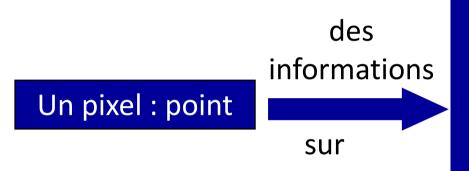
Qu'est-ce que la vision?

- Notre environnement spatial est composé d'objets, à structure 3D
- L'humain sait décrire et interpréter cet environnement (presque parfaitement)
- Or l'information disponible sur la rétine n'est qu'un ensemble de points
- Pour un capteur on parle de pixels (picture elements)

Qu'est-ce que la vision?

 Les éléments « vus » (table, chaise, personnage, ...) n'existent pas sur la rétine!!



la lumière reçue en ce point de la rétine (quantité, contenu spectral/couleur, Orientation de la source de lumière)

Interprétation : résultat du processus visuel et cognitif

Traitement : vision et cerveau

Un peu de texte

Analyse qualitative

La Neurobiologie

- but : comprendre l'anatomie du cerveau et le fonctionnement de l'acquisition et des mécanismes neuronaux des systèmes biologiques.
- outil : l'exploration in situ (modèle, expériences, ..)

La Psychologie Perceptive (ou psychophysique)

- but : comprendre les mécanismes sous-jacent à la perception visuelle, incluant les mécanismes intelligents.
- outil : l'analyse de la réaction à des situations particulières.

•

Analyse quantitative

La Vision par Ordinateur (VO)

but : concevoir et analyser les aspects calculatoires et algorithmiques des processus d'acquisition, de traitement et d'interprétation des images numériques.

outil: la modélisation mathématique et algorithmique.

--> La vision Industrielle

but : construire des machines capables de traiter un ensemble connu de tâches qui s'appuient sur la vision .

outil : la méthodologie de conception de systèmes industriels et l'intégration.

modèle computationnel

Livres en vision par ordinateur

- David A. Forsyth & Jean Ponce. Computer Vision: A Modern Approach, Prentice-Hall, 2002.
- R. Horaud & O. Monga. Vision par Ordinateur: Outils Fondamentaux, Editions Hermès, 1995.
 - http://www.inrialpes.fr/movi/people/Horaud/livre-hermes.html
- Dana H. Ballard & Christopher M. Brown. Computer Vision, Prentice-Hall, 1982.
- Kenneth R. Castleman. Digital Image Processing. Prentice-Hall, 2000.
- Olivier Faugeras. Three-Dimensional Computer Vision, MIT Press, 1993.
- Linda G. Shapiro, George C. Stockman, Linda G Shapiro & George Stockman.
 Computer Vision. Prentice-Hall, 2001.

Revues et conférences

- Revues (les plus anciennes)
 - IEEE Transactions on Pattern Analysis and Machine Intelligence
 - Pattern Recognition
 - Computer Vision and Image Understanding
 - International Journal of Computer Vision
- Conférences (les plus connues)
 - International Conference on Pattern Recognition (ICPR)
 - International Conference on Image Processing (ICIP)
 - International Conference on Computer Vision (ICCV)
 - International Conference on Computer Vision and Pattern Recognition (CVPR)
- Recherche bibliographique (Computer Vision Bibliography)
 - http://iris.usc.edu/Vision-Notes/bibliography/contents.html

I.O . Un peu d'histoire

I.0. Un peu d'histoire

L'histoire de ce domaine (assez récent) qu'est la vision artificielle est généralement découpée en quatre grandes parties :

Cible: observateur humain:

•Transmission (avec le premier client : la presse). Avant l'apparition des satellites : moyens limités (1920) 1 semaine pour 1 image entre Londres et N. York

```
Système Bartlane --> 3heures : image codée 5 niveaux/point → Reconstruction après 1923 → on passe à 15 niveaux
```

I.0. Un peu d'histoire

Des premières images à leur traitement (1950-1970):

<u>Début du traitement d'images</u>: 1950 (avec les premières analyses d'images dans les chambres à bulles). Application : Rayons X, OCR, ...

