



Working Without Words – Image based Retrieval as a Model to overcome Language Hurdles

Jan Simane

Head Librarian
Kunsthistorisches Institut in Florenz
Max-Planck-Institut
Florence, Italy

Traduction:

*Ernst Véronique
Université de Neuchâtel,
Faculté des lettres et sciences humaines, Suisse
(veronique.ernst@gmail.com)*

Meeting:

79 — Aborder les défis du plurilinguisme dans les arts: catalogues, bases de données, collections numérisées et autres matériaux dans le contexte global — Section Bibliothèques d'Art

Résumé :

Dans l'histoire de l'art les images sont le principal véhicule de l'intérêt. Dans un contexte scientifique, les objets multimédia, y compris les images digitales, sont représentées sous la forme d'un code sémantique généré intellectuellement ou automatiquement. Par conséquent, les grandes options pour rechercher et extraire l'information dans les réseaux numériques sont essentiellement basées sur le texte. En utilisant une base de données d'images des armoiries florentines, une caractéristique expérimentale a été développée permettant la recherche et l'extraction non verbale des images numérisées. Cette caractéristique expérimentale permettant la recherche est basée sur la comparaison d'éléments structurels de portions de l'image. Ceci est un exemple de la manière dont les problèmes de plurilinguisme peuvent être résolus, bien que dans un cas vraiment très particulier.

Étudier l'art et examiner des œuvres d'art est une discipline étrange. Les arts visuels sont faits pour les yeux ; les formes, la matière, les styles sont des éléments complémentaires des expressions figuratives. Et les artistes adressent leurs messages tout d'abord et surtout au sens visuel du spectateur. Cependant toute notre information et notre connaissance, nos descriptions et nos appréciations, tous les types de systématisation, de classification et chaque approche analytique de l'art et des expressions artistiques sont basés sur les mots. Ces deux médias très différents, la nature non verbale de l'objet d'une part et la nomination verbale de l'autre représente le dualisme spécifique de notre système d'encodage des phénomènes visuels pour des buts scientifiques. En conséquence de ce fait peu surprenant, nous sommes habitués à traiter des objets visuels à l'aide de représentations verbales comme étant la seule manière de nommer, de différencier et de contextualiser ce que nous voyons. Ce fait est des plus pertinents dans le contexte de la disponibilité

d'une quantité énorme d'images sur internet ou dans des bases de données spécifiques, surtout lorsque des outils de recherche de l'information efficaces deviennent une question cruciale. Ce ne sont pas seulement des millions de nouvelles photographies qui sont mises en ligne quotidiennement, mais également des collections entières d'images analogues qui sont scannées et rétablies dans des bases de données complètes. L'accès à cet important matériel est maintenant beaucoup plus simple et plus rapide et nous utilisons aussi les diverses options de liens hypertextuels et de mise en réseaux afin de rassembler les collections localisées dans des endroits différents et afin de faire de nouvelles trouvailles en accédant à davantage de résultats différents grâce aux outils de méta recherche simultanées. Cependant, dans tous les cas, les outils de recherche sont basés sur du texte et par conséquent sur des métadonnées limitées. Ces limites sont un des résultats des méthodes hétérogènes de description ou, à l'âge de la communication globale, sont dues à la simple pluralité des langues. L'incongruence susmentionnée entre la structure visuelle de l'image et sa représentation verbale, le fameux fossé sémantique, est, dans ces circonstances, impossible à éviter. Ce problème est inexistant quand on a affaire uniquement à du texte. Ici, l'encodage des caractères singuliers est la condition décisive pour trouver des éléments divers et identiques. À cet égard, la structure d'un terme, c'est-à-dire d'une séquence de caractères, et son code numérique forment une unité. Est-ce que ce principe pourrait être appliqué aux images ? La conséquence serait une recherche non verbale, basée sur la forme, une identité sémantique entre l'objet visuel utilisé pour la recherche et celui recherché. Cela semble être une idée prometteuse. Et un modèle intéressant pour s'attaquer au problème du plurilinguisme dans les procédures de recherche verbale.

