情報デザイン専攻

画像情報処理論及び演習II

パターン認識の基礎
形状検出・機械学習

第10回講義
水曜日１限
教室6218

吉澤　信
shin@riken.jp, 非常勤講師
大妻女子大学　社会情報学部

独立行政法人 理化学研究所

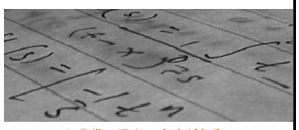
今日の授業内容

www.riken.jp/brict/Yoshizawa/Lectures/index.html
www.riken.jp/brict/Yoshizawa/Lectures/Lec22.pdf

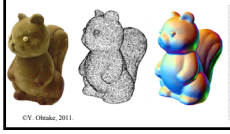
1. 形状検出.
2. パターン認識

形状検出

- ✓ 特徴点・コーナー点・エッジの抽出は勾配やHessian等の微分オペレータを近似して、その固有解析や極値探索を行うのが基本.
- ✓ 円や楕円等の簡単な形状は関数のフィッティングを行うのが基本.
- ✓ 一般の形状検出は曲面再構成等の補間法の問題.

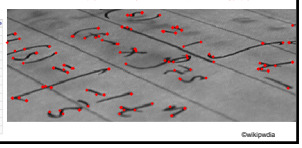


✓ 非常に沢山の方法があるので今回は代表的な方法だけ紹介.



©Y. Ohtake, 2011.

Feature detector	Edge	Corner	Blob
Canny	X		
Sobel	X		
Harris & Stephens / Plessey	X	X	
SUSAN	X	X	
Shi & Tomasi		X	
Level curve curvature		X	
FAST		X	
Laplacian of Gaussian		X	X
Difference of Gaussians		X	X
Determinant of Hessian		X	X
MSER			X
PCBR			X
Grey-level blobs			X



©wikipedia

テンプレートマッチング

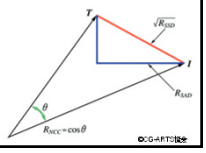
- ✓ T, I ＝$T(i,j)$, $I(i,j)$を並べたベクトル.
- ✓ SAD＝市街地距離.
- ✓ SSD＝ユークリッド距離の2乗.
- ✓ NCC(正規化相互相関)＝正規化されたベクトルの内積＝なす角の余弦.

$$R_{SAD} = \sum_{j=0}^{N-1}\sum_{i=0}^{M-1} |I(i,j) - T(i,j)|$$

$$R_{SSD} = \sum_{j=0}^{N-1}\sum_{i=0}^{M-1} (I(i,j) - T(i,j))^2$$

$$R_{NCC} = \frac{\sum_{j=0}^{N-1}\sum_{i=0}^{M-1} I(i,j)T(i,j)}{\sqrt{\sum_{j=0}^{N-1}\sum_{i=0}^{M-1} I(i,j)^2 \times \sum_{j=0}^{N-1}\sum_{i=0}^{M-1} T(i,j)^2}}$$

$$T = \begin{pmatrix} T(0,0) \\ T(0,1) \\ \vdots \\ T(M-1,N-1) \end{pmatrix}$$



©CG-ARTS協会

テンプレートマッチング2

- ✓ ZNCC(相互相関係数):

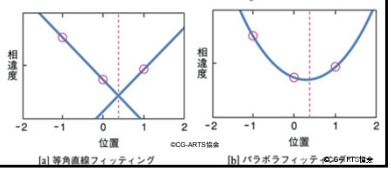
$$R_{ZNCC} = \frac{\sum_{j=0}^{N-1}\sum_{i=0}^{M-1} (I(i,j) - \bar{I})(T(i,j) - \bar{T})}{\sqrt{\sum_{j=0}^{N-1}\sum_{i=0}^{M-1} (I(i,j) - \bar{I})^2 \times \sum_{j=0}^{N-1}\sum_{i=0}^{M-1} (T(i,j) - \bar{T})^2}}$$

- ✓ 画素単位で求められた相違度をフィッティング関数で補間し、フィッティング関数の最小を与える位置をサブピクセルで求める方法もある.

$$where \quad \bar{I} = \frac{1}{MN}\sum_{j=0}^{N-1}\sum_{i=0}^{M-1} I(i,j)$$

$$\bar{T} = \frac{1}{MN}\sum_{j=0}^{N-1}\sum_{i=0}^{M-1} T(i,j)$$

CK, Hotta, ICPR 2006.



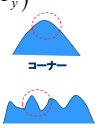
[a] 等角直線フィッティング　©CG-ARTS協会　[b] パラボラフィッティング　©CG-ARTS協会

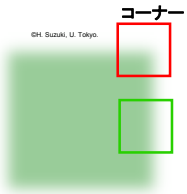
コーナー検出(Harris)

- ✓ Harrisの方法: 共役計量の固有値を使って凹凸を検出.

$$C = \nabla I \nabla I^T = \begin{pmatrix} I_x \\ I_y \end{pmatrix} \begin{pmatrix} I_x & I_y \end{pmatrix}$$

$$= \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}$$





©H. Suzuki, U. Tokyo.

画素値が様々な方向に変化している.

- ✓ ガウス関数の重み付平均を使う.

$$M = \begin{pmatrix} G_\sigma(I_x^2) & G_\sigma(I_x I_y) \\ G_\sigma(I_x I_y) & G_\sigma(I_y^2) \end{pmatrix} = G_\sigma(\nabla I \nabla I^T)$$

$$\approx \sum_{i,j} w_{i,j} \nabla I(i,j) \nabla I(i,j)^T = \sum_{i,j} w_{i,j} \begin{pmatrix} I_x(i,j)^2 & I_x(i,j) I_y(i,j) \\ I_x(i,j) I_y(i,j) & I_y(i,j)^2 \end{pmatrix}$$

det(M): ガウス曲率の近似.
trace(M): 平均曲率の近似.

det(M)/trace(M)やdet(M)-k*trace(M)など.

コーナー検出(Harris)2

$$R = \lambda_1\lambda_2 - k(\lambda_1+\lambda_2)^2 = \det M - k(trace(M))^2$$

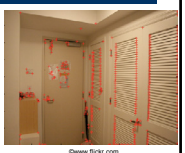
- $\lambda_1\lambda_2$⇒近傍における勾配(方向)の広がり:
  - 勾配が一つの方向に揃っていれば, 0.
  - 様々な方向に広がっていれば, 大きな値.
- $(\lambda_1+\lambda_2)^2$⇒近傍の勾配の和の大きさ:

$$\lambda_1+\lambda_2 = \mathrm{trace}(M) = G_\alpha(I_x^2)+G_\alpha(I_y^2)$$

- kは, これらの二つ値の調整.