Des images de mauvaise qualité et très volumineuses (déjà 700 x 500 pixels sur 8bits par image)

Trois domaines dominants:

Restauration (corriger les défauts liés à l'acquisition) Amélioration (rendre l'image "belle" pour l'affichage) Compression (réduire le volume de l'image)

Du traitement à « l'interprétation » d'images (1970 - 1980) :

- Une évolution naturelle vers l'extraction automatique d'informations
- Apparition de la notion de description structurelle (ensemble des informations contenues dans une image et que l'on peut extraire sans aucune connaissance a priori de l'application).
- Des nouveaux thèmes : seuillage, segmentation, extraction de contours, morphologie mathématique

I.O. Un peu d'histoire

Interprétation (au départ lié à l'IA)

- Un engouement explosif au moment de l'apparition des systèmes experts
- Une déception proportionnelle à l'engouement
- --> Les raisons de l'échec
 - il n'y a pas d'expert pour la vision (règles difficile à cerner)
 - le savoir est trop complexe pour être décrit et résumé par un seul modèle

I.O. Un peu d'histoire

La vision par ordinateur (vers 1980):

- de l'image 2D aux modèles tri-dimensionnels
- Analyse du mouvement
- Vision pour la robotique (mouvement, 3D, détection d'obstacle, trajectoire)
 S'appuie sur le paradigme de Marr
- !! Le bilan : pousse à revoir les buts (peu de problèmes généraux résolus mais a donné de l'élan à beaucoup de domaine
- 3D complet est-il possible ? est-ce même une solution ? De la vision passive à la vision active !!!

I.O. Un peu d'histoire

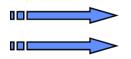
Et maintenant...

• Plus réaliste, conscient des limites.

--- On sait que l'analyse dépend du but poursuivi ---

-->>Un nouvel élan :

- la réalité augmentée
- la TV interactive, les loisirs numériques



- les bibliothèques numériques

Représentation (quels traitement, quel codage ?)

Transmission (codage et réseaux)

il ne s'agit plus d'interpréter pour interpréter mais d'interpréter pour recherche restituer une information (reconstruction basée images, Indexation,)

I.1. Une Théorie de la VO?

I.1. Une Théorie de la VO?

Est-il possible de concevoir un modèle computationnel ?

Peut être mais!!

La VO ne cherche pas à comprendre ou à reproduire la vision biologique

Il s'agit de construire un modèle mathématique et/ou algorithmique vu de l'extérieur le modèle créé, possède des propriétés semblables

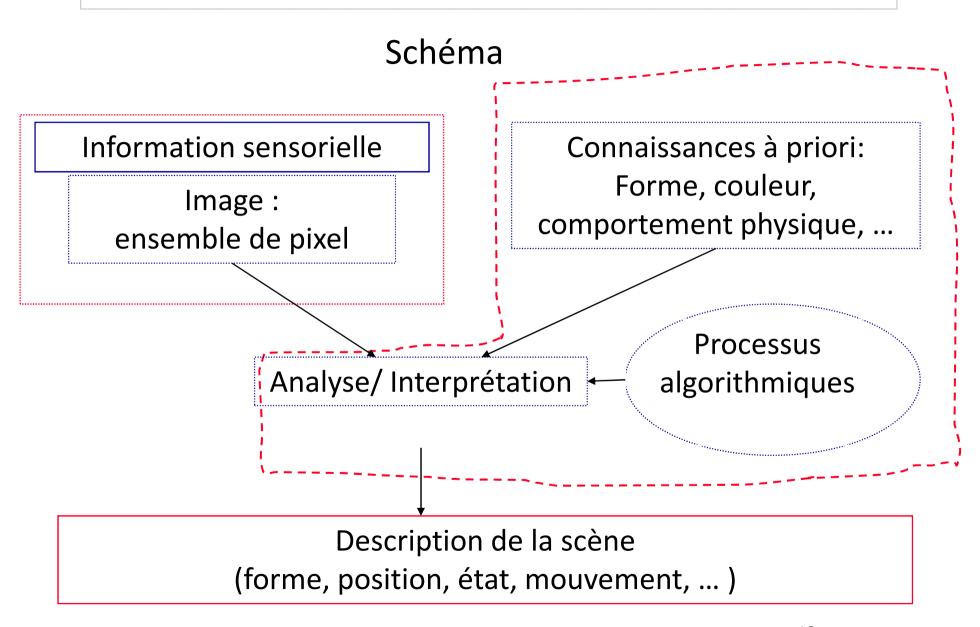
Il ne s'agit pas de machine qui voit (trop tôt)

