

人物抽出手法に基づく 色・距離情報からの 観測方程式モデリング

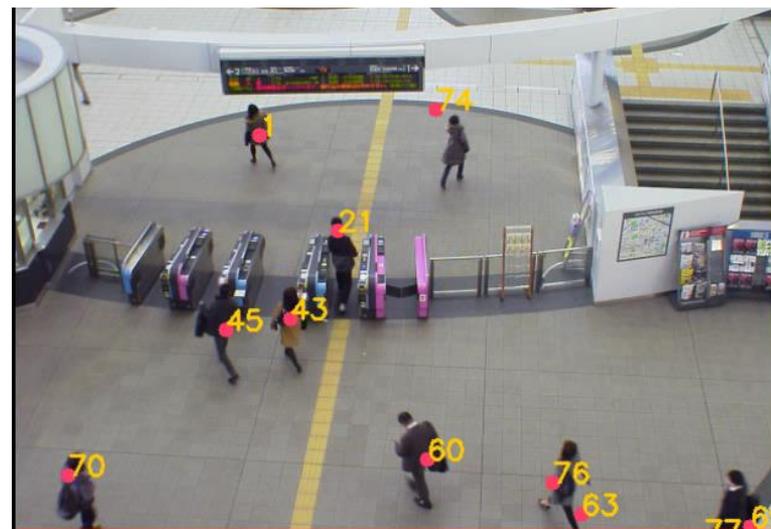
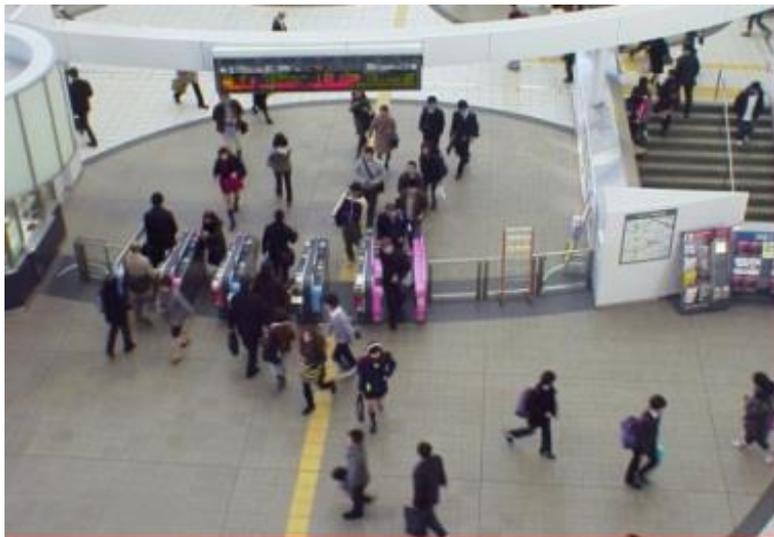
東京大学 中西 航

大学院工学系研究科 社会基盤学専攻
地域/情報研究室



第2回若手研究者の研究会
— 動的变化と不確実性を捉える理論の構築に向けて —
2013/07/14 東京大学

狭域における複数人物逐次追跡

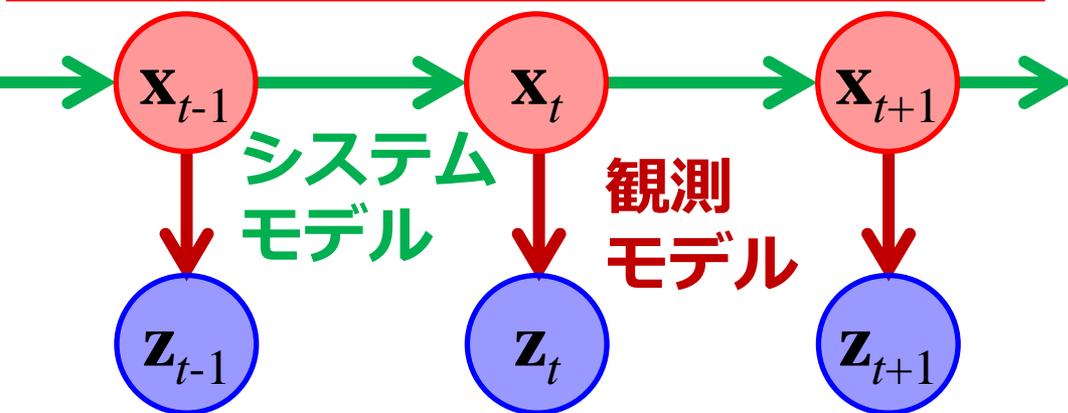


- ◆ いかなる手法でも、各時刻での「見え」と、時刻間の「動き」を捉える

人物逐次追跡の時系列・確率モデリング

- ◆ 一般状態空間モデルによるデータ同化

\mathbf{x}_t : 時刻 t における状態ベクトル
(人物位置・形状) : 観測不可能



\mathbf{z}_t : 時刻 t における観測ベクトル
(色・距離) : 観測可能

システムモデル：
歩行者挙動予測
観測モデル：人物抽出

$\mathbf{z}_{initial}$ から \mathbf{z}_t まで観測し、
事後確率 $p(\mathbf{x}_t | \mathbf{z}_{initial:t})$ を
ベイズ推定

