

平成 22 年 6 月 2 日現在

研究種目：基盤研究 (C)

研究期間：2007～2009

課題番号：19500175

研究課題名 (和文) 意味の曖昧さと多様さを考慮する大規模画像・  
ビデオの自動表現手法に関する研究研究課題名 (英文) A study on large-scale image/videos' automatic representation  
considering the ambiguity and diversity of semantics

研究代表者

戴 瑩 (Dai Ying)

岩手県立大学・ソフトウェア情報学部・准教授

研究者番号：60305290

研究成果の概要 (和文)：本研究は今までの関連研究で対応しなかった画像・ビデオの意味の曖昧さと人間のそれに対する解釈の多義さがあらかじめ定義したドメイン/カテゴリーに対する連想値で反映され、semantic tolerance relation model で表現することを検討した。さらに、その構築した semantic tolerance relation model をドメイン/カテゴリーに対する連想値の計算に導入し、画像/ビデオの意味に基づく自動注釈に活用した。semantic tolerance relation model の導入が画像/ビデオの意味による検索性能の改良に役に立ったことが分散ユーザによる検索結果により検証した。

研究成果の概要 (英文)：For embodying the ambiguity and diversity of image/videos' semantics, the semantic tolerance relation model (STRM) is established. On the basis of it, meanings of image/videos are represented by the associative values with defined domains/categories. Further, the method of calculating the associative values combined with the STRM is proposed, and an associative values-based image/videos' retrieval system toward subjects distributed users is constructed and implemented. The retrieval results made by some show the effectiveness and efficiency of the proposed method compared with those without conducting the STRM.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	1,600,000	480,000	2,080,000
2008年度	1,100,000	330,000	1,430,000
2009年度	800,000	240,000	1,040,000
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野：総合領域

科研費の分科・細目：情報学・感性情報学

キーワード：画像・ビデオ、意味データ構造、自動注釈、検索、あいまいと多様性、連想値、分散ユーザ、分散ネットワーク

## 1. 研究開始当初の背景

今の IT 社会において、機械情報を作成、伝達、蓄積する装置の異常なほどの発達により、我々は、機械情報の洪水に溺れてしまう一方で、価値が高いメディア・データは有効に利用されないままに放置されている。従って、情報洪水を防ぐとともに、コンテンツ資源を有効利用するために、メディア・データの有効な蓄積及び検索方法の研究はとても重要である。

メディア・データの検索技術について、今まで検索手法により次のような分類があった。「ターゲット検索」ではユーザが特定のターゲットにメディア・データを見つけようとする。「カテゴリー検索」ではユーザが一般的なカテゴリーから一つ以上のメディア・データを捜す。「open-ended browsing」では、ユーザが探しているものについてのただ漠然とした考えである。「感性検索」はユーザが入力した印象語などにより一つ以上のメディア・データを提供する。しかし、これらの方法では、画像/ビデオの特性である意味の曖昧さと人間のそれに対する解釈の多義さはまだうまく反映されていない。そして、大規模なメディア・データ検索のために、以下の問題がまだ解決されていない。

- (1) 画像/ビデオの意味の曖昧さに対応できるカテゴリーの体系的な定義及びそれに基づく画像/ビデオの自動登録；
- (2) ユーザの画像/ビデオに対する解釈の多義性とシステムの計算性能間のギャップの埋め；
- (3) ユーザからの質問に含む多様な認知基準の感知；
- (4) 大規模な自動登録と柔軟的な検索手法の実現；

本研究は主に自動注釈のための画像/ビデオからコンテンツ表現データベースを生成するとき生じる問題 (1) と問題 (2) の解決及び、利用者の意味に対する解釈基準の多様性による問題 (3) と問題 (4) の解決を目指す。

## 2. 研究の目的

(1) 画像/ビデオの特性である意味の曖昧さと人間のそれに対する解釈の多義さをうまく反映できるカテゴリーを系統的に定義する方法を検討し、カテゴリーの間には関連性が付けられる。さらに、その関連性に対する最適な計測手法を究明する。さらに、それに基づいて、各カテゴリー間の関係を表明できる semantic tolerance relation model を構築する。

(2) 人間の画像/ビデオに対する解釈の多義さにより、新たにドメインやカテゴリーを追加、削除する要求が生じると考える。そのために、ドメインやカテゴリーの追加、削除に伴う semantic tolerance relation model の再構築手法及びその手法の有効性ともたらず影響を究明する。さらに、画像/ビデオをうまく表現できるドメインやカテゴリーの数の限界を究明する。

(3) 画像/ビデオを意味に基づき自動注釈するための画像/ビデオの有効表現手法とその実現方法を検討する。それにより人間にとって自然的な検索方式を実現する。

(4) 分散ネットワークを用いたコンテンツ共有システムの導入を検討する。Prototype の運行により、上述の考案した方法は大規模画像/ビデオの自動登録及び柔軟検索に有効することを検証する。

