

Repérage d'images ordinaires : analyse des requêtes des chercheurs d'images

Elaine Ménard, Ph. D.
Assistant Professor / Professeure adjointe
School of Information Studies
McGill University

Abstract

For many years, the Web became an important media for the diffusion of multilingual resources. However, linguistic differences still form a major obstacle to scientific, cultural, and educational exchange. Besides this linguistic diversity, a multitude of databases and collections now contain documents in various formats, which may also adversely affect the retrieval process. For example, image collections have many areas of application. The need to retrieve images from a particular collection is shared by several user communities. Until recently, image collections were difficult to access due to limitations in dissemination and duplication procedures. The advent of the Web highlighted the pressing need to develop suitable tools for the description of digital images, since these can be found in the majority of available resources. This research compares image retrieval within two contexts: a monolingual context, i.e., where the language of the query (French) is the same as the indexing language (French); and a multilingual context, i.e., where the language of the query (French) is different from the indexing language (English). This article presents the results of the analysis of the queries used to retrieve a set of ordinary images representing everyday-life objects in a multilingual retrieval context. The analysis of the terms contained in the queries of the image searchers illustrates the terminological, perceptual, and structural trends. In general, participants employ simple queries containing only one word. Moreover, the queries are often generic in nature. The analysis also underlines that participants frequently use functional attributes to retrieve ordinary images. On the structural level, generic relations dominate other types of relations. The contributions of this research project are manifold. At the theoretical level, this research identifies the characteristics essential to the indexing process of digital images. This study also emphasizes the pressing necessity to optimize the methods used for image processing, in order to facilitate the image retrieval and their dissemination in multilingual environments.

Résumé

Depuis quelques années, le web est devenu un média incontournable pour la diffusion de ressources multilingues. Cependant, les différences linguistiques constituent souvent un obstacle majeur aux échanges de documents scientifiques, culturels, pédagogiques et commerciaux. En plus de cette diversité linguistique, on constate le développement croissant de bases de données et de collections composées de différents types de documents textuels ou multimédias, ce qui complexifie également le processus de repérage documentaire. Par exemple, les collections d'images numériques sont aussi nombreuses que diversifiées. Le besoin de repérer une image spécifique dans diverses

collections est devenu une préoccupation partagée par plusieurs communautés. La croissance du web a mis en relief le besoin pressant de se doter d'outils propres à la description des images dans le but de faciliter leur repérage, puisque l'on retrouve celles-ci dans la plupart des ressources disponibles. Cette recherche compare le repérage d'images dans deux contextes linguistiques différents : un contexte monolingue, c'est-à-dire que la langue de la requête (français) est la même que la langue d'indexation (français); et un contexte multilingue où la langue de la requête (français) est différente de la langue d'indexation (anglais). Cet article présente les résultats de l'analyse des requêtes formulées par les participants pour repérer un ensemble d'images ordinaires représentant des objets de la vie quotidienne, en contexte de repérage multilingue. L'examen des termes contenus dans les requêtes des chercheurs d'images révèle les tendances observées sur le plan terminologique, perceptuel et structurel. Les participants emploient généralement des requêtes simples ne contenant qu'un seul mot. De plus, les requêtes sont souvent de nature générique. L'analyse des requêtes indique également que les participants utilisent fréquemment les attributs fonctionnels pour repérer ce type d'images. Sur le plan structurel, les relations de genre à espèce prédominent les autres types de relations. Les contributions de ce projet de recherche sont multiples. Sur le plan théorique, cette étude identifie les caractéristiques essentielles au processus d'indexation de l'image numérique. Cette étude souligne également le besoin pressant d'optimiser les méthodes utilisées pour le traitement d'images, afin d'améliorer leur repérage et leur diffusion d'images dans les environnements multilingues.

Keywords

Image processing; image retrieval, image indexing, multilingual information retrieval; query analysis

Mots-clés

Traitement de l'image; repérage d'images; indexation d'images; repérage d'information multilingue; analyse des requêtes

Problématique

Les collections d'images numériques disponibles sur le web sont aussi nombreuses que diversifiées. Les bibliothèques, les musées, les services d'archives, les centres d'information et de recherche, les hôpitaux, les établissements d'enseignement, les journaux, les services commerciaux et gouvernementaux, de même que les individus ont peu à peu commencé à élaborer leurs propres collections d'images. Ainsi, de plus en plus d'images sont accessibles et peuvent être utilisées pour plusieurs applications précises. Parmi celles-ci, mentionnons la gestion des collections muséales et des galeries d'art, le dessin architectural et d'ingénierie, la décoration intérieure, la télédétection et l'échographie terrestre, les systèmes d'informations géographiques, la gestion de bases de données scientifiques, les prévisions météorologiques, le commerce de détail, la création de mode et de tissus, la gestion des marques de

fabricants et des droits d'auteurs, le maintien de l'ordre et la prévention de la criminalité, l'archivage photographique, le divertissement personnel, l'imagerie biomédicale, les utilisations militaires, touristiques, éducationnelles, etc. (Gudivada et Raghavan 1995, 18). Les images sont employées, entre autres, à des fins personnelles, commerciales, scientifiques et éducatives. Dans certains cas, elles servent d'instrument de travail (médecine, architecture, météorologie, etc.).

