

人を観る技術の先端的研究

片岡 裕雄, *Ph.D.*

産業技術総合研究所

知能システム研究部門 コンピュータビジョン研究グループ

<http://www.hirokatsukataoka.net/>

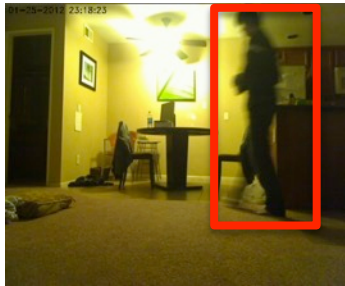
人を観る技術

研究に対する哲学：

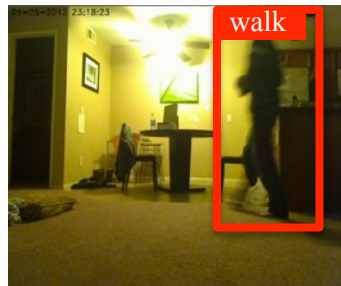
観測により知識を強化するだけでなく、知識により観測を強化

- － 「観測」のみでなく「累積された知識」を相互的に活用する枠組み

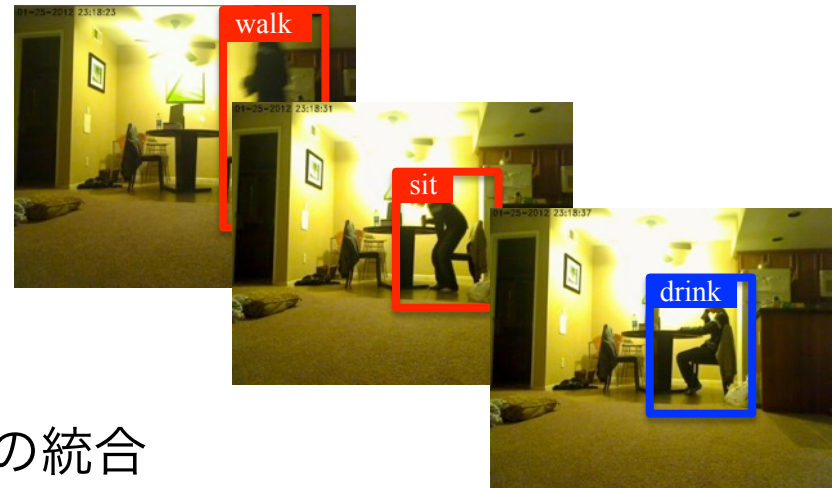
検出



行動認識



行動予測



- － コンピュータビジョンとデータ解析の統合

- 潜在的な知識の導入

- － 検出 – 行動認識 – 行動予測 までを一貫して研究

- 相補的に処理をインテリジェント化

 入力  予測対象

片岡裕雄の研究歴 – 2008 ~ 2015

ヒューマンセンシングを中心としたコンピュータビジョン技術の提案

人物行動解析技術

サッカー映像解析(2008 – 2011)

- 混雑状況下における複数選手追跡
- リアルタイム(30fps)で追跡・俯瞰画像位置特定
- 論文: ACCVW'09, 電気学会論文誌'10, 画像電子学会論文誌'12
- 受賞: **DIA'10若手奨励賞**, **SICE'10論文賞**

歩行者検出・追跡(2009 – 2011)

- 共起特徴量の提案
- Tracking-by-detectionのフレームワークにてリアルタイム(15fps)追跡
- 論文: IJVS'12, IEICE Trans. '14
- 受賞: **電気学会論文誌賞'12**, **IEEE IECON'12 Award**

大規模動線解析(2012 – 2013)

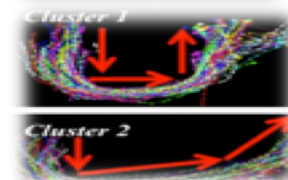
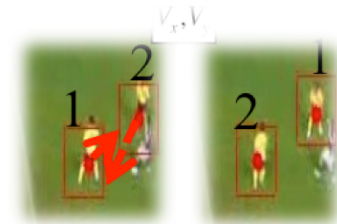
- 環境認識による教師無しクラスタリング
- 約30万人の動線を解析
- 論文: ViEW'12, IAPR MVA'13

人物行動認識のための改良型特徴(2011 – 現在)

- Dense Trajectories + 共起特徴
- Topic Model – LDAによる特徴量抽出
- Convolutional neural networks (CNN)による特徴抽出
- 論文: MIRU'12, SSII'13, **ACCV'14**, and in submission

データ解析を統合した行動予測(2012 – 現在)

- 行動履歴データベースの解析
- ベイジアンフレームワークによる行動予測
- 論文: MIRU'12, SSII'13, and in submission



共同研究

国内外の大学・国立
研究所・企業

【国立研究所】

- 産総研
- 交通安全環境研究所

【企業】

- パナソニック
- デンソーITラボラトリ
- 沖電気
- トヨタ自動車
- Siemens AG

【University】

- ミュンヘン工科大学
- カリフォルニア大学
- 東京大学
- 慶應義塾大学
- 東京電機大学

SIT

'08

Keio Univ. '09 - '14

Univ. of Tokyo '14 - '15

AIST '15 -

UCR '11 - '12

TUM '13 -

修士課程まで (2008 – 2011)

