

007.6
AN

マルチメディアデータベースにおける
視覚的ナビゲーションと検索

12680338

平成12年度～13年度科学研究費補助金(基盤研究(C)(2))研究成果報告書

平成14年3月

横浜国立大学附属図書館



11446393

研究代表者 有澤 博

(横浜国立大学大学院環境情報研究院)

1. はしがき

本報告書は平成 12 年度～13 年度にわたり、科学研究補助金基盤研究 (C) (2)「マルチメディアデータベースにおける視覚的ナビゲーションと検索」に基づく研究成果報告書である。

①本研究の目的

最近のコンピュータ技術の発達は著しく、安価なパソコンでもかなり精密なカラー画像や動画画を表示することが可能になった。データベースの分野においてもこのような背景からいわゆるマルチメディアデータ、すなわち画像、音声、映像などをデータベースで管理し、また音や絵などのデータを高度な検索言語を駆使しなくても探索し、かつ表示したいという要求が高まってきている。特に工学、医学を始め、多くのマルチメディアデータを扱う分野からは、膨大なメディアデータの直観的な扱いができる環境が要望されている。しかしながら、従来型の文字・数値を対象としたデータベースではデータベース構造 (スキーマ) とその中に収められている具体値 (インスタンス) をテーブル等の形で簡単に表現できたのに対して、マルチメディアデータベースでは、データの形態そのものが多様であることに加え、メディアデータ間の関連付けが複雑でかつデータ量も非常に多く、データベース全体をユーザが直観的に把握しにくいという問題がある。特に音声や映像などにおいては、データ自身に時間軸が含まれ、再生時の時間同期も必要である。そこで本研究では、マルチメディアデータベースの中身を非専門家ユーザが直観的に把握できるようにするために、スキーマ図に基づく視覚的なナビゲーションインターフェースを提供するための理論的な基盤の探求とプロトタイプを試作を行った。

② 学術的な特色・独創的な点と意義

本研究ではデータベースを作成するときに必ず作られるスキーマダイアグラムを用い、インスタンスを直接指示したり、ある条件を満たすインスタンスのみを表示したり、音や映像に関してはインスタンス群をまとめてプレーヤーに渡すことで再生したりすることにより、コンピュータの非専門家ユーザに対しても直観的で分かりやすいマルチメディアデータベース閲覧のユーザインタフェースが提供するところに有用性と独創性がある。今までデータベースのブラウザとしては、データのひとまとまり (例えばオブジェクト単位) にウィンドウが割当てられ、それらのウィンドウをただ順番に開いて見て行くだけ、といった程度のものしか考えられていなかった。本研究ではこれらとは根本的に異なり、ユーザにはデータベースの構造を率直にあらわしているスキーマ・ダイアグラムを示し、ユーザは自分の関心のあるタイプを選択し、その上でインスタンス間のつながりをもとにインスタンス群を (ビットマップや映像などのメディアデータを含め) 集合としてまとめて表示できるようにしている。

スキーマダイアグラムを用いたデータナビゲーションは、現在までまったく知られていないが、データベース構造が直接見え、かつ任意の部分を「切り取って」表示できる点で非常に自

由度の高い表示と操作が行えるという点に大きな特色がある。また、本研究の成果として生まれるプロトタイプはパソコン等で動くことを想定しており、現在必ずしも一般化しているとは言えないマルチメディアデータベース管理システムの普及に対しても大きな推進力となると思われる。

研究組織

研究代表者 有澤 博 横浜国立大学大学院環境情報研究院 教授
研究分担者 田村直良 横浜国立大学大学院環境情報研究院 教授
研究分担者 富井尚志 横浜国立大学大学院環境情報研究院 講師

交付決定額（配分額）

（金額単位：千円）

	直接経費	間接経費	合 計
平成 12 年度	2 3 0	0	2 3 0
平成 13 年度	1 3 0	0	1 3 0
総 計	3 6 0	0	3 6 0

研究実績の概要

本研究では申請者らが提案し、開発と改良を行ってきた拡張ERモデルを用いマルチメディアデータベースの内容を直観的に表示するシステムを開発した。拡張ERモデルは、意味データモデルとオブジェクト指向データモデルの長所を併せ持つダイアグラム表現ができ、かつ、データベースに蓄積する最小の要素（エンティティ）やメディアデータ（例えば、静止画を表すpixel map）をそれぞれ独立して画面内に表現できる点で、本研究の目的に合致している。本研究では、データベースの構造（スキーマ）とそこに含まれる個々のデータ（インスタンス）を、相互のつながりも含めて表示できるダイアグラム記法を詳細レベルまで考察を行い、仕様を確定した。本研究のスキーマダイアグラムは、データベース全体のスキーマから、ユーザーが自分の関心のあるデータ集合を選び、スパンニングツリーの形のダイアグラムとして表示するものである。このようなデータ表現の枠組みをグラフィックワークステーションを用いて実装した。ついで、その上にナビゲーションと検索・閲覧の仕組みを作りこんだ。ナビゲーションに関してはダイアグラム全体からの興味のある部分を付け加えたり、逆に興味のない部分の「中抜き」を理論的矛盾がないように行う方法（ビューの作成）を考案した。このため、ユーザーが考えるスキーマに対して元データベースから矛盾なく大量のインスタンス群を変換するアルゴリズムを考案し、実装した。また、大量のデータ、特にインスタンスを限られた画面内で有効に視覚化する手法として、必要なデータのみを選択して動画によって視覚化する機構を作成した。

以上からマルチメディアデータベースのもつスキーマおよびインスタンス情報を直感的・視覚

的にナビゲーションするシステムを確立することができた。

研究発表

1. 今井, 富井, 有澤: “Mediator 概念に基づく身体形状・動作のモデル化とデータベース設計”, 情報処理学会論文誌, Vol.41, No.SIG1(TOD5), pp.100-108, 2000.2
2. 坂木, 今井, 富井, 有澤: “多視点映像を用いた時空間オブジェクトの解析とモデリング”, 電子情報通信学会, 第 11 回データ工学ワークショップ(DEWS2000)論文集, 2000.3
3. 佐藤, 永野, 富井, 有澤, 酒井: “人体骨格 CG モデルを用いた人体動作データベースの構築”, 電子情報通信学会技術研究報告, Vol.100, No.31, pp.73-80, 2000.5
4. 有澤: “時空間データベース”, bit, 共立出版, 2000.7, Vol.32, No.7
5. 有澤: “CG 骨格モデルを用いた人体モデリング”, IMS 成果報告会, 2000.7.11-12
6. 有澤: “人間-機械協調システム最適設計のための人体シミュレータの構築に関する研究”, IMS 成果報告会, 2000.7.11-12
7. 中家, 徐, 富井, 有澤: “次世代放送向け多チャンネル映像のデータベース化”, 情報処理学会研究報告, Vol.2000, No.69, pp.207-214, 2000.7
8. H. Arisawa, S.Imai: “Mediator-based modeling of factory workers and their motions in the framework of Info-Ergonomics”, Human Friendly Mechatronics – Selected Papers of the International Conference on Machine Automation ICMA2000, pp.395-400, 2000.9
9. 佐藤, 永野, 富井, 有澤, 酒井: “骨格モデルを用いた人体動作データベースの構築”, 情報処理学会論文誌, Vol.42, No.SIG1(TOD8), pp.92-102, Jan. 2001
10. 富井, 坂木, 四十物, 鬼東, 有澤: “時空間データベースコンセプトに基づくサッカー映像のモデリングと解析手法”, 画像情報メディア学会技術報告, Vol.25, No.24, pp.27-32
11. 有澤: “映像からの 3次元解析と時空間モデリング”, 画像情報メディア学会技術報告, Vol.25, No.24, pp.33-43
12. 佐藤, 富井, 有澤: “空間マウスを用いた人体骨格モデルの動作定義支援システムの構築”, 第 12 回データ工学ワークショップ(伊豆熱川 DEWS2001)論文集, 2001.3.
13. 川口, 青木, 富井, 有澤: “高速ネットワークグラフィックス表示装置を用いた遠隔仮想水族館の実現手法”, 第 12 回データ工学ワークショップ(伊豆熱川 DEWS2001)論文集, 2001.3.
14. 松田, 越前谷, 富井, 有澤: “XML を用いたシーンの時空間オブジェクト情報の記述”, 第 12 回データ工学ワークショップ(伊豆熱川 DEWS2001)論文集, 2001.3.
15. 家富, 森寄, 富井, 有澤: “マルチメディアデータブラウジングのための視覚的データベースインターフェース AVIS の実装”, 電子情報通信学会論文誌 D-I, 6月号(システム開発論文特集), 2001.
16. 四十物, 坂木, 鬼東, 富井, 有澤: “時空間 MMDB におけるサッカーの戦術記述と ad-hoc 検索”, 情報処理学会研究報告, Vol.2001, No.70, pp.71-77, 2001.7

17. 越前谷, 松田, 富井, 有澤: “XML を用いたマルチメディア構造化文書 STOIC の意味表現能力とその評価”, 情報処理学会研究報告, Vol.2001, No.70, pp.79-86, 2001.7
18. 岸辺, 家富, 富井, 有澤: “スキーマダイアグラムを用いたマルチメディアデータベース検索手法”, 情報処理学会研究報告, Vol.2001, No.70, pp.87-94, 2001.7
19. S.Imai, T.Tomii, H.Arisawa: “HUMAN BODY/MOTION MODELING BASED ON THE “MEDIATOR” CONCEPT”, 4th International Scientific Conference on Prevention of Work-Related Musculoskeletal Disorders(PREMUS 2001), pp.127, Amsterdam, the Netherlands, Oct. 2001
20. H.Arisawa, T.Sato: “HUMAN-BODY MOTION SIMULATION USING BONE- BASED HUMAN MODEL AND POSTURE DATABASE”, 4th International Scientific Conference on Prevention of Work-Related Musculoskeletal Disorders(PREMUS 2001), pp.128, Amsterdam, the Netherlands, Oct. 2001
21. S.Imai, T.Tomii, H.Arisawa: “Motion Modeling and Simulation of Human Workers based on Info-Ergonomics Concept”, 20th International Conference on Conceptual Modeling ER2001, Proc. of the International Workshops on Modeling of Human/ Organizational/ Social Aspects of Manufacturing Activities(HUMACS 2001), Yokohama, JAPAN, Nov. 2001
22. H.Arisawa, T.Sato, T.Tomii: “Human-Body Motion Simulation Using Bone-Based Human Model and Construction of Motion Database”, 20th International Conference on Conceptual Modeling ER2001, Proc. of the International Workshops on Modeling of Human/ Organizational/ Social Aspects of Manufacturing Activities(HUMACS 2001), Yokohama, JAPAN, Nov. 2001
23. 坂木, 四十物, 鬼束, 富井, 有澤: “複数オブジェクト共存空間上の“局面”モデリングとサッカー戦術への適用”, データベースと Web 情報システムに関するシンポジウム(DBWseb2001)論文集, pp.1-8, Kyoto, Japan, December 5-7, 2001.
24. 田中, 中家, 松田, 家富, 富井, 有澤: “大規模マルチメディアを対象としたスクラップアンドビルド型 DBMS の構築”, 第 13 回データ工学ワークショップ(DEWS2002)論文集, 2002.3.
25. 浅野, 田村: “文章セグメントの単一化による多文書自動要約”, 言語処理学会第 8 回年次大会, 2002 年 3 月
26. 田中, 田村: “負データが極端に少ない訓練事例を用いる OCR 誤認識検出”, 言語処理学会第 8 回年次大会, 2002 年 3 月
27. 大村, 田村: “主題構造解析による新聞記事からの気象情報の抽出と応用”, 言語処理学会第 8 回年次大会, 2002 年 3 月
28. 森野, 後藤, 田村: “文脈自由文法に基づく点字楽譜の自動解析の検討”, 電子情報通信学会論文誌, D-I, Vol.J85-D-I, No.5, pp.1-9, 2002 年 5 月

出版物

Hiroshi Arisawa, ADVANCES IN VISUAL INFORMATION MANAGEMENT
KLUWER ACADEMIC PUBLISHERS (2000.5)

研究成果による工業所有権の出願・所得状況

なし

Mediator 概念に基づく身体形状・動作のモデル化とデータベース設計

今井 さやか† 富井 尚志† 有澤 博†

本稿では、人体形状・動作データベースに蓄積・検索するための人体形状・構造と、動作のモデリングの手法について提案する。最少のデータ量で人間の動作の記述を行うために、我々は、運動 Mediator を導入した。Mediator は、ある人間のある動作における標準的な形状・構造や、動作を表現できる概念である。Mediator を利用することで、実世界の人間の動作をデータベースに蓄積・検索し、CG アニメーションとして再構築できる。さらに、我々は Mediator が元の動作を再現できることを評価した。

Human Body/Motion Modeling and Database Design based on the Mediator Concept

SAYAKA IMAI,† TAKASHI TOMII† and HIROSHI ARISAWA†

This paper proposes a modeling methodology of human body and the motions for database storing and retrieval. In order to describe human motions with small amount of data, we introduced "Mediator" which can represent typical figures and structure in a human body and postures in a specific motion. By using Mediator, storing, retrieving and replaying the original motions are possible. We evaluate the expressive ability by comparing a Motion Mediator with the original posture sequence.

1. はじめに

最近の CG 技術の発達はめざましいものがあり、パソコンクラスのマシンに CG ボードを搭載することでも、少し前の高価なグラフィックワークステーション並みの性能を出せるものが出てきた。今後、このような技術の大衆化を背景に、2次元、3次元のリアルタイム CG 描画をベースにしたコンテンツの普及が急速に進むであろう。ところで、コンテンツの種類や量の飛躍的な拡大に伴って必ず生まれてくる問題は、それをいかに整理体系化して、「必要なものを必要なときに取り出せる」ようにするか、すなわち、データベース化の方法論である。

我々は数年に渡り、マルチメディアデータベースの究極的な姿として、3次元的・時空間的な「事象」をそのままデータベースに蓄積し、かつさまざまな視点から検索するための方法論を研究してきた。その成果として実世界の事象すなわち、さまざまな物体についての形状や動きの情報を「あるがまま」「見たまま」に取り込み、蓄積し、さまざまな検索に応えることので

きるリアルワールドデータベースが提案されている¹⁾。

しかし時空間的な「事象」情報を体系化するためには、多くの問題を解決しなければならない。まず我々は実世界の3次元空間データを厳密に取得する方法を持っているわけではない。静止した物体の表面形状やテクスチャなら3次元スキャナでかなり精密に取り込めるが、運動している人体などの動物体に対してはスキャナを使用することはできない。動物体にマーカを取り付けて、モーションキャプチャを行うことはできるが、それは限られた点のみの位置データであり、応用範囲が制限されるだけでなく、対象物体のグラフィック表現との対応が正確に取れていないと、動きの再現はできない。これに対してステレオ(多視点)ビデオ映像は、物体形状や位置などの3次元的な情報を抽出しうる潜在性を持っているが、物体表面のすべての点においてこの解析を行うことは不可能である。我々は、このようにいわば不完全な実世界データから3次元空間を動く物体の形状や動きを抽出し、さらにそれを支える構造的情報をも考慮してデータベースの蓄積、検索に耐えうる解析(モデル化)を行う必要がある。

また、3次元空間内を動く、構造を持ったオブジェクトについてのデータベース化を念頭においた一般的な表現手法(スキーマ化)は今までほとんど提案されて

† 横浜国立大学工学部電子情報工学科
Division of Electrical and Computer Engineering,
Faculty of Engineering, Yokohama National University

おらず、したがって人間のような多くの自由度を持ったオブジェクトの動作に関する検索システムも、個別アプリケーションを除いては、存在していないのが現状である。

そこで本稿では、人間が何かある意図を持って行う動作(例えば、ある工場作業など)の身体形状と動きについての3次元・時間的な情報を大量に蓄積したデータベースの構築を念頭におき、それを現在利用可能な技術を用いて、実現するための一つの手法を提案する。このために解決すべき課題として次のような段階を踏んで考える。

- (1) 3次元スキャナやステレオ映像など、種々のメディア情報から人体の形状・動作の3次元時空間情報を解析、取得する手法
 - (2) (1)で得られたデータを、より一般化するため、少ない情報量で形状や動きを記述する手法
 - (3) (2)で得られたデータを統合的に1つのデータベーススキーマで表現する手法。
 - (4) 必要な時空間検索を記述実現する手法
- (1)について我々はすでにステレオビデオ映像を用いたノンマーカ型のモーションキャプチャを提案し、特徴となる部位の3次元位置情報獲得の可能性を評価した²⁾。(2)は、データベースに蓄積、検索するための時間的、空間的なインデックスの作成につながる技術であり、重要である。この課題を解決するため、形状情報の一般化手法として、*Mediator*(仲介者)という概念がすでに提案されている³⁾。本稿ではこれを拡張して、形状・構造および動きについての*Mediator*概念の提案、評価を行う。さらに、(3)についても基本的な考え方を述べ、データベースとして実現可能であることを示す。(4)については今後に残された研究課題とする。以下本稿では、2章で人間のモデリングについての関連研究と比較をした上で、*Mediator*の導入の必要性について述べる。また、3章で形状・構造と動作の*Mediator*の定義を述べ、その利点、人間のモデル化手法について述べる。さらに、4章で*Mediator*の妥当性の評価を行う。5章では*Mediator*のデータベース表現と考察を行い、6章でまとめを述べる。

2. 身体形状・動作のモデリング手法

2.1 身体動作の表現とモデリング

人間の動作の表現方法としては、人体の腕・足など、形状は変化しないで動きを実現する最小単位となる部品群を考え、それらの形状をまず記述しておき、次いで、時間の経過とともにその部品たちの位置関係がどのように変わっていくかを記述するということが

よく行われている。このような人体モデルを仮定して、人間の動作の解析・モデル化を行った例として、文献^{4), 5), 6)}があげられるが、人体の形状と構造、動作が複雑に絡み合ったモデルを用い、また自然な身体動作の再現に主眼がおかれているため、論理的、体系的に整理された身体動作モデルとはいえない。

また、人間の「ふるまい」すなわち、「ある個人のある意図を持った行動」のモデリングを行った例として、文献⁷⁾があげられる。ここでは、歩行を撮影した映像から被写体のシルエットをとりだし、その重心の動きの特徴から個人を特定するというを行っている。しかし、本来3次元的な動作を2次元的な動作に置き換えて特徴量を抽出した限定的な手法であり、また、個人別の歩行の特徴から個人を識別することを目的としているため、同じ人物の「腕を前後に大きく振っている歩行」や「膝を高くあげた行進」など動作の個別の違いをモデリングすることは行われていない。

一方で人間の身体動作記述をモデル化した例として文献⁸⁾があげられる。ここでは身体動作を定義する手法として、一般的なCGアニメーションを作成する際によく利用されるキーフレーム法の姿勢と、ダンスの振付けが図で描かれている舞踊譜との姿勢とを対応づけることによって、何の動作なのかを記述している。しかし、入力したデータは人体の各部位ごと、動作の一まとまりごとにカプセル化され、データベースに蓄積する人体の動作の差異などの再現、人物を変えた動作の再現、動作を変えた人物の再現などは難しい。

「歩行」動作の生体的解析に焦点を絞ると、特にバイオメカニズムなどでさまざまな解析が行われている。例えば文献^{9), 10)}に示されている研究では、歩行の際にはどのような生体的な特徴が現れるかを検証している。しかしこれらもデータベースに蓄積するような一般化された情報表現の形にはなっていない。

以上のように、人間の構造や動作をモデル化したり、グラフィックとして再現した研究は様々に行われているが、個人別の形状や運動を表現できる一般的な人間モデルの提案はされていない。

2.2 身体形状・動作データベースの目的

以下本稿では身体形状・動作のデータベース化を提案するが、その目的は以下の通りである。

- (1) 「ある人間によってある目的を持って行われた動作」を一まとまりの情報として、何千、何万と蓄積すること
- (2) 人を変えたり、動作を変えたり、同じ人・動作でも何回も行えば違いが出るので、同一性(例えば同じ人によって行われている)を考慮しつ

- つ別個のインスタンスとして表現できること
- (3) 蓄積された情報に対して集合的な立場からの検索と集約演算 (例えばある動作を 100 人にさせた時の前屈角度の平均値を求めるなど) を行えること
- (4) 形状や、動作などの「共通性」を基にした時空間検索が行えること (例えば作業者の目と、手に持っている刃物の距離が 3cm 以内に近付いた時点の画像を検索するなど)

これらの目的を満たすため、特に (3)(4) のような検索を行えるようにするためには、身体形状・動作の特徴を統一的に規格化する必要がある。また、(2) の観点から、タイプ-インスタンスの対応づけが要求され、さらに (1) の観点からは、形状や動作を少ない情報量で表現し、色々な視点でのインデックスとして利用させる必要がある。

我々は以前の研究で、人体部品の形状については、すべての形状データに共通な基本的情報 (基本モデル: 例えば人間の頭は円筒型で近似) を定義し、それを変形させて、実世界データ (例えば 3 次元スキャナを用いて得られた表面形状) に近い形にした形状記述データ (Mediator) を提案し、時空間検索に利用できることを示した³⁾。しかし、人間の動作 (作業) をデータベース化する際には人体の各部品の形状だけでなく、構造や運動を表現できなければならない。そこで以下では Mediator 概念を拡張して、部品のつながり、自由度や可動域はどうなっているか、および、どういう運動を行っているかまでを表現することを試みた*。

3. Mediator 導入の意義とモデリング

3.1 Mediator の意義

身体形状・動作データベースを構築するには、実世界から 3 次元スキャナやステレオビデオカメラなどを用いて取得したデータをもとに、人体形状を 3 次元グラフィックモデル、動作をそのモデルの動きとして蓄積し、またそのデータに基づいて時空間検索が行えなければならない。ところが、スキャナによって得られる表面形状 (ポリゴン群) やステレオ映像解析によって得られる多数の特徴点位置の時系列情報は、一般には非常に膨大な量となり、データベース検索にも実時間表示にも耐えられない。この問題を解決し、2.2 節 (1)~(4) に挙げたデータベースの目的を果たすため、人体形状・構造と動作 (運動) それぞれについて

Mediator を導入する。それぞれ次のように定義される。この内、形状・構造 Mediator については、以前の研究³⁾で人体の個々の部品形状に対してのみ考えられていた Mediator 概念を人体「構造」にまで拡張したものであり、動作 Mediator については、ここで新たに提案を行うものである。

- 形状・構造 Mediator ある人間の体を (頭、肩、上腕などの) 時間的に形状変化しない「部品」と「関節」から成ると仮定し、個々の部品に対してはポリゴンから成るその人間固有の表面形状とテクスチャを、関節に対しては関与する部品とその人間固有の自由度 (Degree of Freedom) および可動域 (Range of Motion) データを割り当てることによって得られた 3 次元グラフィックモデル**。
 - 動作 Mediator ある人間がある時行ったある「動作」について、その途中に必ず現れる特徴的な中間姿勢 (各動作ごとに決められた個数存在する) を開始点からその姿勢に至った相対時刻、および、その時点の形状・構造 Mediator における関節角度によって記述したもの。すなわち、その人間のその動作固有の決められた特定中間姿勢列。
- 形状・構造 Mediator は、個人別に作成され、部品形状や関節の可動域などについては個人の違いを反映する。一方、動作 Mediator は「ある人の行ったある動作」ごとに作成され、それぞれ固有の身体の動きの違いを表現する。これらによって、2.2 節 (1)(2) に挙げた固有の形状・構造・動作をデータベースに表現することが可能となる。さらに Mediator を考える際に重要なことは、必要最低限のデータ量でそのデータの特徴が表現できることと、一般的な情報または Mediator 間に共通な意味情報が示され、データベース上で表現されることである。これらによって、2.2 節 (3)(4) で示したような高度の検索がデータベース上で実現可能になる。そこで次節では Mediator の一般化である「共通モデル」と「共通モデル」に基づく Mediator の作成手法を提案する。

3.2 共通モデルからの Mediator の作成

共通モデルとは、形状・構造、動作それぞれに関してその本質を共通的に表し、規格化するための情報である。次のように定義する。

- 形状・構造共通モデル 人体の構成部品を最も単

* 本稿では以下、「基本モデル」の代わりに「共通モデル」という用語を用いることにする。

** なお、関節に関連する個別の人体の動きを特徴づけるには、医学的に見ると自由度・可動域の規定のみでは不十分で、相互関連 (例えば、上腕をある角度まであげると肩甲骨が回転をはじめなど) についてのパラメータも Mediator 情報として必要であるが、このことについては本論文では立ち入らない。

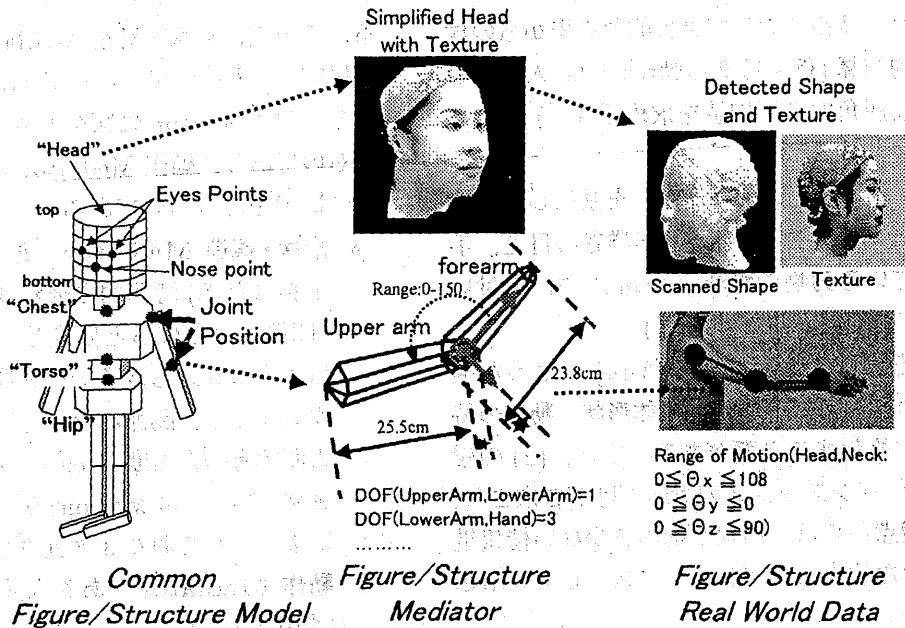


図1 形状・構造の場合の共通モデル-Mediator-実データ
Fig. 1 Common Model, Mediator and Real World Data about a figure and structure

純化されたグラフィックオブジェクトで表し、(例えば、頭を円筒形、肩を円錐台、胸を円柱など)その表面の特定位置に(目、鼻、口など)我々がその物体に対して持つ「意味」情報を貼り付けたもの。さらに人体全体の構造を記述するため、各部品ごとに「関節位置」を定め、部品相互間の連結情報を定義したもの。

- 動作共通モデル 人間の動作に共通して現れ、かつその動作を行う際には必ず通る中間姿勢を形状・構造共通モデルにおける人体部品の相互位置関係を用いて記述したもの。

共通モデルは形状・構造および動作について、その特徴を記述するための最小限の「項目」を示しているとも言え、通常のデータモデルで用いられるのとは違った意味だが一種の「タイプ」であるとも言える。これに対して Mediator は具体的な個人別、動作別の事例に対して共通モデルを変形し、値を付与して得られるもので、共通モデルのインスタンスに相当する。

次に、形状・構造基本モデルから Mediator を作成する手法を考えよう。形状については、3次元スキャナから得られた詳細データに対して、基本モデルを「変形」し「フィッティング」することにより行えることがすでに示されている³⁾。関節の自由度および可動域については、対象者に簡単な動作をさせることにより医学的な観点から測定できることが知られている¹¹⁾。以上二者から Mediator は容易に構成できる。図1に形状・構造の共通モデル-Mediator-実データの関連

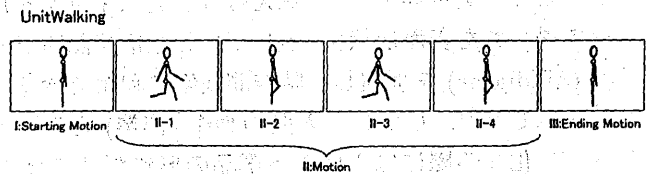


図2 動作共通モデル
Fig. 2 List of Common Motion.

を示す。

一方、共通動作モデルから動作 Mediator を作る場合、共通モデル側にどのような中間姿勢を選ぶかが重要である。これは現段階では人間が動作ごとに設計するしかない。例えば一般的に人間の「歩行」動作は次にあげるような特徴的な中間姿勢列として表現することができる。

UnitWalking:

I:Starting Motion— 両足のつまさきの位置が床上にあり、体幹と水平に位置している姿勢

II:Motion

- II-1:右足の膝が一番高い位置にある瞬間の姿勢
- II-2:右足を一步踏みだし、接地した瞬間の姿勢
- II-3:左足の膝が一番高い位置にある瞬間の姿勢
- II-4:左足を一步踏みだし、接地した瞬間の姿勢
- II-5:(は次のII-1、以下何回か繰り返し)

III:Ending Motion-次の動作へ移る前の姿勢 (Starting Motionと同じ姿勢) 図2に「歩行」の動作共通モデルの概念図を示す。

一般に、ステレオビデオカメラなどを用いて実世界

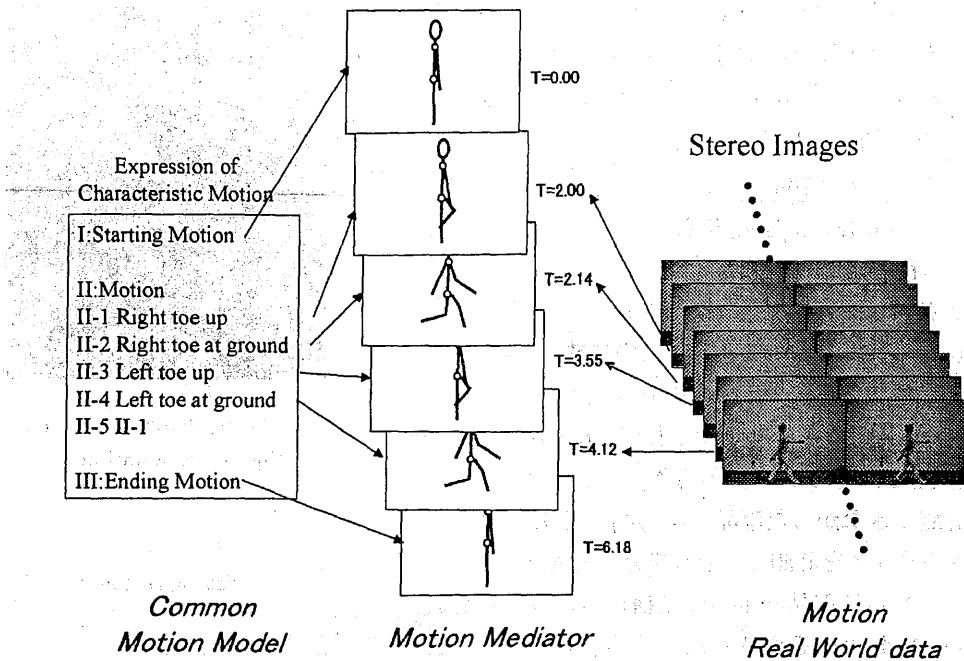


図3 運動の場合の共通モデル- Mediator -実データ
 Fig. 3 Common Model, Mediator and Real World Data about a motion

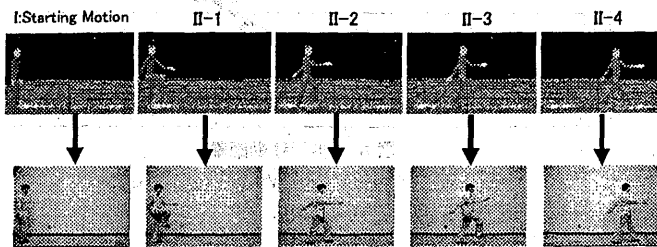


図4 動作 Mediator とその時点の実世界の映像フレーム
 Fig. 4 Motion Mediator and its Real World Data

から取り込まれた映像データから人間の姿勢(各部位の3次元的位置)を逐一精密に抽出することは非常に困難である。しかし、動作共通モデルで指定された特定の姿勢に限れば、あらかじめ決められた少数の特徴点の3次元位置を対応点検出により求めるだけでよく、そこからの姿勢(関節角度)の解析は比較的容易である。したがって、共通モデルから動作 Mediator を作成することが可能となる。図3に運動の場合の共通モデル- Mediator -実データを示す。図3に示した姿勢系列は、動作の数値データを姿勢の図として便宜的に表示したものである。

4. 動作 Mediator の評価

4.1 動作 Mediator の評価手法

今まで述べたように Mediator は非常に少ないデータ量で、形状や動作の個別の違いを統一的な枠組の中で表現することができると期待される。そこで本節で

は3.1で述べた動作 Mediator が実際の動作の特徴をつかんでいるか、すなわち Mediator を用いて元の動作を再現できるかについての評価を次の手順で行った。(人体部品の形状の評価については、文献³⁾で行われているので、本稿では省略した。構造については、人体の部品展開の手法を一般的な人間CGモデル¹²⁾に合わせてあるため、それ以上の評価は行わなかった。)

- (1) 被写体の形状・構造 Mediator を用いて、計算機上に人間グラフィックモデル(人間モックアップ)を作成する。
- (2) ステレオ映像から共通動作モデルの中間姿勢に該当する被写体の関節角度の計測を自動的に行い、動作 Mediator を作成する。
- (3) 動作 Mediator から Mediator 間の時点の姿勢を補間して3次元アニメーションを作成する。
- (4) 上記とは別に、ステレオ映像の左右フレーム列に対し、人手を介して対応点検出を行って精密に動作を抽出し、被写体の形状・構造 Mediator を用いた計算機上の人間モックアップにマッピングして3次元アニメーションを作成する。
- (5) (3)と(4)で作成された2種類の3次元アニメーションの人間モックアップの動作(人体の空間位置や関節角度変化など)を比較することにより、動作 Mediator からの人間の動作の再現性の評価を行う。

4.2 評価のためのサンプルデータ

本稿における評価用データに使用した身体形状・動

作の実世界データのサンプルは以下の通りである。

- ・被験者 … 20 代の男女 8 名 (被験者 1~8)
- ・動作 … 歩行動作 (歩行距離:約 2m,
足踏み (4 歩), 歩行 (6 歩))
- ・撮影機材…VGA カメラ (640 × 480) 2 台
- ・映像フレーム … 60 frame/sec(左右とも)
- ・動作時間 … 歩行 10 歩分の平均 6.1 sec
- ・撮影フレーム数 … 被験者平均 361 frame

4.3 サンプルデータからの Mediator の作成

まずはじめに被写体の部品形状の Mediator については, 3 次元スキャナを用いて被写体の頭と胴体部分の精密な形状を計測し, そのデータから作成した. 3 次元スキャナで計測できなかった部品 (腕, 足) などについては, 被写体のサイズを計測し, それを元に CAD を用いて作成した. さらに部品同士の連結構造を付与し, 形状・構造の Mediator を作成した. この結果できた人間モデルは, 部品数 25 点, 関節数 24 点, 50 自由度であり, 腰部品をルート部品とした階層構造をなしている.

次に, 動作 Mediator については, 2 台の VGA カメラで撮影されたステレオ映像に対して, 空間位置算出プログラム CHASER²⁾を用いて, 図 2 の共通動作モデルに該当する中間姿勢を算出した. 各姿勢は左右の映像フレームに対して, 両足のつまさき, 膝, 腰の位置, 両肩, 両手先の位置の 9 点の部位の位置を検出することで自動的に算出した. しかし, CHASER の計測には精度的に問題があるので今回の実験では, CHASER の計測結果にスムージングを行って, なめらかな動きをするよう修正を加えた各関節角度データを Mediator データとし, 動作 Mediator の姿勢時系列を作成した.

図 4 に動作 Mediator に対応する時点の実世界データ映像と, その時点に対応する動作 Mediator のスナップショットを示す.

4.4 Mediator からの再現アニメーションの作成と再現性の評価

4.1 節の (3) に相当する手法として, グラフィックシミュレーションソフト ENVISION (米国 Deneb 社製) の動作補間機能を使って, 動作 Mediator の中間姿勢から, 正解データの時間間隔と等しくなるよう補間を行って 3 次元アニメーションを作成した. そして, 作成したアニメーションから求めた各時点における関節角度や体の重心の位置と, 手法 (4) において作成した実世界データにおける関節角度や体の重心の位置との比較, 実データと動作 Mediator のデータ量の比較を行った.

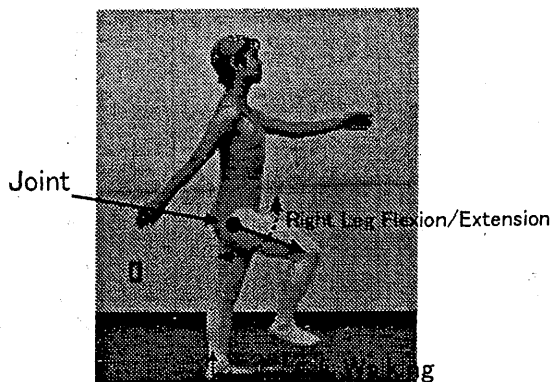


図 5 グラフで示した人体の関節部位の角度と歩行移動距離
Fig. 5 Evaluation Joints

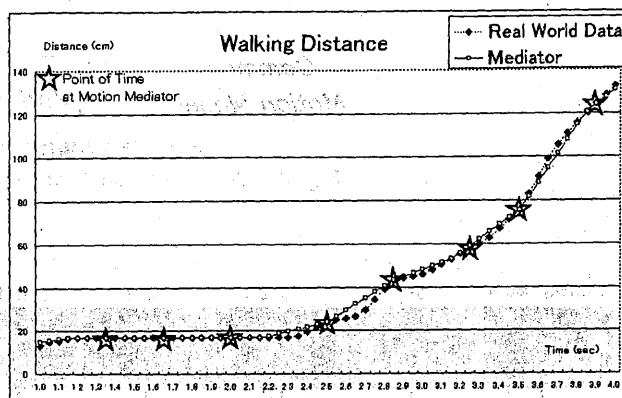


図 6 歩行移動距離
Fig. 6 Evaluation Walking Distance

図 6 に示したグラフは被験者 1 の動作データに対して, 図 5 に示すような, 腰の空間位置を床上に射影した位置の, 歩行開始からの前身方向への移動距離 (1.00~4.00 秒の区間, 4 歩分) の数値データ (図 5 の Walking) である. 図 7 に掲載したグラフの値は, 図 5 の右足股関節前後の振幅角度 (Right Leg Flexion/Extension) の数値データである. グラフ全体では, 動作 (各位置, 角度変化) の特徴をとらえていることがわかる. しかし各時点の角度データを細かく比較すると, 被験者 1 では移動距離の最大誤差 5.6cm, 右股関節角度最大誤差 23.1 度となった. これを被験者 8 人の平均で見ると, 歩行移動距離では被験者 8 人の平均最大誤差約 8.0cm, 右股関節前後の振幅角度では平均最大 25.4 度の誤差が生じた. ただし, 被験者 8 名の誤差が最大になっている時点を見てみると, 動作 Mediator の中間姿勢として選択されていない時点であった. これに対して, 選択されている時点のデータは, 最大で移動距離 3cm, 右股関節角度 10 度に収まっている. したがって非常に大きな数値誤差が生じたのは, 姿勢の補間に使用した, プログラムによるものと考えられる. この

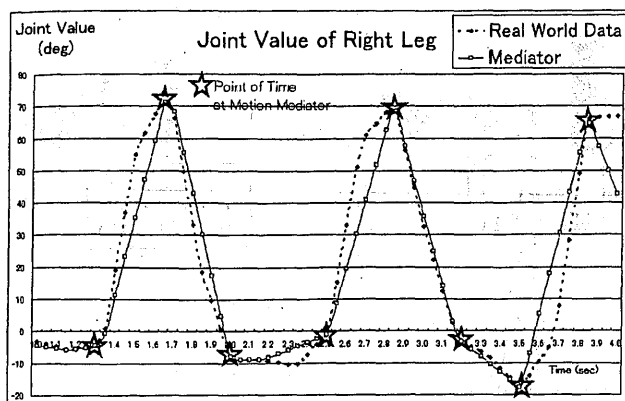


図7 歩行中の右足関節の角度(前後振幅)
Fig. 7 Evaluation Value of Right Leg Joint

補間プログラムでは、片足を接地して、次に足をあげるまでの「ため」を無視しているの、この点を改良すれば、より良い結果が得られるものと期待される。

また、運動を記述するために必要なデータの量は、実データが、動作内容 歩行 10 歩に対して、人体の関節数 24 関節(人体形状・構造 Mediator の関節数と同数)、平均動作解析時点数 361 時点(平均 361 映像フレーム)であるのに対し、動作 Mediator は、動作内容 歩行 10 歩に対して、人体の計測した関節数 9 関節、平均動作 Mediator の時点数 21 時点となり、運動に関しても少ない時系列データでも運動の特徴を示すことが可能であり、このことはデータベースの蓄積にも検索にも有効に働く。

さらに、それぞれの 3 次元アニメーションを比較したところ、動作は極めて良く近似されていることが確認された。したがって、動作 Mediator より作成した人間モックアップの動作の 3 次元アニメーションは被験者の運動の特徴をとらえており、少ない情報(動作 Mediator)でも、個別の動作情報を再現することが可能であることが示された。

5. 身体形状・動作データベースの実現と考察

5.1 Mediator に基づく身体形状・動作データベースの実現

データベースに人間の動作を蓄積するためには、今まで述べたような共通モデル Mediator-実世界モデルの各レベルに属するデータ(データの種別は文字・数値だけでなくステレオ映像や、ポリゴン・テクスチャなど様々なものである)を、その参照関係とともに蓄積できなければならない。さらに、データベースに蓄積されているデータを自由に組合せたり、目的に応じて思い通りの検索を実現したい。本稿で提案する Mediator の意義は、データ表現のためのデータモデルには

特に依存しない。しかし、データの参照関係が単純なメカニズムで定義され、かつ蓄積するデータの単位が小さい方が望ましい。そこでここでは身体形状・動作を蓄積するため、関数型データモデル AIS(Associative Information Structure)¹³⁾を用いて記述した。