Parmi les images disponibles, on distingue en général trois grandes catégories : l'image d'art, l'image documentaire et l'image ordinaire. L'image d'art se caractérise par un traitement documentaire élaboré pouvant inclure plusieurs niveaux de signification (pré-iconographique, iconographique, iconologique). L'image documentaire se retrouve principalement dans les archives historiques, les agences de presse et les archives des médias. Par exemple, les photographies familiales appartiennent à cette catégorie. La principale caractéristique du traitement documentaire de ce type d'images est l'utilisation d'une légende ou toute autre information supplémentaire ayant pour objectif d'identifier l'image au sein d'une collection spécifique. Finalement, l'image ordinaire sert généralement à des fins commerciales ou d'illustration. Par images ordinaires, on entend les images représentant des objets de la vie quotidienne, par exemple : une bicyclette, un marteau, une cafetière, etc. Le qualificatif « ordinaire » distingue ce type d'images avec les images dites d'art, c'est-à-dire dont la valeur première à titre de document visuel est d'être la manifestation d'une expression artistique (Turner 1998, 84). Pour le traitement documentaire de ce type d'images, on tient généralement peu compte de certaines métadonnées descriptives ou analytiques comme le nom de l'auteur ou la date de création, mais on considérera plutôt le contenu visuel de l'image, c'est-à-dire son sujet (Greisdorf et O'Connor 2008, 53).

L'image ordinaire occupe une place importante dans les recherches des internautes. De plus en plus d'individus et d'organisations s'intéressent aux images ordinaires. Par conséquent, l'élaboration de moyens d'accès efficaces à ce type d'images est devenue nécessaire. En effet, l'augmentation du nombre d'images disponibles sur le web met en relief le besoin pressant de se doter d'outils pouvant faciliter leur repérage, puisque l'on retrouve celles-ci dans la plupart des ressources : pages web personnelles, bibliothèques et collections muséales virtuelles, catalogues de services et produits commerciaux, information gouvernementale, etc.

En général, deux catégories de requêtes sont employées pour le repérage d'images. D'une part, l'individu peut soumettre une requête graphique (utilisation d'une image ou d'un dessin) et le système tente de repérer une image similaire en utilisant certains aspects de l'image tels que la forme, la couleur, la texture, etc. Ces caractéristiques physiques de bas niveau peuvent être intéressantes dans le cas du furetage, c'est-à-dire lorsque l'individu n'a pas une idée très précise de ce qu'il recherche (Markkula et Sormunen 2000, 283). Toutefois, on constate que ces caractéristiques extraites automatiquement de l'image n'ont pas nécessairement de sens pour l'individu. En outre, les systèmes basés sur le contenu de l'image (*Content-Based Image Retrieval – CBIR*) utilisés pour l'indexation et le repérage et exploitant les caractéristiques physiques de l'image restent complexes à mettre au point et à utiliser. Pour le moment, les méthodes d'indexation et de repérage basées sur le contenu physique de l'image, bien que pleines

de promesses, demeurent au stade expérimental (Jørgensen 2003, 167). Par conséquent, l'approche de repérage d'images utilisant les requêtes textuelles demeure encore la voie la plus utilisée par les internautes. Ainsi, le repérage de l'image peut se faire à l'aide du texte entourant celle-ci ou encore avec l'information décrivant l'image, c'est-à-dire les métadonnées créées manuellement ou automatiquement au moment de son indexation. Ce type de repérage s'apparente au repérage de documents textuels puisque le succès dépend alors de la correspondance devant être établie entre les termes de la requête et le texte associé aux images. La majorité des moteurs de recherche disponibles sur le web offrant le repérage d'images repose toujours sur ce mode de repérage (Tjondronegoro et Spink 2008, 341).

Deux difficultés peuvent survenir au moment du repérage d'images ordinaires effectué à l'aide d'une requête textuelle. Le premier obstacle concerne le transfert de la représentation visuelle de l'objet à repérer en une représentation textuelle (Gudivada et Raghavan 1995, 22 ; Jansen 2008, 85). Selon Turner (1994, 33), le risque d'ambiguïté et d'erreur au moment de ce transfert demeure important. En effet, la possibilité d'interprétations multiples est caractéristique de la description de ressources visuelles. Il n'y a aucune manière unique de décrire le contenu d'une image, comme le dit si bien l'adage : une image vaut mille mots. En d'autres termes, les chercheurs d'images ne représentent pas nécessairement une même image avec les mêmes concepts, ni avec les mêmes mots. La deuxième difficulté concerne le grand nombre de langues que l'on retrouve sur le web. Étant donné cette multiplicité linguistique, on doit aussi s'attendre à ce que le texte associé aux images existe en de nombreuses langues différentes. Par conséquent, lorsqu'il utilise une requête textuelle, le chercheur d'images fait face à un double défi au moment de repérer celles-ci. D'une part, sa requête doit correspondre au texte associé aux images et d'autre part, la langue de sa requête doit également correspondre à la langue du texte associé aux images.

Cette étude examine les requêtes formulées par les chercheurs pour le repérage d'images ordinaires d'objets de la vie quotidienne et décrit les tendances observées pour le repérage de cette catégorie d'images.

Méthodologie

Une base d'images ordinaires a été constituée à partir des images tirées d'un catalogue en ligne d'une société commerciale canadienne ayant accepté de mettre son catalogue à notre disposition. Pour les besoins de cette recherche, 3 950 images ont ainsi été recueillies. L'indexation de chacune des images a été faite en deux langues : en français et en anglais. En outre, les images de la base ont été indexées de deux manières différentes : en vocabulaire contrôlé et en vocabulaire libre.