人物を検出・追跡する

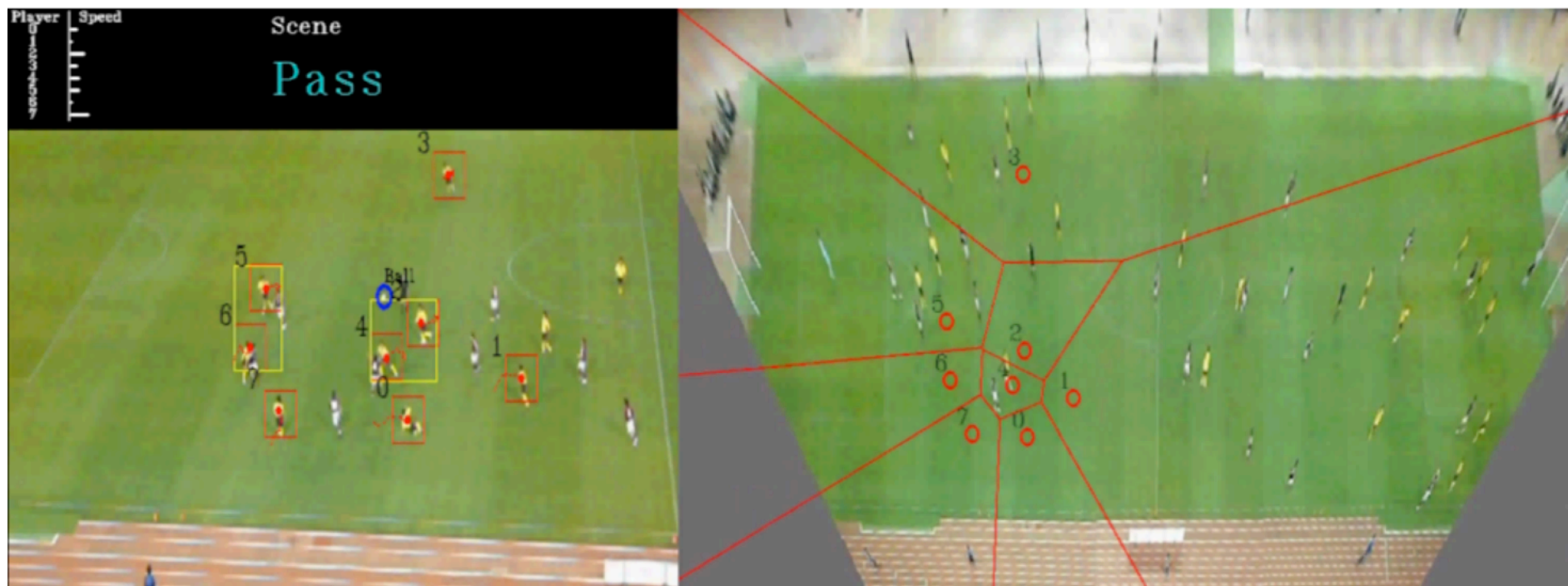
- 所属：慶應義塾大学青木義満研究室 <http://www.aoki-medialab.org/>
- 興味：解析のための検出・追跡, 人物認識の基礎技術の習得
- テーマ 1 : スポーツ映像処理 => 戦術解析
 - 卒業論文, 修士論文
- テーマ 2 : 歩行者検出・追跡 => 自動車の予防安全
 - 結果的には博士論文に

サッカー映像解析 (2008 – 2011)

戦術解析のための複数選手追跡

DIA'10若手奨励賞, SICE'10論文賞

- Color-based Particle Filterによる個人選手追跡
- 混雑状況時には識別器により選出を検出して重心を再配置
- 選手の相対的な位置関係や前のフレームの速度を考慮



リアルタイム(20fps)で複数選手追跡とグローバル位置推定

歩行者検出・追跡 (2009 – 2011)

共起特徴による歩行者検出と歩行者追跡

電気学会論文誌賞'12, IEEE IECON'12 Award, IJVS'12

- 共起特徴 + 識別器による歩行者検出
- Tracking-by-detectionに車両運動モデルや識別器を適用



複雑環境下でも検出・追跡できるフレームワーク

博士課程 (2011 – 2014)

人物の行動を認識する

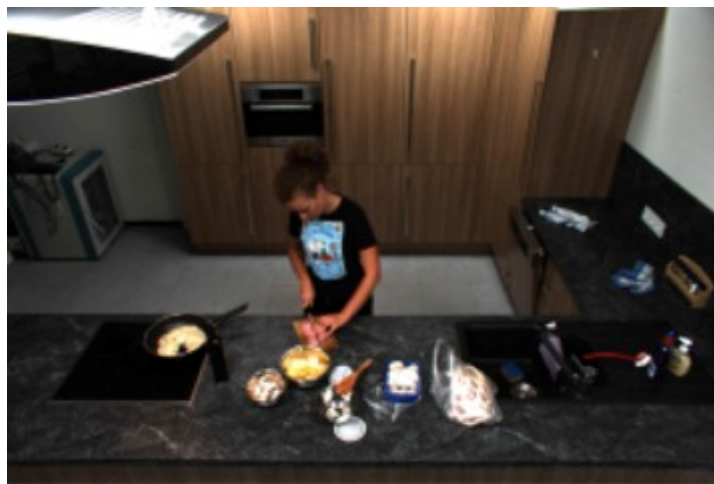
- 所属：慶應義塾大学青木義満研究室 <http://www.aoki-medialab.org/>
- 在外 1：カリフォルニア大学リバーサイド校 (2011 – 2012) <http://www.cris.ucr.edu/index.php>
 - Supervisor: Prof. Bir Bhanu
 - 紹介：斎藤英雄先生
- 在外 2：産業技術総合研究所 (2012 - 2013) https://unit.aist.go.jp/is/ci/group/cvrg_j.html
 - Supervisor: 佐藤雄隆博士， 岩田健司博士， 大西正輝博士
- 在外 3：ミュンヘン工科大学(2013, 2014) <http://campar.in.tum.de/Main/SlobodanIlic>
 - Supervisor: Prof. Nassir Navab, Dr. Slobodan Ilic
- 興味：より詳細な行動の把握， 世界水準の研究へ(博士課程3年うち2.5年は学外)
- テーマ：詳細行動認識 => 日常行動， 手術映像の解析， 歩行者動作

