MAP-MRFに基づく最適化(I)

ビジョン問題：その性質

１．問題の表現

ラベル集合Lの連続性

- 連続（離散）問題
- 組合せ問題

２．目的関数

解 f が探索空間内で制約を受けるか否か

- 制約なし問題
- 制約つき問題

３．最適化アルゴリズム

局所的探索か、それとも大局的探索か

- Local Method
- Global Method

ビジョン問題に対する最適化手法

- 尤度最大化 ⟶ EM法
- 事後確率最大化（MAP:エネルギー最小化）

⟶ 勾配法や変分法，Lagrange未定乗数法
各種のMAP - MRF法

- Relaxation Labeling法
- Iterated Conditional Mode(ICM)法
- Metropolis sampling and Gibbs sampling法
- Mean Field法， Belief Propagation法　など.

最大尤度推定による画像分割

〔資料９参照〕

原画像

初期分割画像
（K-mean法；４クラスタ）

EM法による最尤推定結果
（４混合分布によるモデル化）

勾配法と最小化問題

勾配法の概要（エネルギー最小化）

MAP解 $\boldsymbol{f}^* = \arg\min_f E(\boldsymbol{f})$を求める最も単純な方法は

勾配法．初期値 $\boldsymbol{f}^{(0)}$から始め、

$$\boldsymbol{f}^{(t+1)} \leftarrow \boldsymbol{f}^{(t)} - \mu \nabla E(\boldsymbol{f}^{(t)}),$$

（μはステップ刻み）

を$\nabla E(\boldsymbol{f}^{(t)}) = \boldsymbol{0}$となる点まで繰り返す．

$\nabla E(\boldsymbol{f}^{(t)}) = \boldsymbol{0}$を与えるラベルを$\boldsymbol{f}^*$とすると，

f^*は局所的最小解．

連続ラベルのエネルギー関数最小化

連続ラベル集合Lにおけるラベル$\boldsymbol{f} \in \boldsymbol{F}$の連続復元問題の事後エネルギー$E(\boldsymbol{f})$は,

サイト集合：$S = \{(x, y) \mid 1 \le x, y \le n\}$,

近傍系　　　：$N_{(x,y)}$　サイト(x, y)の近傍集合

とすると,

$$E(\boldsymbol{f}) = U(\boldsymbol{d} \mid \boldsymbol{f}) + U(\boldsymbol{f})$$
$$= \sum_{x,y} (f_{x,y} - d_{x,y})^2 + \lambda \sum_{x,y} \sum_{(k,l) \in N_{(x,y)}} g(f_{x,y} - f_{k,l})$$

$g(\cdot)$はペナルティ関数

$g'(\eta)=2\eta h(\eta)$と仮定すると，サイト(x,y)におけるラベル勾配$\nabla_{f_{x,y}}E(f)$は，

$$\frac{\partial E}{\partial f_{x,y}}=2(f_{x,y}-d_{x,y})+2\lambda\sum_{(k,l)\in N_{(x,y)}}(f_{x,y}-f_{k,l})h(f_{x,y}-f_{k,l})$$

勾配法の局所的最小解条件$\nabla E(\boldsymbol{f}^{*})=0$を満たす$\boldsymbol{f}^{*}$を得るために，以下の反復式を各サイト毎に実行.

$$f_{x,y}^{(t+1)}\leftarrow f_{x,y}^{(t)}-2\mu\{(f_{x,y}^{(t)}-d_{x,y})+\lambda\sum_{(k,l)\in N_{(x,y)}}(f_{x,y}^{(t)}-f_{k,l}^{(t)})h(f_{x,y}^{(t)}-f_{k,l}^{(t)})\}$$

μは，ステップ幅を決定 する定数

オプティカルフロー問題

オプティカルフロー：

観察者の移動とシーン中の物体の移動によって生じる画像面内のベクトル場

→ 領域と境界の分割、形状復元などの重要な情報を含む.

