

SCALE-INVARIANT FEATURE TRANSFORM

Scale-invariant feature transform (SIFT), que l'on peut traduire par « transformation de caractéristiques visuelles invariante à l'échelle », est un algorithme utilisé dans le domaine de la [vision par ordinateur](#) pour détecter et identifier les éléments similaires entre différentes [images numériques](#) (éléments de paysages, objets, personnes, etc.). Il a été développé en 1999 par le chercheur [David Lowe](#).

L'étape fondamentale de la méthode proposée par Lowe consiste à calculer ce que l'on appelle les « **descripteurs SIFT** » des images à étudier. Il s'agit d'informations numériques dérivées de l'analyse locale d'une image et qui caractérisent le contenu visuel de cette image de la façon la plus indépendante possible de l'échelle (« zoom » et résolution du capteur), du cadrage, de l'angle d'observation et de l'exposition (luminosité). Ainsi, deux photographies de la tour Eiffel auront toutes les chances d'avoir des descripteurs SIFT similaires, et ceci d'autant plus si les instants de prise de vue et les angles de vue sont proches. D'un autre côté, deux photographies de sujets très différents produiront selon toute vraisemblance des descripteurs SIFT très différents eux aussi (pouvoir discriminant). Cette robustesse, vérifiée dans la pratique, est une exigence fondamentale de la plupart des applications et explique en grande partie la popularité de la méthode SIFT.

Les applications de la méthode sont nombreuses et ne cessent de s'étendre ; elles couvrent au début du XXI^e siècle des domaines tels que la [détection d'objet](#), la [cartographie](#) et la navigation, l'[assemblage de photos](#), la [modélisation 3D](#), la [recherche d'image par le contenu](#), le [tracking video](#) ou le [match moving](#).

Cet algorithme est protégé aux [États-Unis](#) par un [brevet](#) détenu par l'[université de la Colombie-Britannique](#).



Exemple de résultat de la comparaison de deux images par la méthode SIFT (*Fantasia ou Jeu de la poudre*, devant la porte d'entrée de la ville de Méquinez, par [Eugène Delacroix](#), 1832).

ORIGINES DE LA MÉTHODE

C'est en 1999 que le chercheur [canadien](#) David G. Lowe, professeur à l'[université de la Colombie-Britannique](#), présente lors de la [conférence ICCV](#) une nouvelle méthode d'[extraction de caractéristiques](#) et de détection d'objets dans des [images en niveaux de gris](#). Le nom de *Scale-invariant feature transform* (SIFT) a été choisi car la méthode transforme les données d'une image en coordonnées [invariantes à l'échelle et rapportées à des caractéristiques locales](#). Des améliorations et une description précise de la méthode sont proposées par Lowe en 2004 dans la revue [International Journal of Computer Vision](#). L'Université de la Colombie-Britannique a demandé et obtenu la protection par brevet de cet algorithme aux [États-Unis](#).



Exemple de caractéristiques SIFT. Chaque point-clé est représenté par une flèche dont la direction correspond à la direction principale associée et la norme au facteur d'échelle (illustration : *Fantasia*, par Eugène Delacroix).

L'idée d'utiliser des points d'intérêt locaux remonte aux travaux de [Hans Moravec](#) en 1981 sur la recherche de correspondance entre images stéréoscopiques et aux améliorations apportées en 1988 par Harris et Stephens. À la suite de quoi, différents types de caractéristiques sont proposés, tels que des segments de lignes, des groupements d'arêtes ou de zones. En 1992, l'intérêt de ce type de détecteur est confirmé par les travaux de Harris et le descripteur de coins qu'il propose, améliorant la plupart des défauts du détecteur de Moravec, va connaître un important succès. Il faut cependant attendre 1997 et le travail précurseur de Cordelia Schmid et [Roger Mohr](#) pour établir l'importance, dans le domaine de la vision, des caractéristiques locales invariantes appliquées au problème général de détection et de recherche de correspondance. Si le descripteur qu'ils développent (et qui s'appuie sur le détecteur de coins de Harris) apporte l'invariance à la rotation, il reste cependant sensible aux changements d'échelle, d'angle d'observation et d'exposition. Lowe comblera grandement ces défauts avec son descripteur SIFT.

GÉNÉRALITÉS

La méthode proposée par Lowe comprend deux parties (chacune ayant fait l'objet de recherches plus ou moins indépendantes par la suite) :

- un algorithme de **détection de caractéristiques** et de **calcul de descripteurs** ;
- un algorithme de mise en correspondance proprement dit.

De ces deux aspects, le premier est sans doute celui qui a le plus assuré la popularité de la méthode, à tel point que le sigle SIFT fait plus souvent référence aux « descripteurs SIFT » qu'à la méthodologie globale. Il s'agit tout d'abord de détecter sur l'image des zones circulaires « intéressantes », centrées autour d'un *point-clé* et de rayon déterminé appelé *facteur d'échelle*. Celles-ci sont caractérisées par leur unité visuelle et correspondent en général à des éléments distincts sur l'image. Sur chacune d'elles, on détermine **une orientation intrinsèque qui sert de base à la construction d'un histogramme des orientations locales des contours, habilement pondéré, seuillé et normalisé pour plus de stabilité.** C'est cet histogramme qui sous la forme d'**un vecteur à 128 dimensions** (ou valeurs) constitue le descripteur SIFT du point-clé, et l'ensemble des descripteurs d'une image établissent ainsi une véritable signature numérique du contenu de celle-ci.

Ces descripteurs présentent l'avantage d'être invariants à l'orientation et à la résolution de l'image, et peu sensibles à son exposition, à sa netteté ainsi qu'au point de vue 3D. Ils possèdent ainsi des **propriétés similaires à celles des neurones du gyrus temporal inférieur utilisées en détection d'objet en vision primaire.** Par exemple, ils décriront de façon très semblable les détails d'une image originale et ceux d'une image retouchée par l'application d'une rotation, d'un recadrage ou d'un lissage, par une correction de l'exposition, par occultation partielle ou encore par l'insertion dans une image plus grande.

Une fois le calcul des descripteurs effectué, l'algorithme de mise en correspondance intervient dans le but de rechercher les zones d'une image dite suspecte qui contiennent des éléments visuellement similaires à ceux d'une bibliothèque d'images de référence, c'est-à-dire des descripteurs numériquement proches. Ceci ne peut fonctionner que parce que les descripteurs SIFT sont à la fois **robustes (aux principales transformations affines et aux changements d'exposition ou de perspective)** et discriminants (deux objets différents ont une grande probabilité d'avoir des descripteurs différents).