Lors du processus d'indexation en vocabulaire contrôlé, l'indexeur a reçu comme directive d'utiliser la politique d'indexation spécifiquement conçue pour cette recherche qui préconise l'utilisation du *Nouveau dictionnaire visuel multilingue* (Corbeil et Archambault, 2003) pour l'indexation en vocabulaire contrôlé de chacune des 3 950 images contenues dans la base. Ce dictionnaire visuel regroupe 32 000 mots ainsi que 6 000 illustrations correspondant au type d'images incluses dans la base, c'est-à-dire

des images d'objets de la vie quotidienne. Le processus d'indexation en vocabulaire contrôlé s'est déroulé comme suit. L'indexeur visualisait d'abord l'image à indexer, puis repérait dans le dictionnaire l'image correspondant le mieux à celle-ci. Les termes associés à l'image du dictionnaire étaient ensuite utilisés comme termes d'indexation, en français et en anglais. Entre un et cinq termes d'indexation en vocabulaire contrôlé pouvaient être attribués aux images. Si l'image à indexer ne correspondait à aucune image du dictionnaire, un terme plus général du *Visuel* était alors retenu comme terme d'indexation. Par exemple, l'image d'une « clôture à neige » qui ne serait pas représentée dans le *Visuel* aurait été indexée avec le terme « clôture », qui lui se trouve dans le dictionnaire.

Une politique spécifique pour l'indexation en vocabulaire libre a également été élaborée pour cette recherche. Cette politique prévoyait que l'indexeur visualisait d'abord l'image à indexer et utilisait ensuite ses propres termes pour la décrire. Ce type d'indexation s'apparente à ce que l'on retrouve dans les systèmes de partage d'images de type « Flickr » que l'on trouve sur le web. La politique d'indexation en vocabulaire libre donnait aux indexeurs des directives spécifiques quant à certains aspects (indexation orientée sur l'image, niveau d'exhaustivité et niveau de spécificité), mais n'imposait aucune règle quant au choix des termes et à leur forme. En outre, aucune restriction n'a été imposée à l'indexeur en ce qui concerne le nombre de termes d'indexation à attribuer aux images.

Une fois le processus d'indexation achevé, chaque image était associée à six formes d'indexation (vocabulaire contrôlé français, vocabulaire libre français, combinaison du vocabulaire contrôlé et libre français, vocabulaire contrôlé anglais, vocabulaire libre anglais, combinaison du vocabulaire contrôlé et libre anglais), tel qu'illustré à la *Figure 1*.


Image	Forme d'indexation
	Vocabulaire contrôlé français (VCF) vélo cross
	Vocabulaire libre français (VLF) bicyclette
	Vocabulaire contrôlé et libre français (VCLF) vélo cross bicyclette
	Vocabulaire contrôlé anglais (VCA) mountain bike
	Vocabulaire libre anglais (VLA): BMX bike
	Vocabulaire contrôlé et libre anglais (VCLA) mountain bike BMX bike

Figure 1 - Exemple d'image à repérer et des termes d'indexation associés

La simulation du repérage d'un ensemble de trente images indexées selon chacune des formes d'indexation a ensuite été effectuée auprès de soixante participants, répartis de manière aléatoire en six groupes indépendants de dix participants. Chacun de ces groupes était associé à une des formes d'indexation. Pour des considérations d'éthique, les participants devaient être âgés de 18 ans. De plus, afin d'assurer l'homogénéité du

groupe de participants, trois critères de sélection ont été définis : participants dont la langue maternelle est le français, étudiants de l'Université de Montréal de niveau 1^{er} cycle, et finalement, étant donné la nature de la tâche à effectuer lors de l'expérience, c'est-à-dire le repérage d'images, les participants ne devaient posséder aucune expérience professionnelle dans un domaine utilisant l'indexation et le repérage d'images.

Le repérage de chaque image a été amorcé par une image montrée (*Target Search* ou *Known-Item Search*), c'est-à-dire que le participant tentait de repérer l'image qui lui était d'abord présentée. Cette technique est régulièrement utilisée dans les expériences en repérage d'images et dans nombreuses recherches réalisées dans le cadre du Cross Language Evaluation Forum (CLEF) (Clough et Sanderson 2006, 702). Le chercheur d'images doit alors interpréter l'image montrée et formuler une requête dans une langue donnée. L'amorce de repérage par image montrée mobilise certaines ressources sensorielles et cognitives des chercheurs d'images, ce qui peut avoir une influence sur la formulation des requêtes et par conséquent, sur le repérage. L'ordre des images à repérer était toujours le même pour tous les participants, de manière à ce que l'effet d'apprentissage soit le même pour toutes les conditions de l'étude, c'est-à-dire les six formes d'indexation.