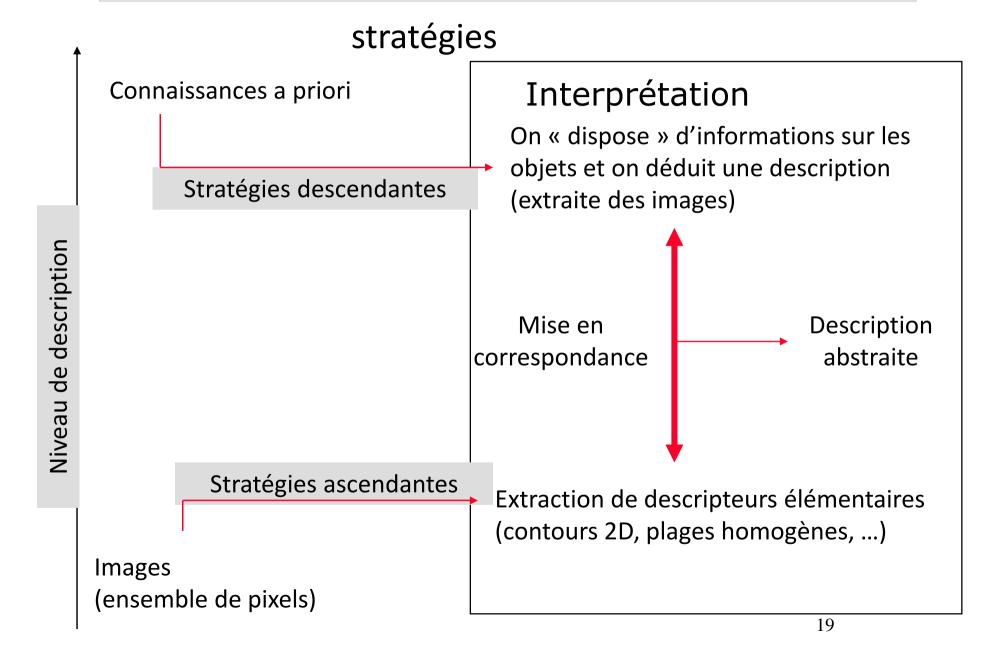
Paradigme en 3 étapes

- Enoncer la théorie : spécifier les données (l'entrée) et élaborer les concepts de bases
- -->exprimer le cadre formel qui est à la base de la théorie.
- Formaliser les concepts modèle mathématique algorithmique (ne pas confondre <u>mathématique</u> et <u>analytique</u>)
- •Réaliser un ensemble expérimental qui permet de vérifier la théorie

Enoncé (Radu Horaud):

- processus de traitement de l'information.
 l'entrée du système est une séquence d'images
- utilise des stratégies bien définies pour atteindre ses buts. Les stratégies font intervenir un certain nombre de connaissances (hypothèses, contraintes liées à l'entrée)
- la sortie est une description de l'entrée en termes d'objets, des état des objets et de relations entre ces objets





Connaissances a priori

• Quantitatives:

- géométrique (forme, formation de l'image)
- physiques (optiques, lumière,...)

- ...

•---> Quantifiable, mesurable (souvent)

• Qualitatives :

- sémantiques (fonction des objets, relations entre les objets, comportement)

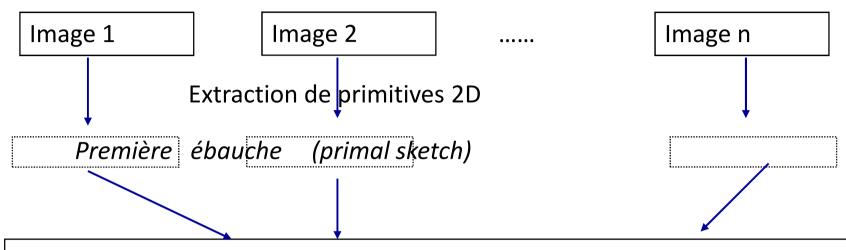
Algorithmes et méthodes

- Algorithmes : processus traduisant des stratégies d'analyse et des méthodes dérivant de théories mathématiques
 - extraction de primitives, de descripteurs
 - représentation et organisation des connaissances
 - comparaison, appariement

• Exigences:

-un système de vision doit être fiable, rapide robuste aux conditions initiales

Paradigme de David Marr



Ebauche 2,5D

Représentation centrée sur l'observateur (le monde 3D vu depuis la caméra)

- Stéréoscopie
- Analyse du mouvement,
- Ombrage, reflets, textures, ...

+ Connaissances 3D

But : construire l'environnement 3D, centré sur la scène

II. Segmentation, reconstruction, reconnaissance

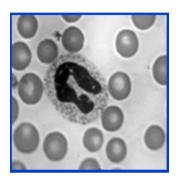
Segmentation, reconstruction, description

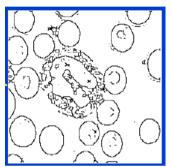


Paradigme de David Marr

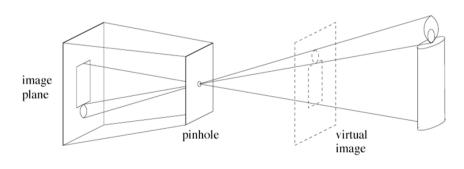


La segmentation est la pierre de base : analyse bas niveau





Reconstruction : calibrage, mises en correspondance









description: quelques algorithmes (squelettisation, ..)

