PHP (pages web dynamiques). Cette interface permet aux utilisateurs de naviguer dans une scène 3D en temps réel et d'observer en détails des reconstructions 3D d'édifices. A chaque moment, notre système peut formuler des requêtes d'images en utilisant différents critères. Une fois reçue la requête de l'utilisateur, le système répond en visualisant les images correspondantes. Les capacités du système sont directement liées aux relations projectives établies entre la représentation 3D de la morphologie de l'édifice et l'ensemble des sources iconographiques 2D spatialisées. La figure 67 illustre le diagramme d'utilisation de notre interface.

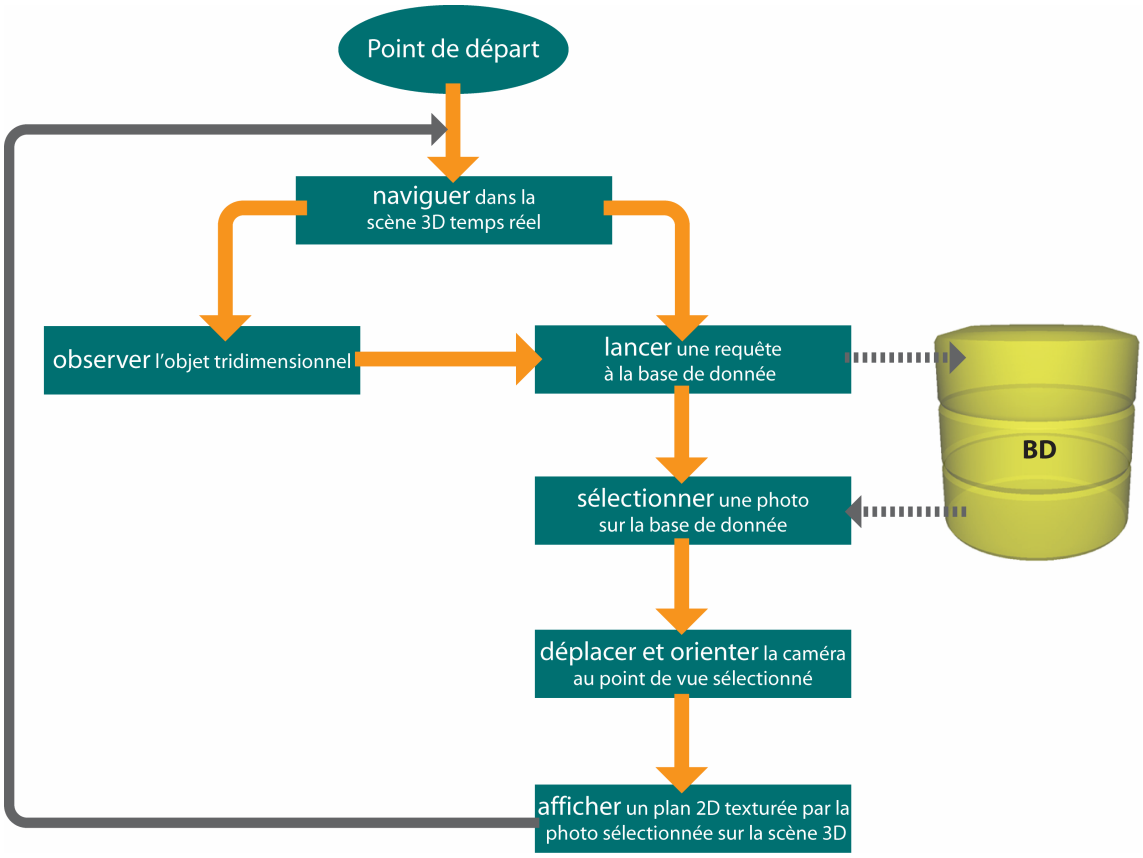


Figure 67 Diagramme de l'utilisation des fonctions d'interface

7.1.3. Fonctions pour la recherche d'images

L'utilisateur de notre système peut formuler des requêtes de différentes manières en exploitant toute l'information contenue dans la représentation tridimensionnelle de l'édifice, sa structuration sémantique ainsi que dans les images spatialisées. Les résultats du processus d'annotation sémantique de la morphologie de l'édifice et des images permet de formuler des requêtes sur :

- Les attributs sémantiques des éléments architecturaux qui sont présent dans les images (résultat de l'annotation sémantique)
- Les éléments présents dans l'image (résultats de la projection des contours des éléments 3D dans les images)
- Le point de vue de l'image dans l'espace 3D (résultat du référencement spatial)

7.2. Recherche d'images basée sur critères spatiaux

Cette méthode de recherche permet la formulation de requête dans la base de données iconographiques en fonction de critères spatiaux. Cette fonction est basée sur la sélection d'informations sur les images à partir du point d'observation dans la scène 3D. Afin de rechercher les sources iconographiques correspondantes à la position et à l'orientation courante de la caméra de navigation, nous utilisons l'intersection entre la pyramide visuelle de la camera de navigation et les pyramides visuelles relatives aux points de vue des images spatialisées (voir la figure 68, gauche). Cette méthode est conçue pour permettre aux utilisateurs de retrouver les images qui contiennent la même portion d'espace qu'il est en train d'observer dans la scène 3D. Ceci permet de visualiser des ensembles d'images ayant un point de vue similaire.

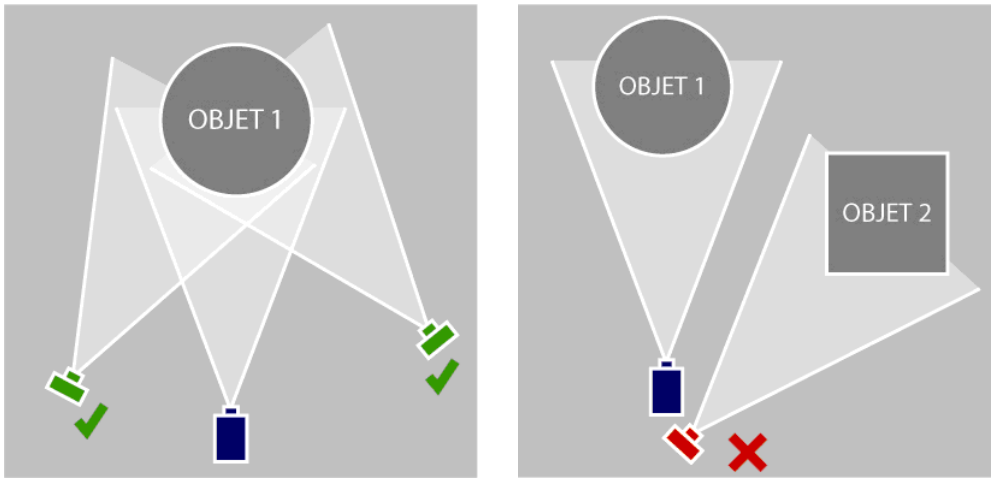


Figure 68 Méthode de détection d'images basée sur l'intersection de pyramides visuelles

Les images dont la pyramide visuelle ne se trouve pas en intersection avec celle relative à la camera de navigation sont exclues de la sélection (voir figure 68, droite).

Nous paramétrons la détection des images similaires au point d'observation en fonction de trois critères distincts :

- l'angle de vue de la camera de navigation (pour contrôler l'échelle de la pyramide visuelle) ;
- la distance entre la position de la caméra de navigation et la position du point de vue de chaque image ;
- l'extension de l'angle entre l'orientation de la caméra de navigation et celle du point de vue de chaque image.