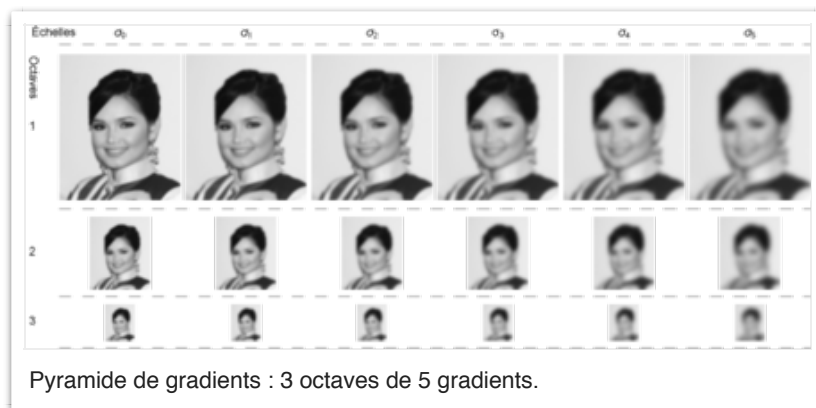
DÉTECTION DES POINTS-CLÉS ET CALCUL DU DESCRIPTEUR SIFT

La première étape de l'algorithme est la **détection des points d'intérêt, dits *points-clés***. Un point-clé (x, y, σ) est défini d'une part par ses coordonnées sur l'image (x et y) et d'autre part par son facteur d'échelle caractéristique (σ). En toute rigueur, il s'agit d'une **zone d'intérêt** circulaire, le rayon de la zone étant proportionnel au facteur d'échelle. Il s'ensuit une étape de reconvergence et de filtrage qui permet d'améliorer la précision sur la localisation des points-clés et d'en éliminer un certain nombre jugés non pertinents. Chaque point-clé restant est ensuite associé à une orientation intrinsèque, c'est-à-dire ne dépendant que du contenu local de l'image autour du point clé, au facteur d'échelle considéré. Elle permet d'assurer l'invariance de la méthode à la rotation et est utilisée comme référence dans le calcul du descripteur, qui constitue la dernière étape de ce processus.

DÉTECTION D'EXTREMUMS DANS L'ESPACE DES ÉCHELLES

La détection s'effectue dans un **espace discret** que l'on appelle *espace des échelles* (*scale space*) qui comporte trois dimensions : les coordonnées cartésiennes x et y et le facteur d'échelle σ . On appelle *gradient* de facteur d'échelle σ (noté L) le **résultat de la convolution d'une image I par un filtre gaussien G de paramètre σ** , soit :

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$



Cette convolution a pour effet de lisser l'image originale I de telle sorte que les détails trop petits, c'est-à-dire de rayon inférieur à σ , sont estompés. Par consé-

quent, la détection des objets de dimension approximativement égale à σ se fait en étudiant l'image appelée *différences de gaussiennes* (en anglais *difference of gaussians*, DoG) définie comme suit :

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma),$$

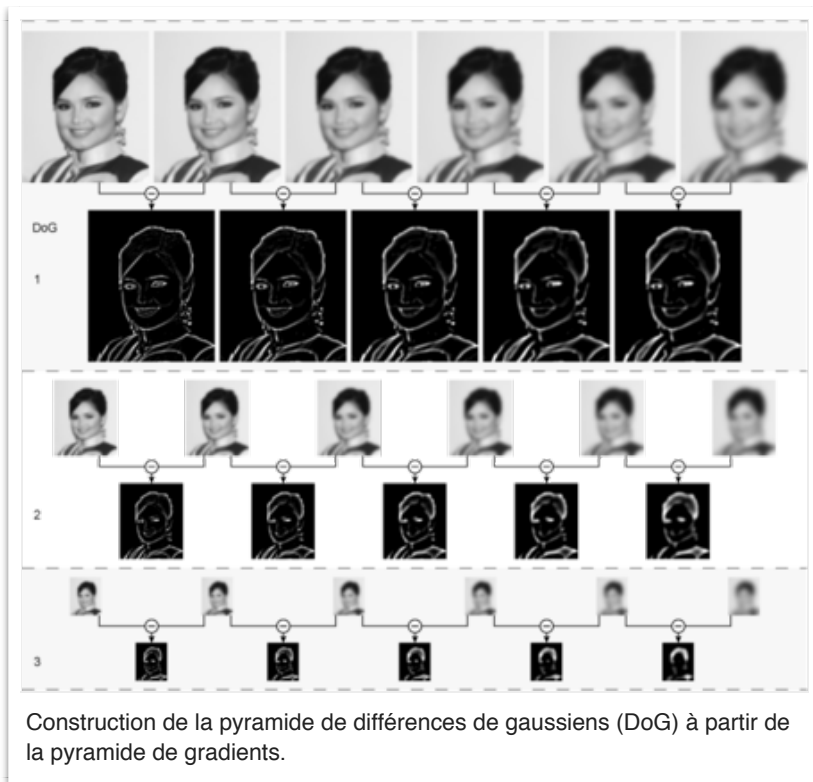
où k est un paramètre fixe de l'algorithme qui dépend de la finesse de la discrétisation de l'espace des échelles voulue.

Dans cette image ne persistent plus que les objets observables dans des **facteurs d'échelle qui varient entre σ et $k\sigma$** . De ce fait, un point-clé candidat (x, y, σ) est défini comme un point où un extremum du DoG est atteint par rapport à ses voisins immédiats, c'est-à-dire sur l'ensemble contenant 26 autres points défini par :

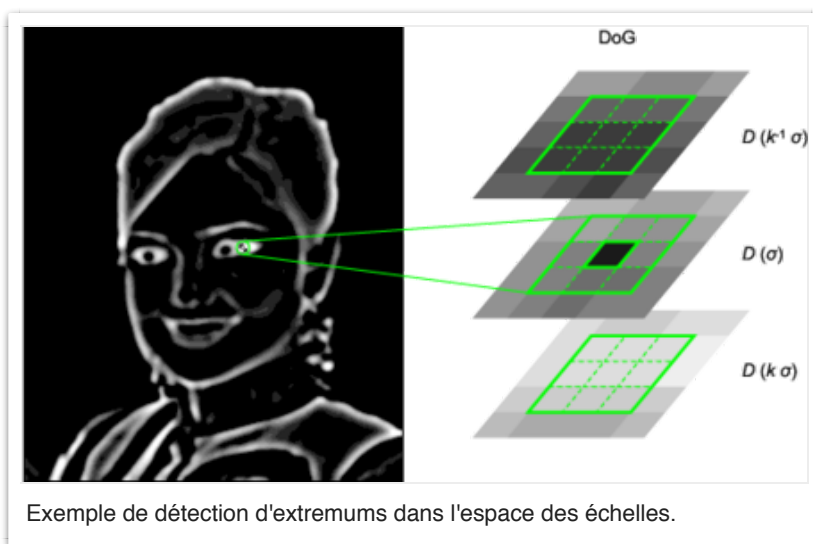
$$\{D(x + \delta_x, y + \delta_y, s\sigma), \delta_x \in \{-1, 0, 1\}, \delta_y \in \{-1, 0, 1\}, s \in \{k^{-1}, 1, k\}\}$$

L'utilisation **d'une pyramide est préconisée pour optimiser le temps de calcul** des images floutées à un grand nombre d'échelles différentes. La base de la pyramide est en général l'image originale et un niveau donné – on parle d'*octave* par analogie avec la *musique* – est obtenu à partir du précédent en divisant la résolution de l'image par 2, ce qui revient à doubler le facteur d'échelle. Au sein d'une même octave, le nombre de convolutions à calculer est constant. Le facteur fixe k dans les formules ci-dessus est calculé pour qu'au final, l'espace discrétisé des facteurs d'échelles considérés corresponde à une progression géométrique

$\{\sigma_0, k\sigma_0, k^2\sigma_0, \dots\}$, avec à chaque changement d'octave une valeur $k^p\sigma_0$ qui devient égale à une quantité de la forme $2^t\sigma_0$. Ce détail – la progression géométrique des facteurs d'échelle – est important pour que les valeurs des DoG à différentes échelles soient comparables entre elles et il évite, observe Lowe, d'avoir à utiliser un facteur de normalisation dans leur calcul.



