

## Indexation des images : descripteurs locaux

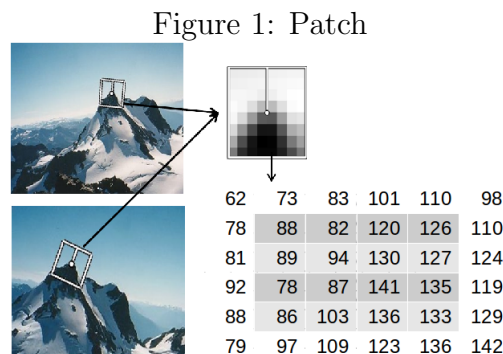
### Exercice 1 : Descripteurs SIFT

Le descripteur SIFT (Scale Invariant Feature Transform) consiste à calculer un histogramme d'orientation de gradients.

1. Dans SIFT, quelles sont les paramètres usuellement utilisés pour la taille de grille et le nombre d'orientations?
2. Quelle est la dimension usuelle d'un SIFT?
3. Décrivez les grandes étapes du calcul du descripteur SIFT?

### Exercice 2(Examen) : Calcul d'un descripteur

Sur la figure est représentée une imagerie(patch) (encadrée par le rectangle) de taille 4x4 pixels synthétisée à partir d'une image réelle. Le voisinage de l'imagerie est représenté (d'où la taille 6x6 pixels)



Dans cet exercice, vous allez effectuer le calcul des premières composantes d'un descripteur SIFT simplifié des patches. (Pour cela on prend les paramètres suivants): la grille spatiale est de taille 2x2 le nombre d'orientations possibles est fixé à 4

1. Quelle est la dimension d'un SIFT qui résulte du paramétrage proposé?
2. Calculer le descripteur SIFT pour ce patch ?

(a) Calculer le module  $I_g$  et l'orientation  $I_{or}$  du gradient en chaque pixel

En pratique, on approximera les dérivées partielles par des différences finies : dans ce cas les dérivées partielles  $I_x$  et  $I_y$  peuvent s'obtenir en calculant la convolution de l'image par un masque :  $I_x = I * M_x$ ,  $I_y = I * M_y$ . On utilisera ici les masques de Sobel 3x3 suivants :

$$M_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad M_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

- i. Montrer que les masques  $M_x$  (resp.  $M_y$ ) sont séparables, c'est à dire qu'ils peuvent s'écrire comme  $M_x = h_y * h_x$ , où  $h_y$  est un masque 1d sur les lignes (taille  $3 \times 1$ ) et  $h_x$  est un masque 1d sur les colonnes (taille  $1 \times 3$ ).
- ii. Calculer  $I_x$  et  $I_y$ .
- iii. En déduire le module du gradient  $I_g = ||\vec{G}|| = \sqrt{I_x^2 + I_y^2}$
- iv. Calculer l'orientation du gradient  $I_{or}$  en chaque pixel
- (b) Calculer gaussienne  $M_\sigma$  pour pondérer le calcul du SIFT en appliquant un masque Gaussien ( $3 \times 3$ )
- (c) Calculer le descripteur SIFT

On décompose le voisinage de point d'intérêt en 4 blocs de  $2 \times 2$  pixels, dans chaque bloc on forme un histogramme d'orientation de gradients, en discrétisant l'orientation en 4 bins.

- i. La formation de l'histogramme est effectué comme suit :

$$H[\Theta^i(x, y)] = H[\Theta^i(x, y)] + ||\vec{G}(x, y)|| * M_\sigma(x, y)$$

où  $\Theta^i(x, y)$  est l'orientation du gradient discrétisée au pixel  $(x, y)$ . Ainsi, on incrémente le bin de l'histogramme correspondant à  $\Theta^i(x, y)$  de la valeur  $||\vec{G}(x, y)|| * M_\sigma(x, y)$

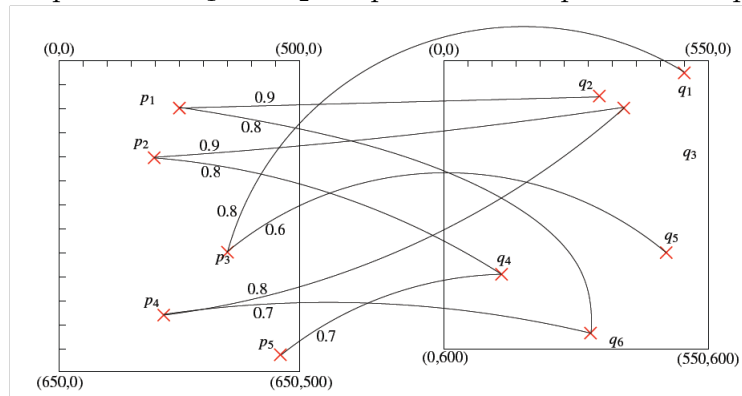
où  $||\vec{G}(x, y)||$  est le module du gradient au pixel  $(x, y)$  et  $M_\sigma(x, y)$  la pondération obtenue au pixel  $(x, y)$  en appliquant un masque Gaussien ( $3 \times 3$ )

- ii. Le descripteur SIFT associé au centre de patch (point d'intérêt), est la combinaison des histogrammes d'orientation de gradient  $H = (H1, H2, H3, H4)$ ,  $H_i = []$ : histogramme d'orientation du gradient de bloc  $i$ .

### Exercice 3 : Mise en correspondance : appariements de points

La figure 2 présente la mise en correspondance initiale effectuée, en utilisant les intensités lumineuses, entre les points d'intérêt des images I1 et I2. Seuls les appariements significatifs ont été conservés : ceux dont la similarité était supérieure à 0.9 ont été supprimés par seuillage. Les données de similarité sont résumées sur la figure 2.

Figure 2: Correspondance  $I_1 \rightarrow I_2$  : représentation spatiale des appariements



1. Donner la matrice de similarité entre points d'intérêt (correspondance entre I1 et I2).
2. Désambigüiser les appariements multiples résultant de la mise en correspondance initiale.
3. Donnez une hypothèse sur la nature de la transformation affine ayant transformé l'image I1 en l'image I2.