(パーティクルフィルタにより計算)

$$\begin{aligned} p(\mathbf{x}_t | \mathbf{z}_{initial:t}) &\propto p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{z}_{initial:t-1}) \quad (\because \text{ベイズの法則}) \\ &= \underbrace{p(\mathbf{z}_t | \mathbf{x}_t)}_{\text{観測モデル}} \int \underbrace{p(\mathbf{x}_t | \mathbf{x}_{t-1})}_{\text{システムモデル}} \underbrace{p(\mathbf{x}_{t-1} | \mathbf{z}_{initial:t-1})}_{\text{1期前の結果}} d\mathbf{x}_{t-1} \end{aligned}$$

- ◆ 一般状態空間モデル

$$\begin{aligned} p(\mathbf{x}_t | \mathbf{z}_{initial:t}) &\propto p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{z}_{initial:t-1}) \quad (\because \text{ベイズの法則}) \\ &= \underbrace{p(\mathbf{z}_t | \mathbf{x}_t)}_{\text{観測モデル}} \int \underbrace{p(\mathbf{x}_t | \mathbf{x}_{t-1})}_{\text{システムモデル}} \underbrace{p(\mathbf{x}_{t-1} | \mathbf{z}_{initial:t-1})}_{\text{1期前の結果}} d\mathbf{x}_{t-1} \end{aligned}$$

- ◆ 非線形のシステムモデル・観測モデルの導入が可能
- ◆ しかし、単純に他分野の手法にならって一般状態空間モデルにおいてモデリングを行うには、いくつかの問題がある
 - ◆ 各モデルが確率分布で表現できればよいという説明がある一方で、確率分布を明示的に持つのは各モデルの誤差項だけ、というモデリングがほとんど

- ◆ 物理法則として知られている非線形システムモデル

- ◆ 例：波の伝播方程式

- ◆ システムモデルの計算結果と、観測値の関係とが明確な観測モデル

- ◆ 例：波高の算出値と実測値

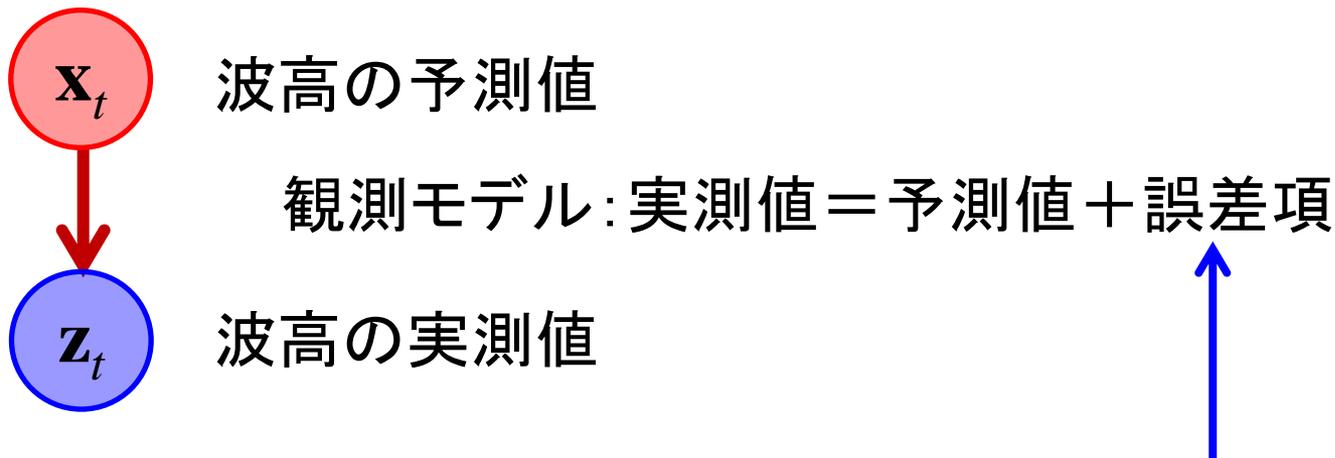
- ◆ この両者を組み合わせると、物理法則が考慮できていない部分について、「観測値を分布として取り込む」という状態空間モデリングの長所が発揮される
 - ◆ 我々が直接知ることが難しい初期条件、対象固有の条件などが反映された状態推定が行われる
 - ◆ 長期予測における蓄積誤差の影響が軽減される
- ◆ システム側に理論式があり、観測側に高精度なセンサがあれば、誤差項を(多くの場合、正規)分布として加えるだけでモデリングが完了する