II.1 Segmentation

II.1 Segmentation

La segmentation est la pierre de base : analyse bas niveau

Se situe juste après l'étape de traitement :

Filtrage (traitement du bruit liés à l'acquisition)
Restauration (corriger les défauts liés à l'acquisition)

Beaucoup d'algorithmes souvent des variantes de la même stratégie

II.1 Segmentation

2 familles de méthodes :

Méthodes à base de détection de contours (Edge based methods)

Méthodes à base de décomposition en régions (region based methods)

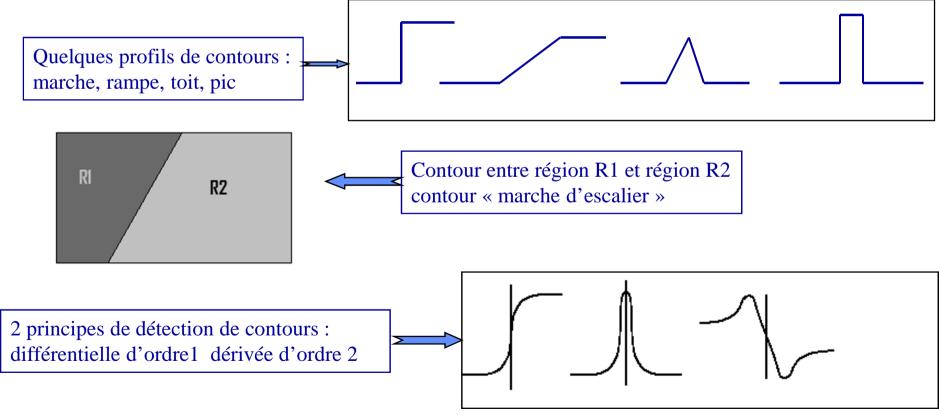
II.1.1 Segmentation à base de frontières Detection de contours : Edge detection

Contours : zones de pixels correspondant à une forte transition. Changement brusque de la valeur image.

- --> Les opérateurs différentiels sont des opérateurs qui mettent en évidences certaines variations spatiales dans les images.
- -->Ils sont utilisés comme traitements de base dans de nombreuses opérations

II.1.1 Segmentation à base de frontières Detection de contours : Edge detection

- * Présentation des opérateurs de convolution
- * Définition d'une convolution :
- * Gradient d'une image
- * Interprétation géométrique du gradient



II.1.1 Segmentation à base de frontières Produit de convolution

Convolution Analogique

L'équation générale de convolution est une application du principe de superposition. Il s'agit de somme de réponses impulsionnelles pour les réponses

$$y(t) = \int_{-\infty}^{+\infty} X(t - u) f(u) du$$

Cette intégrale est appelée *produit de convolution* des fonctions X et f. On la note par : X * f

La convolution est l'opération de traitement de signal la plus fondamentale. Elle indique que la valeur du signal de sortie à l'instant t est obtenue par la sommation (intégrale) pondérée des valeurs transmises du signal d'excitation X(t).

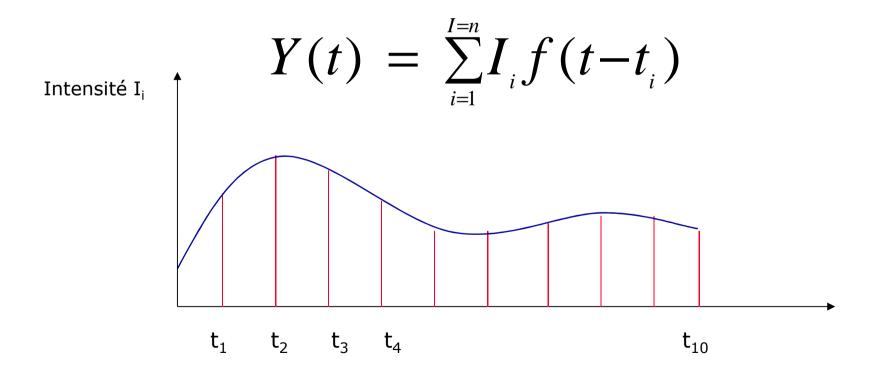
La fonction f est la fonction de transfert du signal. Elle joue le rôle de fonction de pondération. C'est précisément la réponse impulsionnelle f(t).