身体形状・動作のデータベーススキーマとインスタンスを図 8 に示す。AIS モデルは、ER モデルを基にした entity ベースのデータモデルであり、実世界の事象・事象の代理物である entity と、それらの間の関連づけ(association)によって情報を表現している。つまり、ステレオ映像や、ポリゴン形状、テクスチャ、人体構造の記述、運動データなどデータベースに蓄積されるデータはすべて、データベース上の entity として表現される。図 8 において、矩形が entity タイプを表し、小丸はそのタイプに属する entity を示す。また、entity タイプ間のパス(実線)によって対応付けが示されている。タイプ同士の対応づけとして、菱形は実線で結ばれた entity タイプに属するインスタンス集合間の関連(直積)を表し、菱形に矩形がついている時には entity のつながり(対応)もまた entity とみなすことを示す。二重矢印は親のタイプに対する属性タイプを示し、しかも親 entity に対しては属性タイプ中の entity 列が対応することを表している。したがって、蓄積された entity 同士の関係は全て単純な参照関係に分解されている。例えば、実世界のある動作(TIME_DEPENDENT_CAPTURED_DATA)に対応して、運動 Mediator(MOTION_MEDIATOR)が作られ、インスタンスとして蓄積され、その運動 Mediator インスタンスは、中間姿勢(MEDIATOR_INTERMEDIATE_POSTURE)の列から成り、その中間姿勢 1 個は各関節(JOINT_MEDIATOR)ごとの関節角度値(MEDIATOR_JOINT_VALUE)で定義される。このように複雑な関係も、インスタンスレベルの関連として直接表現することができる。

また、このパスをたどることで自由に検索を行うことができる¹⁴⁾¹⁵⁾。例えば「ある運動途中のある時点の関節部位の両端の部品形状」という検索は、ある運動途中のある時点の運動 Mediator の中間姿勢(MEDIATOR_INTERMEDIATE_POSTURE)-その運動 Mediator の関節角度(MEDIATOR_JOINT_VALUE)-ある個人の関節の Mediator(PART_MEDIATOR)-ある個人の部品形状(PART_MEDIATOR)-ある個人の実際の部品形状(TIME_INDEPENDENT_SCANNED_DATA)とパスをたどることにより検索を行うことができる。

5.2 フレキシブルな動作検索に関する考察

前節に示したスキーマインスタンス図では、身体形

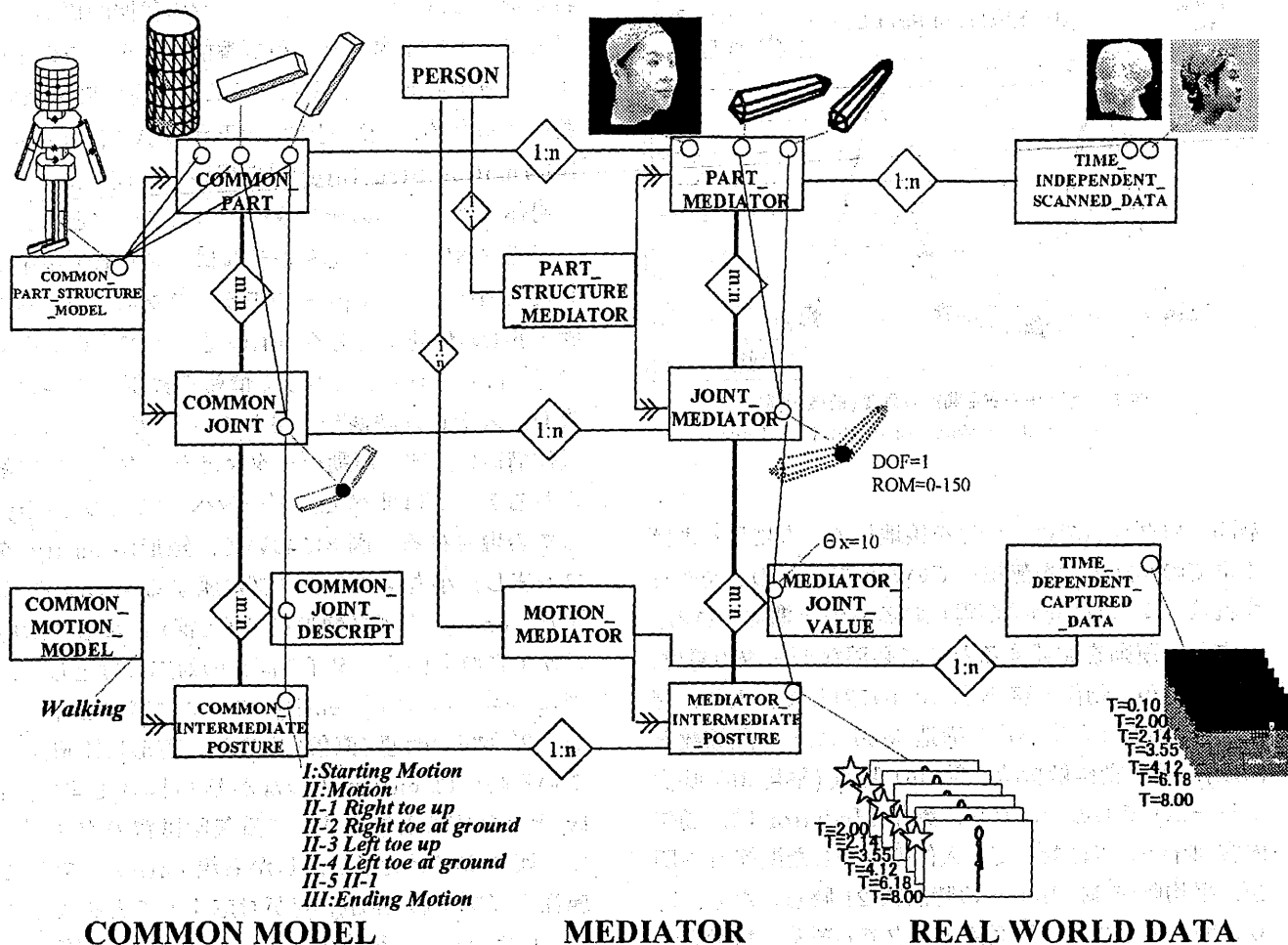


Fig. 8 Database Schema and Instances of Human Figure and Motion

状・構造、動作において、共通モデル、Mediator、実世界データのインスタンスが分かれて蓄積されている。また、形状・構造と動作の Mediator 同士はそれぞれ直交(独立)しているので、検索によって人間と動作を様々に組み合わせることが可能である。例えば、実世界から取り込んだ時の人間とその人間が行った動作の組合せだけでなく、仮想の作業アニメーションを検索結果として作成することも実現できるであろう*。つまり、データベースからの検索によって「Aさんが行った動作」と「Bさんのモックアップ人間」を組み合わせ、「Aさんの動作をBさんにさせた場合の動作」を3次元アニメーションで構築することができる。

また Mediator は、部品形状・構造や動作の個別の特徴を表現しているだけでなく、同じ部品形状・構造、動作の違いも記述していることから、例えば、「歩行」動作すべてのデータに対して、動作 Mediator の姿勢

列データから「行進の一步の距離の平均」「歩行動作の時間平均」「～歩歩いた時の移動距離の最大値」「AさんとBさんの歩幅の差」などといった検索も可能となる。

6. まとめ

本稿では、身体形状・構造と動作の Mediator を提案し、個人の身体形状や3次元空間内の運動のモデル化を行った。さらに、Mediator から作成される人間モックアップと、動作アニメーションが個々の身体形状や動作の特徴を示していることを評価し、本研究で提案された Mediator の妥当性を示した。また、データベースにおけるスキーマ記述を示し、運動の内容の蓄積・検索への可能性を示した。今後の展望としては、本稿で例としてあげた「歩行」以外の運動に対する Mediator の妥当性の評価や、運動 Mediator を利用した検索、それぞれのモデルを作成する要素技術の議論などがあげられる。

謝辞 本研究を行うにあたり、人体の骨のつながり

* INTERMEDIATE_POSUTRE_MEDIATOR と JOINT_MEDIATOR 間の entity の対応を変更すれば良い。

や、人体の動作にともなう関節の骨格の動きなど、横浜市民病院 整形外科部長 酒井直隆 博士に助言を頂いた。謹んで感謝の意を表す。本研究の一部は文部省科学研究費補助金(基盤研究(B)(1) 10558042)および(財)製造科学技術センターIMS国際共同研究プログラム「生産システムにおける人間・機械組織化の研究」による。

参 考 文 献

- 1) 有澤博: “リアルワールドデータベースとその実現技術,” bit, 共立出版, 1996.9.10.11, Vol.28, No.9.10.11.
- 2) 奥井宏昌, 岡本陽介, 今井さやか, 富井尚志, 有澤博: “工場作業データベース設計支援のためのビデオ映像に基づく動作解析”, 情処学報, Vol.98, No.58, 98-DBS-116-58, pp.241-248, 1998.
- 3) 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”, 信学論, D-I, Vol.J82-D-I, No.1, pp.211-222, 1999年1月.
- 4) 筒口拳, 末永康仁, 渡部保日児, 下原勝憲: “3次元シーン内の人物像歩行動作生成システム”, 情処学論, Vol.38, No.4, pp.787-796, 1997年4月.
- 5) 鶴沼宗利, 武内良三: “コンピュータアニメーションにおける感情を伴った人間の歩行動作の生成方法”, 信学論, D-II, Vol.J76-D-II, No.8, pp.1822-1831, 1993年8月.
- 6) 筒口拳, 境野英明, 渡部保日児: “地形適応型歩行動作生成手法を用いた人物の歩行動作アニメーション”, 信学論, D-II, Vol.J77-D-II, No.8, pp.1663-1670, 1994年8月.
- 7) 村瀬洋: “シルエットを用いた歩行動画像からの個人識別”, 信学論, D-II, Vol.J75-D-II, No.6, pp.1096-1098, 1992年6月.
- 8) 平松尚子, 八村広三郎: “身体動作の記述と表示のための実行制御環境の開発”, 情処学論, Vol.40, No.3, pp.939-948, Mar.1999.
- 9) 鈴木良平: “足のバイオメカニクスと歩行解析”, 日整会誌(J.Jpn. Orthop.Assoc.)61, pp.75-86, 1987.
- 10) 多賀巖太郎: “下肢運動のモデル化のシミュレーション”, バイオメカニズム学会誌, Vol.16, No.3, pp.209-214, 1992.
- 11) 中村隆一, 齊藤宏: 基礎運動学 第4版, 医歯薬出版株式会社, 1995.
- 12) Deneb Robotics Corp.: ENVISION Ergonomics Reference Manual, 1999.
- 13) H.Arisawa, T.Tomii, H.Yui, H.Ishikawa: “Data Model and Architecture of Multimedia Database for Engineering Applications,” IEICE Trans. Inf. & Syst., Vol.E78-D No.11, pp.1362-1368, November, 1995.
- 14) 富井尚志, 有澤博: “マルチメディアデータベースにおける映像モデリングと操作言語”, 信学論, Vol.J79-D-II, No.4, 1996.
- 15) T.Tomii, K.Salev, S.Imai, H.Arisawa: “Human Modelling and Design of Spatio-Temporal Queries on 3D Video Database”, Y.Ioannidis and W.Klas eds., *Visual Database Systems 4 (VDB4)*, IFIP, Chapman & Hall Pub. Co., pp.317-336, May 1998.

(平成11年9月21日受付)

(平成11年12月27日採録)

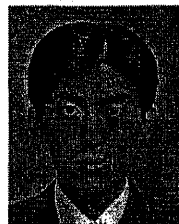
(担当編集委員 片岡 良治)

今井さやか



平成7年横浜国立大学工学部電子情報工学科卒業。平成9年同大学大学院工学研究科博士課程前期修了。現在、同大学大学院工学研究科博士課程後期在学中。主として、データベースにおけるヒューマンモデリングの研究に従事。

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8年同大学大学院工学研究科博士課程前期修了。平成11年同大学大学院工学研究科博士課程後期修了。現在同大学工学部助手。工学博士。主としてマルチメディアデータベース、質問処理などの研究に従事。

有澤 博(正会員)



昭和48年東京大学理学部物理学科卒業。富士通(株)を経て、昭和50年横浜国立大学工学部に奉職。現在同学部電子情報工学科教授。工学博士。1991年アメリカ・オレゴン州立大学計算機学科客員教授。データベース理論、マルチメディアデータベースシステムを研究テーマとしている。

多視点映像を用いた時空間オブジェクトの解析とモデリング

坂木 和則 今井 さやか 富井 尚志 有澤 博
横浜国立大学 工学部 電子情報工学科
〒 240-8501 横浜市保土ヶ谷区常盤台 79-5

E-mail : {sakaki,sayaka,tommy,arisawa}@arislab.dnj.ynu.ac.jp

あらまし

本稿では、時空間データベースに現実世界の情報をできる限り精密に取り込む手法の1つとして、多視点映像を用いて実空間に存在する人間の姿勢や動きの解析する手法を提案する。そこでは、まず初めに人体を関節と剛体から成るパーツ-ジョイントモデルであるとし個別の人間について固有の形状データからモックアップモデルを構成する。その上で、モックアップモデルに仮の姿勢をさせ、多視点映像から抽出された人体の領域とマッチングを行なうことで、その人の姿勢や運動データを抽出した。また、この手法のプロトタイプを並列コンピュータ上で作成することにより、抽出された姿勢が、比較的精度が高く行なえることを示した。

Analyzing and Modeling of Spatio-temporal Objects from Multi Perspective Video Images

Kazunori SAKAKI, Sayaka IMAI, Takashi TOMII, and Hiroshi ARISAWA
Division of Electrical and Computer Engineering, Faculty of Engineering
Yokohama National University
79-5, Tokiwadai, Hodogaya-ku, Yokohama 240-8501 JAPAN

E-mail: {sakaki,sayaka,tommy,arisawa}@arislab.dnj.ynu.ac.jp

Abstract

This paper presents a method of motion and posture analysis of a human in the real world based on multi perspective images, for storing into Spatio-temporal Databases. At first, we consider the body of the human as Parts-Joints model, and construct a mock-up model that reflect individual characteristics of the human body data. Then we offer a matching algorithm between assumed postures of the mock-up and a region of the human in the image. Also, we have implemented a motion analysis system by using parallel computer and evaluated it.

1 はじめに

近年、工業、医学のみならずスポーツなどの分野において、マルチメディアデータの解析や操作が行なわれるようになってきており、そのデータを蓄積、検索できるデータベースの需要が高まっている。特に、オリジナルのマルチメディアデータをそのまま扱うだけでなく、そのデータに含まれる内容をもとにしたコンテンツベースの検索ができることが望まれている。

我々は、現実世界を様々なマルチメディアデータを用いることでできるだけ忠実に記録し、さまざまな検索をかけた結果を現実感豊かに表現できるデータベースとしてリアルワールドデータベースシステムを提案している [1]。その中で、現実世界のオブジェクトを「見たまま」に取り込むには、静止状態であれば立体スキャナを用いることで、その形状や表面のテクスチャを取り込むことができる。また我々は、実世界に現れるオブジェクト—例えば人間の形状や構造、動作(ふるまい)を意味に基づいてスキーマ化し、データベース化する手法を提案した。例えば形状情報に関しては、人体部品の形状データを個別の人の特徴を反映しつつ簡単化されたグラフィックモデル (Mediator) として表現している [2]。また動作に関しても基本となる動作情報をスキーマ化して、個別の動きはそのインスタンスとして運動 Mediator を構成した [3]。しかし、オブジェクトの実時間上での移動、すなわち動きの情報をマルチメディアデータから抽出して計測しデータベース化するためには工夫を要する。例えば、正確な動きの情報を得る方法として、運動している人体などにマーカをつけてその動きを取得することはできるが、限られた部位のデータしか得られない、動作空間の範囲が限定してしまう、対象物体に明示的なマーカをつけなければならないなどの問題がある。この結果、サッカーなどの比較的広範囲の空間を用いて行なわれるスポーツをデータベース化の対象と考えると、マーカを用いた動きの取得には不向きである。

そこで本稿では、サッカーなどの団体スポーツが行われるような比較的大きな世界を対象とし、その中で人間のある目的をもった動きの3次元、時間的情報を取得するため、同一の広い空間を複数台のビデオカメラを用いて非同期で撮った映像群(多視点映像)をもとに、マーカをつけずに解析し、形状

及び運動のモデリングを行なう手法を提案し評価する。また今回述べる手法の適用例として、複数のオブジェクトが協調してある目的をもった動きをするサッカーの試合を複数台のカメラを用いて撮影し、その中で起こる事象の解析とモデリングを行なう。

各時間ごとの人間の姿勢を記述する方法としては、人体の各部分が始めにどの位置にあり、その後、時間の経過にもなるとどのような位置に移るのかを記述することが必要である。従って、試合に出場しているプレーヤーの足の動きに注目し、下半身の各パーツのモックアップモデルがどの位置、角度にあるかを多視点映像と対応づけることで決定する。これにより人間のモックアップモデルの姿勢を決定し、そのシーンにおける被写体の姿勢のデータベース化を行なう。

2 多視点映像からの時空間DBの構成

現実世界の情報をデータベース化するためには、まず対象とする世界がどのような時空間的情報で構成され、どのようなデータによってその情報が表現されるのかについて考える必要がある。

ここでは、データベース化する対象範囲の世界において時間的、空間的にある「ことがら」が始まってから終わるまでを、単位としてシーンと呼ぶ。例えばサッカーならば攻撃と守備の切り替わりを起点とし、攻撃側がゴールを決めるかボールがラインから出るまでを1シーンとする。

このシーンに現れるオブジェクトをデータベース化するには、次のような情報をモデル化し、蓄積する必要がある。

- シーンの全体的な時間や3次元空間情報
- 構成している場の情報
- シーン中に登場するオブジェクトの属性
- オブジェクトの情報

これらの情報の一次取得手段の1つとして、多視点映像は、シーン全体を撮影できる、注目するオブジェクトに専用カメラを設けることができる、時間情報を含むなどの特徴から有効である。その反面、

多視点映像と構成する個々の映像は、空間を画像に投影した一すなわち3次元を2次元にプロジェクトしたデータ系列であり、その3次元情報が失われるため空間情報復元が必要となる。このためには、それぞれの視点の画像間で対応する点(ピクセル)を検出しなければならないが、単純に画像処理だけで対応点検出をすることは困難であり、一般に多視点画像からの空間情報復元には工夫を要する。例えば、静止状態の多視点画像からの空間情報復元として、平行カメラからの奥行きを概算し物体間の大まかな重ね合わせを復元する方法 [4] や、360度方向から物体を撮影する方法 [5] などがある。しかし、サッカーの様に広い空間で選手が自由に動くタイプのシーンでは、解析したい情報・データをもう1度考え直す必要がある。

上に列挙した情報は、1シーン中に不変なデータ(例えばサッカーならば、フィールドや各選手などの情報)と動的に変わるデータ(各選手の位置や動きなどの情報)に大きく分けられる。この2種類のデータをあるシーンから取得する方法は異なり、また、それらのデータを独立に取得することは困難である。そこで、本稿では、シーン中のオブジェクトの構造や形状をあらかじめ取得しておき、その意味情報をもとに空間的な位置・動きの情報を得る手法を提案する。

その適用例として、サッカーシーンの時空間データベース化支援の為に、サッカー選手の脚の動きを解析するシステムを構築し、実証することを試みた。

3 多視点映像からの運動解析手法

3.1 多視点による時空間情報の抽出

広い空間内で動く人間の動きを抽出するには、その人間の空間中の位置を知る必要がある。しかし、人間の動きを抽出するために特定の人間をアップで撮影した映像では、背景が少ないためその位置が分からない。そこで我々は、複数台のカメラの内数台を全体の空間を固定した状態で撮影し、これらのカメラから得られる画像を基に各人間がどこにいるのか大まかな位置を特定した。また、人間の姿勢を抽出するには両腕や両脚がはっきり見える画像を基に解析の方が解析が行ないやすい。そこで、全体を撮影した映像から抽出対象となる人間の姿勢の比較

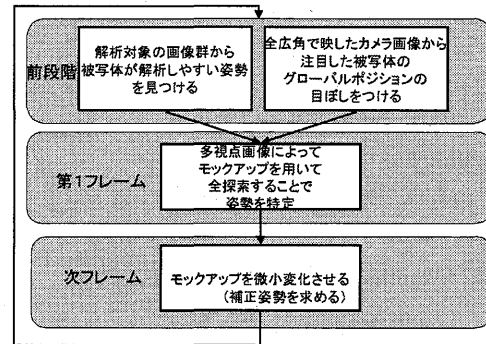


図 1: 多視点映像による時空間情報の抽出法

的解析しやすい場面(基準フレーム)を見つけることとする。つまり、あるシーンにおいて姿勢の抽出を行なう開始フレームは常に初めからではなく解析のしやすい基準フレームから姿勢を抽出し、次のフレームからは、前の姿勢を基に解析を行なう。

以上の前段階を行なった後に、各人間を最低2台のカメラで追うことにより姿勢の解析を行うことを考えた(図1)。つまり、全体の空間を撮影したカメラから得られた、被写体の大まかな位置を含めた近傍でCGで作成した人間の各関節を動かし、各カメラ間に映る画像とのマッチングをとることで姿勢を決定する。これは、比較的時間がかかっても良いとするが、精度が高く姿勢を同定することを目的とする。

次のフレームからはモックアップの構造的・時間的制約を満たしつつ微小変化させることで、比較的短時間に姿勢を決定していく。この方法ではフレームが進むにつれ誤差が増えていくので、数フレーム後にまた前段階に戻り姿勢の補正を行なう。この作業を繰り返すことで姿勢の時系列情報を得る。

本稿では、被写体の解析しやすい姿勢に対してどのようにしたら比較的時間がかかっても精度が高く姿勢を抽出できるのかについて、その方法を提案し検証を行なうこととする。

3.2 Mediator の導入

各視点から得られる画像から人間の姿勢を抽出する際に、その人間に関するデータ(例えば一般的な人間の構造やその人個人の体格)、を一切使わずに

解析を行なおうとすることは困難である。そこで、人間個人に関するデータをあらかじめ知識として持ち、解析対象に対してはそのデータに基づいたモックアップをまず構築し、これと各視点からの画像とマッチングさせることで、そのフレームにおける姿勢を解析する。

この身体形状や動作をモデル化する方法として、我々は Mediator という概念を提案している [3]。

この Mediator は次のように定義される。

- **形状構造 Mediator** 人間の体を形状的に変化しない「部品」と「関節」から成るとし、個々の部品に対してはポリゴンから成る各人間ごとの表面形状とテクスチャを持ち、関節に対しては関与する部品と各人間ごとの自由度と動作域のデータを割り当てることによって得られる 3次元グラフィックスモデル。
- **動作 Mediator** ある人間がある時行つたある「動作」について、その途中に必ず現れる特徴的な中間姿勢を開始点からその姿勢に至った相対時刻、および、その時点の形状構造 Mediator における関節角度によって記述したもの、すなわち、その人間のその動作ごとに決められた特定中間姿勢列。

この Mediator を導入することにより、個人別の形状・動作の特徴を統一的に規格化することができ、また形状や動作を少ない情報量で表現することができる。

そこで形状・構造 Mediator の概念にもとづいて与えたデータをもとにモックアップを作成し、各人間の姿勢時系列を得ることにより動作のモデリングを行なうことができる。

3.3 人体のモックアップの構成

形状構造 Mediator は、元の物体の詳細な形状構造データを基に作成されるため、映像の被写体固有の形状や色を反映している。そこで、この形状構造 Mediator を用いた人間のモックアップを構成しておけば、画像との対応をとりやすい。

この人間モックアップモデルは、時間的に形状が変化しないパーツ と、関節からなるとする。パーツ は、被写体固有の形状データを持ち、関節は各パーツ 間における人間固有の可動域をもっている。

また、一般的にグラフィックスの世界では人体は腰を根とした木構造で表現でき、その木構造に従って各パーツが相対座標で表されるため、上半身と下半身は別々に処理することが可能である。従って本稿では、人体の足の動きに注目し下半身部分についてのみの解析を行なうこととする。

この下半身モックアップは、人体形状モデルと人体階層モデルに基づいて、腰 (pelvic)、左右の大腿 (thigh)、左右の下腿 (thin) のパーツで構成される。各パーツの形状は、円筒近似したものであるが、その長さは対象となる人間の個別の長さをもっており、形状構造 Mediator の概念に基づいている。

この人間モックアップモデルの姿勢を仮想空間上で再現した時、各々の視点の画像との差分が最も少ない—すなわちマッチングがとれれば、それを選手の姿勢と決定する。

3.4 多視点映像と空間の関係

人間モックアップモデルと画像とのマッチングをとるためには、まず空間モデルを考えておく必要がある。空間上で運動する被写体を、複数のカメラで撮影した際の空間モデルを図 2 に示す。

あるカメラ C_1 の焦点を原点として画像の奥行き方向を Y'_1 、画像の水平方向 U_1 を X'_1 、画像の垂直方向 V_1 の逆方向を Z'_1 としたカメラ座標系 O'_1 において O'_1 系の被写体の点を $p'_1 \equiv (x'_1, y'_1, z'_1)$ とする。

p'_1 が画像上の画素の位置 (u_1, v_1) (ただし、 $0 \leq u_1 \leq U_1, 0 \leq v_1 \leq V_1$: U_1, V_1 はそれぞれカメラ C_1 の画像の水平方向と垂直方向の画素数) に結像したとき、カメラ C_1 の焦点距離が f_1 であったならば、透視変換の原理より、

$$\begin{cases} f_1 : y_1 = (u_1 - U_1/2) : x'_1 \\ f_1 : y_1 = (V_1/2 - v_1) : x'_1 \end{cases} \quad (1)$$

となる。従って、画素 (u_1, v_1) に結合される空間中の点 $p'_1 = (x'_1, y'_1, z'_1)$ は $s(0 \leq s < \infty)$ を媒介変数としたカメラ系での透視変換 $C_1(s)$ により

$$\begin{bmatrix} x'_1 \\ y'_1 \\ z'_1 \\ 1 \end{bmatrix} = C_1(s) \begin{bmatrix} u_1 \\ v_1 \\ 1 \\ 1 \end{bmatrix} \quad (2)$$

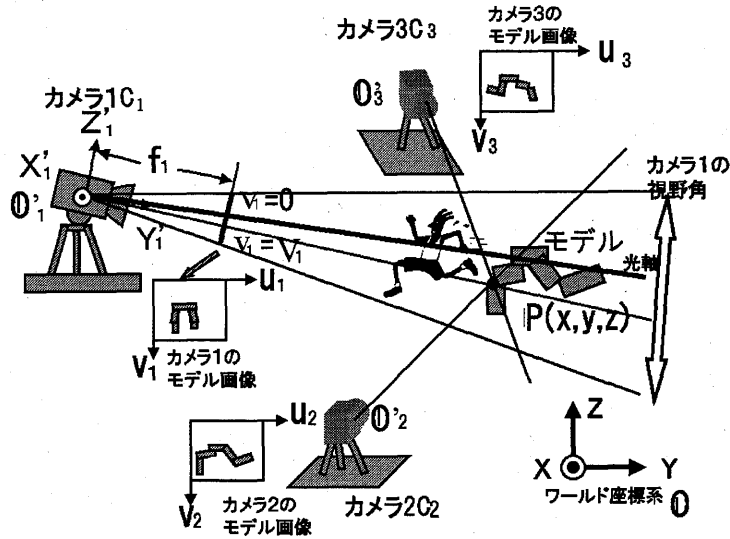


図 2: 複数のカメラで撮影した際の空間モデル

ただし、

$$C_1(s) = \begin{bmatrix} s/f_1 & 0 & -sU_1/2f_1 & 0 \\ 0 & 0 & s & 0 \\ 0 & -s/f_1 & sV_1/2f_1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

に存在する。

ところで、カメラの設置位置やパン (横方向の回転), チルト (縦方向の回転) などの回転操作の影響は、ワールド座標系 O を基準として、それぞれの行列を S 、 Y 、 P とすると

$$S(X_{C1}, Y_{C1}, Z_{C1}) = \begin{bmatrix} 1 & 0 & 0 & -X_{C1} \\ 0 & 1 & 0 & -Y_{C1} \\ 0 & 0 & 1 & -Z_{C1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$Y(\phi_1) = \begin{bmatrix} \cos \phi_1 & -\sin \phi_1 & 0 & 0 \\ \sin \phi_1 & \cos \phi_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$P(\theta_1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta_1 & \sin \theta_1 & 0 \\ 0 & -\sin \theta_1 & \cos \theta_1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

によって変換されるため、カメラは水平を基準に設置したと仮定すると、 O 上の点 $p(X, Y, Z)$ は O'_1

では、

$$p'_1 = P(\theta_1)Y(\phi_1)S(X_{C1}, Y_{C1}, Z_{C1})p \quad (7)$$

である。従って、ワールド座標系 O 系における点 $p(x, y, z)$ は、カメラ C_1 の像に結像し、

$$u_1 = C^{-1}(s_1, f_1)P(\theta_1)Y(\phi_1)S(X_{C1}, Y_{C1}, Z_{C1})p \quad (8)$$

である。上式において、 $f_1, \theta_1, \phi_1, X_{C1}, Y_{C1}, Z_{C1}$ は、カメラを操作した際に変化し決定される値であり、これを「カメラワーク」と呼ぶ。「カメラワーク」のパラメータの決定は、画像全体のオプティカルフローを追跡して算出する方法や、モザイク画像にもとづくアフィンパラメータの同定 [6] によって映像から自動的に決定することもできる。

以上より、同時に複数のカメラ (視点) を用いて同一の点 $p(X, Y, Z)$ を撮影すると、

$$u_i = C^{-1}(s_i, f_i)P(\theta_i)Y(\phi_i)S(X_{Ci}, Y_{Ci}, Z_{Ci})p \quad (9)$$

となる。

ここで、カメラワークが何らかの方法で分かっているとすれば、これを定数 A_i として

$$u_i = C^{-1}(s_i)A_i p \quad (10)$$

となる。

この原理に基づいて、モックアップモデル内座標系におけるモックアップモデル上の点を p_m としたとき

$$u_i = C^{-1}(s_i)A_i T p_m \quad (11)$$

(ただし、 T :ワールド座標系からモデル座標系への変換行列)に結像する。これより、 p_m 上の色をモックアップのテクスチャから求めたとすれば、カメラ C_i のuv上の画素と同一色となる。

3.5 多視点映像からの姿勢決定

各フレームにおける人間モックアップモデルの姿勢の同定に用いる情報としては、画像中の色情報が考えられる。しかし、各画像はカメラごとに全体の明るさが違うので、各画像の1ピクセルごとの色情報が微妙に異なる。また、同じカメラにおいても異なるフレーム間においてその色情報が変わってしまう。

しかし、あるオブジェクトの各パーツなどの色情報には様々な色が混ざっているが、その混じり方には固有の特徴があり、同一カメラで獲得したフレーム列の中で変わらない。そこで、この固有の特徴をオブジェクトごとのHSVの色ヒストグラムとして表す。

また、我々があらかじめ選択した同一色領域の色ヒストグラムを作成し、シーン内における色ヒストグラムの変化を調べたところ全シーンにおいてオブジェクトごとにはほとんど変わらないことが分かった。

そこで各カメラごとにシーン中の数フレームを基に被写体の各パーツや背景(ここでは、フィールドの芝やトラック)について作成する。次に被写体の各パーツにおける色ヒストグラムはモックアップの基準色データとして全頂点に割り当て、これをモックアップのテクスチャデータとする。

また前節より、モックアップ上の任意の点は画像中のどのピクセルに投影されるかが分かる。そこで姿勢の決定の手順は以下の方法が考えられる。

1. 人間モックアップモデルの全頂点をピクセルマップに投影し、映った各ピクセルの位置を取得する。
2. 1で得た各位置の近傍 n ピクセルに関し色分布ヒストグラムを作成する。

3. 基準色データと2で得たヒストグラムを比較し、全てのヒストグラムの中で1番近いものがそのモデルのヒストグラムならば当たりとしその数を数える。

4. 1~3を全ての姿勢で行ない、それ以外の姿勢について当たりの数を数え、最大のものをその時点における姿勢とする。

これにより求めた姿勢をもとに次フレームからは、下半身モックアップを微小変化させてマッチングすることで短時間に姿勢の決定を行なう。

4 試作システムの実装と評価

4.1 実装システムの構成

3章で述べた手法に基づいて、サッカーの試合を5台のカメラを用いて非同期で撮影した。このカメラの内、2台は1人の被写体を追跡し、1台は補助用として用いた。残りの2台はカメラを固定し、フィールド全体を撮影した。

被写体を追跡するカメラの既知のパラメータは、各カメラの位置座標 (X_i, Y_i, Z_i) とカメラ内での焦点距離 f_i 、パン、チルトの角度 θ_i, ϕ_i である。また、今回の実験では、腰の3次元的な位置座標とその選手が向いている大まかな方向はあらかじめ分かっているものとし、各関節角度のみを 5° ずつ 60° の範囲で動かしてマッチングを行なった。

2台のカメラの同期については、今回のサッカー場のような広い空間内では各カメラに同期信号を与えて合わせるができない。そこで、各カメラで撮った映像をもとに被写体が、ボールを蹴った瞬間といった典型的な姿勢をしているところをキーとしてそれぞれのフレームの同期をとることにした。

実験は、2台から得られたほぼ同時時間の画像を基に3章で述べた手法に基づいて解析を20フレームについて行なった。

4.2 試作システムの評価

解析により得られるデータは、各パーツの仮想空間上での位置座標と各関節の相対角度である。これにより、下半身モデルの姿勢が決定される。評価の仕方として、始めに解析したフレームにおける正

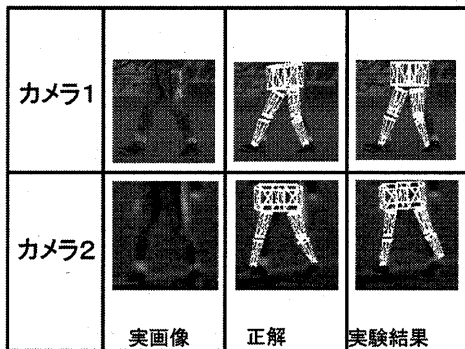


図 3: フレーム No.4 の実験結果

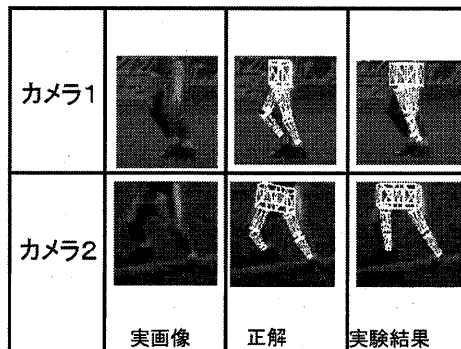


図 4: フレーム No.5 の実験結果

解の姿勢を得るために、3次元のグラフィックシュミレータを用いた。これにより得られた各関節の相対角度と実験の結果えられた相対角度を比較した(表 1)。

全体の結果を通して、今回用いたモックアップには足 (foot) がいないため各パーツの z 軸回転については精度が良くなかった。5回の実験の結果から、視覚的に評価して最も良いと判断されたフレームの内フレーム No.4 と、最も外れていると判断されたフレーム No.5 の出力結果をそれぞれ図 3, 図 4 に示す。フレーム No.1~4 は、2台のカメラともに両足が見えているため比較的マッチしていた。しかし、フレーム No.5 はカメラ 1 の被写体の右足が左足に隠れているため、精度が悪かった。このことから、2台のカメラの両方に被写体の両足が見えていた場合には、比較的精度良く姿勢の同定ができることが分かった。また、更にカメラの台数を増やすことで精度を高められることが予想できる。

5 まとめ

本稿では、多視点映像を基にした時空間DBの構築方法について述べた。手法としては、まず Mediator の概念に基づく形状のモックアップモデルを作成しておき、このモデルが各々のカメラにどの様に映し出されるのかを算出して、撮影された映像との比較を行なった。

また今回は、前半の被写体の解析しやすいつレームに対して比較的時間がかかっても精度が高く姿勢

を同定することに主眼をおくこととし、2台のカメラ間で解析対象の両脚がはっきりと見えているならば比較的精度が高く姿勢を決定できることを示した。

姿勢予測などを用いて、時間軸方向の運動解析を少ない計算量で実現することが今後の課題として残されている。

参考文献

- [1] 有澤 博: “リアルワールドデータベースとその実現技術 1-3”, bit, Vol.28, No.9-11, 1996
- [2] 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”, 信学論 D-I, Vol.j82-D-I, No.1, pp.211-222, 1999.1
- [3] 今井 さやか, 富井 尚志, 有澤 博: “Mediator 概念に基づく身体形状・動作のモデル化とデータベース設計”, 情報処理学会論文誌:データベース, Vol.41, No.SIG1(TOD 5), pp.100-108, Feb.2000.
- [4] Kiyohide Satou, Itaru Kitahara, Yuichi Ohta: “3D Image Display with Motion Parallax by Camera Matrix Stereo”, IEEE Proceeding of MULTIMEDIA'96, pp.349-357, 1996.6
- [5] Andrew Zisserman, Andrew Fitzgibbon, Geoff Cross: “VHS to VRML:3D Graphical

表 1: 関節の相対角度の比較

フレーム No.		腰			左腿			右腿			左脛	右脛
		x	y	z	x	y	z	x	y	z	x	x
1	正解	0	0	-20	-5	-15	-20	-20	0	40	30	0
	結果	0	-5	-5	10	-5	-25	-20	-5	30	5	10
2	正解	0	-5	5	-5	-15	-20	-5	5	-30	40	0
	結果	0	-5	-15	10	-15	-20	-25	-5	5	15	40
3	正解	0	0	-20	-5	-15	-25	-15	15	-5	15	0
	結果	0	-5	-20	-5	-15	-25	-25	15	-5	15	5
4	正解	0	5	-40	15	-15	-20	-20	0	-20	15	30
	結果	0	-10	-25	-10	-5	15	-5	30	-20	10	30
5	正解	0	5	-45	10	-25	-20	-25	10	-25	10	65
	結果	0	0	-25	-15	-30	-30	20	0	-30	5	5

Models from Video Sequences”, IEEE ICM-
CS’99, pp51-57, 1999.6

- [6] Sunghoon Choi, Yongduek Seo, Hyunwoo Kim, Ki-Sang Hong: “Where are the ball and players? :Soccer Game Analysis with Color-based Tracking and Image Mosaick”, I-CIAP’97

以下の頁は著作権者の許諾を得ていないため、公表できません。

抜刷 3 ~ 抜刷 6

抜刷 8 ~ 抜刷 8

抜刷 13 ~ 抜刷 13

抜刷 28 ~ 抜刷 28

抜刷 ~ 抜刷

抜刷 ~ 抜刷

以下の頁は著作権者の許諾を得ていないため、公表できません。

抜刷 4 ~ 抜刷 6

抜刷 8 ~ 抜刷 8

抜刷 13 ~ 抜刷 13

抜刷 28 ~ 抜刷 28

抜刷 ~ 抜刷

抜刷 ~ 抜刷

以下の頁は著作権者の許諾を得ていないため、公表できません。

抜刷 4 ~ 抜刷 6

抜刷 8 ~ 抜刷 8

抜刷 13 ~ 抜刷 13

抜刷 28 ~ 抜刷 28

抜刷 ~ 抜刷

抜刷 ~ 抜刷

以下の頁は著作権者の許諾を得ていないため、公表できません。

抜刷 4 ~ 抜刷 6

抜刷 8 ~ 抜刷 8

抜刷 13 ~ 抜刷 13

抜刷 28 ~ 抜刷 28

抜刷 ~ 抜刷

抜刷 ~ 抜刷

次世代放送向け多チャンネル映像のデータベース化

中家 路也 徐浩源 富井尚志 有澤博

横浜国立大学 工学部 電子情報工学科
〒240-8501 横浜市保土ヶ谷区常盤台 79-5
E-mail:{nakaie, tommy, arisawa}@arislabs.dnj.ynu.ac.jp
横浜国立大学 総合情報処理センター
〒240-8501 横浜市保土ヶ谷区常盤台 79-5
E-mail:haoyuan@ynu.ac.jp

あらまし

本稿では、昨今盛んに行なわれている次世代放送の研究において新たな提案を試みている。一般的にいわれる次世代放送の特徴としては、インタラクティブ性、デジタル放送性、高画質性(ハイビジョンなど)などがあげられる。しかし、近年急激に普及しているインターネットの基盤である高速ネットワークの技術や、映像などの大容量データを蓄積することを可能にしたストレージ技術を用いることにより、次世代放送にさらなる可能性があるのではないかと考えられる。そこで、本稿では次世代放送における多視点カメラによる新しい撮影環境を提案し、その環境によって取得された映像などのデータを高速ネットワークを用いてデータベースに蓄積するシステム、MCVDBを提案しその設計を行なう。また、このシステムにおけるプリミティブである高速ネットワークの検証をおこない、スキーマに基づくデータベースへの蓄積システムについて述べる。

Multi Channel Video Database for Next-Generation Broadcasting

Michiya Nakaie Haoyuan Xu Takashi Tomii Hiroshi Arisawa

Division of Electrical and Computer Engineering, Faculty of Engineering,
Yokohama National University
79-5, Tokiwadai, Hodogaya-ku, Yokohama 240-8501 JAPAN
E-mail:{nakaie, tommy, arisawa}@arislabs.dnj.ynu.ac.jp
Information Processing Center
Yokohama National University
79-5, Tokiwadai, Hodogaya-ku, Yokohama 240-8501 JAPAN
E-mail:haoyuan@ynu.ac.jp

Abstract

This paper presents a new methodology in researching next-generation broadcasting. Generally a feature of next-generation broadcasting is interactivity, digital broadcasting, high definition. As for interactivity, audiences may request to a broadcasting station and using the results, he edits programs. But in recent years, a high speed network technology on a base of rapidly spreading Internet, it is possible to accumulate very large multimedia data like a video data by a technology of storage, so we suppose a new possibility of next-generation broadcasting. So this paper presents a new filming environment for a multi channel camera, and a system of transferring the obtained video data in it on high speed network, then accumulating those data in database. We design the system(MCVDB), and verify high speed network in this system, describe accumulating system on the database schema.

