

クエリフィードバックによる意味情報推奨機構を有する成長型ユビキタス環境データベースの実装と評価

渡 邊 優 作[†] 佐々木 貴 司[†] 富 井 尚 志^{††}

本研究では、現実世界から得られるセンサの値や3次元形状データと意味情報を結びつけることで、現実世界と仮想空間とを対応付けたユビキタス環境の実現を目指す。この際、現実世界に存在するすべての物体に対して適切な意味情報を付加することは、ユビキタス環境を構築・運用するうえでの大きなコストとなる。我々は、このコスト低減を目的とする「クエリフィードバック」という手法を提案する。本手法では、物体が新たに環境に持ち込まれた際に、すでにDBに蓄積されているユビキタス環境のインスタンスデータを対象に、特徴量を用いたクエリによって絞り込みを行い、順位付きで対応すべき意味情報を推奨する。この結果から、利用者は上位に提示された意味情報から1つを選択するだけで、新たな物体の意味情報をDBに登録することができる。本論文では、ユビキタス環境をDB化する際に3次元ポリゴンデータを用いて特徴量を抽出した。また、本手法が有効であることを示すため、運用実験を行い、評価した。この結果、提案手法によって正しい意味情報が上位に推奨されること、および、登録インスタンスが増加するとともに精度が向上することが分かった。

Design and Implementation of Growing Ubiquitous Environment Database with Semantic Information Recommendation Module by Query Feedback

YUSAKU WATANABE,[†] TAKASHI SASAKI[†], and TAKASHI TOMII^{††}

In this study, we aim at fulfillment of ubiquitous environment that connected real world and a virtual world by relating sensor data and 3D shape data to semantic information. On this occasion, the number of treated objects are bigger, costs for construction of such environment and addition of appropriate semantic information to all objects are becoming bigger. We aim at cost reduction by proposing technique of "Query Feedback". When a new object is carried on such environment, the proposed technique narrows a list of the semantic information that related the instance data registered in DB down by an appropriate query which uses feature vector. The result shows ranked semantic information that should relate to the new object. As this result, user can register semantic information of the object by only selecting one of the semantic information in ranked list. In this paper, we use 3D polygon data in modeling of ubiquitous environment data, and we extract feature vector from them. We performed an operative experiment and evaluated it in order to indicate that the proposed technique is effective. As a result, correct semantic information is ranked in higher order by using proposed method. It is also shown that accuracy of the result is improved in the case that a number of registered instance grows higher.

1. はじめに

近年、計算機技術の発達や高速なネットワーク環境

の普及による技術的背景と、u-Japan 政策¹⁾に代表されるような政策的背景の両方のファクタからユビキタス環境の実現が現実味を帯びてきている。日常生活の中でも、スーパーやコンビニエンスストアで非接触のICカードを用いた買い物があたりまえのように利用でき、またRFIDタグが普及し安価になったことから、今後はすべての物体にIDタグが貼付されることも予想される。このユビキタス環境を背景に、井形らのOKAR²⁾では、人が持つICカードのIDや作成された文書データにメタ情報を結びつけることによって、人の持つ知識を有効に活用する研究を行って

[†] 横浜国立大学大学院環境情報学府情報メディア環境学専攻
Department of Information Media and Environment
Sciences, Graduate School of Environment and Informa-
tion Sciences, Yokohama National University

^{††} 横浜国立大学大学院環境情報研究院
Graduate School of Environment and Information Sci-
ences, Yokohama National University
現在、NEC ソフト株式会社
Presently with NEC Soft, Ltd.

いる．このような物体や人（以下，オブジェクト）とその意味情報（メタ情報）を管理し，ユビキタス環境を通じた高度なサービスを提供する研究は今後も発展していくと考えられる．また，時空間情報をデータベース化する研究が以前から行われている．金子らの MOVE³⁾ では，オブジェクトの動きに着目し，基本機能をクラスとして実装することで動きのシミュレーションや視覚化を実現している．また，本研究の先行研究となる現実シーンデータベース⁴⁾ では，意味情報を付加した基本モデル形状とスキャン形状とをマッピングすることにより，仮想 CG 空間上での実空間検索を実現している．これらを背景に我々は，現実空間と仮想空間を RFID などのタグによってマッピングし，オブジェクトや操作に関する意味情報の付加を行うことで，複数利用者で概念を共有・検索できる「概念共有環境 CONSENT」を提案してきた⁵⁾．概念共有環境 CONSENT では，オブジェクトのデータと意味情報を独立して格納することで，オブジェクト・操作のタイプ名称や説明といった単純な意味情報はもちろん，それらに関係を持たせた構造化した意味情報を扱うことが可能である．しかし，意味情報を用いたユビキタス環境下でオブジェクトを管理するためには，そのオブジェクトを DB に登録する際に，少なくとも以下のような情報を付与しなければならない．

- オブジェクトに付けられたタグやセンサの ID
- オブジェクトを表現するデータ
- オブジェクトの意味情報との関連付け

これらを 1 つ 1 つのオブジェクトに対して付加することは，ユビキタス環境を構築・運用するうえでのコストとなる．特に，オブジェクトの意味情報との関連付けを登録することは，次の理由から大きなコストとなる．すなわち，あるオブジェクトに対して利用者が意図した意味情報を，画像や形状データなどのマルチメディアデータの特徴量だけから正確に自動推測することは困難である．このため，オブジェクトに対して正確に意味情報を付与するためには，利用者の手動登録に頼らざるをえない．そこで我々は，すでに登録されたオブジェクトの情報を用いて，新規オブジェクト登録時の意味情報を推奨する「クエリフィードバック」という手法を提案し，ユビキタス環境構築・運用の実現性評価を目指してきた⁶⁾．本手法では新規オブジェクト登録時に取得できるデータとユビキタス環境 DB に蓄積された正しいインスタンスデータを参照するクエリを用いて意味情報の推奨を行う．クエリによって結びつけるべき意味情報を絞り込むことによって意味情報選択のコストを低減するとともに，ユビキタス環

境 DB に蓄積されたインスタンスデータを参照することで利用者の傾向を考慮した空間構築支援が実現できる．また，利用者がクエリによって推奨された意味情報を選択し，ユビキタス環境 DB に登録（フィードバック）することにより，次回登録時にそのデータが反映され，さらに正しい推奨が行える成長性が得られる．本論文ではクエリフィードバックの有用性・成長性を実証するため，提案してきたユビキタス環境 DB 「概念共有環境 CONSENT」上でプロトタイプシステムを実装し評価を行う．以下，2 章において概念共有環境 CONSENT を概説し，3 章で本論文の提案手法であるクエリフィードバックについて詳述する．そして 4 章でクエリフィードバックを概念環境 CONSENT 上に実装し，5 章でその有用性と成長性に対する評価実験を行う．6 章で関連研究をあげ，最後に 7 章でまとめと今後の課題について述べる．

2. 概念共有環境 CONSENT

本章では，我々がこれまでに提案してきたユビキタス環境 DB の 1 つである概念共有環境 CONSENT について概説し，そのオブジェクト登録コストについて考察する．