II.1.1 Segmentation à base de frontières Produit de convolution

* Cas discret:

S'il y a plusieurs traits en x1, x2, x3, ..., d'intensités respectives l_1 , l_2 , l_3 , ...,

H étant le filtre, le signal de sortie est donné par :



II.1.1 Segmentation à base de frontières produit de convolution

Généralisation:

Si le signal d'entrée est une fonction à valeurs discrètes de t, on peut passer de l'intégrale à une sommation :

$$y(t) = \sum_{i=-\infty}^{i=+\infty} x(t-t_i) f(t_i)$$

Si le support de f est un intervalle fini [-m, +m] on aura :

$$y(t) = \sum_{i=-m}^{i=+m} x(t-t_i) f(t_i)$$

II.1.1 Segmentation à base de frontières

produit de convolution : propriétés

Le produit de convolution est

commutatif : f * h = h * f

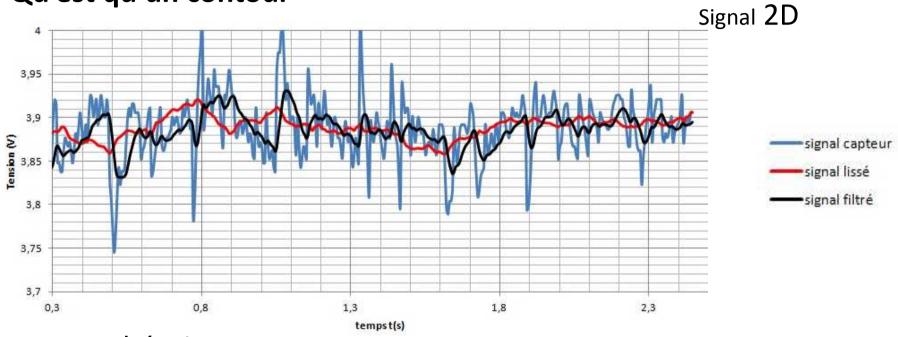
associatif : f * (g * h) = (f * g) * h

distributif : f * (g + h) = f * g + f * h

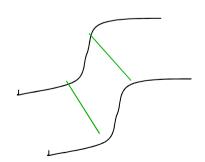
Propriété de dérivation : f * h' = (f * h)'

• Segmentation : détection de contours

Qu'est qu'un contour



Contour théorique



Contour: changement brutal de l'intensité lumineuse.

Segmentation détection de contours

Soit le point p (x, y)

Si on suppose que f est une fonction continue, pour estimer la variation de I=f au point p on a (p(x, y)):

$$dz = \frac{\partial f(x,y)}{\partial x} dx + \frac{\partial f(x,y)}{\partial y} dy$$

dz ou $\Delta f(x,y)$ est appelé Gradient et a pour amplitude :

$$\|\Delta f(x,y)\| = \sqrt{\left(\frac{\partial f(x,y)}{\partial x}\right)^2 + \left(\frac{\partial f(x,y)}{\partial y}\right)^2}$$

Norme euclidienne

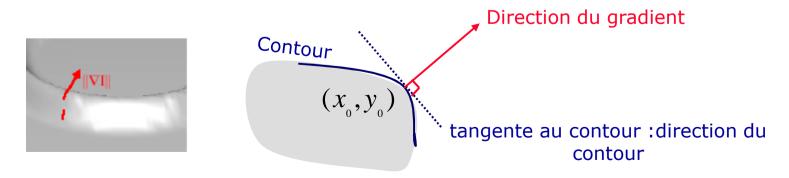
Sa direction est donnée par :
$$\theta = Arc \tan \left(\frac{\partial f(x,y)}{\partial y} \frac{\partial f(x,y)}{\partial x} \right)$$

• Segmentation : détection de contours

Le gradient est un opérateur directionnel.

L'amplitude du gradient de l'image en 1 pixel donne une indication sur l'amplitude de la variation de la valeur image (niveau de gris par exemple) en ce point.

La direction du vecteur gradient est par définition normale au contour.



Dans le domaine discret (image), le calcul des dérivées partielles font l'objet d'une approximation par des différences finies.

$$\frac{\partial f^{+}(x,y)}{\partial x} \quad \text{est \'evalu\'ee par} \quad \frac{\Delta^{+}}{x} \frac{f(x+\Delta x,y)-f(x,y)}{\Delta x}$$

$$\frac{\partial f^{+}(x,y)}{\partial y} \quad \text{est \'evalu\'ee par} \quad \frac{\Delta^{+}}{x} \frac{f(x+\Delta x,y)-f(x,y)}{\Delta y}$$

En passant à la dérivée totale on obtient :

$$\frac{\partial f(x,y)}{\partial x} \approx \frac{\Delta \frac{+}{x} f(x,y) + \Delta \frac{-}{x} f(x,y)}{2}$$

$$\frac{\partial f(x,y)}{\partial x} \approx \frac{f(x+1,y) - f(x-1,y)}{2}$$

Segmentation détection de contours

Dans le domaine discret (image), le calcul des dérivées partielles peuvent être vues comme la convolution de f par une fonction discrète à support fini.