1 はじめに

本稿では、昨今盛んに行なわれている次世代放送の研究において新たな提案を試みている。一般的にいわれる次世代放送の特徴として、インタラクティブ性やデジタル放送、高画質性などがある。しかし、近年急激に普及しているインターネットの基盤である高速ネットワーク(図1)の技術や、映像といった大容量データを蓄積することを可能にしたストレージ技術を用いることにより、次世代放送にさらなる可能性があるのではないかと考えた。

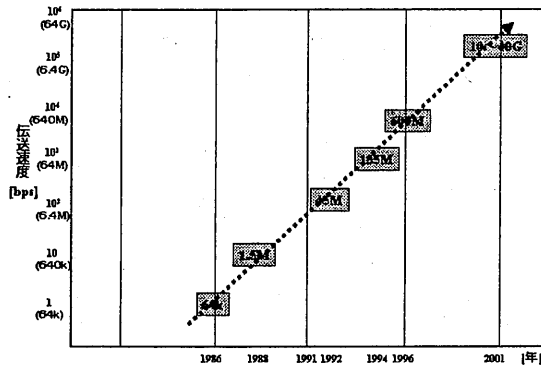


図1: コンピュータ・ネットワークの高速化

そこで次世代放送においては、現在の野球中継の放送ように1チャンネル映像を放送するだけでなく、同時刻における様々な視点から撮影している映像やステレオ映像を用いることにより、球場の中で様々な視点から見ている感覚を味わえるといった、臨場感豊かな放送が求められると予測される。また、多視点カメラ映像を用いて解析処理を行なうことにより、被写体の空間位置を抽出するといった研究も近年盛んに行なわれていて、多視点カメラ映像の取得やデータベース化が求められると考えられる。

しかし、多視点から撮影している映像をスイッチングして、1チャンネル映像を放送するといった現在の放送方法では、同時に複数のチャンネルの映像を放送することが困難であるため、多視点カメラの映像をマルチチャンネルTVとして見るができない。更に、これまでの放送方式では多視点カメラの映像を放送と同時に、データベースに蓄積する必

要性がなかったため、このようなデータベースは実現されていない。

そこで著者らは、多視点カメラによって撮影されている映像を、高速ネットワークと大容量ストレージを用いてデータベース化することにより、マルチチャンネルTVを実現することができると考えた。このシステムを次世代放送向け多チャンネル映像のデータベース化システム(Multi Channel Video DataBase: MCVDB)と呼ぶ。また、MCVDBにおいては映像データをデータベースに蓄積後、そのデータが画像解析などにも用いられると想定しているため、元の情報の欠落がない非圧縮映像データを用いることも必要である。しかし、図1のように、ネットワークの高速化が進んでいるとはいえ、ネットワークの帯域は有限であり、非圧縮映像データ(ここではD1映像)をリアルタイムに伝送するためには最低160Mbpsの帯域が必要となる。このような帯域を多チャンネル分確保することは困難であり、実用的ではない。そこで、限られたネットワークリソースを有効に利用して、大規模な多チャンネル非圧縮映像データを情報の欠落なくデータベースに蓄積する方法が必要となる。また、多視点カメラの映像を蓄積する上では、各カメラのフレームレートが異なることも考えられ、このような場合に撮影時間とのカメラ固有の時間との対応づけをどのようにしておこなうかが問題となる。更に、どのようにして多チャンネル映像の放送を実現するかが課題となる。

論文の構成を述べると、2章では多チャンネル映像のデータベース化及び、マルチチャンネルTVを実現するために必要とされる機能について述べ、MCVDBのシステム設計を行なう。3章において多視点カメラから取得した映像やカメラパラメータといったデータの特徴に触れ、多視点カメラ映像のデータベース実現するためのスキーマを設計する。そして、4章では本システムを実現する上で重要となる高速ネットワークの評価を行ない、最後にまとめを述べる。

2 次世代放送向け MCVDB システム

著者らが提案する、次世代放送向け多チャンネル映像データベース化システムは、多視点カメラの非

圧縮映像データを、高速ネットワークを用いてリモートにあるデータベースに蓄積することによって、マルチチャンネルTVと多視点カメラの映像データベースを実現しようというものである。このシステムを実現するための要件として映像のバッファリング、フレーム補完、ネットワークによる高速伝送、インデックスとの対応づけ蓄積があげられる。著者らはMCVDBシステムを設計する上で、上記の要件をプリミティブとしてかかげ、この要件を解決することにより本システムを実現することを目標にしている。

1. 映像のバッファリングと再送機能:多チャンネル映像においては高速なネットワークを用いて伝送した場合にも、フレームドロップが生じる。そこで、ドロップしたフレームをバッファリングすることが可能であることと、後に再送することが可能であること。ただし、バッファは撮影対象を十分に蓄積するだけの容量があるものとする。
2. フレーム補完機能:映像がフレームドロップした場合には、バッファリングしているフレームを後でデータベースの所定の位置に埋め込む。すなわち、フレームドロップした時点がどこであるかを記録しておき、データベースに蓄積する上で時間軸にそって蓄積を可能にする機能。
3. 高速ネットワーク伝送機能:映像を伝送するための帯域としては、最低でも1チャンネルの映像をフレームドロップなしに伝送できることが可能であること。
4. インデックスとの対応づけ蓄積機能:撮影に用いられるカメラは様々な機種と考えられるため固有のフレームレートを持つ。そのため、各カメラが持つカメラ時間が存在しその時間と撮影時間との対応が記述できその対応に基づき蓄積する機能。

1は各カメラに設置されるディスクが行なう処理である。2、4はデータベース側で行なわれる処理である。3は高速ネットワークに要求される要件である。上記の機能が本システムの上でどのような流れで処理を行なうかを具体的に説明する。

2.1 システム構成

本節では、MCVDBシステムのマルチチャンネル映像放送と、多チャンネル映像のデータベース化の実現方法について述べる。MCVDBのシステム構成は図2のようになっている。

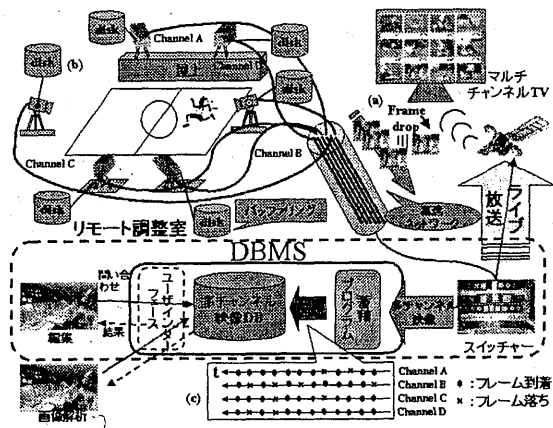


図 2: システム構成図

本システムでは、多視点カメラによって映像を取得することを前提としている。しかし、この多視点カメラによって取得されるデータは映像以外にも、カメラアングルやカメラポジションといったデータも取得されるものとした。というのは、このようなデータはジャイロやGPS(Global Positioning System)を用いることで取得可能であり、本システムでは多視点カメラによって取得した映像データは、解析処理などにも利用されると考えている。そのことから、このようなデータを取得してデータベースに蓄積しておくことは後の解析処理にも有用であると考えた。次に、多視点カメラによって取得された映像などのデータは、高速ネットワークを用いて伝送されるのであるが、ここで一つの仮定を行なう。

多視点のカメラ映像のどれを放送するかの決定はディレクターが行ない、撮影対象の状況を見て多視点カメラ映像、または単視点カメラ映像の選択をする。

例えば、野球の放送ではピッチャーが投げるかまえばに入れば、投球するということがわかり、視聴者の興味がバッターとピッチャーの対決に集中する。このような場面には通常ピッチャーの後部からピッ

チャーをバッターを写している単視点映像による放送が行なわれると考えられる。つまり、このような場面ではディレクターは単視点放送を行なうと予想される。ところが、バッターがバッターボックスを外した場合には、視聴者のバッターやピッチャーに対する興味が薄れると考えられる。そこで、そのような場面では、ディレクターは多視点カメラ映像を用いて球場内を様々な角度から写している映像を選択すると仮定するのである。このようなことから、常に多視点映像が放送に用いられるわけではなく、状況に応じて単視点映像との切替が行なわれると考えられる。そこで著者らは、このような放送の性質を利用して、高速ネットワークの帯域の有限性に対応できると考えた。つまり、多チャンネルの放送ではネットワークリソースの有限性からフレームドロップが生じる(図 2-(a)) ため、そのような映像データをそのままデータベースに蓄積することが実用的とはいえない。そこで著者らは、フレームドロップが生じたものに関してはカメラに附属するディスク(図 2-(b)) を用いてバッファリングを行ない、後で再送することにより有用な映像データベースを実現しようと考えた。そのためには、実際にフレームドロップが生じた時点を記録しておく仕組み(図 2-(c))が必要となる。

3 多チャンネル映像のスキーマ設計

前節でも述べたように、多チャンネルカメラによって取得した様々な情報を逐次データベースに蓄積し、様々な問い合わせに応じることのできるスキーマを設計する必要がある。まず、実世界から多チャンネルカメラによって取得されるデータについての特徴を具体的に述べ、次にスキーマ設計を行なう。

3.1 蓄積する情報

高速ネットワークからデータベースに渡されるデータは、映像などのような原データや撮影された時刻などのインデックスデータであるが、具体的なデータは以下のようなものである。

- 映像

取得される映像は、様々な位置に設置された多チャンネルのカメラによって取得されたものである。また、場合によっては立体映像、つまり、2台のビデオカメラをセットに何組かを用いて取得される。注意すべき点はこれらのカメラが全て同期されて(フレーム間の対応づけがなされて)データベースに蓄積される必要があるということである。

- カメラ情報

本システムにおいてはカメラによって取得される映像データだけでなく、カメラの情報をGPSやジャイロを用いて取得し、放送と同時にデータベースに蓄積される。このような情報の例としては、個々のカメラがある時点にもつ焦点距離などのパラメータや、撮影環境におけるカメラの位置やカメラの角度といったデータなどがあげられ、このような情報をもとにカメラの撮影領域などが算出されデータベースに蓄積される。

3.2 映像データベースのスキーマ

多視点映像のデータをデータベースに蓄積するためのスキーマを設計するためには、エンティティベースのデータモデルが必要である。ここでは、データモデルとして、ERモデルなどの意味データモデルを拡張したAISモデル[1],[2],[3]を用いて設計を行っている。図 3-(a)に、本システムのデータベースで用いるスキーマを示す。このスキーマにおいて、プリミティブとなっているのが撮影時刻とカメラ時間の対応を行なっている部分である。本来、カメラ時間と撮影時刻との対応づけを行なうのが困難である。なぜなら、各カメラのフレームレートの違いなどから、ある撮影時刻に対応するカメラ時間が必ず存在するわけではないからである。このような問題を、撮影時間の最小単位を時区間(3-(b))とすることにより解決している。

3.3 スキーマに基づく蓄積システム

次に、前節で設計したスキーマに沿ってリアルタイムにデータベースに蓄積するシステムについて述べる。MCVDBにおいては、ネットワークから転送される映像や時間情報などのデータを先に設計した

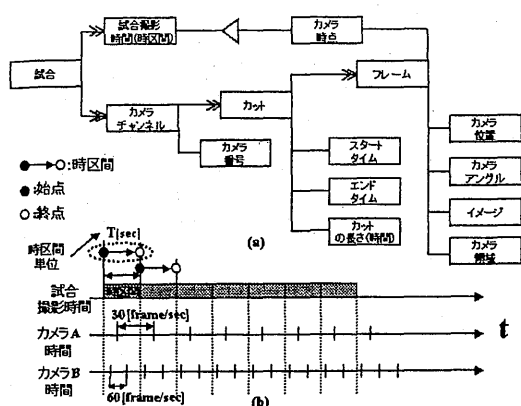


図 3: 多チャンネル映像データベースの例

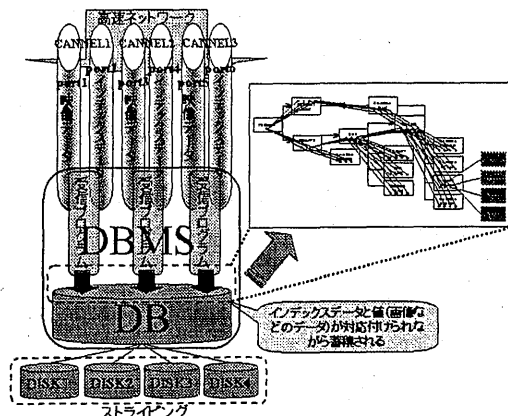


図 4: データ蓄積システム

スキーマに基づきながら蓄積する必要がある。つまり、映像をフレームの列とで考えた時に、そのフレームと撮影された時間との対応(つながり)などを記述し、それをインデックス情報として蓄積するのである。このような対応づけを行なうことにより、様々な検索が可能になるのである。そこで、このような処理を次の図4に示すようなシステムによって、実現できるのではないかと考えた。

そこで、図4を例に、カメラのチャンネルが3チャンネルの場合においてこのシステムについて説明する。3チャンネルの各映像は撮影時間などのデータとともに伝送されてくる。このようなことから、まず映像の各チャンネルに対して1ポートずつ割り振る。また、撮影時間などのインデックスデータについても、各チャンネル毎に対して専用のポートを割り振る。この結果、図4においては全てで6つのポートが割り振られることになる。このようにして、割り振られたポートに送られてきたデータを受信プログラムで受取り、一つの値(データ)を受けとる毎にエンティティを順次生成していく。つまり、インデックスやフレームの数だけエンティティは生成される。また、受信プログラムではそのエンティティ間の繋がり(対応)も蓄積と同時に結び付けていく。例えば、FRAMEというエンティティ(ID:f1)と、IMAGEというタイプのエンティティ(ID:i1)の繋がりを結ぶのである。このように、データの到着時に順次エンティティIDを生成しながら、エンティティ間の繋がりを結ぶという処理を行なうのが受信プログラムで

ある。

以上の処理によって、データがデータベースに蓄積されるのである。

4 高速ネットワークによる伝送と蓄積

MCVDBにおいては多視点カメラから撮影された映像などのデータは高速ネットワークを用いて伝送される。このようなことから、高速なネットワークによる高速伝送を実現する必要がある。

そこで、本稿では高速ネットワークを実現する手法として、マルチメディアデータの伝送に適しているといわれるATM(Asynchronous Transfer Mode)を用いた。そして、伝送速度の測定実験と多チャンネル映像の伝送実験を行なった。実験においては次に示す点について注目し実験を行なった。

- ATMの通信方式としてIPOA(IP over ATM)とLANE(LAN emulation)という二つの方式を用い、この二つの方式を比較する実験を行なった。
- 実験においてはソケットを用いたプログラムを使用し、その通信の際には通信プロトコルとしてTCP/IPを用いた。また、実験においてウィンドウサイズをパラメータの一つとして用い、ウィンドウサイズと伝送速度の関係に注目した。

- リアルタイムに伝送される映像データ(非圧縮データ)を伝送し、映像にフレームドロップが生じずに蓄積できるかを確認した。

まず最初にあげた ATM の通信方式であるが、IPOA では IP パケットを ATM セルにのせるため、TCP/IP または UDP による通信しか行なえない。また、LANE では従来の LAN(Ethernet, FDDI など)の通信方式が行なえるが、オーバーヘッドによる伝送速度の低下などの問題がある。

次に、ソケットのプログラムを用いた伝送を行なう際に、あらかじめ送受信側でウィンドウサイズを決めて伝送速度の測定を行なった。このウィンドウサイズとは TCP/IP プロトコルにおいて一度に送っても良いフレームの個数を示すもので、この値が大きいほどフレームの再送などが起きなかった場合に伝送効率が良くなる。しかし、ネットワークのトラヒックやネットワークインターフェースの性能などによりこのサイズに限界がある。

最後に、このような伝送速度測定実験を行なうことで最適なウィンドウサイズがわかることから、この結果を反映した最適な伝送設定を行ない、多チャンネル映像を伝送してチャンネル数とフレームドロップの関係を測定した。この結果はマルチチャンネル TV を実現する上で重要なサンプルとなると考えられる。というのは、伝送帯域は有限でありその帯域に応じたチャンネル数の決定を行なう必要があると考えられるからである。

また、このような ATM-LAN を構築する上でネットワークで問題となる点をいくつか明らかにした。そこで、この実験を行なった環境およびその伝送速度実験の結果について次に述べる。

4.1 実験

最初に、前述の伝送速度実験におけるネットワーク構成図 5(a) を示す。また、この実験は表 1 にあげる 3 台のワークステーションを用いて行なった。この伝送速度実験に関係のあるパラメータとして MTU(Maximum Transfermission Unit) サイズが、伝送速度に大きく影響することがわかった。この MTU サイズは、フレームの最大長を表すものでこのサイズによっても、伝送速度に大きな変化があった。というのは、今回の実験に先駆けネットワークベンダの協力により、LANE を用いた通信方式では MTU サ

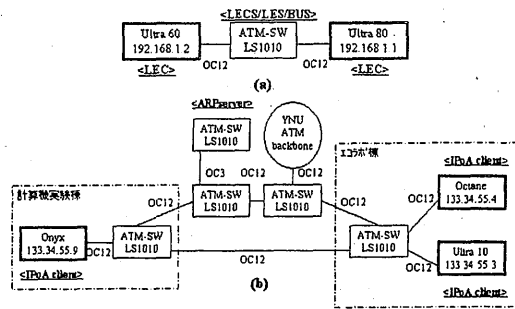


図 5: ネットワーク構成図

イズに大きく依存することが明らかになった。これまで、LANE による通信方式では上記の環境で MTU サイズ 1500[Kbyte] において、伝送速度の最大値は 270[Mbps] であったのに対し、MTU サイズ (ATM スイッチ, ATM-NIC 両者とも) を 9234[Kbyte] に設定し、ATM スイッチを一つはさんだ通信では、最大 500[Mbps] の値を示した。この実験環境のネットワーク構成図 5(b) 示す。このようなことから、MTU サイズを 9234[Kbyte] に設定し上記の環境で LANE 通信方式に適用しようとしたところ、ATM スイッチの OS のバージョンの違いにより相互接続を確立できなかった。

そこで、IPOA を用いた通信方式に変更して実験を行なうことにした。IPOA においても MTU サイズに大きく依存することがわかったため、MTU サイズを ATM-NIC では 9234[Kbyte]、ATM スイッチにおいては 4470[Kbyte] に設定し伝送速度の測定をおこなった。IPOA においては ATM スイッチの OS のバージョンによる問題もなく、相互接続が可能であったため伝送速度の測定を達成することができた。その伝送速度測定による実験結果を図 6,7,8,9,10,11 に示す。この伝送速度の実験結果から、1 チャンネルの映像データ (D1 映像) などのマルチメディアデータを十分に伝送可能な帯域を確保することができると考えられる。

次に、この伝送帯域を用いてリアルタイムに映像の伝送を行ない、フレームドロップが生じずに蓄積することができるかを確かめる実験を行なった。この実験には映像として D1 カメラ (表 2) で撮影している映像データを使用し、先ほどの実験結果をもとに IPOA 方式を用いて伝送し高速大容量ストレージ

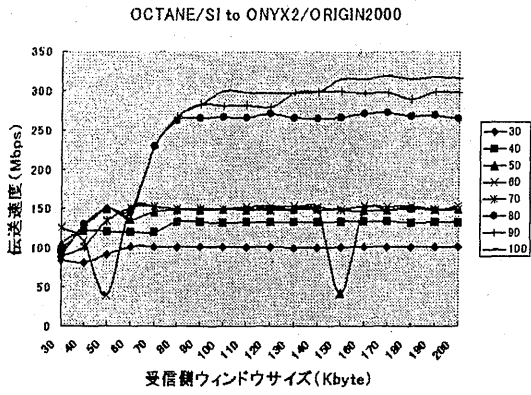


図 6: 測定結果 1

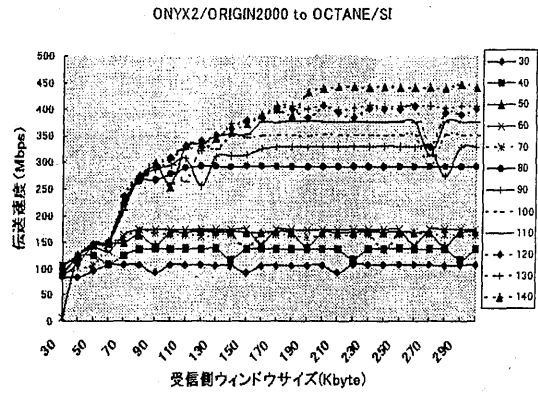


図 9: 測定結果 4

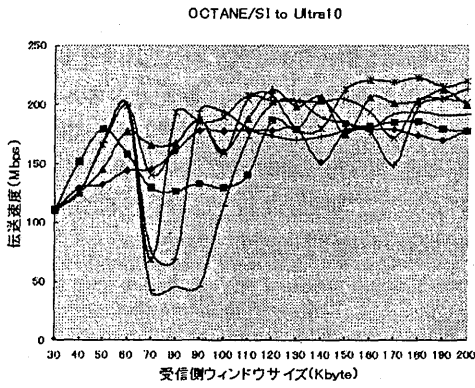


図 7: 測定結果 2

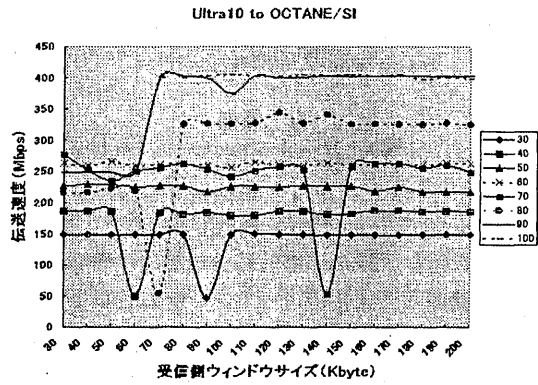


図 10: 測定結果 5

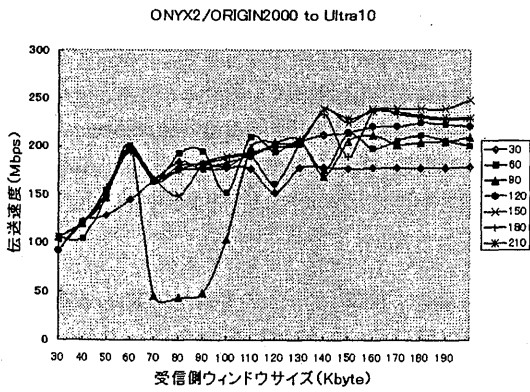


図 8: 測定結果 3

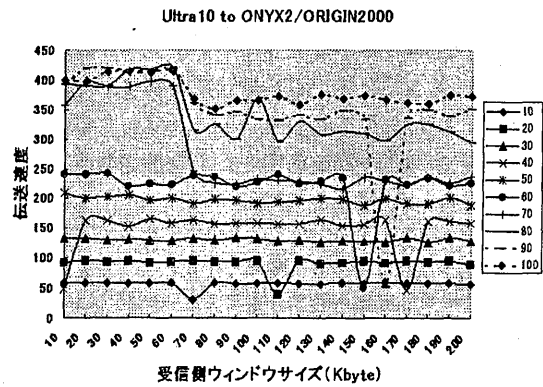


図 11: 測定結果 6

ワークステーション環境	
Machine name	SGI ONYX2/ORIGIN2000
CPU	260[MHz]MIPS RISC R10000(× 8)/ 190[MHz]MIPS RISC R10000(× 48)
Memory size	4[Gbyte]/16[Gbyte]
Disk size	162[Gbyte]/117[Gbyte]
ATM-NIC	FORERunner 200E
ネットワーク環境	
Machine name	SGI OCTANE/SI
CPU	195[MHz] MIPS RISC R10000
Memory size	128[MByte]
Disk size	9[Gbyte]
ATM-NIC	FORERunner 200E
Sun Ultra 10	
Machine name	Sun Ultra 10
CPU	300[MHz]UltraSPARC-III
Memory size	256[Mbyte]
Disk size	2.5[Gbyte]
ATM-NIC	FORERunner 200E
ネットワーク環境	
ATM-Switch	LS1010(LightStream1010)
OS	TOS ver.11.2.8(16MDRAM)

表 1: 実験に用いた環境

frame rate	30[frame/sec]
frame size	699840[byte]

表 2: D 1 映像

である Onyx2/Origin2000 を用いて蓄積した。また、この実験を行なう際には事前に伝送前の映像を用意し、その映像との比較によってフレームドロップが生じているかどうかを調べた。次に、チャンネル数を増加させた場合の実験を行なった。この測定方法としては、ある時間内に届くはずのフレーム数で実際に届いたフレーム数で割った時の値(× 100)を測定値とする。(図 6 の測定で用いたワークステーション)間で行ない、チャンネル数による影響が映像伝送に及ぼすかどうかを確かめた。その結果を図 12 に示す。この図からわかるように、チャンネル数が増えることによってフレームドロップの割合が増加していることがわかることから、ネットワークの帯域に応じてチャンネル数を決定する必要があることがわかる。

5 まとめ

本稿では、次世代放送において求められるマルチチャンネル TV を実現する MCVDB システムの設計を行なった。また、この MCVDB では多チャンネル映像のレコーダとしてデータベースを用いることを

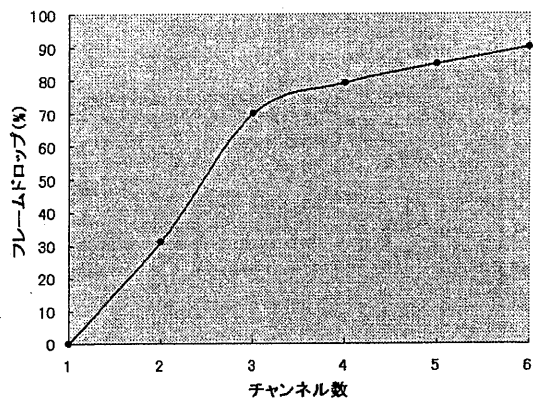


図 12: フレームドロップ

提案した。このシステムにおけるプリミティブの一つである、高速ネットワークの伝送速度測定実験を行なった。さらに、同一ネットワークにおける多チャンネル映像の伝送実験を行ない、多チャンネル映像伝送のフレームドロップ発生率を測定した。これらの実験結果により、非圧縮映像によるマルチチャンネル TV において、要求されるクオリティに対する必要伝送帯域を割り出すことができた。

参考文献

- [1] H. Arisawa, H. Nagae, Y. Mochizuki: "Representation of Complex Objects in Semantic Data Model "AIS" and Implementation of Set Operators", IEICE TRANSACTIONS, vol. E74, No. 1, 1991.
- [2] H. Arisawa, T. Tomii, H. Yui, and H. Ishikawa: "Data Model and Architecture of Multimedia Database for Engineering Applications", IEICE TRANS. INF. & SYST., vol E78-D, No. 11, 1995.
- [3] 石川 英彦, 有澤 博: "データベーススキーマの効果的な図示手法", 情報処理学会データベースシステム研究会, 1996

以下の頁は著作権者の許諾を得ていないため、公表できません。

抜刷 4 ~ 抜刷 6

抜刷 8 ~ 抜刷 8

抜刷 13 ~ 抜刷 13

抜刷 28 ~ 抜刷 28

抜刷 ~ 抜刷

抜刷 ~ 抜刷

骨格モデルを用いた人体動作データベースの構築

佐藤 貴子[†] 永野 志保^{†,☆} 富井 尚志[†]
有澤 博[†] 酒井 直隆^{††}

本稿では人体動作の3次元時空間解析/モデリングシステムの一環として人体骨格モデル (BBHM) を用いた人体動作シミュレーションを行う手法を提案する。人体動作を骨格レベルで取り扱う際に問題となるのが、詳細な人体動作データの取得方法である。これに対し本稿では、動作データを取得することができ、かつ人体運動の特徴を十分に表現可能な構造を持つ、単純化人体モデル (SHM) を用いた動作定義の手法を提案し、BBHMの構造、およびBBHMに対するSHMを用いた動作マッピング手法の妥当性を評価する。

Design and Construction of Human-Body Motion Database Using Bone-Based Human Model

TAKAKO SATO,[†] SHIHO NAGANO,^{†,☆} TAKASHI TOMII,[†]
HIROSHI ARISAWA[†] and NAOTAKA SAKAI^{††}

This paper presents a new methodology of human-body motion simulation with Bone-Based Human Model (BBHM) as a part of spatio-temporal analysis/modeling system. In order to describe detailed human-body motion in skeleton level, bone base modeling is quite useful. Also, we introduce the Simplified Human Model (SHM) as an essence of logical structure of the human-body instead of full bone model. We evaluate propriety of BBHM precision and the way of motion mapping by using SHM.

1. はじめに

近年のコンピュータグラフィックス (CG) 技術の発達により、さまざまなデータを視覚的に表現することが可能となりつつある。なかでも特に、人間のように複雑な形状を持つ3次元物体やその動きをCGを用いて表現する技術の発達はめざましい。しかし、CGでは動画として自然に美しく表現することに主眼を置いているため、CGにおけるモデル化の手法は実世界の「もの」を忠実に再現するという観点で行われているものではなかった。

ところが、実世界の「もの」を忠実に再現したモデルをコンピュータ上に作成することができれば、実世界のオブジェクトやその振舞いに対する非常に正確なシミュレーションや評価を行い、その結果を視覚的

に表現することが可能となる。そこで、我々は、コンピュータ上にグラフィックス的に表現され、3次元的な大きさを持ち、構造の定義されている物体を3Dオブジェクトと呼ぶことにし、3Dオブジェクトを用いた人体動作のシミュレーションの可視化について考察した。既存のソフトウェア、たとえば、エンジニアリングシミュレータにおいても、人間の3Dオブジェクトを用いて、工場作業を行った場合に身体の各部位にかかる負荷の評価などが行われている。しかし、簡単な構造の3Dオブジェクトを用いているため、評価できる内容は限られていた。これに対し我々は、骨格など、人体の複雑な構造に基づいた動作評価を行いたいと考えた。これにより、工場作業の負荷評価を行う際、個人差や年齢差による負荷のかかり方まで評価することが可能である。さらに高度な人間の動作に対するシミュレーション・評価の例としては、ダンスなどの表現芸術において、人間の体の微妙な動きがどのような骨の動きにより表現されているかを解析し、無理がなく表現できるような骨の動きのメカニズムを評価する、といったことがあげられる。また、リハビリテーション医学の分野では、骨格的な個人差が疾患を治療する

[†] 横浜国立大学工学部電子情報工学科

Department of Electrical and Computer Engineering,
Faculty of Engineering, Yokohama National University

[☆] 現在、ソニー木原研究所

Presently with Sony-Kihara Research Center, Inc.

^{††} 横浜市立市民病院

Yokohama Municipal Citizens Hospital

際に影響を及ぼすため、シミュレーションが個人に合わせた効果的な治療方法を見出すことにも役立つ。

一方、これらの動作評価を行ううえでは、データベースの役割が重要となる。人体の形状や構造、動作、特性などを、単に3Dポリゴンの集まりや位置・角度の時系列、数値などで表すだけでなく、統一的なスキーマのもとに蓄積することで、人間が共通に持っている性質と、各個人により異なる特性とを明示的に切り分けて表現できる。これにより、たとえばある動作に関して「この動作を無理なく行うことができる人は誰か」、「平均的な50代の人がこの動作を行った場合、どの部位に多く負担がかかるか」という問合せを行い、ユーザが自由にデータを組み合わせることでシミュレーションを行うことが可能となる。

しかしながら、これまで述べてきたような動作評価を行う場合には、次のような問題が生じる。

- 使用する3Dオブジェクトの構造は、どこまで詳細にすればよいか。
- 複雑な構造を持つ3Dオブジェクトに対し、人体の動作をどのように定義(移植)するか。
- データベース化の際、オブジェクトの形状、構造、動作をどのようにモデリングすればよいか。

以上から、本稿では、動作評価を行ううえで適切な詳細度と考えられる3Dオブジェクトとして、人体の骨格構造および表面形状に基づいたモデル化・データベース化の手法であるBone-Based Human Model(BBHM)を提案する。また、BBHMに動作を定義するための手法として、単純化人体モデル(Simplified Human Model: SHM)による大局的な動作定義、およびマッパーによる動作データの詳細化の手法を導入したうえで、BBHMの妥当性について評価する。

2. BBHMの構成

2.1 BBHMの導入の理由

前章で述べた動作評価を行うためには、それに見合う詳細度で人体をモデル化しなければならない。

人体の動作を表現するための人体の3Dオブジェクト化には、次のような方法があると考えられる。

- (1) 人体の表面形状のみを使って表現する方法
- (2) 人体の表面形状(皮膚)のみならず、骨格的な構造やその動きの制約条件を表現する方法
- (3) 内臓、血管、神経も含めた人体構造を完全に忠実に表現する方法

(3)のような非常に精密な3Dオブジェクトは、手術の事前シミュレーションなどの医学分野への応用を期待できるが、このレベルでの人体のモデル化は、CT

やMRIなどの特殊で医学的な装置を用いても非常に困難であり、現時点では実用的ではない。また、(1)のような人体の外形のみを表現した3Dオブジェクトでは、前章に述べたように、人体の複雑な構造に基づく評価を行うことができない。これらに対し、(2)のような骨格構造と表面形状を表現した3Dオブジェクトでは、関節可動域(関節の可動範囲)に基づく評価など、(1)に比べはるかに精密な生理学的な動作評価が可能になる。たとえば人体は、つま先を前に向けたときと外に向けたときで脚を側方に上げられる量が異なる。この違いは大腿骨と骨盤との衝突により生じるものであり、骨格の形状や構造までをモデル化した3Dオブジェクトでなければ評価できない。このような評価は、ダンスなどの詳細度の高いシミュレーションを可能とする。また、(2)の手法は、(3)と異なり、現実的な手法で3Dオブジェクトを生成可能であると考えられる。

これらから、我々は、(2)の手法が動作評価などを行ううえで最適であると考え、表面形状と内部構造とをあわせ持つ3Dオブジェクトを作成することとした。

ところで、(2)の手法においても、ただ表面形状と内部構造とが表現されていればよいというものではない。骨格の動きは、骨自体の形は変わらず、骨の位置関係のみが変化するため、剛体リンクモデルで表現できる。しかし、表面形状は皮膚そのものの形が変化するため、剛体リンクモデルでは動きを正確に表現することができない。そのため、表皮に関しては、軟体モデル(部品の形状が変化するモデル)で表現することが妥当と考えられる。

さて、人体の詳細なモデル化の例として、VRMLヒューマノイドアニメーション(H-Anim)があげられる³⁾。H-Animは脊椎や手指の関節などについては、人体の骨格構造を反映した関節数を用意している。しかし、そもそも見せることを目的としたモデル化が行われており、構造が詳細であるとはいっても医学的なシミュレーションに耐えうるほどの精度は持っていない。また、表皮についても関節を用いて表現するため、実際の人体の表皮に基づいたモデル化を行っているとはいえない。

2.2 BBHMの提案

動作評価を行うには、内部構造を骨の剛体リンクモデルで、表面形状を軟体モデルで表した3Dオブジェクトが適していると述べた。我々はこのようなモデルとして、BBHMを提案する。

BBHMは、人体の骨形状を表した部品、および軟体モデルで表現された表皮からなる3Dオブジェクト

であり、部品数 155 個、関節数 76 個、126 自由度を持つ。部品どうしは実際の関節の動きに基づいた可動性の関節で接続している。この関節は、呼吸時の肋骨の動きなどを除けば、人体の骨格の動きはおおよそ表現することができる程度に精密に人体の関節を表している。このような人体モデルがあれば、リハビリテーションに関する治療計画や工場などでの作業設計を行ううえで、個人の体格や関節の丈夫さを考慮するほど精密さが問われるような負荷評価を行うことができ、有用である。

しかし、ここで大きな問題となるのは、BBHM に動作を定義することが、実際には非常に困難だということである。剛体リンクモデルで表現された 3D オブジェクトに対する動作の定義は、一般に、すべてのリンク部分に対し、角度を定義することにより行われている。しかし、BBHM は、人体の骨格構造を反映した複雑な構造を持っているため、全関節に対し時系列的に角度を定義することは非常に多くの労力を要する。たとえば、脊椎は人体に 24 個あり、それぞれの間の関節が少しずつ曲がることで身体全体をひねったりそらしたりする。BBHM も脊椎を 24 個の部品で表しているが、このとき、脊椎の間の関節 1 つ 1 つに対して角度を定義することは困難である。

我々はこの問題を、単純化人体モデル (Simplified Human Model: SHM) を用いた動作定義およびマッピングの手法により解決することを考えたが、この手法については 3 章で述べることにする。

2.3 人体構造のデータベース上での表現

BBHM では、人体を骨格レベルで精密にモデリングしなければならないが、この際、従来の人体モデリングの手法を適用することができない、という問題がある。

従来、エンジニアリングシミュレータなどの人間 3D オブジェクトは、人体を形状の変化しない剛体部品に分け、それらに親子関係を持たせて、階層構造で表されていた。これでは、ある部品が親部品を複数持つことはできない。しかし、人体の骨格の構造は複雑であり、たとえば前腕部にある橈骨と尺骨は、肘関節と手首の両方で拘束条件を持つため、階層構造では「合流」を生じてしまい、表すことができない。

ここで、骨格レベルで人体をモデリングする際に最も本質的なことからは、骨と関節とが互いに独立した要素であり、それらのつながりによって人体が構成されているということである。

また、骨格レベルで人体を見ると、肘関節のように、1 つの関節に 3 つ以上の骨が接続しているような場合

も多数見られる。そのため、骨の接続する位置関係や可動範囲などを関節に一意に与えることは妥当でない。

この問題は、1 つの関節は「骨と関節とのつながり」という、より原子的な「接続」により構成されている、ととらえることにより解決できる。この原子的な接続を、connection と呼ぶこととする。たとえば肘関節は、肘関節と上腕骨の接続、肘関節と橈骨の接続、肘関節と尺骨の接続という 3 つの connection で表される。

ところで、このような接続の表現をデータモデリングの観点から見ると、connection は、1 つの骨と 1 つの関節という独立した要素が組み合わせられることにより生じる要素であり、直積により表すことが本質的であるといえる。

さて、これまで述べてきた人体の構造は、人間誰でもに共通である。しかし、人体の構造には誰にも共通でない性質もある。たとえば、骨が関節で接続する位置、可動範囲などは、一般的にどれくらいであるかの目安はあるものの、人により異なった値を持つものである。ここで、「ある 1 人の人」と「ある骨 (あるいは関節, connection)」とを組み合わせると、「ある人のある骨」という 1 つのものを表すことができることから、そのような人により異なる性質を表現するためにも、直積が有効ではないかと考えられる。

以上のように、人体の詳細な構造をデータベース化するためには、直積を分かりやすく表現できるデータモデリング手法に基づく必要がある。

2.4 AIS モデルと直積を用いた表現

人体の骨格構造やその個人差が直積としてとらえられることは示したが、それでは直積を表すのに最も直観的な表現手法とはどのようなものだろうか。

直積は、複数の集合から 1 つずつ要素を取り出し、その要素を組み合わせたものをまた 1 つの要素と見なすという概念である。そのため、直積を表現する際には、各集合から取り出した要素が何であり、それを組み合わせた要素が何になるのか、を分かりやすく表現できなければならない。また、集合から取り出した要素どうしは対等な関係であることから、一方の要素の属性として他方の要素を表すような表現方法は望ましくない。

これらの理由から、我々は、直積を直観的に表すことができ、平坦な構造を持つ、拡張 ER モデル (associative information structure: AIS^{4),5)} が適していると考えた。

AIS は、主体 (entity) と主体間のつながり (association) により情報を表現しており、主体は表す意味によって主体型 (entity type) に基づく集合に属する。

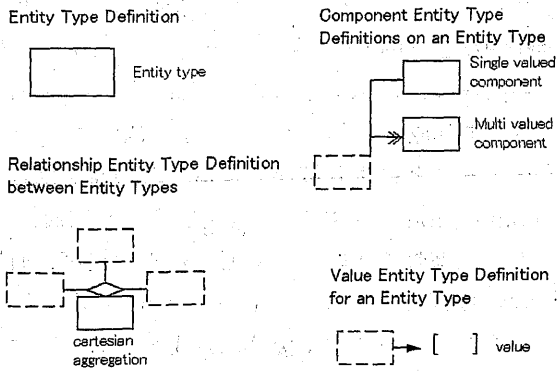


図1 AISダイアグラムの基本要素
Fig.1 Primitives of AIS diagram.

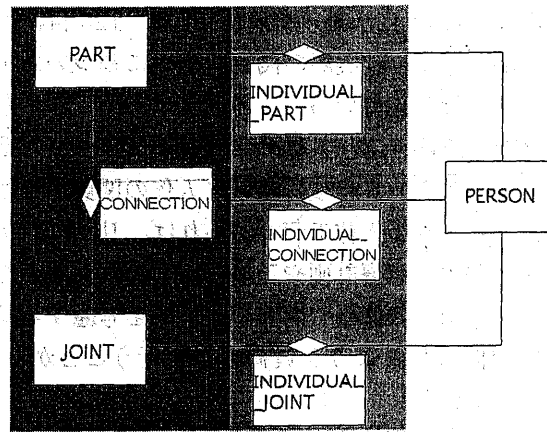


図2 骨格構造のデータベース上での表現
Fig.2 Expression of bone structure on database.

これを主体集合と呼ぶ。主体は内部構造を持たない1つの点であり、現実世界の事物・事象はすべて主体である。また、数値、文字列などの値も主体としてとらえられ、値主体 (value entity) と呼ばれる。

異なる主体集合の直積集合の要素もまた主体であり、関連主体 (relationship entity) と呼ばれる。また、主体間の親子関係を表すものは成分主体 (component entity) と呼ばれ、特に子が複数の主体列になっているものを多値成分、そうでないものを単値成分という。

AISに基づくスキーマは、ダイアグラムにより表記される。ダイアグラムの構成要素を図1に示す。

主体型を矩形で表し、主体型間の直積は主体型どうしをつなぐ線分上に菱形で表す。このとき、関連主体型は、菱形に矩形をつなげることで表す。また、成分主体は、親となる主体型の矩形の辺の延長線から直角に子となる主体型へ線分を伸ばすことで表す。多値成分の場合はこの線分を2重矢印として表す。また、値主体は矢印でその主体のデータ型を示す。

インスタンスを示す場合には、主体はその属する主体型の矩形中に小丸で表し、つながりは主体どうしを結ぶ線分で表す。

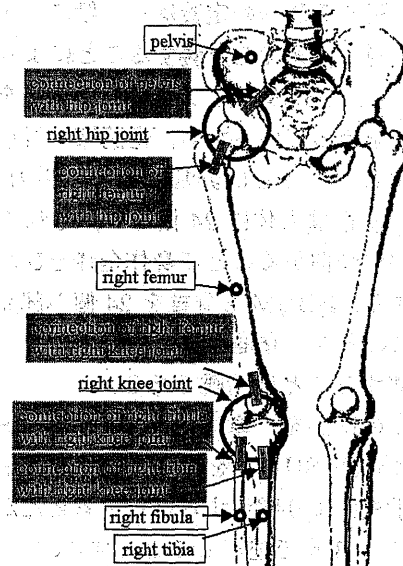


図3 人体の骨格構造
Fig.3 Bone structure of human-body.

2.5 BBHMのスキーマ表現

以上の考え方のもとで、人体構造の骨子を表していると考えられるBBHMをデータベーススキーマで表したものを、図2に、また図3の部位についてインスタンスを示したものを図4に示す。

人体の独立した要素である骨と関節は、それぞれPARTとJOINTという主体型で表され、これらの直積CONNECTIONでconnectionの集合を表す。このPARTとJOINT、CONNECTIONのつながりによって、一般的な人体の構造が表されている。

たとえば、right_femurはright_hip_jointやright_knee_jointとの間にconnectionを持つとか、right_knee_jointはright_femur、right_fibula、right_tibia

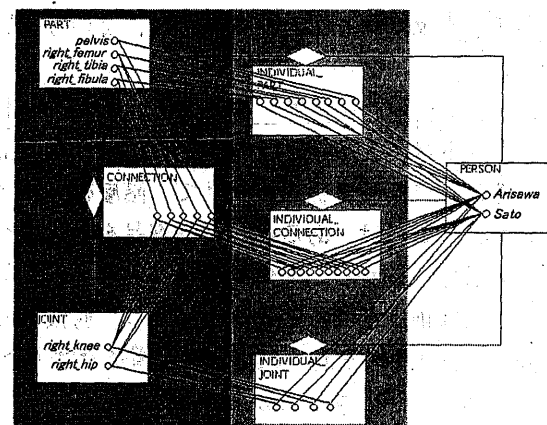


図4 骨格構造のスキーマ表現のインスタンス例
Fig.4 Instance examples of bone structure.