L'étape de détection des points-clés candidats décrite ci-dessus est une variante de l'une des méthodes de *blob detection* (détection de zones) développée par Lindeberg, qui utilise le laplacien normalisé par le facteur d'échelle au lieu des DoG. Ces derniers peuvent être considérés comme une approximation des laplaciens et présentent l'avantage d'autoriser l'utilisation d'une technique pyramidale.



LOCALISATION PRÉCISE DE POINTS CLÉS

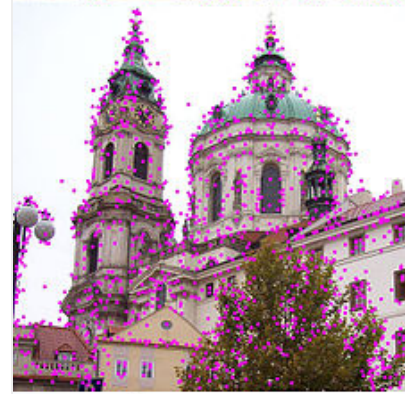
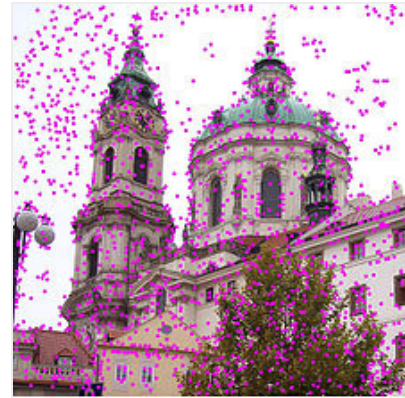
L'étape de détection d'extremums produit en général un grand nombre de points-clés candidats, dont certains sont instables ; de plus, leur localisation, en particulier aux échelles les plus grandes (autrement dit dans les octaves supérieures de la pyramide où la résolution est plus faible) reste approximative. De ce fait, des traitements supplémentaires sont appliqués, pour un objectif double : d'une part, reconverger la position des points pour améliorer la précision sur x , y et σ ; d'autre part, éliminer les points de faible contraste ou situés sur des arêtes de contour à faible courbure et donc susceptibles de « glisser » facilement.

AMÉLIORATION DE LA PRÉCISION PAR INTERPOLATION DES COORDONNÉES

Visant à augmenter de façon significative la stabilité et la qualité de la mise en correspondance ultérieure, cette étape, qui est une amélioration de l'algorithme original, s'effectue dans l'espace des échelles à trois dimensions, où $D(x,y,\sigma)$, qui n'est connu que pour des valeurs discrètes de x , y et σ , doit être interpolé.

Cette interpolation s'obtient par un [développement de Taylor](#) à l'ordre 2 de la fonction [différence de gaussiennes](#) $D(x,y,\sigma)$, en prenant comme origine les coordonnées du point-clé candidat. Ce développement s'écrit comme suit :

$$D(\mathbf{x}) = D + \frac{\partial D}{\partial \mathbf{x}} \mathbf{x} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2} \mathbf{x}$$



Après la détection des extremums dans l'espace des échelles (leurs positions sont indiquées sur l'image du haut), l'algorithme élimine les points de faible contraste (les points restants apparaissent sur l'image du milieu), puis les points situés sur les arêtes. Les points restants sont indiqués sur l'image du bas.

où D et ses dérivées sont évaluées au point-clé candidat et où $\mathbf{x} = (x, y, \sigma)^T$ est un delta par rapport à ce point. Les dérivées sont estimées par [différences finies](#) à partir des points voisins connus de façon exacte. La position précise de l'extremum $\hat{\mathbf{x}}$ est déterminée en résolvant l'équation annulant la dérivée de cette fonction par rapport à \mathbf{x} ; on trouve ainsi :

$$\hat{\mathbf{x}} = -\frac{\partial^2 D^{-1}}{\partial \mathbf{x}^2} \frac{\partial D}{\partial \mathbf{x}}$$

Un delta $\hat{\mathbf{x}}$ supérieur à 0,5 dans l'une des trois dimensions signifie que le point considéré est plus proche d'un des voisins dans l'espace des échelles discret. Dans ce cas, le point-clé candidat est mis à jour et l'interpolation est réalisée à partir des nouvelles coordonnées. Sinon, le delta est ajouté au point candidat initial qui gagne ainsi en précision.

Un algorithme de reconvergence similaire a été proposé dans l'implémentation temps-réel basée sur les pyramides hybrides de Lindeberg et Bretzner.

ÉLIMINATION DES POINTS-CLÉS DE FAIBLE CONTRASTE

La valeur de $D(\hat{\mathbf{x}})$ aux coordonnées précises du point-clé peut être calculée à partir du développement de Taylor de cette fonction, et constitue donc un extremum local. Un seuillage absolu sur cette valeur permet d'éliminer les points instables, à faible contraste.

ÉLIMINATION DES POINTS SITUÉS SUR LES ARÊTES

Les points situés sur les arêtes (ou contours) doivent être éliminés car la fonction DoG y prend des valeurs élevées, ce qui peut donner naissance à des extremums locaux instables, très sensibles au bruit : si l'image devait subir un changement numérique même imperceptible, de tels points-clés peuvent se retrouver déplacés ailleurs sur la même arête, ou même simplement disparaître.

Un point candidat à éliminer, si l'on considère les deux directions principales à sa position, est caractérisé par le fait que sa [courbure principale](#) le long du contour sur lequel il est positionné est très élevée par rapport à sa courbure dans la direction orthogonale. La courbure principale est représentée par les [valeurs propres](#) de la [matrice hessienne](#) \mathbf{H} :

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

Les dérivées doivent être évaluées aux coordonnées du point d'intérêt (x,y,σ) dans l'espace des échelles. Les valeurs propres de \mathbf{H} sont proportionnelles aux courbures principales de D , dont seul le rapport r est intéressant. La trace de \mathbf{H} représente la somme de ces valeurs, le déterminant son produit. Par conséquent, en adoptant un seuil r_{th} sur le ratio des courbures ($r_{th} = 10$ dans la méthode originale de Lowe), un point-clé candidat va être retenu, selon le critère adopté par Lowe, si :

$$R = \frac{\text{tr}(\mathbf{H})^2}{\det(\mathbf{H})} = \frac{(r + 1)^2}{r} < \frac{(r_{th} + 1)^2}{r_{th}}$$

La vérification de ce critère est rapide, ne nécessitant qu'une dizaine d'opérations flottantes seulement. Lorsque ce critère n'est pas vérifié, le point est considéré comme localisé le long d'une arête et il est par conséquent rejeté.

Cette étape est inspirée de la technique de [détection de points d'intérêt](#) par l'opérateur de Harris ; pour le seuillage, une matrice hessienne est utilisée au lieu de la matrice des moments d'ordre 2 ([tenseur](#)).