Ce qui peut être écrit sous la forme :

$$\frac{\partial f(x,y)}{\partial x} \approx \frac{\Delta_{\chi}^{+} f(x,y) + \Delta_{\chi}^{-} f(x,y)}{2}$$

$$\frac{\partial f(x,y)}{\partial x} \Big|_{(i,j)} \approx \frac{1}{2} f * K(i,j) \quad \text{avec} \quad K = [...0 \quad -1 \quad 0 \quad 1 \quad 0...]$$

En supposant que l'opérateur est séparable, en procédant de la même manière on peut écrire :

$$\frac{\partial f(x, y)}{\partial y} \approx \frac{\Delta_{y}^{+} f(x, y) + \Delta_{y}^{-} f(x, y)}{2}$$

$$\frac{\partial f(x, y)}{\partial y}](i, j) = f * K(i, j)$$
avec $K_{y} = [...0 -1 0 1 0...]^{t}$

Filtrage suivi de dérivation

Soit h un filtre donné, on peut écrire : f * h' = (f * h)'

Ainsi on peut appliquer le gradient (la dérivé) d'un filtre pour atténuer le bruit et calculer le gradient de l'image.

Application à la détection de contours

Niveau de gris de l'image au point (x, y) : I(x, y)

Gradient:
$$\nabla I(x,y) = (G_x, G_y) = (\frac{\partial I(x,y)}{\partial x}, \frac{\partial I(x,y)}{\partial y})$$

Module/Norme du gradient :
$$G = \sqrt{G_x^2 + G_y^2}$$

Orientation du gradient : $\theta = arctan(\frac{G_y}{G_x})$

Exemple de calcul (peu précis) : masque de Prewitt

$$y \uparrow \longrightarrow_X$$

-1	0	1
-1	0	1
-1	0	1

-1	-1	-1
0	0	0
1	1	1

Application à la détection de contours

Récapitulatif:

- Un détecteur de contour est un opérateur différentiel
- Peut être réalisée par une convolution (discret :masque)
- La somme des coefficients est nulles
- L'opérateur est antisymétrique (fonction impaire)

Post-traitement

Problématique: 2 points

- Opérateurs différentiels : sensibles au bruit les variations d'intensité dues au bruit sont aussi des points de fort gradient
 - besoin de pré- prétraitements
 - de post-traitements
- 2. Résultat ensemble de point sans structurer le résultat structure

Post-traitement

1. Seuillage ou valeurs de tolérance

But : sélectionner les valeurs significatives

2. Seuillage par hystérésis des maximums locaux

But : limiter la fragmentation des contours obtenus

--> 2 seuils : *Sh* > *Sb*

On garde:

- les maximums locaux de valeur supérieure à S_h
- les maximums locaux de valeur supérieure à Sb appartenant à une composante connexe de maxima locaux contenant au moins une valeur S_h
- Résultat : une image binaire (image des contours)

Retour sur quelques notions

L'image peut être plongée dans le Plan discret : Z²

Point discret: point de Z², correspond au pixel centré en ce point

2 points discrets P et Q sont 4-adjacents

$$|x_P - x_Q| + |y_P - y_Q| = 1$$
 ou $[(x_P - x_Q)^2 + (y_P - y_Q)^2]$

2 points discrets P et Q sont 8-adjacents

$$Max\{|x_P - x_Q|, |y_P - y_Q| = 1\}$$

Point discret 4-connexité 8-connexité

Composante connexe / région 2D

Ensemble k-connexe : ensemble de points discrets E tel que $\forall P, Q \in E$,

∃ un chemin discret *k*-connexe dans *E* d'extrémités *P* et *Q*. Composante connexe d'un ensemble de points discrets : ensemble connexe maximal (ou classe d'équivalence pour la relation d'adjacence).

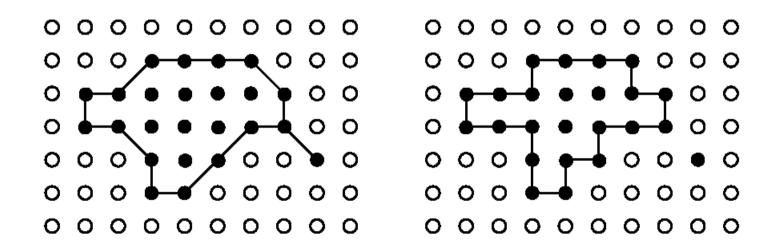
Exemple : ensemble discret composé d'une composante 8-connexe (de 2 composantes 4-connexes)