フロー計算の２つのパラダイム

- 特徴に基づく方法…局所的なベクトル場
- 濃度勾配に基づく方法…画像全体のベクトル場

点(x, y)，時刻 t における画像の明るさ を$I(x, y, t)$とする.
このとき，明るさ不変 の拘束条件を仮定：

（拘束A)
移動する点の明るさは，時刻tと$t+\delta t$の間で不変.

$$I(x, y, t) = I(x+\delta x, y+\delta y, t+\delta t)$$

$I(x+\delta x, y+\delta y, t+\delta t)$の$(x, y, t)$における１次のTayler展開
と拘束A，$\delta t \to 0$より，以下の関係が導かれる.

$$\frac{\partial I}{\partial x}\frac{dx}{dt} + \frac{\partial I}{\partial y}\frac{dy}{dt} + \frac{\partial I}{\partial t} = 0$$

$$\nabla I = \begin{pmatrix} I_x \\ I_y \end{pmatrix} = \begin{pmatrix} \partial I / \partial x \\ \partial I / \partial y \end{pmatrix}, I_t = \frac{\partial I}{\partial t}, \boldsymbol{v} = \begin{pmatrix} v_x(x, y) \\ v_y(x, y) \end{pmatrix} = \begin{pmatrix} dx / dt \\ dy / dt \end{pmatrix}$$

とおくと，オプティカルフロー$\boldsymbol{v}$の基本式は

$$v_x \cdot I_x + v_y \cdot I_y + I_t = \nabla I \cdot \boldsymbol{v} + I_t = 0$$

画像観測の関係上，誤差が含まれる．さらに，オクルージョンや動きの不連続などが存在．

⇨ 事後エネルギーの最小化問題

(尤度エネルギー＋事前エネルギー)

- 尤度（観測データのらしさ）エネルギー：

$$U(I \mid \boldsymbol{v}) = (\nabla I \cdot \boldsymbol{v} + I_t)^2$$

- 事前エネルギー：フロ ーの滑らかさ

⟹ フロー勾配の大きさの2乗で拘束を表現

$$U(\boldsymbol{v}) = \| \nabla v_x \|^2 + \| \nabla v_y \|^2$$

$$= \left(\frac{\partial v_x}{\partial x} \right)^2 + \left(\frac{\partial v_x}{\partial y} \right)^2 + \left(\frac{\partial v_y}{\partial x} \right)^2 + \left(\frac{\partial v_y}{\partial y} \right)^2$$

従って，事後エネルギー$E(\boldsymbol{v}) = U(\boldsymbol{v} \mid I)$は

$$E(\boldsymbol{v}) = U(\boldsymbol{v} \mid I) = U(I \mid \boldsymbol{v}) + \lambda U(\boldsymbol{v})$$

λは重み係数

$$= (\nabla I \cdot \boldsymbol{v} + I_t)^2 + \lambda \{ \| \nabla v_x \|^2 + \| \nabla v_y \|^2 \}$$

(Horn-Schunck’s method)

⟹ 事後エネルギー$E(\boldsymbol{v})$を最小化する$\boldsymbol{v}^*$の計算.

デジタル画像におけるフロー計算

サイト(i, j)における尤度エネルギ　ー：$c(i, j)$

$$c(i, j) = \left[I_x(i, j)v_x(i, j) + I_y(i, j)v_y(i, j) + I_t(i, j) \right]^2$$

サイト(i, j)での事前エネルギー $s(i, j)$は,

$$\begin{aligned} s(i, j) = \tfrac{1}{4}[&(v_x(i, j) - v_x(i-1, j))^2 + (v_x(i+1, j) - v_x(i, j))^2 \\ &+ (v_x(i, j+1) - v_x(i, j))^2 + (v_x(i, j) - v_x(i, j-1))^2 \\ &+ (v_y(i, j) - v_y(i-1, j))^2 + (v_y(i+1, j) - v_y(i, j))^2 \\ &+ (v_y(i, j+1) - v_y(i, j))^2 + (v_y(i, j) - v_y(i, j-1))^2] \end{aligned}$$