詳細行動認識

詳細行動認識(Fine-grained Activity Recognition)

- 視覚的に非常に類似しているが、意味が異なる対象の識別
- 行動認識の場合には時系列の差分を見極める必要がある

MPII cooking activity dataset^[Rohrbach, CVPR2012]



(a)



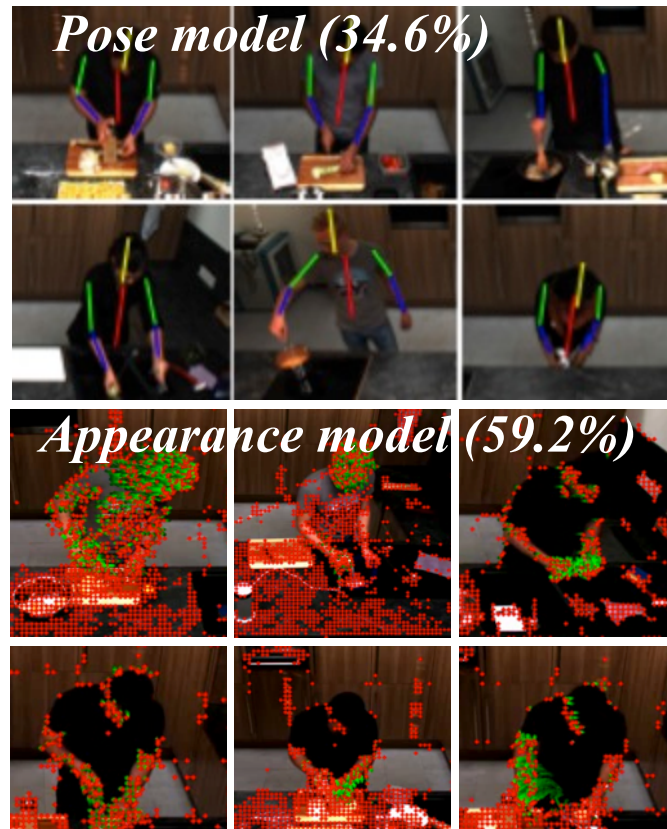
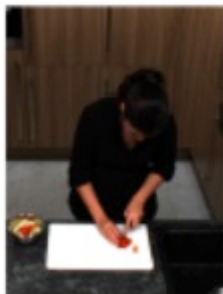
(e)



(f)



(g)



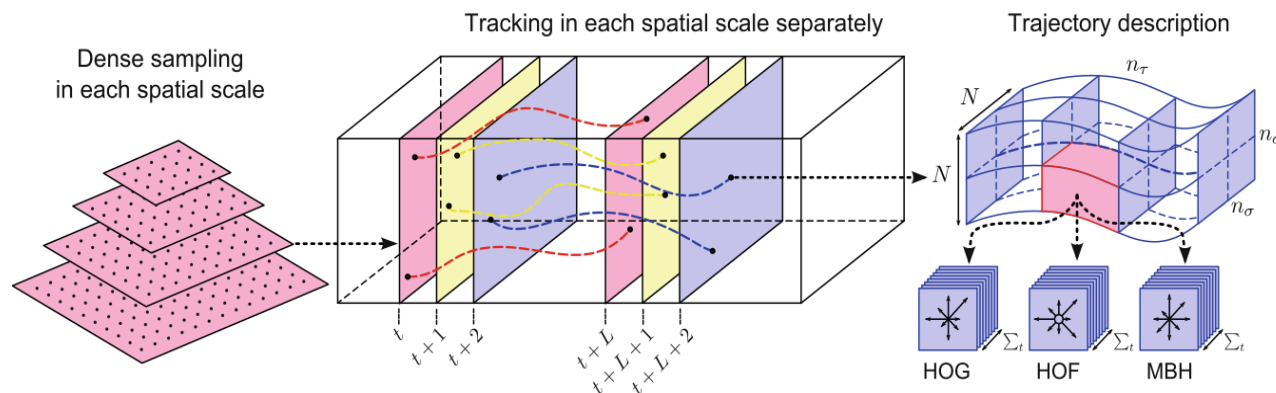
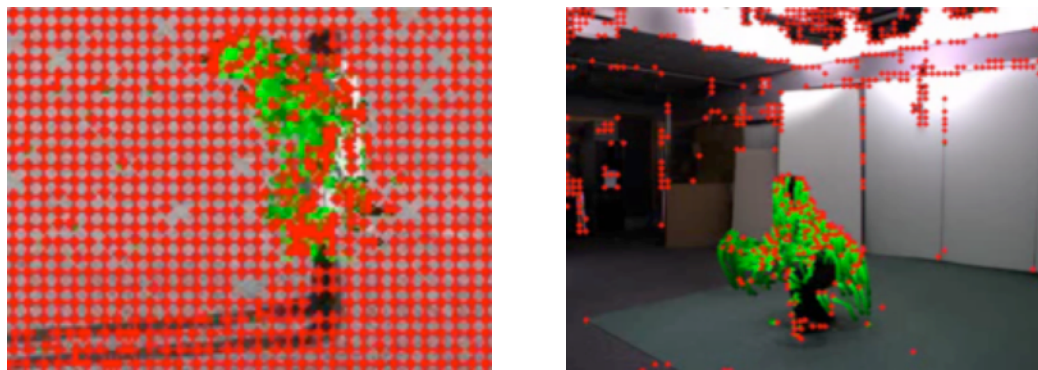
Pose model (34.6%)

