

ASSISTANCE A LA CONCEPTION ARCHITECTURALE ET TECHNIQUE PAR DES METHODES ET OUTILS DE RECHERCHE D'INFORMATIONS PAR L'IMAGE

Jean-Claude BIGNON, Gilles HALIN, Walaiporn NAKAPAN,
Pascal HUMBERT, Marc WAGNER

CRAI (Centre de Recherche en Architecture et Ingénierie) UMR M.A.P. Culture/CNRS, N°694
École d'Architecture de Nancy, 2 rue Bastien Lepage, 54000 Nancy
Tél. : 03.83.40.81.34 - Fax : 03.83.40.81.30
email :{bignon,halin,nakapan,humbert,wagner}@crai.archi.fr

CONTEXTE GENERAL

Le présent article se situe dans le cadre général de nos travaux de recherche sur l'assistance à la conception architecturale et technique.

Du point de vue des sciences cognitives, nous abordons l'activité de conception à partir de deux hypothèses principales :

- L'activité de conception architecturale est une activité de résolution de problème singulière. Elle se caractérise essentiellement par les faits suivants :
 - Le problème est généralement mal posé ou souvent posé incomplètement,
 - Les méthodes pour le résoudre sont nombreuses et mises en œuvre de manière opportuniste par le concepteur.
- La résolution de problèmes se fait largement par des raisonnements à base de références (RBR). Une référence est un dispositif heuristique qui aide le concepteur à résoudre son problème, à mieux le formuler ou à formuler un problème nouveau. Parmi tous les systèmes de références utilisés (mots, musique, ambiance, édifice ou ouvrage modèle...), l'image joue un rôle privilégié.

CONTEXTE PARTICULIER

Le terme d'assistance à la conception renvoie à des méthodes et des attitudes différentes selon que l'on se situe à l'amont du processus (Esquisse, APS) ou à l'aval (APD, PEO...). Afin de

bien borner les limites de notre travail, il est important d'indiquer que nous travaillons actuellement plus spécifiquement sur l'assistance à la recherche d'informations relatives aux produits du bâtiment. Un tel thème concerne généralement une phase déjà avancée du processus de conception.

Dans un tel contexte, notre hypothèse particulière est la suivante.

La recherche d'information fondée sur une approche par critères explicites (nom d'un produit, nom d'une marque, performances particulières...) est bien adaptée à des situations où la formulation de problème est aisée. C'est souvent le cas dans les phases finalisées de la conception où le travail de conception relève moins d'une activité de création que d'une activité de vérification (Calcul de structure, thermique, CCTP...). Elle est en revanche souvent mal adaptée à des stades antérieurs lorsque le concepteur doit faire des choix alors qu'il ne dispose pas de toutes les informations pour le faire ou qu'il doit trouver des solutions répondant à des critères multiples. Dans ce type de raisonnement plus incertain, il est utile de travailler sur des modes de recherche de l'information plus flous et moins focalisés.

Nos travaux visent donc à mettre en place des méthodes et des outils de recherche par l'image qui utilisent au mieux les potentialités de ce média et les aptitudes des architectes à raisonner à partir de figures visuelles. Ils viennent en complément d'autres travaux portant sur la structuration de l'information dans le domaine du bâtiment comme les IFC ou ceux relatifs aux potentialités des NTIC pour aider à la conception ou à la réalisation des édifices.

PERTINENCE D'IMAGE

Les raisons de l'efficacité de l'image sont connues et nombreuses. On citera en particulier les caractéristiques suivantes:

- Sensibilité physiologique importante du visuel ;
- Grande aptitude mnémonique des images ;
- Forte capacité d'encodage de l'information ;
- Traitement parallèle de l'information ;
- Message global instantané ;
- Effet de preuve;
- Séduction de l'iconique;
- Support privilégié pour le raisonnement spatial.

Mais l'image n'est pas sans souffrir de limites qui peuvent conduire à une désorientation informationnelle (le récepteur ne sait pas ce qu'il voit) ou à une mésinterprétation sémantique (le récepteur interprète une information de manière différente de celle voulue par l'émetteur).

Nous évoquons quelques-unes des propriétés qui peuvent entraîner des difficultés d'interprétation dans notre champ d'application.