### 3. 研究の方法

本研究の仕組みが Fig. 1 のようになっている。主に下記の研究を行う。

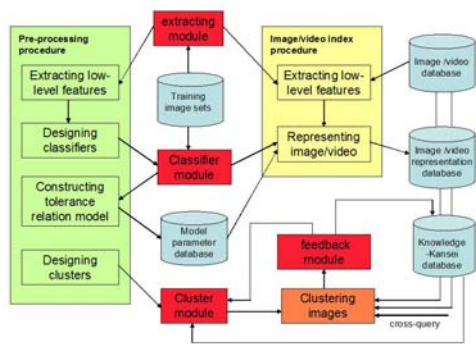


Fig. 1 研究の仕組み

(1) 前処理過程では同じドメインのカテゴリ間の intra-association と違うドメインのクラス間の inter-association の自動計測手法を検討すること、各クラス間の関係を表明できる semantic tolerance relation model の構築を行う。

(2) 画像/ビデオの意味に基づきの自動注釈のための有効表現手法とその実現方法を検討する。

(3) 人間にとって自然的な検索方式を確立する。つまり、異なるドメインのカテゴリに自由にアクセスでき、そして、多様な質問の組み合わせによりふさわしい画像を簡単に検索できる仕組みを確立する。

(4) 画像/ビデオの意味によるコンテンツ表現データベースのデータ構造を確立する。

(5) ユーザの多様質問及びそれに対する検索基準を感知できるも feedback シュールと、多様基準を反映できる知識-感性データベースの構成方式を検討する。

(6) 上述の一連の研究をまとめ、prototype システムの構築を検討する。分散ネットワークを用いたコンテンツ共有システムの導入も検討する。

(7) prototype の実装により、システムの性能を評価する。

### 4. 研究成果

(1) 本研究は画像・ビデオの意味の曖昧さと人間がそれに対する解釈の多義さをうまく反映できるように、画像・ビデオの意味はあらかじめ設定したドメイン/カテゴリとの連想値によるデータ構造で表現すると提案した。さらに、語彙の典型関係を参考する上で、各カテゴリ間関係を表明できる semantic tolerance relation model を構築し、その相互関係を数値化する方法を考案した。その概念図を Fig. 2 のように示した。

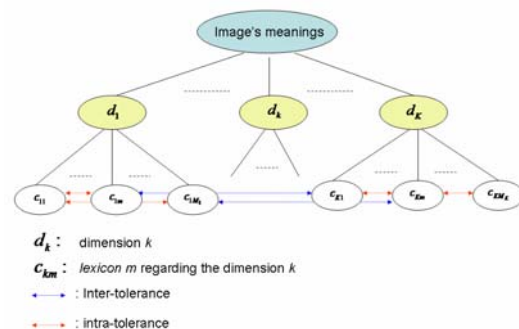


Fig. 2 モデルの概念図

(2) 人間の画像・ビデオに対する解釈の多義さにより、新たにドメインカテゴリを追加、削除する要求が生じると考える。そのために、ドメインやカテゴリの追加、削除に伴う分類器の再構築を効率的に実現できる手法を考察し、少量の学習サンプルでクラスを代表するモデルを構築する方法を検討し、有効性を検証した。

(3) ビデオをショットに分割し、そして、ショットを代表するキーフレームを抽出する。その上で、画像・キーフレームの意味を表現する有効手法として、あらかじめ設定したドメイン/カテゴリとの連想値により画像・キーフレームを自動的に登録する実現方法を検討した。従って、それらの連想値を獲得するために、カラーと形状特徴に基づく Bayesian 分類器や画素に基づく双方向連想

記憶を利用し、構築した semantic tolerance relation model により各カテゴリーに対する連想値を計算するアルゴリズムを開発した。さらに、それらのアルゴリズムが自然 vs. 人口ドメイン、人間 vs. 非人間ドメインと印象ドメインに関わる 40 個のカテゴリーの連想値の計算に実装した。表 1 にはその定義した 40 個のカテゴリーを示した。

表 1 定義した 40 個のカテゴリー

Domains	Lexicons
nature/man-made	ground, landscape, flower, tree, light, building, sunset, restraint, texture_brick, texture_wood, texture_paper, texture_cloth, food_matierals, food, mall, beach, snow_viewing, moutain_lake, flower_arrangement, interior, street, waterfall, mountain, pool, guard, ship, board, rocket, brriage, factory
face/non-face	portrait, face, non_face
impression	active, cool, calm, soft, hard, grayish, clear

実験結果は semantic tolerance relation model の導入がカテゴリーに対する連想値の計算精度の向上に役に立ったことを示した。他に、画素に基づく双方向連想記憶がカラーと形状特徴に基づく Bayesian 分類器より計算精度が高かつ実装しやすいということは実験結果から分かった。Fig. 3 には建物に対する検索精度を示した。+マーク青い曲線は semantic tolerance relation model を導入していない結果である。