✓ Rが大きいところ＝輝度分布に大きな凹凸.

✓ 局所的にRが最大になるところがコーナー.

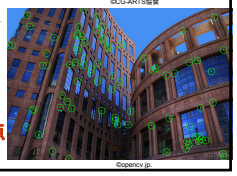

©www.flickr.com



©CG-ARTS協会


©de.academic.ru

←通常非常に沢山のコーナー点が検出されるので、閾値処理などで顕著な点だけを使う(SIFT等).

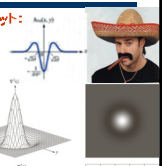

©opencv.jp.

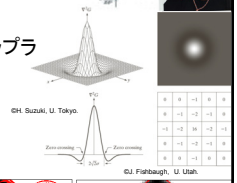
Blob(小塊)検出(LoG)

✓ Laplacian of Gaussian (LoG):
- エッジ=ラプラシアンのゼロ交差:極値探索.
- ノイズを強調してしまう.
- ガウシアンフィルタで平滑化してからラプラシアンを計算.
- この二つのフィルタは次式でまとまる.

$$h_{\log}(x,y) = \frac{x^2+y^2-2\sigma^2}{2\pi\sigma^6}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

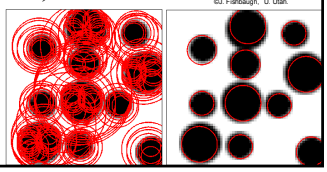
メキシカンハット:



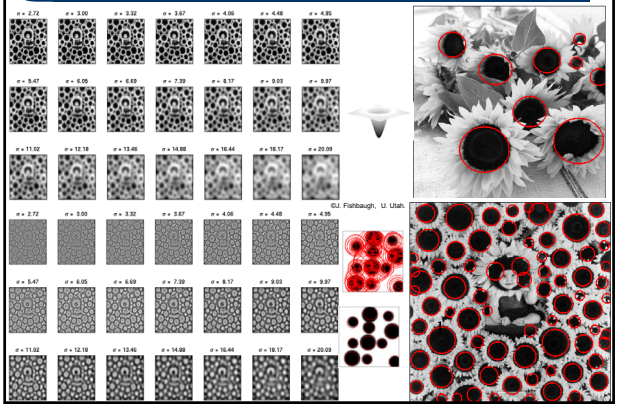

©H. Suzuki, U. Tokyo.
©J. Fishbaugh, U. Utah.

✓ 局所的に円を作成し最適化等の後処理でBlobを検出.

✓ 通常スケールスペース等と一緒に使う.



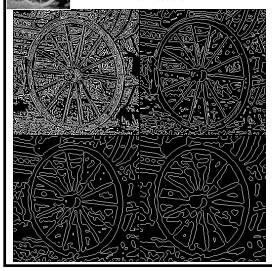
Blob(小塊)検出(LoG)2

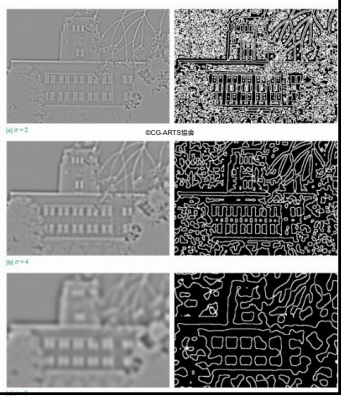

©J. Fishbaugh, U. Utah.

エッジ検出(LoG)

✓ Marr-Hildreth法:
LoG+ゼロ交差検索.


©www.cs.ubc.ca

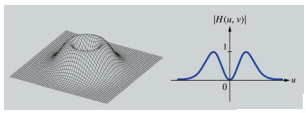


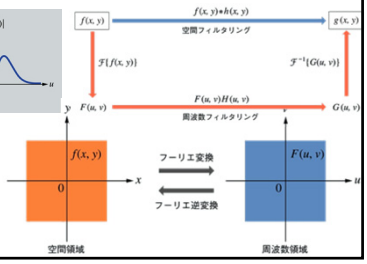

©CG-ARTS協会

LoG:周波数領域

空間 $$h_{\log}(x,y) = \frac{x^2+y^2-2\sigma^2}{2\pi\sigma^6}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

周波数 $$H_{\log}(u,v) = -4\pi^2(u^2+v^2)\exp\left(-2\pi^2\sigma^2(u^2+v^2)\right)$$



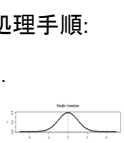


$$f * g = F^{-1}[F[f]F[g]]$$

Canny Edge検出

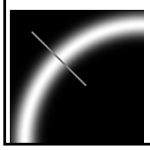
✓ Cannyアルゴリズムの処理手順:
- ノイズ低減と微分.
- 勾配の最大方向の検出.
- 閾値処理.

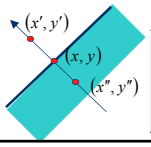
$$g(x,y) = \sqrt{G_x(f(x,y))^2+G_y(f(x,y))^2}$$
$$d(x,y) = tan^{-1}\left(\frac{G_y(f(x,y))}{G_x(f(x,y))}\right)$$

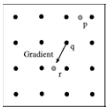



[a] 入力画像
[b] $\sigma=0.7$, $Th_{low}=50$, $Th_{high}=10$
解説:Th_{low}とTh_{high}が逆
©CG-ARTS協会
[c] $\sigma=0.3$, $Th_{low}=50$, $Th_{high}=10$
[d] $\sigma=0.7$, $Th_{low}=20$, $Th_{high}=10$
[d] キャニーのエッジ検出

✓ 勾配方向に沿って、勾配は局所最大か否かをチェック(非最大エッジ抑制).