との間の3つのconnectionで構成されることなどが表される。

さて、人体の構成部品 (PART) と、各個人 (PER-

SON) という独立の主体集合の直積により, ある人のある骨を表すことができる。

骨についてとりあげてみると, PERSONとPARTの直積により INDIVIDUAL_PART すなわち誰かの骨を表すことができる。骨に関する個人別の特性(たとえば長さ)は, INDIVIDUAL_PARTの属性である。たとえば, Satoのright_tibiaの長さは, Satoとright_tibiaにつながる INDIVIDUAL_PARTの要素(entity)の属性として表される。JOINTやCONNECTIONに関しても, 同様にして考える。

以上により骨レベルの人体構造を, 平坦な構造を持つデータモデル AISのスキーマで表現できることが示された。付加属性などの詳細情報を含んだスキーマ解説は付録 A.1 に譲る。

3. SHMによる動作定義とマッパー

これまで述べてきた BBHMは, さまざまな動作を行わせることで, 人体各部の負荷の状態を正確に表現することが可能である。しかしここで, 関節自由度の総数が非常に多いため, 評価したい動作を忠実に BBHMにマッピングすることが困難だという問題が出てくる。

しかし, 医学分野においては, 人体の大局的な動きに対して各関節がどのように動くかの割合が研究されている⁶⁾。この割合から, 関節ごとに, 大局的な動きのうち何%を担うかを示す重み付けパラメータを算出できる。このようにすると, たとえば脊椎では, 脊柱1つ1つの間の角度を脊柱全体の動きから求めることができる。さらに, 腰に対する肩の位置や, 肘, 膝の動きといった表面形状からも分かるような大局的な動きに関しては, 映像から各関節の角度を取得する研究がなされている^{7)~9)}。

これらの知識・技術を用いることで, 評価したい動きを忠実に BBHMにマッピングすることが可能である, と我々は考えた。

そこで本章では, 人体の大局的な動作を定義するためのモデルとして, 単純化人体モデル (SHM), および, SHMを用いて定義した動きを重み付けパラメータなどを用いて BBHMにマップする手法として, マッパーを提案する。SHMの構成を図5に示す。

SHMは28部品27関節からなる剛体関節モデルであり, 全体で59自由度を持つものとする。この関節数は, 人体の大局的な動きを十分に表現することができ, しかも, 映像からの解析で動作データを取得できると考えられる範囲である^{7)~9)}。たとえば, SHMでは上体の回転や曲げ伸ばしを腰・胸にある2つの関節

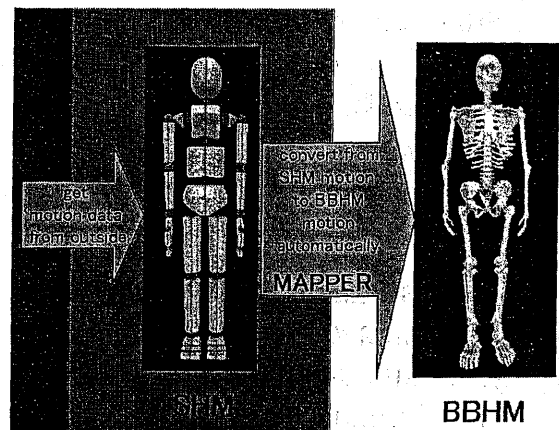


図5 SHMの構成図

Fig. 5 Structure of SHM.

として考えるが, これは実際の人体における腰に対する肩の位置から算出するものとしており, 重み付けパラメータなどから BBHMの脊椎の関節角度を導出できる。

マッパーは, SHMで定義された人体の大局的な動きを, 医学的な知識に基づき, BBHMレベルの詳細な動作データに変換するものである。関節数の少ないモデルから多いモデルに動作データを変換することを考える場合, 当然ながら変換前の動作データが実際の人体の動きを忠実に表現できていたとしても, 変換した後の動作データが実際の動作と異なってしまうのではないかと懸念される。しかし, 医学的知識や重み付けパラメータ, 重さ, 可動域, 筋や腱など, 人体の運動を制約するさまざまな要因を基に個々の関節の動きを算出することにより, 変換後の動作データで実際の骨格の動きを再現することは可能である, と我々は考えている。

しかし, これらの要因は互いに影響を与えたり連鎖して変化したりするため, 非常に複雑であり, 一度にすべての要因に対するモデル化を行うことは困難である。そこで我々は, 段階を踏んで各要因のモデル化を行うことを考え, まず第1段階として, 重み付けパラメータをマッパーで用いることとした。たとえば, 腰椎には関節が5つあるが, 腰を屈曲させると, 一般にこの5つの関節は6:8:9:12:8の割合で曲がるということが知られており⁶⁾, この割合から重み付けパラメータを算出する。重み付けパラメータが個人により決まっていること, 一般的な値や年齢, 病気によりどのように変化するかが知られていること, さらに具体的な測定手法に関する研究が医学分野で行われていること^{10),11)}から, このような非常に簡単な手法によっても, 文献6)に記されている動作の再現までは可能である。

ここで、SHMの関節角度の比から重み付けパラメータを算出する手法、および重み付けパラメータから関節角度を求める手法は、以下のとおりである。

$$P_{ij}(\%) = \frac{R_{ij}}{\sum_n R_{nj}} \times 100 \quad (1)$$

$$B_j = S_i \times P_{ij} \quad (2)$$

- B_j : BBHMのj番目のjoint角度
- S_i : SHMのi番目のjoint角度
- P_{ij} : SHMのi番目のjointからBBHMのj番目のjointへの重み付けパラメータ
- R_{ij} : SHMのi番目のjoint角度がBBHMのj番目のjoint角度に影響する比

先ほどの腰椎の例について実際に計算を行うと、屈曲運動に関して、腰椎第1関節のパラメータは、

$$\frac{6}{6+8+9+12+8} \times 100 = 14.0(\%) \quad (3)$$

また、腰を30°屈曲させた際の腰椎第1関節の角度は、

$$30 \times \frac{14.0}{100} = 4.2(^{\circ}) \quad (4)$$

となる。このようにして算出した、一般的な人体における重み付けパラメータを、付録B.1に示す。

ここで、SHMの動作データのスキーマ表現を図6に示す。SHMの動作データからBBHMの動作データへの変換はあくまでマッパーによるものであり、スキーマ上では両者は完全に独立している。動作は姿勢の時系列と考え、各関節の角度で姿勢を表す。これは、「ある時点」と「ある関節」の組合せごとに角度を与えることで表現できることから、動作もまたTIME_POINTとSHM_JOINTの直積により表すことができる。なお、関節は自由度によっては1方向だけでなく多方向に動くが、これは医学での運動方向の分類にならない、屈曲・伸展(Flexion/Extension)、内転・外転(Adduction/Abduction)、回旋(Rotation)に分けて考える。これらをFLEX, ADC, ROTと略す。また、重み付けパラメータはBBHMのCONNECTIONごとに決まっており、個人により異なる値なので、図2のINDIVIDUAL_CONNECTIONの属性として蓄積される。

この2つのデータベースを用いることで、ある個人が何らかの動作をした場合だけでなく、世代などによる動作の傾向もシミュレーションすることができる。たとえば、冒頭に述べた「50代の人がある動作を行った場合、どの部位に多く負担がかかるか」というシミュ

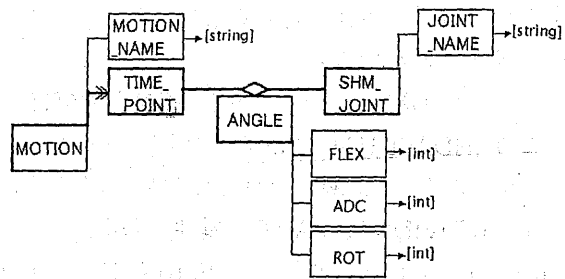


図6 SHMの動作のスキーマ表現
Fig. 6 Scheme of SHM motion.

レーションの手順の例を以下に示す。

- PERSONから、年齢が50~59歳の人を選択する。
- INDIVIDUAL_CONNECTIONをたどり、各connectionごとに重み付けパラメータと関節可動域の平均を求め、マッパーに渡す。
- MOTIONから、シミュレーションしたい動作を選択する。
- TIME_POINTをたどり、すべての関節角度を取り出してマッパーに渡す。
- データベースから取り出した重み付けパラメータと関節角度時系列から、BBHMレベルでの動作時系列に変換し、この動作データと関節可動域平均をシミュレータに渡す。
- シミュレータでマッパーから渡された動作データと関節可動域を比較し、各時点で関節角度が関節可動域の何%を占めているかに応じ骨に着色を施したシミュレーションを行う。

4. BBHMを用いた動作評価の実現と評価

4.1 人体動作評価の目的

本稿における人体動作評価に関する具体的な要求として以下の3点がある。

- (1) ある人の実際の動きに対する詳細なシミュレーション
(スポーツ疾患を持つ人のリハビリテーションの効果シミュレーションなど)
- (2) ある人の動作を他の人が行った場合のシミュレーション
(アマチュアダンサーがプロのダンサーの動きを真似た場合の限界など)
- (3) ある集団の動作の傾向の評価
(高齢者用住宅の空間設計のためのシミュレーションなど)

前章までに提案した概念に基づきこれらの要求を満たすシステムとして、我々はBBHMシミュレーションシステムを米国deneb社のエンジニアリングシミュ

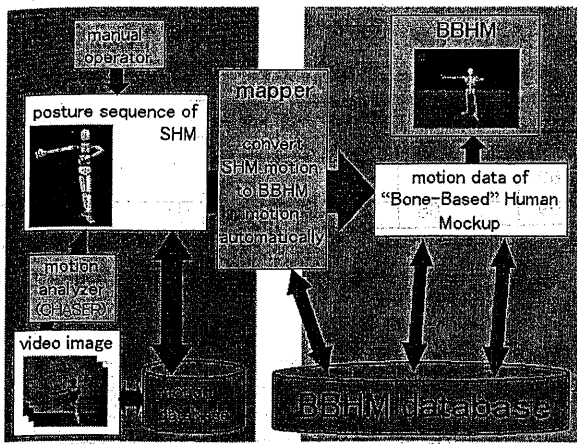


図7 シミュレーションシステムの構成図

Fig. 7 Architecture of simulation system.

レータ ENVISION version 5.0¹⁾上で実装した。なお、今回の実装システムでは、表皮については、不完全であるが形状の変化は考えず、剛体モデルにより表現した。このシステムの流れは図7に示す。

このシステムは、SHMで動作を定義し、その動作データをマッパーを用いてBBHMにマップし、シミュレーション・評価を行うものである。動作データや重み付けパラメータなどの値はスキーマに従ってデータベースに蓄積してあるため、データベースに問合せを行うことにより、ある個人やある集団(40代の人など)が動作を行った場合の負担評価など、さまざまなシミュレーションを行うことが可能である。

我々は実際に行われた動作と架空の動作の両方に対して評価を行いたいと考えている。これらを実現するため、SHMへの動作定義は以下の2つの手法をとることとした。

- (1) 映像から動作データを取得する。
- (2) SHMに対し動作データを3次元マウスで直接定義する。

また、各PARTの大きさや関節可動域、重み付けパラメータなどをデータベース上にあらかじめ蓄積しておき、このデータを用いて個人に合わせたシミュレーションなどを行う。なお、重み付けパラメータ設定機で設定する重み付けパラメータのデフォルト値は文献6)の一般的な人体における特性に基づき設定した。

4.2 BBHMデータの作成

BBHMの骨格部品の形状は、実際の人骨の形状を、非接触3次元スキャナを用いて取得した。この際に使用したスキャナは、376×281(mm)、撮影対象物まで326(mm)、解像度はx方向500 μ m~2mm、y方向313 μ m、z方向50~200 μ mである。

4.3 動作シミュレーション

BBHMシミュレーションシステムでの動作評価の

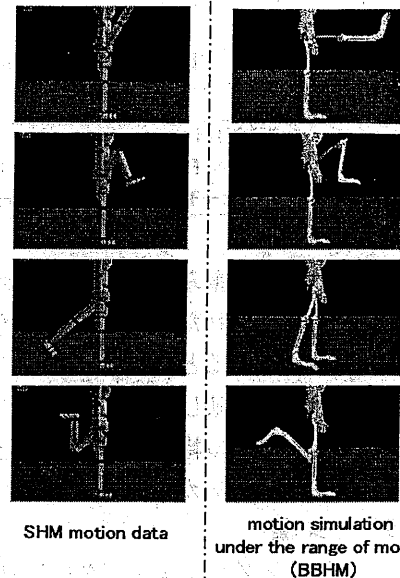


図8 関節可動域の変化に関する評価

Fig. 8 Evaluation of change of range of motion.

例として、動作データと関節可動域との比較を行った。関節可動域を越えた動きは行えないことはもちろんのこと、可動域ぎりぎりまで関節を曲げた方が負荷が大きくなることは容易に予測できる。これより、関節可動域を越えた動作データがマップされるとその部位が着色され、可動域内での動きに変更して表示するような動作評価機能を実装した。この機能を用いて、関節可動域の広いプロのダンサーが行った動きを素人ダンサーが行おうとした場合に負荷がかかる部位のシミュレーションを行った。

また、この評価の際、膝の角度変化に応じて腿の関節可動域が変化する様子についても動作シミュレーションを行った。この様子を図8に示す。

4.4 BBHMシミュレーションシステムの評価

本稿で提案したBBHMシミュレーションシステムの妥当性を評価するためには、少なくとも以下の3点について評価を行う必要がある。

- (1) BBHMの妥当性
実際の人体を十分に表現する能力があるか。
- (2) SHMの妥当性
人体動作を定義する能力があるか。
- (3) マッパーの妥当性
マッパーによる手法が適切なものか。

本稿では、このうち主に(1)について評価を行った。(2)については、医学書¹²⁾における四肢(両手足)および体幹(胴)の運動を考える際の、骨格の分類を参考にしていることから、SHMの構造はほぼ妥当であると判断される。(3)については、現在のところ、一般的な人体の動きに関しては、医学的な調査により得

表1 人体3Dオブジェクトのデータ量の比較
Table 1 Comparison of data quantity of human-body 3D object.

	部品数 (個)	自由度数	ポリゴン数 (個)
BBHM	155	120	79514
MAN_P50 (ENVISION/Deneb)	62	50	3518
Jack (Transom Jack/EAI)	69	68	2239

られたデータ⁶⁾に基づき重み付けパラメータを設定していることから、変換されたデータは一般的な人体の骨格の動きを表現していると考えられる。しかし、個人により異なる特性を反映した骨格の動きの再現については、具体的にそれらの特性を測定するためにどのような手法を用いたらよいか、現在検討中である。

4.4.1 データ量

従来エンジニアリングシミュレータなどで用いられてきた人体3Dオブジェクトと作成したBBHMを比較したものを表1に示す。今回BBHMと比較した3Dオブジェクトは、ENVISIONで用いられているMAN_P50、および米国EAI社の人間動作シミュレーションソフトTransom Jack²⁾で用いられているJackである。

今回作成したBBHMは、ポリゴン数が多いものの、部品数、関節自由度の数という面から見てエンジニアリングシミュレータの人体3Dオブジェクトより詳細に人体の姿勢や動作を表現することができる。また、人体の内部構造と外部構造とを兼ね備えているため、人体を外から見た動きと内部での動きとを同時に見ることが可能である。

4.4.2 動作表現に関する比較

次に、MAN_P50と動作表現について比較を行う。

胴体をひねる場合、3章でも述べたとおり、脊椎の間の各関節の回転する割合は決まっている。一般的な3Dオブジェクトでは、そのような割合はまったく考慮されていないが、これでは腰に疾患を持つ人などの正確な動作評価は不可能であるといえる。MAN_P50においても、胴体のひねりは6関節が同じ割合で回転することで表現されている。しかし、今回作成したBBHMでは、脊椎の間の関節1つ1つに対し回転角度を定義でき、さらにマッパーにより動作データを作成することにより、各関節の回転する割合も表現可能である。胴体全体を90°ひねったときの脊椎の各関節角度を下から順に足していった値を、実データ、BBHM、MAN_P50で比較したものを図9に示す。ただし、実データは文献⁶⁾より得た。また、MAN_P50のデータは関節数が実際の脊椎と比べ非常に少ないので、位置からおおよそ脊椎のどの関節にあたるかを決めた。なお、横軸のLxは腰椎第x関節、Tyは胸椎第y関節を

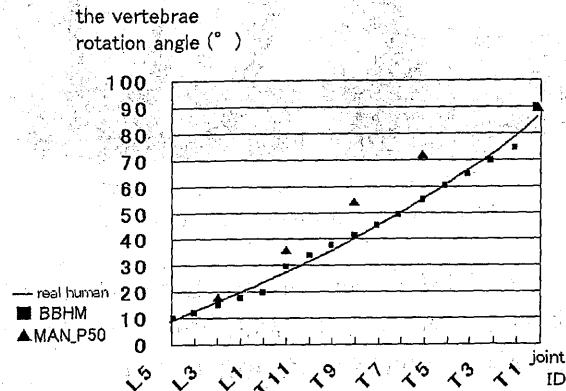


図9 上体を90°ひねった時の各関節角度
Fig.9 Joint rotating angles for 90 degree upper body twisting.

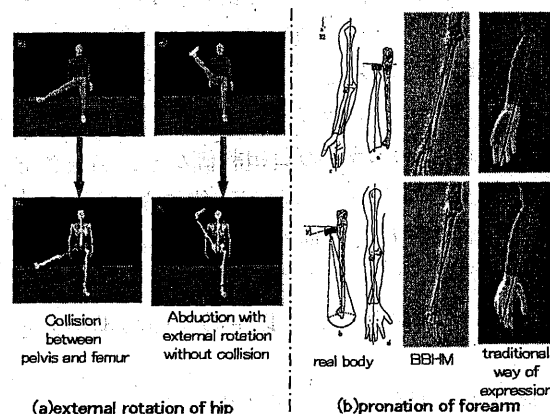


図10 大腿骨と橈骨・尺骨の動き
Fig.10 Motion of femur and radius/ulna.

示している。このグラフより、BBHMはMAN_P50などの一般的3Dオブジェクトに比べ、人体の動きの正確な表現に長けているといえる。

また、骨どうしの干渉検出を行うことにより、関節の動きの骨格的な限界を表現した。これにより、大腿の外転での、外旋を行っているか否かによる可動域の変化も表現できるようになった。衝突検出によるシミュレーションの様子を図10(a)に示す。

さらに、前腕をひねる場合について考えると、従来は前腕を1つの部品として考え、前腕の中心を通る軸を中心とした回転として表現していた。しかし、実際は前腕は橈骨と尺骨という2つの骨からなり、尺骨を軸にして橈骨が回転することにより前腕をひねることになる。BBHMでは、このような動きを表現するこ

とが可能である。この様子を図 10 (b) に示す。

5. ま と め

人体の動きに対する解析・モデリングの手法として BBHM, および BBHM に対する動作定義を行うための SHM とマッパーを提案した。また、データベース上で人体に関する詳細な表現ができ、データベースを用いることでさまざまな人体動作シミュレーションを行うことができることを示した。

今後の課題としては、骨格以外の人体の構成要素(腱, 筋肉など)の実装, インバースキネマティクスに基づく動作解析, 関節可動域以外の負担や疲労度の可視化などがあげられる。腱, 筋肉などといった他の人体の構成要素を表現することにより, より精密な人体動作評価を行うことが可能となり, また疾患や関節可動域を制限する要因を視覚的に理解する手がかりともなる。また, 今回は表皮も剛体モデルで表したが, 形状の変化を軟体モデルとしてどのように表現するかも課題として残されている。さらに, インバースキネマティクスと医学的な知識を併用することにより, 外部から見ただけでは容易に動きの分からないような部位に対する関節角度値の取得が可能となる。このようにして BBHM や人体動作データの取得手法のさらなる詳細化を行うことにより, 多角的な人体動作評価が可能となり, それらの結果の 3D オブジェクトでの視覚的な表現が期待できる。

また, 集団に対する動作評価を行ううえでは, どの動作データが指定された動作を示しているのか, さらに, 比較対象となっている姿勢をとっている時点がどれか, ということを知る必要がある。そのため, 動作データに対するインデクシングの手法を確立することも今後の課題である。

参 考 文 献

- 1) <http://www.delmia.com/solutions/html/ergonomics.htm>
- 2) <http://www.eai.com/products/jack/>
- 3) <http://ece.uwaterloo.ca/~h-anim/>
- 4) Arisawa, H., Tomii, T., Yui, H. and Ishikawa, H.: Data Model and Architecture of Multimedia Database for Engineering Applications, *IEICE Trans. Inf. & Syst.*, Vol.E78-D, No.11, pp.1362-1368 (1995).
- 5) 富井尚志, 有澤 博: マルチメディアデータベースにおける映像モデリングと操作言語, 電子情報通信学会論文誌, Vol.J79-D-II, No.4 (1996).
- 6) カバンディ: 関節の生理学, 医歯薬出版(1986).
- 7) 奥井宏昌, 岡本陽介, 今井さやか, 富井尚志, 有

澤 博: 工場作業データベース設計支援のためのビデオ映像に基づく動作解析, 情報処理学会研究報告, Vol.98, No.58, 98-DBS-116-58, pp.241-248 (1998).

- 8) 坂木和則, 今井さやか, 富井尚志, 有澤 博: 多視点映像を用いた時空間オブジェクトの解析とモデリング, 電子情報通信学会技術研究報告(第 11 回データ工学ワークショップ, DEWS2000), 7B-4 (2000).
- 9) 米元 聡, 有田大作, 谷口倫一郎: 多視点動画映像処理による実時間全身モーションキャプチャシステム—視覚に基づく仮想世界とのインタラクション, 映像情報メディア学会誌, Vol.54, No.3, pp.409-416 (2000).
- 10) 田中直史, 大久保衛, 大槻伸吾: VICON による肩甲骨の動作解析, 日本臨床バイオメカニクス学会誌, Vol.16, pp.217-222 (1995).
- 11) 奥住成晴, 腰野富久: 歩行中の正常腰椎矢状面アライメント, 整形外科バイオメカニクス, Vol.7, pp.155-160 (1985).
- 12) 中村隆一, 齊藤 宏: 基礎運動学第 4 版, 医歯薬出版(1995).

付 録

A.1 データベーススキーマ

今回提案した人体動作データベースでは, 人体の詳細な構造記述をするデータベースと人体の動作を記述するデータベースとはマッパーを介して完全に分離している。人体の詳細な構造記述, すなわち BBHM の構造を記述するデータベーススキーマを図 2 に, 人体の動作を (SHM レベルで) 記述するデータベーススキーマを図 6 に示した。しかし, 図 2 に示したデータベーススキーマは特に根幹となる主要な概念を含む部分のみを示した。そのため, 実際に 3 章で示したような検索を行うためには, このスキーマでは不十分である。

ここでは, BBHM レベルでの人体の構造記述のための詳細なデータベーススキーマを図 11 に示す。

CONNECTION によっては, 1 方向だけでなく複数の方向に曲がるのが可能なものもある。たとえば股関節と右大腿骨の間の CONNECTION では, 前後, 左右, 回転と 3 方向の動きが可能である。CONNECTION の成分主体 DOF は, 関節自由度 (Degree of Freedom) という意味であり, 関節が動くことのできる方向を示す。また, INDIVIDUAL_CONNECTION の成分主体 ROM は, 関節可動域 (Range of Motion) という意味であり, 関節の可動範囲を示す。ROM の成分主体は, 各個人の関節可動域を値主体として持つ。ここで, DOF および ROM の成分主体名の中の FLEX, ADC,

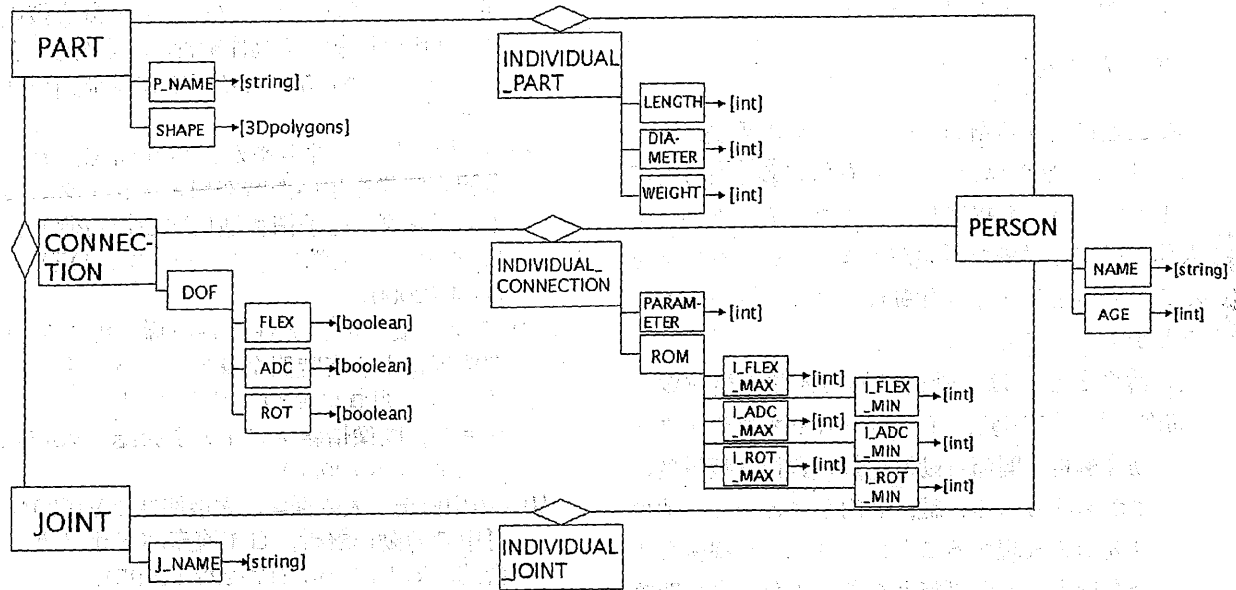


図 11 詳細なスキーマ
Fig. 11 Precise scheme.

表 2 SHM と BBHM での joint の対応
Table 2 Correspondence of SHM joint with BBHM joint.

SHM joint	BBHM joint	normal parameter (%)			SHM joint	BBHM joint	normal parameter (%)		
		Flex	Adc	Rot			Flex	Adc	Rot
back	back_1	14.0	16.1	12.5	R.elbow	R.elbow	100	—	—
	back_2	18.6	25.8	12.5	R.radio-ularnar	R.radio-ularnar	—	—	100
	back_3	20.9	25.8	12.5	R.wrist	R.carpal	100	100	—
	back_4	27.9	25.8	12.5	R.finger	R.MP	100	100	—
	back_5	18.6	6.5	50.0	R.finger_2	R.PIP	100	—	—
chest	chest_1	8.3	8.3	22.9	R.finger_1	R.DIP	100	—	—
	chest_2	8.3	8.3	8.6	R.thumb	R.thumb	100	100	—
	chest_3	8.3	8.3	8.6	R.thumb_2	R.thumb_MP	100	—	—
	chest_4	8.3	8.3	5.7	R.thumb_1	R.thumb_IP	100	—	—
	chest_5	8.3	8.3	5.7	L.scapla	L.scapla	100	100	100
	chest_6	8.3	8.3	5.7	L.shoulder (0~90°)	L.shoulder	100	—	100
	chest_7	8.3	8.3	5.7		L.shoulder (90° 150°)	L.shoulder	—	100
	chest_8	8.3	8.3	5.7	L.scapla	L.scapla	—	100	—
	chest_9	8.3	8.3	5.7	L.elbow	L.elbow	100	—	—
	chest_10	8.3	8.3	5.7		L.radio-ularnar	L.radio-ularnar	—	—
	chest_11	8.3	8.3	5.7	L.wrist	L.carpal	100	100	—
	chest_12	8.3	8.3	14.3	L.finger	L.MP	100	100	—
L.hip	L.hip	100	100	100	L.finger_2	L.PIP	100	—	—
L.knee	L.knee	100	—	100	L.finger_1	L.DIP	100	—	—
L.ankle	L.talocrural	66.7	66.7	—	L.thumb	L.thumb	100	100	—
	L.subtalar	33.3	33.3	—	L.thumb_2	L.thumb_MP	100	—	—
L.toes	L.f.MP	100	100	—	L.thumb_1	L.thumb_IP	100	—	—
R.hip	R.hip	100	100	100	neck	neck_1	14.2	14.2	14.2
R.knee	R.knee	100	—	100		neck_2	14.2	14.2	14.2
R.ankle	R.talocrural	66.7	66.7	—		neck_3	14.2	14.2	14.2
	R.subtalar	33.3	33.3	—		neck_4	14.2	14.2	14.2
R.toes	R.f.MP	100	100	—		neck_5	14.2	14.2	14.2
R.scapla	R.scapla	100	100	100		neck_6	14.2	14.2	14.2
	R.shoulder (0~90°)	100	—	100		neck_7	14.2	14.2	14.2
	R.shoulder (90° 150°)	—	100	—					
	R.scapla	—	100	—					

ROTはそれぞれ屈曲・伸展 (Flexion/Extension), 内転・外転 (Adduction/Abduction), 回旋 (Rotation) の略である。

2章にも示したとおり, PART, JOINT, およびその直積である CONNECTION で一般的な人体の構造を表し, これらと各個人を表す PERSON との直積をとることで各個人の人体の構造を表している. たとえば, 人体のある部位の骨であればおよそこんな形状をしている, という情報はポリゴンとして PART の成分主体 SHAPE に記述される. この形状を, INDIVIDUAL_PART の成分主体 LENGTH や DIAMETER などに示した各個人の形状情報に合わせて修正することにより, 各個人の骨形状をポリゴンとして得ることができる.

B.1 マッパーのアルゴリズム

SHM と BBHM の joint の対応, および文献 6) に示されている個々の関節角度への影響を基に導出した, 35 歳から 49 歳までの人の一般的な重み付けパラメータを, 表 2 に示す. なお, Flex は屈曲・伸展 (Flexion/Extension) の, Adc は内転・外転 (Adduction/Abduction) の, Rot は回旋 (Rotation) の略である.

(平成 12 年 6 月 20 日受付)

(平成 12 年 9 月 27 日採録)

(担当編集委員 有川 正俊)



佐藤 貴子 (学生会員)

現在横浜国立大学大学院工学研究科博士課程前期在学中. マルチメディアデータベースにおけるデータモデリング, および人体骨格モデリングに従事.



永野 志保

平成 12 年横浜国立大学大学院工学研究科博士課程前期修了. 現在 (株) ソニー木原研究所勤務. マルチメディアデータベースにおける人体骨格モデリングに興味を持つ.



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 8 年同大学大学院工学研究科博士課程前期修了. 平成 11 年同大学院工学研究科博士課程後期修了. 現在同大学工学部助手. 工学博士. 主としてマルチメディアデータベース, 質問処理等の研究に従事.



有澤 博 (正会員)

昭和 48 年東京大学理学部物理学科卒業. 富士通 (株) を経て, 昭和 50 年横浜国立大学工学部に奉職. 現在同学部電子情報工学科教授. 工学博士. 平成 3 年アメリカ・オレゴン州立大学計算機学科客員教授. データベース理論, マルチメディアデータベースシステムを研究テーマとしている.



酒井 直隆

昭和 54 年山形大医学部卒業. 医学博士. 米国メイヨー・クリニック研究員, 横浜市立大学整形外科講師を経て, 平成 12 年横浜市立市民病院診療部長. 横浜市立大学医学部・京都市立芸術大学・神奈川衛生短大・洗足学園大各非常勤講師. 臨床では手・膝の外科, スポーツ医学, リウマチが専門で, バイオメカニクスを研究テーマとしている.

時空間データベースコンセプトに基づく サッカー映像のモデリングと解析手法

富井 尚志, 坂木 和則, 四十物 裕司, 鬼束 郷, 有澤 博

横浜国立大学 工学部 電子情報工学科

〒240-8501 横浜市保土ヶ谷区常盤台 79-5

TEL: 045-339-4159

FAX: 045-338-1012

e-mail: {tommy, sakaki, anchan, go, arisawa}@arislabs.dnj.ynu.ac.jp

あらまし 本稿では、サッカーシーンをデータベース化するためのモデリング手法と、データベース設計、およびメディア解析データの蓄積手法について述べる。サッカーのような複雑な時空間に基づく情報をモデル化するためには、情報の規格化と個別情報の表現手法が重大な課題となる。この点について本研究では、規格と個別事例との直積要素をentityとみなし、Mediator概念によってモデル化を行った。また、Mediatorがメディアのどの位置にあるかという情報も直積によって表現できることを示した。さらに、このモデルに基づいてサッカーのシーンを時間非依存情報と時間依存情報とに分類し、それぞれ同様のスキーマによってデータベース化できることを示した。

キーワード 時空間データベース, マルチメディアデータベース, 時空間モデリング, メディア解析

Modeling and Analysis for Multiple Perspective Video of Soccer Game Scenes

Based on a Concept of Spatio-Temporal Databases

Takashi Tomii, Kazunori Sakaki, Yuji Aimono, Go Onitsuka, Hiroshi Arisawa

Yokohama National University

Faculty of Engineering, Department of Electrical and Computer Engineering

79-5 Tokiwadai, Hodogaya-ku, Yokohama 240-8501, JAPAN

TEL: +81-45-339-4159

FAX: +81-45-338-1012

e-mail: {tommy, sakaki, anchan, go, arisawa}@arislabs.dnj.ynu.ac.jp

Abstract This paper presents a methodology for handling soccer game scenes on a database. We offer a new modeling method, along with relevant database design and storage technique for the media analysis data. When modeling complex, spatio-temporal information as the one involved in soccer game scenes, information standardization vs. representation of individual information is a very important issue. As a solution, in this research the Cartesian aggregation of standard and individual element is regarded as an entity, and is modeled through the Mediator concept. Also, we show that such an approach makes possible the representation of information about Mediator's location in the media. Further, employing the model, soccer game scenes information is broken down into time-independent and time-dependent parts, and each is stored in the database with a relevant schema.

key words Spatio-Temporal Database, Multimedia Database, Spatial and Temporal Modeling, Media Analysis

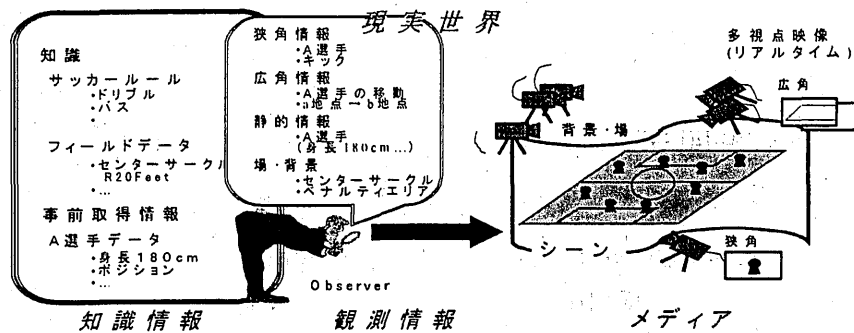


図1 サッカーシーンにおける情報とメディア

1. はじめに

空間的な広がりや時間的な移動を伴う事物・事象の統合的な管理・蓄積を目指したデータベースを時空間データベースと呼ぶ。我々の存在する現実世界は、空間的な要素と時間的な要素をあわせもった時空間であり、時空間データベース応用の中でも現実世界の時空間データベース化・モデル化は重要な課題である¹⁾。

近年の計算機技術の向上や、画像処理技術の発展により、多視点の映像や空間スキャンデータを用いて現実世界を計算機上に取り込む手段が実際に利用可能となってきた。映像やスキャンデータは、現実世界をデータ化するひとつの手段であり、これらのデータを基にして時空間データベースを構築できればその応用分野は計り知れない。中でもスポーツ分野では、放送のための映像データ²⁾だけではなく、実際に行われた試合事例の蓄積やそれらの統計処理と分析など、データベースを基盤とした情報処理技法と映像データを基盤とした解析技術の両者が求められ、さらにそれらを統合化して取り扱う必要が生じる。特にサッカーの試合を例に取れば、取得された映像データは様々な視点・視野角でフィールドを撮影した多視点の映像データとなるだろう。しかしそれだけではなく、そこに映りこんだ複数の選手のフィールド上での位置や姿勢、ボールの軌跡、それら全体があらわす戦術的意味など、様々なレベルの情報が集約されている。このため、時空間データベース化するための興味深い問題や工夫を要する点が数多く存在する。

サッカーの試合シーンのような複雑な現実世界をデータベース上に表現するためには、1) メディアが表す空間的・時間的情報が何かを分析する、2) 一般的・汎用的情報と、個別の情報とを区別し、規格に基づいて蓄積する、3) メディアデータを同一のデータベースに蓄積する、4) メディアデータ内における個別情報の位置を指し示してリンクする、5) メディアには見えない「意味情報」を指し示してリンクする、といった様々なモデル化が必要となる。

本研究では、上記1)、2)に関しては時空間データベースのコンセプトに基づいて汎用・規格化情報と個別のデータとの関連付けについてモデル化を行うことを試みた。この手法は、Mediator 概念⁴⁾と呼ばれ、具体的なデータを必要最小限で表現し、かつ規格化することができるので、集約演算や統計処理が可能となる。また、3)について統合的なスキーマ表現手法について述べる。さらに設計されたスキーマに基づいて4)で求められるデータの関連付けについて述べるが、この点に関しては特に映像メディアデータからの自動解析が有効であると考えられ、

- 映像メディアが映しこんだ場を復元する
- 空間オブジェクトの個別の形状・構造を復元する
- 空間オブジェクトの大局的な動きを復元する
- 空間オブジェクトの局所的な動きを復元する

などを考えることができる。これらの情報がすべて表現されることによって究極の目標でもある5)への実現性を示すことが可能となるだろう。

2. 時空間データベースを用いたサッカーシーンのデータベース化

2.1 サッカーシーンの含む情報とデータベース化

一般にデータベース化することとは、現実世界の情報を、概念的な代理物を用いてモデル化することであるといわれ、「情報」を表す具体的なデータを抽象的なデータ構造(スキーマ)に基づいて計算機に蓄積することである。スポーツの分野で一般的に取り扱われる情報は、選手の身体データなどの人間に関わる情報や、競技場のラインの位置などの空間的なジオメトリ情報、試合のシーンを撮影した映像、そこで行われている時間依存のプレイ(動作)とそのプレイを観察者が見て解釈した情報など極めて多岐に渡る。また、これらの情報は、映像や画像などのメディアデータの中に内容(コンテンツ)情報として現れる。メディアデータを内容に基づいて検索したり、後で再利用するためには、内容情報とメディアとの対応もデータベースに表現しなければならない。

ここで、例えばサッカーの試合を考えた場合、図1のように様々な情報を考えることができる。実際の試合のシーンは、複数の視点からの映像メディアデータとして取得できるだろう。しかし、このシーンは、観察者が背後に持つ「知識情報」にもとづいて、個々の事象・事例が観察者の解釈、すなわち「観測情報」としてとらえられる。そしてこれらすべての情報が現実世界の情報であると言い換えることができる。

これらの情報を対象として、計算機上で何らかの検索や統計などの処理を行いたいという要求は高い。

2.2 データベースを用いる意義

図1に示した情報に関して、画像処理による「被写体抽出」や「オブジェクトの姿勢抽出」を映像データに対して行うことは、応用事例ごとに実現が可能であると考えられる⁹⁾。例えば、競技場全体を視野内に捕らえる広角映像から選手の色領域を抽出できることと、現実空間と画素とのキャリブレーションが行われていれば、フィールド上での選手のグローバルな位置情報を得ることが可能である⁷⁾⁸⁾⁹⁾。一方で、ある個人の選手だけをズームして撮影した狭角・多方向からの映像から手足の位置・姿勢を解析できれば、その選手のローカルな動き情報を得られるかもしれない。このような、多視点映像からの人体の全身の動きをモデルベースで解析する例として文献¹⁰⁾¹¹⁾¹²⁾を挙げることができる。

このように、一人の選手の動きを映像から解析できるとしても、大局的な動きと局所的な動きとは、解析する対象となる映像が異なり、また、解析手法そのものも異なる場合が多い。この場合、解析された結果が同一の選手のものであることを示すためには、解析結果同士にリンクを張るなどの工夫をしなければならない。

また、一般に画像解析を行うためには、解析の事例に応じた処理時間が要求される。元の映像データは、カメラからリアルタイムで生成されるが、これに対して解析処理ごとに異なるタイミングで解析結果データが生成されていく。このようなデータを統合的に取り扱うために

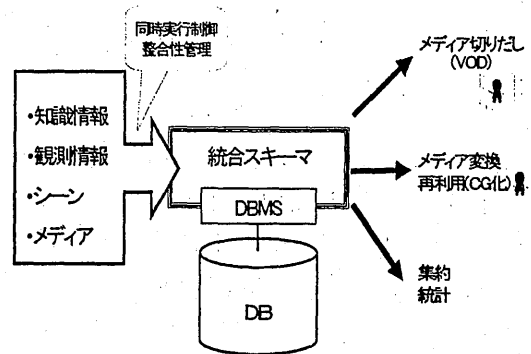


図2 データベース化の意義

は、複数のメディアの同期を取りながら蓄積¹³⁾したり、また、解析結果も元のメディアに対応付けて不整合が生じないように蓄積できなければならない。

上記のような問題点を総合すると、様々な詳細度の解析データとその対象(この場合は選手)の同一性の保証、生成データの時間同期の保証と解析処理プログラム間での同時実行制御などを、工夫されたデータ構造を用いて実現しなければならない。このためのひとつの解として、全てのデータに関する統合スキーマを作成して、データベース上でユニークに管理すれば、利用者の要求に応じた形式でのデータ検索¹⁴⁾やデータ再利用が可能となるだろう(図2)。

このように、意味情報、全てのメディアデータ、そして解析データを統一表現されたデータベースに蓄積することは、現実世界を対象としたシステムに対しては有用であると考えられる。

3. サッカーシーンのモデリングと統合スキーマ表現

3.1 直積 entity に基づくモデリング

前章で述べたように、データベース上で表現しなければならない重要な事項は、一般的な情報と個別の情報とを区別し、対応付けて表現することである。例えばサッカーの場合には、およそ全ての選手に共通の事項と、選手個人ごとや時点ごとに異なる事象とを対応付けてモデ

