

Méthodologies en Imagerie Médicale

Méthodes de recalage
Dominique.Bereziat@lip6.fr
LIP6 - UPMC

25 novembre 2008

Table des matières

1	Introduction	1
1.1	Problématique	1
1.2	Classification des méthodes	2
1.3	Conclusion	4
2	Transformations géométriques	4
2.1	Transformations rigides	5
2.2	Transformations non rigides	7
2.3	Transformations projectives	8
2.4	Rectification	10
3	Principes généraux	11
3.1	Méthodes basées sur les points	11
3.2	Méthodes basées sur les surfaces	13
3.3	Méthodes basées sur les luminances	16
4	Quelques applications	21
4.1	Recalage ponctuel 2D ou 3D : ICP	21
4.2	Mise en correspondance de cartes 3D	24
4.3	Distances tangentes	27
4.4	Recalage 3D-2D : RANSAC	29
4.5	Recalage non affine : mise en correspondance de courbes	32

1 Introduction

1.1 Problématique

Problématique

- Comment comparer deux acquisitions d'une même scène :
- à deux instant différents?

- à deux points de vue différents ?
- avec deux capteurs différents ?
- Ce problème concerne l'ensemble des acquisitions médicales :
 - X,CT-X
 - US,
 - IRM,
 - PET, SPECT,
 - MEG, EEG,
 - simple données vidéos (laparoscopie, endoscopie, ...).
- Les méthodes de recalages diffèrent aussi selon le type d'imagerie et les problématiques : aspects très médicaux.
- Autre secteur applicatif : télédétection \Rightarrow Système d'Information Géographique (SIG).

Un problème très dépendant du contexte

Intra-modal	Inter-modal	Inter-modal (suite)
CT	CT-MR	X - CT
MR	CT-PET	X - MR
PET	CT-SPECT	X - US
SPECT	PET-MR	
US	PET-US	Modality to model
Video	SPECT-MR	CT
X	US - CT	MR
	US - MR	SPECT
	Video - CT	X
	Video - MR	

Tiré de [Maintz and Viergever, 1998].

1.2 Classification des méthodes

Classification des méthodes

- 8 catégories (Elsen, Pol, Viergever, 1993) :
 1. Dimensionnalité des images.
 2. Base de recalage.
 3. Transformation géométrique.
 4. Degré d'interactivité.
 5. Nature/modalité des images.
 6. Méthode d'optimisation.
 7. Sujet.
 8. Objet.
- Dimensionnalité :
 - 2D/2D, 3D/3D : cas les plus habituels (3D pour la tomographie)

- 2D/3D : intervient surtout dans un cadre bi-modal : exemple recalage une acquisition X durant l'opération avec une acquisition CT-X *pre* opératoire.
- Séries temporelles : on parle ici de suivi ou surveillance (développement de tumeur par exemple) *pre* ou *post* opératoire.
- Base de recalage :
 - Extrinsic : objets artificiels attachés au patient.
 - des points référents fixés rigidement sur le patient,
 - certaines méthodes sont invasives et d'autres non.
 - Intrinsic : des invariants dans les images.
 - Point référents (*amers*) anatomiques ou géométriques.
 - Des surfaces référentes (issues de segmentation).
 - Luminescence des voxels.
- Type de transformation : Quelle transformation pour quelle problème? Une description exhaustive suit.
- Degré d'interactivité :
 - automatique : cas idéal.
 - interactive : le recalage est fait manuellement (assisté par l'ordinateur).
 - semi-automatique : généralement, c'est initialisation du procédé de recalage : choix des points référents ou des segmentations.
- Nature/modalité des images :
 - recalage intra-modal : entre acquisitions de même nature, intervient lors d'acquisition distante en temps.
 - recalage inter-modal : entre acquisitions de différente nature (X : radiance rayon X, IRM : irradiance RF, US : impédance acoustique, PET : irradiance X, ...).
 - Comment comparer des acquisition de nature différentes?
 - Topologie des structures observées différentes.
 - Résolutions spatio-temporel des données différentes.
 - Mesures radiométriques différentes.
- ⇒ Le recalage inter-modal est fondamentale en imagerie médicale pour enrichir les informations :
 - Exemple 1 : planification de radio-dose. (CT-)X identifie les tissus qui absorbe les rayons γ ; IMR permet de localiser la tumeur.
 - Exemple 2 : imagerie cérébrale. IRM fournit une description anatomique des structures. PET fournit une description fonctionnelle du cerveau.
 - recalage modal/modèle ou modal/patient : lorsqu'on dispose d'un modèle (un atlas) généraliste ou du patient.
- Méthode optimisation : les paramètres du recalage (de la transformation) sont calculées soit :
 - directement,
 - par minimisation d'un critère d'erreur.
- Sujet : désigne le fait si le sujet est le même sur chaque acquisition. Dans le cas contraire, on cherche à réaliser des modèles des patients : on parle d'atlas.
- Objet : quelle structure fonctionnelle (organe, tissu, ...) est concernée par

le recalage? Les procédés de recalage et de la calibration varient selon le type d'objet. Exemple : on peut poser des structures référentes sur le crâne d'un patient, pas sur le cœur ...

1.3 Conclusion

Conclusion

- Le recalage est une opération fondamentale en imagerie médicale.
- La méthodologie dépend fortement de la problématique médicale, du type d'acquisition, de la structure fonctionnelle ciblée.
- Beaucoup des publications en imagerie médicale concernent cette problématique dans un cas souvent particulier.
- Pour une vue plus médicale :
 - Handbook of Medical Imaging (volume 2) [Sonka and Fitzpatrick, 2000],
 - A Survey of Medical Image Registration [Maintz and Viergever, 1998].
- Dans ce cours : on ne s'intéresse qu'aux aspects méthodologiques du recalage.

2 Transformations géométriques

Transformations géométriques

- Soit Ω le domaine des images, on a $\Omega \subset \mathbb{R}^3$ (on peut aussi se restreindre à \mathbb{R}^2)
- Soit une application telle que :

$$\begin{aligned}\tau : \Omega &\rightarrow \Omega \\ \mathbf{x} &\mapsto \mathbf{x}'\end{aligned}$$

- Soit deux images I_1 et I_2 définies sur Ω
- On a : $I_2(\tau(\mathbf{x})) = I_1(\mathbf{x})$
- Inventaire des transformations géométriques τ possibles

Rappels : transformations affines et linéaires

Definition 2.1 (Transformations linéaires). Soit f telle que $f : \Omega \rightarrow \Omega$. f est linéaire si :

- $\forall (x, y) \in \Omega \times \Omega, f(x + y) = f(x) + f(y)$
- $\forall x \in \Omega, a \in \mathbb{R}, f(ax) = af(x)$

- Les rotations, symétries et changements d'échelles sont des transformations linéaires. Matriciellement : $\mathbf{x}' = M\mathbf{x}$

Definition 2.2 (Transformations affines). Soit f telle que $f : \Omega \rightarrow \Omega$. f est affine si $\forall x \in \Omega, f(x) - f(0)$ est linéaire.

- les transformations linéaires et les translations sont des transformations affines. Matriciellement : $\mathbf{x}' = M\mathbf{x} + \mathbf{t}$

2.1 Transformations rigides

Transformations rigides

Definition 2.3 (Transformations rigides). Les transformations qui préservent les distances sont qualifiées de rigides.

- Ces applications préservent les angles, les lignes droites.
- Ces applications sont linéaires.
- Elles s'écrivent de manière générale :

$$\mathbf{x}' = R\mathbf{x} + \mathbf{t}$$

avec :

- R une matrice orthogonale
- $\mathbf{t} = (\mathbf{t}_x, \mathbf{t}_y, \mathbf{t}_z)^t$ un vecteur de translation

Definition 2.4 (Matrice orthogonale). Rappel : une matrice est orthogonale si $R^t R = I$ ou si $R^{-1} = R^t$ ou si $\det(R) = 1, -1$

- Les rotations et les symétries planaires composent ce groupe de transformations orthogonales.
- Les symétries (correspondant à $\det(R) = -1$) ne sont pas des transformations *a priori* utile en recalage.
- Reste les rotations :

Théorème 1 (Rotations). R est une rotation ssi $\det(R) = 1$ et $RR^t = I$

Complément sur les rotations

- Dans \mathbb{R}^3 : un axe, un angle.
- Dans \mathbb{R}^2 :

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

- Dans \mathbb{R}^3 : selon l'axe $(0, z)$, les points de l'axes sont invariants.

$$R_z = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Pour les autres axes :

$$R_y = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}, R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

- En composant les trois rotations : $R_z \circ R_y \circ R_x$, on peut obtenir une rotation pour un axe quelconque.
- Connaissant l'axe, difficulté de la composition.

- Pour un axe δ quelconque connu, on a :

$$R_{\delta,\theta} = \begin{pmatrix} \delta_x^2 V + C & \delta_x \delta_y V - \delta_z S & \delta_x \delta_z V + \delta_y S \\ \delta_x \delta_y V + \delta_z S & \delta_y^2 V + C & \delta_y \delta_z V - \delta_x S \\ \delta_x \delta_z V - \delta_y S & \delta_y \delta_z V + \delta_x S & \delta_z^2 V + C \end{pmatrix}$$

avec $S = \sin \theta, C = \cos \theta, V = 1 - \cos \theta$

- Réciproquement, on se donne une matrice M telle que $\det(M) = 1$
- Donnez une expression à partir de M qui fournisse l'angle et l'axe de la rotation ...
- On a : $\text{trace}(M) = V + 3C = 1 + 2 \cos \theta$
- On a :

$$R - R^t = \begin{pmatrix} 0 & -2S\delta_z & 2S\delta_y \\ 2S\delta_z & 0 & -2S\delta_x \\ -2S\delta_y & 2S\delta_x & 0 \end{pmatrix}$$

Quaternions

- Ils servent à manipuler les rotations en 3D.
- Généralisation des nombres complexes sur \mathbb{R}^3 .
- Analogie dans le plan : paramétrage des rotations planes grâce aux nombres complexes de module 1 :

$$e^{i\theta} \text{ rotation d'angle } \theta$$

$$e^{i\theta} \cdot e^{i\alpha} = e^{i(\theta+\alpha)}$$

Definition 2.5.

$$q = c + xi + yj + zk \in \mathbb{R}^4$$

$$q = \text{Re}(q) + \text{Im}(q) \in \mathbb{R} \oplus \mathbb{R}^3$$

$$q = c + u \in \mathbb{R} \oplus \mathbb{R}^3$$

Arithmétique des quaternions

- Addition :

$$q + q' = (c + c') + (u + u')$$

- Multiplication :

$$q \times q' = (cc' - u \cdot u') + (u \wedge u' + cu' + c'u)$$

. et \wedge sont respectivement les produits scalaire et vectoriel dans \mathbb{R}^3

- $(\mathbb{Q}, +, \times)$ est un corps non commutatif (comme les complexes).