La simulation du repérage s'est déroulée en milieu contrôlé dans un local de recherche de l'École de bibliothéconomie et des sciences de l'information (EBSI), de l'Université de Montréal. Tous les participants devaient repérer les trente mêmes images, présentées dans le même ordre. Un système de repérage d'images « maison » spécifiquement conçu pour cette expérimentation et relié à la base d'images a été employé pour la simulation du repérage. Ce système, nommé EI@ine, a été programmé avec le langage asp.net 1.1 et utilisait une base de données MySQL. Le système comportait quatre composantes principales : une interface de recherche en français permettant au participant de formuler une requête textuelle, un mécanisme de traduction de la requête du participant du français vers l'anglais, un module de mise en correspondance des termes de la requête du participant (ou la traduction des termes de la requête) avec les termes d'indexation associés aux images et une interface d'affichage des résultats du repérage donnant la possibilité au participant de sélectionner une image satisfaisant sa requête ou de raffiner celle-ci. Toutes les requêtes des participants ont été formulées en français, ce qui a permis de comparer deux contextes linguistiques de repérage : un contexte monolingue, c'est-à-dire que la langue de la requête (français) est la même que la langue d'indexation (français), et un contexte de repérage multilingue, c'est-à-dire que la langue de la requête (français) est différente de la langue d'indexation (anglais). Pour ce deuxième contexte, les requêtes des participants, formulées en français, étaient d'abord traduites en anglais par un système de traduction automatique intégré au système de repérage, puis mises en correspondance avec les termes de l'une des formes d'indexation en anglais. Lors de l'élaboration du système EI@ine, le système BabelFish, développé initialement par le moteur de recherche AltaVista et acquis en 2006 par Yahoo! (Yahoo 2008) a été retenu comme mécanisme de traduction automatique. Aux fins de cette recherche, la traduction automatique des requêtes plutôt que des documents, semblait mieux adaptée puisque généralement, les requêtes contiennent peu de mots à traduire. Par conséquent, la traduction pouvait se faire

rapidement et à un moindre coût. Ainsi, trois groupes de participants ont respectivement été associés à un vocabulaire français (contrôlé, libre et combinaison du contrôlé et du libre), alors que les trois autres groupes étaient associés à un vocabulaire anglais (contrôlé, libre et combinaison du contrôlé et du libre). Aucune vérification n'a été faite à savoir si la traduction obtenue était correcte ou non. Si le système de traduction automatique ne retournait aucun résultat, la requête formulée en français était alors conservée. De plus, le processus de traduction des requêtes a été fait à l'insu des participants qui ignoraient à quelle forme d'indexation était associé le groupe auquel ils appartenaient.

L'interface de recherche du système El@ine comportait deux sections. Dans la partie supérieure, les images à repérer étaient montrées à tour de rôle au participant. Ce dernier avait la possibilité d'agrandir l'image en cliquant sur celle-ci avec la souris. Dans la portion inférieure de l'interface se trouvait une boîte de recherche où le participant formulait une requête textuelle en français pour tenter de repérer l'image qui lui était présentée. Aucune suggestion terminologique n'était faite au participant qui devait transposer l'image montrée dans ses propres mots. Une fois la correspondance établie entre les termes de la requête et les termes d'indexation, l'interface d'affichage présentait de nouveau l'image à repérer au participant dans la partie supérieure, ainsi que les images résultant de la requête formulée. S'il le désirait, le participant pouvait alors formuler une autre requête. Aucun texte n'accompagnait les images affichées. En effet, l'examen de certains moteurs de recherche comme Google Images, Yahoo Search!, Alltheweb et AltaVista, indique que ceux-ci n'offrent pas à leurs utilisateurs la possibilité de visionner le texte associé aux images repérées. Pour cette recherche, il a donc été décidé qu'aucun texte n'accompagnerait les images. De cette façon, le participant ne pouvait être influencé par la soumission d'une terminologie à laquelle il n'avait pas pensé pour la formulation de ses requêtes. Cette manière de procéder, même si elle va à l'encontre des conclusions de certains travaux de recherche préconisant la présentation de l'image et du texte associé au moment de l'affichage des résultats (Pu 2005 ; Clough, Al-Maskari et Darwish 2006 ; Karlgren et Olsson 2006), reste légitime puisqu'elle correspond au type d'affichage des résultats de la plupart de moteurs de recherche web offrant le repérage d'images. Une fois satisfait du résultat obtenu, le participant enregistrait son résultat et passait au repérage de l'image suivante.

L'ensemble des 8 436 requêtes générées en cours d'expérience (1 800 images recherchées), toutes formes d'indexation confondues, ont été examinées. Les résultats de cette analyse sont présentés dans les prochains paragraphes. Les démarches utilisées pour le repérage, de même que les difficultés éprouvées par les participants sont ensuite décrites.

Résultats

Les termes contenus dans les requêtes des participants ont été examinés à l'aide d'une grille adaptée de la littérature et développée expressément pour cette recherche. Le *Tableau 1* résume les principales caractéristiques des termes contenues dans les

requêtes. Cette analyse est présentée de manière globale et non par forme d'indexation puisque les participants ignoraient à quelle forme ils étaient associés.