2.1 三層構造モデル

1 章で述べたように，ユビキタス環境では意味情報とオブジェクトとを結びつけることによって，より高度な日常生活をサポートするサービスの提供が可能となる．そこで我々は実空間に存在するオブジェクトや人がオブジェクトに対して行う操作の意味情報と，それらの表現手段であるマルチメディアデータを独立に定義し，意味情報とマルチメディアデータの間を関連付けるエンティティ（存在エンティティと定義）を実空間のセンサデータと対応付けて DB に蓄積する三層構造モデル化手法^{5),7)} を提案してきた．三層とは“意味情報”，“マルチメディアデータ”，“存在エンティティ”を指す．これらを独立に管理することにより，マルチメディアデータが表現する感覚的な情報だけでなく意味情報をも共有することが可能となり，知識や常識などの空間の概念や実空間の現在の状態である現実状態の共有を実現した．図 1 に概念共有環境 CONSENT の論理スキーマを示し，以下“意味情報”，“マルチメディアデータ”，“存在エンティティ”を格納するテーブル群について詳述する．

Concept Sharing Environment の略である．加えて，この環境の利用者は共有される概念の基盤，すなわち，オントロジにあらかじめ同意する（Consent）ことを意味する．

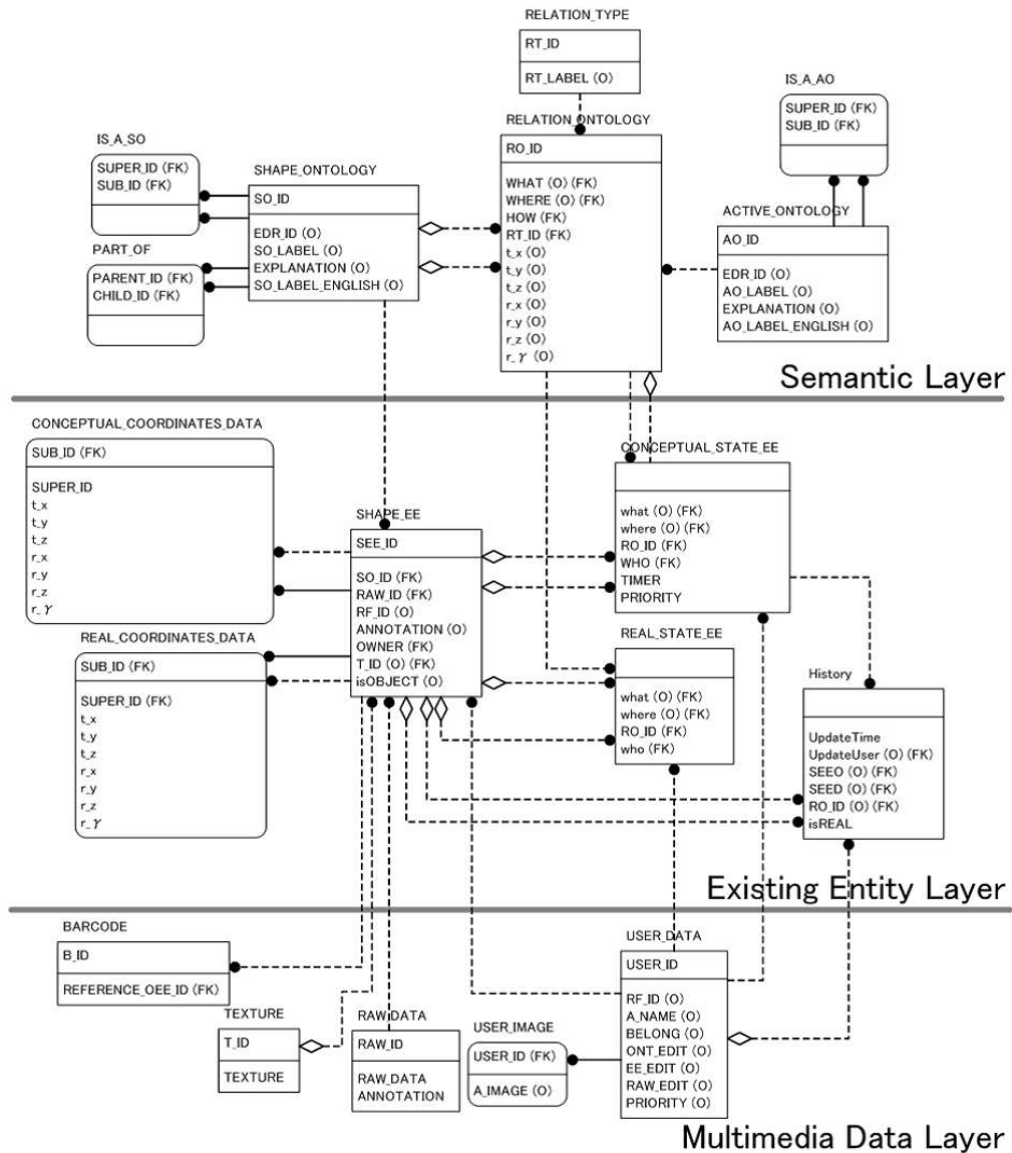


図 1 論理スキーマ

Fig. 1 Logical schema.

2.1.1 意味情報

概念共有環境 CONSENT で扱う意味情報は、オブジェクトに関する意味、操作に関する意味、そしてそれらの関係を表す意味に限定し、それぞれを格納するテーブルで構成する。インスタンスとして、オブジェクトに関する意味には「家具」, 「机」, 「本」など、操作に関する意味には「置く」, 「取る」などが挿入される。それぞれはタイプ名称だけではなく、タイプに関する説明や、外部オントロジへのリンクといった情報も含む。そしてこれらを参照する形で、「机」は「家具」である」という is_a 関係や「本」は「机」に「置く」こ

とができる」といった行動の関係を蓄積する。意味情報のインスタンスは、一般に意味情報のエキスパートによって入力されることを想定しているが、ひな形として既存のオントロジを流用し、サイトごとにカスタマイズすることが可能である⁸⁾。ひな形としては、次のようなものが有効である。まず、一般語彙・概念については、WordNet⁹⁾ や EDR 電子化辞書¹⁰⁾ の概念辞書などを用いることが可能である。また、流通商品を対象とする場合には、JICFS¹¹⁾ や Yahoo!Auction¹²⁾ の分類などが有効であろう。一方、異なるアプリケーションドメインの例として、医療ドメインを DB 化す

る場合¹³⁾には UMLS¹⁴⁾を利用することができる。

2.1.2 マルチメディアデータ

オブジェクトや操作を表現するマルチメディアデータは意味情報と独立して蓄積される。マルチメディアデータには音声、画像、動画、各種センサの値など様々なデータが考えられる。近年では Masunaga らの VWDB¹⁵⁾ やソニー株式会社の PAW¹⁶⁾、さらに仮想社会の Second Life¹⁷⁾ に代表されるように、3次元データを用いた空間表現も一般的に普及している。3次元データを用いた空間表現は、利用者にとって最も直観的である。たとえば、実空間検索の結果提示を考えた際、このファクタは非常に重要である。そこで、概念共有環境 CONSENT でもオブジェクトを表現するマルチメディアデータとして、ポリゴンモデルの形状データを採用した。したがって本論文では、特に形状データ(ポリゴンデータ)を対象として以降の議論を進める。