ASSIGNATION D'ORIENTATION

L'étape d'assignation d'orientation consiste à attribuer à chaque point-clé une ou plusieurs orientations déterminées localement sur l'image à partir de la direction des gradients dans un voisinage autour du point. Dans la mesure où les descripteurs sont calculés relativement à ces orientations, cette étape est essentielle pour garantir l'invariance de ceux-ci à la rotation : les mêmes descripteurs doivent pouvoir être obtenus à partir d'une même image, quelle qu'en soit l'orientation.

Pour un point-clé donné (x_0, y_0, σ_0) , le calcul s'effectue sur $L(x, y, \sigma_0)$, à savoir le gradient de la pyramide dont le paramètre est le plus proche du facteur d'échelle du point. De cette façon, le calcul est également invariant à l'échelle. À chaque position dans un voisinage du point-clé, on estime le gradient par différences finies symétriques, puis son amplitude (c.-à-d. sa [norme](#)) $m(x, y)$, et son orientation $\theta(x, y)$:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$

$$\theta(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right)$$

$\forall (x, y)$ dans un voisinage de (x_0, y_0) .

Un **histogramme** des orientations sur le voisinage est réalisé avec 36 intervalles, couvrant chacun 10 degrés d'angle. L'histo-



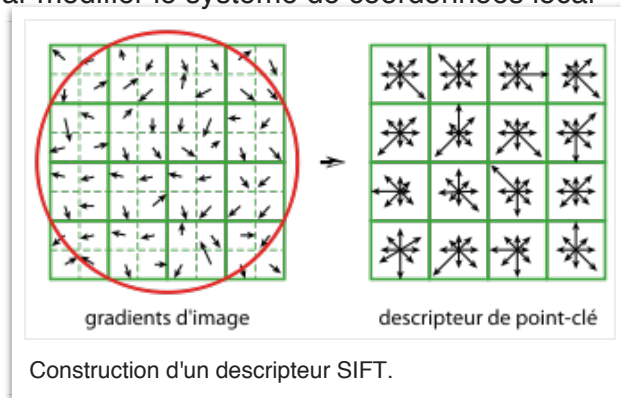
gramme est doublement pondéré : d'une part, par une fenêtre circulaire gaussienne de paramètre égal à 1,5 fois le facteur d'échelle du point-clé σ_0 ; d'autre part, par l'amplitude de chaque point. Les pics dans cet histogramme correspondent aux orientations dominantes. Toutes les orientations dominantes permettant d'atteindre au moins 80 % de la valeur maximale sont prises en considération, ce qui provoque si nécessaire la création de points-clés supplémentaires ne différant que par leur orientation principale.

À l'issue de cette étape, un point-clé est donc défini par quatre paramètres (x, y, σ, θ) . Il est à noter qu'il est parfaitement possible qu'il y ait sur une même image plusieurs points-clés qui ne diffèrent que par un seul de ces quatre paramètres (le facteur d'échelle ou l'orientation, par exemple).

DESCRIPTEUR DE POINT-CLÉ

Une fois les points-clés, associés à des facteurs d'échelles et à des orientations, détectés et leur invariance aux changements d'échelles et aux rotations assurée, arrive l'étape de calcul des vecteurs descripteurs, traduisant numériquement chacun de ces points-clés. À cette occasion, des traitements supplémentaires vont permettre d'assurer un surcroît de pouvoir discriminant en rendant les descripteurs invariants à d'autres transformations telles que la luminosité, le changement de point de vue 3D, etc. Cette étape est réalisée sur l'image lissée avec le paramètre de facteur d'échelle le plus proche de celui du point-clé considéré.

Autour de ce point, on commence par modifier le système de coordonnées local pour garantir l'invariance à la rotation, en utilisant une rotation d'angle égal à l'orientation du point-clé, mais de sens opposé. On considère ensuite, toujours autour du point-clé, une région de 16×16 pixels, subdivisée en 4×4 zones de 4×4 pixels chacune. Sur chaque zone est calculé un



histogramme des orientations comportant 8 intervalles. En chaque point de la zone, l'orientation et la magnitude du gradient sont calculés comme précédemment. L'orientation détermine l'intervalle à incrémenter dans l'histogramme, ce qui se fait avec une double pondération – par l'amplitude et par une fenêtre gaussienne centrée sur le point clé, de paramètre égal à 1,5 fois le facteur d'échelle du point-clé.

Ensuite, les 16 histogrammes à 8 intervalles chacun sont concaténés et normalisés. Dans le but de diminuer la sensibilité du descripteur aux changements de luminosité, les valeurs supérieures à 0,2 sont remplacées par 0,2 et l'histogramme est de nouveau normalisé, pour finalement fournir le descripteur SIFT du point-clé, de dimension 128.

Cette dimension peut paraître bien élevée, mais la plupart des descripteurs de dimension inférieure proposés dans la littérature présentent de moins bonnes performances dans les tâches de mise en correspondance pour un gain en coût de calculs bien modéré, en particulier quand la technique *Best-Bin-First* (BBF) est utilisée pour trouver le plus proche voisin. Par ailleurs, des descripteurs de plus grande dimension permettraient probablement d'améliorer les résultats, mais les gains escomptés seraient dans les faits assez limités, alors qu'à l'inverse augmenterait sensiblement le risque de sensibilité à la distorsion ou à l'occlusion. Il a également été démontré que la précision de recherche de correspondance de points dépasse 50 % dans les cas de changement de point de vue supérieur à 50 degrés, ce qui permet d'affirmer que les descripteurs SIFT sont invariants aux [transformations affines](#) modérées. Le pouvoir discriminant des descripteurs SIFT a pour sa part été évalué sur différentes tailles de bases de données de points-

clés ; il en ressort que la précision de mise en correspondance est très marginalement impactée par l'augmentation de la taille de la base de données, ce qui constitue une bonne confirmation du pouvoir discriminant des descripteurs SIFT.

UTILISATION POUR LA RECHERCHE D'OBJETS DANS DES IMAGES

La problématique de base pour laquelle la méthode SIFT a été conçue est la suivante : peut-on trouver dans une image donnée (dite *image question* ou *image suspecte*), des objets déjà présents dans une collection d'images de référence pré-établie ?

Dans la méthode originale de David Lowe, les points-clés et les descripteurs SIFT sont tout d'abord extraits des images de référence et stockés dans une sorte de base de données. Un objet est identifié dans l'image question en effectuant une comparaison de ses descripteurs à ceux des images de référence disponibles en base de données, fondée simplement sur la [distance euclidienne](#). Parmi toutes les correspondances ainsi établies, des sous-ensembles (*clusters*) sont identifiés, au sein desquels la mise en correspondance est cohérente du point de vue des positions des points, des facteurs d'échelle et des orientations. Les clusters contenant au moins trois correspondances ponctuelles sont conservés. Dans chacun d'eux, on modélise la transformation permettant de passer de l'image question à l'image de référence, et on élimine les correspondances aberrantes par simple vérification de ce modèle. Enfin, Lowe applique un modèle probabiliste pour confirmer que la détection d'une correspondance d'objets entre l'image question et l'une des images de référence n'est pas due au hasard, basé sur l'idée que si de nombreux points n'ont pas pu être mis en correspondance c'est que l'on a peut-être affaire à un [faux positif](#).