Dans la suite, je vous présenterai un projet qui est en cours à la bibliothèque du Kunsthistorisches Institut de Florence et pour lequel la solution mentionnée ci-dessus d'un modèle de recherche non verbale basée sur l'image a été réalisé. Premièrement laissez-moi décrire brièvement ce qu'est aujourd'hui le Kunsthistorisches Institut. L'activité de l'Institut est dédiée à la recherche fondamentale dans le champ de l'histoire de l'art italien. Même si localisé à Florence (Italie), c'est un des plus importants instituts de recherche allemands, spécialisé dans l'art italien et fondé à la fin du 19^{ème} siècle. Ses principaux 'laboratoires' sont la bibliothèque avec une collection d'environ 330 000 volumes et les archives photographiques comprenant environ 600 000 images, tous les deux essentiellement centré sur l'art italien. La Bibliothèque du Kunsthistorisches Institut ne fournit pas seulement une littérature spécialisée mais également des sources primaires importantes pour l'étude de l'histoire de l'art italien. De plus, la bibliothèque est aussi active dans le domaine expérimental. Et c'est ici le sujet de cette présentation. La Bibliothèque possède une collection d'environ 2 800 dessins d'armoiries en couleurs des familles importantes ou nobles de Florence, des confréries et des hôpitaux, complétés d'une annexe, qui consiste en un index des noms. Les dessins, réalisés par un dessinateur anonyme du 19^{ème} siècle, ont été acquis en 1901 et sont devenus un instrument de référence important. Chaque armoirie est dessinée sur une feuille de papier de format A5 environ avec le nom de la famille écrit en-dessous. Durant plusieurs décennies, les spécialistes qui souhaitaient identifier une armoirie dépeinte sur une œuvre d'art ou en lien avec l'architecture consultaient cette collection appréciable. Comme nous le savons tous, les armoiries sont parfois dépositaires de preuve non seulement des datations corrects mais permettent également d'identifier le propriétaire ou le protecteur d'une œuvre d'art ou même d'un bâtiment. En tant qu'emblème sur une tombe ou sur des documents ou des chartes, l'armoire est aussi un témoin authentique et fiable, qui peut contenir des informations importantes concernant le statut et la position sociale de la personnalité ou/et de la famille. Bien que la description des armoiries ne requiert pas un vaste

vocabulaire diversifié, le nombre des différentes variantes et combinaisons des termes est plutôt élevé et suit un système complexe. Les armoiries ne sont pas des œuvres d'art, mais des insignes. Elles doivent être décodées précisément et sans équivoque. Pour cette raison, l'héraldique a développé une terminologie très concise, complètement différente de notre langage de tous les jours ainsi que du langage que nous utilisons pour la description des œuvres d'art. Pour les non-experts, la recherche d'armoiries dans les catalogues ou les index est souvent difficile, en particulier quand le propriétaire des armes est inconnu. Notre projet 'Stemmario' – le terme fait référence à l'italien *stemma* qui peut être traduit par "collection d'armes" – a pour objectif de faciliter les outils de recherche dans une nouvelle base de données et de l'ouvrir à un cercle d'utilisateurs internationaux, qui pourraient ne pas être familiarisés avec les termes techniques et leurs variantes spécifiques dans les différentes langues.

La première étape a été la conversion numérique des 2 800 feuilles. Elles ont été scannées au format de fichier TIFF et les images ont été restaurées dans une base de données. En second lieu nous avons décrit en langue allemande toutes les armoiries selon les standards du vocabulaire héraldique. En plus, nous avons également conçu une division par catégories, un outil qui permet une recherche systématique sur la base d'éléments singuliers représentés sur les armoiries, et ce même si le propriétaire est inconnu. Jusqu'ici, le projet fournit une sorte de base de données d'images conventionnelle, qui a affaire aux désavantages déjà mentionnés de l'indexation verbale et de la complexité terminologique. Enfin, nous avons débuté une phase plus expérimentale dans le but d'éliminer ou du moins de limiter le risque d'une liste non exhaustive de résultats. Pour la réalisation de cette tâche nous avons pu mettre à profit nos excellents contacts avec le réseau du système de l'information multimédia de l'Institut des Sciences et Technologies de l'Information du Conseil National de Recherche à Pise (Italie). Dès le début, nos collègues à Pise ont été intéressés par notre base de données 'Stemmario' en tant que terrain d'expérimentation et nous avons besoin de leur savoir-faire et des technologies développées à Pise pour générer des idées concrètes dans le but de résoudre les problèmes mentionnés précédemment. Le résultat est un simple outil de recherche d'une base de données d'images sans aucune description verbale. Ses principes fonctionnels seront brièvement expliqués ci-dessous : Dans le cas d'une recherche d'objets basée sur une image et non verbale, toutes les informations nécessaires doivent être extraites directement depuis les objets image. Cela se passe sous la forme d'une analyse sémantique de l'objet faite automatiquement et capable de déterminer les caractéristiques formelles de l'image de surface. Récemment, les logiciels utilisés pour isoler des éléments depuis les images et pour l'indexation des structures de l'image ont été améliorés et révisés. Il existe deux options différentes pour la récupération d'information. La première et la plus courante, qui a été développée durant les années 1990, est basée sur le paradigme d'une réconciliation générale avec un exemple, elle a été choisie comme point de départ. Le programme recherche des objets avec des structures totalement similaires. Les résultats, comme illustré par notre exemple qui représente une des étapes durant le processus de développement de notre site web 'Stemmario', ne sont pas satisfaisants et ne peuvent pas être utilisés pour une recherche générale en histoire de l'art ni pour une recherche en héraldique. Une autre possibilité est de rechercher des images similaires, comme cela a été récemment fait sur Google Images. Dans ce cas, les métadonnées verbales sont connectées à un outil de recherche de similarité optique. Les résultats révèlent une similarité au niveau du contenu, mais d'un point de vue formel, ils n'ont rien en commun. Les deux méthodes sont incapables de combler le fossé sémantique dans le cas d'une collection d'armoiries.

