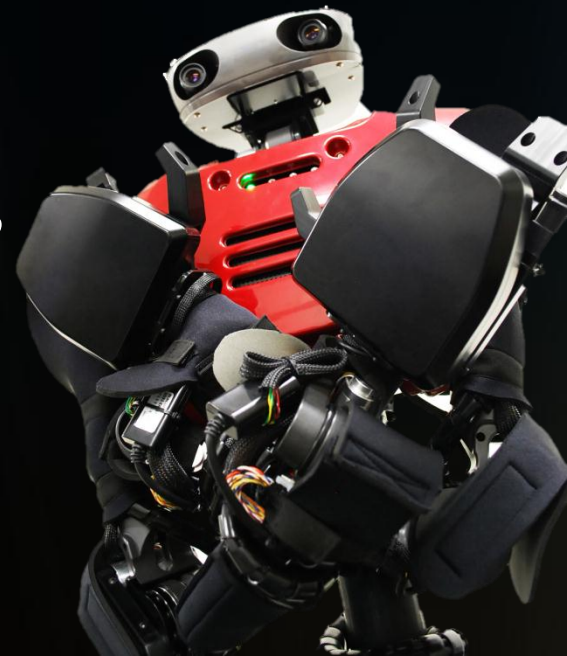


超高速オンライン転移学習

東京工業大学

木村大毅, Kankuekul Pichai,
Aram Kawewong, 長谷川修

2011年07月05日版



Introduction

■ 実世界の物体認識は極めて困難

現在、この問題を根本的に解決する「現実的な」手法は存在しない



実世界の難しさ

- 多種多様の物体
- 高ノイズ
- 新しい物が増える
- オクルージョン

など...

これまでは、Caltech 256 など
対象を個別に学習・認識させる試み



対象毎に、沢山の学習画像を用意



1



2



3



4



5



6



7



8



9



10



11



12



13



14



15



16



17



18



19



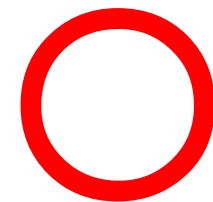
20

しかし、私たちの身の周りには膨大な数のモノがあり、その数は増え続けている



パラダイムシフトが必要！

- ・ 「月」に行くには？
 - 飛行機をどんなに改良しても「絶対」月には行けない



本研究の成果と意義

■ 下記を兼ね備えた、**現実的な**物体認識手法を構築

● **転移学習**を導入

- 実世界を少数の基本的な属性知識の組合せで認識
(例：赤＋球＝リンゴ)

● 命令するだけで「**自力で**・**直ちに**」賢くなる

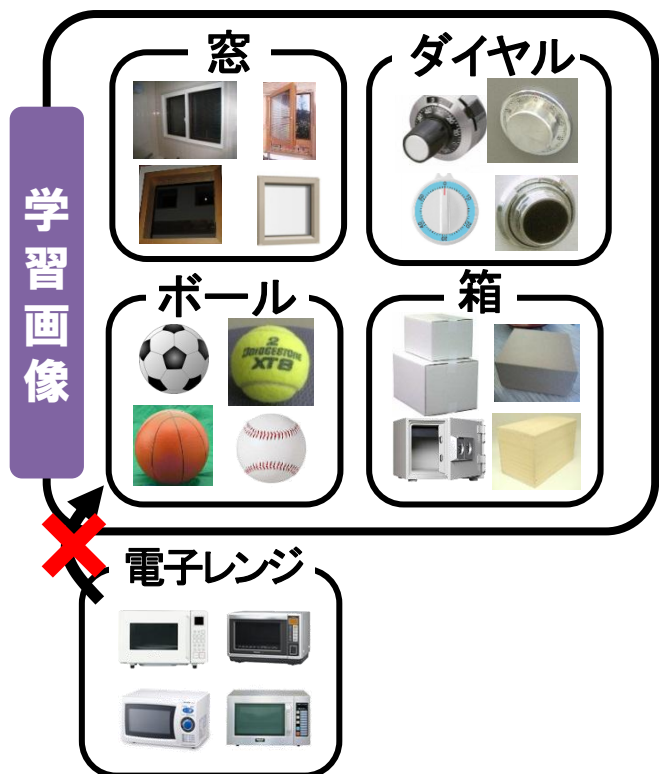
- 学習データはインターネットから自動収集
- **超高速**に**オンライン**追加学習&認識