2.1.3 存在エンティティ

上述の意味情報とマルチメディアデータの間には、多対多の関係が存在する。例として、同じ形でタイプ名称が異なるモノもあれば、同じタイプ名称で異なる形のモノがあるというのが一般的である。そこで、まず現実世界に存在する「モノ」や「実際に行った操作」の1つ1つをエンティティとしてとらえ、「存在エンティティ」と呼ぶことにする。一方、意味情報とマルチメディアデータは独立のものと考え、両者の関係が実在することを存在エンティティによって表すことにする。すなわち、存在エンティティは、ある意味情報(のタイプ)に対して実在するインスタンスである。概念共有環境 CONSENT では、実在のオブジェクトに RFID タグや RFID リーダを取り付け、それらの ID やセンサ値を DB 内の存在エンティティに対応付ける。タグや RFID リーダの ID はオブジェクトの存在エンティティに相当し、そこから得られるセンサ値は操作の存在エンティティに相当する。これにより、現実からのリアルタイムな状態取得を実現できる。

2.2 アプリケーション

本節では、概念共有環境 CONSENT を用いて実装したアプリケーションをいくつか紹介する。最も単純なアプリケーションとして実空間物体検索があげられる。概念共有環境 CONSENT では、オブジェクトの位置情報だけでなく操作された状態を管理しているために、オブジェクトがどこにあって、どのような状態にあるのかという検索が可能である¹⁸⁾。図2に検索時のインタフェースを示す。次に、利用者が理想とする状態を蓄積することで、周囲の環境を整えられるよう



図2 物体検索のインタフェース
Fig.2 Interface of object retrieval.

な環境構築支援アプリケーションを考えている¹⁹⁾。これは2状態の差分をとり、エージェントによって整合を行うか、利用者に通知することにより、理想とする環境の構築支援を行うものである。また、蓄積された利用者の行った操作から、頻出する操作列を取得することで、利用者の支援を行う行動支援アプリケーションを実装している²⁰⁾。頻出する操作列から逸脱した行動を通知することで、忘れ物抑制などの行動支援が行える。

2.3 オブジェクト登録コスト

上述した概念共有環境 CONSENT において、新規オブジェクトを登録・管理することを考えた場合、以下の3つのデータを取得する必要がある。

- (1) オブジェクトに付けられるタグやセンサの ID
- (2) 形状データ
- (3) 意味情報との関連付け

これらのうち、(1)は実空間から容易に取得可能である。つまり、(2)と(3)をいかに低コストで取得するかが概念共有環境 CONSENT 構築のキーとなる。図3に構築コストの概念図を示す。以下では、これらのコストについて考察する。

2.3.1 形状データ登録コスト

前述したように、概念共有環境 CONSENT では実空間検索を直観的に行うために、ポリゴンデータを採用している。ここではポリゴンデータの取得コストを考察する。ポリゴンデータを取得するには以下のような方法が考えられる。

- (1) 3D スキャナによる読み込み
- (2) 人手による作成
- (3) CAD データなど外部データの再利用

(1)の3Dスキャナは安価ではないが、ポリゴンデータを容易に取得可能である。(2)の人手による作成に

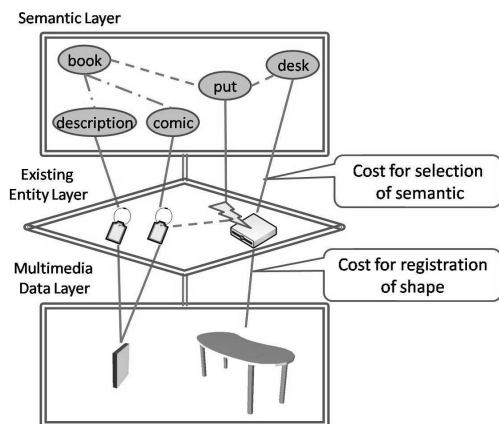


図3 新規オブジェクト登録コスト

Fig. 3 Object registration cost.

については, Second Life¹⁷⁾ のように, エキスパートだけではなく, ポリゴンデータを利用者自らが作成することも可能になってきたと考える. また, 日常で扱うオブジェクトを考えた場合, 形状が似通っているものも多く, (3) のようにポリゴンデータを再利用できる可能性が高い. そこで, あらかじめ利用者が多く使用しそうなポリゴンデータを作成・蓄積しておくことで, 利用者はその中から選択するという手間のみで登録が可能になる. 本論文では, このポリゴンデータの再利用性について評価を行う.

2.3.2 意味情報との関連付けコスト

オブジェクトの登録では, オブジェクトに相当するタイプを意味情報として付加する. 現実のオブジェクトの意味情報(タイプ)を判断し, 手動で登録するのは膨大なコストがかかる. その理由としては大きく2つがあげられる. 1つ目は意味情報の多様性である. ユビキタス環境では様々なオブジェクトを管理対象とする. たとえば, 後述の概念共有環境 CONSENT の運用実験では, 意味情報として商品体系である JICFS¹¹⁾ と Yahoo! Auction¹²⁾ の書籍分類を組み合わせたものを用いたため, is_a 関係といったオブジェクト間の情報を除いても, その要素数は 3,000 を超える. いい換えれば, オブジェクトのタイプが 3,000 以上存在することとなる. このような多数のタイプの中から, 登録オブジェクトのタイプにふさわしい1つを選択するには大きな手間がかかる. 2つ目は利用者の傾向を考慮する困難さである. たとえば, 同じモノを扱う場合でも, タイプとして「計算機」「パソコン」「PC」のどれを与えるかは利用者次第である. しかし, これらを個別に与えてしまうと, 意味情報での統一的な管理が行えなくなってしまう. そのため, 利用者はそのユビキ

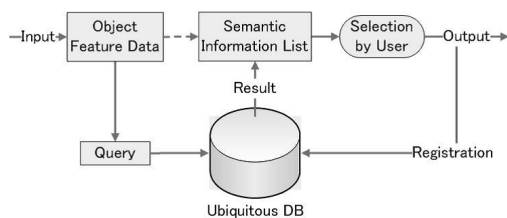


図4 クエリフィードバックの流れ

Fig. 4 Job flow of query feedback.

タス環境下において, つねに意味情報を意識しなければならないというコストが発生する. そこで次章では, これらを考慮した意味情報推奨手法「クエリフィードバック」について詳述する.

3. クエリフィードバック

本章では, オブジェクトと意味情報の結びつけに関する利用者の手間を低減するための提案手法である「クエリフィードバック」について述べる. クエリフィードバックは以下の流れで行われる.