Appearance model (59.2%)

Appearance model (59.2%) > Pose model (34.6%)

Dense Trajectories (DT) [Wang, IJCV2013]

時系列的にも空間的にも密なサンプリング



- スケール変化への対応：ピラミッド構造でフロー取得
- 特徴取得：HOG, HOF, MBH, Traj.
- 特徴表現：Bag-of-words (BoW)

H. Wang+, “Dense Trajectories and Motion Boundary Descriptors for Action Recognition”, in IJCV2013.

H. Kataoka, “【チュートリアル】 動的な人物・物体認識技術 –Dense Trajectories–”, in Slideshare.

<http://www.slideshare.net/HirokatsuKataoka/dt-tutorial>

共起特徴量を用いた改良

DTの特徴取得フレームワークに共起特徴を挿入

H. Kataoka+, “Extended Co-occurrence HOG with Dense Trajectories for Fine-grained Activity Recognition”, in ACCV2014.

– 共起特徴量

- CoHOG[Watanabe, PSIVT2009] (ペアのカウント), ECoHOG (エッジ強度の累積)
- PCAによる次元圧縮

– 特徴改善の効果

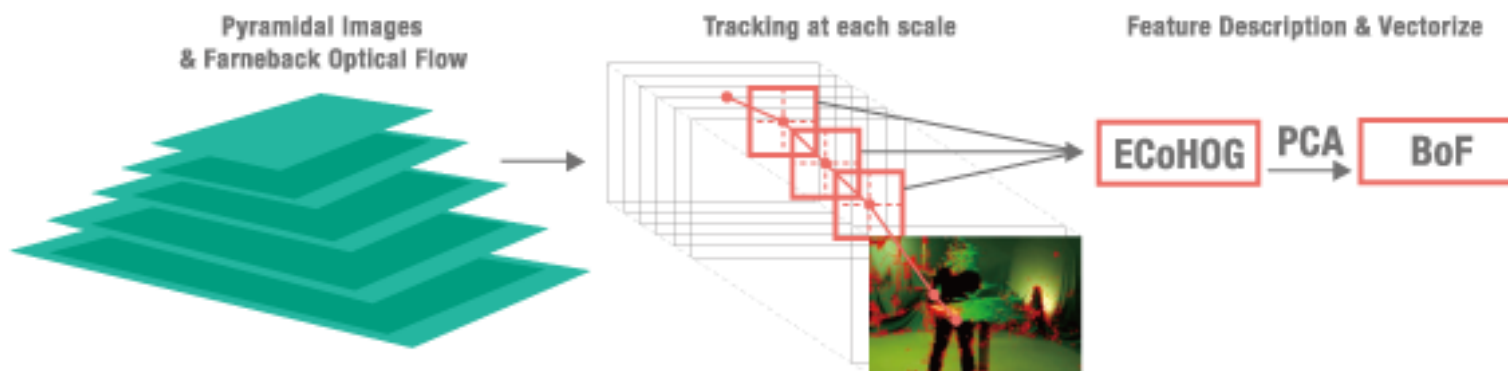
- 密なフロー抽出 + さらに詳細な特徴抽出
- PCAによる効率的な特徴空間生成
- 共起特徴統合 (62.4%) > DT (59.2%)

CoHOG

$$C_{x,y}(i,j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1, & \text{if } d(p,q) = i \text{ and } d(x+p, y+q) = j \\ 0 & \text{otherwise} \end{cases}$$