● 画像以外にも**多様**・**Noisy**・**曖昧**な情報に対応

- 多様な曖昧情報を複合的に利用可能

Transfer learning (転移学習)

- 対象について学習するのではなく、**基本的な概念**（「属性」と呼ぶ）を学習
- 属性の組み合わせで**未学習物体**も認識



ダイヤルと窓がある箱型のもの = **電子レンジ**

人間が定義

Transfer learning (転移学習)

- この関係は**アルファベットと辞書**の关系到類似
 - 英語の場合、26文字の組合せで数十万の単語
 - 前頁の例では「ダイヤル+箱=金庫」もわかる
- さらに、提案手法は**ベクトルデータ全般を属性として超高速オンライン追加学習可能**
 - 画像、音声、各種センサデータ、モータの制御信号など、あらゆるパターンデータが入力可能

提案手法の実施例

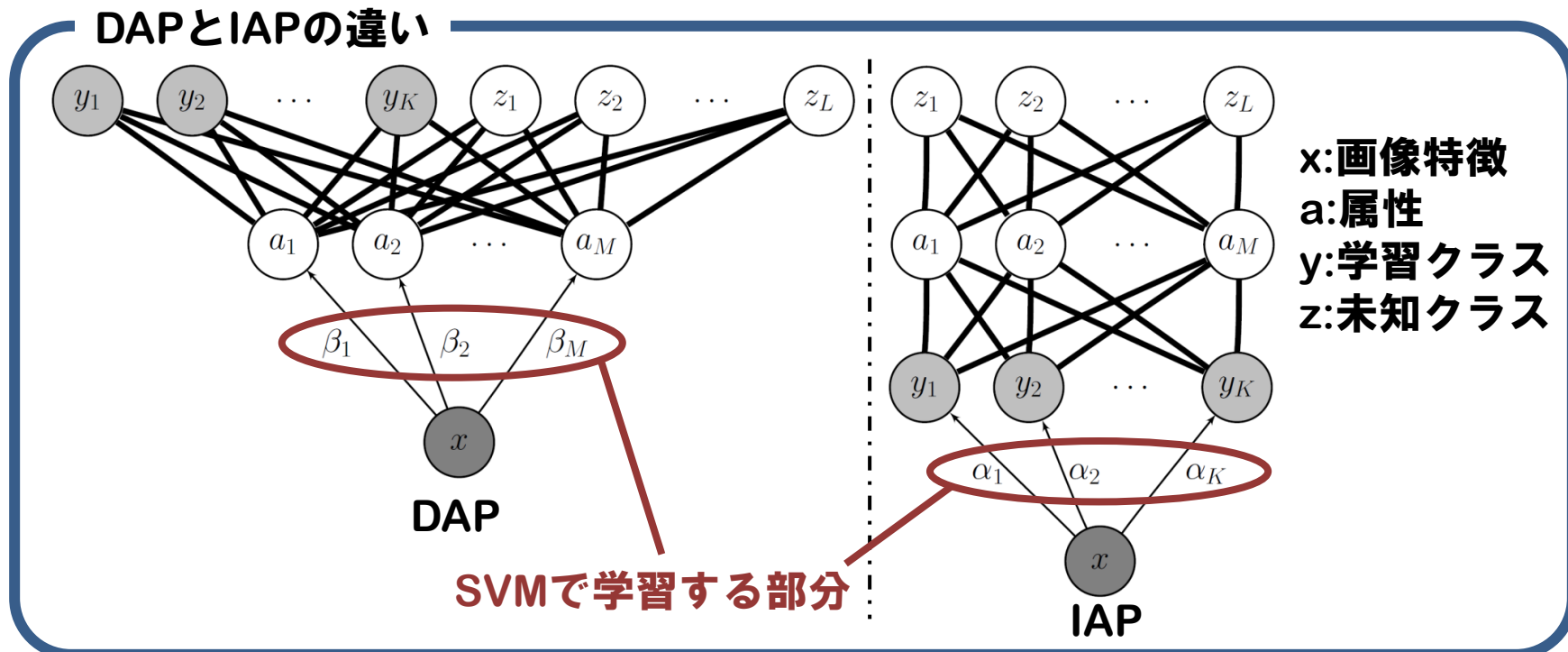