La polysémie iconique. Une image peut véhiculer plusieurs lectures. Son interprétation dépend du point de vue adopté à un moment par le récepteur. Un même objet iconique sert de vecteur à des objets sémantiques différents.

La surcharge sémiographique. Cette situation se rencontre lorsqu'une image représente plusieurs objets différents (image mosaïque, scène composée de plusieurs objets...). La discrétisation de l'image conduit à une défocalisation du regard.

Le déficit informationnel. Pour des raisons stylistiques (cadrage esthétique, représentation métaphorique...) ou techniques (faible résolution de l'image, nombre de couleurs limité...), l'image contient une information réduite qui induit une interprétation complétive.

Dans toutes ces situations, l'effort d'interprétation peut conduire à un écart entre le message édité par l'émetteur et le message interprété par le récepteur. Or du point de vue de la théorie du rendement, une image est pertinente si les effets visuels qu'elle produit suffisent à équilibrer les efforts nécessaires à son interprétation [1].

Afin de diminuer l'effort d'interprétation, il convient donc d'identifier les critères qui favorisent la perception des images et donc leur valeur de pertinence. Dans notre travail, nous avons identifié trois grandes familles de pertinence d'image.

Les critères identifiés ont été définis à partir d'un corpus d'images assujetti à une double contrainte:

- Les images doivent permettre de retrouver des produits du bâtiment ;
- Les images sont extraites du web.

Pertinence morphologique

La pertinence morphologique désigne un premier niveau physique de sélection d'images relatif à leur forme. La définition des caractéristiques de forme renvoie à plusieurs critères :

- Des critères physiologiques (sélectivité de l'œil dans le décodage d'une image, angle de champs visuel...);
- Des critères machines (format, taille et résolution des écrans pour afficher les images...);
- Des critères informatiques (application du format « image » à des éléments d'interface comme les boutons, les bandeaux graphiques...).

La pertinence morphologique prend en compte essentiellement deux aspects : la taille et la proportion de l'image.

Une image est pertinente morphologiquement si elle correspond aux critères suivants :

- Aucune de ses dimensions ne doit être inférieure à 60 pixels. En dessous de cette valeur, l'image n'est plus lisible et ne peut plus transmettre qu'une information pauvre

- graphiquement de type symbole. Par ailleurs, la probabilité que l'image ne soit qu'un bouton, une puce ou une ligne d'interface augmente ;
- Aucune de ses dimensions ne doit être supérieure à 600 pixels. Au-delà de cette valeur, l'image occupe une place trop grande dans l'écran et ne peut plus être perçue qu'en partie ;
 - Sa proportion (largeur/hauteur) doit être dans un intervalle limité situé entre 0,6 et 1.5. Cette fourchette permet d'intégrer des images proches du format photographique courant 24 x 36 vertical ou horizontal. Dès que l'on s'éloigne trop de ce format, la probabilité que l'image ne corresponde qu'à un élément graphique de l'interface (bandeau...) augmente fortement.

Pertinence sémiologique

La pertinence sémiologique s'attache à identifier le sens véhiculé par une image en fonction de sa structure graphique. Compte tenu de notre approche « orientée produit » une analogie visuelle forte entre l'objet réel et sa représentation est importante. On remarquera que cette ressemblance s'appuie sur deux conditions que nous supposons réunies. L'objet représenté doit être connu. Sa représentation (mode, vue...) doit être effectuée dans un univers d'interprétation partagé entre l'émetteur et le récepteur.

Nous avons identifié plusieurs critères qui favorisent cette pertinence. Il n'est pas utile que tous les critères soient remplis. Certains critères peuvent être compensés par d'autres. Mais d'une manière générale, plus une image réunit de critères, plus l'effort d'interprétation sera réduit.

