

Résumé de la thèse de Sid-Ahmed Berrani

Ce travail de thèse s'intéresse aux systèmes de recherche d'images par le contenu (SRIC). Ces systèmes permettent de retrouver au sein d'une base d'images les images les plus similaires *visuellement* à une image requête donnée comme exemple.

Pour cela, ils utilisent des méthodes d'analyse d'images pour décrire *automatiquement* le contenu visuel des images. Il s'agit d'extraire à partir des images des vecteurs numériques de grande dimension appelés descripteurs et de leur associer une mesure de similarité. L'objectif est de ramener la notion de similarité visuelle entre images à une simple notion de proximité entre descripteurs. Retrouver les images similaires à une image requête revient ainsi à effectuer une recherche des plus proches voisins (PPVs) du descripteur de l'image requête dans l'espace de description. Les PPVs permettent de désigner les images les plus similaires à l'image requête.

Lorsque la base d'images est très grande et qu'elle ne tient plus en mémoire centrale, la recherche de PPVs devient extrêmement coûteuse. Elle nécessite alors l'utilisation de techniques d'indexation multidimensionnelle afin de réduire le temps de réponse.

Malheureusement, bien que de nombreuses techniques d'indexation ont été développées dans le domaine des bases de données, elles sont rarement utilisées dans les SRICs. La raison de cela est simple : elles sont complètement inefficaces face à des données réelles de grande dimension. Plusieurs travaux ont montré que leurs performances se dégradent exponentiellement lorsque la dimension des descripteurs augmente (phénomène de la « malédiction de la dimension »). Au delà d'un certain seuil (16 selon Weber *et al.*), leurs temps de réponse deviennent supérieurs à ceux d'une simple recherche séquentielle et exhaustive. Celle-ci est ainsi la meilleure technique en grande dimension. Elle est cependant elle-même extrêmement coûteuse lorsque la taille de la base est très grande.

Les causes de la dégradation des performances des techniques d'indexation sont multiples. D'abord, ces techniques reposent généralement sur des hypothèses irréalistes quant à la distribution des descripteurs (e.g. distribution uniforme) et sont souvent validées sur des descripteurs d'images rudimentaires, voire parfois des données générées aléatoirement. Ensuite, le problème de la grande dimension a été longtemps minimisé. On considérait que la dimension peut être facilement réduite par des méthodes d'analyse de données (e.g. ACP). Or, cette réduction entraîne une baisse significative de la qualité des résultats.

Ce travail de thèse a été initié dans le but de développer une méthode *efficace* de recherche de PPVs qui soit adaptée aux applications d'indexation par le contenu et aux propriétés des descripteurs d'images (grand volume, grande dimension, aucune hypothèse sur la distribution des descripteurs...). Il se positionne à la frontière entre le domaine des bases de données et celui de l'analyse d'images.

La première contribution de ce travail est une étude approfondie des propriétés des espaces de grande dimension et de leur impact sur les performances des techniques de recherche de PPVs. Cette analyse a permis notamment d'expliquer les causes de la dégradation des performances des techniques traditionnelles d'indexation multidimensionnelle lorsque la dimension augmente. Elle a permis également d'analyser les méthodes de réduction de la dimension et leur application à la recherche de PPVs. Cette étude a montré que l'utilisation de ces méthodes –pourtant souvent

présentées comme une solution pour pallier les difficultés liées aux grandes dimensions – n’est pas triviale et qu’elle est sujette à discussion. En particulier, l’imprécision des résultats de la recherche de PPVs induite par une réduction de la dimension par ACP n’est pas proportionnelle à la diminution de l’inertie des données, elle est donc incontrôlable. Une des expériences a montré que cette imprécision peut être de 45 % lorsque la proportion de l’inertie exprimée est supérieure à 90 %.

La seconde et principale contribution est la conception d’une méthode efficace de recherche de k -PPVs. Cette méthode s’appuie sur des travaux récents qui ont montré que le problème de la malédiction de la dimension est particulièrement manifeste lorsqu’il s’agit de retrouver les PPVs « exacts » du descripteur requête. Ces travaux suggèrent de faire un compromis entre la précision de la recherche et le temps de réponse. La méthode proposée permet ainsi d’accélérer considérablement le temps de recherche en échange de petites imprécisions dans les résultats. L’originalité de la méthode réside dans le fait que cette imprécision peut être contrôlée de façon intuitive, en ligne, et à l’aide d’un seul paramètre. Les méthodes de recherche approximative existantes ne permettent, pour certaines, aucun contrôle de la précision, et, pour les autres, ce contrôle est mis en œuvre au travers de nombreux paramètres difficiles à manipuler par un utilisateur non spécialisé. De plus, nombre de ces méthodes ne permettent de rechercher que *le* PPV.

Dans la méthode proposée, le contrôle de la précision s’effectue de façon probabiliste à l’aide d’un seul paramètre appelé le *niveau d’imprécision*. Il s’agit de la probabilité maximale de ne pas retrouver un des k PPVs recherchés. Ce paramètre est intuitif et peut être fixé même en n’ayant aucune connaissance sur le fonctionnement interne de la méthode. Il est de plus fixé en ligne, ce qui permet à l’utilisateur d’adapter la recherche en fonction de ses contraintes. Il peut, par exemple, choisir d’effectuer tout d’abord une recherche grossière mais très rapide en fixant un niveau d’imprécision élevé. Dans le cas où les résultats retournés ne sont pas satisfaisants, il peut alors relancer la recherche avec un niveau d’imprécision plus faible.

Les expérimentations réalisées sur des descripteurs d’images de dimension allant jusqu’à 150 ont notamment montré que la méthode est fiable et qu’elle permet d’obtenir des temps de réponse jusqu’à 15 fois inférieurs à ceux de la recherche séquentielle.

Un autre mode de recherche dit « incrémentiel » a été également proposé. Il s’agit d’une recherche interactive affichant en temps réel les résultats courants tout en fournissant également une indication sur la probabilité de leur amélioration future. L’utilisateur peut alors interrompre la recherche dès que les résultats courants sont satisfaisants.

La troisième et dernière contribution de ce travail est l’application de la méthode de recherche approximative de PPVs proposée à la recherche d’images par le contenu. L’objectif est d’évaluer rigoureusement la méthode de recherche dans le cadre d’une application réelle. Le choix s’est porté sur une application de recherche d’images pour la détection de copies. Dans cette application, de fortes contraintes sont imposées à la fois sur les descripteurs d’images car les images piratées sont le plus souvent retouchées, et sur l’efficacité de la recherche car le système doit également pouvoir traiter un très grand nombre d’images récupérées par exemple du Web et de les confronter aux images de la base. Pour la description des images, les invariants locaux différentiels ont été choisis pour leurs propriétés d’invariance et leur robustesse.

Une étude théorique de la précision a d’abord été réalisée. Elle a montré que l’imprécision de la recherche de PPVs n’a qu’un impact mineur sur les résultats finals. Cette étude a été ensuite corroborée par une évaluation expérimentale étendue qui a montré de nouveau que la recherche approximative de PPVs permet de réduire significativement le temps de recherche par rapport à la recherche séquentielle.