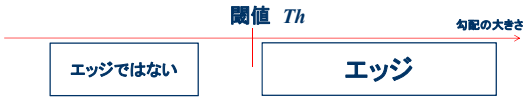




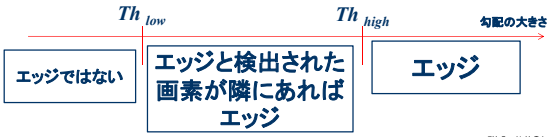


Canny Edge検出2

✓ 単純な閾値処理:

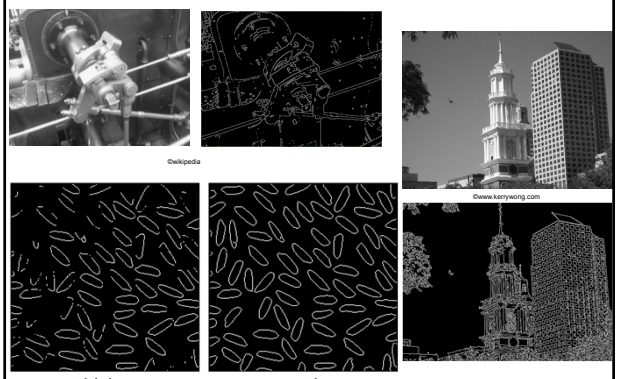


✓ エッジが途切れる問題があるので,



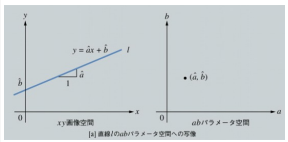
©H. Suzuki, U. Tokyo

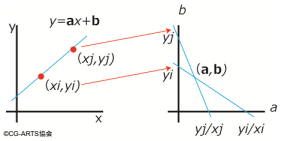
Canny Edge検出3



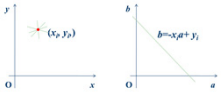
図形検出:ハフ変換

✓ 抽出されたエッジは不連続, 誤検出があるので、その様な結果から形状を検出する方法.





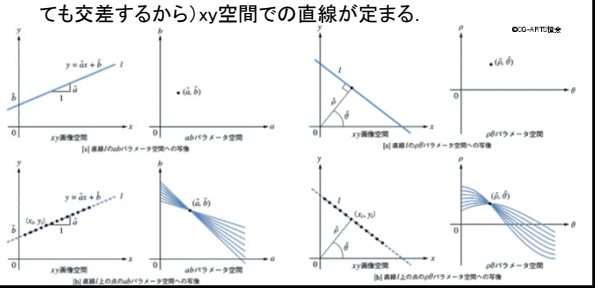
*xy*画像空間では直線だが, *ab*パラメータ空間では点になる



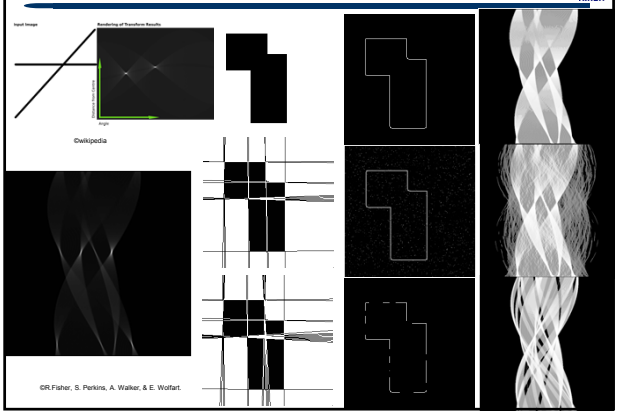
✓ *ab*空間の直線上の各点は, それぞれ*xy*空間で、この点●を通る直線に対応.

図形検出:ハフ変換2

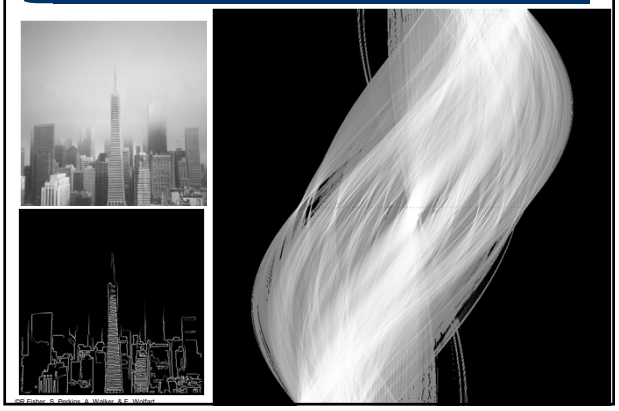
✓ ハフ変換の原理:
- xy空間上の直線を構成している点群の各点をab空間に写像すると, それらの直線は唯一の点で交差する.
- 逆に, このような交点を検出すると, (直線が千切れていても交差するから) xy空間での直線が定まる.



図形検出:ハフ変換3



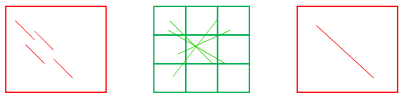
図形検出:ハフ変換4



図形検出：ハフ変換5

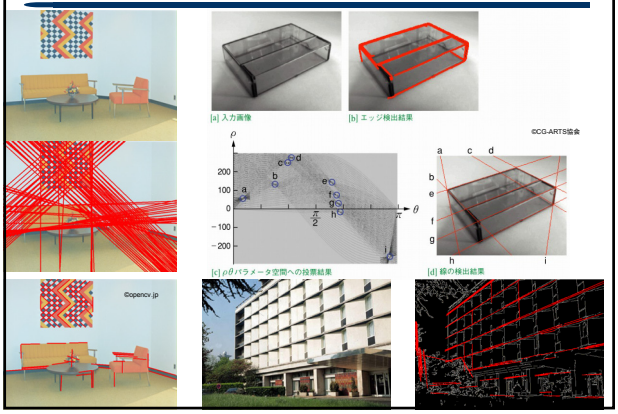
1. エッジ検出を行い，閾値処理により2値画像を用意する．
2. *ab*空間をセルに分割する．セルの値をゼロにしておく．
3. 線の候補となる画素を*ab*空間に写像し，その直線の通るセルに関してはセルの値を1増やす．（投票）
4. 全ての候補についてステップ2を行う．通過回数が記録される．これを投票度数と言う．
5. 投票度数の大きなセルを探索する．これが*xy*空間の直線に対応する．

©H. Suzuki, U. Tokyo.