Propriétés des nombres i, j, k

- Par analogie des nombres complexes, on a :

$$i^2 = j^2 = k^2 = -1$$

- Plus les relations entre i, j, k :

$$\begin{cases} ij = -ji = k \\ jk = -kj = i \\ ki = -ik = j \end{cases}$$

Autres opérations sur les quaternions

- Conjugué de q :

$$\bar{q} = c - u$$

- module de q :

$$q\bar{q} = \bar{q}q = \|q\|^2$$

- Inverse d'un quaternion (si $q \neq 0$) :

$$q^{-1} = \frac{\bar{q}}{\|q\|^2}$$

- En général :

$$q_1 q_2 \neq q_2 q_1$$

Quaternions et rotations

- Les quaternions de norme 1 forme un groupe dont les éléments ont pour conjugué leur inverse.
- Soient $p = (O, v)$ un quaternion pur et $q = (c, u)$ un quaternion unitaire. On a :

$$q = (c, u) = (\cos \theta, \sin \theta \vec{N})$$

et

$$R_q(p) = qp\bar{q}$$

- R est une rotation d'angle 2θ et d'axe $\vec{N} = N_1 i + N_2 j + N_3 k$
- Propriété : $R_{qq'} = R_q \circ R_{q'}$

2.2 Transformations non rigides

Transformations non rigides

- Parmi les transformations affines, les transformations non rigides sont celles qui restent.
- La transformation de changement d'échelle :

$$\mathbf{x}' = S\mathbf{x}, S = \begin{pmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & S_z \end{pmatrix}$$

- S_x, S_y, S_z facteurs d'échelles
- Les transformations affines non-rigides s'écrivent donc :

$$\mathbf{x}' = RS\mathbf{x} + \mathbf{t}$$

avec :

- R rotation,
- S changement d'échelle, t translation.

$$\mathbf{x}' = A\mathbf{x} + \mathbf{t}$$

avec :

- A inversible,
- t translation.

Coordonnées homogènes

- les applications affines de \mathbb{R}^3 ne peuvent pas s'exprimer matriciellement avec des matrices 3×3
- $\mathbf{x}' = \mathbf{x} + \mathbf{t} \neq Ax$
- Introduisons le système suivant : à un point de l'espace \mathbb{R}^3 on associe le quadruplet (x, y, z, w) tel que $\forall w \neq 0$

$$(x, y, z, w) \Leftrightarrow \left(\frac{x}{w}, \frac{y}{w}, \frac{z}{w} \right)$$

- Comment s'exprime la translation dans ce système ?

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

- Toutes les applications linéaires de \mathbb{R}^3 s'expriment matriciellement dans \mathbb{R}^3
- Toutes les applications affines de \mathbb{R}^3 s'expriment matriciellement dans \mathbb{R}^4
- Les applications linéaires L ont pour matrice en coordonnées homogènes :

$$M = \begin{pmatrix} L & 0 \\ 0 & 1 \end{pmatrix}$$

- Les applications affines A ont pour matrice :

$$A = \begin{pmatrix} L & T \\ 0 & 1 \end{pmatrix}, T \text{ translation, } L \text{ composante linéaire}$$

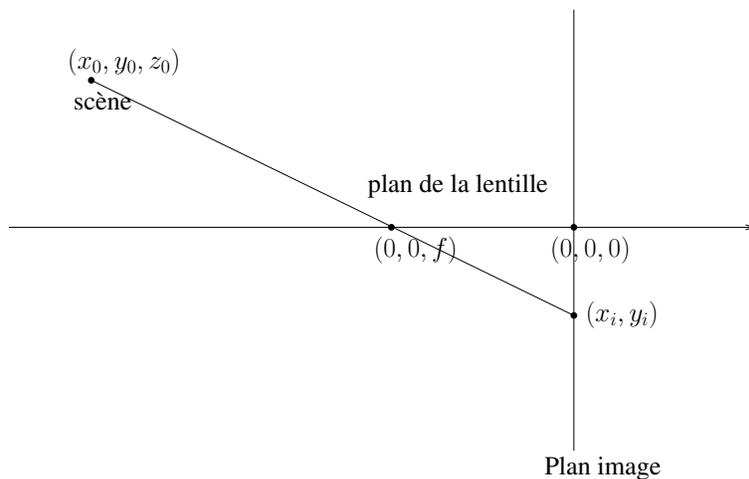
- Il y a d'autres transformations intéressantes qui s'expriment en coordonnées homogènes.

2.3 Transformations projectives

Transformations projectives

- Les projections sont des surjections de \mathbb{R}^n vers un sous-espace de \mathbb{R}^n .
- Les projections ne sont donc pas des applications linéaires (et elles sont non-rigides)
- Contexte image : les projections sont les transformations géométriques de la scène réelle (espace 3D) à l'image (espace 2D).
- Toute image résulte donc d'une projection.
- Deux types de projections :
 - orthogonale (ou cavalière)
 - en perspective.

Projections



Projections

- On a (théorème de Thales) :

$$\begin{cases} x_i = \frac{f}{f-z_0} x_0 \\ y_i = \frac{f}{f-z_0} y_0 \\ z_i = 0 = \frac{0}{f-z_0} \end{cases}$$

- Non linéaire ; fraction de fonction affine
- En coordonnée homogène, on a :

$$\begin{cases} x_i = f x_0 \\ y_i = f y_0 \\ z_i = 0 \\ w_i = f - z_0 \end{cases}$$

Projection perspective en coordonnée homogène

- et donc $(x_i, y_i, z_i, w_i)^t = P(x_0, y_0, z_0, 1)^t$ avec :

$$P = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & f \end{pmatrix}$$

- les applications du type (projections) :

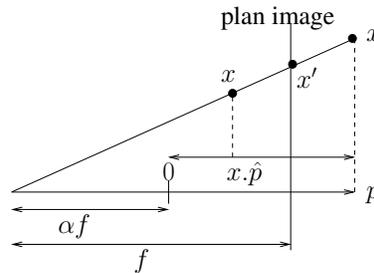
$$\mathbf{x}' = \frac{A\mathbf{x} + \mathbf{t}}{\mathbf{p}^t \mathbf{x} + \alpha}, \quad \mathbf{p}^t = (p_x, p_y, p_z)$$

ont pour matrice homogène :

$$P = \begin{pmatrix} A & \mathbf{t} \\ \mathbf{p} & \alpha \end{pmatrix}$$

Acquisition des images et projection perspective

- Tous les dispositifs d'acquisition en imagerie médicale procède par projection en perspective (exception de l'échographie) : passage du 3D à du 2D.



- $f = \frac{1}{\|p\|}$
- $\hat{p} = \frac{p}{\|p\|}$
- $x' = \frac{fx}{x \cdot \hat{p} + \alpha f}$
- Deux cas :
 - modèle appareil photo : $x \cdot \hat{p} + \alpha f > f$
 - modèle radiographie X : $x \cdot \hat{p} + \alpha f < f$

2.4 Rectification

Rectification

- Imaginons une scène prise à deux points de vue différents :
 - I_1 le plan image correspondant à la première prise, (x_1, y_1) le système de coordonnées associé,
 - I_2 le plan image correspondant à la seconde prise, (x_2, y_2) le système de coordonnées associé,
- Supposons la scène comprise dans un plan.
- Rectification : c'est la transformation géométrique qui permet de passer d'un repère dans une image au repère dans l'autre image.
- On montre que :

$$\begin{cases} x_2 = \frac{a_{11}x_1 + a_{12}y_1 + a_{13}}{a_{31}x_1 + a_{32}y_1 + a_{33}} \\ y_2 = \frac{a_{21}x_1 + a_{22}y_1 + a_{23}}{a_{31}x_1 + a_{32}y_1 + a_{33}} \end{cases}$$

- En pratique :
 - Il faut calculer 9 paramètres : (a_{ij})
 - Identifier des invariants dans les deux images : au moins 5 points.
 - On obtient un système linéaire (de dimension 10) à inverser.

3 Principes généraux

3.1 Méthodes basées sur les points

Méthodes basées sur les points

- Principe : recalculer sur des points “référents”.
- Une fois les points référents identifiés, chercher la transformation géométrique qui “colle” au mieux des points référents.
- La transformation sera recherchée parmi une classe de transformation (on a déjà vu la rectification : transformation en perspective).

Choix des points référents

- des invariants (valable pour certaines structures médicales) :
 - Rarement des points : plutôt des zones identifiables.
 - Difficulté du choix du centre référents dans ces zones
 - Nécessite la supervision par un médecin expert.
- des marqueurs artificiels :
 - Le choix du point référent est indépendant de l’anatomie.
 - On contrôle la taille du marqueur.
 - La détection peut être automatisée et ne nécessite plus d’expertise médicale,
 - exemple : en acquisition bimodale CT/IRM du cerveau, on pose des marqueurs sur le crâne : des capsules remplies de liquide de contraste d’une taille adéquat.

Erreur de localisation des points référents

- Idée : essayer de qualifier en terme d’erreur le choix des points référents.
- \vec{FLE} indique le déplacement entre le point référent choisit et le “vrai” point référent.
- \vec{FLE} : *Fudicial Localization Error*.
- Les erreurs de localisation des points référents induisent *a posteriori* des erreurs de recalage.
- Le \vec{FLE} n’est pas directement observable en général.
- Des erreurs de localisation dans les deux images à recalculer.

Erreur de localisation pour des marqueurs

- La géométrie des marqueurs est connue *a priori*.
- De l’image observée du marqueur, on en déduit un centre qui peut être différent. On peut évaluer le \vec{FLE} .
- Pour localiser le point référent :

$$\mathbf{x} = \frac{\sum_{i=1}^N I_i \mathbf{x}_i}{\sum_{i=1}^N I_i}$$

- $\mathbf{x} = (x, y, z)$
- Peut induire une erreur de localisation.

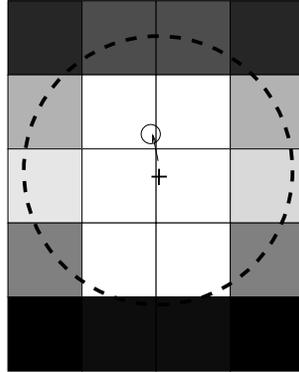


FIG. 1 – Erreur de localisation

Erreur de recalage des points référents

- On définit l’erreur de recalage sur les points référents \mathbf{x}_i d’une image récalée sur les les points référents \mathbf{y}_i d’une seconde image par :

$$\overrightarrow{\text{FRE}}_i = \mathcal{T}(\mathbf{x}_i) - \mathbf{y}_i$$

où \mathcal{T} est la transformation appliqué pour le recalage.