従って，最適フロー推定$\boldsymbol{v}^*$は,

$$\boldsymbol{v}^* = \arg\min_{\boldsymbol{v}} \sum_i \sum_j [c(i, j) + \lambda s(i, j)]$$

最適フロー，$\boldsymbol{v}^* = \left\{ \begin{pmatrix} v_x^*(1,1) \\ v_y^*(1,1) \end{pmatrix}, \begin{pmatrix} v_x^*(1,2) \\ v_y^*(1,2) \end{pmatrix}, \cdots, \begin{pmatrix} v_x^*(n,n) \\ v_y^*(n,n) \end{pmatrix} \right\}$

の必要条件より，$\dfrac{\partial E}{\partial v_x(i,j)} = 0, \quad \dfrac{\partial E}{\partial v_y(i,j)} = 0$

$$
\begin{aligned}
2\left(v_x(i,j)I_x + v_y(i,j)I_y + I_t\right)I_x + 2\lambda(v_x(i,j) - \bar{v}_x(i,j)) = 0 \\
2\left(v_x(i,j)I_x + v_y(i,j)I_y + I_t\right)I_y + 2\lambda(v_y(i,j) - \bar{v}_y(i,j)) = 0
\end{aligned}
$$

ここで，$\bar{v}_x, \bar{v}_y$はv_x, v_yの局所的平均値を表し，

$$
\begin{aligned}
\bar{v}_x(i,j) = 1/4\left\{v_x(i-1,j) + v_x(i+1,j) + v_x(i,j-1) + v_x(i,j+1)\right\} \\
\bar{v}_y(i,j) = 1/4\left\{v_y(i-1,j) + v_y(i+1,j) + v_y(i,j-1) + v_y(i,j+1)\right\}
\end{aligned}
$$

整理すると，

$$
\begin{aligned}
\left[\lambda + (I_x)^2\right]v_x + I_x I_y v_y = \lambda\bar{v}_x - I_x I_t \\
I_x I_y v_x + \left[\lambda + (I_y)^2\right]v_y = \lambda\bar{v}_y - I_y I_t
\end{aligned}
$$

オプティカルフローの解は、Gauss - Seidel法に基づいた連立方程式の反復解として

$$v_x^{k+1} \leftarrow \bar{v}_x^k - \left[\frac{(I_x)\bar{v}_x^k + (I_y)\bar{v}_y^k + I_t}{\lambda + (I_x)^2 + (I_y)^2} \right] I_x$$

$$v_y^{k+1} \leftarrow \bar{v}_y^k - \left[\frac{(I_x)\bar{v}_x^k + (I_y)\bar{v}_y^k + I_t}{\lambda + (I_x)^2 + (I_y)^2} \right] I_y$$

と与えられる.

値の変化がなくなるまで更新を繰り返す.

離散ラベルのエネルギー関数最小化

ラベル集合が離散的な場合，事後エネルギーの最小化は組合せ問題となる.

MRFにおける各ラベルの結合確率を最大化することは困難である.

⇒ greedyな戦略（局所的最適化）：ICM法

すべての変数が同時には関係しない最適化問題である場合，逐次的最適化問題として表現する.

⇒ DP法が有効.

Iterative Conditional Modes法（ICM法） [Besag, J(1986)]

$$\left.\begin{aligned} &\text{密度関数}：p(\boldsymbol{d} \mid \boldsymbol{f}) = \prod_i p(d_i \mid f_i) \\ &P(f_i \mid d_i, f_{N_i}) \propto p(d_i \mid f_i) P(f_i \mid f_{N_i}) \end{aligned}\right\}$$

を仮定.

$$\begin{aligned} &P(f_i \mid d_i, f_{N_i}) \\ &= \frac{P(f_i, d_i, f_{N_i})}{P(d_i)P(f_{N_i})} \\ &= \frac{p(d_i \mid f_i, f_{N_i})P(f_i \mid f_{N_i})P(f_{N_i})}{P(d_i)P(f_{N_i})} \\ &= \frac{p(d_i \mid f_i)P(f_i \mid f_{N_i})}{P(d_i)} \\ &\propto p(d_i \mid f_i)P(f_i \mid f_{N_i}) \end{aligned}$$