- (1) 新規オブジェクトから特徴量を抽出.
- (2) 特徴量に基づいたクエリを生成し, 検索結果の意味情報リストを利用者に提示.
- (3) 利用者はリスト内から正しい意味情報を選択し, オブジェクト情報を DB に登録(フィードバック)する.

この流れを図4に示した. 以下, 各工程について詳しく述べる.

3.1 特徴量の抽出

特徴量として用いる値は, オブジェクト登録時に取得でき, オブジェクトの意味情報を推測可能な値ならば何でもよい. たとえば, オブジェクトの画像を取得し, 画像抽出の技術を用いた特徴量が考えられる. 本論文では, 特徴量として容易に取得できると考えられる商品 ID (Commodity code) や, 実装環境である「概念共有環境 CONSENT」が形状データとしてポリゴンモデルを採用していることを利用して, 表面積・体積・Bounding volume などの形状データ特徴量²¹⁾を用いた.

3.2 クエリの生成と結果提示

特徴量に基づいて, すでに登録済みのオブジェクトを参照する適切なクエリを生成し, 意味情報の絞り込みを行う. すでに登録されたオブジェクトに結びつけられた意味情報は, 利用者の傾向を反映していると考えられる. これを参照することで, 利用者に沿った意味情報の推奨が可能となる. このとき, 参照するデータを利用者各人の登録データのみとするか, 他者の登

録データも利用するかは自由である．本研究では，概念共有環境 CONSENT が集団での概念共有を目的としていることから，他者の登録データも含めて参照するように実装した．また，提示される結果は，特徴量によって提示方法が変化する．たとえば，体積などのスカラー量を特徴量とすれば，その結果としては，体積の近い順に順序付けられたリストが結果として提示される．商品コードなどオブジェクトを特定できるデータの特徴量とすれば，それに対応付けられた意味情報が一意に決定される．

3.3 意味情報の選択と登録

利用者が選択を行うことで，登録されるオブジェクトの意味情報に，より利用者の意思が反映する．このとき，本手法によって提示されたリストの上位に正解となる候補が推奨されていれば，意味情報選択の手間を軽減できる．また，特徴量とオブジェクトをともに登録することで，そのデータも次回登録時には提示される意味情報リストに影響を与えることとなり，より利用者の傾向を加味したリストを提示することができるようになる．この登録されたオブジェクトが次回以降の検索結果の順位に影響することから，本手法をクエリフィードバックと名付けた．

4. 実 装

上述した概念共有環境 CONSENT において，新規オブジェクトの登録を支援するために，本論文の提案であるクエリフィードバックを実装した．以下，実装方法について詳述する．

4.1 特 徴 量

本研究では特徴量として，商品コード，意味的制約，形状データを用いて実装を行った．それぞれの特徴量について以下に述べる．

- 商品 ID (Commodity code)

商品 ID とは，JAN コード²²⁾ や ISBN²³⁾ に代表される，オブジェクトの種類を特定可能な ID を指す．商品 ID はバーコードなどにより，容易に取得可能であるため特徴量に適していると考えた．今回の実装では，取得できる商品 ID と意味情報は 1 対 1 に対応すると考え，商品 ID が取得できた場合は，利用者による選択の工程を省くように実装を行った．もし，商品 ID と意味情報が 1 対 1 に対応しない場合は，利用者を選択をゆだねるように実装すればよい．また，商品 ID は意味情報だけでなく形状データも識別できると考え，2.3.1 項で述べたポリゴンデータの選択の工程も省いた．

- 意味的制約 (Semantic constraint)

概念共有環境 CONSENT では，オブジェクトと操作間の関係を蓄積しているため，その関係を制約として，意味情報リストを絞り込むことができる．たとえば，机の上で登録が行われた場合は“机”に“置く”ことができるものをリストとして提示する．今回の実装では，利用者が登録オブジェクトを初めに置く場所を，ポリゴンデータを用いた仮想空間から指定することにより，意味情報を用いた絞り込みを行えるように実装した．

- 形状データ (Shape data)

ポリゴンデータが持つ形状から得られる値を特徴量として用いる．形状が持つ値は多くあるが，本研究では表面積・体積・Bounding volume の三次元ベクトル (以下，VSB) と D2²⁴⁾ を実装した．D2 とは，ポリゴンにおける 2 頂点の距離をヒストグラムで表した，多次元ベクトルである．

4.2 特徴量に基づいたクエリ

ここでは，上述した特徴量を用いて意味情報を絞り込むクエリを実装する．このクエリは図 1 の論理スキーマに対して発行される．商品 ID を特徴量に用いたクエリを図 5 に示す．図 5 のクエリは，商品 ID である @BARCODE をキーとして，既登録のオブジェクトに結びつけられた意味情報を取得する．商品 ID は BARCODE テーブルによって一意に管理され，新規オブジェクトから取得された商品 ID (@BARCODE) が未登録であった場合は，商品 ID と新規オブジェクトに付与するオブジェクト存在エンティティ ID の対応を登録する．すでに登録されていた場合は，登録されているオブジェクト存在エンティティに付けられた意味情報を新規オブジェクトの意味情報とする．次に，意味的制約を用いたクエリを図 6 に示す．図 6 のクエリは，利用者によって指定された，登録オブジェクトを初めに置く場所 (@WHERE) の持つ意味情報から，その場所で操作可能なオブジェクトの意味情報を取得する．この“何”が“どこ”に“どうする”ことができるという関係は CONCEPT テーブルに格納され，そこを参照することで意味的制約による意味情報の絞り込みが可能となる．最後に形状データを用いたクエリを図 7 に示す．図 7 のクエリは，形状データの特徴量距離から意味情報を絞り込む．今回実装した形状データ特徴量は，VSB と D2 どちらも多次元ベクトルであるため，新規オブジェクトの形状データと DB 内の各形状データを比較して，各々ユークリッド距離を算出し，RAW_DATA テーブルの SIM 属性に格納した．SIM 属性に格納される値はクエリによって表現できる

```

SELECT OBJECT_ONTOLOGY.OO_ID
FROM BARCODE,OBJECT_ONTOLOGY,
      OBJECT_EE
WHERE BARCODE.REFERENCE_OEE_ID
      = OBJECT_EE.OEE_ID
AND OBJECT_EE.OO_ID
      = OBJECT_ONTOLOGY.OO_ID
AND BARCODE.B_ID = @BARCODE

```

図 5 商品 ID を用いたクエリ
Fig. 5 Query with commodity code.

```

SELECT DISTINCT OO2.OO_ID
FROM OBJECT_EE, CONCEPT,
      OBJECT_ONTOLOGY AS OO1,
      OBJECT_ONTOLOGY AS OO2
WHERE OBJECT_EE.OO_ID
      = OO1.OO_ID
AND OO1.OO_ID = CONCEPT.[WHERE]
AND CONCEPT.WHAT = OO2.OO_ID
AND OBJECT_EE.OEE_ID = @WHERE

```

図 6 意味的制約を用いたクエリ
Fig. 6 Query with semantic constraint.