- ◆ 歩行者挙動の法則：
詳細は不明

- ◆ 動画像からの人物抽出：
精度は高くない

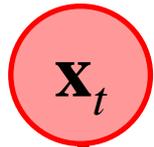
- ◆ 両者とも不確実性の高いモデルしか存在しない
- ◆ この状況において、各モデルが確率分布で表現できるという長所をどのように生かしていくのか？
- ◆ もともとのモデルが不明であるような場合を含めた、より一般的なモデリング方法に関する議論は、これまでほとんど行われていない
- ◆ 本日はこの点について、観測モデルに絞って検討する

- ◆ ある状態 \mathbf{x} からある観測 \mathbf{z} が実現する確率を表した確率分布



- ◆ 「確率分布を持つのは誤差項のみ」という状況
 - ◆ 予測値と観測値の関係は本当は不明であっても、正規分布のようなものを仮定する 경우가ほとんど
- ◆ 基本的に、観測値が得られれば得られるほど状態の推定精度が上がっていくという想定であり、いかに観測精度を上げるかという議論は少ない
 - ◆ 観測頻度、観測の次元を上げるという議論は多い

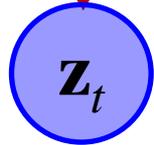
- ◆ 実測値を取得できる(と想定する)ならば一緒



歩行者位置の予測値



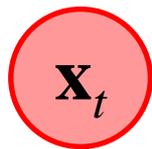
観測モデル: 実測値 = 予測値 + 誤差項



歩行者位置の実測値

- ◆ 実測値をマニュアルで動画像から取得しベイズ推定する研究は、少ないながら存在する(田代+ 2012)
- ◆ しかし、一般には歩行者位置それ自体は観測できず、観測できるものは色や距離の情報にすぎない

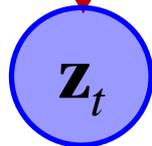
- ◆ 色・距離情報を観測するとするならば・・・



歩行者位置の予測値



観測モデル: 色・距離の値 = $f(\text{予測値}) + \text{誤差項}$

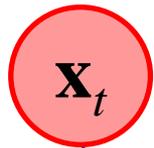


色・距離の観測値

- ◆ この、 $f(\text{予測値}) + \text{誤差項}$ (または $p(\text{観測値} | \text{予測値})$) のモデリングが必要
- ◆ しかしながら、このような観測ベクトルと状態量との関係が定かではない場合については、観測モデルのモデリング方法はほとんど検討されていない

動画像処理における人物抽出手法を参考にする

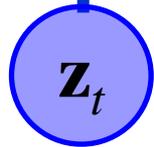
- ◆ 観測モデルとは逆向き



歩行者位置



モデル: 色・距離の値 → 特徴量 → 人物抽出・位置決定



色・距離の観測値

- ◆ モデルの目的は分布を返すことではなく一点を与えること
- ◆ 膨大な数の手法があるが、いずれも「抽出・追跡対象の見え」や「抽出・追跡対象と背景領域との識別」を表す特徴量を定義

→観測モデルとして扱いやすいモデルはどのようなものか？

- ◆ 観測モデルとしての扱いやすさを以下のように想定する
 1. 空間上のあらゆる点で値が算出できる(分布が描ける)
 - ◆ 予測(提案)分布が連続分布であるので、事後分布を計算するためには観測モデルも連続分布であることが望ましい
 2. 人物から外れた地点で高い値にならない
 - ◆ 抽出手法によっては、人物以外のものに対しても高い値を返す場合がある。予測(提案)分布が広がりを持つことを考慮すると、人物付近以外では高い値を返さない観測モデルが望ましい
 3. 遮蔽や近接に頑健
 - ◆ そもそもの抽出性能の高さともいえるが、対象とする環境に依存しない観測モデルが望ましい

- ◆ 特徴量による分類

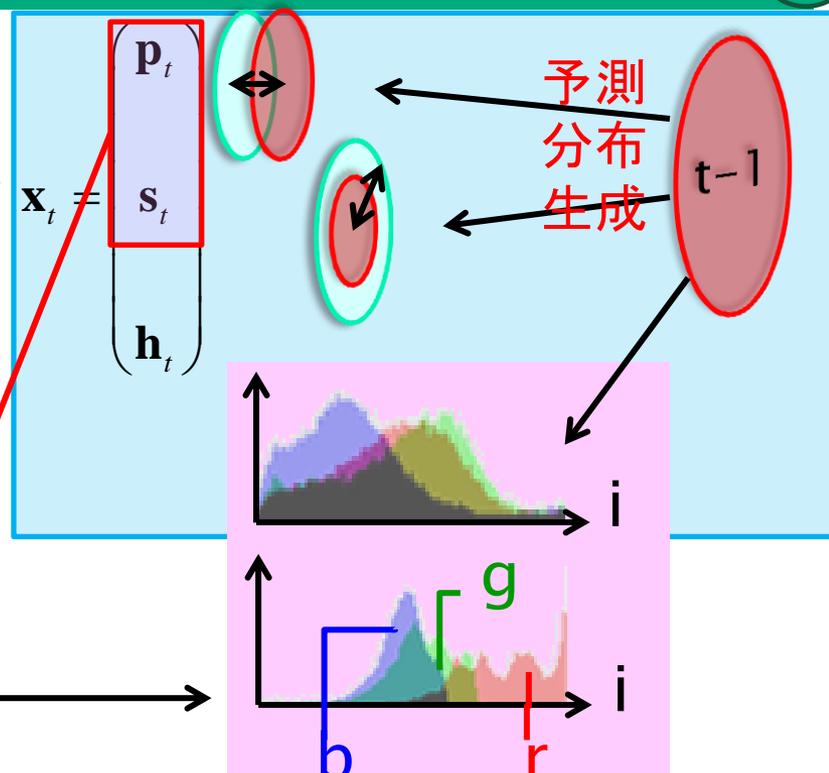
	特徴量設定方法	具体例
点： 画素値	画素値モデリング	PFinder(Wren+ 1997) MOG(Stauffer & Grimson 1999) W4(Haritaoglu 2000)
線： 輪郭	エッジ形状 勾配ヒストグラム	HOG(Dalal & Triggs 2005) Zhao & Nevatia (2003)
面： 領域全体	色ヒストグラム テンプレート 検出器	Nguyen+ (2007) Hagar & Belhumeur (1996) Grabner & Bischof (2006)