Tableau I - Tableau récapitulatif de l'analyse des requêtes

Caractéristiques	Catégories	Sous-catégories	Termes des requêtes (N = 8 436)	
			n	%
Terminologiques	Types de termes	Termes simples	6 683	79,2
		Termes composés	1 753	20,8
		<i>Total</i>	<i>8 436</i>	<i>100,0</i>
	Autres types de termes	Termes complexes	8	0,1
		Abréviations	446	5,3
		Néologismes	169	2,0
		<i>Total</i>	<i>623</i>	<i>7,4</i>
	Niveaux	Termes génériques	6 881	81,6
		Termes spécifiques	1 555	18,4
<i>Total</i>		<i>8 436</i>	<i>100,0</i>	
Perceptuelles	Physiques	Forme	33	0,4
		Dimension	43	0,5
		Couleur	167	2,0
		Texture	13	0,2
		Matériel	109	1,3
		Partie du corps	46	0,5
		<i>Total</i>	<i>411</i>	<i>4,9</i>
	Fonctionnelles	Fonction	1 309	15,5
		Résultat/produit	140	1,7
		Activité/action	484	5,7
		Lieu	296	3,5
		<i>Total</i>	<i>2 229</i>	<i>26,4</i>
	Identificateurs	Marque	74	0,9
		Nom de produit	6	0,1
		Commanditaire	35	0,4
		Lieu géographique	37	0,4
		<i>Total</i>	<i>152</i>	<i>1,8</i>
	Autres caractéristiques	Temporelle	14	0,2
		Numérique	10	0,1
Autre qualificatif		257	3,0	
Personne		42	0,5	
<i>Total</i>		<i>323</i>	<i>3,8</i>	
Structurelles	Relations	Génériques	1 365	16,2
		Partitives	110	1,3
		D'instances	45	0,5
		Associatives	44	0,5
		<i>Total</i>	<i>1 564</i>	<i>18,5</i>

Les trois catégories de caractéristiques des requêtes (terminologiques, perceptuelles et structurelles) sont maintenant examinées en détail.

Caractéristiques terminologiques

Les caractéristiques terminologiques se rapportent aux aspects lexicographiques des termes des requêtes. Notre analyse révèle que les participants à la simulation du repérage ont privilégié l'utilisation de requêtes simples ne contenant qu'un seul terme (79,2 pour cent), alors que les requêtes composées de plus d'un terme ne comptent que pour 20,8 pour cent du total des requêtes enregistrées. De la même manière, peu de requêtes complexes (0,1 pour cent), c'est-à-dire des requêtes contenant plusieurs mots et plusieurs concepts différents, ont été employées par les participants. Sur le plan de l'emploi des abréviations et des néologismes dans les requêtes formulées, des taux d'utilisation de 5,3 pour cent et 2,0 pour cent ont été observés respectivement. Finalement, les taux d'utilisation de termes génériques (81,6 pour cent) et spécifiques (18,4 pour cent) révèlent que dans la majorité des requêtes, les participants privilégient des requêtes de nature générique (par ex., « chandail » plutôt que « chandail de hockey ») pour repérer les images montrées.

Caractéristiques perceptuelles

Les caractéristiques perceptuelles des requêtes sont liées aux aspects physiques, fonctionnels et identificateurs contenus dans les images. Les termes des requêtes ont été examinés selon quatre catégories de caractéristiques (physiques, fonctionnelles, identification et autres caractéristiques). Cet examen démontre que les participants utilisent peu la forme, la dimension, la texture ou les parties du corps pour le repérage d'images, avec des taux variant de 0,2 à 0,5 pour cent. Les caractéristiques couleur et matériel sont par contre un peu plus employées, avec des taux d'utilisation de 2,0 pour cent et 1,3 pour cent respectivement. Par ailleurs, l'analyse des caractéristiques fonctionnelles contenues dans les requêtes indique que la fonction de l'objet représentée par l'image est fréquemment employée par les participants (15,5 pour cent), alors que les autres caractéristiques fonctionnelles (résultat/produit, activité/action et lieu) apparaissent moins souvent dans les requêtes, avec des taux variant de 1,7 à 5,7 pour cent. Parmi les attributs identificateurs, la marque commerciale se distingue avec un taux d'utilisation de 0,9 pour cent, alors que les autres caractéristiques de cette catégorie (nom de produit, commanditaire et lieu géographique) présentent des taux d'utilisation moins importants, variant de 0,1 à 0,4 pour cent. Finalement, aucun des quatre autres caractéristiques perceptuelles examinées (temporelle, numérique, autre qualificatif et personne) ne se démarque vraiment avec des taux d'utilisation variant de 0,1 à 3,0 pour cent.

Caractéristiques structurelles

Les caractéristiques structurelles identifient les relations qui existent entre les termes contenus dans les requêtes. L'examen des quatre types de relations (génériques, partitives, d'instances et associatives) révèle dans un premiers temps que celles-ci n'apparaissent que dans 18,5 pour cent des requêtes des participants. Dans un deuxième temps, l'analyse de ces relations met en relief que les relations génériques dominent (par ex., lampe de chevet) avec un taux d'utilisation de 16,2 pour cent, alors

que les taux d'utilisation des trois autres types de sont beaucoup moins élevés, variant de 0,5 à 1,2 pour cent.

Démarches utilisées pour le repérage

Afin de comprendre la manière avec laquelle les participants ont effectué le repérage d'images, il a été demandé à ces derniers d'expliquer en quelques mots les démarches utilisées pour repérer les images montrées. Le *Tableau II* présente une synthèse des principales démarches décrites par les répondants pour effectuer leurs recherches. Il convient de mentionner ici que plusieurs stratégies primaires ou secondaires ont pu être employées par un même participant.