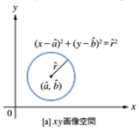
*xy*空間　*ab*空間　*xy*空間

図形検出：ハフ変換6

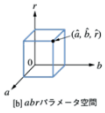


図形検出：ハフ変換7

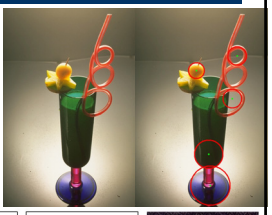
- ✓ 円検出の場合はパラメタは3つ．
- ✓ 一般化ハフ変換もある



[a] *xy*画像空間



[b] *abr*パラメータ空間

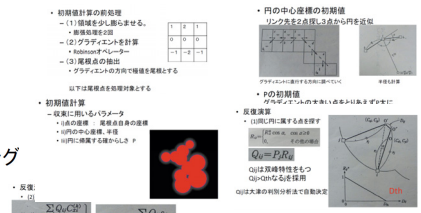


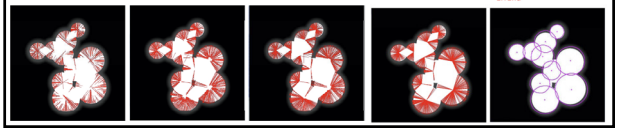
[a] 入力画像　[b] エッジ検出結果　[c] 1セントコイン（銅）に対する投票結果　[d] 25セントコイン（銀）に対する投票結果　[e] 検出結果

©CG-ARTS協会

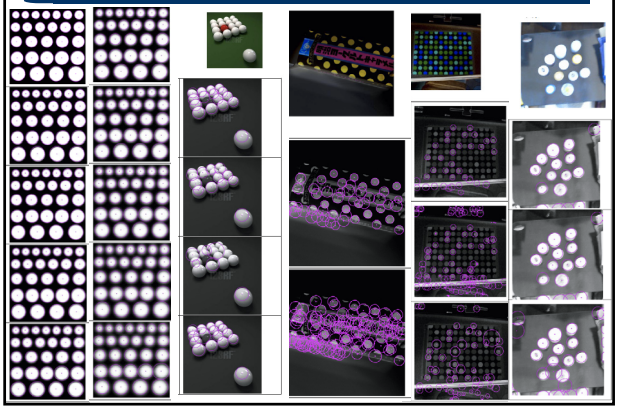
円検出(緩和法)

- 処理の流れ
  - 画像の2値化
  - 初期値計算
  - 反復演算
  - 円のクラスタリング



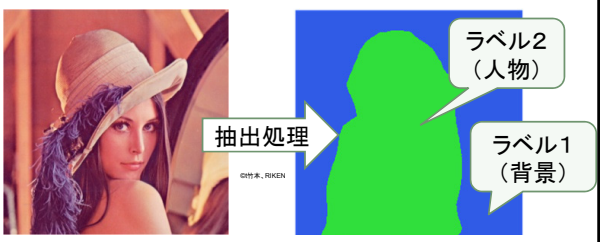


円検出(緩和法)2



領域抽出とは？

- ✓ 領域抽出：画像の領域を分割する処理・対象の領域を切り出して他の領域と区別する事．
- ✓ 画像処理で最も重要な技術．
- ✓ 毎年何百！という新しい方法が提案されている．

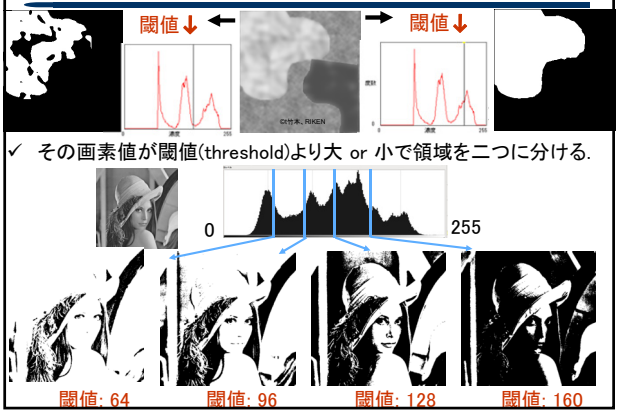


Shin Yoshizawa: shin@riken.jp

領域抽出の例

Shin Yoshizawa: shin@riken.jp

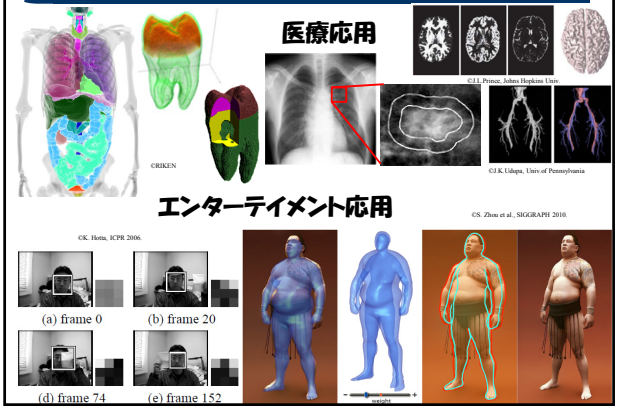
一番簡単な領域抽出：閾値による二値化



✓ その画素値が閾値(threshold)より大 or 小で領域を二つに分ける.

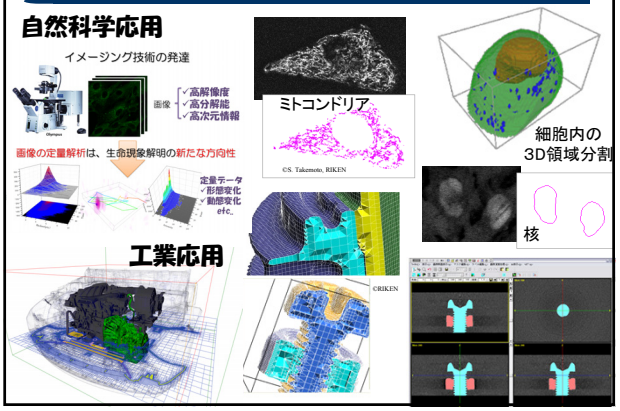
Shin Yoshizawa: shin@riken.jp

何の役に立つのか？



Shin Yoshizawa: shin@riken.jp

何の役に立つのか？2



Shin Yoshizawa: shin@riken.jp

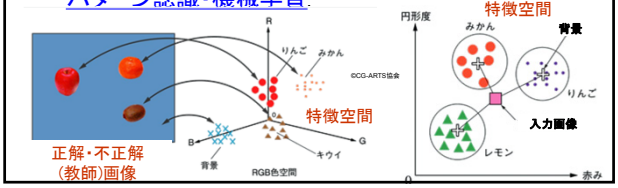
領域抽出法の分類

✓ 教師なし(Unsupervised Segmentation):

入力画像
(領域抽出したい画像)
⇒ **特徴抽出 分類・識別**
⇒ **領域抽出画像**

✓ 教師あり(Supervised Segmentation):

- パターン認識・機械学習.