À chacune de ces étapes, Lowe propose une approche efficace présentée succinctement dans le tableau ci-dessous et dont les principes sont détaillés dans les paragraphes suivants.

Résumé des techniques utilisées dans les différentes étapes de SIFT

Étape	Techniques utilisées	Avantages
Extraction des points-clés (des	Pyramide de gradients, diffé-	Précision, stabilité, invariance

images de référence et de l'image question)	rence de gaussiens, assignation d'orientation	aux modifications d'échelle et à la rotation
<i>Calcul des descripteurs</i> (des images de référence et de l'image question)	Échantillonnage et lissage des plans locaux d'orientation de l'image	Stabilité relative aux transformations affines et à la luminosité
<i>Indexation</i> des descripteurs des images de référence	Arbre kd	Efficacité
<i>Recherche des correspondances</i> avec les descripteurs de l'image question	Plus proche voisin approximatif (<i>Best Bin First</i>)	Rapidité
<i>Identification de clusters</i>	Transformée de Hough et table de hachage	Modèle de transformation fiable
<i>Vérification du modèle</i>	Moindres carrés linéaires	Élimination des fausses correspondances
<i>Validation de l'hypothèse</i>	Inférence bayésienne	Fiabilité

INDEXATION DES DESCRIPTEURS ET RECHERCHE DE CORRESPONDANCES

L'indexation est l'opération de stockage des descripteurs SIFT des images de référence, d'une manière qui facilite l'identification des descripteurs correspondants de l'image question. Lowe utilise un [arbre kd](#) pour indexer les descripteurs puis une méthode de recherche dans cet arbre modifiée par rapport à l'approche classique, appelée *Best bin first*. Cette dernière est capable de trouver les [plus proches voisins](#) d'un descripteur question avec une bonne probabilité de façon très économe en temps de calcul. Cet algorithme se fonde sur deux astuces. Tout d'abord, le nombre de boîtes (feuilles de l'arbre kd) à explorer pour trouver le plus proche voisin d'un descripteur question donné est limité à une valeur maximale fixée, par exemple à 200. Ensuite, les nœuds de l'arbre kd sont explorés dans l'ordre de leur distance au descripteur question, grâce à l'utilisation d'une [file de priorité](#) basée sur un [tas binaire](#). Par distance d'un nœud à un descripteur, on entend distance euclidienne de la boîte englobante des feuilles sous-jacentes à ce descripteur. Dans la plupart des cas, ce procédé fournit la meilleure correspondance, et, dans les cas restants, un descripteur très proche de celle-ci.

De façon à déconsidérer les descripteurs faiblement discriminants (parce qu'ils sont trop souvent présents dans la base de référence, comme par exemple ceux qui sont extraits de motifs d'arrière-plan souvent répétés dans les images), Lowe recherche à la fois le plus proche voisin et le second plus proche voisin de chaque descripteur question. Lorsque le rapport des distances est supérieur à 0,8, la correspondance est éliminée, car considérée comme ambiguës (« bruit »). Par ce critère, Lowe parvient à éliminer 90 % des fausses correspondances en perdant moins de 5 % de correspondances correctes.

IDENTIFICATION DE CLUSTERS PAR LA TRANSFORMÉE DE HOUGH

On appelle *pose* d'un objet le point de vue sous lequel il est photographié, c'est-à-dire sa position dans l'espace par rapport à un repère absolu. En général, la base des images de référence contient différentes poses d'objets identiques. Chaque correspondance individuelle entre points-clés obtenue à l'étape précédente constitue une hypothèse quant à la pose de l'objet sur l'image question par rapport à l'image de référence concernée. Il s'agit donc maintenant de grouper les hypothèses cohérentes entre elles de façon à mettre en correspondance les objets des images et non plus seulement des points isolés. La [transformée de Hough](#) est utilisée pour cela. Elle identifie des groupes de correspondances, que l'on appelle *clusters*, par un système de vote.

Alors qu'un objet rigide possède en général six degrés de liberté dans l'espace (trois rotations, trois translations), une correspondance de points-clés entre deux images ne permet de relier que quatre paramètres (deux paramètres de position, l'angle et l'échelle) et constitue donc une approximation parmi toutes les poses possibles. Le sous-espace des poses ainsi paramétré est segmenté en boîtes de façon assez grossière : Lowe utilise des intervalles de 30 degrés pour l'orientation, un facteur de 2 pour l'échelle et 0,25 fois la dimension maximale de l'image de référence concernée (compte tenu du rapport d'échelles) pour la position. Chaque correspondance vote pour la boîte associée. Les correspondances concernant les facteurs d'échelles les plus élevés, considérées comme plus pertinentes, se voient attribuer un poids double de celles concernant les facteurs les plus bas. En outre, pour éviter les effets de bord lors de l'assignation des boîtes, chaque correspondance vote pour les 2 plus proches intervalles dans chaque dimension, c'est-à-dire qu'elle vote 16 fois.

Lors de l'implémentation, Lowe recommande l'utilisation d'une [table de hachage](#) pour éviter de représenter toutes les boîtes possibles en mémoire dont seulement un petit nombre risque d'être non vide.

Lorsque plusieurs correspondances votent pour une même boîte, c'est-à-dire pour la même pose d'un objet, la probabilité que la correspondance soit correcte est largement supérieure à ce que pourrait donner une correspondance isolée. Ainsi, les boîtes contenant au moins trois entrées sont considérées comme des clusters fiables, et retenus pour la suite de l'analyse.

VÉRIFICATION DU MODÈLE PAR LA MÉTHODE DES MOINDRES CARRÉS

Les clusters obtenus sont soumis à une procédure de vérification par l'application de la [méthode des moindres carrés](#) aux paramètres de la [transformation affine](#) reliant le modèle (image de référence) à l'image question. La transformation affine d'un point modèle $[xy]^T$ en un point image $[uv]^T$ peut s'écrire ainsi :

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix},$$

où $[t_x t_y]^T$ est la translation du modèle, et où les paramètres m_1 , m_2 , m_3 et m_4 modélisent la transformation vectorielle (qui peut être une rotation à l'origine, une similitude, une distorsion, etc.) Pour obtenir les paramètres de la transformation, l'équation ci-dessus est réécrite de manière à regrouper les inconnues dans un vecteur-colonne :

$$\begin{bmatrix} x & y & 0 & 0 & 1 & 0 \\ 0 & 0 & x & y & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ t_x \\ t_y \end{bmatrix} = \begin{bmatrix} u \\ v \\ \vdots \end{bmatrix}$$

Cette équation représente une seule correspondance. Pour obtenir les autres correspondances, il suffit de rajouter pour chacune d'elles deux lignes à la matrice et deux coefficients au second membre. Trois correspondances au minimum sont requises pour espérer une solution. Le système linéaire complet s'écrit alors :

$$A\mathbf{x} \approx \mathbf{b}$$

où A est une matrice $m \times n$ connue (vérifiant généralement $m > n$), \mathbf{x} un vecteur inconnu de paramètres de dimension n et enfin \mathbf{b} un vecteur connu de mesures de dimension m .