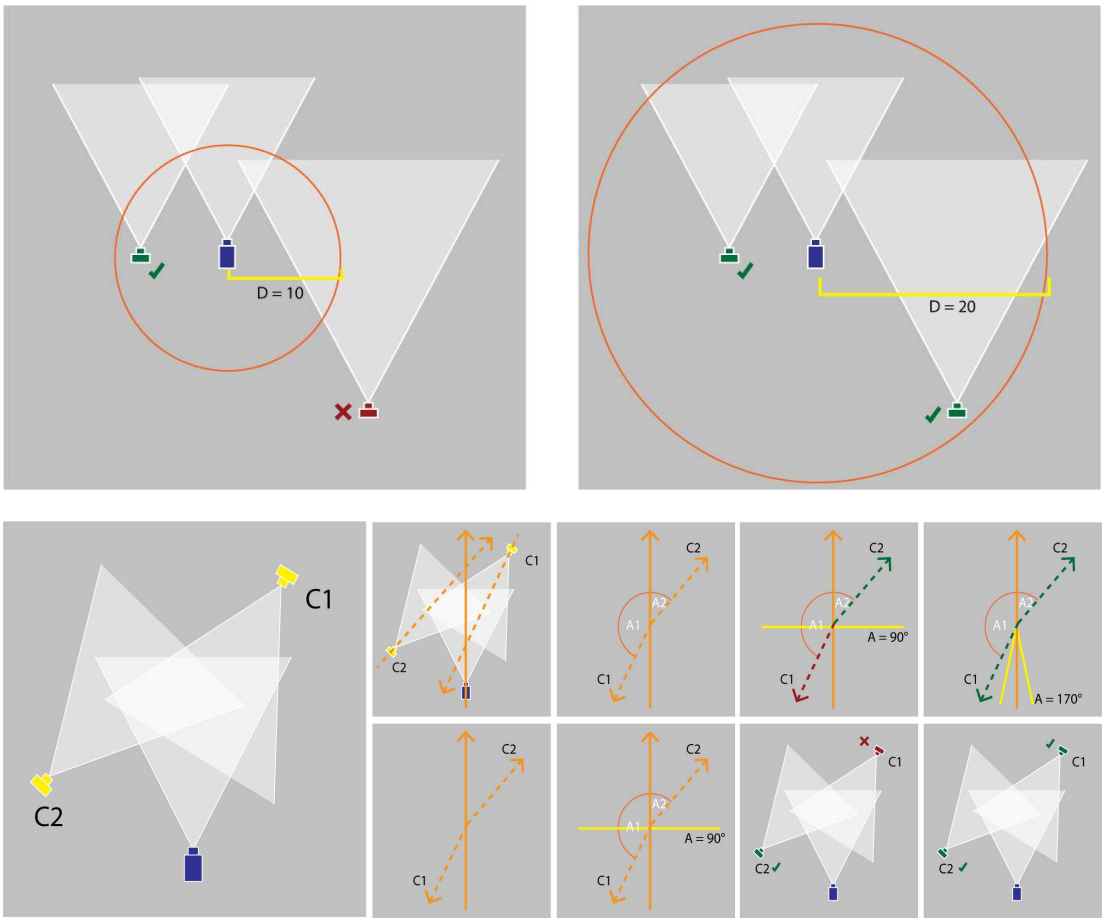


Figure 69 Paramètres du critère spatial. En haut : distance, en bas : différence d'orientation

Notre système permet à l'utilisateur de configurer ces 3 paramètres à tout moment lors de la navigation pour ne pas limiter la flexibilité d'utilisation. L'utilisateur peut choisir l'angle de champ de la camera de navigation entre 50 à 120 degrés (au delà de ces angles, la visualisation serait trop distordue), la distance entre sa position et celle des images à recherches entre 0.5 et 50m (valeur établie en fonction de l'échelle architecturale), la différence d'orientation entre 0 et 180 degrés.

En appliquant cette méthode, l'utilisateur peut affiner progressivement ses requêtes, par contre, si l'utilisateur introduit des paramétrages trop fins, il pourrait obtenir des résultats pas pertinents ou des pertes importantes (voir figure 70).

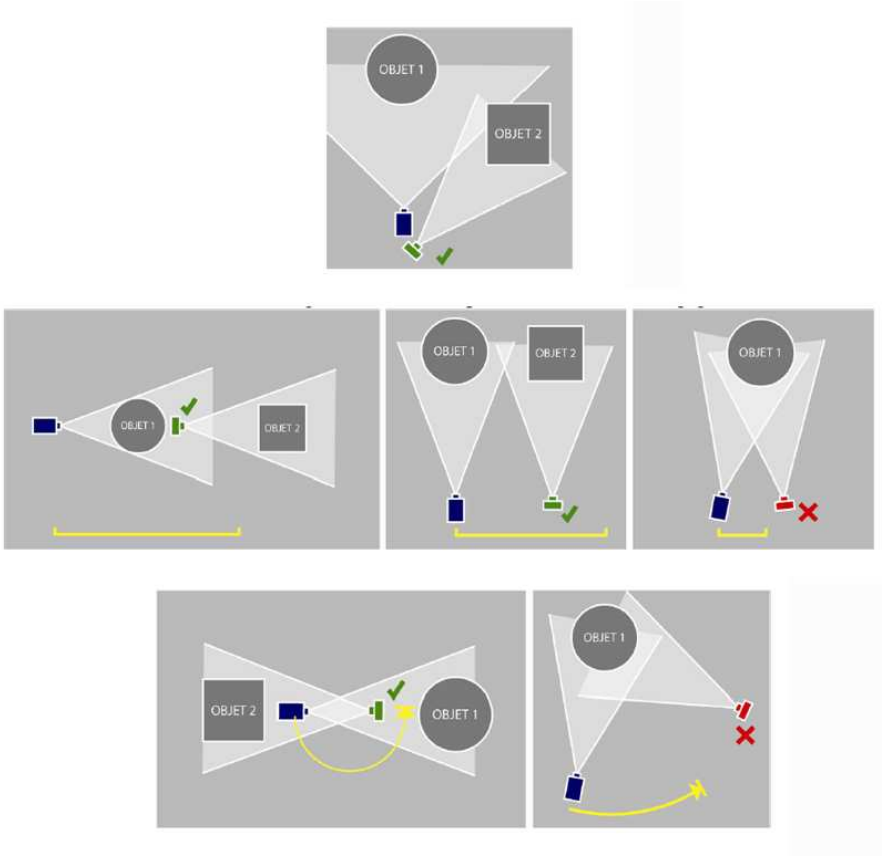


Figure 70 Exemples d'échec de détection : cas où la valeur de la limite introduite par l'utilisateur est trop grande ou trop petite.

Après plusieurs tests, les paramètres par défaut du système (pour une détection efficace et pertinente d'images similaires au point d'observation) ont été défini ainsi: angle de champ de la camera de navigation = 90.6 degrés, distance entre camera de navigation et points de vues des images = 15 m, différence d'orientation entre la camera de navigation et points de vue des images = 90 degré.

Pour donner un exemple d'ajustement de ces trois paramètres, si l'on formule une requête : la détection de toutes les photographies dont les champs visuels recoupent celle de la caméra de navigation, à une distance inférieure à 50 m (valeur maximale) mais dont la limite de l'angle est égal à 0 (très précis) ; la base de donnée affichera seulement les photos qui ont exactement la même orientation que celle de la caméra de navigation (voir la figure 71 gauche).

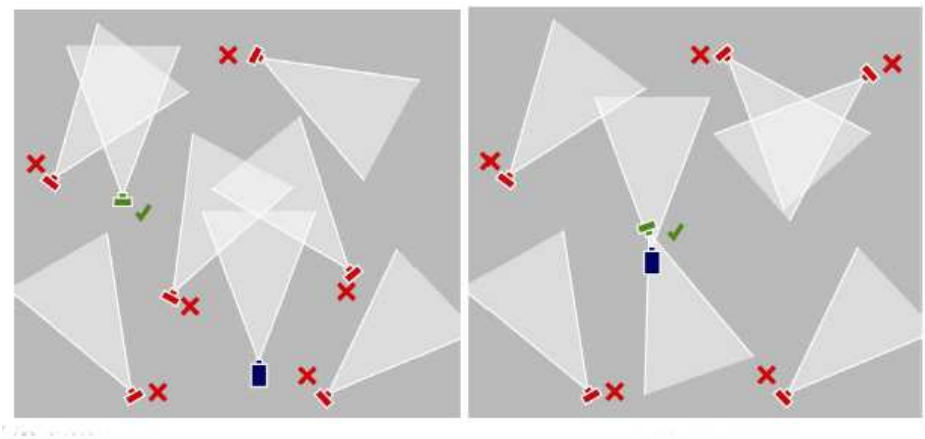


Figure 71 Combinaison de requêtes

On peut tout aussi bien formuler la requête en paramétrant une distance précise et une différence d'orientation limite très étendue (voir la figure 71 droite).

Ces paramètres sont également utilisés pour trier les résultats. Une fois une requête formulée, l'affichage des résultats se base en effet sur un ordre défini par la distance et la différence d'orientation (on calcule seulement l'orientation d'axe X (pitch) et Y (yaw)) entre les points de vue des images et de la caméra de navigation. L'image répondant le mieux à la requête s'affiche donc en première position.

Du point de vue de l'implémentation informatique, quand on clique sur le bouton pour lancer une requête, la liste des identifiants des photos sélectionnées par la procédure ci-dessus décrite est envoyée depuis la scène 3D à une page PHP qui récupère les informations relatives depuis la base de données MySQL. Ce type de recherche est particulièrement utile pour la recherche de sources iconographiques associées aux espaces architecturaux (cours, entrées, galeries, etc.)

7.3. Recherche d'images basée sur critères morphologiques

La possibilité de sélectionner un élément architectural est particulièrement utile quand on explore une scène architecturale 3D. Quand en navigant l'utilisateur s'arrête face à un élément intéressant, la sélection de cet élément est une voie intuitive pour rechercher les sources iconographiques le concernant. Plusieurs recherches proposent des solutions pour visualiser des informations documentaires associées à un élément morphologique [Snavely N. et al., 2006; Herbig U et al., 1998]. Ces méthodes permettent aux utilisateurs de sélectionner des représentations d'éléments architecturaux (comme des portions de nuages de points ou des maquettes 3D) et de lancer une requête pour trouver des photos dans la base de données correspondantes à l'objet sélectionné. Par contre, ces travaux se basent sur une association element-source documentaire définie préalablement.

Dans notre système, à cote des critères spatiaux nous proposons une nouvelle méthode d'accès aux sources iconographiques basée sur la sélection d'éléments de la morphologie de l'édifice.

Compte tenu du fait que la présence de chaque entité 3D de la morphologie de l'édifice est annotée sur les images au travers de la projection de son contour (voir chapitre 6), notre système permet de retrouver toutes les images dans lesquelles cette entité est présente.