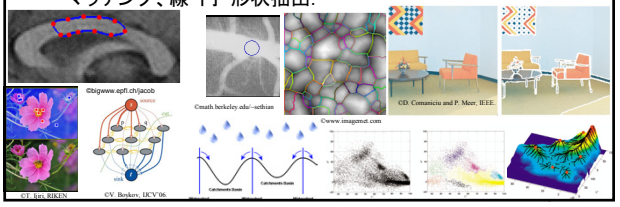


Shin Yoshizawa: shin@riken.jp

領域抽出法の分類

✓ 教師なし(Unsupervised Segmentation):

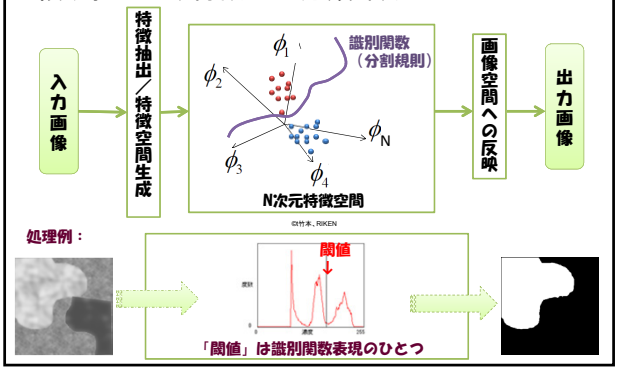
- 領域の輝度値や抽出したい形状に関するエネルギー(目的関数)を最小化・最大化する事で特徴量の分布や滑らかさを基準.
- 領域抽出でよく用いられる方法は**大津の二値化法**, **Snake (Active Contour), Graph Cuts, Mean Shift, Water Shed (Region Growing)**等の方法が有名(目的関数の違いなど沢山の亜種).
- モデルを用いた検出：エッジ抽出、コーナー検出、テンプレートマッチング、線・円・形状抽出.



Shin Yoshizawa: shin@riken.jp

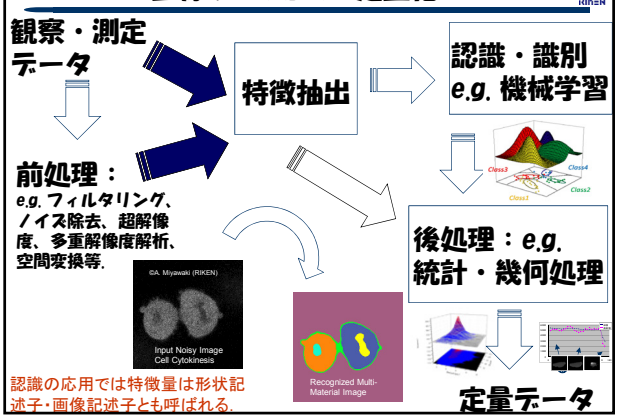
領域抽出処理の流れ

✓ 領域抽出は、特徴量の分類・識別.



Shin Yoshizawa: shin@riken.jp

画像データからの定量化

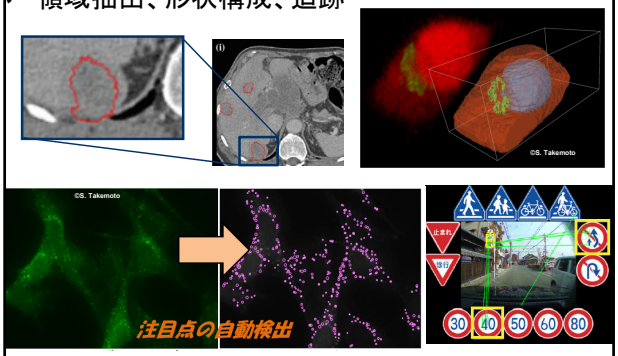


認識の応用では特徴量は形状記述子・画像記述子とも呼ばれる.

Shin Yoshizawa: shin@riken.jp

特徴量を用いた定量解析(認識/識別あり)

✓ 領域抽出、形状構成、追跡…



特徴間の距離にてオブジェクトを認識

Shin Yoshizawa: shin@riken.jp

類似度(相関)＝画像間の距離

類似度＝画像間の距離
＝特徴ベクトル間の距離

$$dist(\mathbf{X},\mathbf{Y}) = \left\| \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_n \end{pmatrix} \right\|$$

様々な距離：
- ✓ ユークリッド: L1, L2, Lmax…
- ✓ カイ2乗, エントロピー,…
- ✓ Bhattacharyya, Minkowski,…
- ✓ Earth Mover's, 相関係数…
- ✓ Etc.

Y. Rubner et al. *Earth Mover's Distance*, IJCV, 2000.

Shin Yoshizawa: shin@riken.jp

特徴量を用いた定量解析(認識なし)

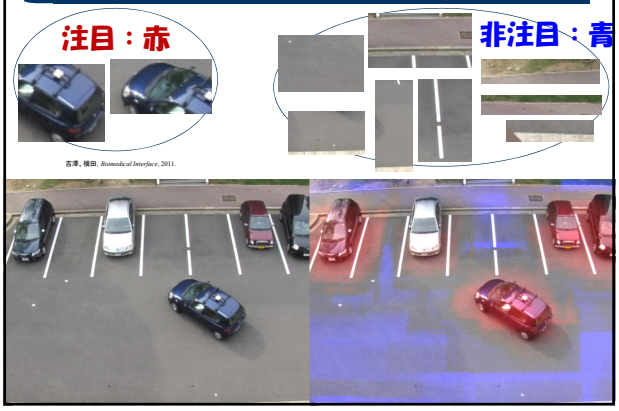
✓ Google等の画像検索：リトリーバル

✓ 物体追跡：パーティクルフィルタ等の尤度計算

✓ 画像内の局所領域間の類似度を物体移動の確率分布に利用

Shin Yoshizawa: shin@riken.jp

特徴量による類似度マップ



吉澤, 横田, *Biomedical Interface*, 2011.

Shin Yoshizawa: shin@riken.jp

特徴量の種類

1. 普遍的な特徴: 基本的に数学の分野別.
   - ✓ 幾何特徴:長さ、面積、曲率、計量テンソル、測度、オイラー数、Index…
   - ✓ 解析特徴:Gradient、Hessian、Lapalcian、フーリエ係数、球面調和関数、Wavelet…
   - ✓ 代数特徴:固有値、行列式、階数…
   - ✓ 統計特徴:平均、分散、頻度、相関係数…
   - ✓ etc. …組み合わせも有り e.g. 曲率のGrad…
2. 問題依存(Ad hoc)の特徴:
   - ✓ 実験的・経験的に提唱された量.
   - ✓ Specificな応用のみで意味がある量:

   e.g. 生物遺伝解析等で用いられている画像記述子(パラメータ).