Ce système n'a en général pas de solution exacte puisqu'il est sur-dimensionné ; mais ce qui nous intéresse est qu'il est possible d'en trouver une **pseudo-solution** $\hat{\mathbf{x}}$ au sens des moindres carrés, c'est-à-dire un vecteur qui minimise l'erreur $\|\mathbf{b} - A\mathbf{x}\|_2$, et qui est aussi solution de l'« équation normale » :

$$A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$$

La solution du système d'équations linéaires, qui s'exprime grâce à la matrice $(A^T A)^{-1} A^T$, dite **pseudo-inverse de Moore-Penrose** de A , est donnée par :

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$$

Elle minimise la somme des carrés des distances entre les positions des points-clés calculées à partir de la transformée des positions du modèle, et les positions correspondantes sur l'image cible.

Les **données aberrantes** peuvent ensuite être écartées en vérifiant la concordance de chaque correspondance avec la solution des moindres carrés. Sont éliminées les correspondances qui se situent au-delà de la moitié de l'intervalle d'erreur utilisé pour la discrétisation de l'espace des poses dans la transformée de Hough. Une nouvelle solution par la méthode des moindres carrés est calculée à partir des points restants, et le processus est itéré plusieurs fois. Il est également possible grâce à la transformation estimée par la méthode des moindres carrés de récupérer des correspondances qui avaient été manquées précédemment. À la fin du processus, s'il reste moins de 3 points après suppression des points aberrants, alors la correspondance est dite infructueuse et rejetée.

VALIDATION FINALE PAR INFÉRENCE BAYÉSIENNE

La décision finale d'admission ou de rejet d'une mise en correspondance d'objets (hypothèse) se fait sur la base d'un modèle probabiliste. Cette méthode détermine en premier lieu le nombre attendu de fausses correspondances, connaissant la taille estimée du modèle, le nombre de points-clés dans la région et la précision

de la correspondance. Une analyse par [inférence bayésienne](#) donne ensuite la probabilité que l'objet trouvé sur l'image de référence soit effectivement présent dans l'image question à partir du nombre de vraies correspondances trouvées. Une hypothèse est acceptée si cette probabilité est supérieure à 0,98. À l'exception d'images présentant de grands écarts de luminosité ou des objets ayant subi des transformations sévères ou non rigides, la détection d'objet au moyen de la méthode SIFT parvient ainsi à d'excellents résultats.

COMPARAISON DES DESCRIPTEURS SIFT AUX AUTRES DESCRIPTEURS LOCAUX

Plusieurs études d'évaluation des performances de différents descripteurs locaux (notamment le descripteur SIFT) ont été menées, dont les principales conclusions sont :

- les caractéristiques SIFT et apparentées (en particulier GLOH) donnent les meilleurs taux de [rappel](#) pour des transformations affines de moins de 50 degrés ; au-delà de cette limite, les résultats deviennent incertains ;
- le caractère discriminant des descripteurs est mesuré en sommant leurs valeurs propres, obtenues par analyse en composantes principales appliquée à ces descripteurs normalisés par leur variance ; cela traduit dans les faits la quantité de variance « capturée » par les différents descripteurs et, partant, leur pouvoir discriminant ; les caractéristiques PCA-SIFT, GLOH et SIFT donnent les meilleurs résultats ;
- les descripteurs à base de SIFT dépassent en performance les autres types de descripteurs locaux en présence de scènes texturées ou, dans une moindre mesure, de scènes structurées ;
- les descripteurs à base de SIFT dépassent en performance les autres descripteurs dans le cas d'images texturées ou structurées ayant subi un changement d'échelle d'un ratio compris entre 2 à 2,5 ou une rotation comprise entre 30 et 45 degrés ;
- les performances de tous les descripteurs, et notamment ceux basé sur l'analyse des contours, tel que [shape context](#), se dégradent en présence d'images ayant subi un floutage substantiel (du fait de l'estompage des contours) ; en

tout état de cause, les descripteurs GLOH, PCA-SIFT et SIFT sont ceux qui présentent les meilleures performances de tous les descripteurs étudiés. Il en est de même dans les situations de modifications d'illumination.

Il ressort de ces études que les descripteurs à base de SIFT, du fait de leur plus grande robustesse et de leur meilleur pouvoir discriminant, sont particulièrement adaptés à la recherche de correspondances.

Des descripteurs plus récents, tel que SURF, n'ont cependant pas été pris en compte dans ces études, quoiqu'une étude ultérieure ait montré que ce dernier avait des performances similaires à celles de SIFT tout en étant sensiblement plus rapide.

Une nouvelle variante du descripteur SIFT, basée sur un [histogramme](#) irrégulier, a récemment été proposée, apportant une nette amélioration des performances. Au lieu d'utiliser des grilles de boîtes d'histogrammes de 4×4 , toutes les boîtes sont ramenées au centre de la caractéristique, d'où une amélioration significative de la robustesse aux changements d'échelle.

Enfin, comparée au descripteur SIFT standard, la technique SIFT-Rank s'est révélée avoir de meilleures performances. Un descripteur SIFT-Rank est généré à partir d'un descripteur SIFT standard, en fixant chaque boîte d'histogramme à son rang dans un tableau de boîtes. La [distance euclidienne](#) entre les descripteurs SIFT-Rank est invariante aux modifications monotones arbitraires dans les valeurs des boîtes d'histogramme, et elle est liée à la [corrélation de Spearman](#).

APPLICATIONS DES DESCRIPTEURS SIFT

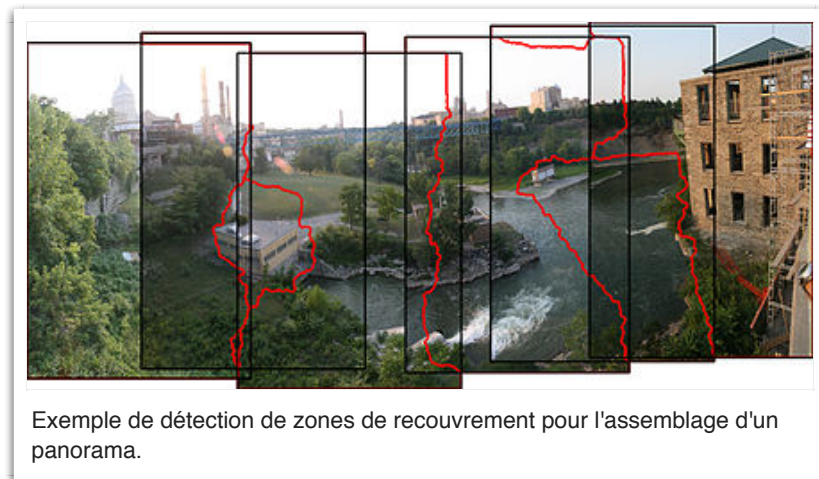
Outre la reconnaissance d'objet, les avantageuses propriétés des descripteurs SIFT (caractère discriminant, invariance à la translation, à la rotation et au changement d'échelle et robustesse aux transformations affines en général (distorsions), aux changements de points de vue 3D ainsi qu'aux changements de luminosité) en font un excellent choix pour d'autres applications dont quelques-unes sont énumérées ici.