- ◆ 前出の条件を全て満たすのは色ヒストグラムと考えられる

\mathbf{x}_t 歩行者位置の
予測位置・形状・色ヒストグラム

観測モデル $p(z|x) \sim$
予測位置・形状内部の
色ヒストグラムの類似度

\mathbf{z}_t 色の観測値



- ◆ 本発表では、状態ベクトルのうち、形状と色ヒストグラムはシステム遷移の間は不変とする。すなわち、

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}) = \begin{pmatrix} \mathbf{p}_{t-1} \\ \mathbf{s}_{t-1} \\ \mathbf{h}_{t-1} \end{pmatrix} + \begin{pmatrix} \mathbf{v}_{t-1} \\ 0 \\ 0 \end{pmatrix}$$

- ◆ 各時刻における各歩行者速度 \mathbf{v} は、マニュアル取得した歩行者位置の差分から事前にマニュアルで与えておく

◆ 特徴量による分類

	特徴量設定方法	具体例
点： 観測点ま での距離	背景距離との差分 前景距離との差分	Harville + (2001) 鈴木 + (2005)
線： 輪郭	点群の形状	中村 + (2005) 橋本 + (2011)
面： 領域全体	距離ヒストグラム	Harville (2004) Veltkamp (2001)

- ◆ 前出の条件を全て満たすのは、
前景距離との差分 または 距離ヒストグラムと考えられる

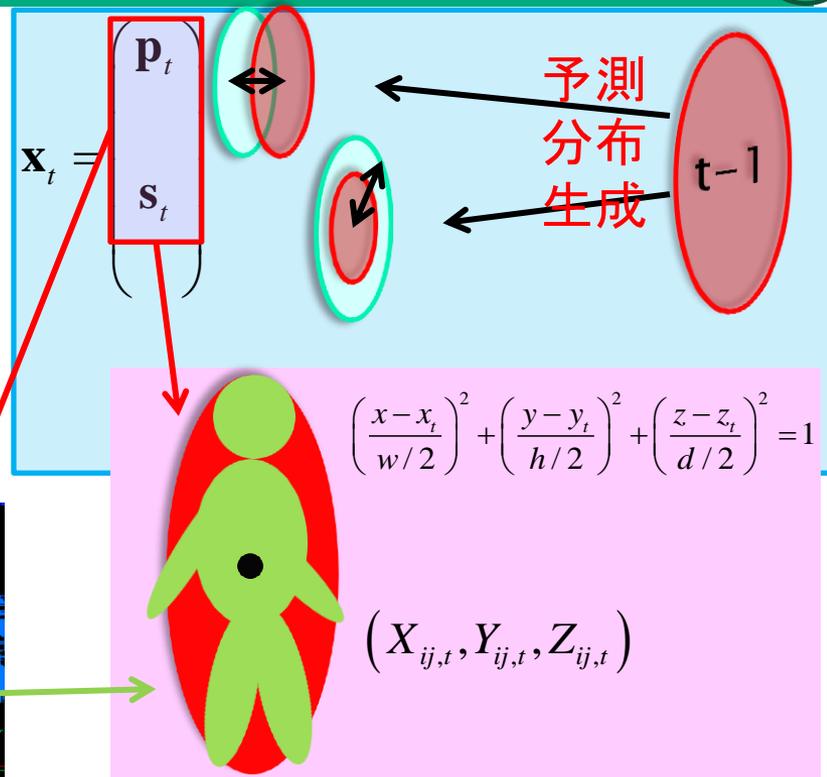
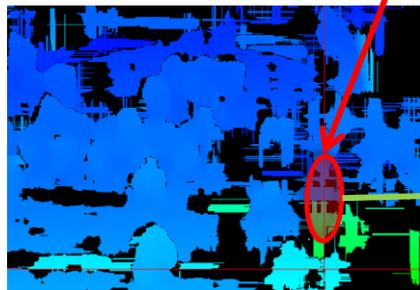
\mathbf{x}_t

