

# Bouclage de pertinence par arbres de décision crédibilistes en indexation d'images par le contenu

P. VANNOORENBERGHE

Laboratoire Perception, Systèmes, Information, FRE 2645 CNRS  
Université de Rouen, Place Emile Blondel, 76821 Mont Saint-Aignan Cedex

Patrick.Vannoorenberghe@univ-rouen.fr

## Résumé

*En fouille et indexation d'images par le contenu, le bouclage de pertinence offre la possibilité à l'utilisateur d'affiner sa requête. Cette phase, qui permet d'améliorer l'adaptabilité du système aux besoins exprimés, peut être réalisée par un classifieur dont l'algorithme d'apprentissage se base sur le marquage par l'utilisateur des images perceptuellement similaires à l'image requête. Dans cet article, nous plaçons pour l'utilisation d'un arbre de décision basé sur les fonctions de croyance qui permet de s'adapter à ce problème d'apprentissage délicat. Cet arbre, remis à jour de façon itérative dans le bouclage, permet ensuite d'évaluer la pertinence de chacune des images de la base. La méthode est appliquée à l'indexation d'une base d'images couleur généraliste ANN composée d'environ de 500 images partitionnée en 11 catégories. Les performances du système sont mesurées à l'aide de courbes de précision/rappel sur une catégorie recherchée.*

## Mots clefs

Indexation, Bouclage de pertinence, Apprentissage, Arbres de décision crédibilistes, Fonctions de croyance.

## 1 Introduction

La création et le développement de bases d'images volumineuses ont été favorisés par l'essor des dispositifs d'acquisition, par les capacités croissantes de stockage et les systèmes de transmission et de diffusion de l'information. De manière à utiliser ces bases, des techniques fiables et efficaces capables de retrouver une information dans la masse de données doivent être développées. Dans cet article, nous nous intéressons aux systèmes d'indexation d'images basés sur le contenu (Content Based Image Retrieval). Dans ce contexte, il s'agit d'extraire un ensemble de paramètres caractéristiques des images de la base pour constituer des signatures qui serviront d'index pour la recherche d'images similaires à une image requête présentée au système d'indexation par l'utilisateur [1]. Ces paramètres sont des caractéristiques de bas niveau extraites de l'image comme par exemple la texture, la couleur [2]. Il peut exister ainsi un fossé sémantique entre certains concepts de haut niveau exprimés par l'utilisateur via l'image requête et les caractéristiques bas niveau calculables sur l'image par l'ordinateur.

De plus, certaines caractéristiques peuvent avoir une importance variable selon l'image requête ou les buts spécifiques de la recherche exprimée par l'utilisateur. Il est alors utile de sélectionner les caractéristiques pertinentes au sens de la requête ou de les pondérer entre elles [3]. Le système doit également offrir la possibilité d'affiner la requête par un mécanisme appelé le bouclage de pertinence. Les différentes actions de l'utilisateur permettent alors au système d'estimer de manière interactive la similarité perceptuelle qui lui est propre, c'est-à-dire de l'améliorer pour mieux répondre à ses attentes [4]. Dans ce contexte, une phase d'apprentissage est souhaitable et les techniques de reconnaissance de formes prennent ici une place importante [5, 6, 7]. Une des principales difficultés de la phase d'apprentissage en bouclage de pertinence provient du faible nombre d'exemples étiquetés [8]. Bien souvent, ce nombre est inférieur à la dimension de l'espace de représentation. Dans [9], les auteurs proposent d'indexer une base d'images médicales à l'aide d'arbres de décision. L'utilisation d'un arbre de décision dans un système de fouille d'images possède deux avantages. Il permet tout d'abord, via une représentation graphique de la règle de décision, d'obtenir une parfaite compréhension de la boucle de pertinence en phase d'apprentissage. Le second avantage provient du fait que l'algorithme d'induction de l'arbre a la capacité d'exploiter un nombre restreint de caractéristiques. Il permet donc de partitionner de plus en plus précisément l'espace des caractéristiques et ainsi d'améliorer les performances en terme de discrimination des images pertinentes. Dans cet article, nous proposons un système de fouille d'images [10] basé sur les arbres de décision crédibilistes [11], formalisme dérivé de la théorie des fonctions de croyances [12].