Cet aspect explicite à fond l'idée d'utiliser la maquette numérique comme moyen d'accès aux informations patrimoniales relatives à la morphologie d'un édifice.

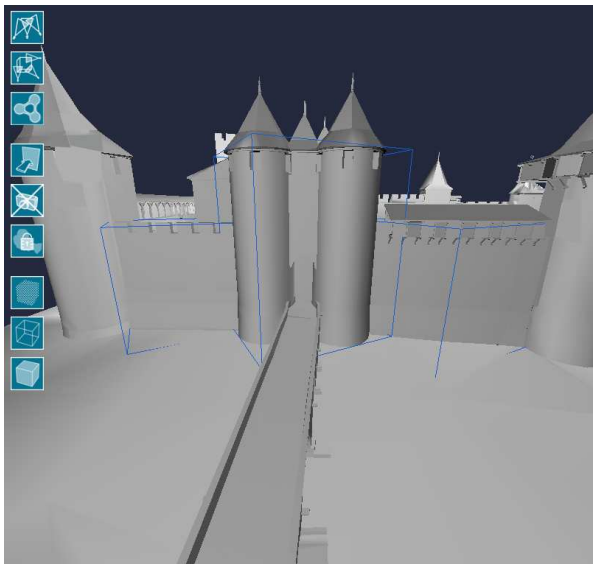


Figure 72 Sélection de multiples objets dans la scène 3D

Dans cette modalité de recherche, l'utilisateur peut sélectionner un objet 3D dans la scène interactive. Une sélection multiple est admise (voir figure 72). Une fois les objets sélectionnés, le système envoie l'identifiant de(s) objet(s) sélectionné(s) à une page dynamique PHP. Cette page affiche alors toutes les images dont la couche sémantique SVG associée contient ce(s) identifiant(s).

Les résultats sont triés en fonction de l'extension de la superficie de la région polygonale SVG (relative à l'entité ou aux entités sélectionnées) associée aux images (voir section 6.4.2.4). L'image dans laquelle l'entité occupe la superficie majeure est affichée en premier. L'extension de la présence d'un élément dans les images est aussi utilisée comme paramètre pour affiner la recherche. L'utilisateur peut en effet indiquer une valeur (par exemple plus de 40% de la superficie de l'image) pour rechercher seulement les images où l'élément sélectionné est en premier plan.



Figure 73 Résultat de la requête basée sur la sélection de multiples objets.

7.4. Recherche d'images basée sur critères sémantiques

La recherche d'images basée sur des critères sémantiques est l'objectif principal de notre travail. Au vu de l'intégration réalisée dans notre système d'informations métriques, morphologiques et sémantiques, cette méthode permet de rechercher des sources iconographiques reliées à des attributs utilisés pour qualifier les entités architecturales qui composent l'édifice. Cela permet à l'utilisateur de retrouver par exemple toutes les images contenant des éléments qui partagent des attributs avec un ou plusieurs éléments sélectionnés dans la scène 3D ou sur une image.

Pour configurer cette recherche, au moment de l'ouverture d'un projet, le système classe tous les attributs sémantiques associés aux éléments de la morphologie (termes de thesaurus utilisés pour décrire les éléments, attributs spécifiques à une analyse, etc.) et crée des listes consultables par l'utilisateur.

A partir de la sélection d'un attribut sémantique, toutes les images contenant des éléments auxquels cet attribut a été associé seront affichées. La requête permet aussi la sélection de plusieurs attributs à la fois. Par exemple, l'utilisateur peut rechercher toutes les sources iconographiques où des éléments de type "tour" sont présents à coté d'éléments de type "hourd".



Figure 74 Résultat d'une requête basée sur critère sémantique.

Ce niveau de précision et de paramétrage n'est jamais été atteint au sein d'une base de données d'informations patrimoniales.

Suite à la formulation de la requête, le système sélectionne les images dans la base de données qui correspondent aux critères et copie chaque couche vectorielle (SVG) associée aux images sélectionnées pour créer un ensemble de fichiers image temporaires. Ces fichiers sont édités à la volée et les régions vectorielles des images correspondantes aux entités sélectionnées sont colorées. Cette fonction permet aux utilisateurs de connaître ainsi la position des éléments recherchés dans les images trouvées. Dans ce type de recherche aussi, les résultats de la requête sont classé en fonction de l'extension des la superficie de la région identifiée sur l'image.

Ce type de recherche peut être particulièrement utile dans un contexte de conservation. En effet, si la description de la morphologie de l'édifice répertorie les phénomènes de dégradation qui affectent l'édifice, on pourrait utiliser ce type de recherche pour trouver par exemple toutes les images où un type spécifique de dégradation est présent, ou encore toutes les images où on peut observer un type spécifique de dégradation sur un type spécifique d'élément architectural.

7.5. Récupération de point de vue image

A tout moment de son utilisation, notre système permet de retrouver le point de vue relié à une source iconographique dans l'espace 3D. Cette fonction permet d'expliciter visuellement la relation entre l'image sélectionnée et la scène 3D.

Quand l'utilisateur sélectionne une image, la procédure de dialogue entre la base de données et la scène 3D est utilisée pour la recherche du point de vue correspondant dans la scène 3D. Il s'agit de reproduire les paramètres d'une caméra à projection perspective ou orthogonale. Les paramètres intrinsèques et extrinsèques de la caméra associée à l'image sont transmis à une fonction interne de la scène 3D qui calcule l'interpolation entre la position et l'orientation courantes de la caméra de navigation et les valeurs relatives à l'image sélectionnée.

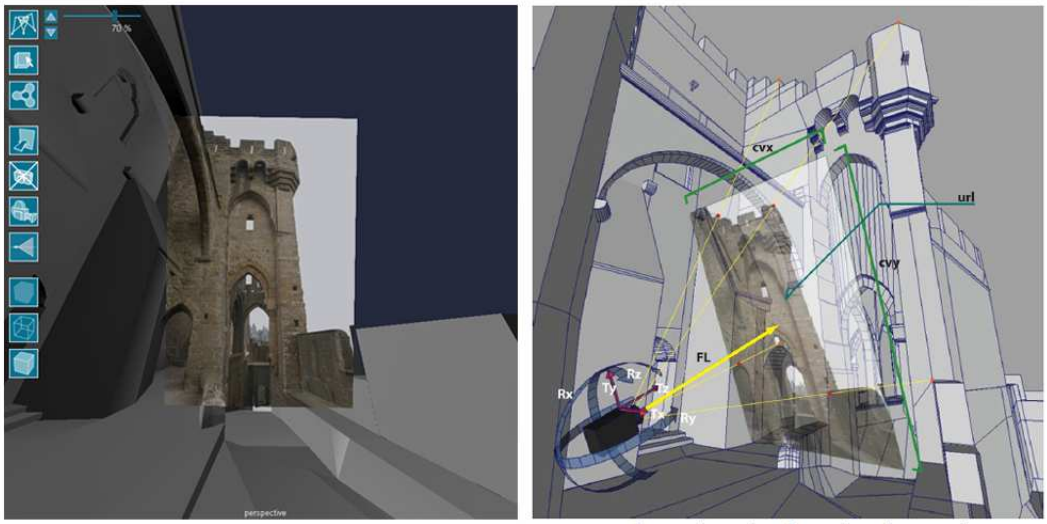


Figure 75 Récupération du point de vue de la source iconographique sélectionnée

Dès que le déplacement (de la position courante au point de vue de l’image) est terminé, le système affiche automatiquement un plan 2D orienté de façon normale à la caméra de navigation, texturé par la photo sélectionnée et calée sur la scène 3D. Pour visualiser correctement l’image à l’intérieur de la scène 3D, on utilise la valeur de distance focale (ou de zoom dans le cas d’une projection orthogonale) de l’image sélectionnée afin de calculer la distance entre le plan 2D et la position de la caméra de navigation.

8. Conclusions et perspectives

Notre travail s'est concentré sur la conception et le développement d'une méthodologie pour l'annotation sémantique de sources iconographiques par référencement spatial. Le point central de ce travail était d'exploiter la relation spatiale entre l'information 3D (maquette numérique) et l'information 2D (sources iconographiques). Notre méthode permet de superposer des éléments visuels 2D sur une scène 3D avec différents niveaux de précision, en relation aux différents degrés de cohérence géométrique existants entre l'édifice réel, sa reconstruction tridimensionnelle et l'ensemble de sources iconographiques bidimensionnelles que l'on peut collecter dans les archives (photos, dessins, peintures, etc.). La relation spatiale établie entre les informations 2D et 3D est enfin utilisée pour véhiculer les annotations sémantiques de la maquette numérique de l'édifice vers les sources iconographiques reliées.

La précision de la segmentation et de l'annotation sémantique des images dépendent principalement du processus de référencement spatial. Et le niveau de segmentation des images dépend essentiellement du niveau de décomposition de la représentation 3D. Ceci permet d'envisager de conduire de multiples segmentations et annotations sémantiques (en 3D comme en 2D) en fonction d'exigences d'analyse et d'étude particulières. De la même manière il est possible d'imaginer de projeter (du 3D vers le 2D) plein d'autres informations qui peuvent enrichir la description sémantique de la morphologie de l'édifice.

