

機関番号：12612

研究種目：若手研究（A）

研究期間：2008～2010

課題番号：20680010

研究課題名（和文） 1000クラスに対応した大規模一般画像認識システムの実現

研究課題名（英文） A Study on Large-Scale Generic Object Recognition System for 1000-class Categorization

研究代表者

柳井 啓司 (YANAI KEIJI)

電気通信大学・大学院情報理工学研究科・准教授

研究者番号：20301179

研究成果の概要（和文）：

本研究では多数のクラスに対応した大規模一般画像認識システムの実現およびそれに関連する研究を行い、以下の成果を得た。

- (1) 1000種類の画像DBの構築,
- (2) 200万枚のWeb画像を用いた画像オントロジーの構築,
- (3) 100種類の食事画像に対応した食事画像認識システムの構築,
- (4) 100種類の動作に対応した動作ショットデータベースの自動構築,
- (5) 視覚概念と位置の関係の分析.

研究成果の概要（英文）：

In this research, we studied on large-scale generic object recognition system and related topics:

- (1) Construction of a 1000-category image database
- (2) Construction of an image ontology database with two million Web images
- (3) Food image recognition system for 100-category foods
- (4) Automatic construction of 100-kind action shot database
- (5) Analysis of relation between visual concepts and locations

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	6,600,000	1,980,000	8,580,000
2009年度	4,500,000	1,350,000	5,850,000
2010年度	7,600,000	2,280,000	9,880,000
年度			
年度			
総計	18,700,000	5,610,000	24,310,000

研究分野：知覚情報処理，メディア情報学

科研費の分科・細目：知覚情報処理・知能ロボティクス

キーワード：一般物体認識，画像認識，Webマルチメディアマイニング

## 1. 研究開始当初の背景

- (1) 研究開始当時の2008年までは、一般物体認識研究は20種類や多くても100種類程度の認識クラスを対象に行われていた。
- (2) 当時は研究レベルで一般物体認識研究

が行われているのみで、実应用到一般物体認識技術が使われたことはなかった。

## 2. 研究の目的

- (1) Web から自動的に一般物体認識のための学習画像を収集する技術 Web 画像マ

マイニングを用いて、1000種類規模の一般物体認識のためのデータセットを構築し、1000種類規模の多数のクラスに対応した実用を目指した物体認識システムを構築する。また、どのような種類のクラスを認識対象カテゴリとして準備すべきかについて Web 上の大量の画像を用いて視覚概念の分析を行う。

- (2) 実験室内だけではなく実応用を想定した一般物体認識技術及びその応用システムを開発する。



### 3. 研究の方法

実際には以下の5つの方法によって研究を実施し、それぞれ成果を得た。

- (1) これまで我々が研究を続けていた、画像認識技術を用いて Web から指定したキーワードに関連する画像を自動的に大量に収集する Web 画像マイニング技術を用いて、1000種類以上のカテゴリの画像を自動収集する。
- (2) 200万枚のキーワード付与された画像を画像共有サイトである Flickr から収集し、視覚概念同士の類似性の分析を行う。
- (3) 実用的な一般物体認識の応用例として、自動的な食事記録の実現を目指した食事画像認識システムを開発する。Web 画像マイニング技術によって、Web 上から 100種類の食事画像を収集し食事画像 DB を構築する。
- (4) 静止画像の認識のみでなく、Web 画像マイニング技術を Web 動画にも拡張し、Web 動画共有サイトからキーワード付きの Web 動画を収集し、指定した動詞キーワードに対応する Web 動画中のショットを自動抽出し、動詞に対応する動画ショット DB を自動構築する。
- (5) Web 上には撮影位置の情報のついた位置情報付き画像が大量に存在している。そうした画像を用いて、視覚概念と位置の関係について分析を行う。

### 4. 研究成果

- (1) 1000種類の画像 DB の構築を行った。ただし、Web 上の Crowdsourcing 技術を用いて自動的に数百万種類の大規模画像 DB を自動構築する研究 ImageNet が海外で発表されたため、本研究での 1000種類 DB は新規性がなく

なってしまったため、外部発表は行っていない。

- (2) 200万枚の Web 画像を用いた画像オントロジーの構築を行った。視覚概念の類似性を視覚特徴、タグ共起特徴、その混合特徴の3種類によって表現し、グラフ表現することを行った。画像オントロジーを用いて検索語拡張を行うことによって、キーワードによる画像 DB 検索の精度が向上することが示された。
- (3) 100種類の食事画像に対応した食事画像認識システムの構築を行い、実際に携帯電話から写真を送ると食事名とおおよそのカロリーを返すシステムを実現した。また、食事画像認識 Twitter Bot として一般公開している。
- (4) Web 上の動画共有サイトからキーワード付き動画を収集することによって 100種類の動作に対応した動作ショットデータベースの自動構築を行った。100種類もの動作ショット動画 DB を自動構築した例は今までなく、国際的にも評価を受けた。
- (5) Web 上の位置情報付き画像を用いて、250種類の視覚概念について画像 DB を構築し、各視覚概念について位置の関係の分析を行った。その結果、位置によって視覚特徴が変化する概念と、変化しない概念、特定の地域のみ集中している概念など、概念によってそれぞれ位置に関する異なる特性を持っていることを発見した。この研究は初期段階であり、続きは現在、基盤研究(B)によって継続して研究を実施している。

### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計4件)

- ① 秋間雄太, 川久保秀敏, 柳井啓司: Folksonomy を用いた画像特徴とタグ共起に基づく画像オントロジーの自動構築, 電子情報通信学会論文誌 D, Vol. J94-D, No.8, pp.1248-1259 (2011).
- ② 川久保秀敏, 柳井啓司: 単語概念の視覚性と地理的分布の関係性の分析, 電子情報通信学会論文誌 D, Vol.J93-D, No.8, pp.1417-1428 (2010).
- ③ 上東太一, 柳井啓司: Multiple Kernel Learning による 50 種類の食事画像の認識, 電子情報通信学会論文誌 D, Vol. J93-D, No.8, pp.1397-1406 (2010).