歩行者位置の
予測位置・形状

観測モデル $p(z|x) \sim$
予測形状と
観測形状の類似度

\mathbf{z}_t

距離の観測値



- ◆ 本発表では、状態ベクトルのうち、形状はシステム遷移の間は不変とする。すなわち、

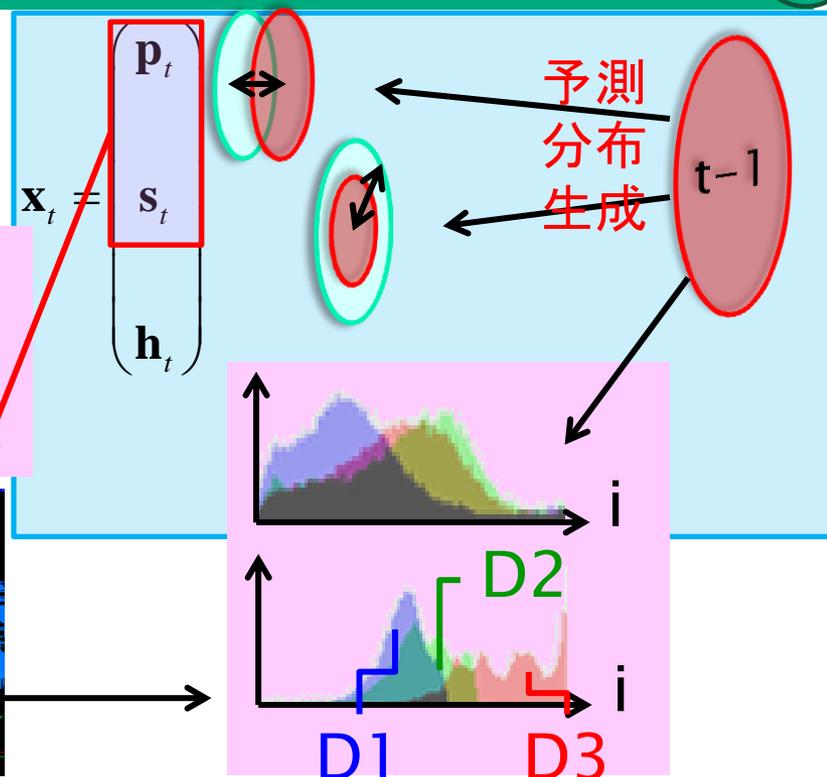
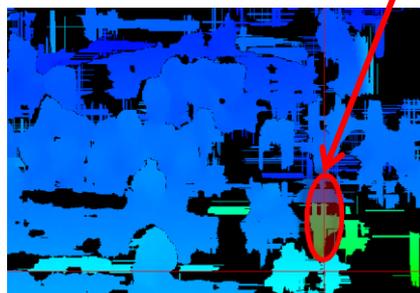
$$p(\mathbf{x}_t | \mathbf{x}_{t-1}) = \begin{pmatrix} \mathbf{p}_{t-1} \\ \mathbf{s}_{t-1} \end{pmatrix} + \begin{pmatrix} \mathbf{v}_{t-1} \\ 0 \end{pmatrix}$$

- ◆ 各時刻における各歩行者速度 \mathbf{v} は、マニュアル取得した歩行者位置の差分から事前にマニュアルで与えておく

\mathbf{x}_t 歩行者位置の
予測位置・形状・距離ヒストグラム

観測モデル $p(z|x) \sim$
予測位置・形状内部の
距離ヒストグラムの類似度

\mathbf{z}_t 距離の観測値



- ◆ 本発表では、状態ベクトルのうち、形状と距離ヒストグラムはシステム遷移の間は不変とする。すなわち、

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}) = \begin{pmatrix} \mathbf{p}_{t-1} \\ \mathbf{s}_{t-1} \\ \mathbf{h}_{t-1} \end{pmatrix} + \begin{pmatrix} \mathbf{v}_{t-1} \\ 0 \\ 0 \end{pmatrix}$$

- ◆ 各時刻における各歩行者速度 \mathbf{v} は、マニュアル取得した歩行者位置の差分から事前にマニュアルで与えておく

- ◆ 前出の一般状態空間モデルへの適用可能性から、いくつかのモデル候補が得られたとき、モデル構造間の比較やモデルパラメータの推定を行いたいが、どちらも困難を伴う