Tableau II - Synthèse des démarches de recherche des participants

Catégories de démarches	Principales stratégies	Nombre de participants (N = 60)		
		n	%	
Démarches primaires	Stratégies terminologiques	Mot précis	51	85,0
		Mot semblable	3	5,0
		Mot général	3	5,0
	Stratégies liées à l'objet	Fonction de l'objet	3	5,0
		Marque de l'objet	2	3,3
		Détails spécifiques de l'objet	2	3,3
		Forme de l'objet	1	1,7
Démarches secondaires	Heuristiques	Concept général	17	28,3
		Synonyme	12	20,0
		Fonction de l'objet	9	15,0
		Description physique	6	10,0
		Marque de l'objet	5	8,3
		Trait précis de l'objet	4	6,7
		Couleur de l'objet	4	6,7
		Défilement des images	3	5,0
		Concept spécifique	3	5,0
		Association d'idées	3	5,0
		Partie de l'objet	3	5,0
		Variante orthographique	2	3,3
		Mot anglais	2	3,3
		Localisation	1	1,7

Ce qui ressort principalement des propos recueillis est que la plupart du temps, deux stratégies spécifiques sont nécessaires pour effectuer une tâche de repérage d'image. D'une part, 85,0 pour cent des répondants indiquent utiliser les mots les plus précis qui leur viennent en tête pour débiter le repérage. Par la suite, si l'image montrée n'est pas repérée, une stratégie secondaire est employée. Parmi les heuristiques les plus utilisées comme stratégies secondaires par les participants, mentionnons l'utilisation de concepts plus généraux pour décrire l'objet (28,3 pour cent), de synonymes (20,0 pour cent) et de termes liés à la fonction de l'objet (15,0 pour cent).

Difficultés rencontrées au moment du repérage

Il a également été demandé aux participants de spécifier s'ils considéraient avoir rencontré des difficultés au moment du repérage. La *Figure 2* présente l'évaluation des difficultés éprouvées par les répondants, par forme d'indexation.

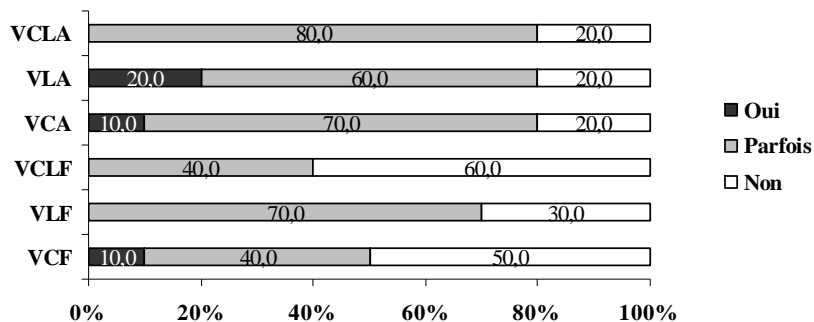


Figure 2 - Difficultés rencontrées par les participants au moment du repérage

L'examen de la figure précédente révèle que les formes d'indexation utilisant un vocabulaire anglais semblent avoir causé plus de difficultés aux participants, au moment du repérage, que les formes d'indexation en français. En effet, la majorité des répondants considèrent avoir rencontré ou parfois rencontré des difficultés au moment du repérage des images indexées en anglais, alors que ces taux sont inférieurs pour les formes d'indexation en français. Seulement 20,0 pour cent des participants considèrent n'avoir éprouvé aucune difficulté pour le repérage d'images indexées en anglais, alors que ces mêmes taux sont plus élevés pour les formes d'indexation en français.

Finalement, afin de mieux cerner les difficultés éprouvées au moment du repérage, il a été demandé aux participants de décrire celles-ci dans leurs propres mots. Le *Tableau III* résume les principaux obstacles rencontrés par les répondants au moment du repérage.

Tableau III - Principales difficultés rencontrées par les participants pendant le repérage

Difficultés rencontrées	Nombre de participants (N = 60)	
	n	%
Difficulté terminologique pour décrire l'objet représenté sur l'image	25	41,7
Résultats non pertinents	14	23,3
Difficulté d'identification de l'objet représenté sur l'image	13	21,7
Impossibilité de choisir parmi trop de résultats semblables	2	3,3
Incapacité à accomplir la tâche	2	3,3
Manque de spécificité dans la recherche	1	1,8
<i>Aucune difficulté spécifiée</i>	16	26,7

Le tableau précédent souligne que les principales difficultés, représentant plus de 85,0 pour cent des cas, sont les difficultés d'ordre terminologique au moment de décrire l'objet représenté par l'image montrée, le manque de pertinence des résultats obtenus et la difficulté à identifier les objets représentés par les images montrées. Il faut toutefois

mentionner qu'un peu plus du quart des participants (26,7 pour cent) considèrent n'avoir rencontré aucune difficulté au moment du repérage.

Discussion

L'examen des termes contenus dans les requêtes des participants s'étant prêtés à l'expérience de simulation de repérage a révélé certaines tendances marquées sur le plan des caractéristiques terminologiques, perceptuelles et structurelles des requêtes. Pour cette recherche, l'analyse des requêtes montre que les participants ont formulé des requêtes contenant en moyenne 1,3 terme. Cette moyenne est légèrement inférieure à ce que l'on retrouve habituellement dans la littérature. En effet, Goodrum et Spink (2001) considèrent qu'en général, les requêtes formulées pour le repérage d'images contiennent en moyenne 3,7 termes. La longueur moyenne des requêtes formulées pour la présente étude peut s'expliquer par le type d'images utilisées qui n'exigent pas une requête très sophistiquée. En effet, les requêtes formulées par les participants tiennent compte du contenu visuel de l'image, c'est-à-dire le niveau pré-iconographique lié aux objets représentés (Panofsky 1955 ; Markey 1984 ; Shatford 1986 ; Svenonius 1994 ; Enser 1995) et ne prennent pas en considération le niveau iconographique ou iconologique des objets représentés. Les termes des requêtes se rapportent donc aux objets ou aux situations identifiables que l'on retrouve sur l'image, ce que Shatford (1986) définit comme le niveau *of*, c'est-à-dire ce qui peut être identifié sur l'image. Le niveau *about* (interprétation des images) avec lequel les actions et les attitudes du matériel à repérer sont interprétés et le niveau de la signification qui réfère à l'interprétation de l'image s'appuyant sur le savoir et l'érudition (Turner 1998, 12) sont exclus du repérage puisque ceux-ci ne s'appliquent pas vraiment aux images ordinaires utilisées pour cette recherche.