- ④ Akio Kitahara and Keiji Yanai: Associating faces and names in Japanese photo news articles, Progress in Informatics, No. 7, pp.73-70 (2010).

[学会発表] (計 22 件)

- ① Kohya Okuyama and Keiji Yanai: A Travel Planning System Based on Travel Trajectories Extracted from a Large Number of Geotagged Photos on the Web, Proc. of Pacific-Rim Conference on Multimedia, Sydney, Australia (2011/12).
- ② Do Hang Nga and Keiji Yanai: Automatic Construction of an Action Video Shot Database using Web Videos, Proc. of IEEE International Conference on Computer Vision (ICCV), Barcelona, Spain (2011/11).
- ③ Hidetoshi Kawakubo and Keiji Yanai: GeoVisualRank: A Ranking Method of Geotagged Images Considering Visual Similarity and Geo-location Proximity, Proc. of the Nineteenth International World Wide Web Conference (WWW), Poster Paper, Hyderabad, India (2011/03).
- ④ Keita Yaegashi and Keiji Yanai: Geotagged Image Recognition by Combining Three Different Kinds of Geolocation Features, Proc. of Asian Conference on Computer Vision (ACCV), Queenstown, New Zealand (2010/11).
- ⑤ Akitsugu Noguchi and Keiji Yanai: A SURF-based Spatio-Temporal Feature for Feature-fusion-based Action Recognition, Proc. of ECCV WS on Human Motion: Understanding, Modeling, Capture and Animation, Crete, Greece (2010/09).
- ⑥ Keita Yaegashi and Keiji Yanai: Geotagged Photo Recognition using Corresponding Aerial Photos with Multiple Kernel Learning, Proc. of IAPR International Conference on Pattern Recognition (ICPR), Istanbul, Turkey (2010/08).
- ⑦ Keiji Yanai and Kobus Barnard: Region-based Automatic Web Image Selection, Proc. of ACM International Conference on Multimedia Information Retrieval (MIR), Philadelphia, USA (2010/03).
- ⑧ Akio Kitahara and Keiji Yanai: Detecting "In-play" Photos from Web Sports News Photos, Proc. of the Pacific-Rim Conference on Multimedia,

Bangkok, Thailand (2009/12).

- ⑨ Taichi Joutou and Keiji Yanai: A Food Image Recognition System with Multiple Kernel Learning, Proc. of International Conference on Image Processing (ICIP), Cairo, Egypt (2009/11).
- ⑩ Akitsugu Noguchi and Keiji Yanai: Extracting Spatio-temporal Local Features Considering Consecutiveness of Motions, Proc. of Asian Conference on Computer Vision (ACCV), Xian, China (2009/09).
- ⑪ Keiji Yanai, Hidetoshi Kawakubo and Qiu Bingyu: A Visual Analysis of the Relationship between Word Concepts and Geographical Locations, Proc. of ACM International Conference of Image and Video Retrieval (CIVR), Santorini, Greece (2009/07).
- ⑫ Keiji Yanai: Web Image Gathering with Region-based Bag-of-Features and Multiple Instance Learning, IEEE International Conference on Multimedia and Expo (ICME), New York, USA (2009/07).
- ⑬ Keita Yaegashi and Keiji Yanai: Can Geotags Help Image Recognition?, Proc. of Pacific-Rim Symposium on Image and Video Technology, Tokyo, Japan (2009/01).
- ⑭ Bingyu Qiu and Keiji Yanai: Objects over the World, Proc. of the Pacific-Rim Conference on Multimedia, Tainan, Taiwan (2008/12).
- ⑮ Akitsugu Noguchi and Keiji yanai: Rushes Summarization Based on Color, Motion and Face, Proc. of ACM MM Workshop on TRECVID Rush Summarization, Vancouver, Canada (2008/10).

[図書] (計 2 件)

- ① Keiji Yanai, Hidetoshi Kawakubo and Kobus Barnard: Entropy-based Analysis of Visual Concepts, In "Multimedia Information Extraction", Wiley-IEEE Computer Society Press, (2012/05).
- ② Keiji Yanai and Bingyu Qiu: Mining Regional Representative Photos from a Consumer-Generated Geotagged Photo Database, In "Handbook of Social Network Technologies and Applications", Springer, (2010/10).

[産業財産権]

○出願状況（計2件）

名称：画像処理方法、その方法を実行するプログラム、記憶媒体、撮像機器、画像処理システム

発明者：柳井 啓司

権利者：電気通信大学

種類：特許

番号：特願2008-106546

出願年月日：2008年4月16日

国内外の別：国内

名称：画像ランキング方法、プログラム及び記憶媒体並びに画像表示システム

発明者：川久保 秀敏, 柳井 啓司

権利者：電気通信大学

種類：特許

番号：特願2010-109454

出願年月日：2010年5月11日

国内外の別：国内

○取得状況（計0件）

〔その他〕

ホームページ等

<http://mm.cs.uec.ac.jp/>

[https://twitter.jp/foodimg\\_bot/](https://twitter.jp/foodimg_bot/)

6. 研究組織

(1) 研究代表者

柳井 啓司 (YANAI KEIJI)

電気通信大学・大学院情報理工学研究科・  
准教授

研究者番号：20301179

(2) 研究分担者

なし

(3) 連携研究者

なし