La solution proposée est en effet parfaitement répondante à des mises à jour de la structuration de la maquette 3D comme à l'ajout de nouvelles sources documentaires.Une maquette informatique du système de recherche d'images se base sur des critères spatiaux, morphologiques et sémantiques, ce qui a été développée sous forme d'application Web. Ses fonctionnalités sont disponibles d'hors et déjà pour les partenaires potentiels qui voudrons commencer à l'expérimenter dans des applications concrètes. Le système donne la possibilité d'accéder à une information riche d'une manière innovante et pourrait facilement être utilisé par des chercheurs, des spécialistes de la documentation et de la conservation du patrimoine ainsi que par le grand public.

En comparant notre système avec d'autres travaux de recherche sur l'annotation sémantique d'images, notre système requiert moins de manipulations que des solutions basées sur des méthodes manuelles (basées sur l'insertion systématique de mots-clés), et il est bien plus précis (en termes de détail et de paramétrage) que des solutions basées sur des méthodes automatiques (reconnaissances des formes).

De plus, notre système associe à l'image toute la richesse informationnelle contenue dans la représentation tridimensionnelle (dimensions, géométrie, etc.).

8.1. Limites

Notre méthodologie d'annotation sémantique d'images se base sur le référencement spatial. C'est la raison pour laquelle la qualité du référencement est vraiment importante. Or, certains types de sources iconographiques sont particulièrement difficiles (parfois impossibles) à spatialiser à cause de leur grande incohérence géométrique par rapport à la réalité. Au delà de l'incohérence, il peut également se présenter le cas d'un manque d'information dans la source 2D ou dans la représentation 3D. Ces manques peuvent causer des problèmes dans les deux sens:

- Un élément architectural peut être présent dans l'image mais pas dans le modèle 3D. Cette incohérence peut apparaître à cause des limites des processus de relevé 3D. Les édifices patrimoniaux peuvent avoir des extensions spatiales très importantes, et les éléments d'un édifice peuvent aussi être intègres à d'autres édifices ou à d'autres éléments urbains. Les limites de la reconstruction tridimensionnelle d'un édifice peuvent exclure des parties d'édifice qui apparaissent dans une image. Comme illustré dans la figure 76 la partie droite du dessin est absente dans le modèle 3D.

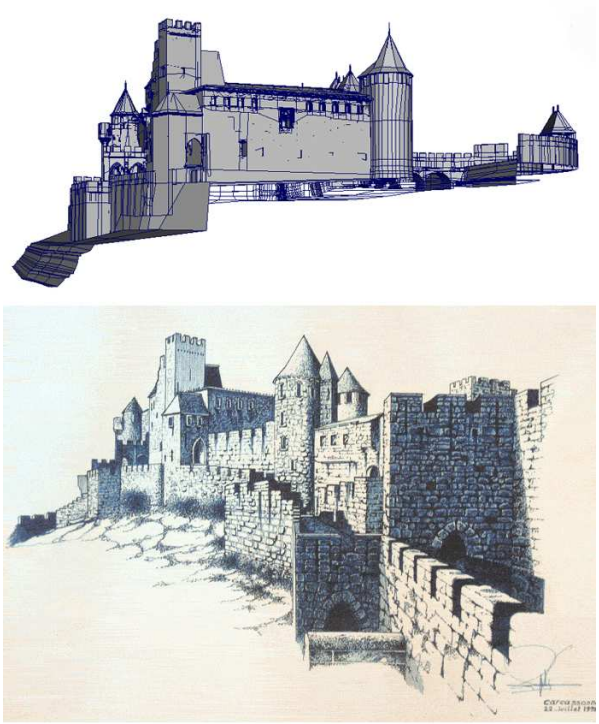


Figure 76 Incohérence entre le modèle 3D et l'image

- Un élément architectural (ou un ensemble d'éléments) peut également être présent dans le modèle 3D et absent dans l'image. Ceci peut se vérifier quand l'information présente dans l'image n'est pas aussi détaillée ou cohérente comme celle contenue dans la représentation 3D.

Par rapport à la projection, certaines sources iconographiques (notamment celle qui ne se basent pas sur des construction perspectives) présentent aussi des représentations qui échappent complètement à une cohérence formelle avec la réalité (en termes de morphologie et de proportions à la fois)(voir figure 77).

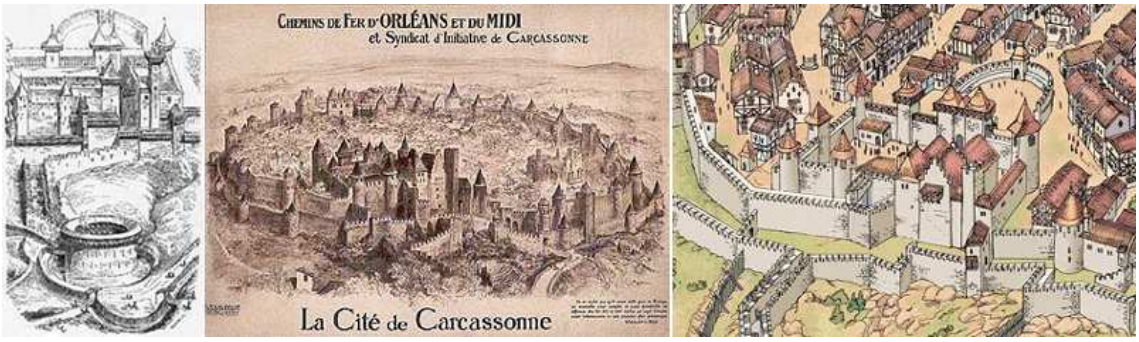


Figure 77 Exemples des représentations échappe à une cohérence formelle avec la réalité (images du château Comtal, Carcassonne).

A cet aspect il faut également ajouter la dimension subjective que chaque auteur d'une source peut ajouter dans l'exécution de l'œuvre (voir figure 78). Des éléments représentés pourraient être complètements inventés ou supposes.



Figure 78 Exemples des images de point de vue subjective (images artistiques du château Comtal, Carcassonne).

Finalement, en ce qui concerne la dimension temporelle, il est évident que notre système ne peut prendre en considération que la mise en relation spatiale de sources 2D et modèles 3D qui représentent un même état temporal de l'édifice. Cette limite ouvre les champs des perspectives de recherche.

.

8.2. Perspectives de recherche

En fonction de la dernière limite identifiée dans le paragraphe précédent, la première perspective de recherche en vu d'une évolution de notre travail concerne la prise en compte d'un référencement temporal des sources iconographiques permettant de décliner des multiples représentations 3D de l'édifice en fonctions des transformations que l'édifice a subi dans le temps. Dans ce contexte, le rôle que le référencement spatial de sources iconographiques peut jouer pour la reconstruction des transformations temporelles des édifices patrimoniaux a fait l'objet d'un article que nous avons publié récemment [De Luca L. et al. 2010].

Une deuxième perspective de recherche concerne l'amélioration de notre méthode de référencement spatial (en particulier pour ce qui concerne les photographies). Comme nous l'avons bien expliqué dans le chapitre 6, pour un référencement géométrique, nous utilisons un algorithme de résection spatial qui nécessite d'un certain nombre de points homologues (en 2D et en 3D) pour déterminer le point de vue de l'image sur le modèle 3D. Or, différentes solutions d'analyse d'images (notamment la détection de lignes caractéristiques, de discontinuités, de parallélismes, etc.) pourraient être utilisées pour conduire un recalage automatique qui exploiterait aussi l'information géométrique contenue dans la maquette 3D (sommets, arêtes, normales, etc.). On pourrait par exemple imaginer de faire intervenir ces automatismes suite à un calage manuel approximatif de l'image sur le modèle 3D afin de bénéficier de la similarité des points de vue.

Une dernière perspective de recherche concerne la relation projective 2D/3D. Dans notre travail nous avons exploitée cette relation seulement dans un sens : on qualifie les images 2D à partir d'une maquette 3D sémantisée. Or, au vu des résultats récents du domaine de la photogrammétrie et de la vision par ordinateur en matière de mise en correspondance, calibration et orientation automatique d'ensemble d'images photographiques, cette relation pourrait bien être exploitée dans le sens inverse : conduire la caractérisation sémantique (segmentation et l'annotation) sur les photos pour structurer (par projection) une représentation 3D (nuages de points, maillages polygonaux, etc.) construite à partir des images.

Bibliographie

[AbdelAziz Y.I. et al., 1971] AbdelAziz Y.I., Karara H.M., Direct Linear Transformation from comparator coordinates into object space coordinates in close range photogrammetry. Proceedings of the ASP/UI.1971

[Abmayr T. et al., 2004] Abmayr T., Härtl F., Mettenleiter M., Heinz I., Hildebrand A., Neumann B., Fröhlich C. Realistic 3D reconstruction – combining laserscan data with RGB color information. Proceedings of ISPRS Internation Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol XXXV, Part B. Istanbul 2004. pp.198-203, 2004

[Akcay H.G. et al., 2007] Akcay H.G., Aksoy S. Automatic Detection of Geospatial Objects Using Multiple Hierarchical Segmentations, in IEEE Transactions on Geoscience and Remote Sensing, 46(7):2097-2111, July 2008.