$P(f_i \mid d_i, f_{N_i})$の最大化

$\Rightarrow$ 条件付エネルギー$V(f_i \mid d_i, f_{N_i})$

の最小化

ただし,

$$V(f_i \mid d_i, f_{N_i}^{(k)}) = V(d_i \mid f_i) + \sum_{j \in N_i} V(f_i \mid f_j^{(k)})$$

⇒反復式　$f_i^{(k+1)} \leftarrow \arg\min_{f_i} V(f_i \mid d_i, f_{N_i}^{(k)})$を利用.

例えば，区分的一定な面の復元問題に対しては，

$$V(f_i \mid d_i, f_{N_i}) = \frac{(f_i - d_i)^2}{\sigma^2} + v_{20} \sum_{j \in N_i} [1 - \delta(f_i - f_j)]$$

各サイトに対して反復式を適用．すべてのサイトが収束するまでICMサイクルを繰り返す.

初期値 $\boldsymbol{f}^{(0)}$としては，通常 $\boldsymbol{f}^{(0)} = \boldsymbol{d}$（観測値）を用いることが多い.

Dynamic Programming法

ポテンシャルエネルギーが,

$$E(f_1, f_2, \cdots, f_m)$$
$$= E_1(f_1, f_2) + E_2(f_2, f_3) + \cdots + E_{m-1}(f_{m-1}, f_m)$$

のように表されると仮定.

$$D_1(f_2) = \min_{f_1} E_1(f_1, f_2)$$
$$D_2(f_3) = \min_{f_2}[D_1(f_2) + E_2(f_2, f_3)]$$
$$\vdots$$
$$D_{m-1}(f_m) = \min_{f_{m-1}}[D_{m-2}(f_{m-1}) + E_{m-1}(f_{m-1}, f_m)]$$ とすると,

最小解 $\min_{f} E(f)$ は，以下のように表される.

$$\min_{f_1, \cdots f_m} E(f_1, f_2, \cdots, f_m) = \min_{f_m} D_{m-1}(f_m)$$

領域分割問題（例１）

ガウス性雑音を含む領域をMLLモデルで表し，画像の領域分割を行う方法(Derin - Elliot法).

$S = \{(i, j) \mid 1 \le i, j \le m\}$, $L = \{l_1, l_2, \cdots l_M\}$として，

事後エネルギーを

$$E(f) = U(f \mid d) = \sum_{c \in C} V_c(f) + \sum_{I \in L} \sum_{(i,j) \in S^{(I)}} \frac{1}{\sigma^2}(d_{i,j} - l_I)^2$$

$$= \sum_{k=1}^{m-1} E_k(f_k, f_{k+1}) = \sum_{k=1}^{m-1} \{ \sum_{c \in C^{k,k+1}} V_c(f) + \sum_{I \in L} \sum_{S_I^k} \frac{1}{\sigma^2}(d_{i,j} - l_I)^2 \}$$