- ◆ 例)モデル構造やパラメータの良し悪しを全体尤度で評価

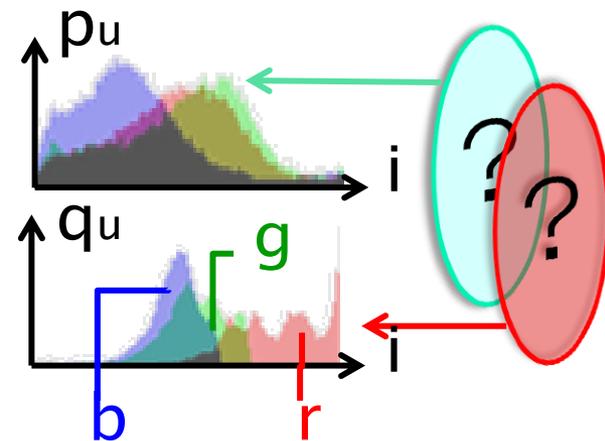
$$\begin{aligned} p(\mathbf{z}_{1:T}) &= p(\mathbf{z}_T | \mathbf{z}_{1:T-1}) p(\mathbf{z}_{1:T-1}) \\ &= \prod_{t=1}^T p(\mathbf{z}_t | \mathbf{z}_{1:t-1}) \\ \log p(\mathbf{z}_{1:T}) &= \log \{ p(\mathbf{z}_T | \mathbf{z}_{1:T-1}) p(\mathbf{z}_{1:T-1}) \} \\ &= \sum_{t=1}^T \log p(\mathbf{z}_t | \mathbf{z}_{1:t-1}) \end{aligned}$$
$$\begin{aligned} p(\mathbf{x}_t | \mathbf{z}_{1:t}) &= \frac{p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{z}_{1:t-1})}{p(\mathbf{z}_t | \mathbf{z}_{1:t-1})} \\ &= \frac{\int p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{z}_{1:t-1}) d\mathbf{x}_t}{p(\mathbf{z}_t | \mathbf{z}_{1:t-1})} \end{aligned}$$

- ◆ パーティクル数 N に対して対数全体尤度は $\sum_{t=1}^T \left(\log \sum_{i=1}^N w_{t,i} \right) - T \log N$
 $w_{t,i}$ は時刻 t における i 番目のパーティクルの重み

- ◆ 従って、観測モデルの返す値 w そのものの大小に左右されるため、単純なモデル間比較やパラメータ推定はできない

- ◆ 複数の予測分布×複数の観測モデルについて、それぞれ事後確率分布を算出する
 - ◆ 予測分布は、マニュアルで取得した人物位置を中心とした正規分布とし、分散の大きさを以下の4種類設定する
 - ◆ $(2.5\text{cm})^2$
 - ◆ $(5.0\text{cm})^2$
 - ◆ $(7.5\text{cm})^2$
 - ◆ $(50\text{cm})^2$
- ◆ 事後確率分布の平均・分散について検討する
- ◆ サンプルはマニュアルで位置を取得した11277個

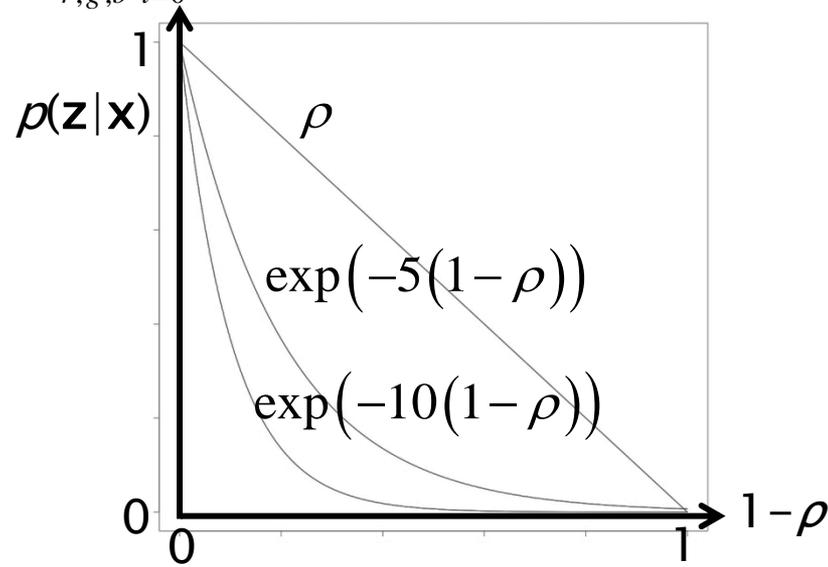
- ◆ ヒストグラムの生成方法2種類
 - ◆ 状態ベクトル内部の画素を一様に扱う
 - ◆ 状態ベクトル内部の画素について、楕円体中心との距離に従って重み付け



- ◆ 評価式3種類

- ◆ ヒストグラムの類似度：
Bhattacharyya Coefficient $\rho = \prod_{r,g,b} \sum_{i=0}^{255} \sqrt{p_u q_u}$ とするとき、