En outre, l'examen des requêtes formulées pour cette recherche confirme les dires de Choi et Rasmussen (2003) considérant que la plupart des requêtes effectuées par les individus appartiennent à la catégorie « générique / identifiable ». Par ailleurs, l'analyse des requêtes de cette recherche corrobore les affirmations de Fukumoto (2004) ayant établi qu'en général les requêtes formulées pour repérer des images suivent un modèle bien défini : introduction d'un ou deux termes de recherche, visualisation des images repérées et retour à la page d'accueil, ou encore les affirmations de Pu (2005) soulignant que de manière générale les individus n'effectuent que des requêtes courtes pour le repérage d'images et que de nombreuses requêtes n'obtiennent aucun résultat. Cependant, il convient de mentionner que ces constatations s'appliquent aux requêtes formulées pour repérer les images utilisées pour cette étude, c'est-à-dire des images ordinaires d'objets de la vie quotidienne, et ne peuvent sans doute pas être indifféremment étendues aux requêtes formulées pour tous les types d'images.

Par ailleurs, rappelons qu'au moment du repérage des trente images, les participants ignoraient avec quelle forme d'indexation les requêtes formulées étaient mises en correspondance. Cette forme d'anonymat au niveau du vocabulaire d'indexation laisse croire que les participants ne pouvaient avoir de stratégie précise pour le repérage des images en ce qui concerne la langue. Autrement dit, les participants qui devaient repérer des images indexées dans une langue différente de celle des requêtes formulées ont

utilisé les mêmes démarches de repérage que les participants devant repérer des images indexées dans la même langue que celle des requêtes. Toutefois, lorsque les participants ont été interrogés sur le niveau de difficulté rencontrée au moment du repérage, les formes d'indexation utilisant un vocabulaire anglais semblent provoquer plus de difficultés au moment du repérage (Figure 2).

Une explication à cet inconfort provient sans doute du processus de traduction des requêtes devant être effectué lorsque la langue de la requête est différente de la langue d'indexation. Parmi les ressources linguistiques habituellement utilisées dans un système de recherche d'information multilingue (RIML), on retrouve les dictionnaires bilingues ou multilingues, les systèmes de traduction automatique et les corpus parallèles ou comparables (Ménard 2006, 257). La recension des écrits menée sur les différentes ressources linguistiques pouvant être employées dans les systèmes de RIML ayant démontré leurs avantages et leurs limites, le choix méthodologique effectué au moment de la conception du système de repérage EI@ine s'est porté sur la traduction des requêtes des participants par un système de traduction automatique intégré à même le système. À l'origine, il était légitime de croire que les requêtes utilisées pour le repérage d'images puissent être correctement traduites par un système de traduction automatique, étant donné que les requêtes contiennent en général peu de termes.

Or, un examen sommaire et partiel de la traduction des requêtes obtenue par le système de traduction automatique intégré dans le système de repérage démontre que souvent, les requêtes sont trop courtes, c'est-à-dire qu'elles ne fournissent pas suffisamment de contexte pour être bien interprétées et sont par conséquent incorrectement traduites. De ce fait, le mécanisme de traduction a eu une influence sur la perception du niveau de difficulté rencontrée par les participants au moment du repérage.

Finalement, lorsque les formes d'indexation sont comparées entre elles et cela, indépendamment de la langue, l'examen du degré de difficulté éprouvée par les participants indique que celui-ci est à la baisse lorsque les images sont indexées avec la combinaison du vocabulaire contrôlé et libre. Ce constat laisse supposer que la cohabitation des deux vocabulaires d'indexation semble intéressante du point de vue du chercheur d'images. En effet, la combinaison des avantages des deux vocabulaires d'indexation considérés séparément semble avoir facilité le repérage.

Conclusion

Depuis quelques années, on constate que les besoins et comportements de recherche des chercheurs d'images se sont considérablement modifiés. Comme le soulignent les travaux de plusieurs chercheurs (Goodrum et Spink 2001 ; Goodrum, Bejune et Siochi 2003 ; Jörgensen et Jörgensen 2005 ; Tjondronegoro et Spink 2008), on assiste à une évolution dans la manière de formuler les requêtes pour le repérage d'images. Par exemple, les individus ont tendance à utiliser des stratégies de recherche plus élaborées et à inclure de plus en plus des termes de recherche touchant les émotions, les thèmes et les relations qui existent entre les objets. Par conséquent, si la manière de

repérer l'image s'est modifiée peu à peu, il convient de se demander si l'indexation de l'image, et plus particulièrement, le vocabulaire contrôlé traditionnellement employé lors du processus d'indexation est bien adapté aux besoins et comportements réels et actuels des chercheurs d'images. L'indexation collaborative en vocabulaire libre que l'on retrouve dans les systèmes de partage d'images notamment, est considérée par plusieurs comme une manière intéressante de détrôner le rôle de l'indexation traditionnelle en vocabulaire contrôlé et de faciliter l'organisation et le repérage d'images sur le web. Cependant, il faudra attendre quelques années avant de voir si le phénomène d'indexation collaborative en vocabulaire libre constitue « la » solution pour optimiser le repérage d'images.