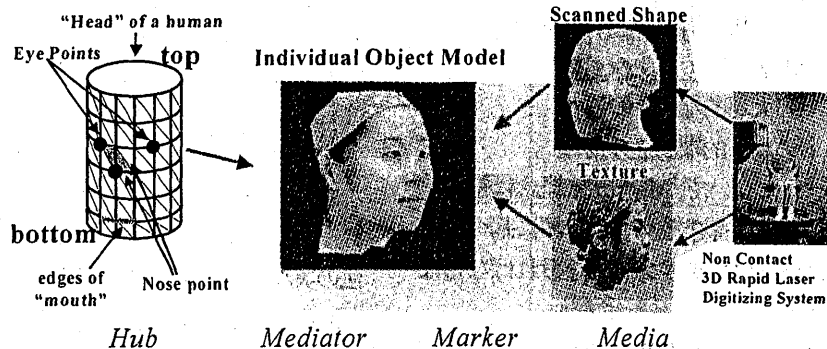


図3 メディエータ概念に基づく個別データの表現

ル化する必要がある。これに対し我々は Mediator 概念によって表現する手法を既に提案している⁴⁵⁾ (図 3)。まず、ある事象や事物の集合があるとき、これら共通の情報は entity である。例えば図 3 では、人間の頭部形状を表現しており、全ての人間に共通の情報として、①頭部の中心を通る軸を想定できる、②軸に対して、半径方向と円周方向からなる直交点によって表面形状を表現できる、③特定の点や高さに「目」や「鼻」などの意味情報を付加できる、といった特徴をもつ。

これに対し、頭部形状や目鼻の位置は個人ごとに異なる。したがって、個別の頭部形状データは、共通情報と個人情報との直積の要素(entity)であると考えられる。この要素を、共通の情報と個別の情報との Mediator (仲介者) であるとよぶ。また、共通の情報を、複数の Mediator が集約されたものという意味で Hub と呼ぶ。Mediator を作成するためには、各メディアデータ内でどの部位が Hub の意味情報であるかを対応付けなければよい。一般にメディアデータには複数の個別情報が存在し、また、ひとつの Mediator は複数のメディアデータから生成される可能性があるため、これもまた直積の entity である。この entity が示す情報は、メディア内における Mediator の部位や位置であるため、この entity を Marker と呼ぶ¹⁵⁾。このようにここで本質となるのは、直積の要素を entity としてとらえることにある (図 4)。Hub と個人・個体・個別事例との直積が Mediator であり、Mediator とメディアデータとの直積が Marker である。

3.2 サッカーシーンにおける情報の分類

前節で述べた Hub-Mediator-Marker-Media のレベルに加え、サッカーのように現実世界の事象そのものを考える際には、時間依存性や局所性を考慮しなければならない。一般に、シーンをモデル化するには次の要素について考慮する必要がある。

1. 時間に依存しない要素
 - i. 場の情報
 - ii. オブジェクトの形状と構造に関する情報
2. 時間に依存する要素
 - i. オブジェクトの大局的な動きの情報
 - ii. オブジェクトの局所的な動きの情報

これらの情報それぞれについて、Mediator 概念に基づくモデリングを考えることができる。この分類に基づいてサッカーのシーンの場合にモデリングしなければならない要素を表 1 に示す。

表 1 において例えば、場の情報のモデリング⁹⁾については、規格情報(Hub)として一般的なサッカーのフィー

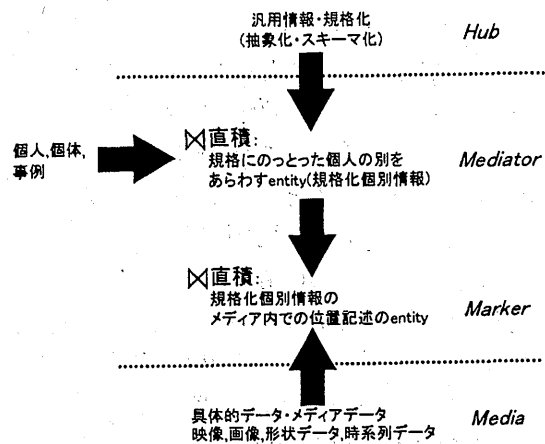


図 4 直積 entity に基づく表現

ルドの情報をモデル化することを考えることができる。これに対して、競技場ごとに異なる個別の情報 (Mediator) として、芝の状態 (テクスチャ) などを、Hub とリンクしておけば「平均的な芝の状態」などの集約が可能となる。また、場を複数台のカメラ画像で取得した際には、芝の状態がそれぞれの画像のどのピクセル群として表現されるかという要素 (Marker) を画像から抽出し、対応付けることが可能であろう。

サッカーのシーンの中でもっとも重要なことからは、個々の選手がいつでもどこにいて、そしてどのような姿勢 (運動) をしたかということである。この場合、オブジェクト (選手) は時間に依存する要素であり、空間的な情報に関して大きく分けて 2 種類の粒度を考えなくてはならない。ひとつはフィールド全体の中での大局的な動きで、もうひとつはオブジェクト内の局所的な動きである。前者はフィールド内の座標値の時系列データとして、後者はオブジェクトの構造に基づく関節時系列データとして表現されるであろう¹⁶⁾。この 2 種類の粒度で表現しておく、例えば CG や VRML などの仮想オブジェクトモデルにマップして、仮想空間データに変換できるため、データベースのデータ再利用という観点では有効である。

表 1 の分類で注目すべき点は、Marker にあたる情報はメディア解析に基づいて自動的に抽出できる可能性があることである。例えば、選手のフィールド内での座標値データは、フィールド全体を広角に撮影した映像から復元可能であるし、選手の姿勢時系列データは、その選手をズームアップで撮影した狭角の映像から解析できるかもしれない。後述する 4 章では、これらの情報を元にデータベースを構築する際の、Marker データの取得方法の実現性について述べる。

3.3 スキーマ表現

3.1 節で述べたように、サッカーシーンの空間的な事象

をデータベーススキーマとして表現する際に本質的になるのは、直積の entity を表すことにある。本節では、表 1 で求めた各要素を直積モデリングに基づくデータベーススキーマ上で表現することを試みる。

まず基本的な概念スキーマを図 5 に示す。図 5 において、矩形は entity の型(type)を表す。またひし形は、線で結ばれる型の具体値(instance)となる entity 同士の直積を表す。したがって、ひし形に隣接して型がある場合には直積の要素がまた entity であることを表す¹⁷⁾。図 5 は図 4 の要素をそのままスキーマ化したものであり、Mediator 概念が直積要素の entity であることを示している。例として場の情報のスキーマを図 6 に示す。

図 5 において、まずデータベース設計者が現実の世界を観察して慎重に分析しなければならない事項は Hub の設計である。特に、直交する属性の集約値 (すなわち特徴量ベクトル) によって汎用情報を記述できれば、それはデータを規格化したことを意味する。個々の Mediator のデータが、この規格にのっとって作成されれば、Mediator データの統一表現が可能となる。

次に Mediator の情報は、Hub と個人・固有の事象との直積で表現される。実際には Mediator の属性情報として Mediator を構成する情報が付加される。Mediator 構成情報は、先に述べた規格化されたデータであると考えられるので、一般には特徴量ベクトルとの直積となる。この点が図 5 では図 4 よりも詳細に記述されている。例えば図 6 の場を記述した例では、個々のサッカー場ごと

にラインの様子や芝の様子が異なる。これらは、場を記述した Mediator の属性である。

一般にそれぞれの個別の事物・事象は、複数のメディアデータ内に現われ、また、ひとつのメディアデータには複数の事物や事象が映りこんでいる。この事実をメディア毎・事物ごとに対応付けて表す要素が Marker であり、Mediator (またはその構成情報) とメディアとの直積要素である。図 6 においては、Mediator の属性ごとに画像データとの対応がとられる。

図 6 の例と同様、オブジェクトの静的情報や時間非依存の要素についても、表 1 の分類に基づいて Hub-Mediator-Marker-Media の対応を直積要素によって表現したスキーマ記述を考えることができる。以上のようなスキーマ表現によって、サッカーのシーンを統合的なデータベースとして構築することが可能となる。

4. メディア解析に基づくデータ取得

前章までに述べた、データベースのモデル化手法、データベーススキーマに基づいて、実際のサッカーシーンからデータベースのインスタンスを生成しなければならない。ここで、特に Marker 情報は、メディアを Mediator データを用いて解析し、対応する位置を画像から抽出するなどして得られるデータである。したがって、このインスタンスはメディア解析の結果をデータベースに蓄積したものであると言い換えることができ、また逆に、解析結果は例えば図 6 のスキーマの Marker の属性値としてデータベースに蓄積することが可能である。

表 1 サッカーシーンのデータベース化のための情報分類

	時間非依存		時間依存	
	場	オブジェクトの静的情報	オブジェクトの大局的動き情報	オブジェクトの局所的動き情報
汎用情報規格化 (Hub)	<ul style="list-style-type: none"> 場の直行成分 タッチライン方向軸 ゴールライン方向軸 部位情報 センターサークル ペナルティエリア 	<ul style="list-style-type: none"> 部品の直交成分 半径方向 円周方向 高さ方向 部位情報 目、鼻 etc 部品間の接続 	<ul style="list-style-type: none"> Low Level な位置情報 (x, y, z, t) 時点 イベントの空間情報 イベント時区間の情報 イベントの意味の直交成分 	<ul style="list-style-type: none"> オブジェクトの構造に基づく情報 (姿勢の時系列表現) 基本動作記述 イベントの意味の直交成分
規格化個別情報 (Mediator)	スタジアムごとの固有値	<ul style="list-style-type: none"> 個人ごとの形状や目 個人ごとの表面形状 個人ごとの関節稼働範囲 	<ul style="list-style-type: none"> 上記の各項目と各オブジェクトとの直積 イベントとの対応 	<ul style="list-style-type: none"> 上記の各項目と各オブジェクトとの直積 イベントとの対応
メディアごとの規格個別情報 (Marker)	GPS 座標値 芝のテクスチャデータの対応など	個人ごとの目鼻の位置と形状データ・テクスチャデータとの対応	個人ごとの空間的位置の時系列情報及びイベントと映像データとの対応	個人ごとの姿勢と映像データ、モーションキャプチャデータとの対応
取得データ (Media)	詳細静止画 動画の各フレーム	立体形状データ 写真データ	広視野角の映像 (一般に多視点映像)	広視野角映像+ズーム 映像 (多視点映像)

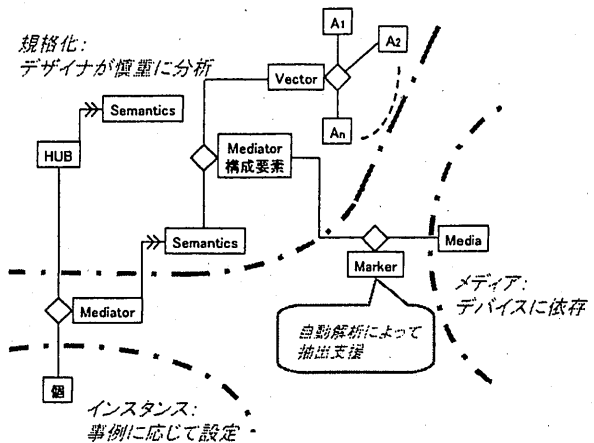


図 5 直積に基づく Mediator の基本概念スキーマ

特に時間依存の情報としては、フィールドを広範囲にわたって撮影した映像から、色領域に基づいて選手の位置を抽出し、その結果を同一のデータベース上に蓄積することが可能である⁹⁾。また、この解析結果や形状の Mediator を足がかり (モデル) として、ズームアップの多視点映像とのマッチングを取ることで、サッカー選手の姿勢情報を得る実験も行っている¹⁰⁾が、この結果も同様のデータベースに蓄積することが可能である。

5. まとめ

本稿では、サッカーシーンをデータベース化するためのモデリング手法と、データベース設計、およびメディア解析データの蓄積手法について述べた。サッカーのような複雑な時空間に基づく情報をモデル化するためには、情報の規格化と個別情報の表現手法が重大な課題となる。

この点について本研究では、規格と個別事例との直積要素を entity とみなし、Mediator 概念によってモデル化を行った。また、Mediator がメディアのどの位置にあるかという情報も直積によって表現できることを示した。さらに、このモデルに基づいてサッカーのシーンを時間非依存情報と時間依存情報とに分類し、それぞれ同様のスキーマによってデータベース化できることを示した。

本稿で示したデータベースは、純粋にメディアデータとオブジェクトとの対応を取っただけのものであり、はじめにの5) で述べた究極の目標である「意味」に基づくデータベース集約について、今後の研究を急いでいる。謝辞 本研究を行うにあたり、サッカーシーンの撮影に際し日本テレビフットボールクラブ (ヴェルディ川崎) の協力を得た。ここに記して謝意をあらわす。また、本研究の一部は、文部科学省科学研究費補助金奨励研究(A) (課題番号 12780202) による。

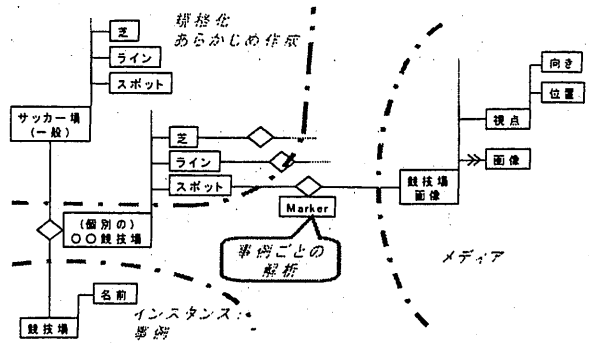


図 6 場の情報のスキーマの例

参考文献

- 1) M.F. Worboys: A Unified Model for Spatial and Temporal Information, *Computer Journal*, 37, 1, pp.26-34, 1994
- 2) A. Del Bimbo, E. Vicario, D. Zingoni: Symbolic Description and Visual Querying of Image Sequences Using Spatio-Temporal Logic, *IEEE Trans. On Knowledge and Data Engineering*, 7, 4, pp.609-622, 1995
- 3) 橋本, 白田, 真野, 飯沢: TV型受信端末におけるダイジェスト視聴システム, 情処学論データベース, 41, SIG3(TOD6), pp.71-84, 2000
- 4) 富井, 小林, 有澤: 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, 信学論, J82-D-1, 1, pp.211-222, 1999
- 5) 今井, 富井, 有澤: Mediator 概念に基づく身体形状・動作のモデル化とデータベース設計, 情処学論データベース, 41, SIG1(TOD5), pp.100-108, 2000
- 6) 安居院, 長尾: 画像の処理と認識, 昭晃堂, 1992
- 7) 瀧, 松本, 長谷川, 福村: サッカー映像からのチームワーク評価方法の検討, 信学技法, PRMU96-10, pp.67-74, 1996
- 8) 林, 富井, 有澤: 映像データベースのための被写体形状の取得, 情処研報, DBS-111-13, pp.95-102, 1997
- 9) A. Zisserman, A. Fitzgibbon, G. Cross: VHS to VRML: 3D Graphical Models from Video Sequences, *IEEE, ICMCS99, Vol.1*, pp.51-57, 1999
- 10) 大垣, 岩井, 谷内田: 動きと形状モデルによる人物の姿勢推定, 信学論, J82-D-II, 10, pp.1739-1749, 1999
- 11) 米元, 有田, 谷口: 多視点動画画像処理による実時間全身モーションキャプチャシステム—視覚に基づく仮想世界とのインタラクション, 映情学誌, 54, 3, pp.409-416, 2000
- 12) 佐藤, 川田, 大崎, 山本: 多視点動画画像から人間動作の追跡と再構成, 信学論, J80-D-II, 6, pp.1581-1589, 1997
- 13) 中家, 徐, 富井, 有澤: 次世代放送向け多チャンネル映像のデータベース化, 情処研報, DBS-122-27, pp.207-214, 2000
- 14) C. X. Chen, C. Zaniolo: SQLST: A Spatio-Temporal Data Model and Query Language, A.H.F. Laender, S.W. Liddle, V.C. Storey (Eds.) *ER2000 Conf, LNCS 1920*, pp.96-111, 2000
- 15) 富井, 有澤: マルチメディアデータベースにおける映像モデリングと操作言語, 信学論, J79-D-II, 4, pp.520-530, 1996
- 16) 佐藤, 永野, 富井, 有澤, 酒井: 骨格モデルを用いた人体動作データベースの構築, 情処学論データベース, 42, SIG 1(TOD 8), pp.92-102, 2000
- 17) H. Arisawa, T. Tomii, H. Yui, H. Ishikawa: Data Model and Architecture of Multimedia Database for Engineering Applications, *IEICE Trans. Inf. & Syst*, E78-D, 11, 1362-1368, 1995
- 18) 坂木, 今井, 富井, 有澤: 多視点映像を用いた時空間オブジェクトの解析とモデリング, 信学会, データ工学ワークショップ論文集, DEWS2000, 7B-4, 2000

映像情報メディア学会
画像情報システム研究会
招待講演

2001. 3. 2

映像からの3次元解析と 時空間モデリング

有澤 博

横浜国立大学 工学部 電子情報工学科
(平成13年4月より横浜国立大学大学院環境情報
研究院「社会環境と情報」研究部門)

背景

✧究極のマルチメディアデータベースとして、
実世界をそのままDB化したい。

→リアルワールドデータベース(RWDB)を提案
(1996~)

✧RWDBを作るには、次の3つが必要。

- ・時空間解析と動作モデリングの研究
- ・時空間データベースエンジン
データモデルの開発、研究
- ・概念、意味などのモデル化と検索(“真の”人工知能)

} 有澤研の
テーマ

✧RWDBコンセプトの応用

学術的、産業的、文化的な役割をもつキラーアプリケーション

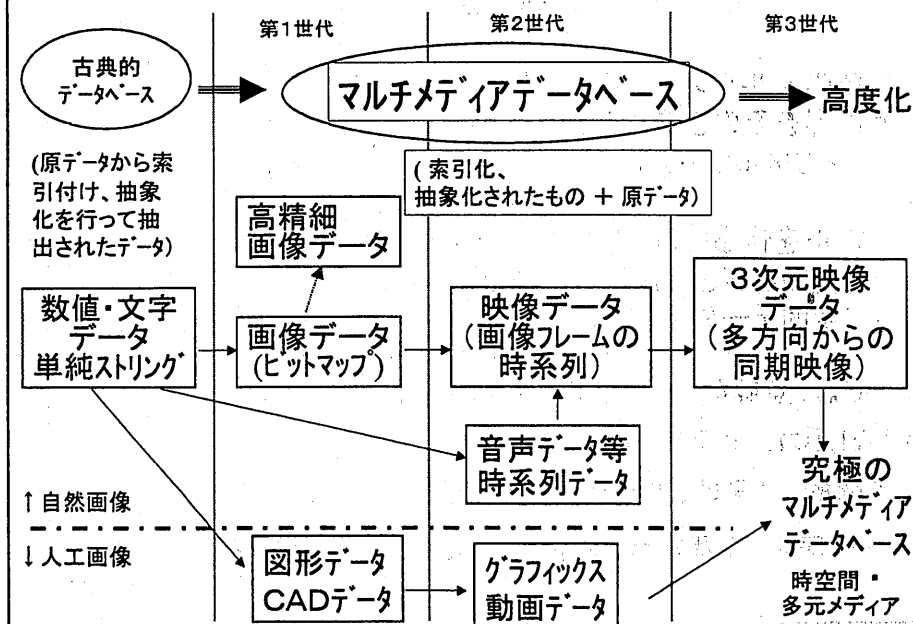
→ 医学、工程設計、スポーツ、芸術、自然現象、放送文化、...

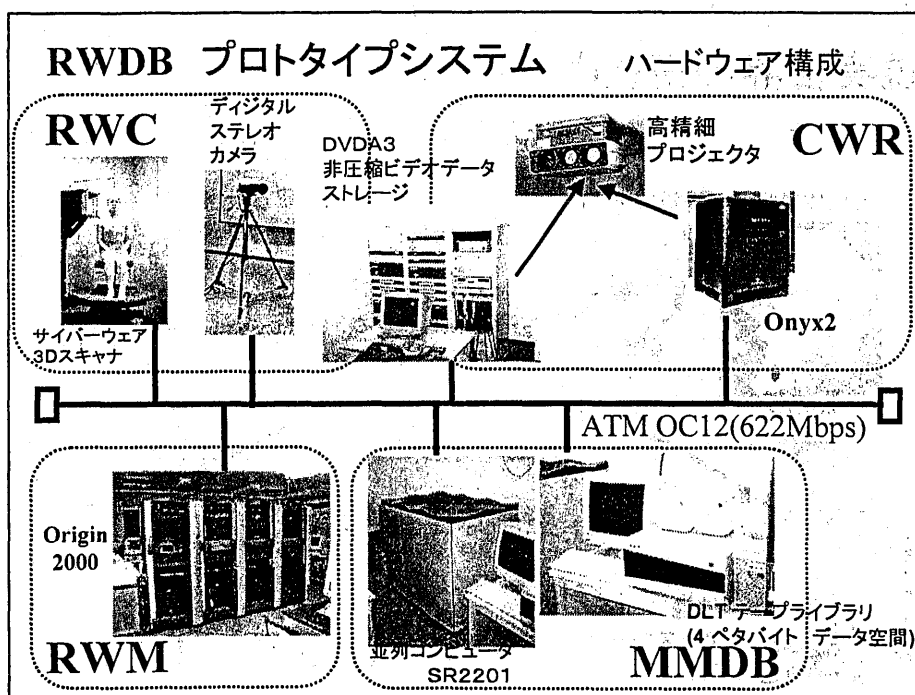
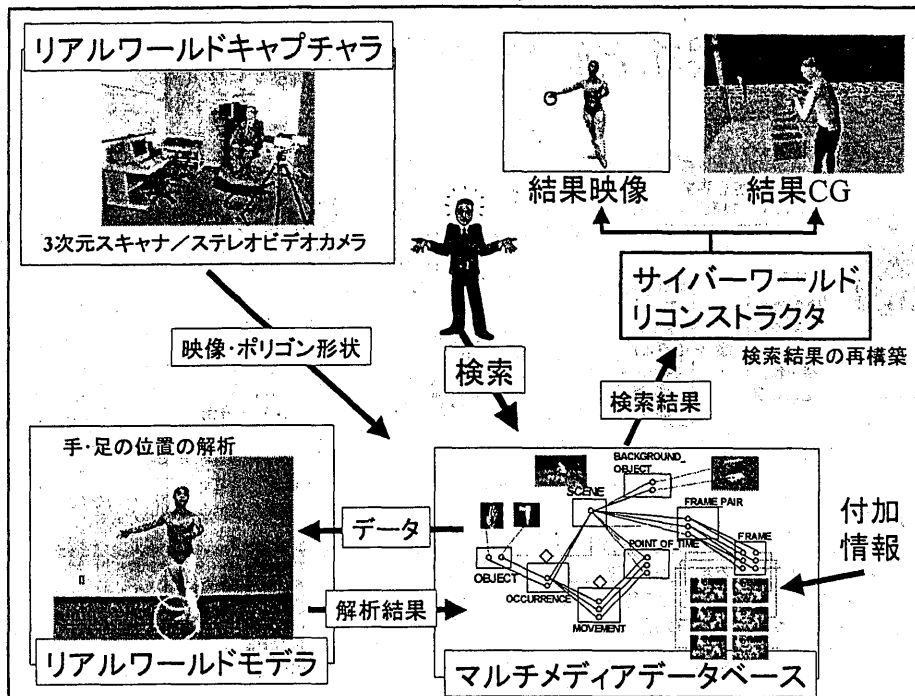
リアルワールド・データベースの目的

- 対象世界の中にある、さまざまな「もの」の3次元形状、時間的変化を「見たまま」「あるがまま」記録する。
- 3次元オブジェクトの形状・動作について、解釈・解析を行って、その結果も原データとの対応を残して、同じデータベースに蓄積する。
- 利用者の見方から自由な検索を行って情報の抽出を行う。
- 検索の結果から再構成されたサイバーワールドを合成された立体映像や精密なVR等を用いて、現実感豊かに再生する。

応用分野...人間の作業空間の最適設計、スポーツ医学、
人間工学シミュレーション、
「名人芸」などの3次元映像DB など

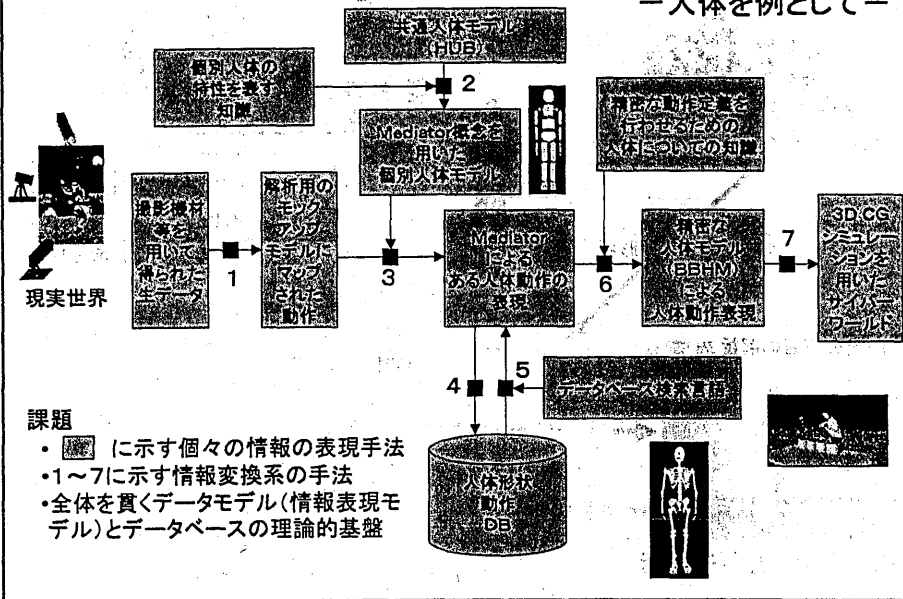
データベースの取り扱うデータの発展





時空間解析と動作モデリング

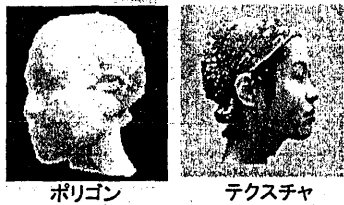
—人体を例として—



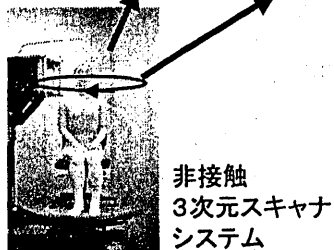
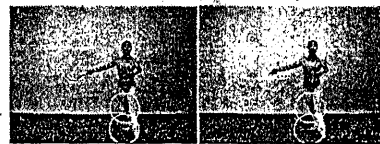
現実世界の情報の取り込み

現実世界の情報をできるだけ詳細に取り込む

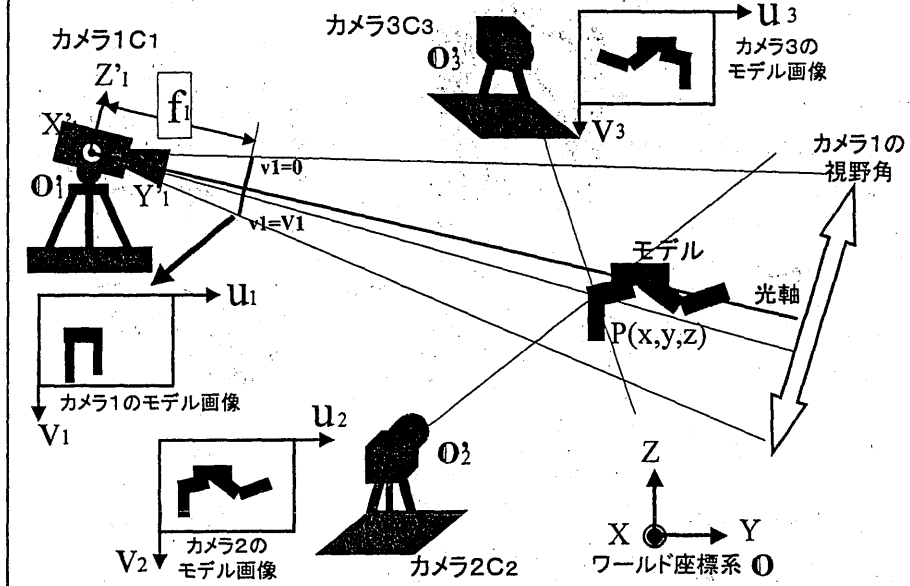
人体形状データ (静的情報)



動作情報 (動的情報)



複数のカメラで撮影した際の空間モデル



メディエータの導入(目的と役割)

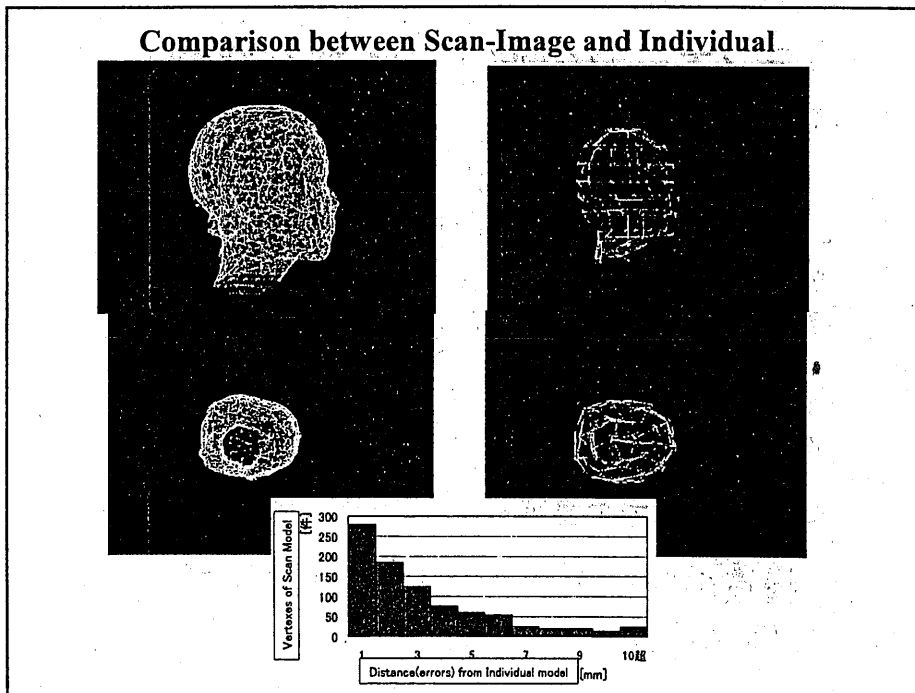
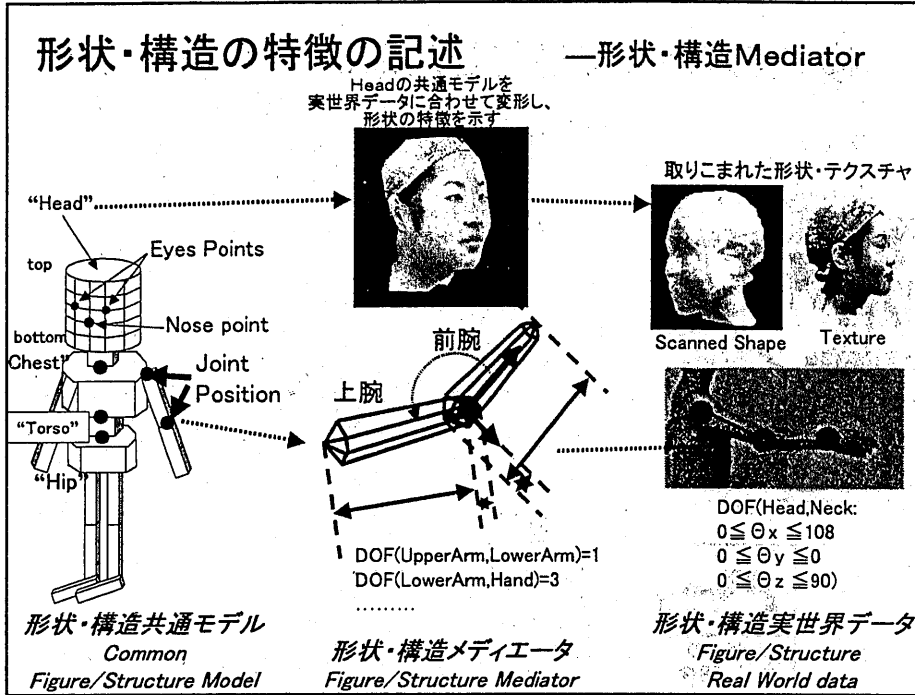
メディエータ:

実世界から得られた3次元オブジェクトの形状、姿勢変化の情報を「本質的にその特徴を記述できる」必要最小限の要素データ群として表現したもの。

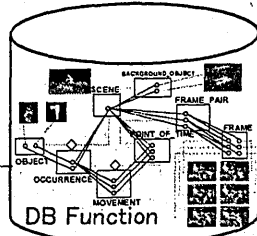
- 直交的な成分に分解して記述
- 「スキーマに基づく検索」を可能にするようなDB表現
- インスタンスごとの違いを表現可能
- オブジェクトが持つ共通の意味や構造を保持

人体データベースでの重要な関連

形状・構造の特徴の記述 → 形状・構造メディエータ
 運動の特徴の記述 → 運動メディエータ



検索とは



関数型データベース

複雑・長大
大容量

→ simple primitives
多値順序関数

AIS model

データベースを参照・利用して
データを検索・再構成し、
新たな形式のデータを生成

関数型操作系

並列化・最適化

宣言的操作

構造化オペレータ

メディア依存処理関数

FP FQL

要求に従って
新たに作成されたデータ

様々な「知識」を
色々な面から検索
(基本要素と「つながり」に分解)



ビデオ映像



CGアニメーション

AISデータベース

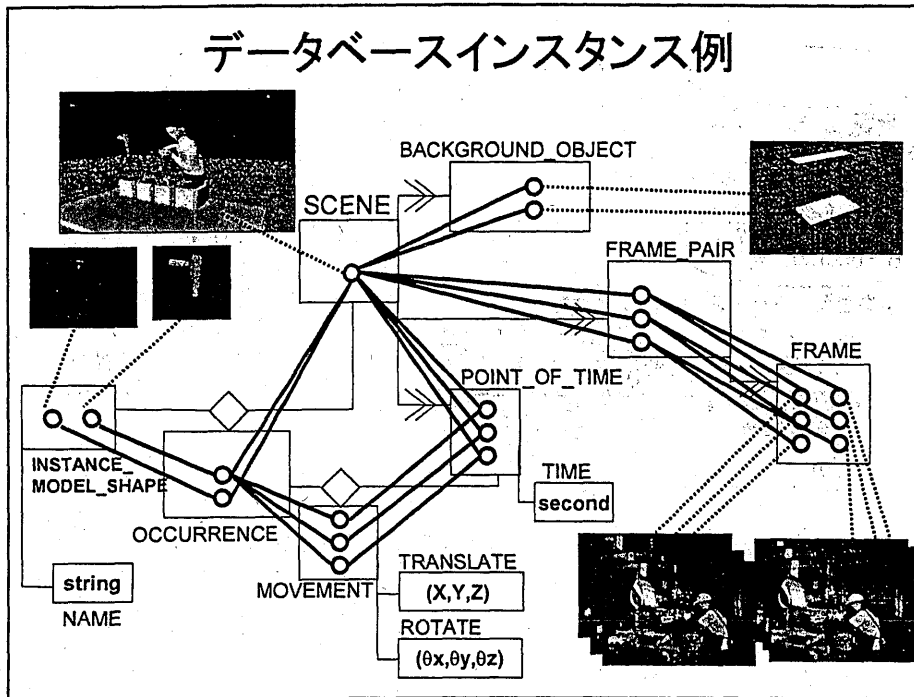
AISモデル

現実世界を主体(entity)と、
それらの間のつながり(association)で表し、
モデル化するデータモデル

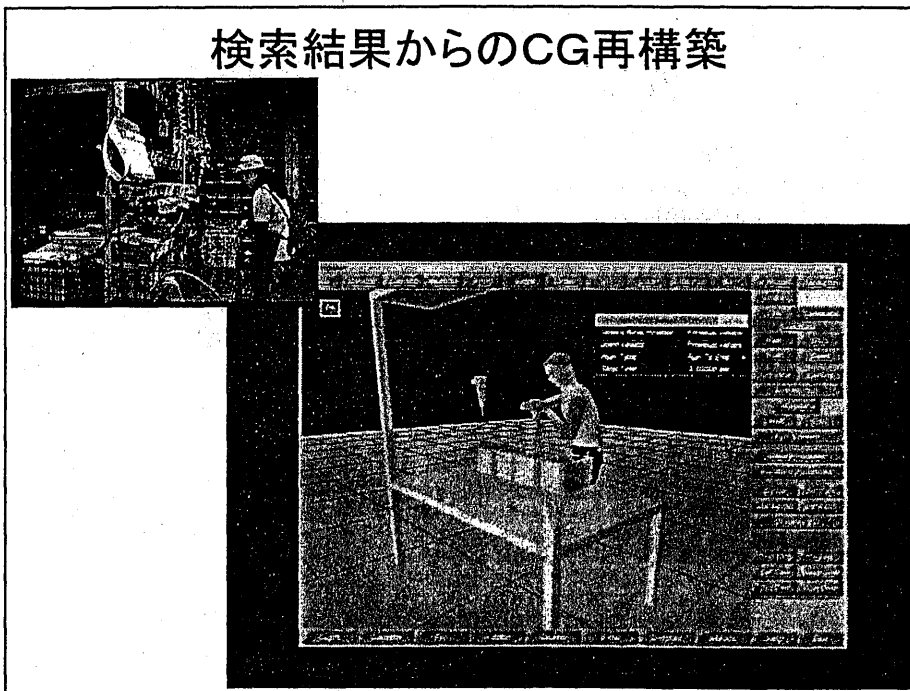
2項関係を用いたフラットな構造
→関数によるデータ表現

P.P.Chenの実体関連モデル(Entity-Relation Model)を拡張

データベースインスタンス例



検索結果からのCG再構築



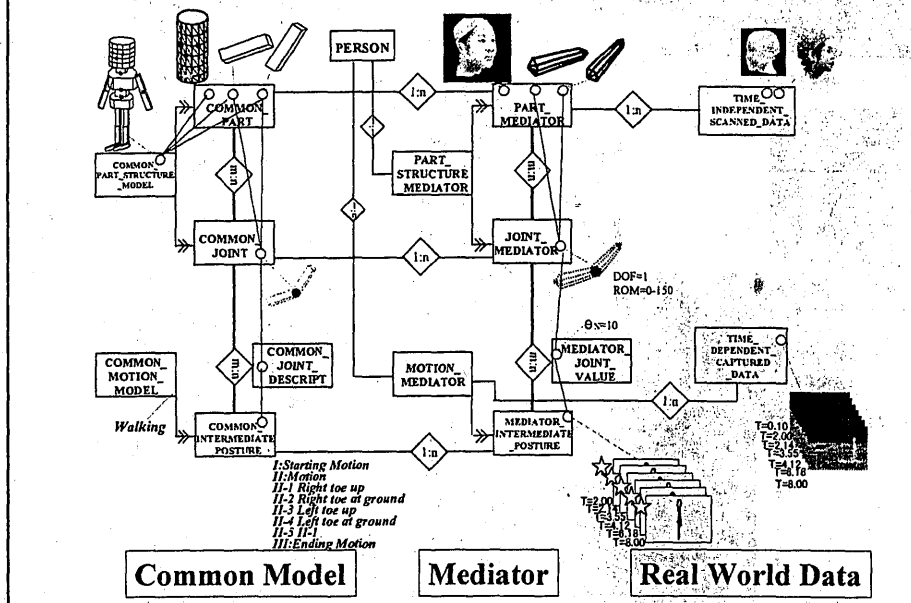
情報人間工学 Info-Ergonomics

—メディエータに基づいた人体と動作のモデル化

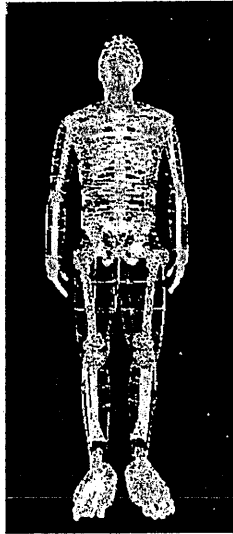
Information Ergonomics:

- 人体の部品形状, 関節構造, 動作のモデル化
- 実世界からの人間の動作のシミュレーションと評価
- リアルワールドデータベースに基づく人間モデルと動作の検索

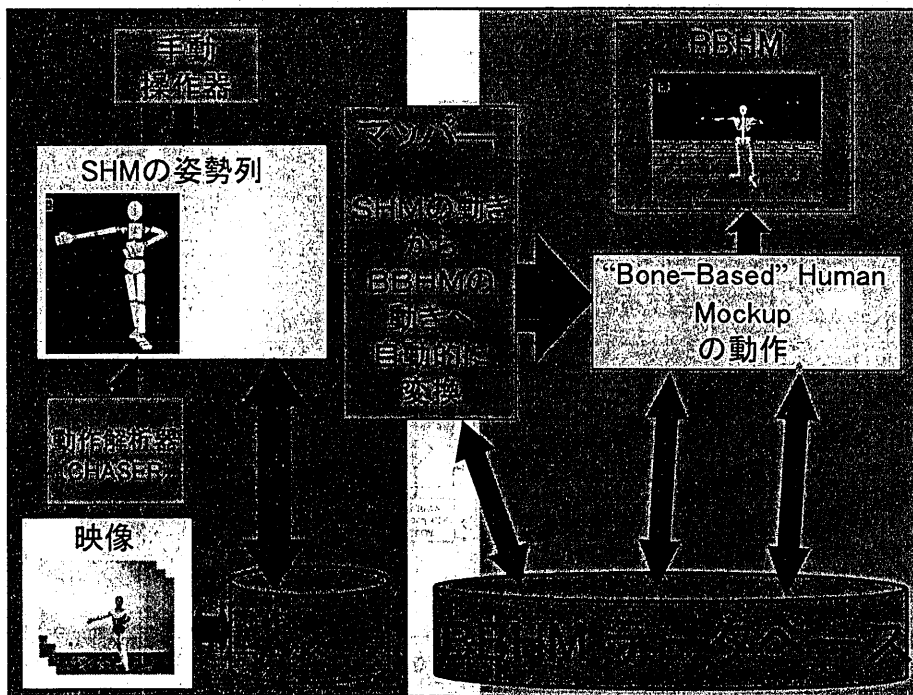
人体形状・構造・動作データベーススキーマ



“Bone-Based” Human Mockup(BBHM)



- 骨と皮により構成
- 骨の長さ、関節可動域を個人に合わせて変えることができる
- 155部品
- 126自由度



今後の課題

- Info-Ergonomics and Human model -
 - 詳細な人体モデルの作成
 - 様々なアプリケーションにおける実現可能性
 - メディエータに基づいた動作解析
 - 力学的解析とモデリング

参考文献

- 今井, 富井, 有澤: "Mediator概念に基づく身体形状・動作のモデル化とデータベース設計", 情報処理学会論文誌, Vol.41, No.SIG1(TOD5), pp.100-108, Feb. 2000
- 佐藤, 永野, 富井, 有澤, 酒井: "骨格モデルを用いた人体動作データベースの構築", 情報処理学会論文誌, Vol.42, No.SIG1(TOD8), pp.92-102, Jan. 2001
- H. Arisawa "Mediator-based modeling of factory workers and their motions in the framework of Info-Ergonomics", Proceedings of the International Conference on Machine Automation ICMA2000, pp.363-370, 2000.9
- 家富, 森崎, 富井, 有澤: "マルチメディアデータブラウジングのための視覚的データベースインターフェースAVISの実装", 電子情報通信学会論文誌D-I, 6月号(システム開発論文特集), 2001.(掲載予定)