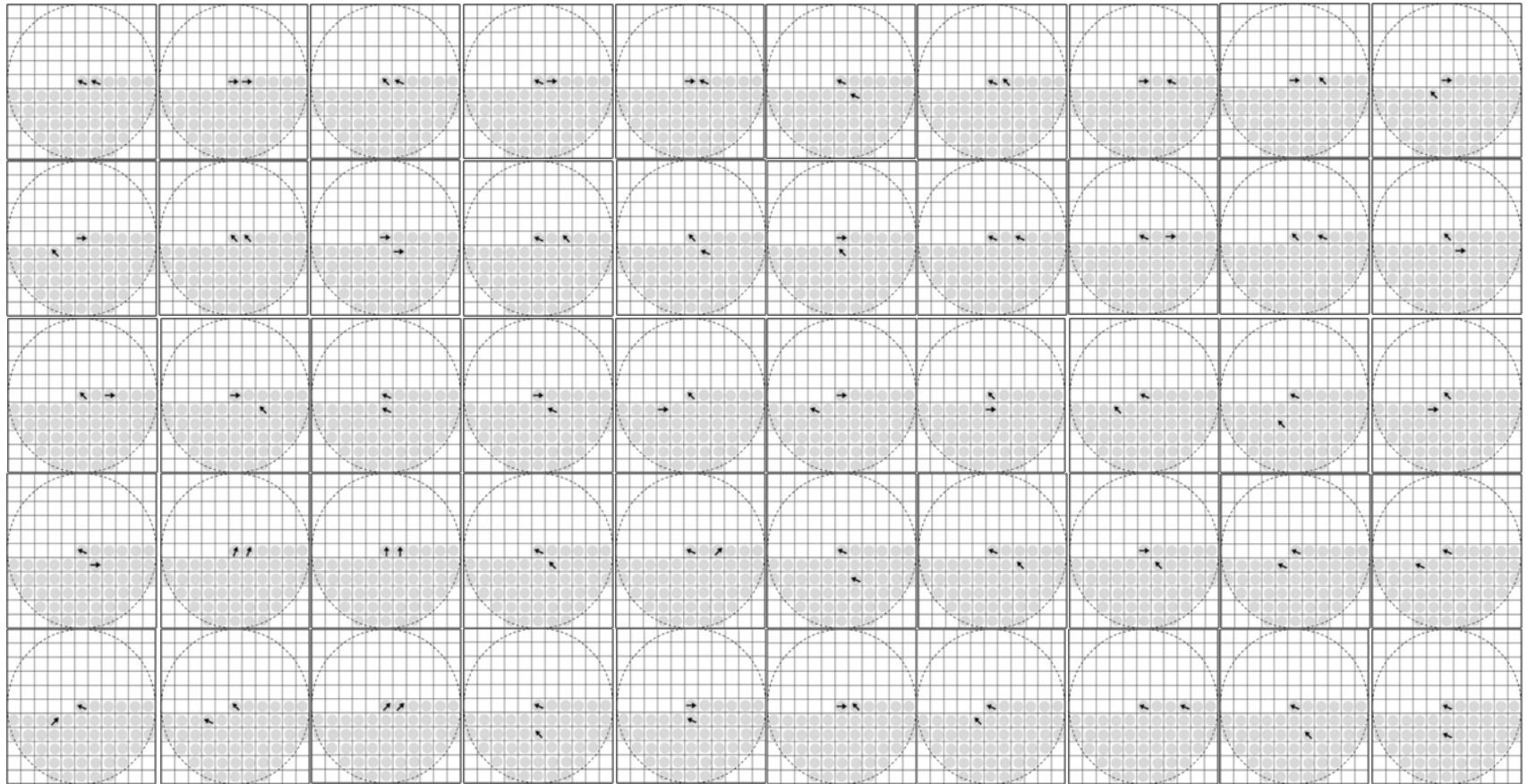
ECoHOG

$$C_{x,y}(i,j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} \|g_1(p,q)\| + \|g_2(p+x, q+y)\| \\ \text{if } d(p,q) = i \text{ and } d(p+x, q+y) = j \\ 0 & \text{otherwise} \end{cases}$$



上位50の頻出特徴量

- 隣接する領域から特徴を抽出しているのがわかる



詳細行動には「密なサンプリング + 詳細な特徴記述」が効果的

学振特別研究員(PD) (2014 – 2015)

人物行動を事前に予測する

- 所属：東京大学下坂正倫研究室 <http://ics.t.u-tokyo.ac.jp/ja/>
 - ウェアラブルセンサや車載センサ等からデータ取得
 - 機械学習によるデータ解析
- 興味：データ解析との統合，画像に映らない情報を推測
- テーマ：特徴マイニング，人物行動予測

BoWの特徴ベクトル

実際に取得されたBoWの特徴ベクトル (DT + BoW)

- ほとんどの値がゼロで構成されている
- unnecessary 部分が多数存在

Bin No. 										
	B1	B2	...						B10	
vec1	0.00	0.00	0.00	0.00	0.00	0.41	0.00	0.00	0.10	0.00
vec2	0.00	0.00	0.00	0.00	0.00	0.41	0.00	0.00	0.20	0.00
vec3	0.00	0.00	0.00	0.00	0.00	0.71	0.00	0.00	0.10	0.00
vec4	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00	0.30	0.00
vec5	0.00	0.00	0.00	0.10	0.00	0.80	0.00	0.00	0.10	0.00
	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.

非ゼロの値は同一のベクトル位置に格納される

→ 冗長部の排除が「精度の向上」や「容量の削減」に繋がる？

H. Wang+, “Dense Trajectories and Motion Boundary Descriptors for Action Recognition”, in IJCV2013.

H. Kataoka, “【チュートリアル】 動的な人物・物体認識技術 –Dense Trajectories–”, in Slideshare.

<http://www.slideshare.net/HirokatsuKataoka/dt-tutorial>

頻出の部分集合を抽出

- Support(支持度)とConfidence(信頼度)という指標を使用

BoWの際の設定

n: BoWのベクトル数 (n=5)