## 2 Apprentissage par arbres de décision crédibilistes

Les approches basées sur l'induction d'arbres de décision sont devenues de plus en plus populaires dans la communauté de l'intelligence artificielle pour résoudre le problème classique de la discrimination. Celui-ci consiste à assigner un vecteur d'entrée  $\mathbf{x}$  à une classe dans un ensemble fini de  $K$  classes  $\Omega = \{\omega_1, \dots, \omega_k, \dots, \omega_K\}$  étant donné la connaissance d'un ensemble d'apprentissage

$\mathcal{L} = \{(\mathbf{x}_i, \omega_i), i = 1, \dots, N\}$  composé de  $N$  formes  $\mathbf{x}_i$  de classes connues. Chaque exemple dans  $\mathcal{L}$  est représenté par un vecteur de dimension  $p$  et sa variable de classe  $\omega_i$  correspondante. L'induction d'un arbre de décision consiste à déduire de  $\mathcal{L}$  un graphe direct et acyclique dans lequel chaque nœud est soit un nœud décision soit une feuille. On associe à chaque nœud décision un test basé sur la valeur des attributs (les composantes de  $\mathbf{x}$ ). Pour les arbres de décision binaires, chaque nœud possède deux successeurs suivant le résultat du test. Ces arbres de décision donnent la possibilité d'interpréter chaque règle en terme d'attribut individuel et ainsi de faciliter la compréhension du processus d'apprentissage.

## 2.1 Apprentissage partiellement supervisé

Récemment, une approche dérivée de ce mécanisme a été proposée dans [13, 11]. Celle-ci est basée sur les fonctions de croyance qui sont réputées pour la gestion d'informations incertaines. Une fonction de croyance *bel* peut être définie mathématiquement par une fonction (ou allocation) de masse, notée  $m$  définie de  $2^\Omega$  dans  $[0, 1]$ , qui vérifie :

$$\sum_{A \subseteq \Omega} m(A) = 1. \quad (1)$$

Ainsi, la masse  $m(A)$  représente le degré de croyance attribué à la proposition  $A$  et qui n'a pas pu, compte tenu de l'état de la connaissance, être affecté à un sous-ensemble plus spécifique que  $A$ . Chaque sous-ensemble  $A \subseteq \Omega$  tel que  $m(A) > 0$  est appelé élément focal de  $m$ .

Par rapport aux arbres classiques, les arbres de décision crédibilistes (Belief Decisions Trees ou BDT) permettent de gérer des ensembles d'apprentissage de la forme :

$$\mathcal{L} = \{(\mathbf{x}_i, m_i^\Omega), i = 1, \dots, N\} \quad (2)$$

où  $m_i^\Omega$  permet d'encoder l'étiquetage de l'exemple  $i$  qui peut être quelconque<sup>1</sup>. Dans [13, 11], la méthode proposée construit pour chaque nœud de l'arbre une fonction de croyance. L'induction de l'arbre de décision suit une stratégie descendante (top-down) pour diviser les nœuds en se basant sur une mesure d'impureté calculée à partir des exemples qui atteignent chaque nœud de l'arbre. Le choix de la sélection des variables est basé sur l'examen de cette mesure calculée à partir d'un critère entropique d'incertitude totale. Lorsque l'arbre classe tous les exemples de l'ensemble d'apprentissage  $\mathcal{L}$ , la procédure est stoppée. Dans [13, 11], l'induction de l'arbre intègre une stratégie d'élagage qui permet de limiter la complexité de la règle de décision et d'éviter le sur-apprentissage. L'arbre est ensuite utilisé pour classer de nouveaux vecteurs. Ainsi, à chaque vecteur  $\mathbf{x}$  est associé une fonction de masse  $m_{\mathbf{x}}^\Omega[\mathcal{L}]$  qui permet de prendre une décision quant à l'appartenance de  $\mathbf{x}$  dans  $\Omega$ . La notation  $m_{\mathbf{x}}^\Omega[\mathcal{L}](A)$  est utilisée pour représenter le degré de croyance attribué au sous-ensemble  $A \subseteq \Omega$ .

<sup>1</sup>La fonction de masse  $m_i^\Omega$  permet de modéliser une connaissance partielle sur l'étiquette du vecteur  $\mathbf{x}_i$ .

En reconnaissance des formes, il quantifie le degré d'appartenance de  $\mathbf{x}$  à  $A \subseteq \Omega$  estimé à partir de l'observation des données d'apprentissage contenues dans  $\mathcal{L}$ .