空間マウスを用いた人体骨格モデルの動作定義支援システムの構築

佐藤 貴子 上島 崇宏 富井 尚志 有澤 博

横浜国立大学 工学部 電子情報工学科

〒 240-8501 横浜市保土ヶ谷区常盤台 79-5

E-mail : {sugar, mal, tommy, arisawa}@arislabs.dnj.ynu.ac.jp

あらまし

本研究では、人体骨格モデルの動作定義支援システムの設計と実装を行った。人体の構造は複雑な構造を成すために、骨格に基づいたモデル (Bone-Based Human Model : BBHM) での動き記述が最も本質的であると考えた。しかし、BBHM は部品点数が多いため、動きを与える事が容易ではなかった。そこで構造を単純化したモデル (Simplified Human Model : SHM) を使い、グラフィックシステム上で SHM を動かすことで、BBHM に動きをマップするシステムを設計・実装した。特に直観的な操作の為、本研究では 3 次元測定デバイスである空間マウスを用いて SHM を動かすシステムを実装・評価した。

1 はじめに

近年、計算機技術が発達するに従い、コンピュータグラフィックス (CG) を用いたデータの視覚化が盛んに行われるようになってきている。特に、人体のような複雑な構造を持つ 3 次元物体の CG による視覚化技術は発展している。

一方、コンピュータ上に現実世界の「もの」の性質を忠実にモデル化することにより、精密なシミュレーションや評価を行う技術が天文学、核融合、気候学など各方面で研究が進められている。

対して、我々はこれまで、現実世界の事物・事象をそっくりそのままとりこみ、蓄積する、リアルワールドデータベース (Real World Database : RWDB) を提案し [1]、研究・開発を行っている。

RWDB では、現実世界の「もの」に関する抽象化された情報が蓄積されている。RWDB と、シミュレーション技術とを組み合わせることで、架空の世界を現実世界にあたかも存在したかのように構築することができる。

ここで重要なことは、現実世界の「もの」をモデル化して得られたオブジェクトの詳細化・精密化である。例えば人体の場合、抽象化のモデルとして形

状や構造の非常に単純化されたものを用いれば、物体との衝突検出などの限られたシミュレーションまでは可能であるが、ある姿勢をとった時の人体各部にかかる負荷を求めたい、といった要求には耐えられないことは容易に想像できるであろう。そこで、特に人体についての詳細なシミュレーションに耐えるモデルとして、これまで、骨格に基づいたモデル (Bone-Based Human Model : BBHM) を提案し [2]、実際にコンピュータ上に作成して人体の姿勢の再現性などの評価を行ってきた。

ここで、シミュレーションを行うためにはモデルに対して姿勢・動作を定義しなければならないが、この時、以下のような問題が生じる。

1. 人体の運動機能に反しない姿勢の定義
人体の単純な動作でも、複数の関節が連動して動くので、姿勢定義ではこの点を考慮して行わなければならない。
2. 3次元空間中での直観的操作
従来のグラフィックスシステムに見られるような、マウスにより 3 方向の移動・姿勢操作に分けて移動や回転を指定する方法はある程度の技術を要する。従って仮想空間内で直観

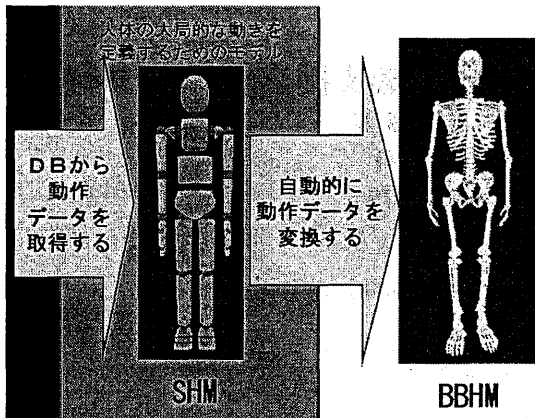


図 1: BBHM への動作定義の流れ

的な操作環境を提供することが有効な手段である。

1. の問題に対しては、人体の姿勢に関する制約は、医学的に人体の関節が互いに連動して動く際の動き方が知られているということを利用して、BBHM に対する姿勢定義を、構造のより単純化されたモデル (Simplified Human Model: SHM) を用いて行い、SHM の姿勢を自動的に BBHM の姿勢へ変換する、という手法をとり [2]、この変換を行うものをマッパーと呼ぶ。これを図 1 に示す。

また、2. の問題に対しては、3 次元的な位置・回転の合計 6 自由度を検出できるセンサデバイスを利用した、空間マウスによる姿勢・動作の教示を実現する。

このようなシステムは、工場における作業工程の最適設計や、スポーツ医学でのリハビリテーション計画などにおいて非常に有効であると考えられる。

本稿では、2. の問題に対して直観的な動作指定の方法として、空間マウスを用いて、仮想空間内に配置された簡易化人間モデル (SHM) に動作を指示し、その動作をリアルタイムに BBHM にマッピングするシステム的设计および構築を行う。

2 姿勢の再構築

我々が提案している人体動作データベースの中に

- 現実世界に存在する、人間も含めたオブジェ

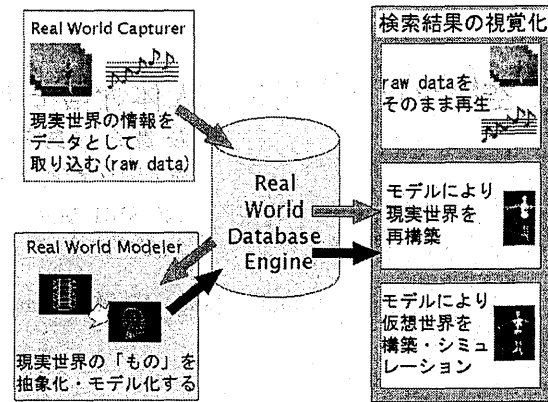


図 2: データベースからの検索結果の表示

クトの形状や動作に関する情報を抽象化して得られたモデル

- 抽象化の際に現実世界の情報を取り込むために用いられた映像や音声、計測機械の数値などの抽象化されていないデータ (raw data)

の両者が蓄積されている。これに対してデータベース利用者がとりだす情報・データは以下のようなレベルがあるといえる。

1. データベース利用者が検索によって要求した raw data そのもの (例えば物体をカメラで撮影した映像、物体をスキャナで取り込みポリゴンで表したデータなど) を単に再生する
2. 現実世界のあるシーンをグラフィックモデルを用いて CG で再現し、様々な角度から見る。
3. 現実世界の「もの」をグラフィックスで再現し、それらを組み合わせて新たな仮想世界を構築してシミュレーションを行う。

この概念図を図 2 に示す。

人体動作のシミュレータとしては、衝突検出やいくつかの人体負荷評価を行うことのできる米 delmia 社の ENVISION/Ergo [3]、様々な人体負荷評価を行うことのできる米 EAI 社の Transom Jack [4]、医療向け数値解析が行える米 Musculographics 社の SIMM [5] などが挙げられる。しかし、これらのシミュレータでは人体を単純にモデル化しているために限られたシミュレーションや評価しか行えない、

あるいは逆に非常に詳細に人体をモデル化しているものの医療分野に特化しており、他分野で利用するには向いていないといった理由から、統合的で精密な人体動作シミュレータであるとはいえない。この問題を解決するためには、

- 複雑で多くの自由度を持つ関節の動きを直観的で精密に指示できること
- 互いに束縛し合いながら動く複数の部品の局所的なシミュレーションができること

が必要となる。

これをふまえ、我々は、SHM とマッパーを介した姿勢変換、および空間マウスを用いた直観的な動作教示手法により動作定義支援システムを考えた。

3 動作定義支援システムのための要素技術

人体の姿勢・動作のシミュレーションを行うためには、まず、人体を表したモデルに対して、シミュレーションを行いたい姿勢・動作を与えなければならない。しかし、骨格に着目すると、複数の関節が連動して動く部位が多く、実際の人体の運動機能を反映したシミュレーションを行うためには、この連動を考慮せずに人体骨格モデル (Bone-Based Human Model: BBHM) の各関節の動きを独立して定義すべきではない。

また、人体の姿勢・動作は3次元空間内で行われるものであり、直観的に操作が行える必要がある。

そこで、人体の運動機能に反しない姿勢定義を行うため、BBHM と比べ非常に構造が簡略化されたモデル (Simplified Human Model: SHM) を使って姿勢を定義し、SHM の姿勢を実際の人体の運動機能に基づいて変換することで BBHM の姿勢を定義するという手法を提案している。また、3次元空間内での直観的な姿勢教示のためのデバイスとして、空間マウスを用いることを考えた。

3.1 人体の運動機能に従った姿勢定義

BBHM が人体の非常に詳細な姿勢を表現可能なモデルであるのに対し、SHM は人間がとらえる人体の姿勢であり、人体の「概念的な姿勢を定義する

モデル」にあたる。例えば、人間の認識では「肘から先を 90° ひねる」という動作であっても、実際の人体においては肘から手首の間に尺骨と橈骨という2本の骨があり、尺骨のまわりを橈骨が楕円錐を描くようにして動くことにより、肘から先が回転する。このように、実際の人体の動きと人間が認識する人体の動きには相違があるため、この間のギャップを埋める必要がある。このギャップを埋める役割を果たすのがマッパーである。

人体の大局的な姿勢に対して骨格レベルでの個々の関節の曲がる割合が決まっている [6] ことを利用すれば、SHM で定義した大局的な姿勢を基に、BBHM の個々の関節の角度を算出することが可能である。

3.1.1 SHM

SHM は人体を大きく 28 個の部位に分け、それらを関節で結び、全体で 59 自由度を持たせたものである。これは、BBHM が骨形状をした部品 155 個、120 自由度であることと比較して、非常に構造が簡略化されていると言える。しかし、人体の大局的・概念的な姿勢を表現するには十分な構造を持っている。

さらに SHM は、BBHM に姿勢を教示する際、ユーザが直接操作する、いわば「インターフェース」のような役割を持つため、単に姿勢を表現するだけでなく、ユーザに人体の構造等を分かりやすく示すことができる必要がある。そこで、各関節を「関節を表現するためのパーツ」(以降 関節パーツと呼ぶ) で表した。また、関節毎に、決まっている自由度を表現するため、自由度毎に関節パーツの形状を変化させた。関節パーツの自由度毎の形状を図 3 に示す。また、パーツに対しては前半面と後半面を塗り分け、中心線を入れることで、現在パーツがどの向きに向いているかをユーザが把握しやすいようにした。この例を図 4 に示す。

3.1.2 マッパー

マッパーは、SHM で定義された大局的・概念的な姿勢・動きから、BBHM の非常に詳細な姿勢・動きへと変換するものである。この際、関節数の少ないモデルから多いモデルへ姿勢・動作データを変換することを考える場合、当然ながら変換前の動作データが実際の人体の動きを忠実に表現できていた

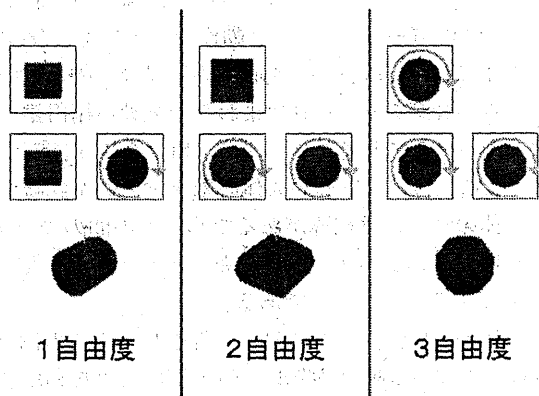


図 3: SHM の関節パーツの形状

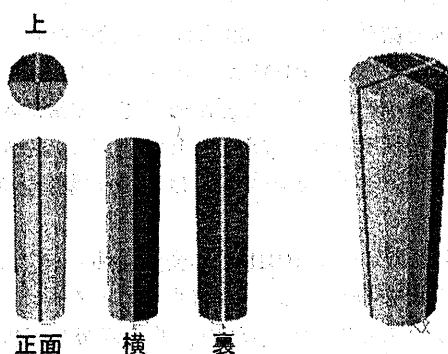


図 4: SHM のパーツの塗り分け

としても、変換した後の動作データが実際の動作と異なってしまうのではないかと懸念される。そのため、人体の運動機能、例えば連動して動く部位の連動の様子などをマッパーでどのようにして忠実に表現するかが問題となる。

ここで、連動して動く部位については、医学分野において、人体の大局的な動きに対して各関節がどのように動くかの割合が研究されている [6]。この割合から、関節ごとに、大局的な動きのうち何%を担うかを示す重み付けパラメータを算出できる。このようにすると、例えば脊椎では、脊椎1つ1つの間の角度を脊柱全体の動きから求めることが可能である。

このように、医学的知識や重み付けパラメータ、各部位の重さ、可動域、筋や腱など、人体の運動を制約する様々な要因を基に各関節の動きを算出することにより、変換後の動作データで実際の骨格の動きを再現することは可能である、と考えられる。

しかしながら、これらの要因は互いに影響を与えて変化してしまうなど、非常に複雑であるため、これらの要因全てをモデル化することは困難である。そこで、筋や腱などによる力学的負荷などはシミュレーションできないものの、医学的な測定に基づき、少なくとも骨格的な動き方は実際の人体に忠実に表されている、重み付けパラメータによりマッパーでの変換を行った。詳細なアルゴリズムについては4章に示す。

3.2 空間マウスによる直観的操作

通常のマウスでは、2次元平面内の移動しか表現することができない。そのため、指定した平面内での2次元移動を複数の平面(少なくとも2つ)に対して行うことで3次元移動を行う、3軸の中から1軸を指定し、その軸に対する回転を繰り返して3次元的回転を行うなどの手法がとられ、直観的な操作を行うことができなかった。

これに対し、3次元空間内での操作を行うためのデバイスが、仮想現実感 (Virtual Reality: VR) などの分野で開発されている。このようなデバイスは、大きく以下の3つに分けることができる。

1. 従来のマウスの上部にボタンやトラックボール等をつけることで、奥行き方向の移動、および回転を表せるようにしたデバイス
2. 糸のついたポインタや磁気式センサ等により、3次元空間内での位置・回転を入力できるデバイス
3. 人体の各部に光学式センサや磁気式センサを取りつけ、各部の位置情報から姿勢を推定するデバイス (モーションキャプチャ)

1. のようなデバイスは、確かに従来のマウスとは異なり3次元的な位置・回転を同時に指示することが可能であるが、操作の直観性に欠けている。

また、3. のようなデバイスでは、ユーザ自身が取れるような姿勢しか入力することができないため、

ユーザより体の軟らかい人のモデルを動かそうとしても最大限まで動かすことができない。さらに、空中や水中での姿勢を作成するのが困難である、といった問題がある。

これに対し、2.のようなデバイスでは、直観的に3次元空間内の移動・回転を指示することができる。また、ユーザ自身の体の柔軟性などに左右されず、いわば「手とり足とり」姿勢を教示することが可能となる。

そこで、我々は、2.のようなデバイスにより、直観的な姿勢の教示を実現することを考えた。

4 空間マウスを用いた姿勢定義システムの構築

4.1 空間マウスによる動作定義手順

空間マウスにより、3次元仮想空間内のオブジェクトを操作するためには以下のコンポーネントが必要であると考えられる。

- 位置や角度を測定し、計算機中に取り込むデバイス
- オブジェクトを選択するかしないかを示すボタン
- 3次元空間における空間マウスの位置を示すポインタ

これらを備えた空間マウスにより、図5の手順で操作を行う。

4.2 動作定義支援システムの設計

システムの流れを図6に示す。

このシステムは、SHMで動作を定義し、その動作データを Mapper を用いて BBHM にマップし、シミュレーション・評価を行うものである。動作データや重み付けパラメータなどの値はスキーマに従ってデータベースに蓄積してあるため、データベースに問い合わせを行うことにより、ある個人やある集団（40代の人など）が動作を行った場合の負担評価など、さまざまなシミュレーションを行うことが可能である。

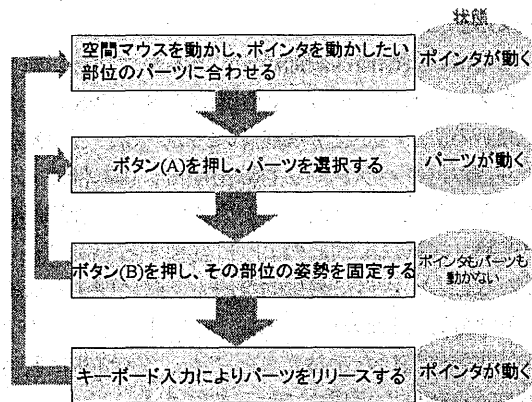


図 5: 空間マウスを用いた姿勢定義手順

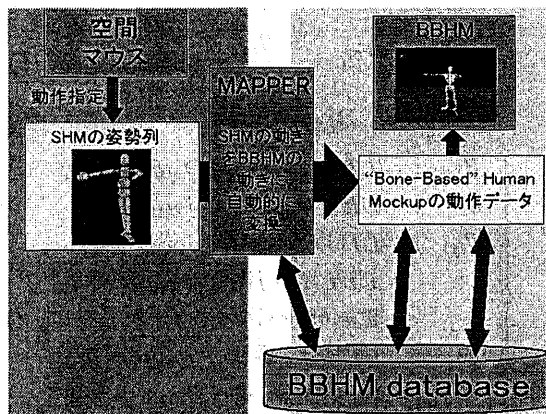


図 6: 空間マウスを用いた姿勢定義手順

我々実際に行われた動作と架空の動作の両方に対して評価を行いたいと考えている。これらを実現するため、SHM への動作定義は以下の2つの手法をとることとした。

1. 映像から動作データを取得する
2. SHM に対し動作データを3次元マウスで直接定義する

また、各 PART の大きさや関節可動域、重み付けパラメータなどをデータベース上にあらかじめ蓄積しておき、このデータを用いて個人に合わせたシミュレーションなどを行う。なお、重み付けパラメータ設定機で設定する重み付けパラメータのデフォルト

ト値は文献 [6] の一般的な人体における特性に基づき設定した。

例えば、腰椎には関節が5つあるが、腰を屈曲させると、一般にこの5つの関節は6:8:9:12:8の割合で曲がるということが知られており [6]、この割合から重み付けパラメータを算出する。重み付けパラメータが個人により決まっていること、一般的な値や年齢、病気によりどのように変化するかが知られていること、さらに具体的な測定手法に関する研究が医学分野で行われていること [7][8] から、このような非常に簡単な手法によっても、文献 [6] に記されている動作の再現までは可能である。

ここで、SHMの関節角度の比から重み付けパラメータを算出する手法、および重み付けパラメータから関節角度を求める手法は、以下の通りである。

$$P_{ij}(\%) = \frac{R_{ij}}{\sum_n R_{nj}} \times 100 \quad (1)$$

$$B_j = S_i \times P_{ij} \quad (2)$$

- | | | |
|---|------------|--|
| { | B_j : | BBHMのj番目の joint 角度 |
| | S_i : | SHMのi番目の joint 角度 |
| | P_{ij} : | SHMのi番目の joint から
BBHMのj番目の joint への
重み付けパラメータ |
| | R_{ij} : | SHMのi番目の joint 角度が
BBHMのj番目の joint 角度
に影響する比 |

先ほどの腰椎の例について実際に計算を行うと、屈曲運動に関して、腰椎第1関節のパラメータは、

$$\frac{6}{6+8+9+12+8} \times 100 = 14.0(\%) \quad (3)$$

また、腰を30度屈曲させた際の腰椎第1関節の角度は、

$$30 \times \frac{14.0}{100} = 4.2(^\circ) \quad (4)$$

となる。

従来、形状の変化しない複数の部品が関節でつながっているモデル(以降剛体リンクモデルと呼ぶ)において、大局的な動きから各関節の動きを求める場合、例えば複数部品からなるアームの先端の位置からアームの各関節の角度を求めるような場合、逆運動学や逆力学が用いられることが多かった。しかし、この手法を人体に対し適用した場合、人体の各部位を一律の“部品”として関節角度を算出するた

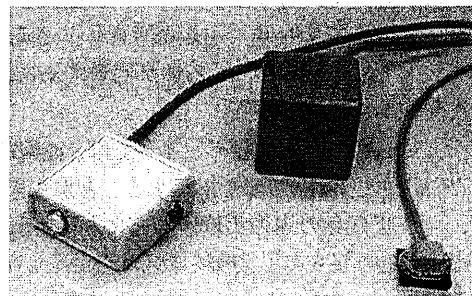


図7: 空間マウス

め、上述の、複数部位の動きの連動のような人体固有の運動機能を表現することができない。このことから、特に骨格レベルでの姿勢・動作の算出を行うためには、マッパーによる姿勢・動作の変換を行うことが適していると考えられる。

4.3 実装システム

今回実装したシステムは、SGI社製グラフィックワークステーション ONIX2 Reality(CPU:R12000 300MHz × 2, メモリ:2GB, OS:IRIX6.5) 上で実装した。SHM および BBHM の表示には米国 delmia 社製エンジニアリングシミュレータ ENVISION TR Version 5.2 [3] を用いた。また、空間マウスにはソリッドレイ社製 3SPACE FASTRAK を用いた。

空間マウスを図7に示す。

4.4 実行例

前節で述べたシステムにおいて、実際に空間マウスを用いて、SHMに動作定義を行い、リアルタイムにBBHMに同じ動作をマッピングする操作を行った。図8にその実行中の様子を示す。また、図9に実行画面を示す。

5 まとめ

本稿では、空間マウスを用いて、SHMに動作の定義を行い、その姿勢・動作をBBHMにリアルタイムにマッピングを行う動作定義支援システムの設計と実装を行った。

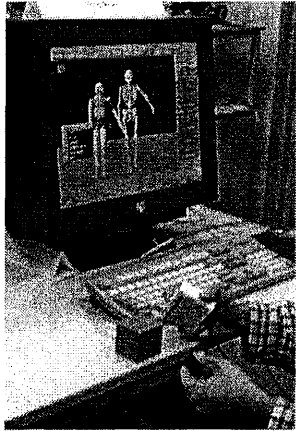


図 8: 空間マウスによる姿勢定義

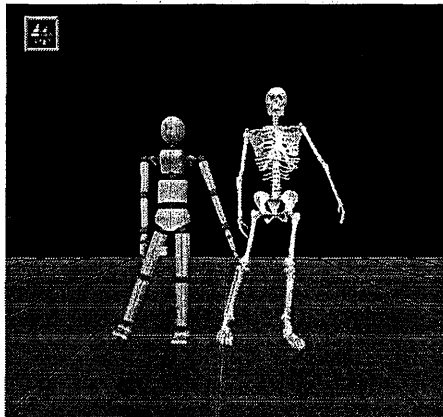


図 9: 姿勢定義の実行画面 (オブジェクトの操作中)

これにより、直観的な動作定義が可能となり、かつ、人体の運動機能に基づいた運動の指定が可能となる。

今後の課題としては、空間マウスを用いて動作を詳細化する手法、データベースへの蓄積のためのスキーマ定義、さらに詳細な姿勢マッピングを行うためのマッパーの詳細化などが挙げられる。

謝辞

本研究を行うに当たり、人体の骨格・関節の構造や、人体の動作に伴う関節の動きなど、横浜市民病院 整形外科部長 酒井直隆 博士に助言を頂いた。謹

んで感謝の意を表す。また、本研究の一部は、文部科学省科学研究費補助金基盤研究 (B)(1) (課題番号 12558026) および (財) 製造科学技術センター IMS 国際共同研究プログラム「生産システムにおける人間・機械組織化の研究」による。

参考文献

- [1] 有澤 博: “リアルワールドデータベースとその実現技術 1-3”, bit, Vol.28, No.9-11, 1996.
- [2] 佐藤 貴子、永野 志保、富井 尚志、有澤 博、酒井 直隆: “骨格モデルを用いた人体動作データベースの構築”, 情報処理学会論文誌:データベース (TOD 8), to be published 2001.1.
- [3] <http://www.delmia.com/solutions/html/ergonomics.htm>
- [4] <http://www.eai.com/products/jack/>
- [5] <http://www.musculographics.com/>
- [6] I. A. KAPANDJI: “関節の生理学”, 医歯薬出版株式会社, 1986.
- [7] 田中 直史, 大久保 衛, 大槻 伸吾: “VICON による肩甲骨の動作解析”, 日本臨床バイオメカニクス学会誌, Vol. 16, pp.217-222, 1995.
- [8] 奥住 成晴, 腰野 富久: “歩行中の正常腰椎矢状面アライメント”, 整形外科バイオメカニクス, Vol.7, pp.155-160, 1985.

以下の頁は著作権者の許諾を得ていないため、公表できません。

抜刷 4 ~ 抜刷 6

抜刷 8 ~ 抜刷 8

抜刷 13 ~ 抜刷 13

抜刷 28 ~ 抜刷 28

抜刷 ~ 抜刷

抜刷 ~ 抜刷

XML を用いたシーンの時空間オブジェクト情報の記述

松田 一章 越前谷 健二 富井 尚志 有澤 博

横浜国立大学 工学部 電子情報工学科

〒 240-8501 横浜市保土ヶ谷区常盤台 79-5

E-mail : {cross, echizenn, tommy, arisawa}@arislabs.dnj.ynu.ac.jp

あらまし

本研究では、シーンの内容が登場者の意味・意図を記述することのできる情報キャリアの設計を行う。現実の世界で行われた事象には、登場者の意図などの意味的な要素のために、様々な解釈がある。これらの意味をデータベースにすべて蓄積すれば、リッチな意味表現を持つ DB が実現できる。一方、VRML などの空間記述言語は、意味的情報を記述するタグ付けが存在しない為、DB からの研究によって得られたシーンデータには、意味的情報を付加することができない。そこで、本研究では DB の一部を切り出して作成したシーン情報に意味的情報を付加できる伝達媒体「時空間オブジェクト情報キャリア」を設計した。また、時空間オブジェクト情報キャリアを XML 文書で表現することでその実現可能性を検証した。

1 はじめに

マルチメディアコンテンツがコンピュータのみならずモバイル端末へと配信される現在であるが、それらは体系化されることなくそれぞれ独自の記述方式に従って、様々なホストの中に存在している。最近特に注目されているのは三次元時空間オブジェクト表現、すなわち、空間内に存在する人間や人間を取り巻く「もの」の時間的な変化を表現し、マルチメディアコンテンツとする技術である [1][2]。

三次元時空間オブジェクト表現には次の3つのレベルがあると我々は考える。

1. 三次元時空間のオブジェクト群と、そのふるまいによって創られる時空間的なイベントを、一般的、かつ、その意味(セマンティクス)までを定義して記述するレベル。
2. 上のレベルの極一部分を取り出し、あるいは再構成して得られるある一つの閉じた時空間的なシーンの記述を行なうレベル。

3. 上のレベルをグラフィック記述言語や VRML などを使用して記述した、言わば、シンタクティカルに表現したレベル。

これらのレベルは全く事なる表現のレベルであると考えられる。

1のレベルはデータベースに記述されるべき情報である。このレベルではシーンを構成するオブジェクト情報と、意味や概念を表している entity 情報とが相互に絡み合っており、また、それらが表す世界が非常にリッチであるため「複雑な構造」をしているものである。3のレベルは三次元表示用データで、特定のプレイヤーを用いてリアルタイムに再生出来るものであり、SMIL、VRML などの記述言語、ないしは 3D CAD、CG 作成ツールなどの商用システムなどが提案され、実装されている。これまではデータベース、三次元表示用データという2つのレベルで考えられており、データベースの一部を取り出した三次元表示用データではそのセマンティクス(データベースで表現されていたそのデータの意味、または検索の際の意図)が欠落し、更にデータ

ベースとは完全に切り離されたものとなっていた。そこで我々は、2のレベルを導入し、単体で時空間的な完全な一つのシーンとなり、更には要素データの由来を背後のデータベースに遡って調べることを可能とするシーン記述用オブジェクトを提案する(図1)。このシーン記述用オブジェクトを以下、時空間オ

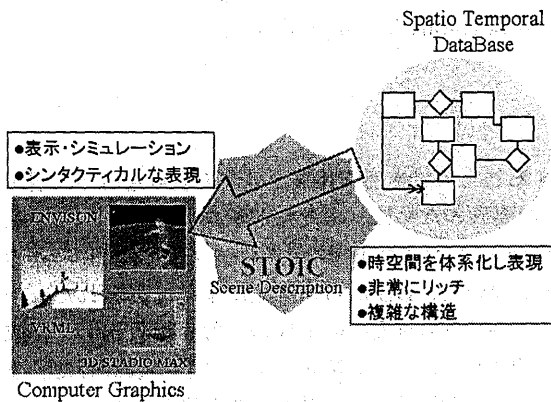


図 1: STOIC の位置付け

ブジェクト情報キャリア (Spacio Temporal Object Information Carrier: STOIC) と呼ぶ。STOIC はそれ単体でデータベースに表現された世界の完全な一つのシーンとなるものであり、また、取り出された結縁を知ることができるためデータベースへの再検索などが期待できる。

本稿ではSTOICのオブジェクト記述方式や明示的な意味注釈の記述方式を設計すると同時にそれを新規のタグ付表現としてXML[3]を用いて記述することを試みた。さらにその有効性と実現性を示すためにXML文書で表現されたSTOICをVRMLに変換し、伝達されたシーンを表示するシステムを実装した。STOICを記述する道具としてXMLを用いたのは、その表現能力の高さと、自己記述型のテキスト文書でありポータビリティに富んでいる特徴を考慮した結果である。以下、2章DBに表現される情報と3次元表現機の表現能力、3章時空間オブジェクト情報伝達子の設計、4章STOICの実現手法と実証、5章でまとめを述べる。

2 DBに表現される情報と三次元表示用データの表現能力

2.1 DBのセマンティクス表現

データベースにおけるセマンティクスとは事物・事象(の集まり)のある一面の解釈を、一つのentityとみなし、それに明示的な注釈を付加したものであると我々は考える[4][5]。そして、データベースを構成するオブジェクトとそれを取り巻くセマンティクスは相互に絡み合い、非常にリッチで「複雑な構造」をなす。

サッカーの例で言えば、「二人の選手がパスをつぎながら、三人のディフェンダーの間を抜いたが、オフサイドになった」という事象を例にして考える(図2)。このシーンは一つの試合の中で起こった事実だが、この解釈・意味は多岐に及ぶ。例えば、「攻撃側の選手A,Bの間をA→B、B→Aという順序でボールが移動した」は「二本のパス」であるという注釈を加えることができる。そして、これら「二本のパス」の意図は、「ワンツーパス」というもう1レベル高次の意味を持ち、「ディフェンダーの動きを攪乱しゴールに向かってボールを移動させる」という意図があった。この意図もまた一つのセマンティクスである。これに加え、ディフェンダーの動きやその解釈を加えると一つのシーンに表れるセマンティクスは多岐に及ぶだろう。このように意味的なつながりと言うものは限りなく広がっていくもので、それらをモデリングするデータベースは「複雑な構造」で「リッチな意味」を記述したものとなる。

2.2 三次元表示用データの表現能力と問題点

三次元表示機 (VRML Viewer, 3D STUDIO MAX, ENVISION...etc) 用のデータの表現方法は、最もローレベルなポリゴンを構成する頂点の座標値というものから、ポリゴンが部品化、構造化された三次元オブジェクトによる表現、他にも様々なものがあるが、どれも単なるサイバーワールドの時系列でしかなく、シーン、または三次元オブジェクトがどういった意味を持っているかは表現されていない。

シーンやそれを構成する三次元オブジェクトの意味が表現されていないとはどういうことか。前節で

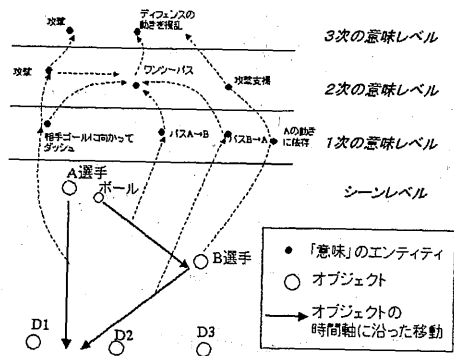


図 2: サッカーの 1 シーンにおけるオブジェクトと意味の entity

述べたサッカーシーンデータベースを例にとると、このシーンは様々な意味を持っているためいろいろな検索要求に適合する。例えば、「サッカーの試合」、「ワンツーパスがあったシーン」、「オフサイドトラップが成功したシーン」などといった検索である。しかし、三次元表示データはどの検索のときも「人間を表す CG オブジェクトが 5 つとボールを表すオブジェクトが動いている」という同じものである。つまり、このデータが上記のどんな検索要求を満たしてデータベースから取り出されたのかという事も表現されていない。更に、「サッカーの試合」として検索を行なった結果として CG データが検索者に渡ったとしても検索者はその CG から「サッカーを行なっている選手の名前」などといったシーンを構成するオブジェクト (またはシーンそのもの) のセマンティクスを得ることはできない。検索者が「サッカーを行なっている選手の名前」というセマンティクスを必ずしも知っているとは限らない。そして、検索者がそういった情報を望む事は決して少なくない。

また他にも問題点は存在する。三次元表示機用のデータ表現での目的とするところは「正確な視覚化」であるために、必ずしも表現している三次元オブジェクトの構造等が実際の構造や機能を忠実に再現したものであるとは限らない。人間の手腕の表現を例に挙げる。人間の手腕は橈骨と尺骨から成り、

実際、手腕の捻りはその両者の動きで成立しており、医学の分野などにおいては忠実に表現する必要がある [6]。しかし、これは単純な階層構造では表現できないため、一般に三次元表示機では橈骨と尺骨の他に疑似的なパーツを用意しそのパーツを動かすという事で橈骨と尺骨の動きを表現することが多い。このように、計算機を用いて視覚化するためにシタクティカルに表現する事と実際の「もの」を忠実に表現する事は全く別である。

このように三次元表示用データの表現能力はデータベースのそれより格段に低いだけでなく、時空間データベースの 1 シーンとしては不十分であると我々は考える。

3 時空間オブジェクト情報キャリアの設計

3.1 シーン記述用オブジェクトの導入

前章では時空間データベースと三次元表示用データという 2 つのレベルの表現能力のギャップについて述べた。一般に三次元表示用データにはシーンやシーンを構成するオブジェクトのセマンティクスを包含していない。このため、三次元表示用データをデータベースの一部を再構成した「シーン」として利用するには問題がある。

この問題を解決するために、「タイプとインスタンス」、「オブジェクトとパーツ」というような論理的、あるいは物理的な概念的構造を厳密に記述でき、更に要素データの由来をデータベースに遡って調べることができるシーン記述用オブジェクトが必要であると考えられる。このようなシーン記述用オブジェクトを時空間オブジェクト情報キャリア (Spacio Temporal Object Information Carrier: STOIC) と呼ぶ。STOIC はデータベースから三次元表示用データへ情報を伝達するものであり、この STOIC の導入により、時空間を非常に厳密、正確に表現したデータベースとその視覚化のための三次元表示用データとのギャップを埋めることができる。さらに、データベースとのリンクを持つという STOIC の特徴により STOIC に表現されるシーンを用いた再検索を期待できる。

3.2 STOIC(時空間オブジェクト情報キャリア)

前節ではシーン記述用オブジェクト STOIC を提案した。STOIC はシーンやシーンを構成するオブジェクトのセマンティクスが包含され、それ単体でデータベースに表現された世界の完全な一つのシーンとなることができ、更にデータベースから取り出された後もデータベースとのリンクを保つためのものである。それには、時空間オブジェクトとそのオブジェクトのセマンティクスが記述できる構造でなくてはならない。

コンピュータグラフィックスで3次元オブジェクトを表現しようとするとその表現手法は様々である。例えば全ての表面形状を点と面で表現する方法、簡単な基本図形を組み合わせる方法などである。これは見た目を表現しているだけで、そのオブジェクトの本質を表現していない。また、様々なエンジニアリングシミュレータではオブジェクトをそれを構成するパーツの階層構造として表現することがある。そのメリットとしてパーツ同士の親子関係を記述することによって、例えば人間で言うと、肩関節を動かして腕をあげると、肩関節に連結する上腕だけでなく、上腕から下に連結している部品群(前腕、手、指)も同時に移動させることが簡単に表現できるからである(図3)。しかし、この構造表現もその本質を表していない。なぜなら、図3のように人間は必ずしも腰が基準となっているわけではなく、そのパーツ同士には束縛しあう制約はあるにしろ、それが階層構造になっているとはいえない。また、このような階層構造を用いて表現をしようとしたときに親が二つ以上ある場合などは表現できない。よって、3次元オブジェクトの構造表現において、パーツごとの階層構造で表現することは本質的でないといえる。

そこで本論文では、佐藤らが提案した、パーツとジョイントとコネクションという概念を用いて3次元オブジェクトの構造を表現する[6]。ここでいうパーツとはオブジェクトの一部分で自由度があるオブジェクトであるとする(例えば人間の胴体など)。ジョイントとはパーツとパーツを結んでいる場所の概念的なものである(例えば人間の肘など)。また、コネクションとはパーツとジョイントを1対1で結んでいるものである(図4)。このような表現手法を

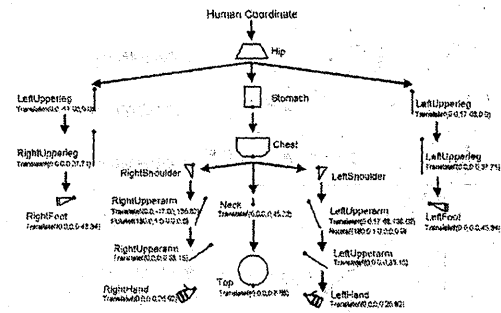


図3: 人体の階層構造モデル

用いることによって、本質的に階層構造ではないものも表現でき、また、親が二つ以上ある場合も表現する事が可能である。

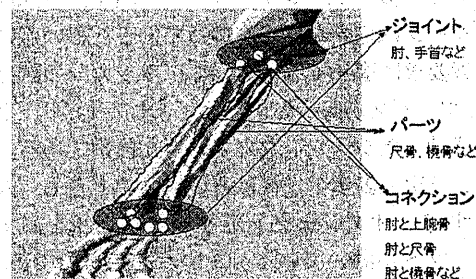


図4: パーツ、ジョイント、コネクションの概念

STOIC ではパーツはその形状を記述するために、エレメントパーツという概念を用意する。エレメントパーツとはパーツの単なる形状を表現する部分で、それ自体に意味はない、いわば視覚的に表現するためのものである。エレメントパーツはいくつかの簡単なモデルまたはポリゴンの組合せでその形状を表現し、また、それらの形状に対して、位置・角度・色の情報を付加することによって、オブジェクトの視覚的な表現を記述する。形状を表すモデルはすべての点の情報を与えることなくその本質を表すパラメータを与えることによって形状を表せるためのもので今回は箱型・円錐・円柱・球を用意した。

また、それらに対し x 軸 y 軸 z 軸に対しての拡大・縮小率を定義することにより様々な形状を表現できるようになっている。また、球のモデルでは、単なる球を表現するだけでなく、その球の表面上の点を中心からの距離を拡大・縮小することができ、それによって、様々な凹凸を表現することができるようになっている。このように頂点データで表現するといった離散的な記述ではなく連続的な記述を行なうことによって、形状の本質的な表現ができると考える。そして精密な表示を行ないたいときはポリゴン数を多く設定し離散化し、また、形状がさほど重要ではない時には、そのポリゴン数を減らし離散化を行えば良い。つまり、オブジェクトの形状は連続的に表現するべきであり、計算機を用いて視覚化する際(シンタクティカルな CG データに変化する時)に離散化されるべきであると考え。よって、精密なモデルと粗いモデルを別々に作成する必要がなくなる。

以上、STOIC における 3 次元オブジェクトの表現をまとめたものを図 5 に示す。

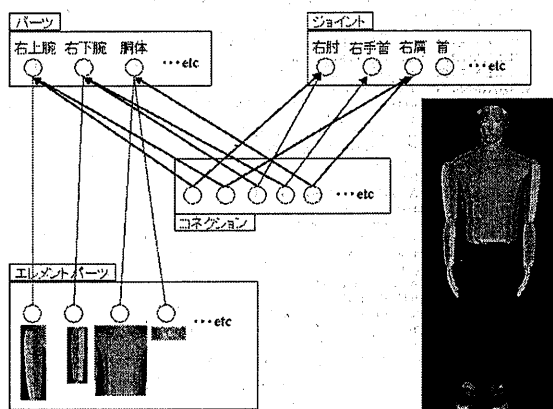


図 5: STOIC での 3 次元オブジェクトの表現

STOIC においてオブジェクトの動きはある時点におけるオブジェクトの姿勢と位置の時系列によって表現する。この時の姿勢とはオブジェクトの構造で定義したコネクションに対して回転角度を与えることにより、オブジェクトを構成するそれぞれのパーツを回転・移動させ表現されるものである。また、その姿勢の時系列で表現された空間的な移動にその動きの意味を付加し動作と定義する。以上のよ

うにして、時空間を視覚的に表現するデータとそのシーンに含まれている意味データを関係付けある閉じたひとつのシーン (STOIC) を表現した。

4 STOIC の実現手法と実証

4.1 XML による STOIC の実装

3 章では時空間データベースの 1 シーンを表現する枠組の必要性を示し、その枠組として STOIC を提案した。本章では STOIC の実装について述べる。本論文では STOIC の構造を XML によって実装した。XML を用いる理由を次に挙げる。

1. 半構造データであり表現能力が非常に高い
2. 自己記述型のテキスト文書である
3. 昨今のポピュラーな言語であり様々なツール、API が用意されている

(1) STOIC は対象とするデータベースによって異なるデータ構造を持つ。そのため、構造自体を DTD (Document Type Definition) や XML スキーマによって記述できる XML は有効である。また、DTD や更にそれを凌ぐ XML スキーマの表現能力は非常に高く STOIC を実現するのに十分であると考え。

(2) STOIC はデータベースから離れ様々なホストに存在する事になる。(1) で述べた様に STOIC のデータ構造は唯一ではないため、それ自身がどのような構造をしているかを記述してある必要がある。さらに、その構造がユーザーに直観的であるようにテキスト文書である事は有効である。

(3) 次節で述べるが STOIC を視覚化するために様々な三次元表示用データのフォーマットに変更する必要がある。その際に STOIC を構文解析、変換する必要があり、そういった API が充実した XML は STOIC の実現に非常に有効である。