- $\overrightarrow{\text{FRE}} = \text{Fiducial Registration Error}$

- Principe de base : trouver \mathcal{T} qui minimise les $\overrightarrow{\text{FRE}}_i$ pour tous les points

référents. *i.e.* $\underset{\mathcal{T}}{\operatorname{argmin}} \sum_{i=0}^N \|\mathcal{T}(\mathbf{x}_i) - \mathbf{y}_i\|^2$

- Amélioration : $\underset{\mathcal{T}}{\operatorname{argmin}} \sum_{i=0}^N w_i \|\mathcal{T}(\mathbf{x}_i) - \mathbf{y}_i\|^2$

- w_i pondération des termes “sûrs”. Typiquement $w_i = \frac{1}{\|\text{FLE}\|^2}$

Cas des transformations linéaires

- $\mathcal{T}(\mathbf{x}) = A\mathbf{x}$, donc minimiser en A :

$$\sum_{i=0}^N w_i \|A\mathbf{x}_i - \mathbf{y}_i\|^2 \quad (1)$$

- C’est un problème de régression linéaire !
- Une solution exacte est donnée par la méthode des “moindres carrés pondérés” :

Moindres carrés pondérés

L’élément qui minimise en A l’équation 1 est :

$$\hat{A} = \left(\sum_{i=1}^N w_i \mathbf{x}_i^t \mathbf{x}_i \right)^{-1} \sum_{i=1}^N w_i \mathbf{x}_i^t \mathbf{y}_i$$

Cas des transformations affines

- $\mathcal{T}(\mathbf{x}) = A\mathbf{x} + T$, donc minimiser en A :

$$\sum_{i=0}^N w_i \|A\mathbf{x}_i + T - \mathbf{y}_i\|^2 \quad (2)$$

- Astuce : passer en coordonnées homogènes !
- $X_i = \begin{pmatrix} x_i \\ 1 \end{pmatrix}$, $Y_i = \begin{pmatrix} y_i \\ 1 \end{pmatrix}$, $B = \begin{pmatrix} A & T \\ 0 & 1 \end{pmatrix}$
- Minimiser en B :

$$\sum_{i=0}^N w_i \|B\mathbf{X}_i - \mathbf{Y}_i\|^2 \quad (3)$$

Inversion de la matrice

- La matrice peut être mal conditionnée (points trop alignés).
- On peut utiliser une décomposition en valeurs singulières pour avoir un calcul plus stable de l'inverse (voir [Press, 1995]).
- Théorème : toute matrice A telle que peut être écrite sous la forme :

$$A = U \begin{pmatrix} w_1 & & & \\ & w_2 & & \\ & & \dots & \\ & & & w_N \end{pmatrix} V^t \text{ avec } UU^t = VV^t = I$$

- Si A est carré alors U et V sont orthogonales i.e. on a $U^{-1} = U^t$
- Application : $A^{-1} = V \text{diag}(1/w_i)U^t$

Le problème de la correspondance entre les points

- La minimisation de l'équation (1) implique une correspondance *de fait* entre les \mathbf{x}_i et les \mathbf{y}_i .
- En pratique, il n'y a souvent que quelques points éparses. Possibilité de :
 - faire la correspondance manuellement,
 - automatiquement par un critère de voisinage (sans ambiguïté si les points sont bien séparés).
- Si le nombre de point augmente (cas des surfaces) :
 - la mise en correspondance manuellement n'est plus possible (trop lourde),
 - le critère de voisinage peut être ambiguë ...

3.2 Méthodes basées sur les surfaces

Méthodes basées sur les surfaces

- Ces méthodes sont l'extensions des précédentes à un ensemble de points référents plus important.
- On a maintenant des structures référentes :
 - en 2D : des courbes ou des régions,
 - en 3D : des surfaces ou des volumes.

- Ces structures sont issues des images et proviennent :
 - d'algorithmes de segmentation,
 - d'algorithmes de contours actifs ou surfaces actives,
 - ...
- Exemple : en imagerie du crâne/cerveau, on peut extraire le visage comme surface référente.

Mesure de similarité

- De manière similaire aux méthodes ponctuelles, on définit une mesure de similarité entre deux surfaces X, Y :

$$d(\mathcal{T}(X), Y) = \sqrt{\sum_{j=1}^N w_j d(\mathcal{T}(\mathbf{x}_j), Y)^2} = \sqrt{\sum_{j=1}^N w_j \|\mathcal{T}(\mathbf{x}_j) - \mathbf{y}_j\|^2}$$

avec $\mathbf{y}_i = \mathcal{C}(\mathcal{T}(\mathbf{x}_j), Y)$, w_j pondération sur le point j .

- C'est une mesure de distance.
- Comme pour les méthodes ponctuelles : la nécessité de mettre en correspondance \mathbf{x}_i avec \mathbf{y}_i .
- D'autres distances hors norme quadratique peuvent être utilisées.
- Si on prend $d(\mathcal{T}(\mathbf{x}_j), Y)$ la distance de \mathbf{x}_j au point le plus proche sur Y , on économise la mise en correspondance \mathcal{C} .

Minimisation de la mesure de similarité

- Rappel : trouver \mathcal{T} qui minimise $d(\mathcal{T}(X), Y) = \sum_i d(\mathbf{x}_i, Y)$
- Problème de minimisation de fonctionnelle non linéaire.
- Cadre mathématique courant : descente de gradient (impose de calculer la différentielle).
- Une alternative pour la recherche de transformation rigide : réduire la minimisation non linéaire à une minimisation ponctuelle (voir l'algorithme ICP *Iterative Closest Point* plus loin dans ce cours 4.1).

Calcul des distances point/ensemble

- Un ensemble : $A = \{\vec{a}_i | i = 1 \dots N\} \in \Omega$,
- $d(\vec{p}, A) = \min_{i=1 \dots N} d(\vec{p}, \vec{a}_i)$,
- d = une distance quelconque (i.e. qui vérifie $d(x, x) = 0$, $d(x, y) > 0$ et l'inégalité triangulaire),
- distance point/segment de droite :

$$d(\vec{p}, [\vec{r}_1, \vec{r}_2]) = \min_{u+v=1} \|u\vec{r}_1 + v\vec{r}_2 - \vec{p}\|$$

- distance point/ensemble de segments de droite :

$$L = \{[\vec{r}_1^i, \vec{r}_2^i]\}_{i=1 \dots N}$$

$$d(\vec{p}, L) = \min_i d(\vec{p}, [\vec{r}_1^i, \vec{r}_2^i])$$

- point/triangle :

$$d(\vec{p}, [\vec{r}_1, \vec{r}_2], \vec{r}_3) = \min_{u+v+w=1} \|u\vec{r}_1 + v\vec{r}_2 + w\vec{r}_3 - \vec{p}\|$$

- point/ensemble de triangles : $T = \{t_i\}, d(\vec{p}, T) = \min_i d(\vec{p}, t_i)$
- point/courbe : on approxime la courbe par un ensemble de segments de droite pour se ramener au cas "distance point/ensemble de segment de droite".
- point/surface : on approxime la surface par un ensemble de triangles pour se ramener au cas "distance point/ensemble de triangles".

Un mot sur les méthodologies de résolution

- Soit $f(\vec{u}) = \|\vec{r}(\vec{u}) - \vec{p}\|^2$
- $\min f(\vec{u}) \Leftrightarrow \frac{\partial f}{\partial \vec{u}} = 0$
- Pour une fonction f conventionnelle, $\frac{\partial f}{\partial \vec{u}} = \nabla f = (\frac{\partial f}{\partial x_i}, i = 1, 2, \dots)^T$
- Si f est linéaire (cas des transformations affines/rigides) ∇f est constant \Rightarrow inversion directe du système ou moindre carré.
- Si f est non linéaire et deux fois dérivable : on peut utiliser la méthode de Newton (convergence exponentielle si condition initiale).

Méthode de Newton (généralisée à \mathbb{R}^N)

- On définit la matrice *Hessienne* de f :

$$\nabla \nabla^T(f) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

- Alors, le schéma numérique suivant :

$$\vec{u}_{k+1} = \vec{u}_k - [\nabla \nabla^T(f)(\vec{u}_k)]^{-1} \nabla f(\vec{u}_k)$$

converge à vitesse exponentielle vers une solution de $\nabla f = 0$ (dans un voisinage de la racine \vec{a} et si $\nabla f(\vec{a}) \neq 0$).

- Remarque : la méthode de Newton est l'application du théorème du point fixe à la fonction $\vec{u} \mapsto \vec{u} - (\nabla f)^{-1}(\vec{u}) \nabla f(\vec{u})$.

Courbe/surface implicite

- Très souvent, courbes et surfaces peuvent être définies implicitement :

$$\mathcal{S} = \{\vec{r} \in \mathbb{R}^3 | g(\vec{r}) = 0\}$$

(voir par exemple les contours actifs géodésiques).

- Distance d'un point \vec{p} à une courbe implicite $g = 0$?

$$\operatorname{argmin}_{\vec{r} | g(\vec{r})=0} \|\vec{p} - \vec{r}\|^2 \tag{4}$$

- \Rightarrow utilisation du multiplicateur de Lagrange.

Multiplicateur de Lagrange

Proposition 1. *Le problème d'optimisation :*

$$\min_{g(u)=0} f(u) \quad (5)$$

est équivalent à résoudre :

$$\begin{cases} \frac{\partial}{\partial \vec{u}} \left(f(\vec{u}) + \sum_{i=1}^N \lambda_i g(\vec{u}) \right) = 0 \\ \frac{\partial}{\partial \lambda_i} \left(f(\vec{u}) + \sum_{i=1}^N \lambda_i g(\vec{u}) \right) = 0 \end{cases} \quad (6)$$

où f est une fonction de \mathbb{R}^N .

- Les variables auxiliaires λ_i sont appelées les multiplicateurs de Lagrange.
- L'équation (6) se simplifie :

$$\begin{cases} \nabla f(\vec{u}) + \sum_i \lambda_i \nabla g(\vec{u}) = 0 \\ \frac{\partial g}{\partial \lambda_i}(\vec{u}) = 0, i = 1 \dots N \end{cases}$$

- À nouveau, on peut utiliser, selon f et g , des moindres carrés ou une méthode de Newton pour résoudre ce système.

3.3 Méthodes basées sur les luminances

Méthodes basées sur les luminances Le point à ce niveau du cours

- On a discuté de ce qu'il faut comparer dans les images (point, courbe, surface) et comment on les compare (transformation géométrique).
- On s'est placé dans un espace de représentation (point, courbe, ...) indépendant des modalités d'acquisition.
- Pour travailler directement sur les luminances (l'information radiométrique) des pixels, il faut pouvoir les comparer.
- Il faut donc définir une notion de *similarité* entre deux images.

Cas du recalage intra-modal

- Entre acquisition de même modalité, on peut comparer plus facilement les radiométries (parfois, une correction peut être nécessaire).
- Cas de figure très similaire à l'estimation du mouvement (voir ACIMED) :
 - Méthodes fréquentielles (par T.F. et consorts);
 - Block Matching;
 - Le flot optique (dont Lucas-Kanade).
- Des différences subsistent :

- Les phénomènes qui induisent le transport de la luminosité dûs au recalage sont différents de ceux du flot optique car non induit par des phénomènes liés à une dynamique temporelle. (Une exception : recalage en imagerie cardiaque ou de tissus déformé par l'activité cardiaque!).
- Autre différence : recalage 3D/2D.
- Contrainte forte sur le champ de déformation : pour le recalage, c'est un difféomorphisme.