Shin Yoshizawa: shin@riken.jp

特徴量の種類2

同じ特徴(e.g.勾配:Gradient)でも…

1. 連続で滑らかな特徴: $\nabla I(x,y) = (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})$
2. 離散化・簡略化した特徴:

$(\frac{I(x+1,y)-I(x,y)}{dx}, \frac{I(y+1)-I(x,y)}{dy})$ 前進1次

$(\frac{I(x+1,y)-I(x-1,y)}{dx}, \frac{I(x,y+1)-I(x,y-1)}{dy})$ 中心2次

差分近似

Sobel

Prewitt

オペレータ

$(I*\frac{\partial g}{\partial x}, I*\frac{\partial g}{\partial y})$

$g = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{x^2+y^2}{2\sigma^2})$

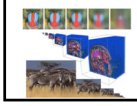
スケールスペース近似

©Y. Ohtake, SGP'05.

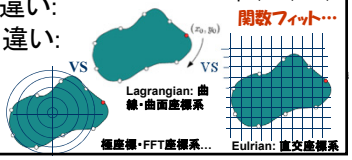
Moving-LS, RBF, Polynomials, Conv., Fourier, Béziers, B-Splin, MPU, SLIM,…

関数フィット…

3. 対象・座標系による違い:
4. 解像度・スケールの違い:



- ✓ 多重解像度表現
  - -Gaussian Pyramid
  - -Laplacian Pyramid
  - -Wavelets/Fourier,…
  - -etc.
- ✓ マルチスケール表現

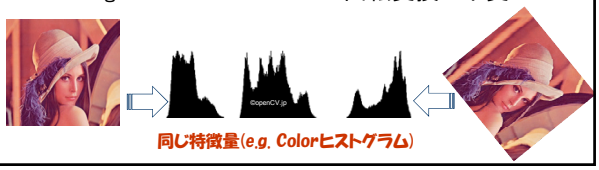


Shin Yoshizawa: shin@riken.jp

特徴量の種類3

- ✓ パターン認識ではSaliency, HLAC, PARCOR, SIFT, HOG, Shape-lets, テクスチャー統計量, 関数展開系(フーリエ, Wavelets, 球面調和, Zernike関数, KL展開, 固有関数展開などの係数列)….
- ✓ 重要な要素:不変量、性質(加法性:画像を足したら、特徴量も足される).

e.g. Rotational Invariants: 回転変換に不変

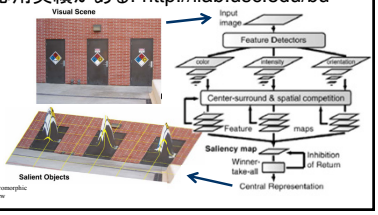


同じ特徴量(e.g. Colorヒストグラム)

Shin Yoshizawa: shin@riken.jp

Saliency

- ✓ 位相・幾何変化にロバストな方法の一つ:顕著性(Saliency).
- ✓ Visual Attention: 人間の脳内では視覚情報の空間的局所領域に対する顕著性に対して特定の解析システム(visual map encoding)が存在 [Human Neurobiol,4:1985].
- ✓ Saliency Model: L. Itti, C. Koch, and E. Niebur [IEEE PAMI, 20(11):1998]によって提案された顕著性計算モデルで2次元時系列解析にて既に多数の応用実績がある: http://ilab.usc.edu/bu
  - 方向、カラー特徴
  - 多重解像度解析
    - 大きさの異なる特徴
  - スケール間解析
    - 画像の自己相似度
    - 各スケール特徴統合



©USC, iLab C++ Neuromorphic Vision Toolkit Overview

Shin Yoshizawa: shin@riken.jp

Saliency2

- ✓ 顕著度・特徴点を抽出.

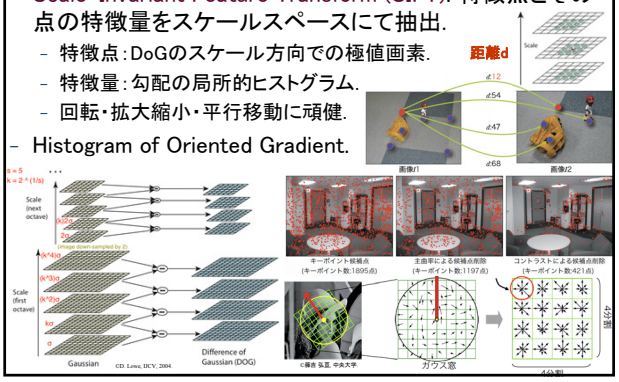
©USC, iLab C++ Neuromorphic Vision Toolkit Overview

Shin Yoshizawa: shin@riken.jp

SIFT & HOG

- ✓ Scale-Invariant Feature Transform (SIFT): 特徴点とその点の特徴量をスケールスペースにて抽出.
  - 特徴点:DoGのスケール方向での極値画素.
  - 特徴量:勾配の局所的ヒストグラム.
  - 回転・拡大縮小・平行移動に頑健.
- Histogram of Oriented Gradient.

距離d



CD. Lowe, IJCV, 2004.

HLAC

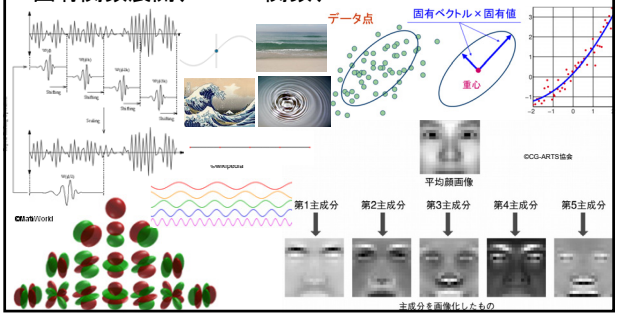
- ✓ 高次局所自己相関(HLAC: Higher-order Local Auto-Correlation)特徴: 局所パターンの自己相関を積分.
  - 積分ベース:(局所)画像の大きさによらず一定次元の特徴.
  - 有名な産総研(旧電総研)の大津先生と栗田先生が開発.
  - 位置不変性 & 加法性.
  - パターンの組み合わせ=次数.
  - 異常検出などの応用.

$$x(\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_N)=\int_p f(\mathbf{r})f(\mathbf{r}+\mathbf{a}_1)\cdots f(\mathbf{r}+\mathbf{a}_N)d\mathbf{r}$$

$\mathbf{r}$: 参照ベクトル　$\mathbf{a}_i$: 変位ベクトル
f : 輝度値パターン　p : 画面（積分範囲）

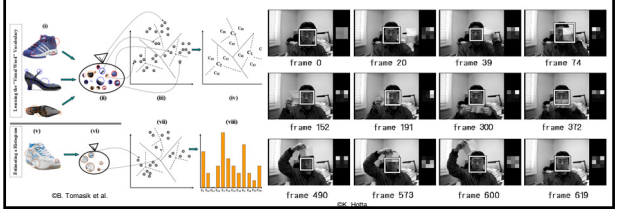
関数展開系

- ✓ 周波数を特徴(係数⇒ベクトル)とする:様々な基底がある.
- ✓ フーリエ係数、KL(Karhunen-Loeve)展開、球面調和関数、固有関数展開、Zernike関数、Wavelets….



