

# 新しい概念の発見を目指して

マルチメディアデータから新しい概念を発見する高階モデル学習  
東京工業大学 井上 中順

コンピュータによる  
**認識・理解**

Deep Learning  
大規模学習

コンピュータによる  
**発見・創出**

抽象度の高い学習  
異種データ統合

情報社会の AI  
イミテイティブ  
問題の単調的な解決

未来の実現に  
必要な挑戦

未来社会の AI  
クリエイティブ  
より人間らしい解決

## 未知のもの・ことを「学習」

1年4ヶ月の成果: コンピュータにとって見たことのないものの学習

### Zero-Shot Learning

学習とテストで検出対象が異なる学習

例) 外部データから類似度を計算

$$\hat{f}_{\text{airplane}}(x) = \alpha(\text{airplane, bird}) f_{\text{bird}}(x) + \alpha(\text{airplane, car}) f_{\text{car}}(x)$$

↑ 入力映像      ↑ 深層学習から得られるスコア  
学習: {bird, car}    テスト: {airplane}

### Visual Concept Vectors

物体間の視覚的類似度を測るベクトル

例) 重み係数を画像特徴量で最適化

$$\psi(\text{swallow}) = \pi_3 v(\text{swallow}) + \pi_2 v(\text{bird}) + \pi_1 v(\text{animal})$$

↑ テキストデータから得られた Word Vectors

学習結果の例 共起関係を捉えているものが多い

Musician:  +  +  + ...

Stadium:  +  +  + ...

Baby:  +  +  + ...

Running:  +  +  + ...

TRECVID 映像データセットでの評価

Representation	MeanAP(%)
Word Vecotors	8.26
Visual Concept Vectors	10.07

\* TRECVID: 映像解析の国際ワークショップ  
30種類の物体・動作・シーンを Zero-Shot 学習  
学習: GoogLeNet on ImageNet, Skipgram on News