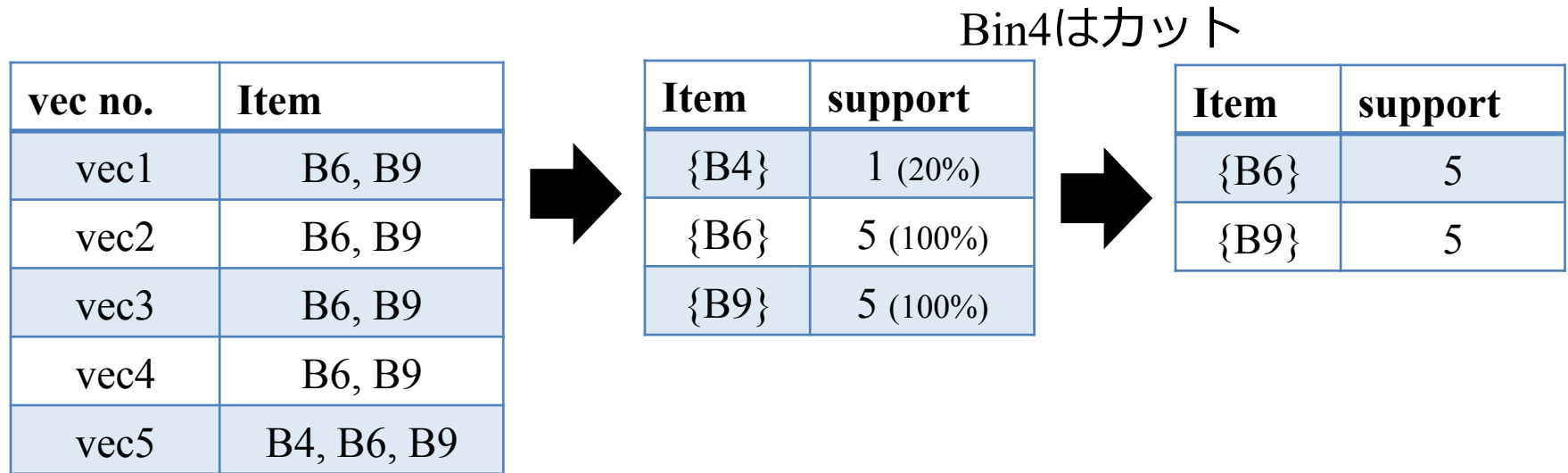
X, Y: BoWの非ゼロのベクトル位置(B1, B2 ... B10)の集合

(X ∪ Y): XとYをどちらも含む特徴ベクトル (e.g. “(B4 B6) ∪ B9”)

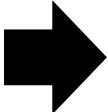
$$support = \frac{(X \cup Y).count}{n}$$

$$confidence = \frac{(X \cup Y).count}{X.count}$$

アソシエーション分析の流れ



組合せを増やしてsupportを計算



Item	support
{B6, B9}	5 (100%)

(“B6”, “B9”)
が頻出要素として抽出

Dense Trajectoriesの部分集合

密な空間の頻度による分解

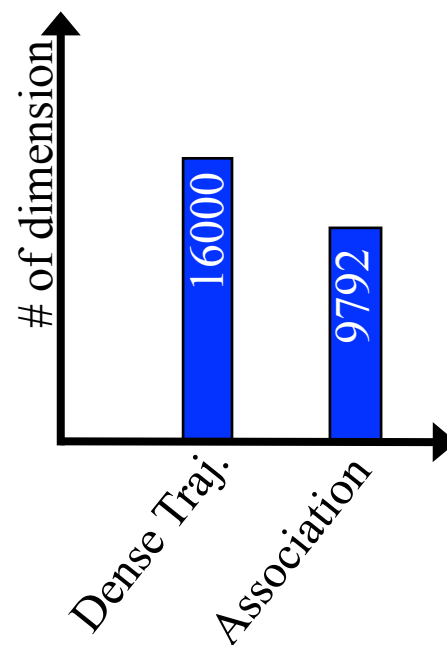
- 頻度の低いフロー(e.g. 背景, 稀な特徴)は除去される
- 共起性も見ているので意味のある空間(e.g. 手振り)に落としこみやすい

密な空間から最小動作単位に分離
(適応的な組み合わせが効果的)

Dense Trajectories



背景領域は除去されている

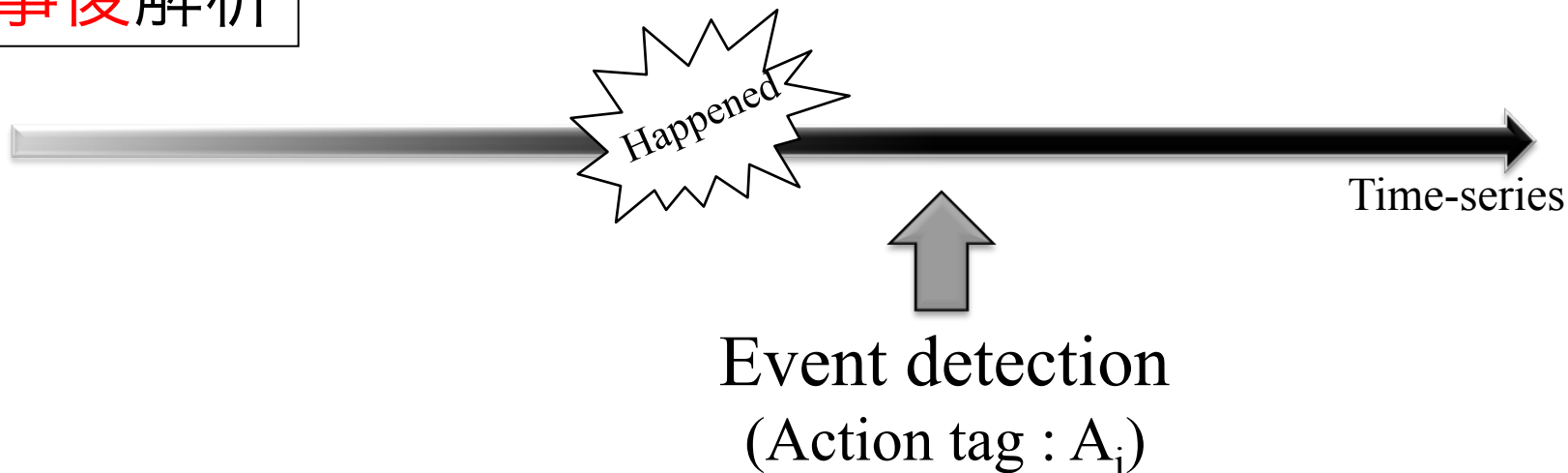


次元を削減しつつも精度を向上(特徴マイニングの重要性)