以上の理由により本論文では STOIC を XML を用いて実装した。また、3 章で述べた STOIC の構造を DTD を用いて実装した

(図 6,7)。

実装した DTD に基づいて今回は 1 人の人間 (名前: ergo) が歩行しているシーンを記述した (図 8)。

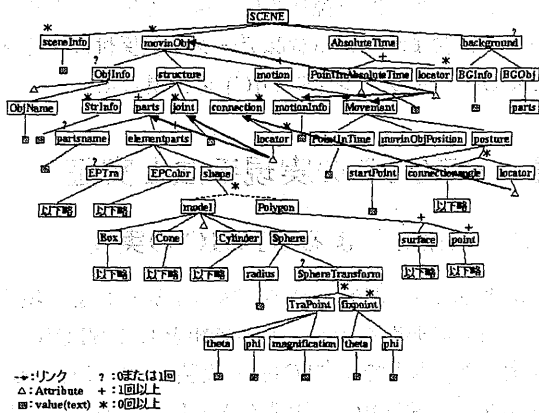


図 6: STOIC の構造 (シーン)

4.2 STOIC の視覚化

前章までに、STOIC がそれ単体で 1 つの閉じた時空間的なシーンを記述することができ、また、XML を用いることにより、計算機上で表現可能なことを示した。そこで本章では、STOIC を視覚的直感的にも表現できることを示し、その有用性や実現性を評価するために、STOIC を三次元表示機用のデータに変換するシステムを構築する手法について述べる。

今回は、XML で表された STOIC を VRML (Virtual Reality Modeling Language) に変換することにした。VRML を用いた理由として以下のようなことが挙げられる。

- テキストベースなので特別なオーサリングツールがなくても誰でも簡単に VRML ファイルを作成することができる。
- VRML を表示するブラウザは様々なところで無料で手に入るなので誰でも手軽に VRML を利用することができ、世の中によく普及している

前に述べた XML 文書を VRML に変換することを考える。VRML では、オブジェクトの形状 (パーツ) 定義と、定義されたオブジェクト群の移動 (平行移動、回転) の記述によって最小構成される。また、一般的に VRML ではパーツの集合から成る集

```
<?xml encoding="Shift_JIS"?>
<ENTITY % remote-resource-semantic. att
"role CDATA #IMPLIED
title CDATA #IMPLIED
show (embed | replace | new) #IMPLIED
actuate (auto | user) #IMPLIED
behavior CDATA #IMPLIED">
<ENTITY % local-resource-semantic. att
"content-role CDATA #IMPLIED
content-title CDATA #IMPLIED">
<ELEMENT scene (sceneInfo*, movinObj*, AbsoluteTime,
background?)>
<ELEMENT sceneInfo (#PCDATA)>
<ELEMENT movinObj (ObjInfo, structure, motion+)>
<!ATTLIST movinObj
id ID #REQUIRED>
<ELEMENT ObjInfo (ObjName?, ObjAge?, ObjHeight?,
ObjWeight?)>
<ELEMENT ObjName (#PCDATA)>
<ELEMENT ObjAge (#PCDATA)>
<ELEMENT ObjHeight (#PCDATA)>
<ELEMENT ObjWeight (#PCDATA)>
<ELEMENT structure (StrInfo*, parts+, joint*, connection*)>
<ELEMENT StrInfo (#PCDATA)>
<ELEMENT parts (partname?, elementparts+)>
<ELEMENT partname (#PCDATA)>
<ELEMENT elementparts (elementpartsTransform?,
elementpartsColor, shape)>
<ELEMENT elementpartsTransform (elementpartsPosition?,
elementpartsRotation?, elementpartsMagnification?)>
<ELEMENT elementpartsPosition (x, y, z)>
<ELEMENT elementpartsRotation (Rx, Ry, Rz, angle)>
<ELEMENT angle (#PCDATA)>
<ELEMENT elementpartsMagnification (x, y, z)>
<ELEMENT elementpartsColor (R, G, B)>
<ELEMENT shape ((model | Polygon)*)>
<!ATTLIST model
visualize (yes | no) #IMPLIED>
<ELEMENT model (Box | Cone | Cylinder | Sphere)>
<ELEMENT Box (width, height, length)>
<ELEMENT width (#PCDATA)>
<ELEMENT height (#PCDATA)>
<ELEMENT length (#PCDATA)>
<ELEMENT Cone (bottomRadius, height)>
<ELEMENT bottomRadius (#PCDATA)>
<ELEMENT Cylinder (radius, height)>
<ELEMENT radius (#PCDATA)>
<ELEMENT Sphere (radius, SphereTransform?)>
<ELEMENT SphereTransform ((TraPoint, fixpoint*))>
<ELEMENT TraPoint (theta, phi, magnification)>
<ELEMENT theta (#PCDATA)>
<ELEMENT phi (#PCDATA)>
<ELEMENT magnification (#PCDATA)>
<ELEMENT fixpoint (theta, phi)>
<ELEMENT Polygon ((surface, point+)+)>
<ELEMENT surface (surfacenumber)>
<ELEMENT surfacenumber (#PCDATA)>
<ELEMENT point (x, y, z)>
<ELEMENT joint (#PCDATA)>
<ELEMENT connection (locator*)>
<!ATTLIST connection
id ID #REQUIRED
xml:link CDATA #FIXED "extended"
inline (true | false) "true"
%remote-resource-semantic. att;
%local-resource-semantic. att;
<ELEMENT locator EMPTY>
<!ATTLIST locator
xml:link CDATA #FIXED "locator"
href CDATA #REQUIRED
%remote-resource-semantic. att;
<ELEMENT motion (motionInfo+, Movement+)>
<ELEMENT motionInfo (#PCDATA)>
<ELEMENT Movement (PointInTime, movinObjPosition,
posture)>
<ELEMENT PointInTime (#PCDATA)>
<ELEMENT movinObjPosition (X, Y, Z)>
<ELEMENT posture (startPoint, (connectionangle, locator*))>
<!ATTLIST posture
xml:link CDATA #FIXED "extended"
inline (true | false) "true"
%remote-resource-semantic. att;
%local-resource-semantic. att;
<ELEMENT startPoint (X, Y, Z)>
<ELEMENT connectionangle (Rx, Ry, Rz, angle)>
<ELEMENT AbsoluteTime (PointInAbsoluteTime, locator*)+>
<ELEMENT PointInAbsoluteTime (#PCDATA)>
<!ATTLIST PointInAbsoluteTime
id ID #REQUIRED
xml:link CDATA #FIXED "extended"
inline (true | false) "true"
%remote-resource-semantic. att;
%local-resource-semantic. att;
<ELEMENT background (BGInfo*, BGObj*)>
<ELEMENT BGInfo (#PCDATA)>
<ELEMENT BGObj (parts+)>
```

図 7: シーン DTD

```

<?xml version="1.0" encoding="Shift_JIS"?>
<!DOCTYPE scene SYSTEM "scene.dtd">
<scene>
<sceneInfo>Man is walking.</sceneInfo>
<movinObj id="Ergo1">
<ObjInfo><ObjName>Ergo</ObjName></ObjInfo>
<structure>
<parts>
<partsname>head</partsname>
<elementparts>
<elementpartsTransform>
<elementpartsRotation>
<Rx>1</Rx><Ry>0</Ry><Rz>0</Rz>
<angle>4.71239</angle>
</elementpartsRotation>
</elementpartsTransform>
<elementpartsColor>
<R>0.3</R><G>0.3</G><B>0.3</B>
</elementpartsColor>
<shape><Polygon>
<surface><surfacenumber>1</surfacenumber></surface>
<point>
<x>0.0649784</x><y>0.0668291</y><z>0.768533</z>
</point>
<point>
<x>0.0544968</x><y>0.0482136</y><z>0.759036</z>
</point>
<point>
<x>0.0435708</x><y>0.0757098</y><z>0.743848</z>
</point>
<surface><surfacenumber>2</surfacenumber></surface>
<point>
...
</point></Polygon></shape>
</elementparts>
</elementparts>
</parts>
</parts>
<joint>neck</joint>
<joint>RightShoulder</joint>
<joint>LeftShoulder</joint>
...
<connection id="head-neck">
<locator href="id(Ergo1).child(1,structure)
.children(1,parts)" role="parts"/>
<locator href="id(Ergo1).child(1,structure)
.child(1,joint)" role="joint"/>
</connection>
<connection id="body-neck">
<locator href="id(Ergo1).child(1,structure)
.child(2,parts)" role="parts"/>
<locator href="id(Ergo1).child(1,structure)
.children(1,joint)" role="joint"/>
</connection>
<connection id="body-rightshoulder">
...
</connection>
</structure>
<motion>
<motionInfo>advance</motionInfo>
<motionInfo>walking</motionInfo>
<Movement>
<PointInTime>0</PointInTime>
<movinObjPosition>
<X>0</X><Y>0</Y><Z>0</Z>
</movinObjPosition>
<posture>
<startPoint>
<X>-0.0571652</X><Y>0.631467</Y><Z>-
0.0367109</Z>
</startPoint>
<connectionangle>
<Rx>0</Rx><Ry>0</Ry><Rz>0</Rz>
<angle>0</angle>
</connectionangle>
<locator href="id(head-neck)" />
<connectionangle>
<Rx>0</Rx><Ry>0</Ry><Rz>0</Rz>
<angle>0</angle>
</connectionangle>
<locator href="id(body-neck)" />
...
</posture></Movement></motion>
</movinObj>
</AbsoluteTime>
<PointInAbsoluteTime id="2001/02/04/05:43:45.00">
00:00:00.00</PointInAbsoluteTime>
<locator href="id(Ergo1)" title="toMovinObj"/>
<locator href="id(Ergo1).child(1,motion)
.child(1,motionInfo)" title="toMotionInfo" role="start"/>
<locator href="id(Ergo1).child(1,motion)
.children(1,Movement)" title="toPointInTime" />
...
</AbsoluteTime>
</scene>

```

図 8: XML インスタンス (ergo の歩行)

合オブジェクト (例えば、人間) を定義する。その理由は集合オブジェクト全体を移動させる際に便利だからである。

まず、オブジェクトの形状 (集合オブジェクト、パーツ) について考える。XML(STOIC) において集合オブジェクトは<movinObj>タグに相当する。そこで、VRML においては Transform ノードに変換する。集合オブジェクトを構成するパーツ形状は<elementpars>タグ以下に記述されているので、一つの<elementpars>タグごとに一つの Transform ノードに変換する。これは集合オブジェクトの Transform ノードの子ノードに追加する。Transform ノード以下については、<EPCOLOR>タグ (パーツの色データ) を appearance ノードに、<shape>タグを geometry ノードに変換する。以上の手順をすべてのパーツに適応し VRML におけるオブジェクトの形状定義が記述できる (図 9)。

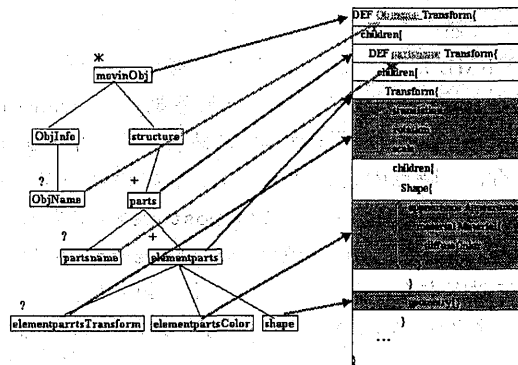


図 9: オブジェクト形状の変換

次に、オブジェクトの移動を XML 表現から VRML 表現に変換することを考える。一般に VRML では先に述べた集合オブジェクトの移動とそれを構成するそれぞれのオブジェクトの移動で表現する。形状定義の際に集合オブジェクトごとに、Transform ノードを、その children(子ノード) にパーツ要素の Transform ノードを用意し、その構造を表現しているため、VRML での動きの表現はこれらの Transform ノード中の translation フィールド、rotation フィールドの値を時点ごとに変化させることによって表現できる。つまり、集合オブジェクトの

Transration ノードで集合オブジェクト全体の動きを、それぞれのパーツの Transration ノードでパーツの動きを記述する。以上の事を踏まえ、XML から VRML への変換を考える。

最初に集合オブジェクトの移動について述べる。集合オブジェクトの移動は XML 文書の時点<Movement>要素の中の<movinObjPosition>要素中に示されている。これを図 10 に示すように変換する。

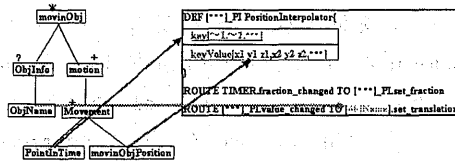


図 10: 集合オブジェクト移動の変換

次に集合オブジェクトを構成するパーツの移動について述べる。

VRML において、パーツの移動はパーツごとにその rotation フィールド (回転移動)、と transration フィールド (平行移動) に値を与える事で記述する。一時点の集合オブジェクトの移動は、STOIC から VRML への変換は<posture>タグ要素の<conectionangel>の値をその<connection>タグ要素に対応する rotation フィールドの値とし、すべてのコネクションについて行なう事で成立する (図 11)。

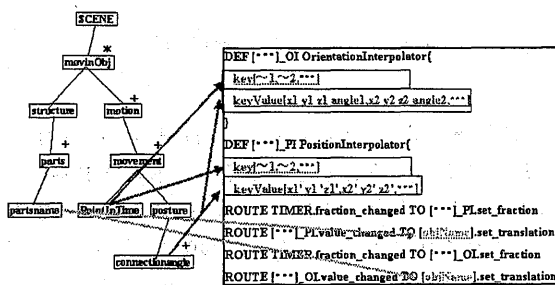


図 11: オブジェクト構成パーツ移動の変換

4.3 STOIC の実証

前節では XML によって実装された STOIC (あるシーン) を視覚化する為の方法として XML から VRML への変換について述べた。この変換方法を用いて、実際に図 8 のシーンを VRML データに変換した (図 12)。

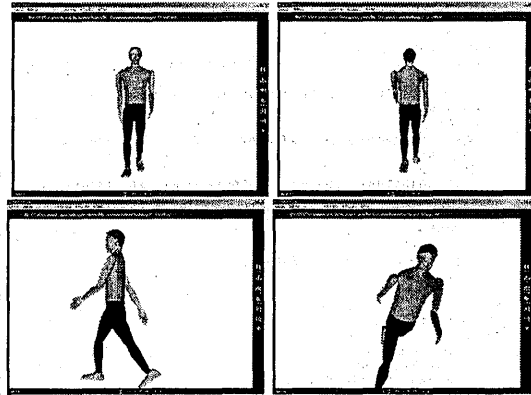


図 12: 変換された VRML データの表示結果

5 まとめと課題

本研究では、シーンの内容が登場者の意味・意図を記述することのできる情報キャリアの設計を行った。現実の世界で行われた事象には、登場者の意図などの意味的な要素のために、様々な解釈がある。これらの意味をデータベースにすべて蓄積すれば、リッチな意味表現を持つ DB が実現できる。一方、VRML などの空間記述言語は、意味的情報を記述するタグ付けが存在しない為、DB からの研究によって得られたシーンデータには、意味的情報を付加することができない。そこで、本研究では DB の一部を切り出して作成したシーン情報に意味的情報を付加できる伝達媒体「時空間オブジェクト情報キャリア」を設計した。また、時空間オブジェクト情報キャリアを XML 文書で表現することでその実現可能性を検証した。

今後の課題として、STOIC の表現方法 (オブジェクト、意味、時間等)、実際のシーンへの適応による検証、STOIC のデータベースとのリンク機構の実装とそれを利用した再検索についての考察、更に

様々な空間記述方式への変換アルゴリズムの開発などが挙げられる。また、STOICの構造の実現手法としてDTDよりも柔軟な表現が可能なXMLスキーマの導入について検討が必要である。

参考文献

- [1] MPEG7
<http://www.darmstadt.gmd.de/mobile/MPEG7/>
- [2] Synchronized Multimedia Integration Language (SMIL) <http://www.w3.org/TR/REC-smil/>
- [3] Extensible Markup Language (XML)
<http://www.w3.org/XML/>
- [4] H.Arisawa,H.Nagae,Y.Mochizuki:“Representation of Complex Objects in Semantic Data Model“AIS” and Implementation of Set Operators”,IEICE TRANSACTIONS,vol.E74,No.1,1991.
- [5] H.Arisawa,T.Tomii,H.Yui,and H.Ishikawa: “Data Model and Architecture of Multimedia Database for Engineering Applications”, IEICE TRANS.INF. & SYST.,vol E78-D,No.11,1995.
- [6] 佐藤貴子,永野志保,富井尚志,有澤博,酒井直隆: 骨格モデルを用いた人体動作データベースの構築, 情報処理学会論文誌: データベース, Vol.42 No.SIG1(TOD8),p92-102,Jan.2001

マルチメディアデータブラウジングのための視覚的データベース インタフェース AVIS の実装

家富 誠敏[†] 森寄 慎之[†] 富井 尚志[†] 有澤 博[†]

An Implementation of Database Visual Interface AVIS for
Multimedia Data Browsing

Masatoshi IETOMI[†], Noriyuki MORISAKI[†], Takashi TOMII[†], and Hiroshi ARISAWA[†]

あらまし 画像、映像、コンピュータグラフィックスなどを扱うマルチメディアデータベースであればこそ、その中身を視覚的にブラウジングすることにより個々のデータやデータ間の関係（つながり、リンク、相互参照）などの状況を直観的に把握したいという要求は強い。本論文では、特にデータ間の関係を視覚化することに着目し、データベーススキーマ及びインスタンスを意味データモデルのダイアグラム表現を用いてデータベースの概念構造そのままに視覚化するインタフェースの設計及び実装手法について述べる。その際、大量のデータ、特にインスタンスを限られた画面内で有効に視覚化する手法として、必要なデータのみを選択して視覚化する機構を実現している。

キーワード 視覚的インタフェース、データベースインタフェース、意味データモデル、ダイアグラム表記

1. ま え が き

近年、計算機技術の発達及び社会の高度情報化が進むにつれ、データベースで取扱いを要求されるデータは、マルチメディア化、大容量化、そして複雑化してきている。映像情報を例に挙げると、映像の中に何が映り込んでいるかを解析し、被写体、主題などをインデックスとして様々な検索を行い、検索結果を新しい映像として再構成できるような映像データベースシステムを構築することが考えられる。またコンピュータグラフィックスを例に挙げれば、オブジェクト（物体）ごとの形状、テキストチャに加えて各時点での姿勢（空間における位置や向き）などをデータベースに蓄積しておき、オブジェクト名をキーとした検索により任意のオブジェクトに関するデータのみを取り出し、これをコンピュータグラフィックスとして再生することなどが考えられるであろう。これらの場合、データベースでは結果として非常に大量かつ多種多様なデータが互いにリンクされているような相互参照の構造を取り

扱うことが必要になる。

ところで、大容量の画像や映像、コンピュータグラフィックスなどを扱うマルチメディアデータベースであればこそ、その中身を視覚的にブラウジングすることによって、データベースの内容を直観的に把握したいという要求が生まれるのも当然のことである。例えば映像データベースにおいて、大量のフレームデータの中から「ある被写体が写っているフレームのみを、被写体がどこに写っているかという情報とともに一覧したい」またはデータベースに蓄積されたコンピュータグラフィックスデータの中から「あるシーンにおける各オブジェクトの動き（各時点ごとの姿勢）をオブジェクトの形状、テキストチャの情報と組み合わせで一覧したい」といった注目する部分のデータだけを取り出して自由に表示し、ブラウジングできる機能があれば、マルチメディアデータベースの中身を直観的に把握できる有用なツールとなると期待される。しかし、これら映像データベースやコンピュータグラフィックスのデータベースに対してみても、データのモデリング（スキーマ化）の手法が必ずしも確立していないことや、データベースの内容を見たいユーザが必ずしもデータベースの構造を熟知しているとは限らないこと

[†] 横浜国立大学工学部電子情報工学科, 横浜市

Division of Electrical and Computer Engineering, Faculty of Engineering, Yokohama National University, Yokohama-shi, 240-8501 Japan

から、このような表示やブラウジングを行うためには、単純にデータの集合や列を表示するだけでなく、データベース中のデータ同士の関係（リンクや参照、意味的なつながり）までを含めて視覚化することが重要である。

これに対して旧来よりデータベースの概念構造、特にスキーマを記述する手法として、ERモデル等の意味データモデルに基づく図式的表現（ダイアグラム）が有用であることが知られている。これを独自に拡張した拡張ERダイアグラムでは、スキーマだけでなくインスタンスも含めてデータベースの概念構造をそのまま記述することができ、このダイアグラムを用いて視覚化を行えば、意味的なデータ間のつながりも直観的に把握することができると期待される。

本研究ではこのような考えに基づき、データベースの概念構造をダイアグラムを用いて視覚化し、データブラウジングを行うことができるインタフェースの実装を行った。その際、特に大量データを扱う上で問題となる、限られた画面を有効に利用してデータを表示する手法として、利用者が必要としているデータのみを視覚化して画面に表示する機構を実現した。これによって大規模なスキーマ、大量のインスタンスに対しても、利用者の求める情報を効果的にブラウジングすることを可能にしている。

本論文では、まず2.でデータベースの視覚化及び視覚的インタフェースに関する一般的な研究を紹介し、次に3.で我々が新たに提案する視覚的インタフェースにおいて要求される事項とその実現手法を議論し、4.で実装したインタフェースAVISの機能について述べる。また5.でAVISを用いたデータベースの視覚化例を示し、6.でまとめを述べる。

2. データベースの視覚化研究

近年、情報視覚化 (Information Visualization: IV) の研究 [1]~[4] は特に盛んであり、データベースの視覚化や視覚的インタフェースに関する研究も報告されている。

- データベース視覚化・視覚的検索に関する研究
 - ー 各種のグラフを用いて視覚化 [5], [6]
 - ー 自己組織化マップ [7] を用いて視覚化 [8]
 - ー クラスタリング結果をカラーアイコンとして視覚化 [9]
- 視覚的データ定義に関する研究
 - ー スキーマ図 (オブジェクト図 [10], ER ダイア

グラム [12] など) を描くことにより、スキーマ定義を行う [13], [14]

文献 [5] 及び [6] では、データベースから検索言語を用いて視覚化対象のデータを切り出し、これを自動的にグラフにマップする機構を提案している。また文献 [8] では、特に検索エンジンの検索結果 (web 文書) に着目し、これを自己組織化マップを用いて視覚化し対話的検索に応用している。更に文献 [9] では、複数の属性値をもつデータ集合を属性値でクラスタリングし、属性値により色が変わるアイコンを利用して、データを視覚的に表現している。しかし、これらの研究では、単一のデータ集合やデータとその属性、または一部分のデータ間の関係については視覚化しているが、データベース全体にわたるデータ間の関係 (データ間のリンクや参照、つながり) の視覚化は行っていない。そこで、本論文では特に、データベース中のデータ間の関係に着目し、データベースを概念構造そのままに、すなわちデータ及びデータ間の関係をそのまま視覚化することを目的とした。これによりデータ間に複雑な相互参照をもつマルチメディアデータに対しても、直観的にデータベースをブラウジングすることが可能になる。

データベースの概念構造そのままを視覚化し、この上でデータ操作を行うというアプローチは文献 [13] でも行われており、いくつかの製品化 [14] も行われているが、これらはスキーマ情報のみしか取り扱っておらず、スキーマに比べてはるかに複雑かつ膨大なインスタンスの視覚化は行っていない。インスタンスを視覚化するためには、単にデータを表示するだけでなく、大量のデータを限られた画面内に有効に表示する工夫が必要である。本研究では、スキーマだけでなくインスタンスも含めて、データベースの概念構造をそのまま視覚化するインタフェースの構築を目指す。

3. 視覚的インタフェースの設計

データベースに対する直観的なブラウジングを実現するためには、データベースを視覚化して画面に表示する視覚的インタフェースが有用である。このような視覚的インタフェースを実現するためには、少なくとも次の要求を満たさなければならないと考えられる。

- データベースの直観的な視覚化
- 必要なデータのみをブラウジングする機構
- マルチメディアデータの画像、映像、グラフィックス動画などによる表示・再生

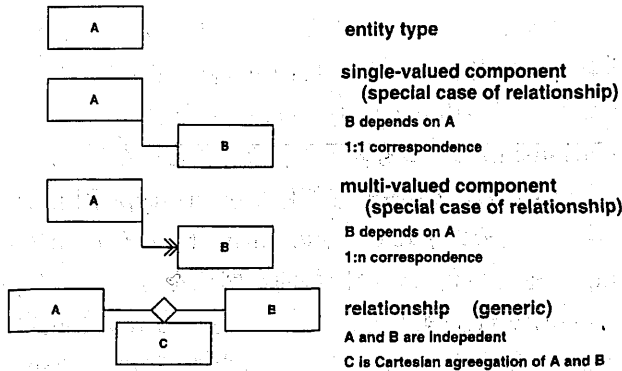


図1 ダイアグラムの構成要素
Fig.1 Primitives of diagram.

以下、それぞれの要求を実現するため、必要とされる機能とその実現手法について考察する。

3.1 データベースの視覚化

データベーススキーマを視覚的に記述する方法としては、旧来よりデータベース設計時などに用いられてきたダイアグラム表記法 [11], [12] がある。とりわけ意味データモデルの一種である ER モデル [12] で用いられる ER ダイアグラムはデータベースの概念構造を表すのに有用であることが知られている。我々は ER ダイアグラムを拡張した表記法として拡張 ER ダイアグラム (AIS ダイアグラム) [15] を提案し、かつ本ダイアグラムが特にマルチメディアデータを扱うのに適していることを既に示している [16]。

拡張 ER ダイアグラムは図 1 に示す基本構成要素から成り立っている (拡張 ER ダイアグラムについては付録参照)。図 2 はこのダイアグラムを用いてスキーマを表現した例である。ここでは entity-type (共通の性質をもつ entity の集合) が長方形のボックスで表され、それらの間の関係 (relationship と呼ぶ^(注1)) がボックスを結ぶ線で示されている。

ダイアグラムを用いることにより、entity-type 及び entity-type 間の関係 (意味的なつながり合い) を図示することができる。これにより、スキーマ情報をデータベースの概念構造そのままに視覚化することが可能になる。このようなダイアグラムを用いてスキーマを視覚化することが、利用者の直観的スキーマ理解のために有効である。

更に、本論文ではスキーマだけでなく、インスタンスも概念構造そのままに視覚化を行うが、特にスキーマとインスタンスとの対応関係を明確に表現するため、インスタンスもまたダイアグラムを用いて視覚化する

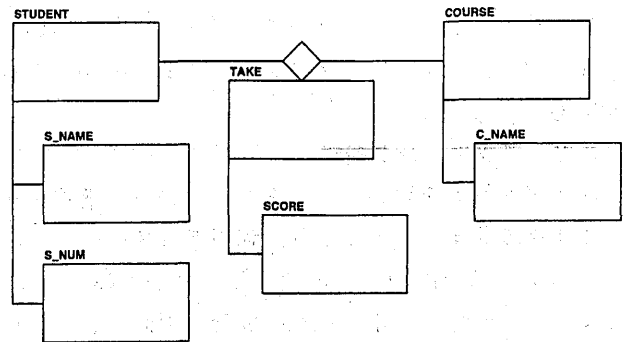


図2. スキーマ図
Fig.2 Schema diagram.

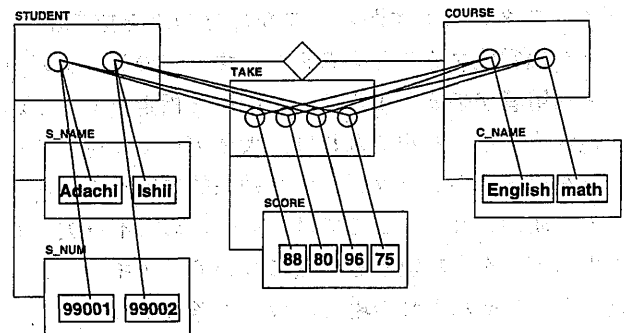


図3 インスタンス図
Fig.3 Instance diagram.

ことを考えた。図 3 は独自にダイアグラムを用いてインスタンスを視覚的に表現した例である。現実世界の事物・事象を表す entity は、自分が属する entity-type のボックスの中に丸 (値をもたない場合) 若しくは四角 (値をもつ場合) で表示される。また、各 entity 間の意味的なつながりは entity 間を結ぶ直線で表現される。ダイアグラムでは、entity 間のつながりがデータベース内で明示的に表現されているが、データ間の参照構造が入り組んでいるマルチメディアデータベースでは、このつながりを明示することは非常に重要である。個々の entity とそのつながりを視覚的に表現することによって、インスタンス全体の構造を直観的に把握することが可能になるからである。映像やグラフィックスのデータベースにおいても、各々のデータがどのようにつながり合っているのか、全体としてどのような構造になっているのかを理解し、効果的なブラウジングを行うことができる。

視覚的インタフェースを用いて、データベースの概念構造を視覚的・直観的に把握するためには、このよ

(注1) : relationship-type 等と呼ばれることもあるが、本論文では単に relationship と呼ぶ。

うに意味データモデルに基づきダイアグラムを用いたスキーマ図、インスタンス図を表示することが大変有効であると考えられる。

3.2 ブラウジング機構

ダイアグラムを用いてデータベースの概念構造そのまま視覚化することにより、利用者はスキーマ及びインスタンスを直観的に把握することが可能になる。しかし、データ量が非常に多い場合、画面に表示された大量のデータに対して、利用者が有効なブラウジングを行うことは非常に困難であるといえるだろう。これを解決するため、本インタフェースではスキーマ及びインスタンス全体から必要な部分のみを抜き出し、これを視覚化することを考えた。すべてのデータの中から利用者が必要とするデータのみを指定して、これを視覚化することができれば、大量データに対してもデータベースのブラウジングを効果的に行うことができる。ここでは、そのために不可欠である表示データを制限する手法について、スキーマとインスタンスそれぞれの場合について考察する。

3.2.1 表示するスキーマの制限

利用者がデータベースをブラウジングする際、特に大規模で複雑なスキーマであればこそ、スキーマの注目している一部分のみを抜き出してブラウジングしたい、という要求が考えられる。例えば、映像データベースに対しては「フレーム画像とそこに写っている被写体の関係だけを見たい」や「シーンとシーンを構成するフレーム画像の関係だけを見たい」などが考えられる。このような要求を満たすためには、スキーマから必要とする部分だけを選択して表示する機構が必要となる。スキーマに対する表示制限としては以下の種類が考えられる。

- entity-type の表示制限
 - relationship (entity-type 間の関係) の表示制限
- entity-type の表示制限としては、利用者が着目している entity-type のみを画面に表示して、その他の entity-type の表示を制限することが考えられる。これにより、必要な entity-type のみを画面に表示することが可能となる。

relationship の表示制限としては、まず relationship 自体が entity-type と entity-type のつながりを表現する概念なので、entity-type が表示制限されたとき、これにつながる relationship は意味をなさないの自動的表示を制限されるべきである。しかし、このような表示制限を行うと、entity-type 間の relationship

がほとんど表示されなくなってしまうケースが考えられる。極端な例では、直接 relationship でつながっていない二つの entity-type のみを表示した際には、relationship は一つも表示されないことになる。だが、スキーマにおいて完全に無関係な entity-type 同士というものは極めてまれで、他の entity-type や relationship を通じて何らかの間接的な関係をもつことが多い。entity-type 若しくはそれにつながる relationship を表示制限することは、このような間接的なつながりが見えなくなってしまうという問題が生じる危険性がある。

これを解決するためには、この間接的な relationship を必要に応じて仮想的に生成し、表示することが考えられる。仮想的な relationship を生成する手順としては以下のものが考えられる。

(1) まず relationship で結ぶ二つの entity-type を指定する。

(2) 次に、この仮想的な relationship がどの道筋(パス)によって、つながりをたどるのかを指定する。

例えば、映像データベースとして、SCENE (シーン)、FRAME (フレーム)、SUBJ (被写体) の三つの entity-type から成り立ち、SCENE と FRAME、FRAME と SUBJ 間に relationship が存在するスキーマがあるとき、SCENE と SUBJ 間に仮想的な relationship を生成する際には、まず仮想的な relationship で結び付ける entity-type として SCENE と SUBJ を指定し、relationship のたどる道筋(パス)を SCENE → FRAME → SUBJ と指定することにより、SCENE から FRAME を通って SUBJ に通じるのと等価な仮想的 relationship が、SCENE と SUBJ 間に生成される。

このようにして仮想的な relationship を生成することにより、必要な entity-type のみを表示し、それらの間があたかも直接に relationship で結ばれているように見せることができる。これによって、スキーマから着目している entity-type 及び relationship だけを抜き出して、表示することが可能になる。

3.2.2 表示するインスタンスの制限

スキーマに対して必要な部分だけを抜き出して視覚化できるのと同様に、インスタンスに対しても利用者が着目する部分のみを抜き出して視覚化できれば、ブラウジングを行う上で更に有効であるだろう。例えば、映像データベースに対して、すべてのフレームの中から、ある被写体が写っているフレームのみを抜き出し

て視覚化する等である。このために必要とされるインスタンスに対する表示制限としては以下の種類が考えられる。

- entity の表示制限
- (entity 間の) つながりの表示制限

entity に対する表示制限としては各々の entity-type 内で特定の entity のみを選択し、表示することが考えられる。特定の entity を指定する方法は以下の種類が考えられる。

- (1) 手動 (マウスのドラッグ操作等) で選択する
- (2) entity が値をもつとき、値に対する条件式で選択する
- (3) 指定した entity-type と指定した道筋 (パス) を通ってつながりのあるものだけを選択する

特に(3)の選択方法は特殊で、既に(1), (2)によって選択を行った entity-type に対して、これに属する entity から直接つながっている他の entity-type の entity だけを残すというものである。これを用いれば、ある特定の entity からつながっている一塊の entity の集合をインスタンス全体から取り出すことができる。例えば、あるシーンを表す entity からつながるフレーム、被写体などすべてのデータを取り出す (そのシーンに関するデータのみすべて抜き出す) などである。

また、つながりの表示制限としては、つながり自体が entity と entity を結び付ける概念なので、entity が表示制限されたとき、これに結びつくつながりは自動的に表示が制限される。ただし、スキーマにおいて前述の仮想的な relationship が生成された場合は、この仮想的な relationship に相当するつながりが新たに表示される。すなわち、スキーマ上で仮想的な relationship が生成されると、それによってインスタンス上でも、各々の entity-type に属する entity から指定された道筋で間接的につながりをたどった「仮想的なつながり」が表示される。

図4は実際にこのような表示制限を行って、部分的に取り出したデータのみを視覚化した例である。これらの表示制限を組み合わせて必要なデータだけをデータベースから取り出すことにより、大規模かつ巨大なデータベースであっても、必要なデータのみを視覚化して閲覧することが可能になる。

3.3 マルチメディアデータの表示・再生

マルチメディアデータ、とりわけ画像、映像、グラフィックスといった視覚的データであればこそ、視覚的インタフェース上でデータの表示・再生を行いたい

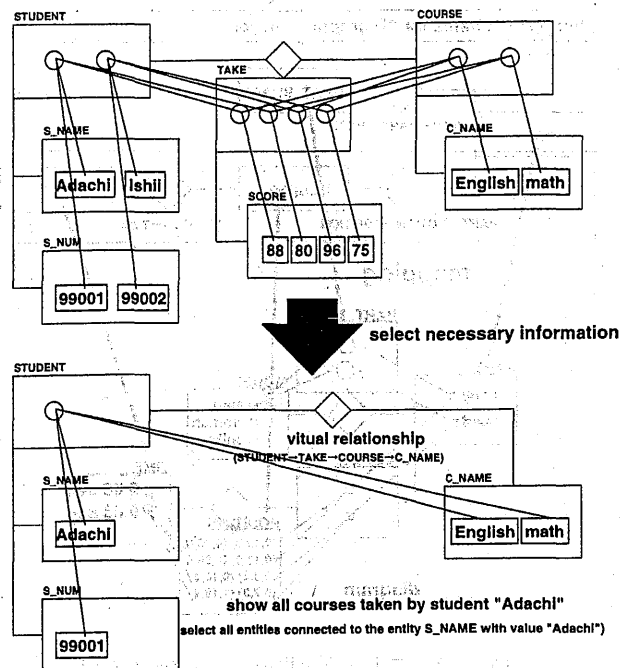


図4 表示制限を用いたブラウジング
Fig. 4 Browsing by restriction.

という要求は強い。例えば、映像を構成するフレームがデータベースに蓄積されている場合、これを時間軸に沿って映像として再生したい、またはデータベースにグラフィックスデータとして物体の3次元形状とその動き (姿勢の時系列) のデータがあるときに、動き付きのコンピュータグラフィックスとして再現したい、などといった要求が考えられる。このような表示・再生を行うためには、再生を行うプレーヤー (映像プレーヤー、グラフィックスプレーヤー等) に対して、プレーヤー側で必要とするデータをデータベースから取り出して受け渡す必要がある。このような機構として、ここではデータベースから再生に必要なデータを取り出し、テンプレートを用いてプレーヤーにマッピングする手法を考えた。テンプレートは、プレーヤーごとに決まった形をもつ一種のスキーマであり、ダイアグラムによって記述される (ここでは、テンプレートを構成するスキーマをテンプレートスキーマ、データベースのスキーマをデータベーススキーマと呼び区別する)。テンプレートスキーマを構成する entity-type は再生に必要なパラメータに対応している。例えば、映像プレーヤーの場合、テンプレートスキーマの entity-type としては、フレーム列、再生レートなどが、グラフィックスの場合、パーツの形状、位置 (x, y, z 座標と向き) とその時点などを表す entity-type が存在する。プ

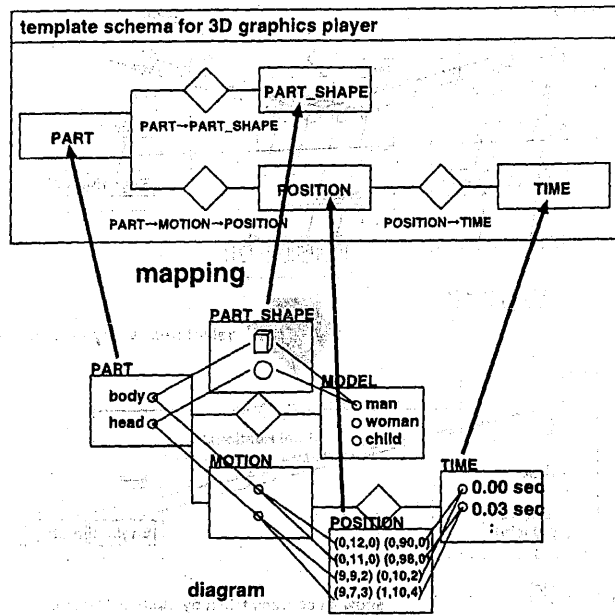


図5 テンプレートを用いたパラメータ指定
Fig.5 Specifying parameter with template.

レーヤを用いてマルチメディアデータを再生するためには、このテンプレートスキーマに対して、データベーススキーマの対応する entity-type をマッピングする。マッピングは、まずテンプレートスキーマ上の entity-type とデータベーススキーマ上の entity-type とを相互に対応づけ、次に同様にして relationship を対応づけることによって行う。対応する relationship がデータベーススキーマ上に存在する場合は問題ないが、存在しない場合は仮想的な relationship を新たに生成して対応づけを行う。

図5は、3Dグラフィックスプレーヤーテンプレートに対するマッピングの例を表している。実際に再生される際には、マッピングされたデータベーススキーマ上の entity-type に属する entity の値が、再生パラメータ値としてプレーヤーに引き渡される。プレーヤーへのパラメータ値の引き渡しは現在データベーススキーマのダイアグラム上に表示されている entity に対してのみ行われるので、前述の表示制限を用いて、再生したい entity のみを抜き出して画面に表示しておくことによって、自由に再生するグラフィックスを変化させることができる。このようにしてテンプレートを用いたマッピングにより、マルチメディアデータの表示・再生が可能になる。

4. 視覚的インタフェース AVIS

これまで、データベースに対する視覚的かつ直観

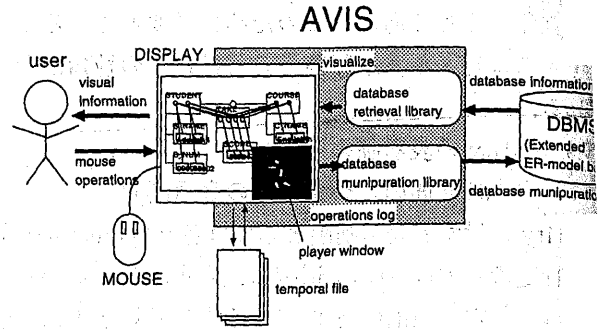


図6 AVIS システム構成図
Fig.6 Architecture of AVIS system.

的操作の実現を目的とし、必要な視覚的インタフェースの機能について考察してきた。これを実現するのとして、我々が開発した視覚的データベースインタフェースが AVIS (Advanced Visualizer of Instance and Schema) である。AVIS のシステム構成を図に示す。

Web ブラウザからネットワークを通じてデータベースに対する操作を行えるように、AVIS は JAVA (JAVA アプレット) で実装されている。開発は Sun ワークステーション Ultra sparc 1 (solaris 2.5.1) 上で、JDK1.1.5 を用いて行った。また DBMS としては、並列計算機日立 SR2201 (HI-UX/MPP, 32 プロセッサ) 上で独自に実装した拡張 ER ダイアグラムに基づく DBMS [17], [18] を用いている。詳細についてはここでは省略するが、これにより高速なデータベースアクセスを可能にしている。

図7は、AVIS によってスキーマ及びインスタンスを視覚的に表示した例である。AVIS ではスキーマ及びインスタンスを拡張 ER ダイアグラムを用いて視覚化するが、実装に際しては、更に有効なブラウジングを行うため、以下の機能が追加されている。

4.1 自由なレイアウト変更

もともと AVIS では、データベース設計者がデータベーススキーマを設計した際に定めた初期レイアウトに従って entity-type の画面配置を行う。初期レイアウトは、そのデータベーススキーマを最もわかりやすく、かつスキーマの意図を反映するように設計者が指定したもので、このような初期レイアウトでスキーマ及びインスタンスが表示されることにより、利用者はデータベース設計者の設計思想を感じ取りながらデータベースの内容を直観的に把握することができる。

しかし、実際に AVIS を利用し、データベースから必

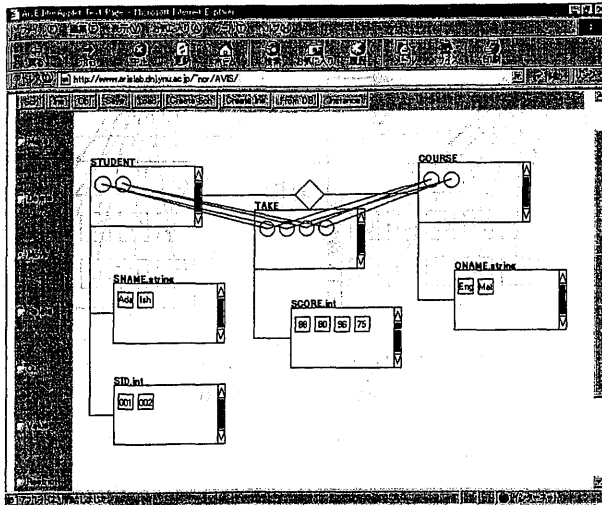


図7 スキーマ・インスタンス表示例
Fig. 7 An example of displaying schema and instance.