と分解．ただし，f_k, f_{k+1}は$k, k+1$行のラベルを表し，

$S_I^k = \{(k, j) \mid f_{k,j} = l_I, 1 \le j \le m\}$を表す．

⇒ DPによる最小化

輪郭線抽出問題（SNAKES)(例２）

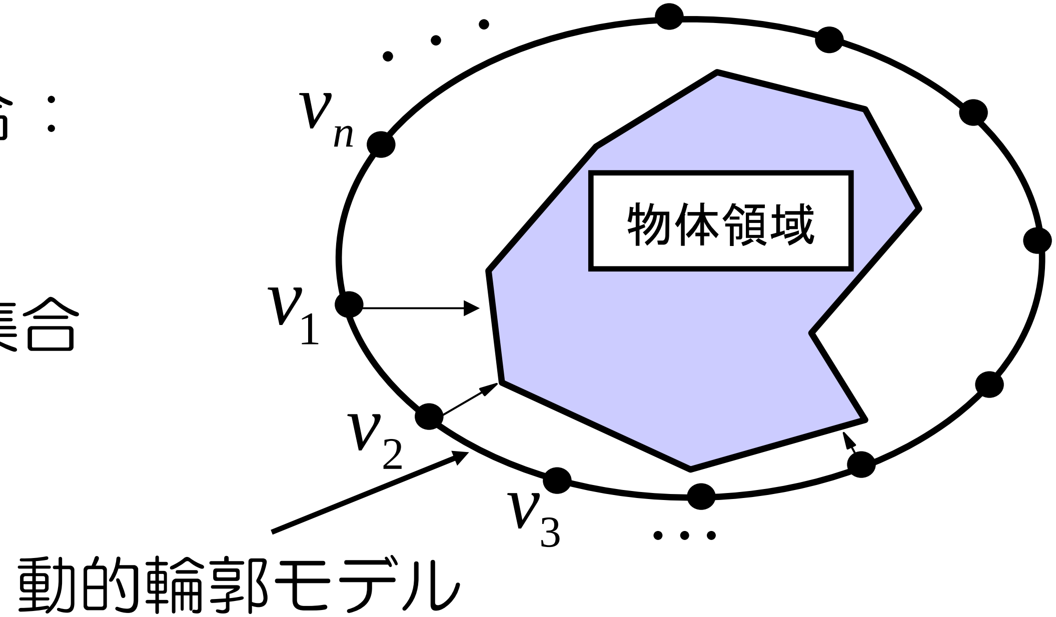
動的輪郭モデル(Active Contour Model)　[Kass et.al.]

エネルギー（評価関数）の導入とその最小化を行うために，輪郭線をパラメータ表現.

サイト（制御点）集合：

$$\boldsymbol{v} = (v_1, v_2, \cdots, v_n)$$

不規則サイト集合



エネルギー関数：E_{snakes}

$\Rightarrow$ 事後エネルギーとして定式化可能.

$$
\begin{aligned}
E_{snakes}(\boldsymbol{v}) &= E_{\mathrm{int}}(\boldsymbol{v}) + E_{ext}(\boldsymbol{I} \mid \boldsymbol{v}) \\
&= \sum_{i=1}^{n} E_{\mathrm{int}}(v_i) + \sum_{i=1}^{n} E_{ext}(\boldsymbol{I} \mid v_i)
\end{aligned}
$$

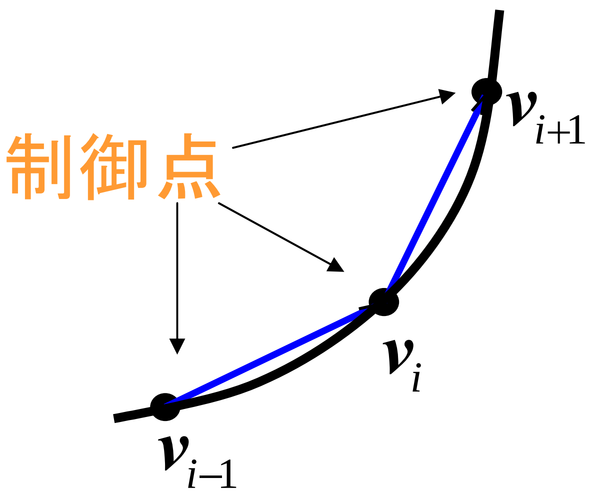
$E_{\mathrm{int}}(\boldsymbol{v})$：　配置 $\boldsymbol{v}$ の事前（内部）エネルギー