現在までのアプローチ

しかし、今までのアプローチはほとんどが事後解析

事後解析

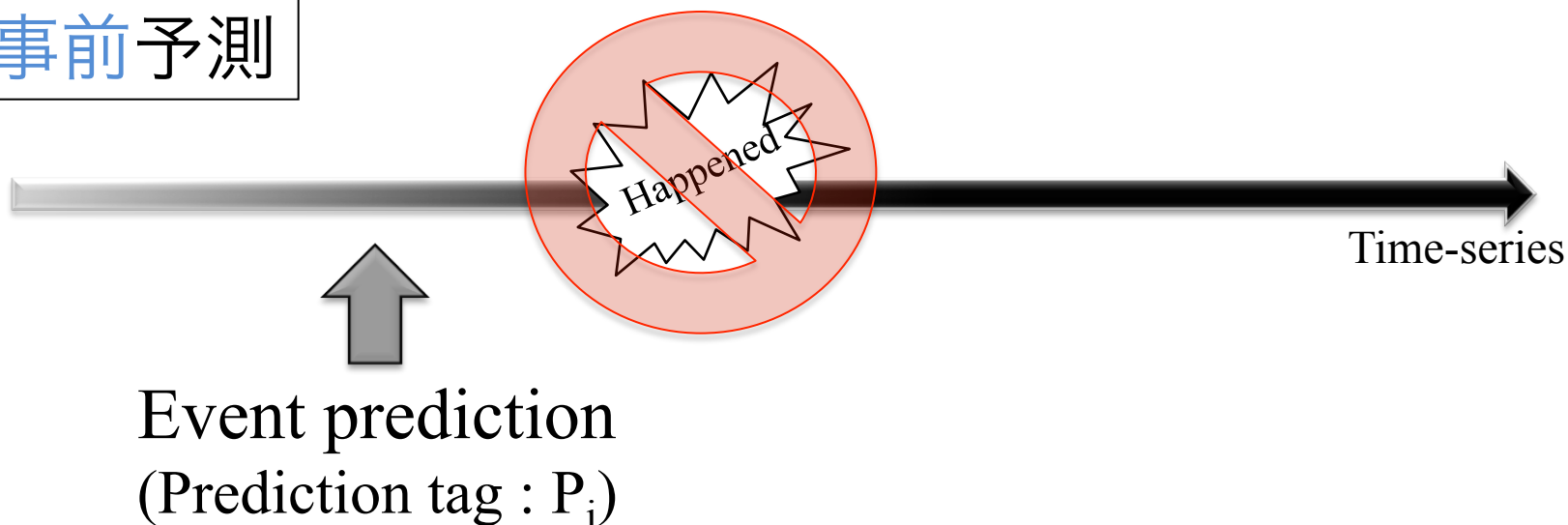


It's too late!!

これからのアプローチ

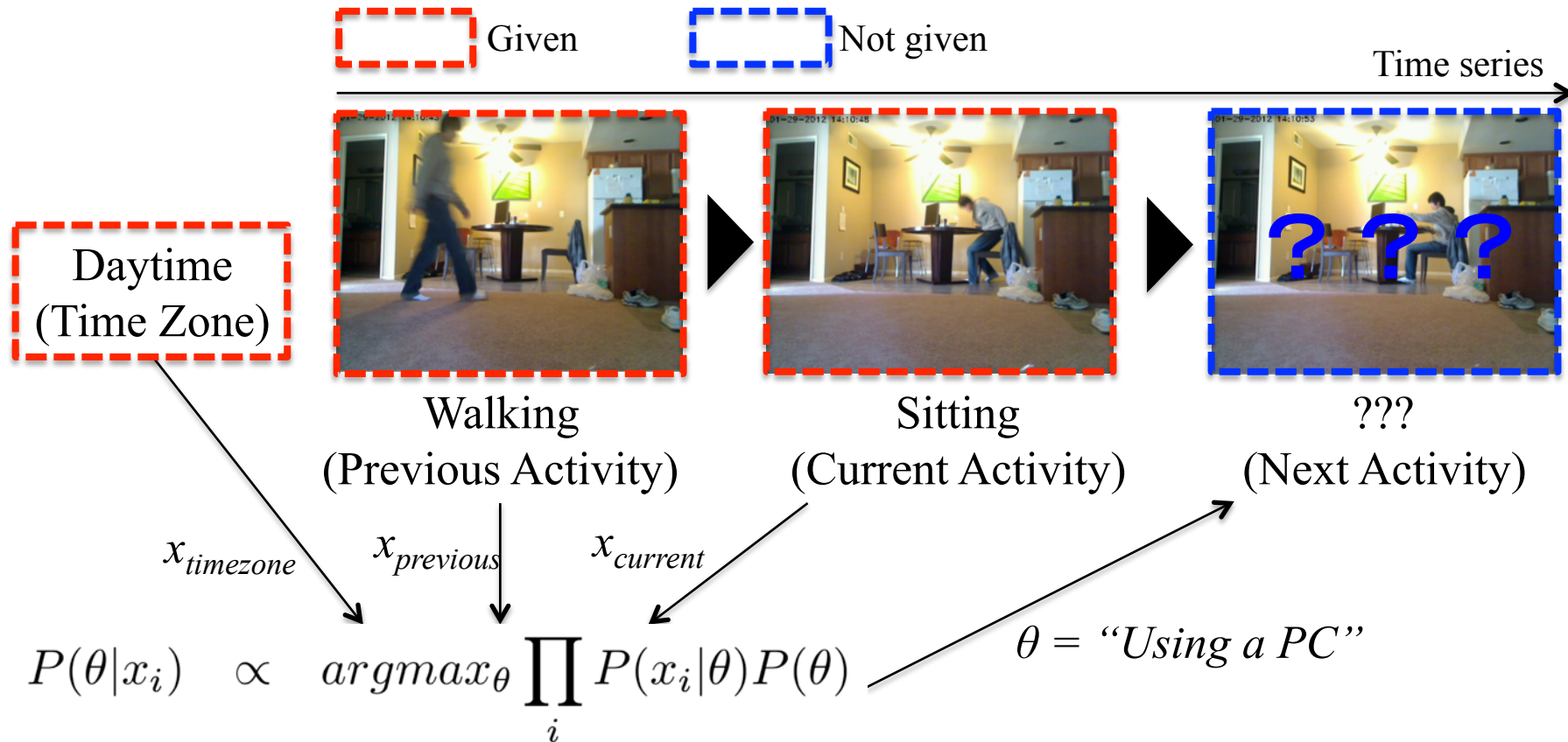
一方、事前に情報をしないと阻止できない場面がある

事前予測



Activity prediction and it's prevented

行動予測の課題



- Action sequence

“Walk” => “Sit” の段階で “Using a PC”を予測

- Time zone (補助的な要素)