[Aksoy S., 2006] Aksoy S., Dogrusoz E., Modeling urbanization using spatial building patterns, in Proceedings of 4th IAPR International Workshop on Pattern Recognition in Remote Sensing, Hong Kong, August 20, 2006.

[Almagro A. et al., 1996] Almagro A., Patias P., Waldhäusl P. The CIPA Otto Wagner Pavilion Test. International Archives of Photogrammetry and Remote Sensing Vol.XXXI Part B5, pp. 463 – 470, 1996

[American Society of Photogrammetry, 1980] American Society of Photogrammetry, Manual of Photogrammetry, Fourth Édition, 1980

[Attene M. et al. 2007] Attene M., Robbiano F., Spagnuolo M. and Falcidieno B., Semantic Annotation of 3D Surface Meshes based on Feature Characterization, Lecture Notes in Computer Science, Vol. 4816 (SAMT'07 Procs.), pp. 126-139, 2007

[Attene M. et al., 2006] Attene M., Biasotti S., Mortara M., Patane G., Falcidieno B. : Computational methods for understanding 3D shapes. Computers & Graphics 2006; 30(3): 323-33, 2006

[Azzi Y., 2008] Azzi Y.:Images et contenus : classification et sémantisationde sources documentaires visuelles, MIAjournal, 2008

[Bhagavathy S. et al., 2007] Bhagavathy S., Manjunath B.S., Modeling and detection of geospatial objects using texture motifs, IEEE Transactions on Geoscience and Remote Sensing, 2007

[Blaise J.Y. et al., 2004] Blaise J.Y., De Domenico F., De Luca L., Dudek I: Acquisition de données vs gestion de connaissances patrimoniales : le cas des vestiges du théâtre antique d'Arles, 2004

[Blaise J.Y. et al., 2007] Blaise J.Y., Dudek I.:Terminologie du corpus architectural : matière, forme ou idées ? Dans Actes de la Conférence TIA'07 - 7ème Conférence TIA 2007, Terminologie et Intelligence Artificielle, France, 2007

[Blaise J.Y., 2003] Blaise J.Y. Le formalisme objet appliqué à l'étude de l'ediice patrimoniale: Problèmes de modélisation et d'échanges de données sur le réseau Internet. Thèse de doctorat de l'Université d'Aix-Marseille III, Spécialité Automatique et Informatique, 2003

[Boehler W. et al., 2002a] Boehler W., Marbs A. 3D Scanning Instruments. Proceedings of the CIPA WG 6 International Workshop on Scanning for Cultural Heritage Recording. Corfou, 2002

[Boehler W. et al., 2002b] Boehler W., Heinz G., Marbs A., Siebold M. 3D Scanning Softawre : an introduction. Proceedings of hte CIPAWG 6 International CIPA WG Workshop on Scanning for Cultural Heritage Recording. Corfu, 2002

[Bohler W. et al., 2003] Boehler W., Bordas V., Marbs A. Investigating Laser Scanner Accuracy. Proceedings of The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, Part 5/C15, pp. 696-701, 2003

[Borghese A. et al. 2000] Borghese A., Ferrari S. A Portable Modular System for Automatic Acquisition of 3-D Objects. IEEE Trans. on Instrumentation and Measurement, Vol. 49, No.5, 200. pp. 1128- 1136, 2000

[Carbonnel M., 1968] [Carbonnell, 1968] Carbonnell M. L'histoire et la situation presente des applications de la photogrammetrie a l'architecture. Etude sur la photogrammetrie appliquee aux monuments historiques. Application of photogrammetry to historic monuments. ICOMOS, Saint-Mande, France, 4-6, VII, 1968

[Cipolla R. et al. 1999] Cipolla R., Robertson D., Boyer E. Photobuilder – 3D models of Architectural scenes from uncalibrated Images. Proceedings of the Conference on Multimedia Computing and Systems, June 1999

[Cipolla R. et al. 1999] Cipolla R., Robertson D., Boyer E. Photobuilder – 3D models of Architectural scenes from uncalibrated Images. Proceedings of the Conference on Multimedia Computing and Systems, June 1999

[Clark R.N., 2002] Clark R.N., Dynamic Range and Transfer Functions of Digital Images and Comparison to Film. ClarkVision.com online article, 3 July 2005

[Cosmas J. et al., 2001] Cosmas J., Itagaki T., Green D., Grabczewski E., Weimer F., Van Gool L.J., Zalesny A., Vanrintel D., Leberl F., Grabner M., Schindler K., Karner K.F.,Gervautz M., Hynst S., Waelkens M., Pollefeys M., DeGeest R., Sablatnig R.,Kampel M.: 3d murale: A multimedia system for archaeology. In: Virtual Real.,Archaeol. and Cult. Herit. 297-306, 2001

[Cross J. et al.,] Cross J., Baber C., Woolley S.I., Layered Annotations of Digital Images for Data Collection in the Field, iswc, pp.154, Seventh IEEE International Symposium on Wearable Computers, ISWC'03, 2003

[Damisch H., 1994] Damisch, H., The Origin of Perspective, Translated by John Goodman. Cambridge, Mass.: MIT Press, 1994

[d'Angelo P., 2007] d'Angelo P., Radiometric alignment and vignetting calibration, 2007

[De Luca L. et al., 2010] De Luca L., Busarayat C., Stefani C., Renaudin N., Véron P., Florenzano M., An Iconography-Based modeling Approach for the Spatio-Temporal Analysis of Architectural Heritage. Proceedings of IEEE International Conference on Shape Modeling International. Aix-en-Provence, 21-23 june 2010

[De Luca L., 2006] De Luca L., : Relevé et multi-représentations du patrimoine architectural Définition d'une approche hybride de reconstruction d'édifices Thèse de l'ecole doctorale Arts et Metiers ParisTech, Aix-en-Provence , Mars 2006

[Dornaika F. et al., 2009] Dornaika, F.; Hammoudi, K, Extracting 3D Polyhedral Building Models from Aerial Images using a Featureless and Direct Approach. Proc. IAPR/MVA, 2009

[Edelsbrunner H., 2001] Edelsbrunner H. Geometry and Topology for Mesh Generation, volume 6. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2001

[EmGard K.L. et al., 2008] Emgard K.L., Zlatanova S.,. Design of an integrated 3D information model. In Urban and regional data management: UDMS annual 2007, Coors, Rumor, Fendel & Zlatanova (Eds.), Taylor & Francis Group, London, UK, pp. 143-156, 2008

[Falcidieno B. et al., 1998] Falcidieno B., Spagnuolo M., A Shape-abstraction paradigm for modelling geometry and semantics. Proceedinf of the international Conference on Computer Graphics. IEEE Computer Society Press, Los Alamitos, CA, 1998

[Faugeras O., 1993] Faugeras O. et al., Camera selfcalibration: theory andexp eriments, G. Sandini (ed.), Proc 2nd ECCV, Vol. 588 of Lecture Notes in Computer Science, SpringerVerlag, Santa Margherita Ligure, Italy, pp. 321–334, 1992

[Faugeras O., 1993] Faugeras O., Three-Dimensional Computer Vision : A Geometric Viewpoint. MIT Press, Cambridge, Massachusetts, 1993

[Feiner S. et al. 1997] Feiner S., Macintyre B., Hollerer T., Webster A., A touring machine: Prototyping 3d mobile augmented reality systems for exploring the urban environment. In Proc. IEEE Int. Sym. on Wearable Computers, 74-81, 1997

[Fergus R. et al., 2003] Fergus R., Fei-Fei L., Perona P.,Zisserman A. Learning Object Categories from Google's Image Search. In Proceedings of the Tenth IEEE international Conference on Computer Vision - Volume 2 (October 17 - 20, 2005). ICCV. IEEE Computer Society, Washington, DC, 1816-1823, 2005

[Fitzgibbon A.W. et al., 1998] Fitzgibbon A.W., Zisserman A. Automatic 3D Model Acquisition and Generation of New Images from Video Sequences. Proceedings of European Signal Processing conference (EUSPICO '98). Rhodes, Greece. pp. 1261-1269, 1998

[Flickner M. et al., 1995] Flickner M., Sawhney H., Niblack W., Ashley J., Qian Huang, Dom B., Gorkani M., Hafner J., Lee D., Petkovic D., Steele D., Yanker P., Query by Image and Video Content. the QBIC system. IEEE Computer, 28(9), 23-32, 1995

[Flickr], http://www.flickr.com

[Floater M. et al. 1998] Floater M., Iske A. Thinning Algorithms for Scattered Data Interpolation. BIT Numerical Mathematics, Vol. 38, No 4, 1998. pp. 705-720, 1998

[Foley J., 2000] Foley J. et al., Getting There : The Ten Top Problems Left, IEEE Computer Graphics and Application, Vol. 20, No 1, pp. 66-68, 2000

[Gaiani M., 1999] Gaiani M. : Translating the architecture of the world into virtual reality and vice-versa: 7 years of experimentation with conservation and representation at OFF. Proceedings of Heritage Applications of 3D Digital Imaging. Ottawa, Canada, 1999

[Galer M. et al., 2005] Galer M., Horvat L, Digital Imaging. Elsevier. ISBN 024051971X, 2005

[Gary R. B. et al., 2002] Gary R. B. et al. Technical Graphics Communication. McGraw-Hill Professional, 2002. ISBN 0073655988, p.330, 2002