Recalage intra-modal : mesure de similarité par SSD

- Soit \mathcal{T} , la transformation de recalage recherchée.
- Soient A et B , la paire d'image à recaler.
- Trouver \mathcal{T} qui minimise la somme des différences au carré (SSD) :

$$SSD = \frac{1}{N} \sum_i (\phi_1(A(\mathcal{T}(i))) - \phi_2(B(i)))^2$$

ϕ_i , sont des corrections radiométrique pour ramener A et B dans la même métrique.

- Ce type de similarité permet de gérer les changements d'illumination globale (il est optimal), avec ϕ_i bien choisis.

Intra-modal : critère d'uniformité du ratio d'image

- Pour deux régions A et B , trouver la transformation qui minimise le RIU (Ratio Image Uniformity).
- Calcul du RIU :
 - $I = A \cap B'$, $N = |I|$, $B' = \mathcal{T}(B)$
 - On calcule le rapport des images : $R(i) = \frac{B(i)}{A(i)}$, $\forall i \in I$
 - Avec $\mu_R = \frac{1}{N} \sum_{i \in I} R(i)$ et $\sigma_R = \frac{1}{N} \sum_{i \in I} (R(i) - \mu_i)^2$, on définit : $RIU = \frac{\sigma_R}{\mu_R}$
- Minimisation de $\sigma_R \Rightarrow$ favorise les plus grandes régions pour I .
- Permet de gérer les changements globales d'illumination : $\mu_R \neq 1$, mais le critère minimise σ_R .
- On peut échanger les rôles de A et B dans R pour éviter des valeurs trop grandes pour R .
- Utilisé en recalage PET-PET et IRM-IRM

Inter-modal : mesure de similarité par partitionnement uniforme

- Une première mesure de similarité inter-modalité (avec des valeurs radiométriques de nature différentes).
- Hypothèses :
 - Chaque tissu est caractérisé par une valeur de n.d.g.,
 - ces valeurs sont conservées sur les deux images à recaler.
- Sur l'image A , on trouve $n_A(a)$ voxel de couleurs a .

- Sur l'image $B' = \mathcal{T}(B)$, on doit trouver idéalement $n_A(a)$ voxels de couleur a , localisés aux mêmes endroits que sur l'image A .
- Considérons la mesure : $PUI = \sum_{a \in \mathcal{A}} \frac{n_A(a)\sigma_{B'}(a)}{N\mu_{B'}(a)}$ avec $N = |A \cap B'|$,
 $\mu_{B'}(a)$ moyenne des voxels de B' co-localisés avec les voxels de valeur a dans A (idem avec $\sigma_{B'}$ en écart-type).
- Utilisée en recalage IRM-PET.

Recalage inter-modal : mesure de similarité par corrélation

- Principe : rechercher des structures spatialement corrélées dans les paires d'images.
- Trouver \mathcal{T} qui minimise la corrélation :

$$\text{Corr} = \frac{\sum_i (A'(i) - \bar{A}')(B(i) - \bar{B}(i))}{(\sum_i (A'(i) - \bar{A}')^2 \sum_i (B(i) - \bar{B})^2)^{\frac{1}{2}}}$$

- avec $A' = A \circ \mathcal{T}$, $\bar{A}' = \sum_i A'(i)$, $\bar{B} = \sum_i B(i)$.
- S'il existe une relation linéaire entre A et B (i.e. si \mathcal{T} est linéaire) ce type de similarité est optimal.

Méthodes basées sur l'analyse des histogrammes

- Il s'agit de trouver la transformation \mathcal{T} qui aligne au mieux les histogrammes des images (après transformation).
- Critère global bien adapté au recalage rigide.
- Histogramme normalisé \Leftrightarrow densité de probabilité.
- Comment comparer deux histogrammes ?
 - Tests statistiques (si les lois sont connus a priori).
 - Théorie de l'information.
- Idée initiale (théo. info.) : utiliser le principe

$$X \perp Y \quad \Leftrightarrow \quad \mathbb{P}(X, Y) = \mathbb{P}(X)\mathbb{P}(Y)$$

- Interpréter des histogrammes 2D n'est pas chose aisée (voir transparent Exemple) : trouver un critère simple (un scalaire).

Comparaison des lois

- Étude de la loi jointe : on peut calculer l'histogramme joint d'une paire d'image :

$$\begin{aligned} I &: \Omega \rightarrow [0, \dots, n] \\ J &: \Omega \rightarrow [0, \dots, m] \\ \text{Hist}(I, J)(i, j) &= \sum_{x \in \Omega} \mathbb{1}_{I(x)=i, J(x)=j} \end{aligned} \quad (7)$$

- On détermine les fréquences d'apparitions des couples (i, j) dans la paire d'images (I, J) .

- Si $I = J$, alors $\mathbb{1}_{I(x)=i, I(x)=j} = 1 \Leftrightarrow i = j$. $Hist(I, I)$ est une matrice diagonale, laquelle est l'histogramme de I .
- I et J doivent être définis sur le même ensemble. Cas inter-modal : on peut être amené à rééchantillonner les images (si les résolutions diffèrent).

Information mutuelle

- Information mutuelle (\mathbb{I}) : mesure de la distance entre la densité de la loi jointe p_{XY} et le produit des lois de X et Y , $p_X p_Y$:

$$\mathbb{I}(X, Y) = \mathbb{E} \left[\frac{\mathbb{P}_{X,Y}}{\mathbb{P}_X \mathbb{P}_Y} \right] = \sum_{x,y} \log \frac{\mathbb{P}_{XY}(X=x, Y=y)}{\mathbb{P}_X(X=x) \mathbb{P}_Y(Y=y)}$$

- Cette métrique s'appelle la distance de Kullback-Leibler.

Proposition 2. \mathbb{I} est reliée à l'entropie de X , Y et (X, Y)

- En effet :

$$\begin{aligned} \mathbb{I}(X, Y) &= \sum_{x,y} \mathbb{P}_{XY}(x, y) (\log \mathbb{P}_{XY}(x, y) - \log \mathbb{P}_X(x) - \log \mathbb{P}_Y(y)) \\ &= \sum_{x,y} \mathbb{P}_{XY}(x, y) \log \mathbb{P}_{XY}(x, y) - \sum_x \mathbb{P}_X(x) \log \mathbb{P}_X(x) \\ &\quad - \sum_y \mathbb{P}_Y(y) \log \mathbb{P}_Y(y) \end{aligned}$$

- Rappel : $\mathbb{P}_X(x) = \sum_y \mathbb{P}_{X,Y}(x, y)$ par définition des lois marginales.
- Finalement on a :

$$\mathbb{I}(X, Y) = -\mathbb{H}(X, Y) + \mathbb{H}(X) + \mathbb{H}(Y)$$

où \mathbb{H} est l'entropie (d'une variable aléatoire).

- Autres propriétés :
 - $\mathbb{I}(X, Y) = 0 \Leftrightarrow X \perp Y$,
 - $\mathbb{I}(X, Y) \geq 0$,
 - $\mathbb{I}(X, Y) = \mathbb{I}(Y, X)$, $\mathbb{I}(X, X) = \mathbb{H}(X)$
- Théorème du traitement des données :

$$\forall f, g \quad \mathbb{I}(f(X), g(Y)) \leq \mathbb{I}(X, Y)$$

Le traitement des données n'apporte pas d'information supplémentaire.

Entropie

- Entropie (Shannon) : mesure de la "quantité d'information" absente dans une variable aléatoire.
- Définition :

$$\mathbb{H}(X) = -\mathbb{E}_X [\log (\mathbb{P}(X))]$$

- V.A. continue :

$$\mathbb{H}(X) = - \int_{\Omega} \log(p(x)) p(x) dx$$

avec p densité de la loi de X .

- V.A. discrète :

$$\mathbb{H}(X) = - \sum_x \log(\mathbb{P}(X = x))\mathbb{P}(X = x)$$

- Remarque : on peut utiliser le logarithme à base 2 (utile pour mesurer la quantité d'information en nombre de bits pour une donnée numérisée).
- Entropie conditionnelle :

$$\mathbb{H}(X|Y) = \mathbb{E}_{X|Y} \log(\mathbb{P}(X|Y)) = - \sum_x \mathbb{P}_X(x) \log \mathbb{P}_X(x)$$

$\mathbb{H}(X|Y)$ mesure l'information restante dans X , connaissant Y .

- Rappel (probabilité conditionnelle) : la formule

$$\mathbb{P}(X = x|Y = y) = \frac{\mathbb{P}_{XY}(x, y)}{\mathbb{P}_Y(y)}$$

permet de calculer en pratique la probabilité conditionnelle à partir des lois (X, Y) et X .

- Entropie jointe : $\mathbb{H}(X, Y) = -\mathbb{E} \log \mathbb{P}(X, Y)$.
- Attention : on ne peut évaluer en 0 l'application $x \mapsto x \log(x)$ (à cause du logarithme). \Rightarrow prolongement par continuité : poser $x \log(x) = 0$ pour $0 \leq x < \epsilon$.
- Propriétés de l'entropie :
 - $\mathbb{H}(X) \geq 0$,
 - $\mathbb{H}(X) = 0$ si $X = \delta$: toute l'information est concentrée en 1 point.
 - $\mathbb{H}(X)$ est maximal si $X \sim \mathcal{U}([a, b])$: tout l'information est dispersée. L'entropie croît avec la longueur de l'intervalle $[a, b]$ (i.e. $b - a$).
 - $X \perp Y \Rightarrow \mathbb{H}(X, Y) = \mathbb{H}(X) + \mathbb{H}(Y)$.
 - $\forall X, Y \quad \mathbb{H}(X, Y) = \mathbb{H}(X|Y) + \mathbb{H}(Y) = \mathbb{H}(Y|X) + \mathbb{H}(X)$

Retour sur l'information mutuelle

- Ainsi, on a aussi :

$$\mathbb{I}(X, Y) = \mathbb{H}(X) - \mathbb{H}(X|Y) = \mathbb{H}(Y) - \mathbb{H}(Y|X)$$

- L'information mutuelle est nulle lorsque l'information contenu à la fois par X et par Y est vide.
- L'information mutuelle est maximale lorsque $X = Y$.
- \Rightarrow on cherche donc à maximiser l'information mutuelle de façon à aligner les images.
- L'entropie répond fortement sur les Dirac. L'information mutuelle va répondre fortement sur les Dirac localisés aux mêmes endroits dans la paire d'images et cela indépendamment de la signification physique des valeurs des pixels.