- Il doit exister une similarité de couleur entre l'image et les couleurs habituellement dominantes de l'objet représenté. Un toit jaune n'évoquera pas immédiatement une toiture en tuiles de terre cuite ;
- Un objet mis en situation de contraste de lumière, de couleur ou de forme dans une image sera perçu comme plus important dans la scène. Un petit crochet rouge sur un fond à dominante verte sera pris en compte plus fortement ;
- L'objet doit plutôt être représenté en entier. Plus il sera tronqué, plus il faudra interpréter les parties manquantes ;
- La représentation d'un objet sous un angle à partir duquel il n'est généralement pas perçu oblige à un repositionnement virtuel qui augmente la difficulté à le reconnaître ;
- La vue doit permettre à l'objet représenté d'occuper une surface importante dans l'image. Plus la surface relative diminue, moins l'objet devient signifiant dans la scène ;
- Un produit est plus facilement reconnu si l'image ou il est présenté intègre des éléments de son contexte d'usage ou de mise en œuvre. Ceci est particulièrement vrai pour tous les accessoires

Pertinence contextuelle

Au-delà du décodage sémiologique, l'interprétation d'une image implique des processus inférentiels reposant sur des informations non codées dans l'image généralement appelées contexte [2]. La notion de contexte définit tout autant le contexte immédiat de l'image (le texte qui lui est lié, les autres images avec lesquelles elle est en relation) que le contexte d'interprétation du lecteur.

Nous avons neutralisé ce dernier aspect en faisant l'hypothèse que nous travaillons dans un domaine d'expertise partagée dans lequel il existe une parenté forte des univers de référence à partir desquels sont interprétées les images.

Une image est pertinente contextuellement si elle correspond aux critères suivants :

- Une image doit pouvoir être attachée à au moins un mot clef d'un thésaurus du domaine ;
- Plus le mot-clef est proche de l'image plus elle est pertinente. La proximité peut être physique (distance entre une image et un texte) ou sémantique (degré d'interprétation du concept représenté par le mot-clef).

NOS METHODES ET OUTILS

Extraction et indexation d'images à partir du WEB

Internet et plus particulièrement le Web représente une source importante d'informations où les images sont nombreuses et variées. Aujourd'hui il existe de plus en plus de fabricants qui proposent leur catalogue illustré de produits sur Internet. L'analyse de ces sites et l'extraction de leurs images, à l'aide des critères de sélection émis précédemment, vont nous permettre d'approvisionner régulièrement notre base d'image.

Le processus d'extraction d'images à partir du Web comporte une sélection des images et une indexation de leur contexte. La sélection suit un arbre de décision. L'indexation associe aux mots du contexte des termes de notre thésaurus.

La sélection d'images

La sélection d'une image est définie dans un arbre de décision [3]. Chaque nœud de l'arbre représente une question (un critère) permettant la sélection de l'image. Ici, l'arbre est binaire, car les réponses aux questions sont soit «oui », soit «non ». Si une image vérifie tous les critères dans l'ordre des nœuds parcourus, elle est jugée pertinente pour la recherche d'informations techniques. Voici la liste ordonnée des nœuds de l'arbre :

1. « La page où se situe l'image est-elle à une distance proche de la racine du site parcouru ? »
2. « La page où se situe l'image est-elle en français ? »
3. « L'image est-elle dans une page intéressante (présentation de catalogue) ? »
4. « La forme de l'image est-elle bonne ? »
5. « L'image a-t-elle un contexte ? »
6. « Le contenu du contexte est-il intéressant ? »

Ces critères ne couvrent actuellement qu'une partie des propriétés de pertinence énoncées plus haut. Au stade actuel de nos travaux, une validation humaine de la pertinence finale demeure encore nécessaire,.

Le résultat de l'extraction est une liste d'images associées à leurs contextes.

L'indexation

L'indexation procède à l'analyse des contextes extraits précédemment, afin de déterminer les termes du thesaurus qui feront partie de l'indexation de l'image. Le processus d'indexation s'appuie sur la technologie des n-grammes [4], il suit les étapes suivantes :

- À chaque terme du thesaurus est associée sa représentation en tri-grammes et bi-grammes ;
- Les contextes de chaque image sont analysés afin d'en extraire des groupes nominaux. À chacun de ces groupes nominaux est alors associée sa représentation en tri-grammes et bi-grammes ;
- Une fonction de mise en correspondance évalue la distance, à l'aide des représentations en tri-grammes, entre chaque groupe nominal contenu dans les contextes et les termes du thesaurus ;
- Un tri est alors effectué pour sélectionner les termes du thesaurus les plus pertinents.