[Gerald F et al., 1998] Gerald F, Brommer, Exploring Drawing. Worcester, Massachusetts: Davis Publications, 1988

[Gernsheim H., 1986] Gernsheim H., A concise history of photography,.Courier Dover Publications. pp.3-4. ISBN 0486251284, 1986

[Godse A.P.,1980] Godse, A.P., Computer graphics. Technical Publications. pp. 29. ISBN 8184315589,1980

[Gong Y. et al., 1994] Gong Y., Zhang H., Chuan H.C., Sakauchi M., Animage database system with content capturing and fast image indexing abilities. In: Proceedings of IEEE Int. Conf. on Multimedia Computing and Systems, 1994

[Goulette, 1999] Sémantique formelle de l'espace : Application au raisonnement spatial qualitatif en architecture. Intellectica : Revue de l'Association pour la Recherche Cognitive. N°29, pp. 9-34, 1999

[Gruber T.R., 1993] Gruber T.R.: Toward Principles for the Design of Ontologies Used for Knowledge Sharing. In Formal Ontology in Conceptual Analysis and Knowledge Representation, edited by Nicola Guarino and Roberto Poli, Kluwer Academic Publishers, in press. Substantial revision of paper presented at the International Workshop on Formal Ontology, March 1993

[Grussenmeyer P. et al., 2001] Grussenmeyer P., Hanke K., Streilein A. Photogrammetrie architecturale. Chapitre du livre, Photogrammétrie numérique. Éditions Lavoisier-Hermès, 2001

[Grussenmeyer, P. et al., 1999] Grussenmeyer P., Koehl M., Nour el Din M.:3D Geometric and semantic modelling in historic sites,XVII. CIPA International Symposium, October 3-6, 1999

[Guarino N., 1998] Guarino N.M. : Formal Ontology and Information Systems. In N. Guarino (ed.), Formal Ontology in Information Systems. Proc. of the 1st International Conference, Trento, Italy, 6-8 June 1998

[Guyonnet F., 2000] Guyonnet F., Aude Cite de Carcassonne Château Comtal, Document Final de Synthèse, intervention archéologique, mars-avril 2000

[Halaschek-Wiener C. et al., 2005] Halaschek-Wiener C., Jennifer G., Andrew S., Michael G., Bijan P., Jim H.: PhotoStuff -- An Image Annotation Tool for the Semantic Web, 2005

[Hartley R., 1997] Hartley R. In defense of the eight-point algorithm. PAMI. Vol. 19/6, pp. 580-593, 1997

[Heikkila J. et al., 1997] Heikkila J., Silven O. A Four-Step Camera Calibration Procedure with Implicit Image Correction. Proceedings of IEEE Computer Vision and Pattern Recognition, . pp. 1106-1112, 1997

[Herbig U. et al., 1997.] Herbig, U., Waldhäusl, P., APIS – Architectural photogrammetry information system. IAPRS, 32(5C1B), pp. 23-27., 1997, accessed December 2004

[Hirsch R., 2000] Seizing the Light: A History of Photography, McGraw-Hill, 2000

[Hu Z. et al. 2006] Hu Z., Uchimura K.:Fusion of Vision,GPS and 3D Gyro Data in Solving Camera Registration Problem for Direct Visual Navigation",International Journal of ITS,4,1,pp.3-12 , December 2006

[Hyman I.C., 1974]. Hyman I.C., Brunelleschi in Perspective. Englewood Cliffs, New Jersey: Prentice-Hall, 1974

[Johansson B. et al., 2002a] Johansson B., Kahl F.M.:Detecting Windows in City Scenes. In Proc. of International Workshop on Pattern Recognition with Support Vector Machines, ICPR, 2002

[Johansson B. et al., 2002b] Johansson B., Cipolla R.. A System for Automatic Pose-Estimation from a Single Image in a City Scene. In Proc. of International Conference on Signal Processing, Pattern Recognition, and Applications, Crete, Greece, 2002

[Kadobayashi R. et al.,2005] Kadobayashi R., Tanaka K.:3D Viewpoint-based Photo Search and InformationBrowsing, Annual ACM Conference on Research and Development in Information Retrieval, Salvador, Brazil, in poster session, Pages: 621 - 622 , 2005

[Kang H. et al., 2000] Kang H., Shneiderman B., Visualization Methods for Personal Photo Collections: Browsing and Searching in the PhotoFinder, Proc. IEEE International Conference on Multimedia and Expo, New York: IEEE, pp. 1539-1542, 2000

[Karathanassi V. et al., 2000] Karathanassi V., Iossifidis C., Rokos D., A texture-based classification method for classifying built areas according to their density, International Journal of Remote Sensing, vol. 21, no. 9, pp. 1807–1823, September 2000

[Kemp M., 1992] Kemp M., The Science of Art: Optical Themes in Western Art from Brunelleschi to Seurat. Yale University Press, 1992

[Kleinbauer W.E. et al., 1982] Kleinbauer W.E. , Thomas P. Slavens, Research Guide to the History of Western Art, Sources of information in the humanities, no. 2. Chicago: American Library Association: 60-72, 1982

[Kodabayashi R. et al, 2005] Kadobayashi R., Tanaka K., 3D viewpoint-based photo search and information browsing. In Proceedings of the 28th Annual international ACM SIGIR Conference on Research and Development in information Retrieval (Salvador, Brazil, August 15 - 19, 2005). SIGIR '05. ACM, New York, NY, 621-622, 2005

[Kraus K. et al., 1997] Kraus K., Waldhaul P. Manuel de photogrammétrie : principes et procédés fondamentaux. Hermes, 1997

[Kuchinsky A. et al., 2002] Kuchinsky A., Pering C., Creech M.L., Freeze D., Serra B., Gwizdka J.: FotoFile: A Consumer Multimedia Organization and Retrieval System, Proc. ACM Conference on Human Factors in Computing Systems (CHI '99), pp. 496-503, 1999

[Kumar R. et al., 1994] Kumar R., Hanson A. Robust Methods for Estimating Pose and a Sensitivity Analysis. CVGIP-Image Understanding. Vol. 60, No. 3, 1994. pp 313-342, 1994

[Kumar S. et al., 2004] Kumar, S., Herbert, M.: Discriminative fields for modeling spatial dependencies in natural images. In: NIPS, 2004

[Lee S.C. et al., 2002] Lee S.C., Jung S.K., Nevatia R., "Automatic Pose Estimation of Complex 3D Building Models", [PDF], IEEE Workshop on Application of Computer Vision, 148-152, 2002

[Lieberman H., 2000] Lieberman H., An agent for integrated annotation and retrieval of images. Paper presented at Workshop on Personal Photo Libraries: Innovative Designs. University of Maryland, 1 June 2000

[Littlefield R., 2006] Littlefield R., Theory of the "No-Parallax" Point in Panorama Photography, 2006

[Lowe D.G., 1999] Lowe D.G., Object recognition from local scale-invariant features [archive], Proceedings of the International Conference on Computer Vision 2: 1150–1157, 1999

[MacEachren A.M. et al., 2005] MacEachren, A.M. et al., Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. Cartography and Geographic Information Science, 32(3), 139-160, 2005

[Manferdini A.M. et al., 2008] Manferdini A.M., Remondino F., Baldissini S., Gaiani M., Benedetti B.: 3D modeling and semantic classification of archaeological finds for management and visualization in 3D archaeological databases. Proc. of 14th Int. Conference on Virtual Systems and MultiMedia (VSMM), pp. 221-228, Limassol, Cyprus, 2008

[Marbs A. et al., 2001] Marbs A., Heinz G., Boehler W. The potential of non-contact close range laser scanners for cultural heritage recording. Proceedings of CIPA International Symposium. Potsdam, Germany, 2001

[Mayer H., 1999] Mayer H., Automatic object extraction from aerial imagery — a survey focusing on buildings, Computer Vision and Image Understanding, vol. 74, no. 2, pp. 138–149, May 1999

[Maynard P., 2005] Maynard, P., Drawing distinctions: the varieties of graphic expression. Cornell University Press. pp. 22. ISBN 0801472806, 2005

[McReynolds T. et al., 2005] McReynolds T., David B., Advanced graphics programming using openGL. Elsevier. pp. 502. ISBN 1558606599, 2005

[Mencl R. et al., 1998] Mencl R., Mueller H. Interpolation and Approximations of surfaces from Three- DimensionalScattered Data Points. State of the Art. Proceedings of Eurographics Conference. Lisbon, Portugal, 1998

[Mueller P. et al., 2006] Mueller, P., Wonka, P., Haegler, S., Ulmer, A., Van Gool, L., Procedural modeling of buildings. Proc. of ACM SIGGRAPH 2006 / ACM Transactions on Graphics, 25(3), pp. 614-623, 2006

[Nubes], http:// www.map.archi.fr/nubes

[Omega] Déclaration d'invention présentée au nom du CNRS et faisant l'objet d'une demande de brevet déposée par le Cabinet Plasseraud le 16 juin 2009 sous le numéro FR 09 54 054 : OMEGA – ORIENTATION MODULE (inventeurs : Livio De Luca, Matthieu Deveau), 2009