Application au recalage

- Construire l'histogramme du couple (I, J) (images à recaler (équation (7)).
- Normalisation de l'histogramme, calcul des deux lois marginales.
- Méthode type block-matching avec l'I.M. entre l'image I et l'image J composée avec la transformation τ à maximiser en τ selon le critère d'information mutuelle maximale.
- Méthode assez coûteuse car le calcul de l'histogramme est en $\mathcal{O}(n)$, $n =$ nombre de pixels. (Coût similaire avec une covariance).
- Articles à lire pour approfondir le sujet : [Viola and Wells, 1997], [Maes et al., 1997]

Exemple en recalage MR/PET

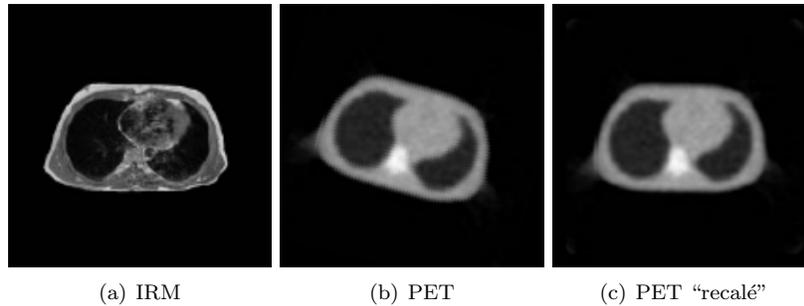


FIG. 2 – Recalage MR/PET

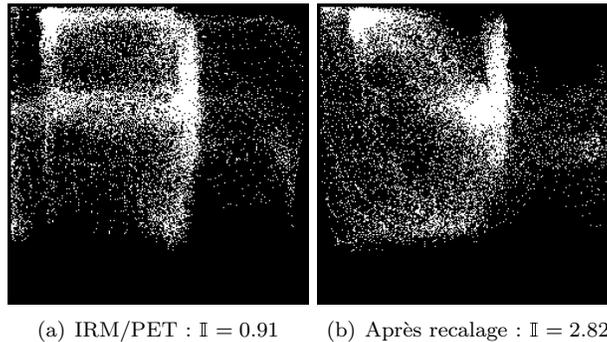


FIG. 3 – Histogrammes joints et \mathbb{I} associé.

4 Quelques applications

4.1 Recalage ponctuel 2D ou 3D : ICP

Recalage ponctuel 2D ou 3D : ICP

- Algorithme de recalage de points proposé par Besl *et al.*, A Method of Registration of 3-D Shapes, *in* Trans. on PAMI, vol 14, No 2, Feb. 1992,
- ICP = Iterative Closest Points,
- Le terme ponctuel est réducteur : on peut l'étendre à des structures géométriques plus complexes.
- Procédé qui permet simultanément de :
 - mettre en correspondance des points/structures d'une image à l'autre
 - de calculer la transformation de recalage.
- Méthode simple, efficace et robuste : très employée, beaucoup de variantes dans la littérature.

ICP

- Recaler un ensemble P par rapport à un autre ensemble R .
- Ces ensembles peuvent être constitués de primitives géométriques diverses (point, segment de droite, courbe, surface).
- On se donne une distance entre un point \vec{p} et un ensemble X :

$$d(\vec{p}, X) = \min_{\vec{x} \in X} \|\vec{x} - \vec{p}\|$$

- On se donne un opérateur *de plus proche voisin* : $Y = \mathcal{C}(P, X)$ = ensemble des points/primitives de P les plus proches de X .
- En d'autres termes :

$$\vec{y} \in Y \iff d(\vec{y}, X) = \min_{\vec{x} \in X} \|\vec{x} - \vec{y}\| \mid \vec{y} \in P$$

Algorithme (en coordonnées homogènes)

1. Démarrage avec $P_0 = P$
2. Étape k :
 - (a) Calcul des points les plus proches :

$$Y_k = \mathcal{C}(P_k, R) \tag{8}$$

- (b) Calcul du recalage; minimiser la fonction de coût :

$$\begin{cases} A_k = \operatorname{argmin}_A \|Y_k - AP_0\| \\ d_k = \min_A \|Y_k - AP_0\| \end{cases} \tag{9}$$

- (c) Recalage : calculer l'image P_{k+1} :

$$P_{k+1} = A_k P_0$$

- (d) Fin si $|d_k - d_{k-1}| < \tau$ sinon $k \leftarrow k + 1$

Estimation de la transformation rigide

- L'ICP fonctionne avec des transformations rigides!
- les transformations affines donnent de moins bons résultats.
- Moindre carré pour transformations rigides :
 - Utilisation de quaternion (formule directe),
 - Moindre carré + décomposition SVD de la matrice homogène des transformations affines,
 - Le plus simple : estimation globale
 - de la T translation des barycentres des nuages de points P et R ;
 - de la rotation des barycentres de $P + T$ et R .

Un mot sur l'opérateur \mathbf{C}

- Pour appliquer l'équation (9), les deux vecteurs Y et R doivent être en correspondance.
- La MEC est assurée par l'équation (8) qui n'est nécessairement pas une bijection puisque en toute généralité, $\text{card}(R) \neq \text{card}(P)$.
- L'algorithme est très simple (en $\mathcal{O}(n^2)$, on peut faire mieux) :
 1. Pour point \vec{p}_i de P , trouver \vec{r}_j dans R qui minimise $d(\vec{p}_i, \vec{r})$,
 2. établir la correspondance $i \mapsto j$.

Convergence de la méthode

Proposition 3. *L'algorithme ICP est monotone décroissant sur la distance donnée par l'équation (9).*

- Preuve : Soit $d_k = \|Y_k - P_{k+1}\|^2$ et $e_k = \|Y_k - P_k\|^2$.
- On a toujours $d_k \leq e_k$:
 - En effet, par définition de P_{k+1} , on a $d_k = \|Y_k - P_{k+1}\|^2 = \min_A \|Y_k - AP_0\|^2$.
 - Le minimum est toujours atteint, on a donc pour tout A : $\min_A \|Y_k - AP_0\|^2 \leq \|Y_k - AP_0\|^2$.
 - Cette inégalité est vraie pour $A = A_{k-1}$ en particulier, donc $\|Y_k - AP_0\| = \|Y_k - P_k\| = e_k$.
 - Montrons maintenant que $e_{k+1} \leq d_k$:
 - On a $e_{k+1} = \|Y_{k+1} - P_{k+1}\|$
 - $Y_{k+1} = \mathcal{C}(P_{k+1}, R)$ donc par définition :

$$\|Y_{k+1} - P_{k+1}\| \leq \|Y - P_{k+1}\|$$

- C'est vrai en particulier pour $Y = Y_k$, donc $e_{k+1} \leq d_k$
- Finalement on a :

$$d_{k+1} \leq e_{k+1} \leq d_k \leq e_k, \forall k$$

CQFD.

Exemple de fonctionnement

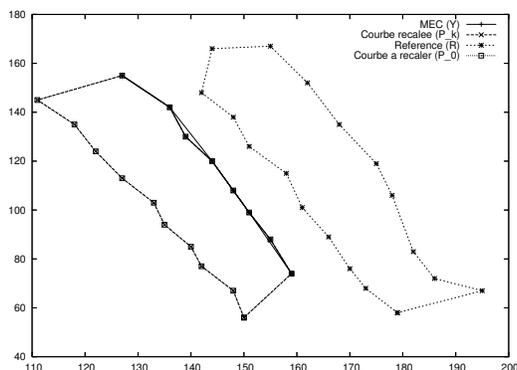


FIG. 4 – Initialisation

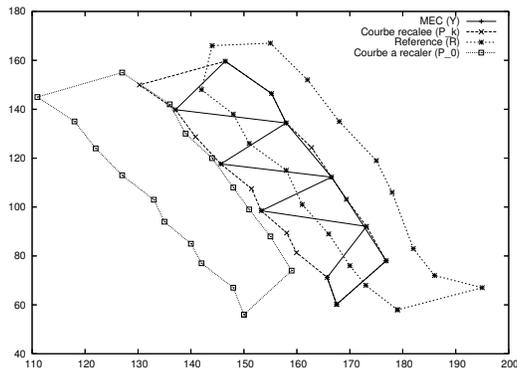


FIG. 5 – Itération 1

Conclusion sur l'ICP

- Méthode robuste et rapide (quelques itérations avec une dizaine de points, quelques dizaine d'itérations avec une centaine de points) ;
- Bien adapté au recalage rigide, mal au non rigide.
- l'opérateur \mathcal{C} est coûteux, on peut l'améliorer (tri en $\mathcal{O}(n \log(n))$),
- Adaptable très facilement dans \mathbb{R}^n

4.2 Mise en correspondance de cartes 3D

Mise en correspondance de cartes 3D Cadre du problème

- Les données sont des cartes de distance 3D : un point \mathbf{x} de \mathbb{R}^3 est représenté comme un point d'une surface explicite.
- $\mathbf{x} = (x, y, z(x, y))^T$
- Problématique : recalage deux cartes de distance 3D.
- On aborde le problème dynamiquement (lien fort avec l'estimation mouvement) : on calcule le vecteur de déplacement (vitesse) entre les deux

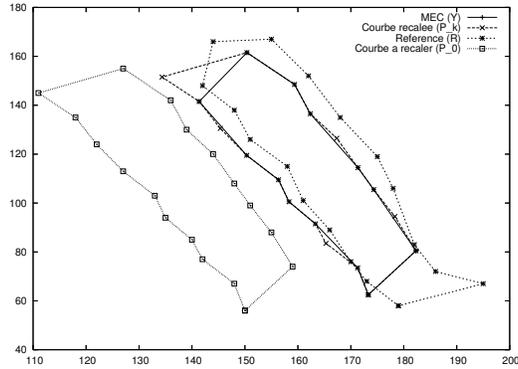


FIG. 6 – Itération 2

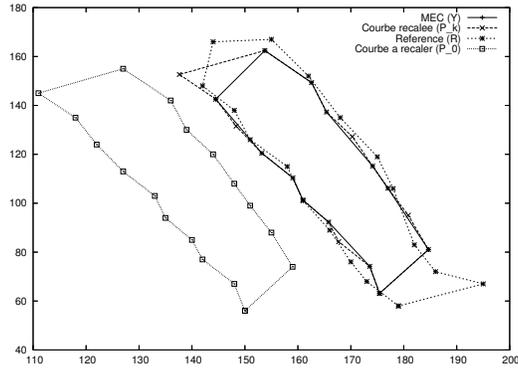


FIG. 7 – Itération 3

cartes.