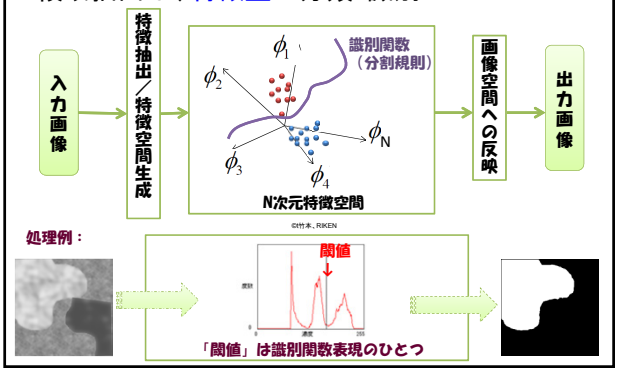
重要:Bag of Features/Keypoints

- ✓ 入力画像、教師画像や評価局所Window内の全ての画素での特徴量を使うのではなく、キーポイントでの特徴量をヒストグラムなどの統計量として使うアプローチ(Joint～).
- ✓ 「見え」の変化や物体の局所変形・移動などに頑健.
- ✓ 同じカテゴリーの複数画像から作成すると代表特徴となる.
- ✓ キーポイントはランダム、SIFT等で特徴量は多種.



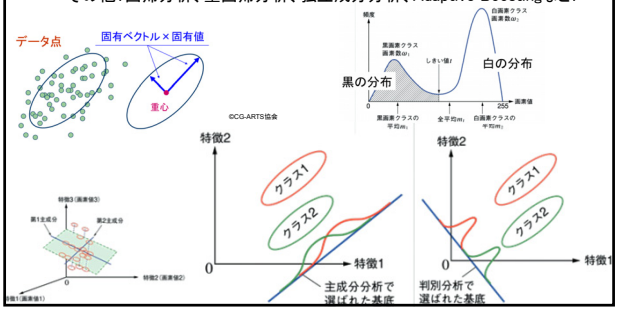
領域抽出処理の流れ

- ✓ 領域抽出は、特徴量の分類・識別.



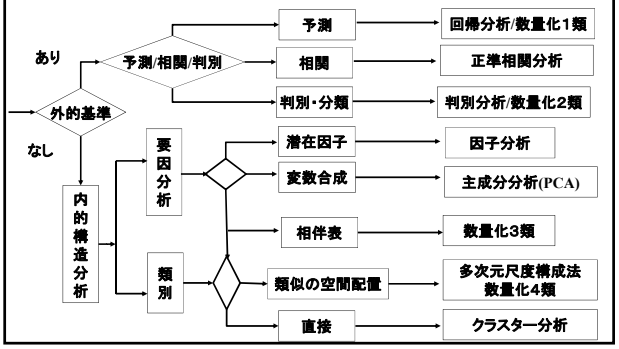
識別・分類・判別

- ✓ 多変量解析:多変量の統計的解析法.
  - 主成分分析(PCA):共分散行列(分散の二乗和)の固有解析.
  - 線形判別分析: 大津の二値化法と同じでクラス間・クラス内分散を用いる.
  - その他:回帰分析、重回帰分析、独立成分分析、Adaptive Boostingなど.



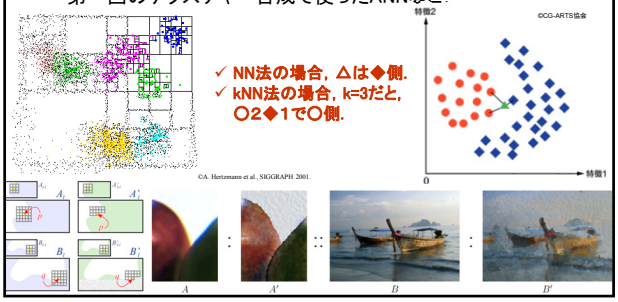
識別・分類・判別2

- ✓ 多変量解析手法の分類:
  - 大津 展之, "*HLAC特徴を用いた学習型汎用認識*", トリケップス, 2013.



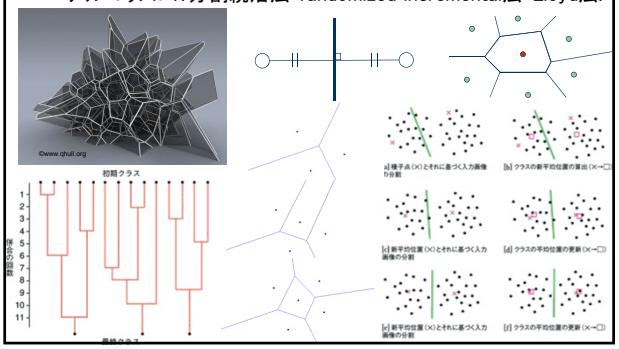
識別・分類・判別3

✓ kNN (k Nearest Neighbor)法：近傍へ分類.
- 近傍k個の学習用データを検索する.
- 帰属するサンプル数が最も多いクラスへ識別する.
- 第一回のテクスチャー合成で使ったANNなど.



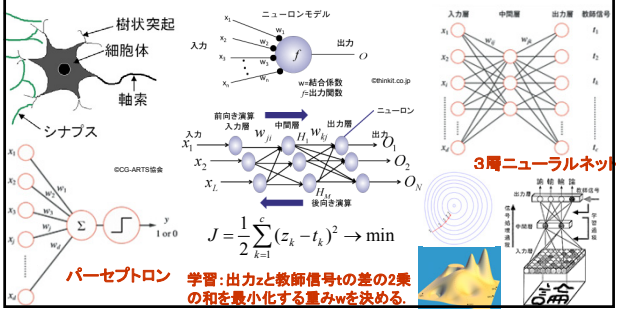
識別・分類・判別4

✓ k-means：平均で分ける⇒重心Voronoi図になる.
- アルゴリズム：分割統治法・randomized incremental法・Lloyd法.



識別・分類・判別5

✓ ニューラルネットワーク：脳の学習と識別のメカニズムをシステム化. シナプス結合を形成した人工ニューロンが、学習によって結合強度を変化させ、問題解決能力を持つ.