が、多次元のユークリッド距離を求めるクエリは複雑になりすぎるため、SIM 属性に格納するように実装した。その後、それぞれの形状データを持つオブジェクト存在エンティティを参照し、それらが持つ意味情報ごとに SIM を平均化して、その値が高い順に意味情報リストを提示する。なお、3.2 節で言及したように、いずれのクエリも利用者情報を含んでいない。すなわち、他者が過去に登録した情報も参照して絞り込みを行っている。利用者個人ごとに傾向がある場合、これらのクエリに利用者情報による絞り込みを付加することで、個人ごとのカスタマイズが可能である。

4.3 オブジェクト登録の流れ

実装したオブジェクト登録のフローチャートを図 8 に示す。新規オブジェクトを登録する際、商品 ID が取得できる場合は、その商品 ID が結びつけられている存在エンティティの意味情報とポリゴンデータを用いて登録を完了する。商品 ID が取得できない場合、利用者は登録しようとするオブジェクトを初めに置く場所を、仮想空間上から指定する。その後、図 9 に示すユーザインタフェースを用いて、登録の支援を行う。利用者はオブジェクトに該当する形状データを、図 9 の領域 (a) から選択する。領域 (a) に示される文字列

```

SELECT AVG(SIM),
      OBJECT_ONTOLOGY.OO_ID
FROM RAW_DATA, OBJECT_EE,
      OBJECT_ONTOLOGY
WHERE RAW_DATA.RAW_ID
      = OBJECT_EE.RAW_ID
AND OBJECT_EE.OO_ID
      = OBJECT_ONTOLOGY.OO_ID
GROUP BY OBJECT_ONTOLOGY.OO_ID
ORDER BY AVG(SIM)

```

図 7 形状データを用いたクエリ
Fig. 7 Query with shape data.

群は、形状データを選択するために付けた便宜的なもので、意味情報との直接的な関係はない。選択時、新規オブジェクトに相当する形状データがリスト上にない場合は、3D 形状スキャナ²⁵⁾ やポリゴンデータ作成ソフト²⁶⁾ を用いて取得することも可能である。システムは形状データの特徴量を計算し、その特徴量や初めに指定した意味的制約を用いて、意味情報を絞り込むクエリを DB に対して発行する。このクエリの結果を図 9 (b) の領域に表示し、利用者はそこから関連付けたい意味情報を選択する。意図する意味情報がリストされていない場合は、全意味情報から選択することも可能である。

5. 評価実験

概念共有環境 CONSENT の実現性を示すために、ポリゴンデータの再利用性の評価と 4 章で実装したクエリフィードバックの評価を行った。まず、実験の準備として本研究室の環境を再現した概念共有環境 DB を構築した。クエリフィードバックはすでに登録されているインスタンスデータを参照するため、初期の段階である程度の数のインスタンスをあらかじめ登録しておく必要がある。そこで、実験準備としてその時点で実験対象空間（研究室）に存在していた 450 個のオブジェクトに関する情報（存在エンティティ 450 個と、それぞれに対応する意味情報および形状データ）を手動で登録した。次に、この環境を 3 カ月間日常生活の中で運用して、研究室に新たに増えたオブジェクトを登録した。登録されたオブジェクトの一部を表 1 に示す。図中の Hypernym は登録された意味情報の上位分類を示す。ここで登録されたオブジェクトの意味情報が、各特徴量での推奨を用いた場合に、何位に順位付けされるかを測定することで、クエリフィードバックの有用性の評価と、クエリフィードバックを組み込

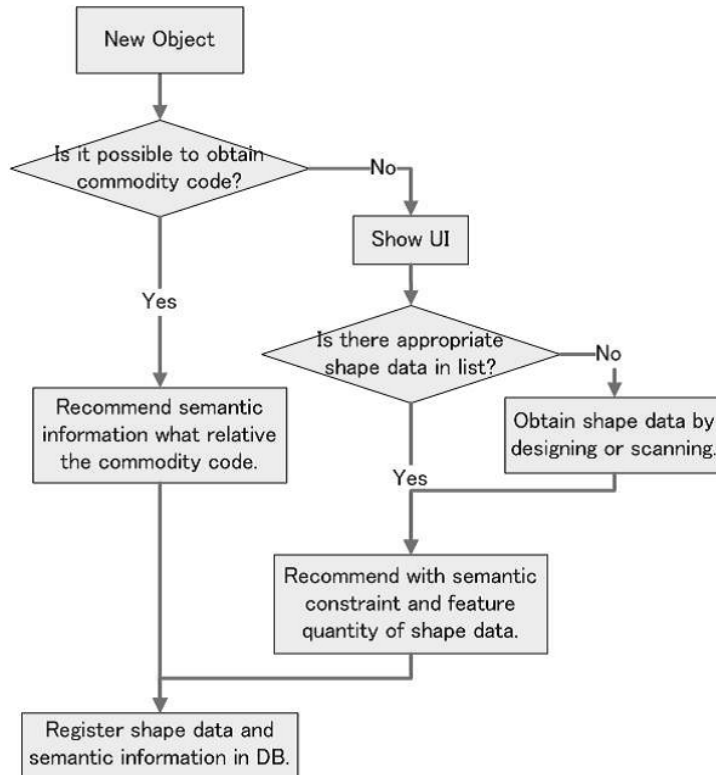


図 8 オブジェクト登録フローチャート
Fig. 8 Flowchart of object registration.

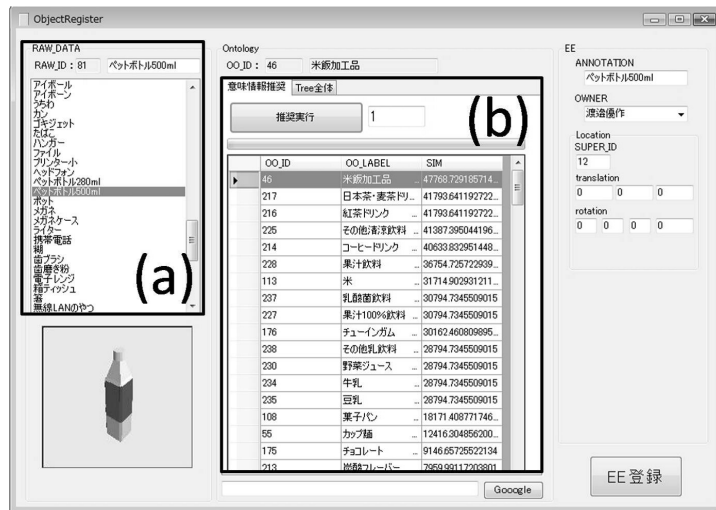


図 9 オブジェクト登録インターフェイス
Fig. 9 Interface of object register.

んだコピキタス環境 DB の成長性に関する評価を行った。実験に関する情報を以下に示す。なお、運用実験で結果的に用いられた形状データは、オブジェクトの寸法を測定し、人手によって作成したものがほとんどであった。2.3.1 項でも述べたように、今後は 3D ス

キャナや外部で作成された 3 次元データを利用した場合も考慮する必要がある。

- 実験室環境
 - 広さ：169 m²
 - 利用者数：14 人

表 1 登録オブジェクトのインスタンス
Table 1 Registered instances.