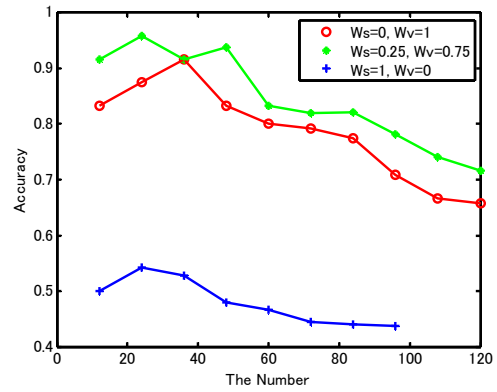


Fig. 3 building に対する検索精度

(4) 上述の成果に基づいて、分散ユーザによる分散環境をベースした大規模画像・ビデオの意味自動登録と検索システム構築の仕組みを検討し、その platform を開発した。さらに、上述データ構造で得た画像・ビデオの意味 metadata は画像の意味の曖昧さと解釈の多義さをうまく反映できることを 5 人の被験者の cross-domain/category query による画像・ビデオの検索結果により検証した。そのうち、5 人の被験者のターゲット画像における検出率（ターゲット画像の意味を反映する画像が検索されたトップ 24 の画像にあるのは検出という）の平均は 82% であり、その有効性を示した。Fig. 4 は 5 人の被験者による検出率を示した。

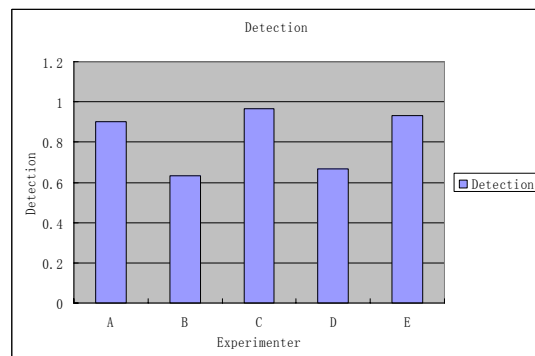


Fig. 4 被験者による検出率

## 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2件)

- ① Ying Dai, “Class-based image representation for Kansei retrieval considering semantic tolerance relation”, Journal of Japan society for fuzzy theory and intelligent informatics, Vol. 21, No. 2, 2009, pp. 184-193. 査読あり
- ② Ying Dai, “Semantic tolerance relation-based image representation and classification”, IJSEIA Journal Vol.2 No.1, 2008, pp.33-42. 査読あり

[学会発表] (計7件)

- ① Feng Guo, Ying Dai, Shaozi Li, “Large-scale image/videos’ automatic annotation and retrieval based on the distributed users”, Proc. of IEEE ICIS 2009, pp. 48-52, Nov. 2009, China. 査読あり
- ② Ying Dai, “Representing meanings of images based on associative values with lexicons”, Proc. of IEEE IJH-MSP 2009, pp. 1222-1227, Sep. 2009, Japan. 査読あり
- ③ Ying Dai, Dawei Cai, “A mechanism for large image/videos’ automatic annotation considering semantic tolerance relation”, Proc. of IEEE ISSNIP 2008, pp. 133-138, Dec., 2008, Australia. 査読あり
- ④ Ying Dai, “Image/Video’ s automatic annotation considering semantics’ tolerance relation”, Proc. of IEEE SMC 2008, pp. 3417-3424, Oct. 2008, Singapore. 査読あり

- ⑤ Ying Dai, “Image/video’ s semantics assignment using bidirectional associative memories”, Proc. of SCI & ISIS 2008, pp. 1759-1764, Sep. 2008, Japan. 査読あり
- ⑥ Ying Dai, “Semantic tolerance relation-based image representation and classification”, IEEE Proc. of IPC 2007, pp.62-67, Oct. 2007, Korea. 査読あり
- ⑦ Ying Dai, “Multi-domain-based automatic image representation using semantic tolerance relation model”, IEEE Proc. of PACRIM’ 07, pp. 505-509, Aug. 2007, Canada. 査読あり

[図書] (計1件)

- ① Ying Dai, etc. (Eds.), IGI Global, Kansei Engineering and Soft Computing, 2010, 350

[その他]  
特になし

## 6. 研究組織

### (1) 研究代表者

戴 瑩 (Dai Ying)

岩手県立大学・ソフトウェア情報学部・  
准教授

研究者番号：60305290

### (2) 研究分担者

蔡 大維 (Cai Dawei)

岩手県立大学・ソフトウェア情報学部・  
准教授

研究者番号：30281041

Basabi Chakraborty

岩手県立大学・ソフトウェア情報学部・  
准教授

研究者番号：90305293

### (3) 連携研究者

( )

研究者番号