■ 従って、こんなことが出来る



Previous methods

■ DAP & IAP [Lampert et al., CVPR09]

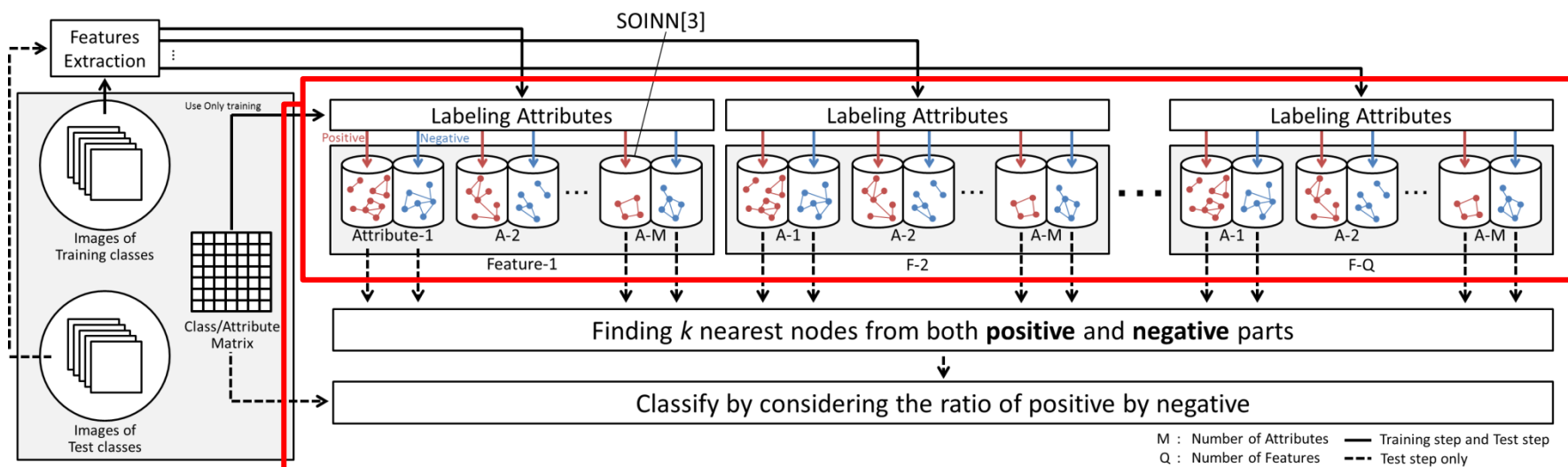
- SVMを用いて属性と画像特徴の関連性を学習し、それを用いて認識



Previous methods

■ AT-SOINN[Kawewong et al., IJCNN11]

- **SOINN**を用いて属性と画像特徴の関連性を学習し、それを用いて認識



SOINNの数 = $2 \times$ 属性の数 \times 特徴の数

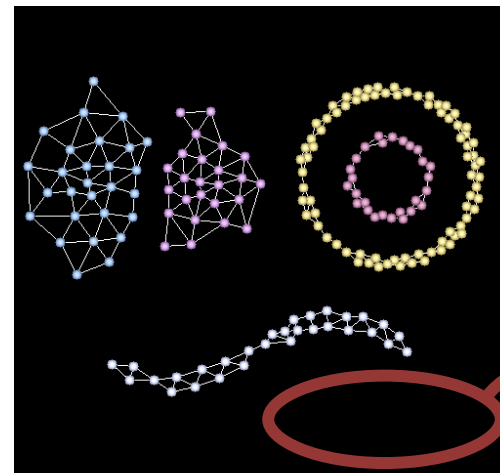
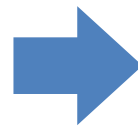
SOINN [Shen, Hasegawa, NN06]