Frontière d'une région 2D

Définition classique

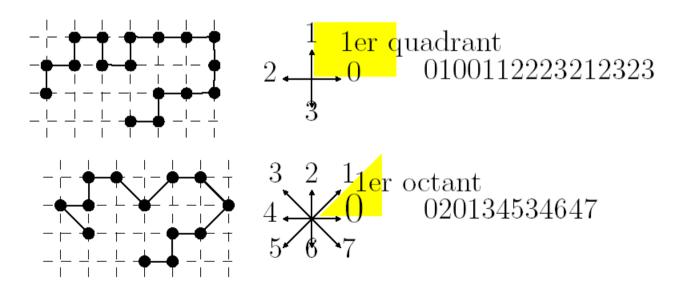
La frontière d'une région 8-connexe (respectivement 4-connexe) *R* est l'ensemble des points de *R* dont au moins un des 4-voisins (resp. 8-voisins) n'est pas élément de *R*.

⇒ La frontière est composée de chemins 8-connexes (resp 4-connexes).



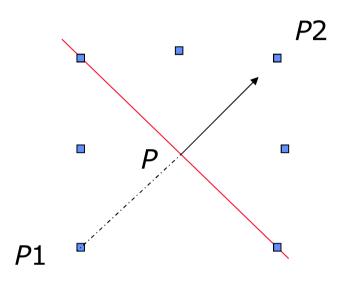
Chemin discret

Un chemin discret k-connexe est une suite de points discrets $(P_0, P_1, ..., P_n)$ tels que $\forall i \in 1...n, P_{i-1}$ et P_i sont k-adjacents. Code de Freeman : $(P_0, ..., P_n)$ est représenté par $(P_0, d_0, ..., d_{n-1})$. La direction d_i code le déplacement élémentaire de P_i à P_{i+1} .



Affinage des contours

Extraction des maxima locaux de la norme du gradient



Gr : norme du gradient en P

Gr1: norme du gradient en P1

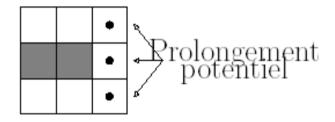
Gr2 : norme du gradient en P2

Présence d'un maximum local : $Gr \ge Gr1$ et $Gr \ge Gr2$

Image des contours

Fermeture des contours

- Idée : suivre une ligne de crète dans l'image de la norme du gradient à partir de chaque extrémité de contour.
- **1-** Repérer les points extrémité (énumération des configurations possibles)



2- Choix entre les points candidats : on explore tous les chemins possibles à partir de chaque point candidat. Le poids d'un chemin peut être défini comme la somme de la norme du gradient en chacun de ses points.

Détection de contours : Modèle de Canny-Deriche

Modélisation du problème

- Recherche d'un filtre de convolution f qui optimise les critères :
- critère de détection --> robustesse au bruit ,
- Critère de précision --> bonne localisation ,
- Unicité de la réponse -->non duplicité des contours

Détection de contours : Modèle de Canny-Deriche

- Modèle:
 - Un contour est un seuil d'amplitude a auquel on ajoute un bruit gaussien n de moyenne E=0

et de variance σ 02, ce qui donne :

$$\rightarrow$$
 I(x) = aU(x) + n(x)

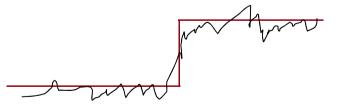
où U(x) est la fonction de heaviside unité :

$$U(x) = 0 \text{ si } x < 0$$

$$U(x) = 1 \text{ si } x \ge 0$$

Le contour effectif est situé en x = 0

- Modèle préconisé par Canny :
 - ⇒ Modèle : contour en escalier (saut quantitatif)



Contour idéal
Contour réel

critère de détection : robustesse au bruit

f antisymétrique $\Rightarrow f(t) = -f(-t)$

$$I_1(x_0) = (\int_{-\infty}^{+\infty} I(t).f(-t)dt = (\int_{-\infty}^{+\infty} (aU(t) + n(t)).f(-t)dt$$

$$O_{s}(0) = a \int_{0}^{\infty} f(-t)dt = (-a \int_{0}^{\infty} f(t)dt) = (a \int_{-\infty}^{0} f(t)dt)$$

Rép. bruit :

$$O_{n}(0) = \left(a \int_{-\infty}^{+\infty} f(-t)n(t)dt\right)$$

⇒ Robustesse au bruit (suite)