$E_{ext}(\boldsymbol{I} \mid \boldsymbol{v})$：観測画像 $\boldsymbol{I}$ の配置 $\boldsymbol{v}$ に対する尤度(画像）エネルギー

$$E_{\mathrm{int}}(v_i) = \alpha\,|\dot{v}_i| + \beta\,|\ddot{v}_i|$$

$$E_{ext}(\boldsymbol{I}\,|\,v_i) = -\gamma \cdot |\nabla_{v_i} I(x,y)|$$

$$\dot{v}_i = v_i - v_{i-1},\quad \ddot{v}_i = \dot{v}_{i+1} - \dot{v}_i$$



$\nabla_{v_i} I(x,y)$:

サイトv_iの位置での

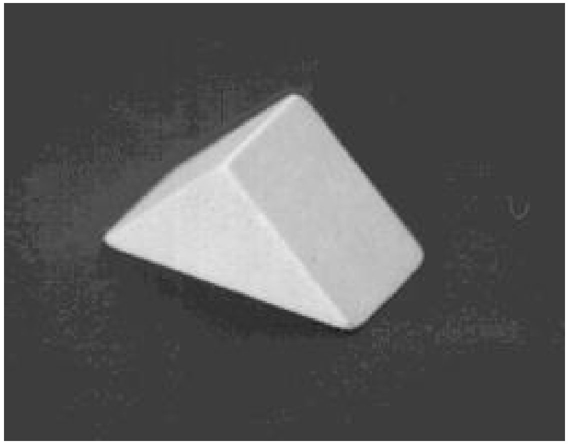
画像Iの濃度勾配

$$E_{snakes}(\boldsymbol{v}) = E(v_1, v_2, \cdots v_n)$$

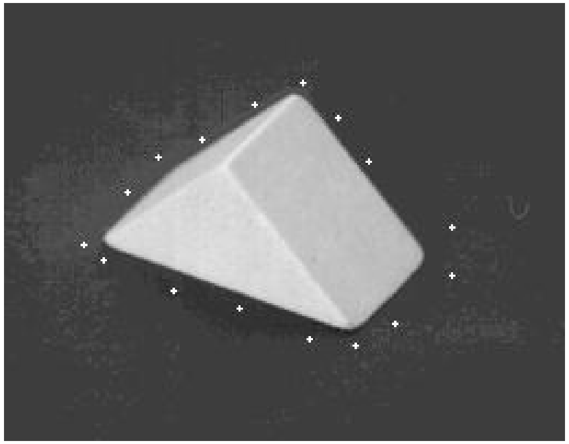
$$= E_1(v_1, v_2, v_3) + E_2(v_2, v_3, v_4) + \cdots + E_n(v_n, v_1, v_2)$$

と表現可能. ⇒ DPの利用.

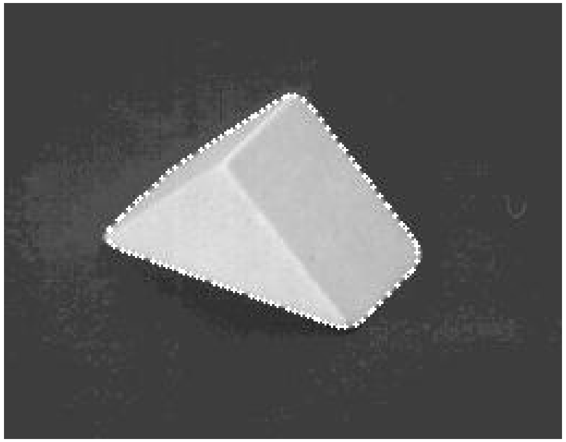
動的輪郭線処理の例



積み木画像



初期制御点配置
(手動による指定）



収束後制御点配置

弛緩法（エッジ強調問題の例）

弛緩法(Relaxation Labeling)

エッジの強さと方向に基づく確率的反復法

エッジ(サイト）集合 $S=\{1,2,\ldots.,n\}$

エッジラベル集合 $\Lambda=\{\lambda_1,\lambda_2,\ldots,\lambda_m\}$

サイト i がラベル λ_u を持つ確率 $P_i(\lambda_u)$

$$\sum_{\lambda_u} P_i(\lambda_u)=1,\ \ i\in S$$

$$P_i(\lambda_u)\geq 0,\ \ \forall i,\lambda_u$$

適合関数(compatibility function)