■ 自己増殖型ニューラルネットワーク

- **オンライン**かつ**追加的**に学習可能
- クラス数や分布の形などの**事前知識が不要**
- 実世界での運用を考慮した**ノイズ耐性**
- **マルチモーダルなパターン**情報を記録



入力情報



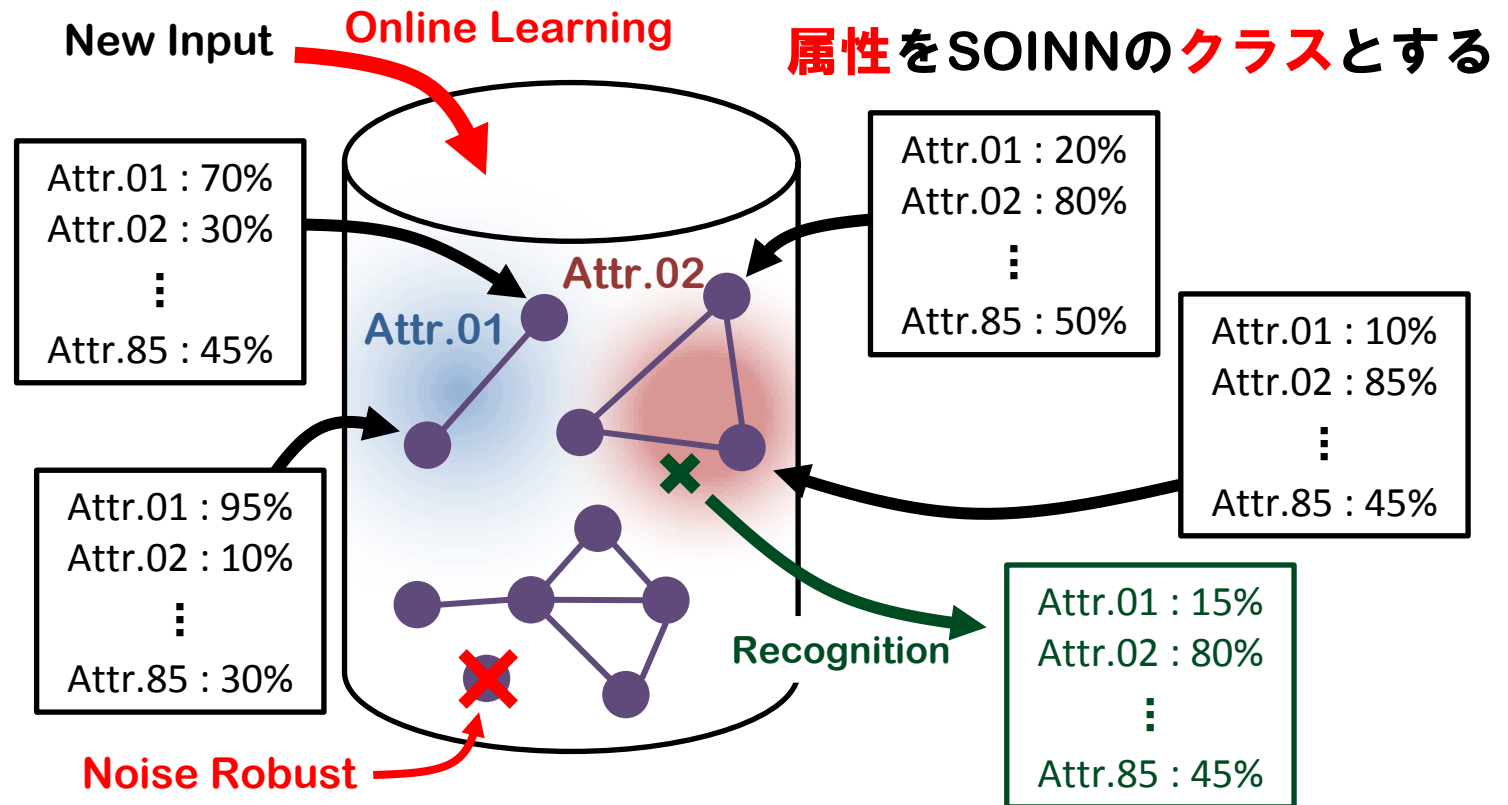
SOINN内の情報

強力な
ノイズ耐性

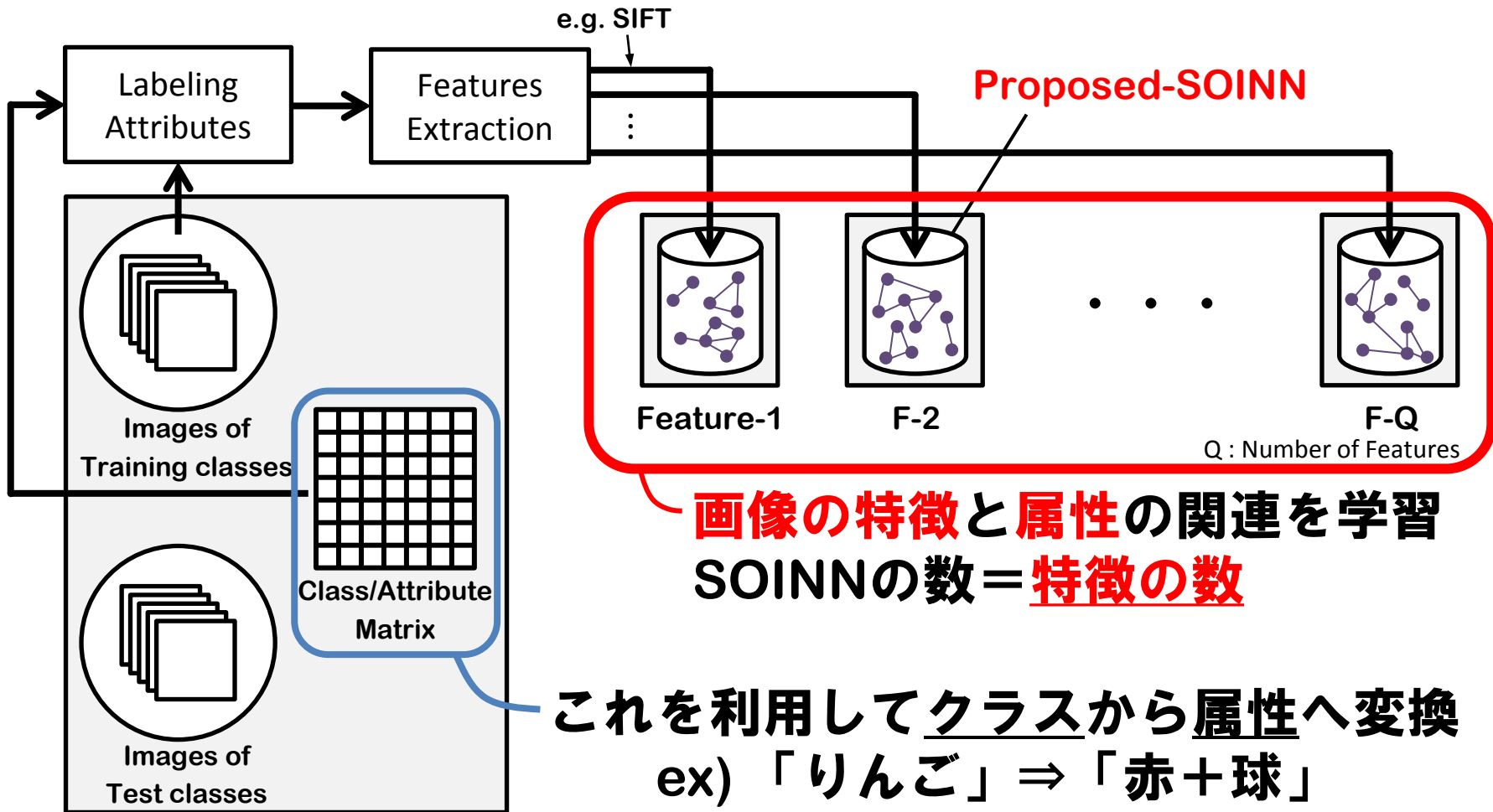