– Un résultat de cinématique :

$$\frac{d\mathbf{x}}{dt} = -\mathbf{v} - \omega \wedge \mathbf{x} \quad (10)$$

avec :

$$\left\{ \begin{array}{l} \wedge : \text{produit vectoriel} \\ \mathbf{v} : \text{vitesse linéaire instantanée} \\ \omega : \text{vitesse angulaire instantanée} \end{array} \right.$$

– Cette équation se développe donc sous la forme :

$$\left\{ \begin{array}{l} \frac{dx}{dt} = -\mathbf{v}_x - \omega_y z + \omega_z y \\ \frac{dy}{dt} = -\mathbf{v}_y - \omega_z x + \omega_x z \\ \frac{dz}{dt} = -\mathbf{v}_z - \omega_x y + \omega_y x \end{array} \right.$$

– D'autre part, on peut écrire la dérivée totale de z :

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial z}{\partial t} \quad (11)$$

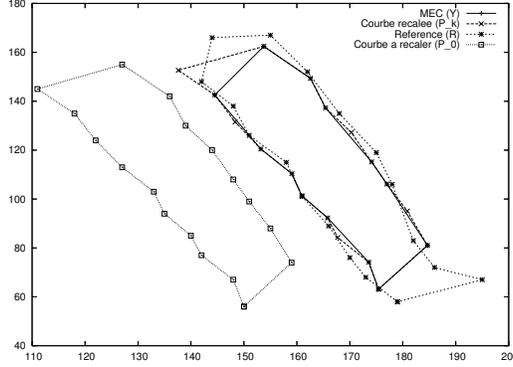


FIG. 8 – Itération 4

- Remarquons que $\frac{dx}{dt} = \frac{\partial x}{\partial t}$ et que $\frac{dy}{dt} = \frac{\partial y}{\partial t}$ (que l'on note respectivement \dot{x}, \dot{y}).
- On note également $\frac{\partial z}{\partial x} = z_x$ et $\frac{\partial z}{\partial y} = z_y$
- Regroupons les équations (83) et (11) pour trouver :

$$\begin{cases} z_x \mathbf{v}_x + z_y \mathbf{v}_y - \mathbf{v}_z - \\ (y + z_y z) \omega_x + (x + z_x z) \omega_y + (z_y x - z_x y) \omega_z + \\ = \frac{\partial z}{\partial t} \end{cases} \quad (12)$$

Bilan

- les termes z_x, z_y et $\frac{\partial z}{\partial t}$ sont des données connues : dérivées spatiales temporelles de $z = \text{IMAGE}$
- les termes x, y et z sont évidemment connus.
- 6 paramètres donc sont à estimer : $X = (\mathbf{v}_x, \mathbf{v}_y, \mathbf{v}_z, \omega_x, \omega_y, \omega_z)^T = (\mathbf{v}, \omega)^T$
- l'équation (12) est une forme linéaire en X et s'écrit :

$$AX = b \quad (13)$$

avec : $A = ((z_x, z_y, -1, -y - z_y z, x + z_x z, z_y x - z_x y)(x_i, y_i)_{i=1 \dots N})$ et $B = (\frac{\partial z}{\partial t}(x_i, y_i)_{i=1 \dots N})$

Résolution

- La relation s'applique en tout point (x_i, y_i) de la carte 3D.
- Avec $N = 6$ points, on peut inverser (13) (sans dépendance linéaire entre les équations).
- Avec un plus grand nombre de points, on utilise les moindres carrés pour un résultat robuste : la solution directe est donnée par :

$$X = (A^T A)^{-1} A^T B$$

4.3 Distances tangentés

Distances tangentés

- Méthodologie initiée par un papier de Simard *et al* dans *Neural Inf. Proc. Systems* 1992.
- Contexte : reconnaissance de caractère (OCR).
- Comment représenter les imgettes (de caractères) indépendamment des transformations affines (orientation, échelle et position) ?
- Plus précisément : un critère de mesure entre les formes invariant par transformation affine.
- Reconnaissance de caractère : base de formes référentes (*template*), on discrimine via une mesure de distance entre deux formes.
- La distance usuelle (euclidienne) entre deux formes ne convient pas ! (pas d'invariance aux transformations affines).

Formulation des distances tangentés

- Soit une image I définie sur un domaine borné de \mathbb{R}^n et à valeur dans un domaine bornée de \mathbb{R}^p

$$I : \Omega \rightarrow \mathcal{I}$$

- Soit T une transformation affine de \mathbb{R}^n , paramétrée par θ .
- Par exemple, pour $n = 3$, θ est un vecteur à 10 composantes.
- Considérons l'ensemble défini par :

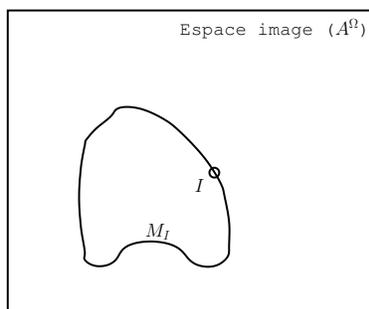
$$M_I = \{T_\theta(I) | \theta \in \Theta\}$$

où Θ est l'ensemble de toutes les valeurs possibles pour le paramètre vecteur θ .

- M_I est une variété de dimension Ω^Θ .

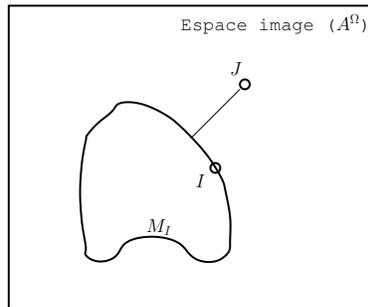
Que représente M_I ?

- M_I = ensemble de toutes les transformations affines possibles pour I .
- Un élément de M_I est donc une image.
- $J \in M_I \Leftrightarrow \exists \theta_0 | J = T_{\theta_0}(I)$



Comparaison de deux images

- Comment comparer deux images I et J ?
- Par exemple : en mesurant la distance de J à M_I



- Rappel : $\exists I_0 \in M_I | dist(J, M_I) = dist(J, I_0)$ avec $dist$ une distance quelconque (dont euclidienne) de l'espace image.

Distances tangentés : interprétation géométrique

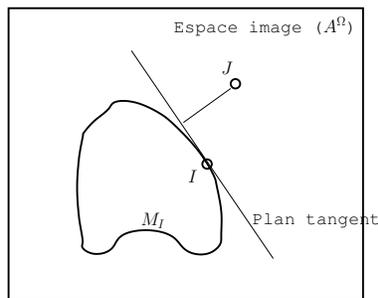


FIG. 9 – Le plan tangent est une approximation ...

- Pourquoi cette distance est-elle invariante par transformation affine ?
- Parceque toutes les transformées de I par T_θ sont à distance nulle de M_I !
- La distance $D(I, J) = dist(M_I, J)$ mesure donc les transformations non affines entre I et J .
- Le calcul de la distance d'un point à un ensemble n'est pas aisée. Particulièrement dans le cas où l'ensemble est de grande dimensions, puisque :

$$\begin{aligned} D(I, J) &= \min_{L \in M_I} \|J - L\|^2 \\ &= \min_{\theta \in \Theta} \|J - T_\theta(I)\|^2 \end{aligned}$$

Calcul de la distance

- La trop grande taille du problème empêche le calcul d'une solution exacte.
- La fonctionnelle n'est pas linéaire :

$$T_{\theta}(I)(\mathbf{x}) = I(A\mathbf{x} + t)$$

- Approximation (linéarisation) du problème : développement de Taylor à l'ordre 1.

Applications

- En reconnaissance des formes (OCR, atlas) :
 ⇒ les DT sont couplés à un procédé de classification (réseau de neurones, classifieur bayésien, k-plus proches voisins, ...) et à une base d'apprentissage.
- En recalage :
 ⇒ Il faut calculer :

$$\theta_{\text{registration}} = \underset{\theta \in \Theta}{\operatorname{argmin}} \|J - T_{\theta}(I)\|^2$$

4.4 Recalage 3D-2D : RANSAC

RANSAC

- Algorithme proposé par [Fischler and Bolles, 1981].
- Cadre d'application initial : alignement de données en cartographie.
- RANSAC : RANdom SAMple Consensus.
- Problématique : exhiber un ensemble de points qui vérifie au mieux un modèle (formulation générique).
- En particulier : trouver la transformation entre un ensemble 2D parmi un ensemble 3D.

RANSAC : les données

- Soit \mathbb{M} un modèle à n paramètres (un ensemble de points, une transformation paramétrique, ...).
- Le modèle vérifie une (ou plusieurs propriétés) pour un ensemble de points $\{q_i\}_i$ si :

$$\mathbb{M}(q_i) = 0$$

- Exemples :
 - \mathbb{M} peut représenter une droite \mathcal{D} d'équation $y = ax + b$:

$$\mathbb{M}(x, y) = y - ax - b, \quad (x, y) \in \mathcal{D} \Leftrightarrow \mathbb{M}(x, y) = 0$$

- Distance à un objet \mathcal{O} :

$$\mathbb{M}(q) = \operatorname{dist}(q, \mathcal{O})$$

- Soit P un ensemble de p points tel que $p > n$.

RANSAC : algorithme

- \mathbb{M} a n paramètres (notons les Θ) : il faut n points $Q = \{q_1, \dots, q_n\}$ pour inverser le système :

$$\mathbb{M}(q_i) = a \Leftrightarrow \Theta = \mathbb{M}_Q^{-1}(a)$$

- Donc : $Q \Leftrightarrow \Theta \Leftrightarrow \mathbb{M}_{(\Theta)}$

Algorithme :

1. $S1 = \{ \text{tirage aléatoire de } n \text{ points dans } P \}$
2. $M1 = \mathbb{M}(S1)$
3. $S1^* = \{p \in P \mid \|M1(p)\| < \epsilon\}$ *L'ensemble $S1^*$ est appelé "consensus"*
4. Si $|S1^*| \geq t$:
 - (a) alors : $M1^* = \mathbb{M}(S1^*)$. Fin de l'algo.
 - (b) sinon : retour en 1 (au plus k fois).

Quelques remarques ...

- Si $p \in S1$ alors $M1(q) = 0$,
- Le paramètre ϵ conditionne la taille de $S1^*$:
 - $\epsilon = 0 \Rightarrow S1^* = S1 \Rightarrow |S1^*| = n$
 - $\epsilon = \infty \Rightarrow S1^* = P \Rightarrow |S1^*| = p$
- Selon les valeurs de ϵ on peut choisir un ensemble de consensus plus ou moins "lisse" (à la différence des approches de moindre carré).
- L'ensemble $S1^*$ représente les points les plus proches du modèle $M1$ à l'erreur ϵ près.
- L'étape 4.(a) est réalisée par moindres carrés (système est surdéterminé).
- Le modèle $M1^*$ représente un plus grand nombre de points (donné par ϵ) que le modèle $M1$.
- L'algo cherche donc à trouver un bon échantillonnage de points de P qui vérifient le modèle \mathbb{M} (à l'erreur près).
- En itérant le procédé (k) jusqu'à l'infini, on trouvera cet ensemble (s'il existe).
- Il faut donc évaluer les chances d'avoir consensus au bout de k itération pour vérifier la faisabilité de la méthode en pratique.
- Algorithme à 3 paramètres : ϵ , t et k . Comment les choisir ?