Mentionnons en terminant que les résultats de cette étude sont intimement liés à la catégorie d'images utilisées. Ainsi, on peut penser que l'analyse des requêtes formulées pour le repérage d'autres types d'images pourrait compléter judicieusement les résultats de cette recherche. Par exemple, l'examen des caractéristiques des requêtes pour le repérage d'images documentaires liées à un domaine spécifique (actualités sportives, imagerie médicales, etc.) ou d'images artistiques (objets muséologiques, œuvres célèbres, etc.) semblent des pistes de recherche nécessaires pour obtenir une meilleure compréhension de la manière avec laquelle l'utilisateur du web formule ses requêtes pour le repérage d'images.

Sources consultées

- Choi, Youngok, and Edie M. Rasmussen. "Searching for Images: the Analysis of User's Queries Image Retrieval in American History." Journal of the American Society for Information Science and Technology 54.6 (2003): 498-511.
- Clough, Paul, Al-Maskari, Azzah, and Darwish, Kareem. Providing Multilingual Access to FLICKR for Arabic Users. 2006. Web Page. URL:
<http://clef.iei.pi.cnr.it/2006/working_notes/workingnotes2006/cloughCLEF2006.pdf>
- Clough, Paul, and Mark Sanderson. "User Experiments With the Eurovision Cross-Language Image Retrieval System." Journal of the American Society for Information Science and Technology 57.5 (2006): 697-708.
- Corbeil, Jean-Claude, and Ariane Archambault. Le Nouveau Dictionnaire Visuel Multilingue. Montréal: Québec Amérique, 2003.
- Enser, Peter G. B. "Progress in Documentation Pictorial Information Retrieval ." Journal of Documentation 51.2 (1995): 126-70.
- Fukumoto, Toru. An Analysis of Image Retrieval Behavior for Metadata Type and Google Image Database. 2004. Web Page. URL:
<http://www.nier.go.jp/fukumoto/fukumoto_ICCE2004.pdf>.
- Goodrum, Abby, and Amanda Spink. "Image Searching on the Excite Web Search Engine ." Information Processing and Management 37.2 (2001): 295-311.
- Goodrum, Abby A., Bejune, Matthew M, and Siochi , Antonio C. A State Transition

Analysis of Image Search Patterns on the Web. 2003. Web Page. URL:

<<http://www.springerlink.com/media/6p8qd4b65j3vpne80ecn/contributions/0/e/w/7/0ew7r6lelaafcqv8.pdf>>.

Greisdorf, Howard F., and C. O'Connor, Brian. Structures of images collections: from Chauvet-Pont d'Arc to Flickr. Westport, Conn.: Unlimited Libraries, 2008.

Gudivada, Venkat N., and Vijay V. Raghavan. "Content-Based Image Retrieval Systems." IEEE Computer 28.9 (1995): 18-22.

Jansen, Bernard J. "Searching for Digital Images on the Web." Journal of Documentation 64.1 (2008): 81-101.

Jørgensen, Corinne, and Peter Jørgensen. "Image Querying by Image Professionals." Journal of the American Society for Information Science and Technology 56.12 (2005): 1346-59.

Jørgensen, Corinne. Image Retrieval : Theory and Research. Lanham MD: Scarecrow Press, 2003.

Karlgren, Jussi, and Olsson, Fredrik. Trusting the Results in Crosslingual Keyword-Based Image Retrieval. 2006. Web Page. URL: <http://clef.iei.pi.cnr.it/2006/working_notes/workingnotes2006/karlgrenCLEF2006.pdf>.

Markey, K. "Inter-Indexer Consistency Tests: a Literature Review and Report of a Test Consistency in Indexing Visual Materials." Library and Information Sciences Research 6 (1984): 155-77.

Markkula, Marjo, and Eero Sormunen. "End-User Searching Challenges Indexing Practices in the Digital Newspaper Photo Archive." Information Retrieval 1.4 (2000): 259-85.

Ménard, Elaine. "La recherche d'information multilingue." Documentation et bibliothèques 52. 4 (2006): 255-61.

Panofsky, Erwin. Meaning in the Visual Arts : Papers in and on Art History. Garden City, N.Y.: Doubleday, 1955.

Pu, Hsiao-Tieh. "A Comparative Analysis of Web Image and Textual Queries ." Online Information Review 29.5 (2005): 457-67.

Shatford, Sara. "Analyzing the Subject of a Picture: a Theoretical Approach." Cataloging & Classification Quarterly 6.3 (1986): 39-61.

Svenonius, Elaine. "Access to Nonbook Materials: the Limits of Subject Indexing for Visual and Aural Languages ." Journal of the American Society of Information Science 45.8 (1994): 600-6.

Tjondronegoro, Dian, and Amanda Spink. "Web Search Engine Multimedia Functionality." Information Processing & Management 44.1 (2007): 340-57.

Turner, James. Determining the Subject Content of Still and Moving Image Documents for Storage and Retrieval : an Experimental Investigation . Toronto: Toronto University, 1994.

---. Images en mouvement : Stockage, repérage, indexation. Sainte-Foy: Presses de l'Université du Québec, 1998.

"Yahoo! Babel Fish - Text Translation and Web Page Translation." 20 May 2009

<<http://babelfish.yahoo.com/>>.