Proposed SOINN

■ AT-SOINN : 1つのノード=1つのクラス

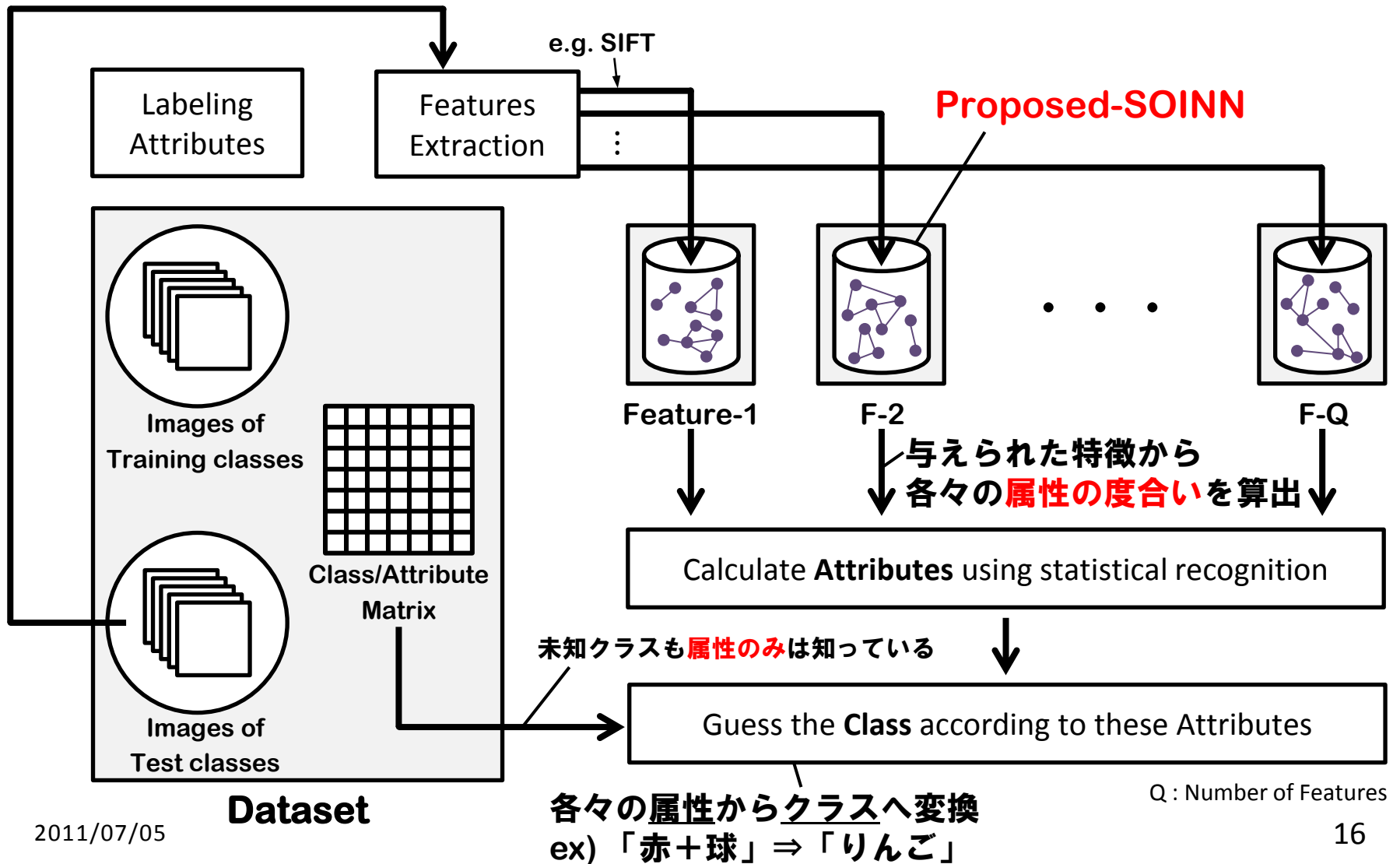
■ 提案: 1つのノード=**複数**のクラス (度合い)