## 2.2 Co-Apprentissage

Dans certaines applications notamment pour les systèmes d'indexation d'images [5], un grand nombre d'échantillons non étiquetés peuvent être disponibles tandis que les exemples étiquetés sont parfois proportionnellement en nombre limité. Ceci peut être dû au fait qu'il est parfois difficile ou voire même impossible d'en obtenir sous certaines circonstances ou que l'étiquetage requiert une connaissance experte coûteuse. Pour palier ce problème, plusieurs auteurs ont proposées différentes solutions de façon à gérer ce type d'informations partielles. Dans cette partie, on suppose que l'ensemble d'apprentissage disponible  $\mathcal{L}$  est composé d'un ensemble d'apprentissage partiellement étiqueté noté  $\mathcal{S}$  et d'un ensemble d'apprentissage non étiqueté noté  $\mathcal{U}$  défini par :

$$\mathcal{U} = \{(\mathbf{x}_u, m_u^\Omega), u = 1, \dots, U\} \quad (3)$$

avec  $m_u^\Omega$  la fonction de masse vide qui permet de représenter l'ignorance totale ( $m_u^\Omega(\Omega) = 1$ ). La disponibilité de l'ensemble  $\mathcal{U}$  pose le challenge de son utilisation de manière à augmenter les performances en généralisation d'un classifieur qui ne serait appris que sur l'ensemble  $\mathcal{S}$ . Dans [14], on propose de ré-étiqueter l'ensemble  $\mathcal{U}$  de façon à faire l'apprentissage sur l'ensemble complet  $\mathcal{L} = \mathcal{S} + \mathcal{U}$ . La méthodologie proposée peut se décomposer en deux étapes. A partir des données contenues dans  $\mathcal{S}$ , un premier classifieur crédibiliste est construit à partir des données d'apprentissage contenues dans  $\mathcal{S}$ . La règle de décision ainsi construite est alors utilisée pour estimer pour chaque vecteur  $\mathbf{x}_u$  de l'ensemble  $\mathcal{U}$  une fonction de masse, notée  $\hat{m}_{\mathbf{x}_u}^\Omega[\mathcal{S}]$ . Dans [14], la règle utilisée est l'extension aux fonctions de croyance de la règle des plus proches voisins [15]. L'ensemble  $\mathcal{U}$  ainsi constitué peut être utilisé conjointement à l'ensemble  $\mathcal{S}$  pour construire la règle de décision finale. Les données d'apprentissage disponibles rassemblent désormais les exemples de  $\mathcal{S}$  préalablement étiquetés par l'expert et les données de  $\mathcal{U}$  dont l'appartenance est estimée par la première règle de décision. Dans cet article, un arbre de décision crédibiliste est utilisé pour construire la règle finale.

## 3 Méthodologie proposée

Dans un système de fouille basé sur le contenu, on suppose que la base à indexer est composée de  $N$  images où chaque image  $\mathcal{I}$  est représentée par un vecteur  $\mathbf{x}_i$  de dimension  $p$  qui permet d'obtenir une signature caractérisant l'image en terme de contenu géométrique et/ou photométrique. Puisqu'aucune information n'est disponible dans la phase de requête à la première itération, on propose d'utiliser les plus proches voisins de  $\mathbf{x}$  dans l'ensemble des vecteurs de la base d'images. Cet algorithme permet de quantifier la pertinence de chacune des  $N$  images de la base lors

$m_i^\Omega(A)$	IP	INE	EP	EI
$\{\omega_1\}$	1	0	0.6	0.8
$\{\omega_2\}$	0	0	0.4	0
$\Omega$	0	1	0	0.2

Tableau 1 – *Étiquetage de pertinence (IP : Image Pertinente, INE : Image Non Étiquetée, EP : Étiquetage Probabiliste, EI : Étiquetage Incertain).*

de la première itération en fonction de la distance entre  $x$  (l'image requête) et les vecteurs  $x_i$  (les images de la base) dans l'espace des caractéristiques. Pour les itérations suivantes, la phase de recherche se scinde en deux étapes qui sont respectivement : l'étiquetage des images pertinentes par l'utilisateur et le bouclage de pertinence.

### 3.1 Étiquetage de pertinence

A chaque itération de la boucle, l'utilisateur ne dispose que d'un nombre limité d'images pour l'étiquetage de pertinence. Le formalisme des fonctions de croyance permet d'envisager un étiquetage des images pertinentes qui peut revêtir plusieurs formes (ordre de préférence, étiquetage possibiliste, probabiliste, certain, ...). Pour obtenir un maximum de flexibilité, on propose de quantifier la pertinence de chacune des images retournées par le système par une fonction de croyance, connue pour la modélisation de données imprécises et incertaines. Dans la phase d'apprentissage, on distingue la classe des images pertinentes, notée  $\omega_1$  et la classe des images non pertinentes, notée  $\omega_2$ . Un exemple de ce type d'étiquetage est proposé au tableau 1 où l'utilisateur peut choisir entre différents types d'étiquetage. Cette fonction, notée  $m_i^\Omega$ , quantifie la pertinence de l'image  $i$  vis-à-vis de l'image requête.