On obtient ainsi pour chacune des images extraites une indexation contenant un vecteur pondéré de termes du thesaurus. Ces images et leur indexation vont être le support au processus de recherche d'images

La recherche interactive et progressive d'images

La recherche interactive et progressive d'images repose sur l'utilisation d'un bouclage de pertinence [5] composé de visualisation, choix, et analyse de choix. Il permet à l'utilisateur de mettre son besoin en correspondance avec l'information présentée par application d'un processus cognitif reposant sur une suite de raffinements successifs.

Le processus de recherche

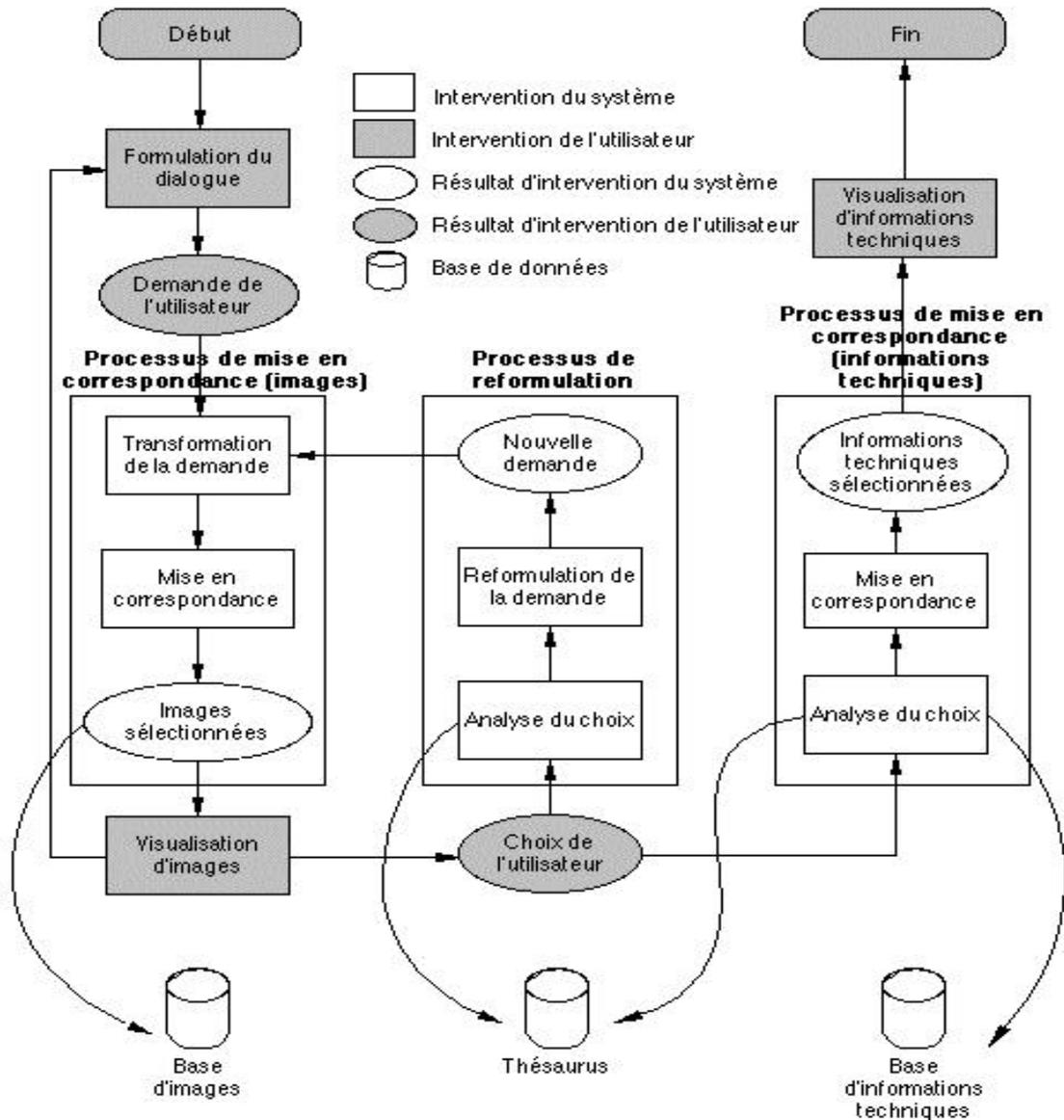
Le processus de recherche peut être décrit de manière suivante :