Hypernym	Semantic information
OA 機器	電話機
アイスクリーム類	ファミリーアイス
いす(椅子)・腰掛け	ひじ掛けいす
ファイル・バインダー	ファイル
コンピュータとインターネット	システム設計, 開発
コンピュータ資格試験	情報処理技術者試験
スープ	インスタント味噌汁・吸物

- 運用実験期間
 - 開始:平成 18 年 9 月 25 日
 - 終了:平成 18 年 12 月 20 日
- 意味情報
 - オブジェクトの意味情報(タイプ): 3,602 個
 - 操作の意味情報(タイプ): 17 個
 - オブジェクトと操作の関係: 62,851 個
 - 運用実験中に使用されたオブジェクトの意味情報(タイプ): 112 個
- 存在エンティティ
 - 運用実験前: 450 個
 - 運用実験後: 1,038 個

5.1 ポリゴンデータの再利用性の評価

実験準備および実験中に登録されたオブジェクトのうち、ポリゴンデータの取得方法とそのコストを表 2 に、取得コストの割合の変化を図 10 に示す。図 10 の横軸は登録オブジェクト数、縦軸は取得コストごとの占める割合を示す。表 2 に示すように、再利用ができず人手で作成したものは比較的高いコスト(High)、商品コードからその形状を一意に特定した場合は最もコストが低い(Low)と見なした。また、再利用は可能だが人による選択を要したものをその中間(Middle)として位置付けた。図 10 に示したとおり、準備段階の開始直後はほとんどの形状データが新たに登録されたもの(High)によるが、準備が進むにつれて、形状データの再利用率が上がっている。運用実験を開始した後(図 10 の横軸 450 以降)も、High の全体に占める割合は減少していき、登録オブジェクト数が 1,000 を超えた時点では約 9%と低く、ポリゴンデータの再利用性を示せたと考える。再利用できたもののうち、Middle が圧倒的に多く、Low は最終的に約 4%しかなかった。これは、商品コードが付いていても、ポリゴンデータを選択する手間と比べて大きな差がないため、利用者は商品コードを用いずに登録したためであった。

5.2 クエリフィードバックの有用性の評価

クエリフィードバックの有用性を評価するため、3カ

表 2 ポリゴンデータの取得方法

Table 2 Acquisition method ratio of polygon data.

Cost	How to obtain shape data
Low	Reuse
Middle	Specify with commodity code.
High	Selection by user with UI.
	3D Scanner, Design by expert.

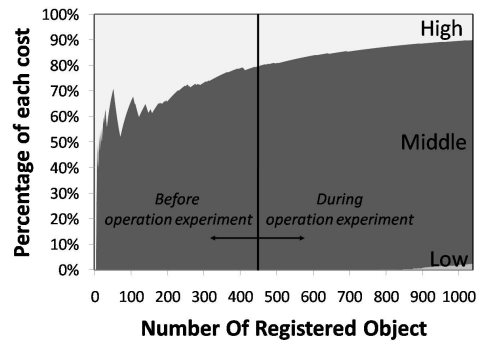


図 10 取得コストごとの割合の変化

Fig. 10 Change in the rate of each cost.

月間の運用実験後の DB において以下の実験を行った。

- (1) 登録されたオブジェクトに対して、意味情報を絞り込む際に用いる特徴量の組合せを変えた場合、登録された意味情報が何位に推奨されるかを測定する。
- (2) (1) をすべての登録されたオブジェクトに対して行う。

ここで、特徴量に商品 ID を用いた場合は、今回意味情報が一意に特定されるよう実装したため、評価の対象外とした。結果を図 11 に示す。図 11 のグラフの横軸は各特徴量を用いた場合に、実際に登録された意味情報が何位に推奨されたかを表す。また、縦軸は登録されたオブジェクトの中で、横軸の順位以内に正解の意味情報が推奨された割合を表す。たとえば、形状データ特徴量である VSB のみを用いた場合、約 40%が 20 位以内に推奨されたことが分かる。支援なしの場合はオントロジのすべての要素の中から正解の意味情報を選択するため、意味情報が推奨される順位が高くなるほど、その順位以内に推奨される割合は比例的に低くなる。推奨を用いた場合、順位付けされたリストから選択するという手間を考慮すれば、より高い順位に多く推奨されることが望ましい。すなわち、結果が図 11 の矢印の方向にあるほど望ましいといえる。図 11 の結果から、クエリフィードバックを用いない場合に比べて、用いた場合の方が適切な意味情報を絞り込むことを示した。意味的制約のみを用いた場合は支援なしの場合と比べて意味情報を絞り込むことはできたが、提示された意味情報に順位付けがなされていないため比例的なグラフになった。これに対し、形状デー

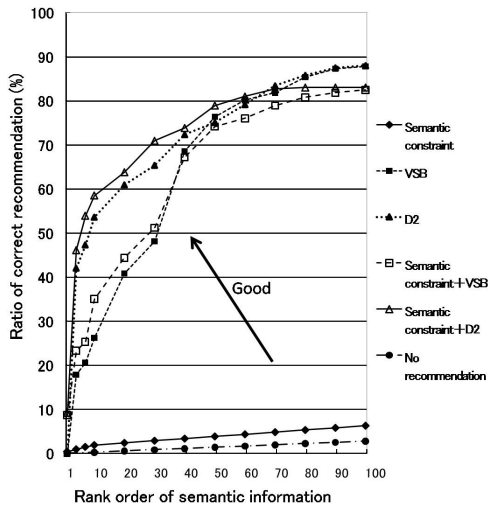


図 11 Recommended order every feature quantity.

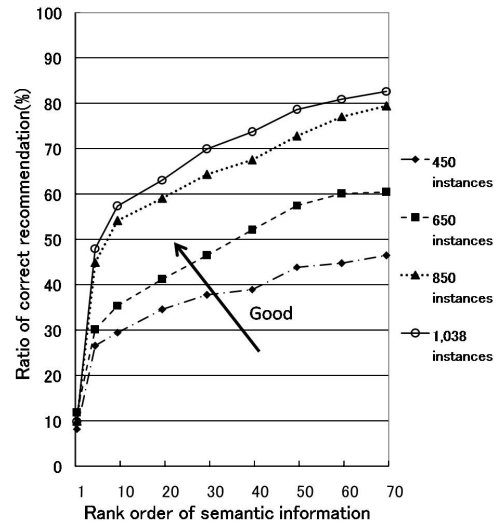


図 12 Recommended order each number of instances.