### 3.2 Bouclage par apprentissage d'un BDT

La phase de bouclage consiste à faire l'apprentissage d'un arbre de décision crédibiliste à partir des images étiquetées par l'utilisateur. L'arbre peut être ensuite utilisé pour classer de nouvelles images. A chaque feuille d'un arbre crédibiliste est associée une fonction de croyance permettant de quantifier l'appartenance d'un vecteur inconnu dans  $\Omega$ , l'espace de pertinence des images. Le faible nombre d'images  $n$  ne nous permettant pas d'obtenir de bonnes performances en généralisation, l'idée consiste à utiliser les images non étiquetées pour la phase d'apprentissage. Afin de prendre en compte ces images, on utilise un mécanisme de co-apprentissage qui consiste à réétiqueter les images non annotées par une règle de décision. Pour cette tâche, l'algorithme des plus proches voisins est utilisé. Ainsi, l'arbre de décision est appris à partir des données de  $\mathcal{L}$  constitué de  $\mathcal{S}$  lui-même et de l'ensemble  $\mathcal{U}$  ré-étiqueté. Nous noterons  $\hat{m}_{x_i}^\Omega[\mathcal{L}]$  les fonctions de masse estimées par l'arbre de décision à partir des vecteurs  $x_i$  de la base. Pour chacune des  $N$  images de la base, on calcule la fonction de probabilité pignistique associée qui permet d'ordonner les

images présentées à l'utilisateur à la prochaine itération.

## 4 Résultats

Nous avons testé le système dans le cadre d'une base d'images généraliste ANN composée d'environ de 500 images partitionnée en 11 catégories de 25 à 50 images. Quelques images de cette base sont présentées à la figure 1. Des attributs (Couleurs Lab et Filtres de Gabor) sont calculés de façon à indexer la base. Quelques images de la base sont présentées à la figure 1.

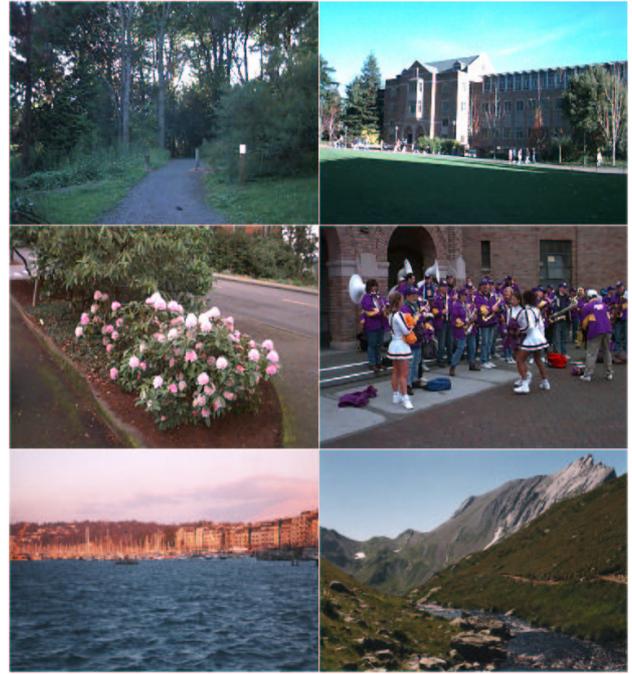


Figure 1 – *Images de la base généraliste ANN comportant 500 images couleur classées en 11 catégories*

Les performances du système sont mesurées à l'aide de courbes de précision/rappel sur une catégorie recherchée comme celle présentée à la figure 2.

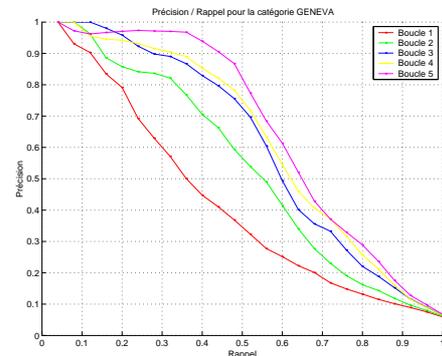


Figure 2 – *Rappel / Précision de la catégorie 'Geneva'*

Sur cette figure, les 5 courbes correspondent aux performances précision/rappel pour la catégorie 'Geneva' pour

1,2,3,4 et 5 bouclages en annotant 5 images à chaque itération. Comme attendu, le bouclage de pertinence permet d'améliorer, pour une valeur de rappel donné, la précision du système. La figure 3 illustre quelques images de la catégorie 'Cherries' pour lesquelles les performances sont satisfaisantes.