LOCALISATION DE ROBOT EN ENVIRONNEMENT INCONNU

Le but de cette application est de fournir une solution robuste et fiable au problème de localisation de robots en environnement inconnu. Un système de **stéréoscopie** trinoculaire est utilisé pour estimer les positions 3D des points-clés déterminés. Un point-clé n'est pris en compte que s'il apparaît simultanément sur les trois images du système trinoculaire, et à condition que les disparités entre ces images ne soient pas handicapantes, afin de générer le moins possible de données aberrantes. En mouvement, le robot se localise en comparant les caractéristiques vues avec celles de la carte 3D qu'il a en mémoire ; il enrichit au fur et à mesure cette carte de nouvelles caractéristiques et met à contribution un **filtre de Kalman** pour mettre à jour leurs positions tridimensionnelles.

L'ASSEMBLAGE DE PANORAMAS

La mise en correspondance de points d'intérêt par la méthode SIFT peut être utilisée comme première étape d'un processus d'**assemblage de photos** entièrement automatisé. Cette application consiste à



Exemple de détection de zones de recouvrement pour l'assemblage d'un panorama.

construire une image panoramique unique à partir de prises de vues successives d'un même sujet (souvent un paysage) décalées les unes des autres par un déplacement dans l'espace. La technique proposée par Brown et Lowe en 2003 a l'avantage de fonctionner sur une collection d'images, sans qu'aucune connaissance a priori ne soit nécessaire sur l'endroit précis de la prise de vue de chacune de ces images, ni même leur ordre ; elle peut aussi exclure de l'assemblage des images qui n'auraient rien à voir avec un panorama donné et même reconstituer en même temps plusieurs panoramas indépendants.

L'idée de l'application est la suivante. Les descripteurs SIFT sont extraits de chacune des images à étudier et mis en correspondance les uns avec les autres (recherche des k plus proches voisins comme dans la méthode SIFT générale). On

construit ensuite un graphe dont les sommets sont les différentes images et l'on ajoute des arêtes entre les images qui ont le plus de points-clés en correspondance. Cela signifie qu'elles ont potentiellement des éléments du décor en commun. Pour en être sûr, on essaie de déterminer la transformation géométrique permettant de passer de l'une à l'autre, ou [homographie](#), en utilisant un [RANSAC](#) entre les points. En cas de succès, la paire d'images est soumise à une vérification supplémentaire basée sur un modèle probabiliste et destinée à s'assurer que les correspondances ne sont pas dues au hasard. Seules sont conservées dans le graphe les arêtes qui passent le test. Les composantes connexes correspondent alors à autant de panoramas indépendants, qui sont assemblés séparément. Pour cela, une technique dite d'« ajustement de faisceaux » (*bundle adjustment*) est utilisée pour déterminer les paramètres objectifs de la caméra lors de la prise de vue de chaque image, par rapport au panorama global. Une fois que cet ajustement est fait, il est théoriquement possible d'obtenir la valeur de chacun des pixels du panorama global à partir des images de départ. Pour la construction finale, on applique une pondération et un lissage liés à la nécessité d'harmoniser l'exposition des images de départ, de telle sorte que les grandes étendues uniformes (cieux, etc.) ne présentent pas trop de variations dans leur luminosité tout en s'assurant que les détails ne seront pas perdus. Cette technique sophistiquée s'appelle assemblage multi-bande (*multi band blending*).

La technique a donné naissance au logiciel [Autostitch](#).

MODÉLISATION, RECONNAISSANCE ET SUIVI D'OBJETS 3D

Cette problématique consiste à déterminer la position exacte d'une caméra à partir des images qu'elle acquiert, simplement en analysant l'orientation et la position d'un objet précis apparaissant dans son champ de vision. Les applications en sont diverses ; la plus populaire étant sans doute la [réalité augmentée](#), qui consiste à intégrer de nouveaux objets au sein d'une scène filmée, par exemple une bouteille sur une table dès que ladite table apparaît dans le champ de la caméra, et cela si possible en temps réel.

Une façon de traiter cette application a été proposée par Gordon et Lowe en 2006. Elle comprend une phase d'apprentissage qui permet d'établir une modélisation 3D de l'objet témoin à partir de l'analyse de différentes prises de vues 2D

de celui-ci. Les descripteurs SIFT sont utilisés pour mettre en relation les points-clés entre différentes images et construire un modèle 3D de l'objet (mise en correspondance dite « 2D-to-3D »). Une fois cette phase effectuée, on extrait de chaque image filmée les descripteurs SIFT et on les met en correspondance avec ceux des images de référence, de manière à obtenir les correspondances « 2D-to-3D » de l'image filmée. La position de la caméra est alors déterminée par une technique de RANSAC améliorée (Levenberg-Marquardt) qui minimise l'erreur de projection.

DESCRIPTEURS SIFT 3D POUR LA RECONNAISSANCE DE MOUVEMENTS HUMAINS

Des extensions au descripteur SIFT aux données spatio-temporelles de dimension 2+1, dans le contexte de reconnaissance de mouvements humains dans des séquences vidéo, ont été proposées dans la littérature scientifique. Le calcul d'histogrammes dépendants de positions locales est étendu, dans l'algorithme SIFT, de 2 à 3 dimensions pour adapter les caractéristiques SIFT au domaine spatio-temporel. Pour réaliser une reconnaissance de mouvements humains dans une séquence vidéo, un échantillonnage de vidéos d'apprentissage est effectué soit à des points d'intérêt spatio-temporel, soit à des positions, instants et échelles aléatoirement choisis. Les régions spatio-temporelles autour de ces points d'intérêt sont décrites au moyen du descripteur SIFT 3D. Ces descripteurs sont enfin rangés en clusters formant un [modèle de sac de mots](#) spatio-temporel. Les descripteurs SIFT 3D extraits des vidéos de test sont enfin comparés à ces « mots ».

Les auteurs confirment l'obtention de meilleurs résultats grâce à cette approche à base de descripteur 3D que par les autres approches utilisant de simples descripteurs SIFT 2D ou les magnitudes de gradients.

ANALYSE DU CERVEAU HUMAIN AU MOYEN D'IMAGES RMN 3D

La [morphométrie](#) basée sur des descripteurs d'image (FBM, *feature based morphometry*), est une technique d'analyse et de classification d'images 3D du cerveau humain obtenus par [résonance magnétique](#). Grâce à la technique SIFT, l'image est vue comme un assemblage géométrique de structures indépendantes (zones d'intérêt). En utilisant un modèle probabiliste assez fin et en se basant sur

l'observation de différentes classes de sujets dûment étiquetés, par exemple des sujets sains d'un côté et des sujets atteints de la [maladie d'Alzheimer](#) de l'autre, la technique FBM parvient à regrouper les structures similaires d'une image à l'autre et à les classer selon l'étiquetage. Un système ainsi entraîné devient capable d'étiqueter automatiquement une image étrangère : sur un ensemble de test constitué d'environ 200 images IRM du cerveau humain, la détection des cas sains et des cas atteints de la maladie d'Alzheimer est correcte dans 80 % des cas, précisent les auteurs.

