

Champs de Markov en Vision par Ordinateur

2: Simulation et optimisation

Solutions

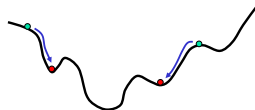
- On ne veut pas seulement modéliser. Il faut **calculer** la valeur d'une estimée.
- Les modèles ne sont pas simples: souvent ils demandent de **grandes ressources en temps de calcul et en mémoire**.
- Les espaces sont **énormes** et il y a beaucoup de **minima locaux**.
- Exemple : le recuit simulé peut prendre des heures même sur les images assez petites.

Estimées de MAP

- Il y a beaucoup des algorithmes différents, mais ils se regroupent dans trois catégories:
 - Variationnels;
 - Stochastiques;
 - Graphiques.

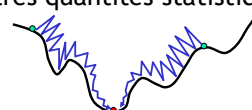
Méthodes Variationnelles

- elles **descendent à long du gradient**.
- **Rapides**, mais normalement on trouve seulement un **minimum local**.
- Dépendantes de **l'initialisation**.



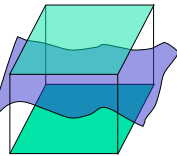
Méthodes Stochastiques

- elles utilisent l'**échantillonnage** pour simuler la probabilité.
- **Très lentes**, mais on trouve le **minimum global** (au moins en théorie).
- On peut calculer le **moyenne** (ou d'autres quantités statistiques).



Méthodes Graphiques

- elles utilisent des algorithmes combinatoires sur les graphes.
- Ni trop lentes, ni trop rapides.
- On trouve le minimum global pour les champs binaires.
- Un garanti sur l'énergie pour les champs non-binaires.
- Il y a des limites sur la forme de la probabilité.
- Espace discretisé



Méthodes Variationnels : En bref

- On considère une configuration dépendant de temps : $X(t)$
- On change S selon le gradient de l'énergie.

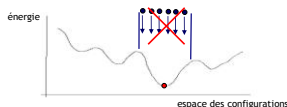
$$\frac{\partial X(t)}{\partial t} = - \frac{\delta U(X)}{\delta X}$$

- Beaucoup de variations sur cette thème.
- Problème : ils trouvent les minima locaux et dépendent de l'initialisation.



Configuration initiale

Lorsque l'on ne dispose pas d'une « bonne » configuration initiale

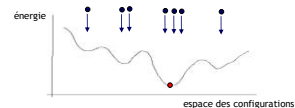


optimisation locale à éviter !

Configuration initiale

Techniques d'optimisation globale

- exploration exhaustive souvent impossible
- utilisation d'échantillonneurs permettant de dégager des configurations d'intérêt



Méthodes Stochastiques: les échantillonneurs MCMC

Intéressantes car

- énergie quelconque, non continue
- espace de configuration continu et/ou discret
- ne dépend pas du passé (pas besoin de mémoire)

Méthodes Stochastiques: les échantillonneurs MCMC

Le principe :

- Consiste à construire une chaîne de Markov $(X_t)_{t \in \mathbb{N}}$ à temps discret sur l'espace d'état, de mesure stationnaire π (spécifiée par notre modèle énergétique)

Méthodes Stochastiques: les échantillonneurs MCMC

Le principe :

- Les transitions correspondent à des perturbations de la configuration qui sont proposées selon des «**Noyaux**».
- La chaîne est construite de manière à être ergodique afin d'assurer la convergence vers la mesure cible.

Méthodes Stochastiques: les échantillonneurs MCMC

Echantillonneur d'Hasting (1970)

A l'instant t , dans la configuration $x_t = x$,

- on choisit un noyau de proposition Q_i avec une probabilité p_i
- on tire selon $Q_i(x \rightarrow \cdot)$ une nouvelle configuration y proche de la configuration x : $y = x + \Delta(x)$ ($Q_i(x \rightarrow y) = P(X_{t+1} = y | X_t = x)$)
- on tire aléatoirement $p \sim U_{[0,1]}$
 - si $p < \frac{h(y) Q_i(y \rightarrow x)}{h(x) Q_i(x \rightarrow y)}$ alors on accepte $x_{t+1} = y$
 - sinon $x_{t+1} = x_t = x$

$$h(x) = \exp - U(x)$$

Les noyaux de proposition

uniformes (algorithme de Metropolis, 1953)

non uniformes

- selon des connaissances a priori sur le problème
- selon la configuration courante (noyaux de régularisation)
- selon les données

Méthodes Stochastiques: le recuit simulé

Echantillonneur d'Hastings comme technique d'optimisation:

$$h(\cdot) \rightarrow h(\cdot)^{\frac{1}{T_t}}$$

le paramètre de relaxation T_t (appelé «température») est une suite décroissante de valeurs convergeant vers 0 quand t tend vers l'infini

Recuit Simulé : Relaxation Stochastique

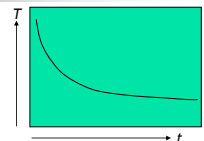
$$\Pr_T(x) = Z(T)^{-1} \exp - \frac{U(x)}{T}$$

- Quand $T \rightarrow \infty$, \Pr_T devient **uniforme**.
- Quand $T \rightarrow 0$, \Pr_T se concentre sur les **maxima globaux** de π .
- Engendrer une séquence de configurations avec $T \rightarrow 0$.

Recuit Simulé : Descente de Température.

- On prouve que, si :

$$T(t) \geq \frac{k}{\log(t)}$$



- Puis la configuration quand $T=0$ sera le **minimum globale**.
- Mais il faut attendre !**
- Plus souvent : $T(t) = kC^t$, $C < 1$.

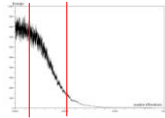
Le recuit simulé en pratique

La décroissance d'énergie en pratique: 3 phases

1- **plateau initial**: à haute température, la majorité des configurations proposées est acceptée (bonne estimation de la température initiale)

2- **phase de forte décroissance énergétique**: le processus explore les différents modes de la densité, et devient de plus en plus sélectif

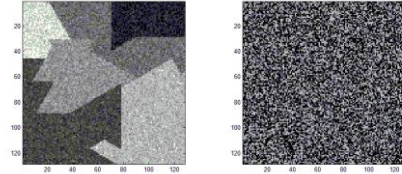
3- **phase d'ajustement**: la configuration courante est proche de l'optimum global, le taux d'acceptation est très faible



Un exemple en segmentation d'images

espace des configurations : $\{1,2,3,4,5,6\}^{nb \text{ pixels}}$

énergie : attache gaussienne + modèle de Potts



Recuit Simulé : Problèmes.

- En pratique, on **doit** utiliser une loi de descente de température **sous-optimale**.
- La **théorème de convergence** peut **donner l'impression** que tous ira bien, mais...
- Expérience avec les **algorithmes graphiques**, qui trouvent le minimum global **dans un temps fini**, montre que **les lois sous-optimales sont...sous-optimales**.
- Convergence en **100 - 1000** itérations.