Bruit Blanc stationnaire:

$$E(n(x)) = 0$$
; Variance = $\sigma^2(n)$

Critère de détection : rapport (signal/bruit)

$$E(n(x))=0$$
; Variance $=\sigma^2$

Le bruit est caractérisé par son écart type



Robustesse ≡ réponse signal / écart-type bruit

Robustesse au bruit (suite)

$$C_{1} = \frac{Signal}{Bruit} = \frac{a}{\sigma} \frac{-\infty}{\sqrt{\int_{-\infty}^{\infty} f^{2}(t)dt}}$$

or
$$\frac{a}{\sigma}$$
 est constante

Le filtre recherché doit maximiser ce rapport



$$C_{1} = \frac{\int_{-\infty}^{0} f(t)dt}{\left(\int_{-\infty}^{\infty} f^{2}(t)dt\right)^{\frac{1}{2}}}$$

Critère de localisation :

Idée: limiter le déplacement par l'opération du lissage

Le point contour doit être aussi près que possible de la position théorique donc proche du point (0,0)

<u>Possibilité</u>: optimiser l'écart type de la position des différents points possibles



Maximiser l'inverse de l'espérance de la distance entre le point théorique et le point estimé.

Critère de localisation :

$$C = \frac{a}{\sigma} \frac{|f'(0)|}{\left(\int_{-\infty}^{\infty} f'^{2}(x) dx\right)^{\frac{1}{2}}}$$

$$C_{2} = \frac{|f'(0)|}{(\int_{-\infty}^{\infty} f'^{2}(x) dx)^{\frac{1}{2}}}$$

Ce critère peut être retrouvé en étudiant

En étudiant le « déplacement » du maximum de la réponse en se plaçant en $x_0 = 0$

$$Max(I(x)) \rightarrow I'(X)$$

$$I'(x) = \frac{d}{dx} \int_{-\infty}^{+\infty} f(x)I(x-t)dt = 0$$

$$I'(x) = \int_{-\infty}^{+\infty} f'(x)I(x-t)dt = 0$$
Deux composante s:
$$\begin{cases} signal \\ bruit \end{cases}$$

• Critère d'unicité

Le filtre doit fournir une réponse unique à un contour

Ce qui peut se traduire par : si (x_{max}) la réponse du filtre f au bruit alors :

f doit être tel que : (xmax) est une fraction de la taille du support du filtre -> (x_{max}) = kW (support de f = [--W, W])

Bruit blanc stationnaire donc

$$x_{\text{max}} = kW = 2\Pi \left(\frac{\int_{-\infty}^{\infty} f'^{2}(x) dx}{\int_{-\infty}^{\infty} f''^{2}(x) dx} \right)^{\frac{1}{2}}$$
 (Rice 1944)

• Filtre de Canny

En combinant les 3 critère, Canny propose un filtre f qui maximise le produit :

$$\Omega = C_{1}C_{2}$$

Sous la contrainte de l'unicité : $(x_{max}) = kW$

Problème variationnel -> Euler-Lagrange

Qui se traduit par une équation différentielle dont la solution est

$$f(x) = a_1 e^{\alpha x} \sin \omega x + a_2 e^{\alpha x} \cos \omega x$$
$$+ a_3 e^{-\alpha x} \sin \omega x + a_4 e^{-\alpha x} \cos \omega x + c$$

Avec les conditions aux limites :

$$f(0)=0$$
; $f(W)=0$; $f'(0)=s$; $f'(W)=0$

Filtre de Rachid Deriche

Généralisation du filtre de Canny,

f est de support infini (w $\rightarrow \infty$)

On obtient :
$$a_1 = a_2 = a_3 = a_4 = c = 0$$

Problème variationnel -> Euler-Lagrange

Qui se traduit par une équation différentielle

$$f(x) = se^{\alpha|x|} \sin \omega x$$

Avec a qui définit la longueur du filtre : compromis (détection-localisation)

Détection de contours : Modèle de Canny-Deriche

Modèle :

 Un contour est un seuil d'amplitude a auquel on ajoute un bruit gaussien n de moyenne E = 0

et de variance σ 02, ce qui donne :

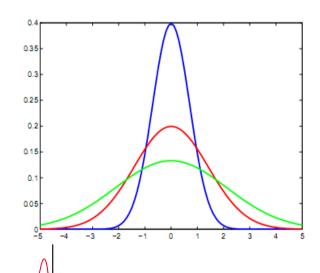
$$\rightarrow$$
 I(x) = aU(x) + n(x)

où U(x) est la fonction de heaviside unité :

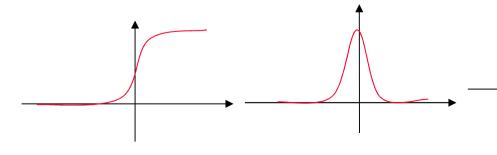
$$U(x) = 0 \text{ si } x < 0$$

$$U(x) = 1 \text{ si } x \ge 0$$

Le contour effectif est situé en x = 0



61



Revenir plus tard sur les contours actifs (level sets) Chapitre à venir