VARIANTES ET EXTENSIONS

RIFT

RIFT est une variante de SIFT adaptée aux images texturées sur lesquelles la notion d'orientation principale n'a pas vraiment de sens, tout en étant invariante aux rotations. Le descripteur est construit en utilisant des patches circulaires normalisés divisés en anneaux également circulaires et d'égale largeur ; un histogramme d'orientation de gradient est calculé à partir de chacun de ces anneaux. Pour garantir l'invariance à la rotation, l'orientation est mesurée pour chaque point du centre vers l'extérieur.

G-RIF

G-RIF (ou *Generalized robust invariant feature*) est un descripteur de contexte regroupant les orientations des contours, les densités des contours et les informations de teintes dans une forme unifiée, combinant information perceptive et codage spatial. Une nouvelle méthode de classification est utilisée, basée sur un modèle probabiliste utilisant les votes des descripteurs locaux dans un voisinage.

SURF

Présenté comme une solution de hautes performances capable de d'approcher et même de dépasser les schémas précédents du point de vue répétitivité, distinctivité et robustesse, [SURF](#) (ou *Speeded Up Robust Features*) est un

détecteur de points d'intérêts invariant aux changements d'échelle et aux rotations. Afin de réduire les temps de traitement, il fait usage d'images intégrales pour le calcul de la convolution, et reprend les points forts des meilleurs détecteurs et descripteurs l'ayant précédé, en utilisant des mesures à base de matrices hessiennes rapides pour le détecteur et un descripteur basé sur les distributions. Il décrit une distribution de réponses d'ondelettes de Haar dans le voisinage du point d'intérêt. Seulement 64 dimensions sont utilisées, ce qui permet de réduire le temps de calcul des caractéristiques et de recherche des correspondances. L'étape d'indexation est basée sur le signe du Laplacien, d'où l'accélération de la recherche de correspondances et l'amélioration de la robustesse du descripteur.

PCA-SIFT et GLOH

PCA-SIFT et GLOH sont des variantes de SIFT. Le descripteur PCA-SIFT est un vecteur de gradients d'image dans les directions x et y calculé à l'intérieur de la région de support. La région du gradient est échantillonnée en 39×39 positions, générant un vecteur de dimension 3042. Cette dimension est réduite à 20 par la méthode d'analyse en composantes principales. GLOH (pour *Gradient location-orientation histogram*) est une extension du descripteur SIFT dont la robustesse et le caractère distinctif ont été améliorés. Comparé à SIFT, le descripteur GLOH est calculé à partir d'une grille de positions log-polaires avec trois boîtes en direction radiale (de rayons 6, 11 et 15) et huit en direction angulaire, soit 17 boîtes de position. La boîte centrale n'est pas divisée dans les directions angulaires. Les orientations des gradients sont quantifiées en 16 boîtes, d'où 272 histogrammes de boîtes. Ici aussi la taille du descripteur est réduite au moyen de l'analyse en composantes principales. La matrice de variance-covariance de l'ACP est estimée sur les patches d'image collectés de différentes images. Les 128 plus importants vecteurs propres sont utilisés pour la description.

Autres méthodes

Une équipe formée autour de Wagner a développé deux algorithmes de détection d'objet spécialement pensés pour tenir compte des limitations techniques des [photophones](#), mettant en œuvre le [détecteur de coins FAST](#) pour la détection de caractéristiques. Les deux algorithmes distinguent deux phases : une phase de préparation hors-ligne, au cours de laquelle les caractéristiques sont calculées pour différents niveaux d'échelles, et une phase de traitement en ligne, au cours de laquelle seules les caractéristiques du niveau d'échelle fixé de la photographie issue du [photophone](#) sont calculées. De plus, des caractéristiques sont obtenues à partir d'une [imagerie](#) d'une taille fixe de 15×15 pixels, formant un descripteur SIFT réduit à 36 dimensions. Cette approche a été ensuite améliorée par l'intégration d'un [arbre de vocabulaire évolutif](#) dans le pipeline de détection. Il s'ensuit une meilleure reconnaissance des objets sur les photophones. Le principal inconvénient de cette approche est son important besoin de [mémoire vive](#).

NOTES ET RÉFÉRENCES

NOTES

RÉFÉRENCES

- ^(en) Cet article est partiellement ou en totalité issu de l'article de Wikipédia en [anglais](#) intitulé « [Scale-invariant feature transform](#) » (voir [la liste des auteurs](#))

ANNEXES

ARTICLES CONNEXES

- [Autostitch](#)
- [Détection de zones d'intérêt](#)

- [Speeded Up Robust Features \(SURF\)](#)

BIBLIOGRAPHIE


PUBLICATIONS

- ^(en) YuanBin Wang, Zhang Bin et Yu Ge, « The Invariant Relations of 3D to 2D Projection of Point Sets », dans *Journal of Pattern Recognition Research*, vol. 3, n° 1, 14 juin 2008, p. 14-23 [[texte intégral](#) (page consultée le 27 septembre 2010)]
- ^(en) David G. Lowe, « Distinctive Image Features from Scale-Invariant Keypoints », dans *International Journal of Computer Vision*, vol. 60, n° 2, 2004, p. 91-110 [[texte intégral](#) (page consultée le 27 septembre 2010)]
- ^(en) Krystian Mikolajczyk et Cordelia Schmid, « A performance evaluation of local descriptors », dans *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, n° 10, 2005, p. 1615-1630 [[texte intégral](#) (page consultée le 27 septembre 2010)]
- ^(en) Svetlana Lazebnik, Cordelia Schmid et Jean Ponce, « Semi-Local Affine Parts for Object Recognition », dans *Proceedings of the British Machine Vision Conference*, vol. 2, septembre 2004, p. 959-968 [[texte intégral](#) (page consultée le 27 septembre 2010)]

IMPLÉMENTATIONS DE SIFT

- ^(en) Rob Hess, « [SIFT Library](#) », 2010. Consulté le 27 septembre 2010.
- ^(en) David G. Lowe, « [Demo Software: SIFT Keypoint Detector](#) », 2005. Consulté le 27 septembre 2010.
- ^(en) Yan Ke et Rahul Sukthankar, « [PCA-SIFT: A More Distinctive Representation for Local Image Descriptors](#) », 2004. Consulté le 27 septembre 2010.
- ^(en) Jean-Michel Morel et Guoshen Yu, « [ASIFT: A New Framework for Fully Affine Invariant Comparison](#) », 2009, (ISBN 2105-1232). Consulté le 27 septembre 2010.
- ^(en) Andrea Vedaldi et Brian Fulkerson, « [VLFeat.org](#) », 2005. Consulté le 27

septembre 2010.

- ^(en) Wan-Lei Zhao, « [Lip-vireo](#) », 2010. Consulté le 27 septembre 2010.
- Depuis sa version 2.2 (décembre 2010), [OpenCV](#) intègre une implémentation de SIFT.
-  [Portail de l'imagerie numérique](#)



: les textes sont disponibles sous [licence Creative Commons paternité partage à l'identique](#) ; d'autres conditions peuvent s'appliquer. Voyez les [conditions d'utilisation](#) pour plus de détails, ainsi que les . En cas de réutilisation des textes de cette page, voyez . Wikipedia® est une marque déposée de la [Wikimedia Foundation, Inc.](#), organisation de bienfaisance régie par le paragraphe du code fiscal des États-Unis.