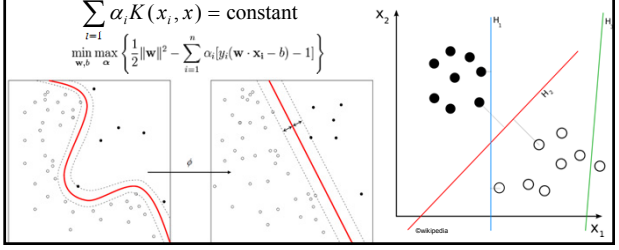
$$J = \frac{1}{2}\sum_{k=1}^{c}(z_k - t_k)^2 \rightarrow \min$$

識別・分類・判別6

✓ SVM(Support Vector Machine)：データ点との距離が最大となる分離(超)平面を求める.
- カーネル関数K(): 線形、多項式、放射基底関数など.
- K()が線形ならラグランジュの未定乗数法や非線形(2次)計画法を使って計算.

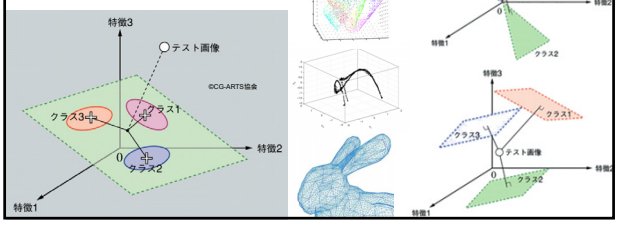
$$\sum_{i=1}^{N} \alpha_i K(x_i, x) = \text{constant}$$

$$\min_{\mathbf{w},b} \max_{\alpha} \left\{ \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{n} \alpha_i [y_i(\mathbf{w}\cdot\mathbf{x}_1 - b) - 1] \right\}$$



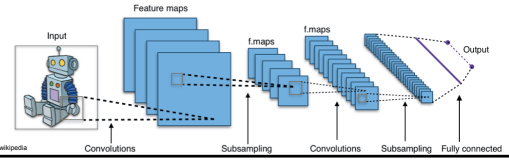
識別・分類・判別7

✓ 部分空間法：広すぎる特徴空間を簡略化.
- 部分空間・射影の選び方で様々な方法がある.
- 最も簡単な部分空間：主成分分析.
- 関数展開した高周波基底のカット.
- 空間の特徴を保持した簡略化.
- Dimension Reduction.



特徴抽出＋識別・分類・判別8

✓ ニューラルネットの復活：2010年代以降(Ciresan et al. CVPR' 12)にCNN (Convolutional Neural Network)等のDeep Learning と呼ばれる方法群が注目されて現在主流の機械学習の枠組みとなっている。
- 特徴抽出と識別・分類・判別を同時に行う、つまり、特徴も学習する方法群.

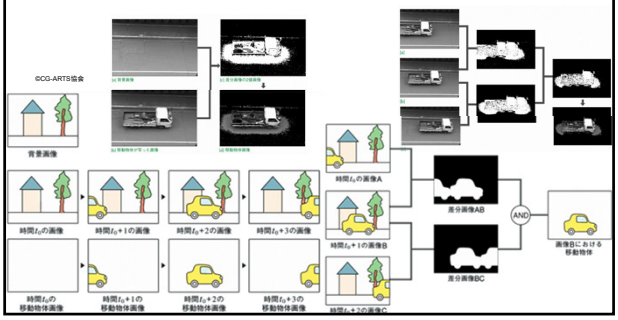


Shin Yoshizawa: shin@riken.jp

動画のパターン認識

✓ 基本は静止画のパターン認識法を高次元として適用する:
- 背景差分、オプティカルフロー、パーティクルフィルタ、確率論等.

✓ 背景・フレーム間差分：時間微分の差分近似.



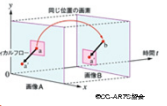
Shin Yoshizawa: shin@riken.jp

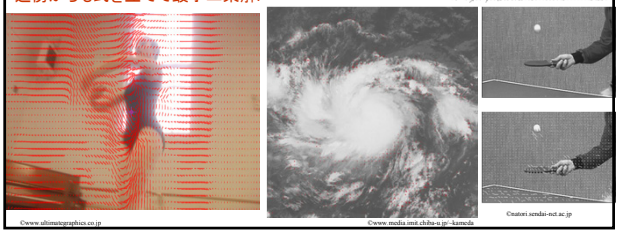
動画のパターン認識2

✓ オプティカルフロー: 移動物体の運動解析.
- ブロックマッチング法：テンプレートマッチング.
- 勾配法：

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} = 0,$$

近傍からも式を立てて最小二乗解.

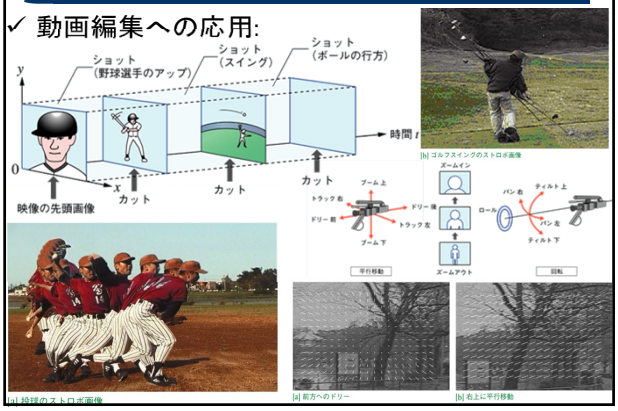




Shin Yoshizawa: shin@riken.jp

動画のパターン認識3

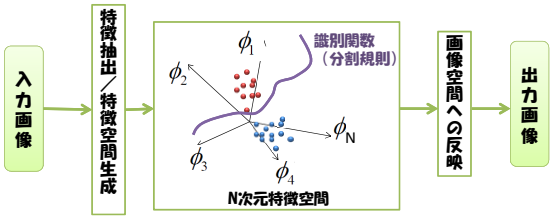
✓ 動画編集への応用:



Shin Yoshizawa: shin@riken.jp

まとめ

✓ パターン認識(領域抽出)は、特徴量の分類・識別.



✓ 目的に応じた特徴量の選択と計算方法の選択.
✓ 識別関数はSVMなどで可能な限り単純化するのが主流
→部分空間やBag of Featuresなど特徴量の拡張で認識の問題を解く.
✓ 2010年代以降はCNNなどのDeep Learningが主流.

Shin Yoshizawa: shin@riken.jp

おわりに、

みなさん良く頑張りましたd(>_・)

今日で本講義は終わりです.

みんな最後まで来てくれてありがとー
o(≧∇≦)o

またいつかお会いしましょう！
ゞ(^-^)ゞ