- L'utilisateur visualise les images. Ces images sont présentées sous forme d'imagettes à l'intérieur d'une mosaïque ;

- L'utilisateur donne son avis sur chacune des images qui lui sont présentées. Trois possibilités de choix lui sont proposées : choisir l'image, rejeter l'image, ou laisser un avis « indifférent » ;
- Après validation de son choix, l'utilisateur laisse le système analyser celui-ci ;
- Le système, grâce à l'analyse du choix de l'utilisateur, va sélectionner de nouvelles images qu'il va proposer de nouveau à l'utilisateur.

Ce processus continue jusqu'à ce que l'utilisateur juge que les images proposées sont toutes suffisamment pertinentes. Il demande alors au système de lui proposer les produits correspondant à ces images.

Schéma du processus de recherche



L'analyse des choix

L'analyse des choix repose sur un calcul d'un poids de pertinence [6] pour chacun des termes présents dans les images choisies et rejetées. Plus le terme est présent dans les indexations des images choisies plus son poids de pertinence sera proche de 1 et réciproquement, plus le terme est présent dans les indexations des images rejetées plus sa valeur sera proche de -1.

La propagation de ces poids dans le thesaurus, qui joue ici le rôle de la connaissance qu'a le système du domaine du bâtiment, permet une évaluation du besoin de l'utilisateur. Un parcours en largeur d'abord des liens « générique/spécifique » sélectionnant les concepts du thesaurus dont le poids est supérieur à un seuil, permet la formulation d'une nouvelle requête. Cette requête, tout comme les indexations, a la forme d'un vecteur pondéré de termes du thesaurus.

Le modèle vectoriel [5] peut alors être utilisé comme modèle de mise en correspondance entre la requête, ainsi obtenue, et les indexations des images de notre base. Cette mise en correspondance donne comme résultat une liste pondérée d'images qui peut être triée afin de ne présenter à l'utilisateur que les plus pertinentes (poids le plus fort).

Applications

Deux applications ont été réalisées à partir des méthodes présentées précédemment. La première application sert à approvisionner la base d'images et la deuxième utilise ces images pour rechercher des informations techniques.

Le robot d'extraction et d'indexation d'images à partir du Web.

La première application est un robot spécialisé dans l'extraction d'images de produits du bâtiment à partir du Web. Écrit en Java, il extrait et indexe les images de produit du bâtiment dans les pages Web en parcourant et analysant les documents HTML des sites des fabricants de produits.

Ce robot est composé de deux parties principales :

- L'extraction et l'indexation automatique,
- Le contrôle de pertinence morphologique.

Ces deux parties font de ce robot logiciel, une application semi-automatique. Le robot fonctionne de la manière suivante :

À partir d'un répertoire de sites Web de fabricants de produits, le robot analyse chacun des sites ainsi que tous les liens s'y trouvant référencés. Le calcul d'une distance par rapport à la racine du site détermine les pages à parcourir. Seules les pages susceptibles de présenter des produits du bâtiment sont analysées. Par exemple, les pages dont le nom contient des mots proches de « historique », « adresse », ou « usine », ... ne sont pas analysées, car il y a de forte chance qu'elles ne contiennent pas de présentation de produit. Puis, le robot extrait les images en respectant le processus de sélection et d'indexation défini précédemment.

À cette étape, l'intervention d'un administrateur semble nécessaire afin de contrôler la pertinence des images extraites. Cette intervention humaine permet de s'assurer que les images candidates vérifient les principes de pertinence sémiologique.