単項適合関数　$r_i(\lambda_u) = C_1 - V_1(\lambda_u)$

２項適合関数　$r_{ij}(\lambda_u, \lambda_v) = C_2 - V_2(\lambda_u, \lambda_v)$

$V_1(\lambda_u)$　単一サイトクリークポテンシャル

$V_2(\lambda_u, \lambda_v)$　ペアサイトクリークポテンシャル

ラベル配置 f のゲイン $G(f) :(= C_0 - E(f))$

$$G(f) = \sum_{i \in S} \sum_{\lambda_u \in \Lambda} r_i(\lambda_u) P_i(\lambda_u) +$$

$$\sum_{i \in S} \sum_{\lambda_u \in \Lambda} \sum_{j \in S, j \neq i} \sum_{\lambda_v \in \Lambda} r_{ij}(\lambda_u, \lambda_v) P_i(\lambda_u) P_j(\lambda_v)$$

事後エネルギー$E(f)$の最小化 $\Rightarrow$ $G(f)$の最大化

$G(f)$の最大化

勾配$\boldsymbol{q} = -\nabla E(f) = \{q_i(\lambda_u) \mid i \in \boldsymbol{S}, \lambda_u \in \boldsymbol{\Lambda}\}$を利用.

各$q_i(\lambda_u)$は、

$$q_i(\lambda_u) = \frac{\partial G}{\partial P_i(\lambda_u)} = r_i(\lambda_u) + 2\sum_{j \in S}\sum_{\lambda_v \in \Lambda} r_{ij}(\lambda_u, \lambda_v) P_j(\lambda_v)$$

ここで、$P_i^{(t)}(\lambda_u) = P_i(\lambda_u)$とし、サイト$i$が時刻$t$で

ラベルλ_uを持つ確率とする.

$P^{(t)}$を、$P^{(t)}$と$q^{(t)}$に基づいて更新.

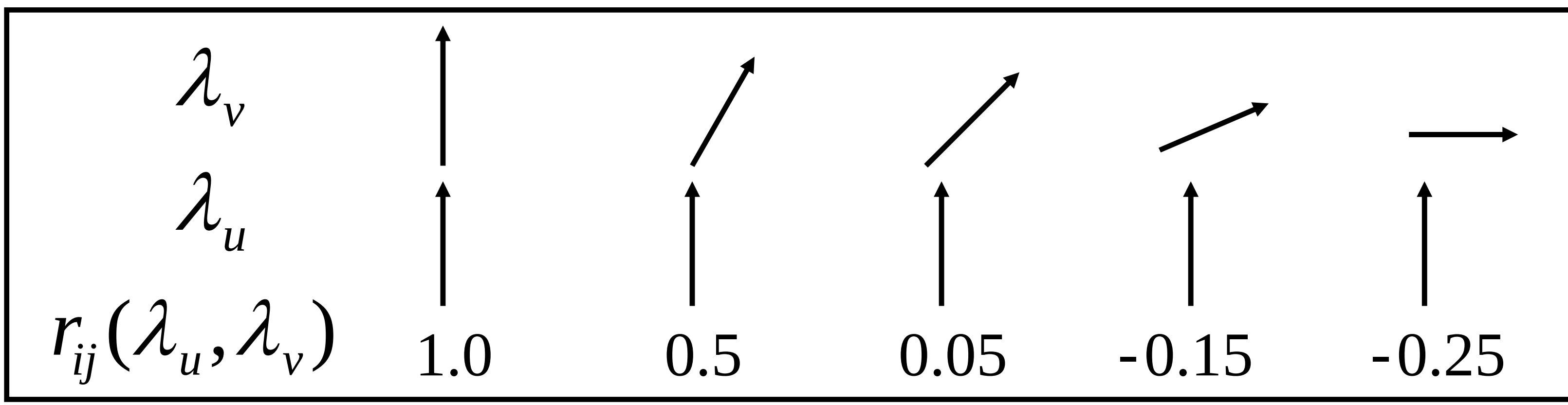
$$P^{(t+1)} \leftarrow \Phi(P^{(t)}, q^{(t)})$$