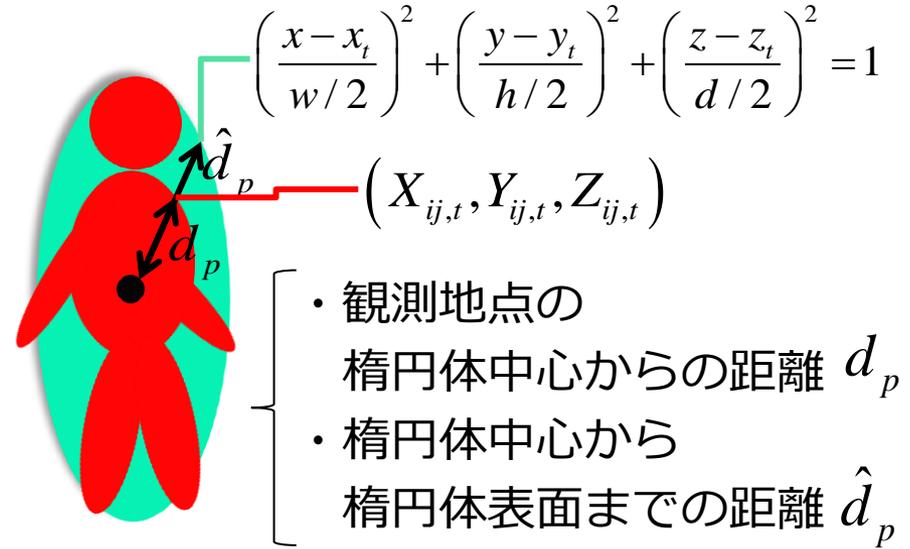
$$p(\mathbf{z} | \mathbf{x}) = \begin{cases} \rho \\ \exp(-5(1-\rho)) \\ \exp(-10(1-\rho)) \end{cases}$$



◆ 変数の設定方法2種類

◆ 実距離 $\frac{1}{I} \sum_p \left(|d_p - \hat{d}_p| \right)^\lambda$

◆ 距離の比 $\frac{1}{I} \sum_p \left(\frac{|d_p - \hat{d}_p|}{\hat{d}_p} \right)^\lambda$



◆ 評価式3種類

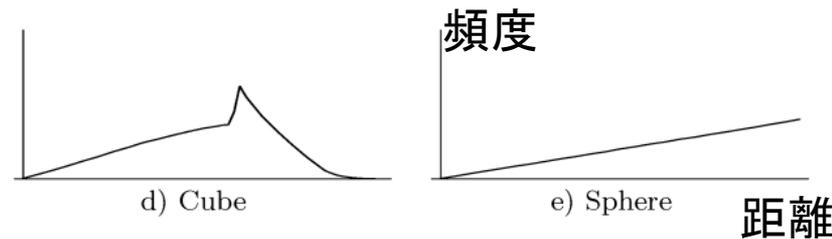
- ◆ $\lambda = 0.5, 1.0, 2.0$ の3種類

◆ Shape Distributionに基づく2種類の方法

- ◆ D1：定点と状態ベクトル内部の点との距離をサンプリング
- ◆ D2：状態ベクトル内部の2点間の距離をサンプリング

→ 適当に区切った距離帯によりヒストグラム生成

- ◆ D2の例(左：立方体、右：球)

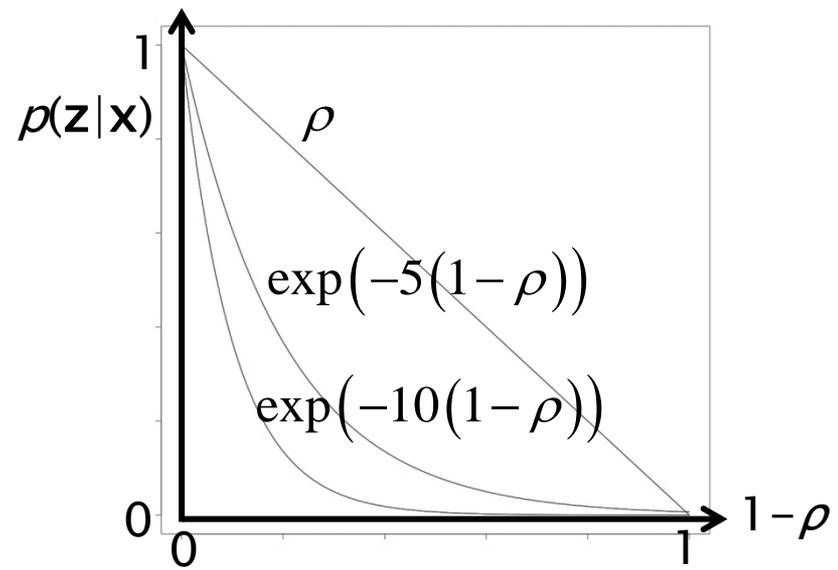


◆ 評価式3種類

- ◆ ヒストグラムの類似度：
Bhattacharyya Coefficient

$$\rho = \sum_{i=0}^{39} \sqrt{p_u q_u} \text{ とするとき、}$$

$$p(\mathbf{z} | \mathbf{x}) = \begin{cases} \rho \\ \exp(-5(1-\rho)) \\ \exp(-10(1-\rho)) \end{cases}$$



数値実験結果(色ヒストグラム)