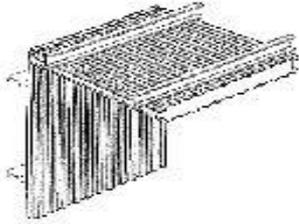
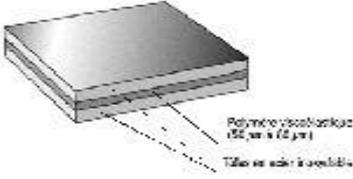
L'outil d'aide à la recherche d'informations techniques par l'image

Cette application utilise le processus de la recherche interactive et progressive d'images qui a été présenté dans cet article. Le processus a été mis en œuvre la manière suivante :

- A. Une première fenêtre demande à l'utilisateur soit de formuler une première demande en choisissant une fonction constructive, soit d'obtenir les premières images à partir d'un tirage aléatoire.
- B. Le premier ensemble d'images est présenté sous la forme d'une mosaïque d'image.
- C. L'utilisateur visualise ces images (cf. Figure 1) et donne son avis sur chacune d'elle : « oui », « non », « peut-être ».
- D. Après cette étape, l'utilisateur a la possibilité de continuer le processus en demandant au système de nouvelles images (E) ou de l'arrêter en demandant les produits correspondant à son choix (G).
- E. Le système analyse les choix afin de construire une nouvelle requête pour sélectionner de nouvelles images.
- F. Les images déjà choisies et les images les plus pertinentes du nouvel ensemble d'images sélectionné sont présentées à l'utilisateur (C).
- G. L'analyse des choix permet au système de construire une nouvelle requête pour sélectionner dans la base des produits ceux qui illustrent ce choix.

Interface de recherche par l'image

CONCLUSION

 <p> <input type="radio"/> oui <input type="radio"/> peut-être <input checked="" type="radio"/> non </p>	 <p> <input type="radio"/> oui <input type="radio"/> peut-être <input checked="" type="radio"/> non </p>	 <p> <input type="radio"/> oui <input checked="" type="radio"/> peut-être <input type="radio"/> non </p>
 <p> <input type="radio"/> oui <input type="radio"/> peut-être <input checked="" type="radio"/> non </p>	 <p> <input checked="" type="radio"/> oui <input type="radio"/> peut-être <input type="radio"/> non </p>	 <p> <input type="radio"/> oui <input type="radio"/> peut-être <input checked="" type="radio"/> non </p>
<input type="button" value="Images suivantes"/>		<input type="button" value="Rechercher"/>

Ces deux applications sont aujourd'hui utilisées à titre expérimental par le CRIT¹. Son utilisation pédagogique et professionnelle nous montre que l'image dans un processus de

¹ CRIT : Centre de Ressources et d'Informations Techniques. <http://www.crit.archi.fr>

recherche interactif et progressif permet à un acteur du bâtiment de confronter et de préciser son besoin en informations techniques en le projetant sur des représentations concrètes de produits.

Ce processus n'est réalisable que si les images sélectionnées sur le Web intègrent des critères de pertinence relevant de trois aspects (morphologiques, sémiologiques, contextuels) afin de permettre au dialogue entre le système et l'utilisateur d'être le plus cohérent possible et d'engendrer le moins d'effort d'interprétation.

Les critères morphologiques et contextuels peuvent être facilement automatisés mais les critères sémiologiques supposent encore une intervention humaine forte. On ne peut envisager aujourd'hui d'avancer dans un processus plus fortement automatisé que par l'utilisation des techniques d'analyse d'images (reconnaissance de forme, analyse de couleurs...). Cette nouvelle étape devrait permettre alors d'avancer dans la généralisation d'outils qui pourraient trouver d'autres domaines d'application comme l'aide à la conception architecturale dans les phases amonts des processus de création ou la veille technologique.

BIBLIOGRAPHIE

- [1] Sperber D., Wilson D. La pertinence: communication et cognition ED de minuit 1989
- [2] Reboul, H. and Moescler, J., La Pragmatique aujourd'hui. Une nouvelle science de la communication, Seuil, Paris (1998).
- [3] Mitchell, Tom M.: Decision Tree Learning, Machine Learning. McGraw-Hill, New York (1997), pp. 52-80.
- [4] Hallab, M., Lelu, A. : Proxilex : un outil d'approximation orthographe à partir des fréquences des n-grammes. Hypertextes hypermédiés et internet, 5e conférence internationale H2PTM'99, Paris (1999), pp. 201-209
- [5] van Rijsbergen, C.J.: Information Retrieval. 2nd edition Butterworths, London (1979).
- [6] Halin, G., Créhange, M., Kerekes P.: Machine learning and vectoriel matching for an image retrieval model: EXPRIM and the system RIVAGE. Proceedings of the ACM 13th International Conference on Research and Development in Information Retrieval, Brussels (1990), pp. 99-114.