Φは、更新操作を表し、個々のアルゴリズム毎に定義される。(多数のアルゴリズムの存在）

•確率の更新例（Rosenfeld et. al.,1976）

$$P_i^{(k+1)}(\lambda_u) = \frac{P_i^{(k)}(\lambda_u)[1+q_i^{(k)}]}{\sum_{\lambda_u}\left\{P_i^{(k)}(\lambda_u)[1+q_i^{(k)}]\right\}}$$

$$q_i^{(k)} = \sum_j d_{ij}[\sum_{\lambda_v} r_{ij}(\lambda_u,\lambda_v)P_j^{(k)}(\lambda_v)]$$

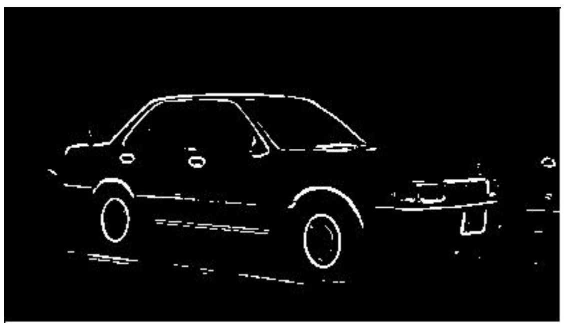


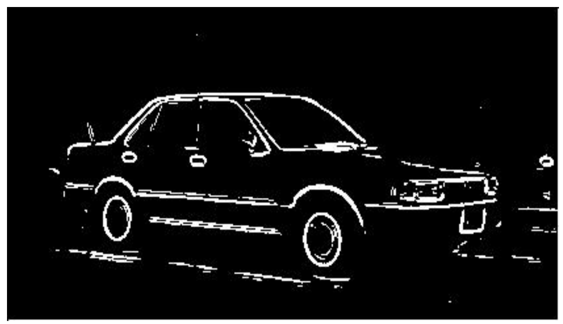
ラベル間の適合度(compatibility)係数の例

弛緩法によるエッジ強調例

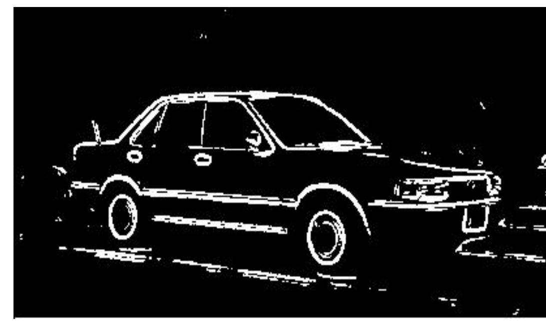
原画像

微分画像





繰返し：２回、閾値：0.7



繰返し：３回、閾値：0.7

繰返し：４回、閾値：0.7

繰返し：５回、閾値：0.7

繰返し：８回、閾値：0.7

主な参考文献

S.Z.Li, Markov Random Field Modeling in Computer Vision, Springer-Verlag, (1995)

Horn, B.K.P and Schunck, B.G., "Determining Optical Flow," Artificial Intelligence, vol.17, nos.1--3, pp.185-- 203, (1981).

Besag, J. ; “On the statistical analysis of dirty pictures”, *Journal of the Royal Statistical Society*, *Series B*, 48, pp.259—302 (1986).

M. Kass, A. Witkin, and D. Terzopoulos; “Snakes: Active contour models” In Proc. 1st Int. Conf. on Computer Vision, pp. 259--268, (1987).

A. Rosenfeld, R.A. Hummel and S.W. Zucker, Scene Labelling by Relaxation Operations, IEEE Trans. on SMC, 6, pp.420--433, (1976)

上田修功, ベイズ学習[I] ―統計的学習の基礎―, 電子情報通信学会誌, vol.85, No.4, pp.265-271, (2002)