Histogram	Variance of Proposal Distribution	Observation Model	Mean of residual	Mean of variance
Uniform	2.5 ²	Linear	2	11
		Exp5	2	11
		Exp10	3	11
	5 ²	Linear	6	40
		Exp5	6	42
		Exp10	9	35
	7.5 ²	Linear	10	87
		Exp5	11	91
		Exp10	15	69
	50 ²	Linear	48	1006
		Exp5	67	1082
		Exp10	74	329
Weighted	2.5 ²	Linear	2	11
		Exp5	2	11
		Exp10	3	11
	5 ²	Linear	6	40
		Exp5	6	42
		Exp10	9	36
	7.5 ²	Linear	10	87
		Exp5	10	92
		Exp10	15	70
	50 ²	Linear	47	2724
		Exp5	66	1170
		Exp10	74	341

数値実験結果(前景までの距離)

Variable	Variance of Proposal Distribution	Observation Model	Mean of residual	Mean of variance
Distance $\frac{1}{I} \sum_p (d_p - \hat{d}_p)^\lambda$	2.5 ²	$\lambda=0.5$	1	12
		$\lambda=1.0$	1	12
		$\lambda=2.0$	1	12
	5 ²	$\lambda=0.5$	3	44
		$\lambda=1.0$	3	44
		$\lambda=2.0$	3	44
	7.5 ²	$\lambda=0.5$	4	98
		$\lambda=1.0$	4	98
		$\lambda=2.0$	5	98
	50 ²	$\lambda=0.5$	23	3306
		$\lambda=1.0$	23	3332
		$\lambda=2.0$	22	3365
Ratio $\frac{1}{I} \sum_p \left(\frac{ d_p - \hat{d}_p }{\hat{d}_p} \right)^\lambda$	2.5 ²	$\lambda=0.5$	1	11
		$\lambda=1.0$	1	11
		$\lambda=2.0$	1	11
	5 ²	$\lambda=0.5$	3	44
		$\lambda=1.0$	3	44
		$\lambda=2.0$	4	43
	7.5 ²	$\lambda=0.5$	5	98
		$\lambda=1.0$	6	97
		$\lambda=2.0$	7	95
	50 ²	$\lambda=0.5$	25	3204
		$\lambda=1.0$	26	3154
		$\lambda=2.0$	27	3078

数値実験結果(距離ヒストグラム)

Histogram	Variance of Proposal Distribution	Observation Model	Mean of residual	Mean of variance
D1	2.5 ²	Linear	1	11
		Exp5	2	11
		Exp10	2	11
	5 ²	Linear	4	42
		Exp5	5	42
		Exp10	6	41
	7.5 ²	Linear	7	94
		Exp5	8	94
		Exp10	10	92
	50 ²	Linear	30	3929
		Exp5	36	3734
		Exp10	43	3350
D2	2.5 ²	Linear	1	11
		Exp5	2	11
		Exp10	2	11
	5 ²	Linear	4	42
		Exp5	4	42
		Exp10	5	41
	7.5 ²	Linear	7	94
		Exp5	7	93
		Exp10	8	91
	50 ²	Linear	29	3975
		Exp5	31	3906
		Exp10	32	3838

- ◆ 全体的にはどのモデルを選んでもあまり差はない
 - ◆ 予測分布に対して不自然な挙動をするモデルは無い
 - ◆ モデル形式による差異も目立たない
 - ◆ 前景までの距離を用いる場合、 λ の値による違いは僅か
 - ◆ 色情報では予測分布の分散が $(7.5\text{cm})^2$ まで、
距離情報では $(50\text{cm})^2$ でも事後分布の分散が増加する
- 現状では、モデル自体の定性的な傾向、
観測情報の密度や計算速度により選ぶのが妥当
- ◆ モデル間にいくつかの相違点も存在する
 - ◆ ヒストグラムを用いた観測モデルでは、指数型の方が残差の平均値が大きくなり、分散の平均値が小さくなる
 - ◆ 色情報より距離情報のほうが、予測分布の分散が増加したときの事後分布の残差の平均は増加しにくい

- ◆ 従来議論が行われていなかった、一般状態空間モデルにおける観測モデルのモデリングに関する基礎的検討を行った
 - ◆ 人物抽出手法におけるモデルの一般状態空間モデルにおける位置づけを行った
- ◆ 観測モデル候補それぞれの特性をもとに、利用するモデルを決定し、モデリングを行った
- ◆ それぞれの観測モデルが、予測分布に対してどのような事後分布を算出するか、72通りの数値実験を行った
 - ◆ 全体としては、観測モデルによって算出される事後分布に大きな差異がないことが確認された

- ◆ モデルのさらなる評価
 - ◆ 事後分布の分布形状の確認
 - ◆ 双子実験
- ◆ 場の状況に応じたモデル選択手法と、
モデル選択結果による場の状況解釈手法への転換
 - ◆ 各観測モデルの定性的な傾向が分かっている
→場の状況を場合分けし、観測モデルを使い分けられないか？
 - ◆ 良いモデルを定量的に評価可能であれば、
選択されるモデルから場の状況を場合分けできないか？
- ◆ そのためには
 - ◆ 定量的な比較方法の検討
 - ◆ 追跡成功率を上げるモデルの把握
- ◆ システムモデルに関しても全く同様の議論があり得る