Choix des paramètres : le paramètre ϵ

- Pas de borne connue sur l'erreur ϵ (dépend des données et du modèle).
- Une méthode simple pour donner une idée de l'ordre de grandeur pour ϵ :

1. Perturber les données (d'un biais δ)
2. Regarder l'impact sur le modèle :

$$\mathbb{M}(q + \delta) = \epsilon(q, \delta)$$

- Le choix reste empirique.

Choix des paramètres : les paramètres t et k

- Calculons l'espérance de converger (trouver un consensus) au bout de k itérations.
- Soit K variable aléatoire donnant le nombre d'itération avant de trouver un *consensus*.
- $\mathbb{P}(K = k) = \text{proba d'un consensus à l'étape } k \times \text{proba d'un non consensus aux étapes précédentes} :$

$$\mathbb{P}(K = k) = b \times (1 - b)^{k-1}$$

- Probabilité b d'avoir un consensus :

$$\begin{aligned} b &= \mathbb{P}(|p \in P, \|M1(p)\| < \epsilon| \geq t) \\ &= \mathbb{P}(\|M1(p)\| < \epsilon)^t = w^t \end{aligned}$$

- et w probabilité qu'un point appartienne au modèle $M1$ (à l'erreur ϵ près).
- donc $\mathbb{P}(K = k) = w^t(1 - w^t)^{k-1}$
- L'espérance :

$$\begin{aligned} \mathbb{E}(K) &= \sum_k k \times \mathbb{P}(K = k) \\ &= \sum_k kw^t(1 - w^t)^{k-1} \end{aligned}$$

- On se rappelle que (si $|a| < 1$) :

$$\sum_{i>0} a^i = \frac{a}{1 - a} \tag{14}$$

- En dérivant l'équation (14), on a :

$$\sum_{i>0} ia^{i-1} = \frac{1}{(1 - a)^2} \tag{15}$$

- Finalement $\mathbb{E}(K) = \frac{w^t}{(1 - (1 - w^t))^2} = \frac{1}{w^t}$
- On peut aussi calculer facilement la variance.
- On dérive l'équation (15) :

$$\sum_{i>1} i(i - 1)a^{i-2} = \frac{2}{(1 - a)^3}$$

et on a :

$$\begin{aligned} \mathbb{E}(K^2) &= w^t \sum_{k>0} k^2(1 - w^t)^{k-1} \\ &= w^t \left(\sum_{k>1} k(k - 1)(1 - w^t)^{k-1} + \sum_{k>0} k(1 - w^t)^{k-1} \right) \\ &= \frac{2 - (1 - w^t)}{(1 - w^t)^2} \end{aligned}$$

- Finalement : $\sigma(K) = \sqrt{\mathbb{E}(K^2) - \mathbb{E}(k)^2} = \frac{\sqrt{1 - w^t}}{w^t}$

w	t=1	2	3	4	5	6
0.9	1.1	1.2	1.4	1.5	1.7	1.9
0.8	1.3	1.6	2.0	2.4	3.0	3.8
0.7	1.4	2.0	2.9	4.2	5.9	8.5
0.6	1.7	2.8	4.6	7.7	13	21
0.5	2.0	4.0	8.0	16	32	64
0.4	2.5	6.3	16	39	98	244
0.3	3.3	11	37	123	412	-
0.2	5.0	25	125	625	-	-

FIG. 10 – Valeur de $\mathbb{E}(K)$ en fonction de w et t .

Ransac : lien entre k et t

- Lorsque $w \ll 1$ et $t \gg 1$, on a $\sigma(K) \sim \mathbb{E}(K)$. Ex : $w = 0.5$, $t = 4$, $\mathbb{E}(K) = 16$ et $\sigma(K) = 15.5$.

Le paramètre t

- Le nombre d'essais k est donc choisi selon les valeurs de t et w (table).
- t doit être plus grand que n , mais pas trop (sinon k sera trop grand).
- t doit être suffisamment grand pour prendre en compte les points proche du modèle.
- En d'autres termes, les points non choisis doivent idéalement avoir une grosse erreur : maximiser l'erreur de modèle des points non choisis.
- Le choix de t est lié à ϵ
- Remarquons que la quantité w est connue et dépend de ϵ :
 - $w = \text{proba d'un consensus} = \frac{|S1^*|}{|P|}$ (par définition de $S1^*$).

Applications

- Initialement proposé pour résoudre un problème de localisation en 3D.
- Reconnaissance/localisation de forme dans l'espace.
- En recalage ?
- Le modèle \mathbb{M} est très généraliste, pour un problème de recalage affine par rapport à un ensemble de points R , on peut écrire :

$$\mathbb{M}(P) = \|P - AR\|^2$$

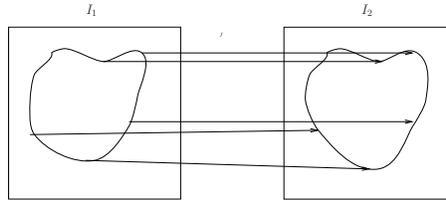
A sont les paramètres (matrice 2x2, 3x3 ou 4x4).

- Recalage nD/nD ou 2D/3D par exemple.

4.5 Recalage non affine : mise en correspondance de courbes

Mise en correspondance de courbes

- Mettre en correspondance *point à point* des formes analogues d'un plan à l'autre [Cohen et al., 1992].
- Le mouvement est déduit du vecteur de mise en correspondance (vecteur de déplacement).



- Soient deux images I^1 et I^2 à mettre en correspondance.
- On cherche une fonction f telle que :

$$\begin{cases} f : I_1 \longrightarrow I_2 \\ f(M) = M', M \in I_1, M' \in I_2 \end{cases}$$

- On a directement :

$$\vec{w} = \frac{\overrightarrow{MM'}}{\delta t}$$

- Sujet très vaste : un exemple de mise en correspondance de courbes fermées : bonne exemple de construction de contraintes.
- Méthode applicable après un processus de segmentation.
- Quel(s) critère(s) utiliser ?
- La courbure !
- Si le mouvement est rigide : la courbure est conservée (invariance par transformations affines).
- Les extrema de courbure (points caractéristiques) seront très certainement mis en correspondance.
- Points de courbure similaires : problème d'ouverture ! Une contrainte supplémentaire.

Définition

- Si le mouvement n'est pas rigide : la courbure n'est pas conservée. Des hypothèses sont nécessaires :
 - Déformations de faible amplitude,
 - Extrema de courbure sont conservées.
- Contrainte supplémentaire : régularité de la mise en correspondance (pénaliser les "croisements").
- C_P, C_Q deux courbes à mettre en correspondance. $P \in C_P, Q \in C_Q$
- Paramétrages des courbes : $s \in [0, 1]$ pour C_P et $s' \in [0, \alpha]$
- Fonction de mise en correspondance :

$$\begin{cases} f : [0, 1] \longrightarrow [0, \alpha] \\ f : s \mapsto s' \end{cases}$$

- Mise en correspondance des paramétrages (et non des points). f est la plus simple possible.

Formulation des contraintes

- Conservation de la courbure : le point M sur C_P à une courbure le plus possible de son correspondant sur $C_Q : C_Q(f(M))$:
 - Notons κ_P et κ_Q les courbures de C_P et C_Q ,
 - Courbure proche : $|\kappa_P(s) - \kappa_Q(s')|$ aussi petit que possible $\forall s \in [0, 1], s' = f(s)$.
 - Formulation en terme d'énergie :

$$E_\kappa(f) = \int_0^1 (\kappa_P(s) - \kappa_Q(f(s)))^2 ds$$

- Contrainte de régularité :
 - Régularité sur les vecteurs de mise en correspondance et non f .
 - $g(s) = \overrightarrow{P(s)Q(f(s))}$. Régularité de g
 - $\left\| \frac{\partial(Q(f(s)) - P(s))}{\partial s} \right\|^2$ aussi petit que possible $\forall s \in [0, 1]$.
 - Formulation en terme d'énergie :

$$E_r(f) = \int_0^1 \left\| \frac{\partial(Q(f(s)) - P(s))}{\partial s} \right\|^2 ds$$

Modèle

- Au final : minimiser en f la fonctionnelle suivante :

$$E(f) = E_\kappa(f) + RE_r(f) \quad (16)$$

$$R \in \mathbb{R}^{+*}$$

- Résolution par formulation variationnelle (même principe que pour le modèle de Horn).

Résolution numérique

- Première étape : calcul des équations d'Euler-Lagrange associé au problème de minimisation de (16) :

$$\begin{cases} f''(s)|Q'(f(s))|^2 = -\kappa_P(s) \langle N_P(s), Q'(f(s)) \rangle \\ \quad - \frac{1}{R} [\kappa_P(s) - \kappa_Q(f(s))] \kappa_Q'(f(s)) \\ f(0) = 0, f(1) = \alpha \end{cases}$$

- Deuxième étape : discrétisation de l'équation en f . Le terme $\|Q'(f)\|^2$ n'est pas linéaire. Nous le considérons comme non-dépendant de f : il est vu comme une donnée externe.

$$Af = L_f$$

- A discrétisation de f'' avec les coefficients $a_{ij} = \|Q'(f)\|^2$
- L_f regroupe les autres termes.

- Troisième étape : schéma d'évolution et discrétisation

$$\frac{\partial f}{\partial t} + Af = Lf$$

- Discrétisation "temporelle" ($\frac{\partial f}{\partial t} \sim \frac{f_t(s) - f_{t-1}(s)}{\tau}$) :

$$\begin{cases} (Id + \tau A)f_t = f_{t-1} + \tau Lf_{t-1} \\ f_0 \end{cases}$$

Initialisation

- Le processus itératif doit être initialisé.
- Premier choix : $f_0 = Id, (f_0(s) = s)$
- L'énergie E n'est pas convexe. Au mieux, on calcule un minimum local de l'énergie. Il faut être près de la solution recherchée (être dans la bonne "cuvette"). L'identité est potentiellement un mauvais choix.
- Une initialisation proche de la solution : apparier les points selon leurs courbures en respectant une certaine régularité.

Algorithme de calcul de f_0

- Pour chaque point $P \in C_P$ faire :
 - Calculer les points les plus proches de P sur C_Q .
 - Définir un ensemble $\{S_i\}_i$ de points candidats.
 - Pour chaque point S_i , chercher le point $Q \in C_Q$ qui minimise la quantité :

$$\int_{\delta S_i} (\kappa_P - \kappa_Q)^2 ds$$

- Prendre f_0 telle que $f_0(P) = Q$

Pré-traitements

- Le modèle réclame le prétraitement d'un certain nombre de données.
 1. Il faut être capable d'obtenir des courbes (qui ont segmenté correctement les régions à mettre en correspondance),
 2. Il faut calculer la courbure sur les courbes discrètes. Sujet très délicat !