Figure 3 – Quelques images illustrant la diversité au sein de la catégorie 'Cherries'

Pour les catégories à tendance unimodale dans l'espace des caractéristiques, la précision du BDT est toutefois moins bonne qu'attendue.

## 5 Conclusion

De part sa conception et son formalisme, la méthodologie présentée possède plusieurs avantages. Un des avantages provient de l'utilisation d'arbres de décision dans la phase d'apprentissage qui permet d'interpréter les résultats de la boucle de pertinence et de faciliter la compréhension du mécanisme par l'utilisateur. L'algorithme d'apprentissage de l'arbre permet la sélection automatique des caractéristiques importantes. Les fonctions de croyance quant à elles offrent la possibilité pour l'utilisateur d'étiqueter les images pertinentes sous différentes formes (ordre de préférence, étiquetage possibiliste, ...). Le co-apprentissage permet de prendre en compte les images non étiquetées de la base pour améliorer les performances en terme d'apprentissage de la boucle de pertinence.

## Références

- [1] W.M. Smeulders, M. Worring, S. Santini, A. Gupta, et R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12) :1349–1380, December 2000.
- [2] R. Schettini, G. Ciocca, et S. Zuffi. *A Survey on Methods for Colour Image Indexing and Retrieval in Image Databases*. J. Wiley, R. Luo, L. MacDonald, 2001.
- [3] B. Ko, J. Peng, et H. Byun. Region-based image retrieval using probabilistic feature relevance learning. *Pattern Analysis and Applications*, 4(2/3) :174–184, 2001.
- [4] C. Shyu, C. Brodley, A. Kak, A. Kosaka, A. Aisen, et L. Broderick. ASSERT, a physician-in-the-loop content-based image retrieval system for HRCT image databases. *Computer Vision and Image Understanding*, 75(1/2) :111–132, August 1999.
- [5] T.S. Huang, X.S. Zhou, M. Nakazato, Y. Wu, et I. Cohen. Learning in content-based image retrieval. Dans *2nd IEEE International Conference on Development and Learning, ICDL'02*, Cambridge, Massachusetts, 2002.
- [6] S. Antani, R. Kasturi, et R. Jain. A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video. *Pattern recognition*, 35 :945–965, 2002.
- [7] J. Peng. Multi-class relevance feedback content-based image retrieval. *Computer Vision and Image Understanding*, 90 :42–67, 2003.
- [8] X.S. Zhou et T.S. Huang. Relevance feedback for image retrieval : A comprehensive review. *Multimedia Systems*, 8(6) :536–544, 2003.
- [9] S.D. MacArthur, C.E. Brodley, et C.R. Shyu. Relevance feedback decision trees in content-based image retrieval. Dans *IEEE Workshop on Content-Based Access of Image and Video Libraries, CBAIVL'00*, pages 68–72, South California, USA, 2000.
- [10] P. Vannoorenberghe. Fouille d'images par Arbres de Décision Crédibilistes. Dans *Colloque RFIA'2004, Atelier Analyse de données, Statistique et Apprentissage pour la Fouille d'Images*, pages 55–60, Toulouse, 2004.
- [11] P. Vannoorenberghe et T. Dencœur. Handling uncertain labels in multiclass problems using belief decision trees. Dans *Proceedings of IPMU'2002*, pages 1919–1926, Annecy, France, 2002.
- [12] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [13] T. Dencœur et M. Skarstein Bjanger. Induction of decision trees from partially classified data using belief functions. Dans *Proceedings of SMC'2000*, pages 2923–2928, Nashville, USA, 2000. IEEE.
- [14] P. Vannoorenberghe. Reasoning with unlabeled samples and belief functions. Dans *Proceedings of the International Conference on Fuzzy Systems FUZZ-IEEE'2003*, Saint-Louis, Missouri, USA, May 2003.
- [15] T. Dencœur. A neural network classifier based on Dempster-Shafer theory. *IEEE Transactions on Systems, Man and Cybernetics, Part A : Systems and humans*, 30(2) :131–150, 2000.