Day time

行動予測のアプローチ

時系列行動タグ + 要素として時間帯を入力へ

- 時間帯 (Time Zone)
 - “morning”, “day time”, “night”
- 前行動, 現行動
 - “walk”, “bend”, “stand”, “sit”...
- 次行動
 - “use a PC”, “read”, “meal”...

解析手法

- Naïve Bayes

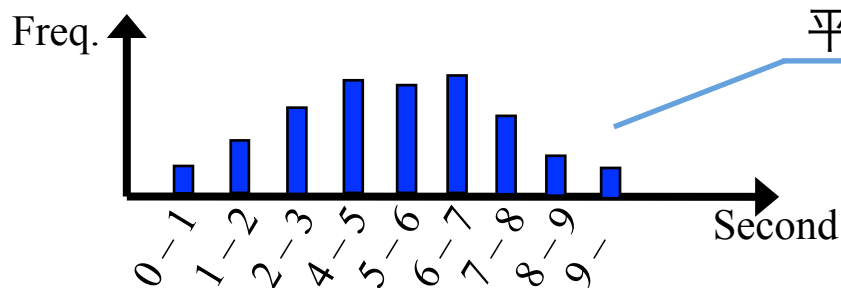


行動履歴DB

1	morning	walk	sit	book
2	morning	walk	stand	book
3	day	walk	sit	pc
4	night	bend	sit	meal
5	night	walk	sit	pc
6	night	walk	sit	pc
7	night	bend	sit	meal
8	night	walk	stand	book
9	night	walk	sit	pc
10	morning	walk	sit	book
11	morning	bend	sit	meal
12	morning	walk	sit	book
13	day	walk	sit	pc
14	day	walk	stand	pc
15	day	walk	sit	pc

結果

Time	Action	Attributes
01-29-2012 14:10:48	walk	Time Zone -> day Current Action -> walk Previous Action -> Intention -> The person will use a pc. Estimated Intention
01-29-2012 14:10:55	sit	Time Zone -> day Current Action -> sit Previous Action -> walk Intention -> PC (0.82) Read (0.11) Predicted activity
	walk	Time Zone -> 2 Current Action -> walk Previous Action -> Intention -> The person will read a book.
	read	Time Zone -> 2 Current Action -> read Previous Action -> take Intention -> Read (1.00) PC (0.00)



平均約5秒，最大で約20秒の予測

背景に潜む知識の使用により81.0%の精度，最大約20秒先を予測

産総研 特別研究員 (2015 – 20xx)

より高度な人物の行動認識・解析・予測へ

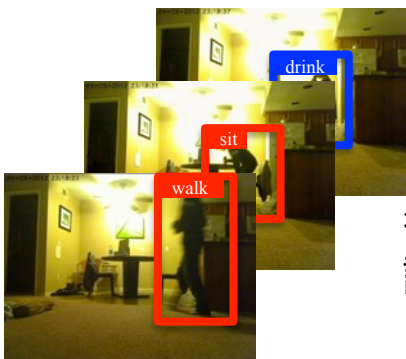
- 所属：産業技術総合研究所 コンピュータビジョン研究グループ
- 興味：人物行動解析技術の高いレベルでの拡張
- テーマ 1：行動認識とデータ解析による高度な人物行動解析
 - データベースや解析技術による知識側の強化
 - 予測だけに留まらないコンピュータビジョンの拡張
- テーマ 2：3D VisionやDeep CNNによる情報量の追加や高精度化
 - 物体情報(3D Vision)の統合による行動領域の拡張
 - Deep Learningによる詳細解析や特徴バリエーションの増加

観測と知識

“観測” が “知識” を強くする

“知識” が “観測” を強くする

行動認識(観測)



行動
認識

観測

歩行 – 「物を落とす」 – 「片付け」 – 座る – 食事

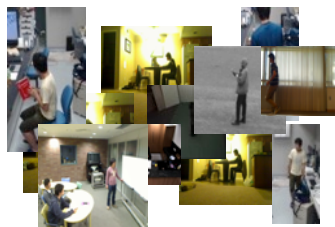
観測から知識強化

モデルと何か違う？

比較

予測, 意図
相違, 共起...

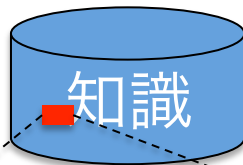
データ解析(知識)



Large scale videos

蓄積

Activity DB



モデル

s08-d02-cam-002	120	303	56
s08-d02-cam-002	358	438	56
s08-d02-cam-002	441	559	21
s08-d02-cam-002	561	629	16
s08-d02-cam-002	644	860	55

知識モデル

モデリング 1:

直立 – ... – 調理

モデリング 2:

歩行 – 座る – 食事

...

モデリング N:

...

今後

コラボレーションによる研究の深化と応用へ

- 多くの共同研究により知見や技術を拡張させるに至っている
- 企業だけに限らず，大学間の連携も行っている

「観測と知識の相補的な深化」を今後も追求