タ特徴量を用いると、形の類似性も考慮して順位付けするため良い結果となった。このとき、形状データ特徴量のみでは今回の実験では D2 の方が VSB より良い形状データ特徴量であるといえる。VSB を用いた場合は 40 位前後で、D2 を用いた場合は 70 位前後で意味的制約を用いた場合と用いない場合が逆転している。これは、実験準備の段階で一部のデータに意味的制約の特徴量として与えた意味情報があらかじめ登録されていなかったことが原因であった。たとえば、「パーソナルコンピュータ」を登録したとき、はじめに置く場所として「部屋」が選択されていた。この「パーソナルコンピュータ」と「部屋」の間に「置ける」の意味情報を事前に登録していなかったため、意味的制約を用いた場合、提示されなくなってしまった。しかし、一般には VSB と D2 両者とも、より高い順位において意味的制約を加えた場合の割合が高くなっていることから、意味的制約を用いた方が結果が良くなっていると考えられる。結論として形状データ特徴量に D2 を用いて、かつ意味的制約を組み合わせた場合が最も良い結果となった。しかし、本手法を用いた場合に適切な意味情報が提示される割合はどの特徴量を用いたときにも、ある値で飽和している。これは、本手法が DB 内のインスタンスデータを基に意味情報を絞り込むため、インスタンスデータに結びつけられていない意味情報は提示されないためである。インスタンスデータに結びつけられていない意味情報に関する提示は今後考えていく必要がある。

5.3 成長性に関する評価

クエリフィードバックは DB 内のインスタンスデー

タを用いて意味情報を絞り込むため、DB 内のインスタンスデータが増えると提示されるリストの精度は向上すると考えた。そこで、これを実証するため以下の実験を行った。特徴量については 5.2 節の結果より最も評価の良かった、D2 と意味的制約を組み合わせたものを用いた。

- (1) 運用実験前の状態の DB において、登録された意味情報が何位に提示されるかを測定する。
- (2) 運用中に登録された意味情報の中からランダムに選択、登録して DB のインスタンスを増やした状態で (1) と同様に測定を行う。
- (3) (2) を DB のインスタンスの数が運用実験後の状態になるまで行う。

結果を図 12 に示す。グラフの見方は図 11 と同様である。パラメータは DB 内の存在エンティティのインスタンス数を表す。図 12 の結果からインスタンス数を増やしていくことで、より高い順位に適切な意味情報が推奨されている割合が高くなることが分かった。これにより、登録数が増えることにより適切な意味情報を絞り込めるという成長性を示したといえる。

6. 関連研究との比較

本章では、ユビキタス環境を想定したオブジェクト登録支援の関連研究を紹介し、本研究との比較を行う。

6.1 状態遷移図を用いたモノの推定

前川らの研究^{(27),(28)}はオブジェクト(モノ)に温度センサ・照度センサ・焦電センサ・三軸加速度センサを取り付けたユビキタス環境を想定している。ユビキタス環境に持ち込まれたモノを推定するために、あら

かじめモノがとりうる状態、状態間を遷移するセンサ値の条件を定義し、モノごとに状態遷移図を作成しておく。これと、モノに取り付けられたセンサから取得される値を比較することで、そのモノが何であるかを推定する。この手法は、利用者に対する手間として、モノにセンサを取り付けるのみで済むという点で優れている。しかし、モノの推定には状態遷移が不可欠であるため、“机”といった、比較的静的なモノの推定ができない。また、状態変化がセンサによって得られたとしても、ある程度のセンサ値を蓄えたうえで始めて推定に至る。本論文で提案したクエリフィードバックでは、登録時に意味情報が決定するため、すぐにユビキタス環境の恩恵を受けることが可能である。

6.2 Spot & Snap

米澤らの研究²⁹⁾も前川らと同様に、オブジェクトにセンサが取り付けられているスマートオブジェクトサービスと呼ばれる環境を想定している。この研究では、センサを登録したいオブジェクトに取り付け、ライトで照射する手法を採用している。取り付けられたセンサには照度センサが含まれており、センサ群の中から最高値を示しているセンサを容易に特定できる。このセンサが持つ ID とライト照射の際に同時に取得した画像を対応付けることで、スマートオブジェクトサービスを構築している。この手法の利点は、やはり登録時の利用者の手間が少ないという点にある。センサを取り付け、ライトで照らすだけなので、計算機を扱わない初心者でも容易に登録が可能である。また、携帯電話などの普及している端末との親和性も高い。この手法では、オブジェクトのメタ情報の取得は行わず、画像によるオブジェクトの表示のみを用いる。米澤らはメタ情報がない欠点よりも、オブジェクトの画像とセンサのみで提供できるサービスの有用性に注目している。しかし、メタ情報を扱わないサービスでは、語彙による検索ができないなどの欠点があり、我々の意味情報のようなメタ情報の必要性も認めている。

6.3 DSW の開発

意味情報に基づいて形状データ群を統合して、検索可能にした研究例として Albertoni らの Digital Shape Workbench (DSW) の開発³⁰⁾ があげられる。DSW では、形状データに適用可能な意味情報を、独立したオントロジとしてあらかじめ記述している。この意味情報に基づいて、形状データとその取得モダリティ、そこから計算される部分的特徴量を含めてデータベース化する。また、これらのデータとオントロジの間に参照関係を持たせるフレームワークを提案している。これらによって、様々な手段で取得された形状データ

をオントロジの要素をキーに検索することができ、セマンティックウェブの枠組みで形状データを取り扱うことが可能となる。この結果、ウェブ上や各種情報リポジトリに散在する多種の形状データを検索エンジンなどでまとめて取り扱え、意味情報をキーとした形状データの再利用が実現できる。以上より、意味情報と形状データを対応付ける研究として、本研究と一致する。しかし一方、本研究ではさらに、形状以外の環境から得られるデータ (RFID やバーコードなど) も追加して、現実世界でのオブジェクトの状況を加味した形状データの再利用を試みた。また、利用者からのフィードバックが精度向上に有効であることを示した。このように、ユビキタス環境下でのデータ取得と再利用可能性、およびデータベースの成長可能性を提案し、評価している点で Albertoni らの研究と異なる。

7. まとめと今後の課題

本論文では、ユビキタス環境におけるオブジェクトと意味情報の結び付けに関する利用者の手間を低減させるための手法であるクエリフィードバックを提案した。また、提案手法を概念共有環境 CONSENT 上に実装し、評価実験を行うことで本手法の有用性とユビキタス環境 DB の成長性を示し、その構築・運用のコスト削減を実現した。今後の課題として、意味情報の絞り込みにおける DB 内のインスタンスの数をを用いた重み付けや、まだインスタンスデータに結びつけられていない意味情報の提示、形状データ以外を用いたクエリフィードバックの適用による存在エンティティの登録コスト削減などがあげられる。

謝辞 本研究は平成 17 年度および 18 年度横浜国立大学大学院環境情報研究院共同研究プロジェクトの助成を受けて行った。また本研究の一部は、文部科学省科学研究費補助金 (課題番号 17700097) の支援による。

参考文献

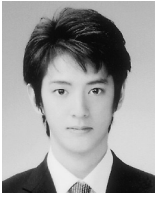
- 1) 総務省: u-Japan 政策 HP. http://www.soumu.go.jp/menu_02/ict/u-japan/index.html/
- 2) 井形伸之, 小櫻文彦, 片山佳則, 津田 宏: セマンティックグループウェア: RDF を用いた KnowWho の実現, 第 5 回セマンティックウェブとオントロジー, 人工知能学会研究会資料 (SIG-SWO-A303) (2004).
- 3) 金子邦彦, 黒木 進, 牧之内顕文: アニメーションデータベースシステム MOVE における動きの格納と検索法, 電子情報通信学会論文誌, Vol. J79-D-1, No.10, pp.834-842 (1996).