[Ono A. et al. 1996] Ono A., Amano M., Hakaridani M., Satou T., Sakauchi M., A flexible content-based image retrieval system with combined scene description keyword. In: Proceedings of IEEE Int. Conf. on Multimedia Computing and Systems, pp. 201-208, 1996

[Palladio A., 1992] Palladio, A., I Quattro Libri dell'Architettura Edizioni Studio Tesi., Pordenone, 1992

[Pang A.T. et al., 1997] Pang, A.T., Wittenbrink, C.M. & Lodh, S.K., Approaches to Uncertainty Visualization. The Visual Computer, 13(8), 370-390, 1997

[Perrault C., 1684] Perrault, C. Les dix livres d'architecture de Vitruve, Pierre Mardaga, 1684

[Petridis K. et al., 2006] Petridis K, Anastasopoulos D, Saathoff C, Timmermann N, Kompatsiaris I, Staab S. M-OntoMat-Annotizer: Image annotation linking ontologies and multimedia low-level features. In: KES 2006 10th Intnl. conf. on knowledge based, intelligent information and engineering systems, 2006

[Pierrot-Desseilligny M., 2007] Pierrot-Deseilligny M., "MicMac, un lociel de mise en correspondance d'images, adapte au contexte geograhique" to appears in Bulletin d'information de l'Institut Geographique National, 2007

[Pollefeys M. et al., 2000] Pollefeys M., Van Gool L., Vergauwen M., Verbiest F., Cornelis K., Tops J., Koch R.: Visual modeling with a hand-held camera. Int. J. of Computer Vis.59(3) 207-232, 2004

[Quintrand P. et al., 1985] Quintrand P., Autran J., Florenzano M., Fregier M., Zoller J., La CAO en architecture. Hermes, Paris, 1985

[Rand G. et al., 2005] Rand G., Litschel D., Davis R., Digital Photographic Capture. Elsevier. ISBN 0240806328, 2005

[Remondino F., 2003] Remondino F. From point cloud to surface : the modeling and visualizaion problem. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, part 5/W10. ISPRS International Workshop on "Visualization and Animation of Reality-based 3D Models". Tarasp-Vulpera, Switzerland, 2003

[Riya], http://www.riya.com

[Rodden K.,1999] Rodden, K., How do people organise their photographs? In BCS IRSG 21st Ann. Colloq. on Info. Retrieval Research, 1999

[Rosen M.J. et al., 2002] Rosen M.J. and David L. Devries Photography & Digital Imaging. Kendall Hunt. ISBN 0757511597, 2002

[Saathoff C., 2006] Saathoff C.: Constraint reasoning for region-based image labelling. In: 3rd IEE international conference of visual information engineering. Special session on semantic multimedia, 2006

[Salemi G. et al., 2005] Salemi G., Achilli V., Cervato C. Data acquisition for cultural heritage navigation: integration of panoramic imaging, terrestrial laser scanning and anaglyphs. Proceedings of the ISPRS working group V/5. Panoramic Photogrammetry Workshop. Berlin, Germany, 2005

[Shamir A., 2006] Shamir A. Segmentation and shape extraction of 3D boundary meshes. In: Eurographics 2006 - State of the art reports, p. 137-49, 2006

[Shen H.T. et al. 2000] Shen H.T., Ooi B.C., Tan K.L., Giving Meanings to WWW Images. In: Proc. ACM MM2000, pp 39-48, 2000

[Shneiderman B. et al., 2000] Shneiderman B., Kang H., Direct Annotation: A Drag-and-Drop Strategy for Labeling Photos. In: Proc. International Conference Information Visualisation, IV2000, 2000

[Shotton J. et al., 2006] Shotton, J.D.J., Winn, J., Rother, C., Criminisi, A.: TextonBoost for Image Understanding: Multi-Class Object Recognition and Segmentation by Jointly Modeling Texture, Layout, and Context, IJCV(81), No. 1, January 2009

[Shum HY.. et al., 2000] Shum H.Y., Kang S.B.. A Review of Image-based Rendering Techniques. IEEE/SPIE Visual Communications and Image Processing (VCIP). Perth, pp. 2-13, 2000

[Sivic J. et al., 2003] Sivic J., ZissermanI A., Video Google: A text retrieval approach to object matching in videos. In Proc. Int. Conf. on Computer Vision, 1470-1477, 2003

[Snavely N. et al., 2006] Snavely N., Seitz S.M., Szeliski R., Photo Tourism: Exploring Photo Collections in 3D. SIGGRAPH Conf. Proc, 2006

[Snavely N. et al., 2007] Snavely N., Seitz S.M., Szeliski R., Modeling the World from Internet Photo Collections. International Journal of Computer Vision, 2007

[Srihari R.K. et al., 2000] Srihari R.K., Zhang Z, Rao, A Intelligent indexing and semantic retrieval of multimodal documents. Information Retrieval, 2, 245-275, 2000

[Stiny G. et al., 1978] Stiny G., Mitchell W.J., The palladian grammar. Environment and Planning B: Planning and Design, 5, pp. 5-18, 1978

[Stiny G., 1980] Stiny G., Introduction to shape and shape grammars. Environment and Planning B: Planning and Design, 7, pp. 343- 351, 1980

[Strothotte T. et al., 1999] Strothotte, T., Masuch, M., Isenberg, T., Visualizing Knowledge about Virtual Reconstructions of Ancient Architecture. Dans Proceedings of the International Conference on Computer Graphics. IEEE Computer Society, p. 36, 1999

[Suh B. et al., 2004] Suh B., Bederson, B.B., Semi-automatic photo annotation strategies using event based clustering and clothing based person recognition. *Interact. Comput.* 19, 4 (Jul. 2007), 524-544, 2007

[Szeliski R., 2005] Szeliski R., Image Alignment and Stitching, Retrieved 2008-06-01, 2005

[Thomson J. et al., 2005] Thomson, J. et al., A typology for visualizing uncertainty. Dans Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. pp. 146-157, 2005

[Tsai R.Y., 1986] Tsai R.Y., A versatile camera calibration technique for high accuracy 3D Machine vision metrology using off-the-shelf TV cameras and lenses ,IEEE Journal of Robotics and Automation, Vol. RA3, No. 4, August 1987

[Tzonis A. et al., 1986] Tzonis A., Lefaivre L., Classical Architecture - The Poetics of Order. The MIT Press, Cambridge, 1986

[Unsalan C. et al., 2005] C. Unsalan and K. L. Boyer, A theoretical and experimental investigation of graph theoretical measures for land development in satellite imagery, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 4, pp. 575–589, April 2005

[Vu H., et al. 2009] Vu H., Keriven R., Labatut P., Pons J.P., Towards high-resolution large-scale multi-view stereo. Conference on Computer Vision and Pattern Recognition (CVPR), Miami, June 2009

[Waldhäusl P. et al., 1994] Waldhäusl P., Ogleby C. 3 x 3 Rules for simple photogrammetric documentation of architecture. In: J.G.Fryer (Editor), International Archives of Photogrammetry and Remote Sensing, Vol.XXX, Part5, pp 426 – 429, 1994

[Ward G., 2006] Ward G., Hiding seams in high dynamic range panoramas. Proceedings of the 3rd symposium on Applied perception in graphics and visualization. 153. ACM International Conference Proceeding Series. ACM. doi:10.1145/1140491.1140527. ISBN 1-59593-429-4, 2006

[Welling W., 1978]. Photography in America, page 23, 1978

[Wenyin L. et al., 2001] Wenyin L., Dumais S., Sun Y., Zhang H., Czerwinski M., Field B., Semi Automatic Image Annotation., Proc. Interact '01, pp.326-333, 2001

[Winn J. et al., 2005] Winn, J., Criminisi, A., Minka, T., Categorization by learned universal visual dictionary. Int. Conf. of Computer Vision, 2005