Le développement des applications, qui permettent de rechercher des éléments singuliers dans une image, marque un réel progrès, car dans ce cas l'utilisateur peut sélectionner lui-même les éléments de l'image qui sont les plus ou les moins importants dans sa recherche. En isolant les parties utiles des objets image par une analyse de la structure sémantique de l'image générée automatiquement cela permet d'affiner la recherche et conduit à des résultats nettement plus précis. À cause de cet avantage, nous avons décidé de choisir cette méthode pour la base de données des armoiries florentines. La pertinence des armoiries en tant qu'images possédant une dimension significative est basée sur le répertoire des symboles et leur arrangement sur le blason. Puisque le même élément se retrouve sur plusieurs variations, la description et la classification des armoiries consiste principalement à nommer les éléments singuliers. Par conséquent, l'identification des armes dépend de la classification correcte de tous les éléments dépeints sur le blason. Le point de départ de l'analyse sémantique est la segmentation, ou la décomposition de l'objet entier en entités homogènes (ou unités). Nos partenaires à Pise ont décidé d'utiliser une méthode de segmentation en deux étapes : la première consiste en la division de la surface selon sa morphologie chromatique. Les segments ne sont pas nécessairement divisés en zones compacts d'une seule et même couleur, mais elles sont générées automatiquement selon l'intensité d'une couleur, indépendamment de leur position sur l'écu. Ensuite de quoi, les champs isolés sont sujets à une seconde segmentation, basée sur le fameux K-Means-Algorithm, qui représente les objets isolés comme des agrégats ; dans notre cas cela signifie une représentation en tant que subdivisions fermées et compactes, manifestant non seulement l'intensité différente des diverses subdivisions, mais aussi leur position sur la surface. Les résultats de la segmentation sont montrés à l'utilisateur sous la forme d'objets image. Maintenant, l'utilisateur peut poursuivre sa stratégie de recherche en décidant de la pertinence des éléments singuliers. En cliquant sur un seul des segments il peut parcourir la base de données à la recherche d'autres éléments ayant une structure similaire.

Pour illustrer cette présentation par un exemple concret : le très populaire chef Capo d'Angiò est un label rouge (une ligne horizontale avec trois ou parfois même quatre pendants, généralement appelés points ou plus rarement gouttes). Ce symbole héraldique avait été introduit par Charles d'Anjou en 1246. En tant que guelfe il défendait les intérêts du pape en bataillant contre les Staufer en Sicile. Par conséquent, son symbole héraldique apparaît sur les armoiries de nombreuses familles nobles dans le but d'étendre leur loyauté à la famille et à sa position politique. Une recherche affinée d'un symbole image isolé à travers la segmentation aboutit à plus de 25 résultats d'armes ayant des compositions de blason très différentes et complexes, mais elles ont toutes été mises ensemble de manière non verbale grâce au contexte sémantique, qu'elles ont toutes en commun comme décrit ci-dessus.

Conclusion : sans l'ombre d'un doute la méthode de recherche d'images sur la base de segments isolés à partir d'autres images est une solution très spécifique pour les cas spéciaux comme l'héraldique homogène officiellement d'une grande portée. Dans les collections d'images très hétérogènes les résultats seraient certainement beaucoup plus décevant. Cependant, le projet a révélé une manière d'éviter les descriptions verbales compliquées des armoiries et d'obtenir des résultats utiles même en l'absence de connaissance au sujet de leur relation historique ou sémantique. Et aucune spécification de langue n'est requise, bien sûr. En regard des sciences de l'information le projet a réussi à combler le fossé sémantique mentionné ci-dessus en isolant tous les éléments constitutifs d'une image, en les représentant en tant qu'images et en surpassant le médium traditionnel de la description verbale. Puisque la quantité d'images disponibles pour les arts croît de

manière constante et que les bases de données qui y sont liées apparaissent de plus en plus dans des contextes supranationaux voire même globaux, la solution présentée ici pourrait aider à surpasser les problèmes de langage. Il est évident que de telles méthodes totalement automatisées joueront un rôle décisif dans le futur.

À propos de l'auteur :

Né en 1959. Il a étudié l'histoire de l'art, l'archéologie et la philosophie à l'Université d'Heidelberg et à l'Université de Vienne. Master of Art (1988) et Ph.D. (1991) à l'Université d'Heidelberg. De 1991 à 1994 il a été assistant au Hessisches Landesmuseum de Darmstadt. 1994/1995 Université des sciences appliquées (Göttingen/Frankfurt am Main) : Master en bibliothéconomie scientifique. Depuis 1996, il est bibliothécaire en chef au Kunsthistorisches Institut de Florence, le Max-Planck-Institut. Il est membre du comité d'accompagnement du projet international artlibraries.net. Il est membre de la commission permanente de la Art Libraries Section de l'IFLA, et président depuis 2009.