- 4) 富井尚志, 小林みな子, 有澤 博: 仮想 CG 空間へのマッピングによる現実シーンデータベースの設計, 電子情報通信学会論文誌, Vol.J82-D-I, No.1, pp.211-222 (1999).
- 5) 富井尚志: マルチメディアデータベースに基づく高度コミュニティ空間の実現, 電子情報通信学会誌, Vol.89, No.6, pp.511-517 (2006).
- 6) 佐々木貴司, 富井尚志: クエリフィードバックによる意味情報推奨機構を有する成長型ユビキタス環境データベース, 電子情報通信学会データ工学研究専門委員会主催第 18 回データ工学ワークショップ (DEWS2007), *Proc. Data Engineering Workshop, DEWS2007*, D7-9 (2007).
- 7) 南 博康, 賀来健一, 富井尚志: 意図共有を実現するオフィス仮想環境の設計, 電子情報通信学会技術研究報告, Vol.104, No.345, pp.1-7 (2004).
- 8) 渡邊優作, 富井尚志: 高度コミュニティ空間におけるオントロジ言語 OWL を用いた意味層の構築, 情報科学技術フォーラム (FIT2006) 情報科学技術レターズ, Vol.5, pp.59-62, LD-005 (2006).
- 9) Princeton University Cognitive Science Laboratory: Word Net. <http://wordnet.princeton.edu/>
- 10) 情報通信研究機構: EDR 電子化辞書. http://www2.nict.go.jp/r/r312/EDR/J_index.html
- 11) 流通システム開発センター: JICFS. <http://www.dsri.jp/company/jicfsifdb/top.htm>
- 12) Yahoo Japan Corporation: Yahoo! Auction. <http://auctions.yahoo.co.jp/>
- 13) 石江哲也, 砂子一徳, 富井尚志: 関心領域情報付き画像所見を蓄積する WB-PET DB の構築と検索による診断支援, 情報処理学会論文誌: データベース, Vol.46, No.SIG8, pp.1-13 (2005).
- 14) Bodenreider, O.: Medical Ontology Research: A Report to the Board of Scientific Counselors of the Lister, *Hill National Center for Biomedical Communications* (2001).
- 15) Masunaga, Y. and Watanabe, C.: Design and Implementation of a Multi-modal User Interface of the Virtual World Database System (VWDB), *Proc. 7th International Conference on Database Systems for Advanced Application (DASFAA)*, pp.294-312 (2001).
- 16) 松田晃一, 三宅貴浩: パーソナルエージェント指向仮想社会 PAW (第 2 版) の構築と評価, 情報処理学会論文誌, Vol.41, No.10, pp.2698-2707 (2000).
- 17) Linden Lab: Second Life. <http://secure-web12.secondlife.com/>
- 18) 渡邊文隆, 賀来健一, 富井尚志: 概念状態と現実状態を統合するユビキタス空間の実現性に関する検証, 電子情報通信学会技術研究報告, Vol.105, No.338, pp.1-6 (2005).
- 19) 麦嶋慎也, 富井尚志: 概念共有環境 CONSENT における現実状態と概念的状態の整合支援手法の提案, 情報科学技術フォーラム (FIT2006) 講演論文集, D-041, pp.93-96 (2006).
- 20) 小川悌知, 賀来健一, 渡邊文隆, 佐渡山英史, 富井尚志: 概念共有環境 CONSENT における行動パターンの取得, 電子情報通信学会データ工学研究専門委員会主催第 17 回データ工学ワークショップ (DEWS2006), *Proc. Data Engineering Workshop, DEWS2006*, 4B-i9 (2006).
- 21) Bustos, B., Keim, D., Saupe, D., Schreck, T. and Vranić, D.: Feature-based similarity search in 3D object databases, *ACM Computing Surveys*, Vol.37, No.4, pp.345-387 (2005).
- 22) 流通システム開発センター: JAN コード. <http://www.dsri.jp/company/jan/>
- 23) 日本図書コード管理センター: ISBN コード. <http://www.isbn-center.jp/>
- 24) Osada, R., Funkhouser, T., Chazelle, B. and Dobkin, D.: Shape Distributions, *ACM TOG*, Vol.21, No.4, pp.807-832 (2002).
- 25) cyberware: cyberware 社. <http://www.cyberware.com/>
- 26) Autodesk: 3ds max. <http://www.discreet.com/products/3dsmax/>
- 27) 前川卓也, 柳沢 豊, 服部正嗣, 岡留 剛: ユビキタス環境におけるコンテキストウェアのためのモノの表現, 日本データベース学会 Letters, Vol.5, No.2, pp.45-48 (2006).
- 28) 前川卓也, 柳沢 豊, 岡留 剛: Tag and Think: センサネットワークを前提としたモノ自身とその状態の推定, 第 13 回ユビキタスコンピューティングシステム研究会, Vol.2007, No.14, pp.211-218 (2007).
- 29) 米澤拓郎, 榊原 寛, 中澤 仁, 永田智大, 高汐一紀, 徳田英幸: 日常物とセンサノードの関連付け手法の提案, 第 13 回ユビキタスコンピューティングシステム研究会, Vol.2007, No.14, pp.179-186 (2007).
- 30) Albertoni, R., Papaleo, L., Pitikakis, M., Robbiano, F., Spagnuolo, M. and Vasilakis, G.: Ontology-Based Searching Framework for Digital Shapes, *OTM Workshops 2005*, Vol.3762, pp.896-905, Springer (2005).

(平成 19 年 6 月 20 日受付)

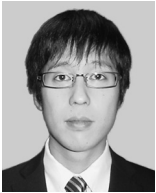
(平成 19 年 9 月 18 日採録)

(担当編集委員 白石 陽)



渡邊 優作 (学生会員)

平成 18 年横浜国立大学工学部電子情報工学科卒業。同年同大学院博士課程前期入学，現在に至る。マルチメディアデータベース，時空間データベースの研究に従事。



佐々木貴司

平成 17 年横浜国立大学工学部電子情報工学科卒業。平成 19 年同大学院博士課程前期修了。マルチメディアデータベース，時空間データベースの研究に従事。同年 NEC ソフト株式会社に入社，現在に至る。



富井 尚志 (正会員)

平成 6 年横浜国立大学工学部電子情報工学科卒業。平成 11 年同大学院博士課程修了。博士 (工学)。同大学工学部助手，同大学院環境情報研究院・講師，助教授を経て，現在，同大学院准教授。マルチメディアデータベース，時空間データベースの研究に従事。電子情報通信学会，映像情報メディア学会，日本データベース学会各会員。