Calcul de la courbure

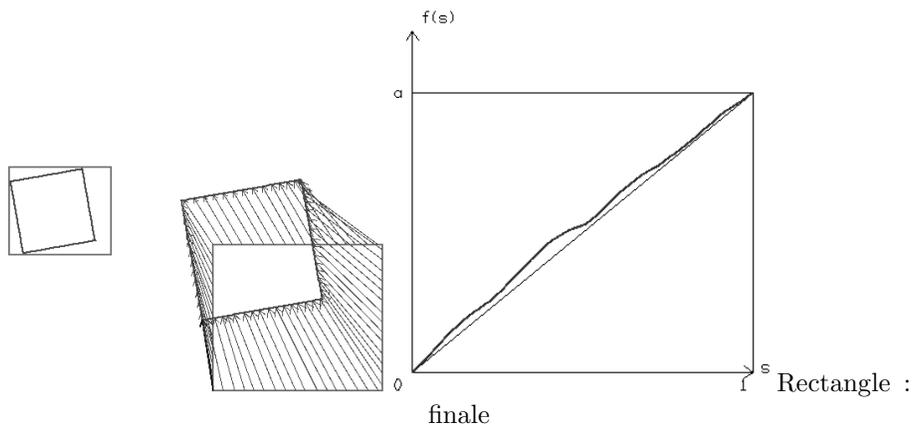
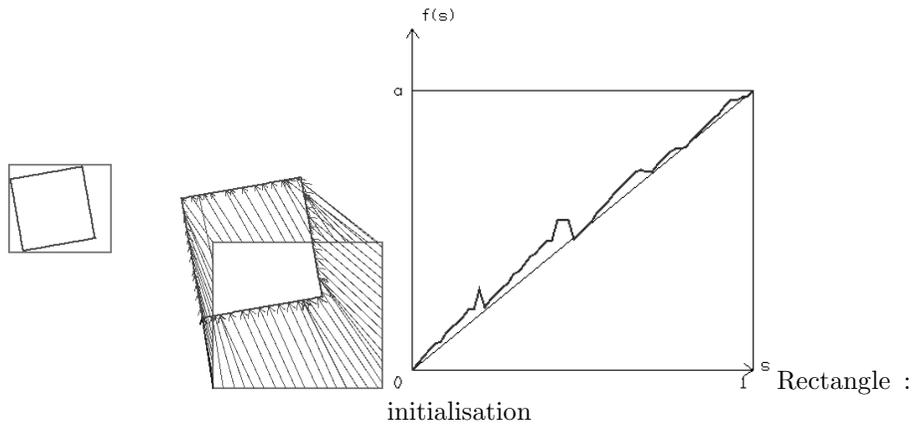
- Mathématiquement : Si $\mathcal{C}(s) = (x(s), y(s))$ une courbe paramétré sur $[0, 1]$, alors :

$$\kappa(s) = \frac{x'(s)y''(s) - x''(s)y'(s)}{(x'^2(s) + y'^2(s))^{3/2}}$$

- Les erreurs de discrétisation ont de trop fortes répercussions sur la courbure (moment d'ordre 2)

Deux méthodes :

- Lisser la courbe avec un filtre gaussien : risque de perte des points singuliers.

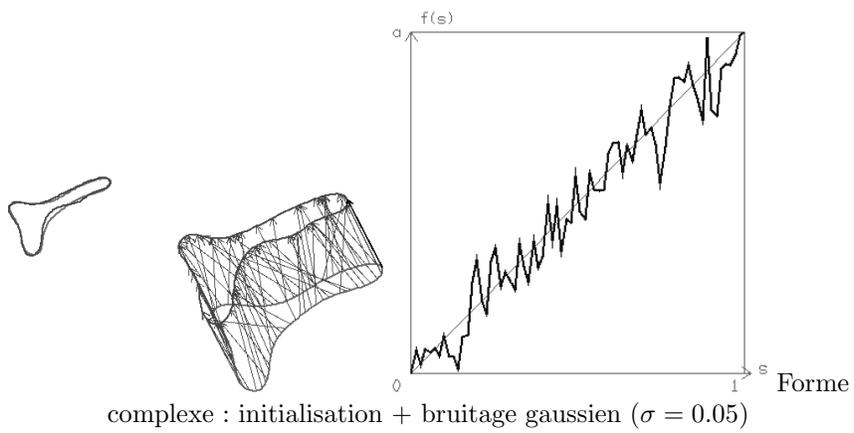
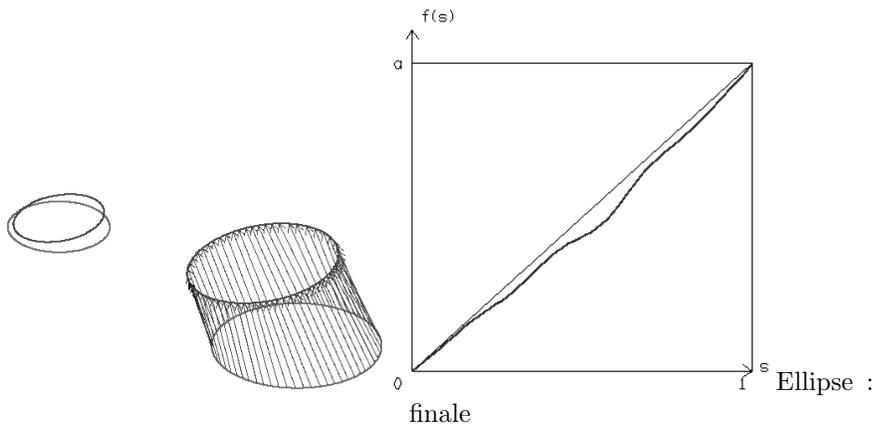
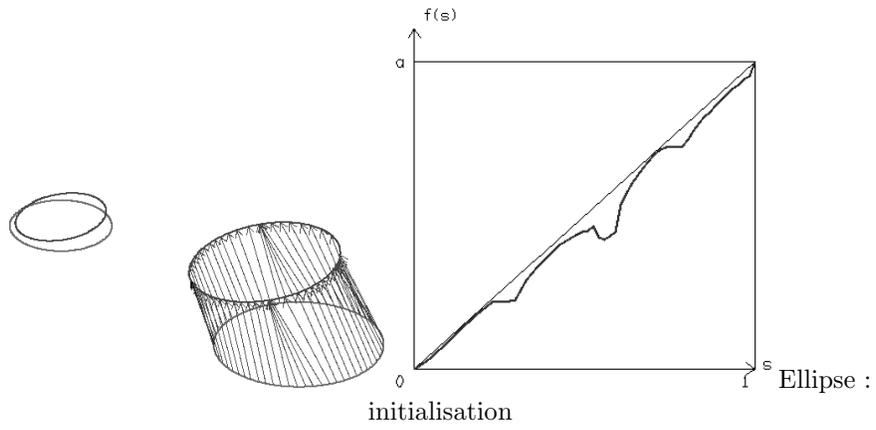


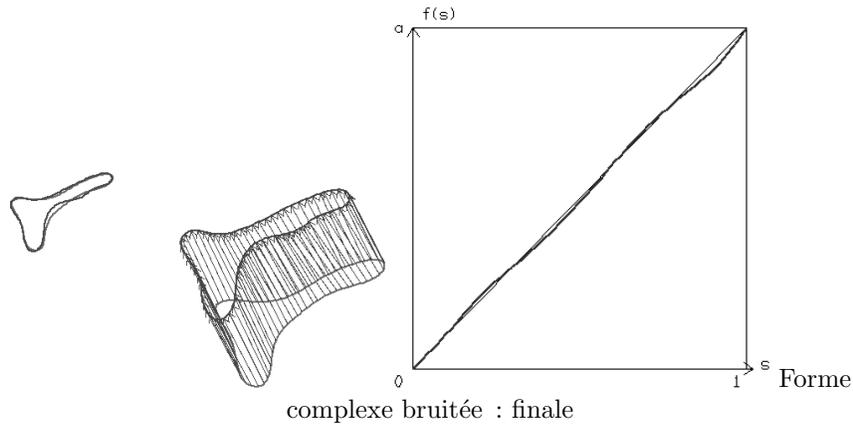
- approximation polynomiale de la courbe par splines cubiques (le meilleur choix). Courbe explicite et régulière. Courbure calculable explicitement.

Résultats

Références

- [Besl and McKay, 1992] Besl, P. J. and McKay, N. D. (1992). A method for registration of 3-d shapes. *Transactions on Pattern Analysis and Machine Intelligence*, 14(2). Premier article sur l'ICP.
- [Brown, 1992] Brown, L. G. (1992). A survey of image registration techniques. *ACM Computing Surveys*, 24(4) :325-376.





- [Cohen et al., 1992] Cohen, I., Ayache, N., and Sulger, P. (1992). Tracking points on deformable objects using curvature information. In *Proc. of 2nd European Conference on Computer Vision*, Santa Margherita Ligure, Italy.
- [Fischler and Bolles, 1981] Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus : A paradigm for model fitting with applications to image analysis and automated cartography. *ACM Graphics and Image Processing*, 24(6). Article sur la méthode RANSAC.
- [Lester and Arridge, 1999] Lester, H. and Arridge, S. (1999). A survey of hierarchical non-linear medical image registration. *Pattern Recognition*, 32 :129–149.
- [Maes et al., 1997] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P. (1997). Multimodality image registration by maximisation of mutual information. *IEEE Trans. on Medical Imaging*, 16(7) :187–198.
- [Maintz and Viergever, 1998] Maintz, J. B. A. and Viergever, M. A. (1998). A survey of medical image registration. *Image Medical Analysis*, 2(1) :1–37.
- [Press, 1995] Press, W. H. (1995). *Numerical Recipes in C : the art of programming*. Number 0-521-43108-5. Cambridge University Press. <http://www.fizyka.umk.pl/nrbook/bookcpdf.html>.
- [Simard et al., 1992] Simard, P.-Y., LeCun, Y. A., and Denker (1992). Efficient pattern recognition using a new transformation distance. *Advances in Neural Informatic Processing Systems*, 5 :50–58. Premier article sur les distances tangentes.
- [Sonka and Fitzpatrick, 2000] Sonka, M. and Fitzpatrick, M. (2000). *Handbook of medical imaging. Medical image processing and analysis*, volume 2. SPIE Press, Washington, USA.
- [Van Den Elsen et al., 1993] Van Den Elsen, P., Pol, E.-J. J., and Viergever, M. (1993). Medical image matching : A review with classification. *IEEE Engineering in Medicine and Biology*, 12 :26–39.

- [Viola and Wells, 1997] Viola, P. and Wells, W. (1997). Alignment by maximisation of mutual information. *International Journal of Computer Vision*, 24(2) :137–154.
- [Zhang, 1993] Zhang, Z. (1993). Le problème de la mise en correspondance : l'état de l'art. Technical Report RR 2146, INRIA.
- [Zitová and Flusser, 2003] Zitová, B. and Flusser, J. (2003). Image registration methods : a survey. *Image and Vision Computing*, 21 :977–1000.