Table des illustrations

Figure 1 L'état actuel du palais du Louvre contient des bâtiments du 17e siècle 16
Figure 2 Documentation photogrammétrique du système APIS [Herbig U. et al.,1998.] 17
Figure 3 Andrea Palladio, Thermes d' Agrippa et Panthéon, RIBA, Londres, Vol. VII, fol. 2. 18
Figure 4 L'état hypothétique du château Comtal de Carcassonne au milieu du XIV^{e} :..... 19
Figure 5 Une vieille photo sur une carte postale présente une partie du château Comtal du Carcassonne. 21
Figure 6 Métadonnées de format EXIF d'une photo numérique 22
Figure 7 Dessin de vue générale sur le château comtal depuis le sud réalisée par Viollet-le-Duc. 23
Figure 8 Coupe de la porte d'entrée du château Comtal de Carcassonne 24
Figure 9 Illustration du dispositif de la perspective de Brunelleschi. 25
Figure 10 Un exemple de dessin en perspective du château Comtal de Carcassonne 26
Figure 11 Un exemple de photographie stéréo (porte de Lanhydrock House, Cornwall) 27
Figure 12 Un séquence d'images extraite d'une séquence vidéo 28
Figure 13 Un exemple d'application de la technique de Camera tracking 28
Figure 14 Tableau résumant les différents types de projection parallèle 29
Figure 15 Élévation de l'aile ouest du château Comtal de Carcassonne. Mélange entre une orthophotographie et un dessin au trait. 30
Figure 16 Un exemple de l'image isométrique de la fouille archéologique 30
Figure 17 Des exemples d'images à projection oblique 31
Figure 18 Des exemples d'images panoramiques 33
Figure 19 L'interface de PhotoFinder 36
Figure 20 L'interface de M-OntoMat-Annotizer 37
Figure 21 Annotation sémantique d'images d'architecture patrimoniale 37
Figure 22 Les résultats d'une détection sémantique 39
Figure 23 Les résultats de la détection sémantique de TextonBoost [Shotton J. et al., 2006] 40
Figure 24 Un exemple de détection sémantique à partir d'images aériennes [Akcay H.G. et al., 2007] ... 40
Figure 25 Les résultats de la détection sémantique d'images aériennes à l'échelle urbaine utilisant des patterns de bâtiment [Aksoy S., 2006] 41
Figure 26 Le processus de la détection sémantique du ARIA 42
Figure 27 L'interface de MiAlbum pour la recherche basée sur texte, pertinence de résultat 43
Figure 28 Différence entre les approches actuelles de sémantisation d'image (en haut) et l'approche proposée (bas) 46
Figure 29 Les trois premières étapes de l'approches proposée : un modèle 3D est réalisé à partir d'une acquisition de l'édifice réel ; le modèle 3D est caractérisé du point de vue sémantique ; les images sont référencées spatialement par rapport au modèle 3D. 47
Figure 30 La quatrième étape de l'approche proposée : projection (du modèle 3D vers les images 2D) d'une couche sémantique. 48
Figure 31 Un exemple de description sémantique multicouche : à chaque décomposition du modèle 3D corresponde une projection de couche sémantique différente sur les images. 49
Figure 32 Schéma représentant les trois techniques de restitution photogrammétrique. De gauche à droite : la restitution à partir d'une seule image, d'un couple d'images et d'un bloc multi-image. 57
Figure 33 Un scanner laser « temps de vol » et résultat produit (nuage de point.) 58
Figure 34 Illustration schématique des deux principes de fonctionnement des dispositifs de balayage laser 3D. À gauche le principe du temps de vol, à droite celui de la triangulation. [Marbs A. et al., 2001]. 58
Figure 35 Consolidation d'un nuage de points issu d'un balayage laser 3D. Dans [De Luca L., 2006]. 60

Figure 36 Un scanner laser 3D à triangulation optique et le résultat obtenu (maillage polygonal d'un élément sculpté)..... 61
Figure 37 Modèle 3D réalisé par des procédures de modélisation interactive. Château Comtal de Carcassonne..... 63
Figure 38 Exemples de modélisation à partir d'images. 64
Figure 39 Graphe de description sémantique associé au modèle 3D [De Luca L., 2006]..... 66
Figure 40 Classification sémantique d'objets archéologiques. Dans [Manferdini A.M. et al.,2008]..... 66
Figure 41 L'interface d'un système pour la modélisation géométrique et sémantique d'objets urbains [Grussenmeyer, P. et al., 1999]..... 67
Figure 42 Des multiples représentations du modèle 3D élaborées à partir du même modèle de description sémantique. 68
Figure 43 Des exemples de structuration sémantique du modèle élaborés en organisant les éléments architecturaux en fonctions des concepts de « entité morphologique », « groupe finalisé » et « repère ». 69
Figure 44 Termes de descriptions associés à la morphologie de l'édifice organisés selon les relations hiérarchiques établies dans le graphe de description..... 70
Figure 45 L'interface du système de [Kadobayashi R. et al.,2003], capable à référencer manuellement les information 2D et 3D 72
Figure 46 Planification des photographies du règle 3x3..... 73
Figure 47 Calibration et orientation des caméras..... 74
Figure 48 Le hall du château Neuschwanstein : l'image de luminance du scanner est mélangée avec les valeurs RVB correspondantes acquises par la caméra. [Abmayr T. et al., 2004]..... 75
Figure 49 Acquisition laser panoramique à Venise. À gauche, images de l'acquisition laser. À droite, les images filtrées correspondantes. De haut en bas : l'image en vraies couleurs, l'image d'intensité en nuances de gris, la carte de profondeur. [Salemi et al., 2005]..... 76
Figure 50 Système de OMEGA, une méthode de référencement automatique d'image. 77
Figure 51 L'interface de Photosynth 77
Figure 52 À gauche, images de différents points de vue avec le modèle adapté à un bâtiment. À droite, Le résultat de l'estimation de la position 78
Figure 53 L'interface pour le téléchargement d'images sur serveur..... 81
Figure 54 L'interface pour le référencement manuel..... 82
Figure 55 L'interface pour le référencement manuel. Variation de l'angle de champ de la caméra..... 82
Figure 56 L'interface pour le référencement manuel. Rotation du plan image..... 83
Figure 57 L'interface pour le référencement manuel. Projection orthographique. 83
Figure 58 Le modèle géométrique d'une caméra à projection perspective. 84
Figure 59 L'interface pour la résection spatiale d'images. En haut : le processus de sélection des points d'entrée sur le modèle 3D et sur l'image, en bas : le résultat du référencement. 85
Figure 60 Schéma de la sémantisation d'images..... 88
Figure 61 Reconstruction des points de vues de images dans une scène 3D. 90
Figure 62 Projection des contours des éléments 3D sur l'image. A gauche, l'image originale, à droite l'image vectorielle SVG. 91
Figure 63 Image semantisée SVG associant des comportements interactifs au passage de la souris..... 92
Figure 64 Dialogue entre les parties du système pour les processus de téléchargement, de référencement spatial et de projection des couches sémantiques..... 93
Figure 65 Dialogue entre les parties du système 96
Figure 66 Données relatives à un point de vue relatif à une source iconographique 98
Figure 67 Diagramme de l'utilisation des fonctions d'interface 100
Figure 68 Méthode de détection d'images basée sur l'intersection de pyramides visuelles..... 101
Figure 69 Paramètres du critère spatial. En haut : distance, en bas : différence d'orientation..... 102

Figure 70 Exemples d'échec de détection : cas où la valeur de la limite introduite par l'utilisateur est trop grande ou trop petite. *103*
Figure 71 Combinaison de requêtes *104*
Figure 72 Sélection de multiples objets dans la scène 3D *105*
Figure 73 Résultat de la requête basée sur la sélection de multiples objets. *106*
Figure 74 Résultat d'une requête basée sur critère sémantique. *106*
Figure 75 Récupération du point de vue de la source iconographique sélectionnée *107*
Figure 76 Incohérence entre le modèle 3D et l'image *110*
Figure 77 Exemples des représentations échappe à une cohérence formelle avec la réalité *111*
Figure 78 Exemples des images de point de vue subjective *111*

Liste de publications

Publications dans des revues avec comité de lecture

De Luca L., Busayarat C., Stefani C., Véron P., Florenzano M., "A semantic-based Platform for the digital analysis of the architectural heritage", Computers & Graphics. (en cours de publication). Elsevier, 2010

Publications dans des actes de congrès avec comité de lecture

De Luca L., Busayarat C., Stefani C., Renaudin N., Véron P., Florenzano M., "An Iconography-Based modeling Approach for the Spatio-Temporal Analysis of Architectural Heritage", Proceedings of IEEE International Conference on Shape Modeling International, June 21-23, 2010, Arts et Métiers ParisTech, Aix-en-Provence, France.

De Luca L., Busayarat C., Stefani C., Véron P., Florenzano M., "NUBES : describing, analysing, documenting and sharing digital representations of heritage buildings", Proceedings of Focus K3D conference on Semantic 3D Media and Content 11th and 12th February, 2010 INRIA Sophia Antipolis - Méditerranée, France.

Busayarat C., De Luca L., Véron P., Florenzano M., "Semantic annotation of heritage building photos based on 3D spatial referencing Proceedings of Focus K3D conference on Semantic 3D Media and Content 11th and 12th February", 2010 INRIA Sophia Antipolis - Méditerranée, France.

Busayarat C., De Luca L., Véron P., Florenzano M., "An on-line system to upload and retrieve architectural documents based on spatial referencing" , IDMME - Virtual Concept 2008 Conference, Beijing, China, oct 2008

Busayarat C., De Luca L., Véron P., Florenzano M., "A real-time 3D interactive interface for a spatial referencing system" , Proceedings of the International Conference on Computer Graphics and Visualization, Amsterdam, Netherlands, jui 2008

De Luca L., Busayarat C., Stefani C., Véron P., Florenzano M., "NUBES : une plateforme intégrée pour la description, l'analyse, la documentation et le partage de représentations numériques d'édifices architecturaux", dans David Peyceré , Florence Wierre , Ed. "Architecture et archives numériques. L'architecture à l'ère du numérique : un enjeu de mémoire" - «Architecture and digital archives. Architecture in the digital age: a question of memory», Edition Infolio 2008 - ISBN : 978-2-88474-146-00.

De Luca L., Busayarat C., Stefani C., Véron P., Florenzano M., "An integrated framework to describe, analyze, document and share digital representations of architectural buildings", in: of VAST 2007: futures technologies to empower the heritage professionals, Brighton, UK, 26-30 november 2007.

Busayarat C, "Conception d'un système de référencement spatial de photographie de relevé", in : MIA journal, Vol. 02 (72DPI), 2006 www.map.archi.fr/mia/journal/