Proposed Method (Training phase)



Proposed Method (Test phase)



Experiments

■ Lampertら[1]の動物の画像で比較実験

● 50種類の動物

- ライオン、シマウマ、豚、シロクマ、イルカなど

● 85種類の属性（基本的な概念）

- 肉食動物、尻尾がある、速く走る、賢いなど

● 6種類の画像特徴

- SIFT、SURF、pHOG、rgSIFT、LSS、CQのそれぞれのヒストグラム（Bag-of-Features）

※上で挙げた従来手法以外にも多くの論文で使用されている

Experiments

例) 未知の豚を当てる方法



Results

オンライン学習

バッチ学習

| | 提案手法 | AT-SOINN[2] | DAP[1] | IAP[1] |
|----------------|--------|-------------|--------|--------|
| 認識率 | 26.82% | 26.96% | 40.51% | 27.83% |
| 学習時間 | 7分 | 6時間 | >70日 | |
| 認識時間 | 1分半 | 4時間 | >2日 | |
| データ量 (ノード数) | 1,825 | 347,082 | - | |
| | | 99.47%削減 | | |
| SOINNの数 | 6 | 1020 | - | |

なお学習時間と認識時間は、特徴検出の時間を含めない

Results

オンライン学習

バッチ学習

| | 提案手法 | AT-SOINN[2] | DAP[1] | IAP[1] |
|--------|------|-------------|--------|--------|
| 曖昧な属性※ | ○ | × | × | × |
| 属性の追加 | ○ | △ | × | × |

※曖昧な属性とは、**連続値**での属性の定義が出来るという意味である

提案手法は、

54枚／秒で学習、**16ミリ秒／枚**で認識

更に、**曖昧な属性**や**柔軟な属性の追加**が可能

今後の課題

Multimodal 情報の利用

■ 曖昧な実世界パターン情報全般を属性として利用

- 人に近い、音声、温度、圧力、質量などの感覚情報もコンピュータやロボットに「経験」を通じて学習させることで...



人の持つより高次の概念の共有と理解

- 「ほんわか」「しゃきしゃき」などを理解させる
 - 「蝙蝠、人、ロボットであるとはどのようなことか？」

インターネットの活用

■ その場で、直ちに賢くなる

例) Google™

YAHOO!®

You Tube
Broadcast Yourself

● 高速性やオンライン学習性を活用

- 複雑な実環境では、有効な属性を事前に想定するのは困難
- 有効な属性を、人や環境とのインタラクションを通じて決定

● ネットの情報は極めてノイジー

- **SOINN**の**ノイズ耐性**を活用

● 大規模評価実験の実施を予定

主要参考文献

- [1] C. H. Lampert, H. Nickisch, and S. Harmeling, “Learning to detect unseen object classes by between-class attribute transfer”, CVPR 2009.
- [2] A. Kawewong, Sirinart Tangruamsub, Pichai Kankuekool and Osamu Hasegawa, “Fast Online Incremental Transfer Learning for Unseen Object Classification Using Self-Organizing Incremental Neural Networks”, The 2011 International Joint Conference on Neural Networks (IJCNN).
- [3] F. Shen, O. Hasegawa, “An Incremental Network for On-line Unsupervised Classification and Topology Learning”, Neural Networks 2006.