要なデータだけを取り出してきて視覚化する際、あらかじめ定められたレイアウトでしか表示できないのでは、利用者の自由なブラウジングを実現することができない。そこでAVISでは、必要に応じてentity-typeの画面配置を自由に変更することができる。

このとき、entity-typeを移動すると、これにつながるrelationshipも自動的に移動するようになっている。これは、たとえentity-typeの画面上での物理配置を移動したとしても、entity-type間の論理配置(relationshipによる結び付き)は変更されないようにするために、entity-typeの位置を変更しても、スキーマとしての論理構造は決して変更されない。

4.2 レイアウト作成機能

AVISでは、必要に応じて画面レイアウトを自由に変更できるが、その際、次のような自動または半自動レイアウト作成機能を用いることができる。

[単値成分及び多値成分(特殊なrelationship)で結ばれたentity-typeの半自動整列]

例えば、学生を表すentity-typeであるSTUDENTに対し、学生の名前を表すS_NAME、または学生の番号を表すS_NUMなどのentity-typeのように、属性としての意味合いをもち、特定のentity-typeに從属する性質をもつentity-typeが存在する。これらのentity-typeは一般に、從属関係の親となるentity-type(この場合はSTUDENT)に対して、複数の子となるentity-type(S_NAMEやS_NUM)が存在しており、特定の親entity-typeに從属している子entity-type群

を整列して表示したいという要求が考えられる。このような配置を次のように半自動で行うことができる。

まず、画面上でレイアウトを行う領域を指定する。この領域内でentity-typeの自動配置を行うので、配置したいentity-type群が十分に入る大きさであることが必要である。次に親entity-typeの配置位置を手動で指定する。領域内での親entity-typeの配置位置によって、子entity-type群の配置位置が変わってくる。最後に配置したい子entity-type群を指定する。これによって、自動的にentity-typeの配置がなされるが、そのアルゴリズムは次のようなものである。

レイアウトを行う指定された領域を縦と横にそれぞれ2分割し、計四つのエリアに分割して考えたとき、親entity-typeが存在しているエリア(複数のエリアにまたがって存在している場合は、最も多く占有しているエリア)をエリアAとする。またエリアAの対角をエリアB、エリアAの横に接するエリアをC、縦に接するエリアをDとする。このとき、まずエリアB(対角エリア)に対してentity-typeの自動配置を行う。各entity-typeはエリアBの横中央に縦一列で等間隔に並べて配置される。間隔は通常entity-typeの縦幅*1.5が標準値であるが、場合によっては変更することができる。子entity-typeを縦一列に並べることによって、親entity-typeに対して從属しているentity-type群であることを表す。エリアBに配置できなかった余りは、同様にしてエリアCに並べて配置される。以下、エリアD、エリアAと配置可能な分だけ配置していく。

以上のアルゴリズムで、從属関係(特殊なrelationship)で結ばれたentity-typeを半自動的に配置する。[relationshipの自動配置]

学生を表すentity-typeであるSTUDENT、科目を表すentity-typeであるCOURSEに対して、これらのつながりを表すTAKEなどのようにentity-type間を結び付けるものとしてrelationshipが存在する。relationshipの配置は次のように自動化できる。

まず、relationshipで結び付ける二つのentity-typeをAとBとし、これらは既に画面上の座標 (A_x, A_y) 及び (B_x, B_y) に配置されているとする。このとき、

- $|A_x - B_x| \geq |A_y - B_y|$ (横長)ならば、relationshipを表すひし形を $((A_x + B_x)/2, A_y)$ に配置し、 $((A_x + B_x)/2, A_y)$ から (A_x, A_y) まで、 $((A_x + B_x)/2, A_y)$ から (B_x, A_y) まで、 (B_x, A_y) から (B_x, B_y) までを結ぶ線を描く。

- $|A_x - B_x| < |A_y - B_y|$ (縦長) ならば, relationship を表すひし形を $(A_x, (A_y + B_y)/2)$ に配置し, $(A_x, (A_y + B_y)/2)$ から (A_x, A_y) まで, $(A_x, (A_y + B_y)/2)$ から (A_x, B_y) まで, (A_x, B_y) から (B_x, B_y) までを結ぶ線を描く。

といったように relationship を自動的に配置する。また, relationship を表すひし形に隣接して表示される entity-type は, $|A_x - B_x| \geq |A_y - B_y|$ (横長) ならばひし形の上か下, $|A_x - B_x| < |A_y - B_y|$ (縦長) ならばひし形の右か左に隣接して配置されるが, デフォルトは下または左で, 他の entity-type と重ならない方向が自動的に選択されて配置される。

このような機能により, 自動または半自動でレイアウト作成をサポートしている。

4.3 entity-type のサイズ変更

entity-type を表す方形は, デフォルトでは内部に entity が横に 5 個, 縦に 2 段並べられる大きさで表示されるが, この entity-type のサイズは必要に応じて変更することができる。entity-type のサイズを変更することにより, 内部に表示される entity の個数が自動的に変化する。entity-type を表す方形を, 横に広げることにより横に表示される個数が, 縦に広げることにより縦に表示される段数を増やすことができ, entity-type ごとに利用者の要求に合わせて最適なサイズを指定することができる。

4.4 スクロールバー

インスタンスを視覚的に画面に表示する際, 大量のインスタンスが存在する場合には, 画面内にすべてのインスタンスを表示することができない。これを解決するため, AVIS ではスクロールバーを用いて画面に表示するインスタンスを切り換えるようにしている。スクロールバーは各 entity-type ごとに存在し, これをスクロールすることによって, entity-type 内に表示される entity が順に切り換わっていく。

スクロールバーは entity-type の縦方向にのみ存在し, 横に並んだ entity を段ごと上下にスクロールする。例えば, entity-type 内に entity が横 N 個 \times 縦 M 段に表示されているとき, 下にスクロールすると, 一番上段の N 個が消え, その分だけ全体として上にずれ, すきまの空いた一番下段に新しく N 個の entity が表示される。entity の個数が多い場合, このようなスクロールバーを用いて表示する entity の切り換えを行う。

このようなスクロールバーによる表示切り換え, または

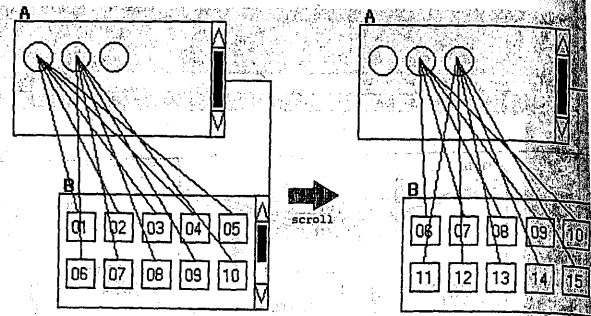


図8 スクロールバー
Fig.8 Scroll bar.

entity の表示制限を行った場合, entity 間のつながりは, 必然的に画面に表示されている entity 同士の間しか表示されない (図8参照)。

4.5 対応が明白なつながり表示の省略

また AVIS では, スキーマとの対応関係を視覚的表現するため, スキーマ図に重ねてインスタンス図表示される。インスタンス図では, 原則として entity 間のつながりを線分によって示しているが, 図7右の entity-type COURSE と CNAME のように, 1つの entity-type に属する entity 同士が, 先頭から順番に必ず 1対1に対応する場合 (単値成分の場合) に限って, 線で結ばなくとも対応が明白であるので entity 間のつながり表示の描画を省略している。このような場合, 前述のスクロールバーを用いて, 一方の entity-type をスクロールすると, 他方の entity-type も自動的にスクロールされる。

4.6 entity の並べ換え

entity-type に属する entity は, entity-type を表す方形内で, 左から右, 上から下という順序をもって表示される (entity が横 N 個, 縦 M 段で表示される場合, まず最上段の左から $1 \sim N$, 次に2段目の左から $(N+1) \sim 2N, \dots$, 最後に M 段目の左から $((M-1) * N + 1) \sim M * N$ という順序で表示される)。初期配置では, entity はデータベースから取り出された順序で表示されるが, 必要に応じて次の中から表示順序を選択することができる。

- entity のもつ値の昇順または降順
数値, 文字など値の大小によって表示する entity の順序を並べ換える。

- 他の entity-type とつながっている順
指定した entity-type の entity との間でつながりが交錯しないように entity の順序を並べ換える。

- データベースから取り出された順序に戻す

初期配置のデータベースから取り出された順序に entity を並べ直す。

並べ換えは、画面に表示されていない entity に対しても行われ、スクロールした際、新たに出てくる entity もこの順序で表示される。

このような機能を用いて視覚化されたデータベーススキーマ及びインスタンスに対して、AVIS では簡単なマウス操作を用いて表示するデータに制限を加えることにより、必要なデータのみを選択して視覚化することができる。各種の操作は AVIS の画面左にあるコマンドアイコンを選択することにより、簡単なマウス操作を用いて行う。データベースのスキーマ及びインスタンスから必要な部分のみを抜き出して視覚化し、これに対してブラウジング、または表示・再生といった操作を行うことによって、データベースの中身を直観的に把握することができる。

5. AVIS の表示例

ここまで、データベーススキーマ及びインスタンスを概念構造そのままに視覚化し、視覚的なデータベース操作を実現するインタフェースとして AVIS の設計

及び実装手法について述べてきたが、実際のデータベースに対して AVIS がどの程度効果的なのかを示すため、コンピュータグラフィックスを用いた人間動作データベースを AVIS で視覚化した例を示す。

図 9 はある人間動作データベースのスキーマ及びインスタンスを表示した例である。AVIS ではこのようにデータベースのスキーマ及びインスタンス情報を視覚化し、利用者に提示する。また図 10 は AVIS の機能を用いて必要なデータのみを抜き出し、これをコンピュータグラフィックスで再生した例である。もとのデータベースには entity-type 「MODEL」に人間のモデルとして3種類のもが登録されているが、これよりモデル「man」を選択し、これに動きを与えて再生している。図 11 は人体を構成するパーツ「PART」のうち、下半身パーツ (waist, foot, leg, toe など)のみを選択し、これに動きを与えて再生した例である。また図 12 は「MODEL」として「child」を選択し、これに図 10 と同じ動きを与えて再生した例である。

このように、再生したいデータ (モデルパーツなど) をダイアグラム上で表示制限を用いて選択することによって、その結果を反映して様々なグラフィックスを自由に再生することができる。人間動作データベース

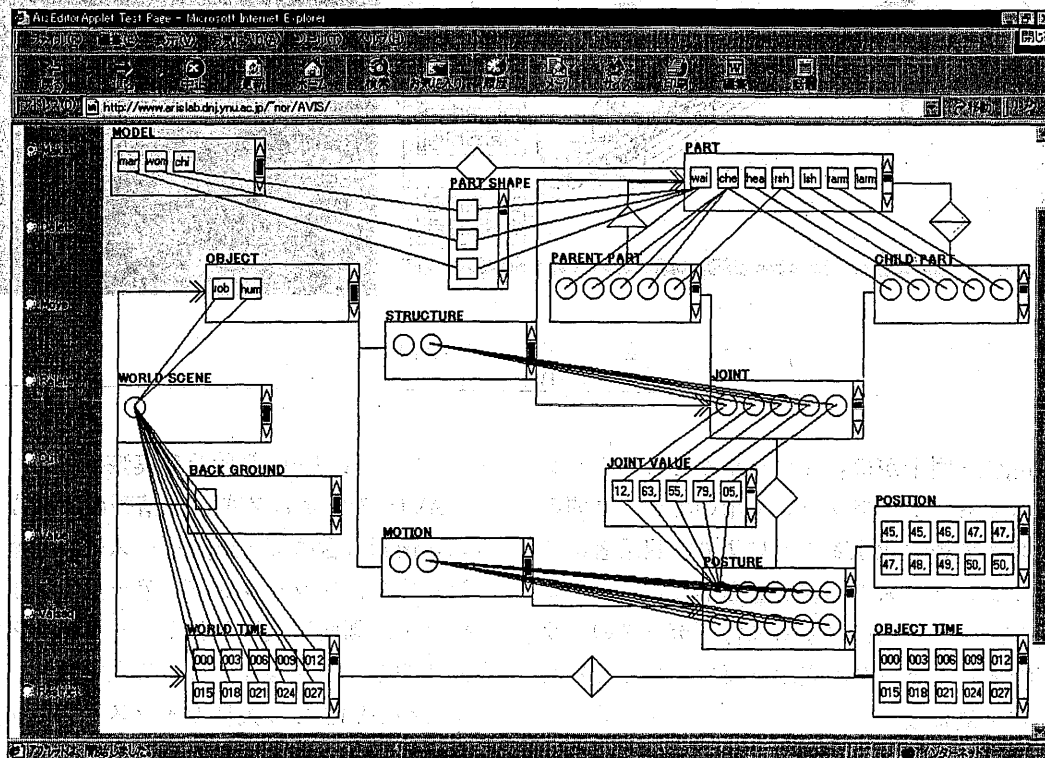


図 9 データベースの視覚化例

Fig. 9 An example of database visualization.

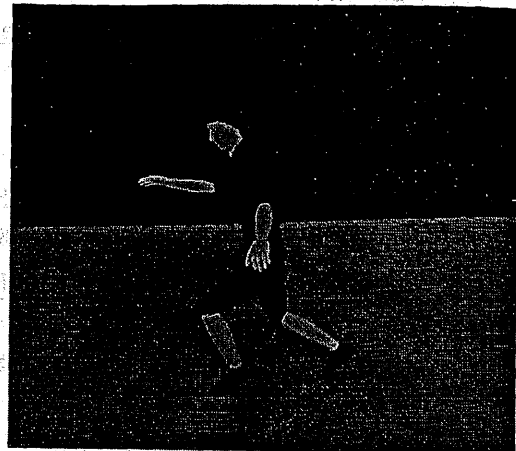
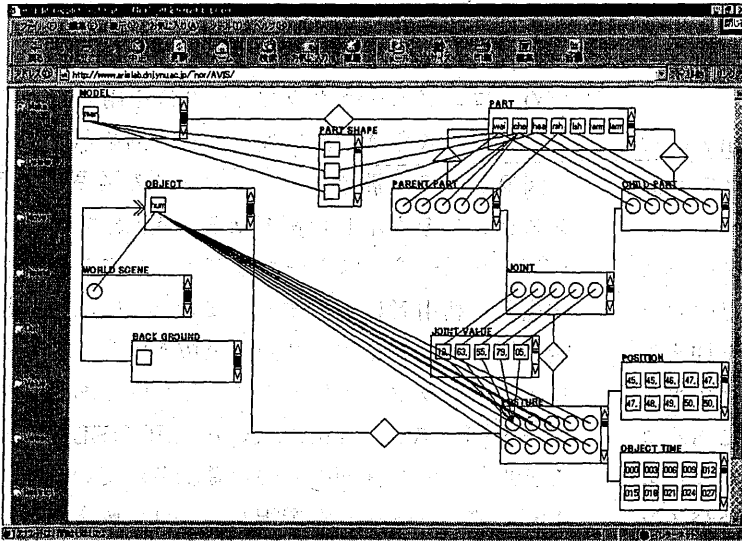


図 10 ブラウジング結果 1
Fig. 10 Browsing result 1.

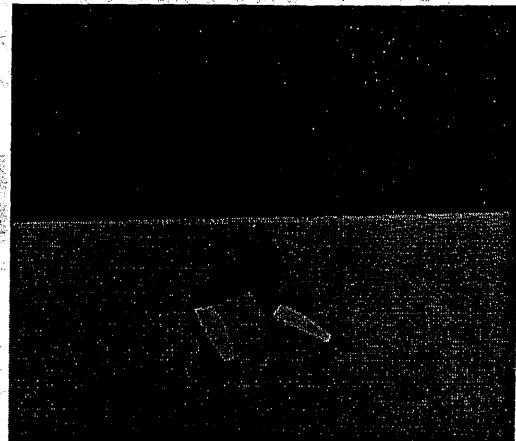
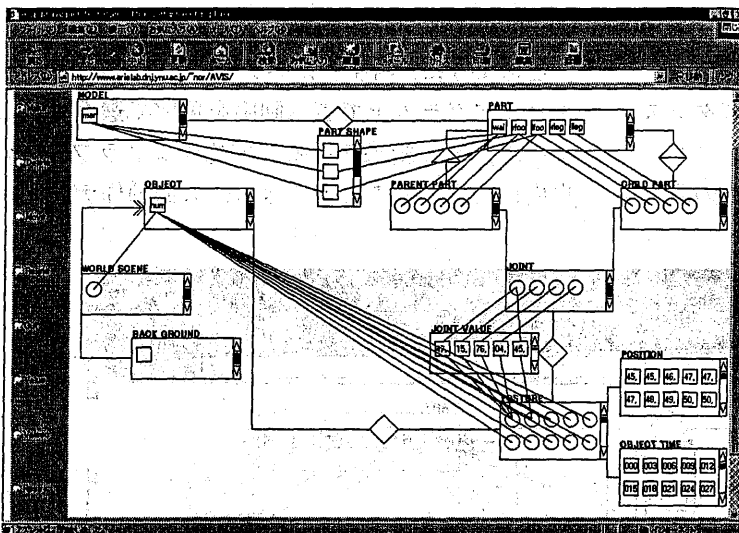


図 11 ブラウジング結果 2
Fig. 11 Browsing result 2.

のように、entity 同士が複雑なつながりをもつようなデータベースでも、ダイアグラムを用いてデータ間のつながりを含めて視覚化することにより、利用者はスキーマ及びインスタンスの情報を直観的にとらえることができ、データベースに対する操作（データのブラウジング、指定したデータの表示・再生）などを容易に行うことが可能になる。

6. む す び

本論文では、データベースを視覚化するインターフェー

スとして AVIS の設計及び実装手法について述べた。

AVIS の特徴はスキーマ情報のみでなく、インスタンス情報も含めてデータベースの概念構造そのままに視覚化し、ブラウジング、データの表示・再生などの操作を、画面に表示されたデータベース情報に対する直観的なマウス操作で行える点である。これを実現するため AVIS では大量のインスタンス情報を利用者の要求に応じて自由に取り出し、必要なデータのみを画面に表示する機構をもつ。これにより AVIS は大規模かつ複雑なデータベースに対しても、利用者のデータ

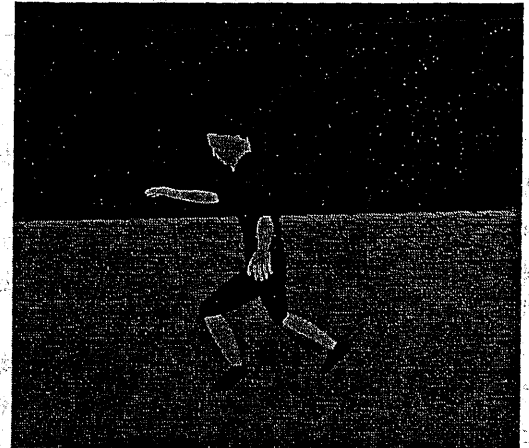
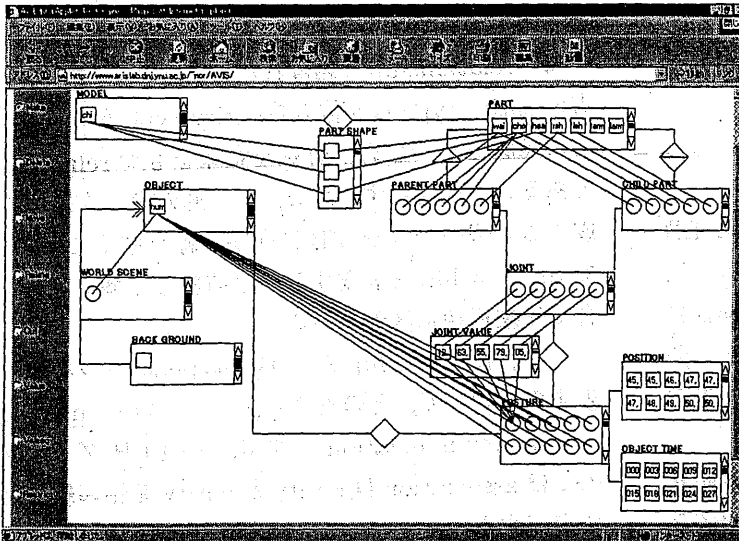


図 12 ブラウジング結果 3
Fig.12 Browsing result 3.

ベース情報の直観的把握と操作を実現している。

今回は AVIS を用いたデータブラウジングのみに焦点を絞って述べてきたが、AVIS 自体はデータベースの概念構造をダイアグラムを用いて視覚化するツールであり、これを用いてデータベースに対するデータの変更や検索を行う技術も研究中である。更に使いやすいデータベース視覚化インタフェースを実現するため、今後も研究を続けていくつもりである。

文 献

- [1] S.K. Card, "Visualizing retrieved information: A survey," IEEE Comput. Graphics & Appl., vol.16, no.2, pp.63-67, 1996.
- [2] 磯部成二, 黒川 清, 塩原寿子, 飯塚哲也, "視覚化による多次元データ分析システム: INFORVISER," 情処学論, vol.40, no.5, 1999.
- [3] 武田浩一, 野美山浩, "テキスト情報の可視化を利用した情報検索," 情報処理, vol.41, no.4, pp.343-350, 2000.
- [4] 古畑理香, 藤代一成, 市川哲彦, 竹島由里子, "GADGET/IV: 情報可視化の半自動設計支援環境," 第 11 回データ工学ワークショップ (DEWS 2000), 2A-3, 2000.
- [5] 尾下真樹, 牧之内顕文, "オブジェクト指向データベースの半自動可視化環境," 信学技報, DE2000-5, 2000.
- [6] M. Derthick, J.A. Kolojchick, and S. Roth, "An interactive visual query environment for exploring data," Proc. ACM Symposium on User Interface Software and Technology (UIST), pp.189-198, Banff, Canada, 1997.
- [7] T. Kohonen, "The self-organizing maps," Proc. IEEE, vol.78, no.9, pp.1464-1480, 1990.
- [8] 佐野綾一, 波多野賢治, 田中克己, "自己組織化マップを用いた Web 文書の対話的分類とその視覚化," 情処学研報, vol.98, no.57, 98-DBS-116(1)-5, pp.33-40, 1998.
- [9] D.A. Keim and H.-P. Kriegel, "VisDB: Database exploration using multidimensional visualization," IEEE Comput. Graphics & Appl., vol.14, no.5, pp.40-49, 1994.
- [10] J. Rumbaugh, M. Blaha, W. Premerlani, F. Eddy, and W. Lorensen, Object-Oriented Modeling and Design, Prentice Hall, New York, 1991.
- [11] C.W. Bachman, "Data structure diagrams," DATABASE (ACM SIGBDP Newsletter), vol.1, no.2, pp.4-10, 1969.
- [12] P.P. Chen, "The entity-relationship model: Toward a unified view of data," ACM Trans. Database Systems, vol.1, no.1, pp.9-36, 1976.
- [13] A. Grau, J.K. Filipe, M. Kowsari, S. Eckstein, R. Pinger, and H.D. Ehrich, "The TROLL approach to conceptual modelling: Syntax, semantics and tools," Proc. International Conference on Conceptual Modelling, pp.277-290, 1998.
- [14] <http://www.platinum.com/products/appdev/erwin.ps.htm>
- [15] H. Arisawa, H. Nagae, and Y. Mochizuki, "Representation of complex objects in semantic data model "AIS" and implementation of set operators," IEICE Trans., vol.E74, no.1, pp.191-203, Jan. 1991.
- [16] H. Arisawa, T. Tomii, H. Yui, and H. Ishikawa, "Data model and architecture of multimedia database for engineering applications," IEICE Trans. Inf. & Syst., vol.E78-D, no.11, pp.1362-1368, Nov. 1995.
- [17] 家富誠敏, 富井尚志, 有澤 博, "関数型データベースシステムの試作および評価," 電子情報通信学会第 8 回データ工学ワークショップ論文集, 1997.
- [18] 河 信司, 市間健太郎, 高橋賢一郎, 家富誠敏, 有澤 博, "スクラップアンドビルド型マルチメディア DBMS の構

成方式,” 情処学研報, 98-DBS-116-62, 1998.

付 録

拡張 ER ダイアグラム (AIS ダイアグラム)

拡張 ER ダイアグラム (AIS ダイアグラム) は ER ダイアグラムを拡張したダイアグラム表記法である。拡張 ER ダイアグラムでは ER ダイアグラムと同様に entity-type (主体型 or 実体型) と relationship (関連) という二つの概念によって対象世界を記述 (モデル化) する。

entity-type は、同じ性質をもつ entity (現実世界の事物・事象) の集合を表す概念である。人や物体などといった形あるもの、科目や講義などといった抽象的な事柄などの現実世界をモデル化する際に注目する事物・事象はすべて entity として表現される。このとき同種の entity の集まりが entity-type となる。拡張 ER ダイアグラムでは entity-type は方形によって表記され、これに属する entity は entity-type を表す方形内に点 (AVIS では丸) として描かれる。entity がもつ値は entity を表す点に隣接して表記される (AVIS では値を四角で囲むことによって表示。このとき点 (or 丸) の表示は省略する)。

対して relationship は entity-type 間の意味的なつながりを表す概念である (正確には entity-type に属する entity 集合同士の直積の部分集合を表す)。entity-type と entity-type を relationship で結び付けることによって、これらの entity-type 間に何らかの関係があることが示される。例えば、学科と学生という二つの entity-type があつたときに、これらの間を「所属」という relationship で結ぶことによって、学生と学科の間には所属するという関係があることが表される。拡張 ER ダイアグラムでは relationship の性質によって、いくつかの表記を使い分けている。これが ER ダイアグラムに対する拡張 ER ダイアグラムの最も大きな拡張点である。通常 ER ダイアグラムでは relationship をひし形とひし形から (relationship で結び付けられる) entity-type へと伸びる線分によって表現するが、拡張 ER ダイアグラムでは relationship を以下のように表記する。

- relationship を、また一つの entity-type として再定義したことを明示するため、relationship を表すひし形の下に方形 (ボックス) を隣接して描く (ただし、改めて明示する必要がないときは、方形の表記

を省略することもある)

- relationship で結ばれた entity-type 同士が、一方の entity-type B が他方の entity-type A の属性としての意味をもつ場合、特殊な意味をもつ relationship であることを明示するため、ひし形の代わりに A を構成する 4 辺の一つから B に向かう線を描くことによって relationship を表現する (単値属性と多値属性の場合で先端形状に区別がある)

また、relationship が entity-type 間のつながりを表すのに対して、その実際の例、すなわち entity 間のつながりを association と呼ぶ。拡張 ER ダイアグラムでは association は entity と entity を結ぶ線分として表記する。

拡張 ER ダイアグラムではこのような表記法によって、現実世界をモデル化する枠組み (スキーマ) と実際の例 (インスタンス) を表現する。拡張 ER ダイアグラムの特徴は、relationship 表現を詳細にすることにより、より現実世界の意味を反映したモデル化を行うことができ、かつ属性としての意味づけを明示することによってカプセル化などの情報隠蔽技術を用いた視覚化が可能になることである。これらの特徴は、マルチメディアデータベースのように大規模・複雑なスキーマを表現する際、特に有効であると考えられる。

(平成 12 年 8 月 25 日受付, 12 月 21 日再受付)



家富 誠敏 (学生員)

平 8 横浜国大・工・電子情報卒。平 10 同大学院工学研究科博士課程前期了。現在同大学院工学研究科博士課程後期在学中。マルチメディアデータベースのアーキテクチャ及び視覚的インタフェースに興味をもつ。



森 慎之

平 11 横浜国大・工・電子情報卒。現在同大学院工学研究科博士課程前期在学中。マルチメディアデータベースの視覚的インタフェースに興味をもつ。



富井 尚志 (正員)

平6 横浜国大・工・電子情報卒。平8 同
大大学院工学研究科博士課程前期了。平11
同大学院工学研究科博士課程後期了。現在
同大学工学部助手。工博。主としてマルチ
メディアデータベース、質問処理などの研
究に従事。



有澤 博 (正員)

昭48 東大・理・物理卒。富士通(株)を
経て、昭50 横浜国立大学工学部に奉職。現
在同学部電子情報工学科教授。工博。1991
アメリカ・オレゴン州立大学計算機学科客
員教授。データベース理論、マルチメイ
アデータベースシステムを研究テーマとし

ている。

時空間MMDBにおけるサッカーの戦術記述と ad-hoc 検索

四十物裕司[†] 坂木和則^{††} 鬼束郷^{††} 富井尚志^{†††} 有澤博^{†††}

[†]横浜国立大学大学院 工学研究科 電子情報工学専攻

^{††}横浜国立大学大学院 環境情報学府

^{†††}横浜国立大学大学院 環境情報研究院

〒240-8501 横浜市保土ヶ谷区常盤台 79-7

E-mail: {anchan, sakaki, go, tommy, arisawa}@arislabs.dnj.ynu.ac.jp

あらまし

本稿では、実世界などの時空間における物体のふるまい(3次元的な位置変化)をデータベースに蓄積し、そこで生じている出来事(イベント)に基づく検索を行うデータベースの構築を目的とするが、特に、複数の人間が共存する時空間を対象として、局面(ある観点で見た時、同じ状況と判断される場面)を検索する手法を提案する。複数の人間が共存する時空間をデータベース化するためには、人の位置や動きといった点の座標の情報を蓄積するだけでなく、人間が移動できる範囲などといった空間的な領域の時間変化をデータベースに蓄積しておくことが自由度の高い検索を実現するためには必要であると考えられる。

そこで本稿では、時空間セルと呼ばれる、空間、領域と時区間との積を用いて領域の時間変化を記述する手法を提案し、サッカーの試合を具体例として取り上げ、時空間セル概念に基づく時空間モデリングと時空間検索手法について述べる。

Description of Soccer Tactics and Ad-hoc Query in Spatio-Temporal MMDB

Yuji Aimono Kazunori Sakaki Go Onitsuka Takashi Tomii Hiroshi Arisawa

Division of Electrical and Computer Engineering, Faculty of Engineering,
Yokohama National University

[†] Graduate School of Environment and Information Sciences, Yokohama National University

^{††} Faculty of Environment and Information Sciences, Yokohama National University

79-7, Tokiwadai, Hodogaya-ku, Yokohama 240-8501 JAPAN

E-mail: {anchan, sakaki, go, tommy, arisawa}@arislabs.dnj.ynu.ac.jp

Abstract

This paper proposes a new We aim at realizing the database which accumulates the behavior (motion of objects in 3-dimensional space) in spatiotemporal space, such as real world, to retrieve an event information. We propose the technique to retrieve the aspect (similar situation in a certain viewpoint) especially for spatiotemporal space where two or more human beings exist. In order to construct database for information of such spatiotemporal space, we should consider accumulating not only points coordinates information such as one's position but also region changes on progress of time is required to realize high flexibility query. Therefore, in this paper, we propose the technique to describe changing region, we call it "spatiotemporal-cell", expressed using the product of space region and time-interval. And taking up the soccer game as an example, we show spatiotemporal information modeling and retrieval technique based on concept of "spatiotemporal-cell".

1 はじめに

近年、マルチメディアデータベースを用いて、コンピュータグラフィックスやバーチャルリアリティなどを扱う研究が盛んに行なわれている。特に実世界などの時空間における物体の3次元的な振る舞いを蓄積し、多様な検索を行なうことができるデータベースを時空間MMDBと呼ぶが、このような時空間MMDBが実現されれば、スポーツにおける人体の動きの解析、医学におけるリハビリテーション計画、生産技術文やでの作業者と工作機械の協調作業の解析・調査など、幅広い応用分野での利用が考えられる。

ところで、一般に、時空間における事物・事象に対する人の解釈や検索したいと考える内容は多様であることから、時空間MMDBを実現するためには、検索者にとって自由度の高い検索であるad-hoc検索、特に時空間での出来事(イベント)を用いた検索に応えられる柔軟性、適応性が必要である。現在、モーションキャプチャシステムやGPS(Global Positioning System)、多視点映像からの画像解析などを用いれば、物体の3次元空間における位置データを自動的に取得することができるが、これによって得られた位置データをデータベースに蓄積するだけでは、これらの値をインデックスとした、非常に簡単な検索しか行うことができない。特に、対象とする時空間に複数の物体が存在するような場合では、物体相互の位置関係やそれらの協調動作を考慮した検索を行うことが必要になるだろう。

ここでは、時空間に複数の物体が共存する共存型時空間を対象として、多様な検索を行うデータベースシステムの実現を目的とするが、共有型時空間において、同じ局面(ある観点で見た時同じ状況と判断される)を探し出すという検索を考えれば、複数の物体によって構成される領域、各物体が他の物体の干渉を受けずに動くことができる領域などといった、空間における複数の物体の位置関係によって決定される領域の時間的変化をデータベースに蓄積し、検索に利用することが有用であると考えられる。

このような検索を行うために本論文では、空間領域の時間的変化を表すものとして、時空間セルという概念を提案する。時空間セルとは時間と空間の次元をあわせ持つセルによって、領域の時間変化を表現したものであり、これを用いて局面の特徴を記述

することによって、該当するパターンの局面を複数の物体の位置関係やその時間的推移を用いて検索することができる。

本稿では、時空間セルを用いたデータベースの構築とデータベース検索手法について述べる。またこのような検索手法が、サッカーの戦術に基づく空間処理で有効であることを示す。

以下、2章では複数の人間が存在する共存型時空間のデータベース化を実現するために必要とされる要素について述べ、3章では、本稿で提案する時空間セルの定義とモデル化について述べる。4章では時空間セル概念に基づく検索手法について、及び共存型時空間の例としてサッカーの試合を挙げて、5章でまとめを述べる。

2 共存型時空間の構成要素

我々が提案する、複数の物体(人間など)が共存する時空間(共存型時空間)を扱うデータベースシステムは、GPSや多視点映像からの画像解析によって得られる人間の3次元空間における位置データとそのデータから得られる解析データをスキーマに従って蓄積することでad-hocな時空間検索を実現しようというものである。このシステムのプロトタイプを図1に示す。

共存型時空間の情報をデータベース化するためには、まず対象とする世界がどのような時空間的情報で構成され、どのようなデータによってその情報が表現されるのかについて考える必要がある。

共存型時空間の構成要素として少なくとも次の要素が考えられる。

1. 共存型時空間全体の情報
データベース化する対象世界の時間的・空間的な情報
2. 共存型時空間に存在する物体(オブジェクト)の情報
人間やロボット、ボールや道具などの対象時空間に存在するものの動きに関する情報
3. 共存型時空間を構成するイベントの情報
走る、ドリルを持つ、DFラインを上げるといった対象時空間に存在するオブジェクトによって起こされる事象・出来事に関する情報

共存型時空間情報をデータベース化するためには、以上のような各種、各レベルの情報をすべてモデル化し、蓄積する必要がある。その中でも3のイベントは共存型時空間に対して、検索者が興味をもつ要素であり、特に複数の人間が協調して行うといった協調イベントに関する興味は大きい。

そこで、共存型時空間情報を有効に蓄積することができ、直観的、かつ自由に検索することができる共存型時空間データベースを構築するためには、協調イベントも検索できるような新しいモデル化手法と検索手法の確立が必要不可欠である。

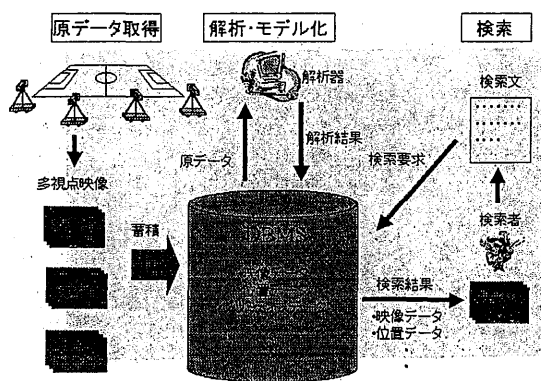


図 1: システムのプロトタイプ

3 時空間セルを用いた領域の表現

我々は共存型時空間をデータベース化の対象として、個別のイベントだけではなく、複数の人間で構成される高度なイベントに基づく ad-hoc 検索を実現することを最終目標としている。

その目標を実現する際に重要となる情報は位置に関する情報ではなく領域に関する情報である。というのは、位置に関する情報が蓄積されていた場合には、その位置や軌跡といった低レベルの情報に関する検索しかできないのに対して、領域は複数の物体で囲まれた領域というのを考えるとより高度なイベントに基づく検索が可能になるからである。

このような検索を考えた場合、単に物体の位置情報が入っているだけでは検索することができない。というのは、上のような検索を行なうためには選手

の位置の変化だけではなく、選手の周りの領域の変化も重要だからである。

そこで、我々は上記のような検索に応えるためには領域を考えることが有効であると考えた。この支配領域の時間変化を記述することができれば、単に選手の位置情報が蓄積されているだけではできないようなイベントに関する ad-hoc な検索が実現できると考えた。

そこで、この支配領域の時間変化を記述するために、時空間セルを提案する。時空間セルとは選手の支配領域と任意の時区間の積で表されるもので、この時空間セルをデータベースに蓄積しておくことで支配領域の時間変化を記述することができ、位置情報だけではできないような高度なイベントに基づく ad-hoc 検索に応えることが可能になると考えられる。

3.1 時空間セルの大きさ

3.1 節で述べた人間の持つ支配領域と 3.2 節で述べたアクションタイムとの積で、表現される要素を時空間セルと呼ぶ。この要素は、共存型時空間における様々な局面や協調イベントを検索する際のプリミティブであり、概念図を図 2 に示す。

この時空間セルはアトリビュートとして、始点と終点における支配領域とそのセルの体積 (STI: Space Time Integration) を持っており、ある時点 t における人間の支配領域の面積を $S(t)$ としたとき、時区間 $T(t_s, t_e)$ における STI は次式のように表される。

$$STI_A(T) = \int_{t_s}^{t_e} S(t) dt$$

ただし、本稿では、この時空間セルの価値を評価することが主題であることから、簡単のために、時区間の補間には線形補間を適応し、この補間法に関する評価は今後の課題としておく。

4 時空間セルを用いたサッカーデータベースの設計

本章では、共存型時空間の例としてサッカーを考え、様々な局面や協調イベントに関する検索を実現するため、時空間セルを用いたサッカーデータベースを設計し、その上での検索手法について述べる。

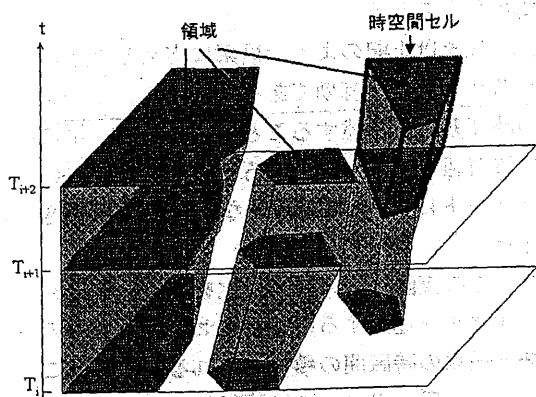


図 2: 時空間セル

4.1 時空間セルにおける空間分割

時空間セルは空間と時間の積で表現すると述べたが、そこでの空間の捉え方について、本節では詳しく述べる。複数の人間が存在する共存型時空間で人間は必ず、同じ時空間に存在する他の人間との兼ね合いで動いている。例えば、チームスポーツにおいて、各選手はチームの勝利のために相手選手や味方選手の動きに連動して動いている。また、電車やバスといった閉空間においても、他の人間の位置を見て、ある程度の距離をとって自分の位置を決定する。

こう考えると、共存型時空間において、人間はある種の支配領域を持っていると考えられる。例えば、サッカーにおいて、ある選手がボールを持てば相手選手はボール所持選手に自由にプレーする領域を与えないようにし、味方選手は相手の支配領域から離れてボールをもらおうとするが、これは置き換えれば、相手選手はボール所持選手の支配領域を減らそうとし、味方選手は相手選手の支配領域から離れて、自分の支配領域を増やそうとしていると考えることができる。また、電車内においても混んでいない時には、他人から離れて自分の位置を決めるが、これも自分の支配領域を増やそうとしていると捉えることができる。

このように、共存型時空間において、人間はある種の支配領域を持っており、その支配領域の変化がその時空間における様々な情報を反映していると考えられる。

そこで本稿では、その人間の支配領域の変化を捉

えるために、共存型時空間は対象時空間に存在する複数の人間の支配領域の総和であると捉え、時点における空間をそれぞれの人間の支配領域に分割する。

人間の支配領域とは誰よりも早く到達できる範囲のことで、これは各個人の空間的な位置関係、および、個人の運動能力によって決定される。一般に、数理的に空間中にいくつかの点が与えられたとき、各点の支配領域はボロノイ領域 [?] で表現できる。このボロノイ領域を拡張することで、集団で移動する個人の支配領域を定式化する手法が確立されている [1]。そこで本稿では、その手法を用いて人間の支配領域のモデル化を行い、支配領域の例を図 3 に示した。

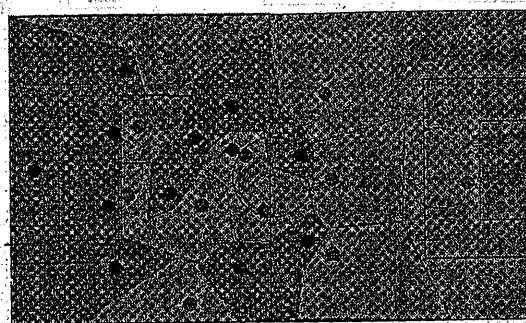


図 3: 支配領域

4.2 時空間セルにおける時間分割

本節では、本稿で提案する共存型時空間に関する多様な局面や協調イベントを直観的、かつ自由に検索するためのプリミティブである時空間セルを生成する際の時間の捉え方について述べる。

時空間セルは、様々なイベント検索をする際の最小要素である。従って、その時空間セルを生成する時区間はイベントに基づいて分割される最小時区間であり、これをアクションピリオドと呼ぶ。このアクションピリオドは、対象時空間における中心的なイベントに基づいて分割される。例えば、サッカーであれば、選手がボールに触るというイベントに基づいて分割し、ある選手がボールを触ってから他の選手が触るまでの時区間がアクションピリオドとなる。

というのは、例えば、サッカーであれば、ボールを持つ選手が変わるごとに周りの選手も連動して動きが変わる、というように多くの人間はその中心的

なイベントを変わり目として行動を変えることが多いからである。

さらに、対象時空間に存在する全ての人間が行動し直すといった大きく局面が変わる時区間をシーンと呼ぶ。つまり、シーンの区切りが対象時空間の全ての人間のイベントの変わり目となるということで、シーンはアクションピリオドよりも大きなイベントの変化に基づいて分割される時区間となる。

そこで、共存型時空間を複数の時区間に分割する際に、次の2つのレベルの時区間を定義する。これら2つの時区間は対象とする共存型時空間によってイベントの種類が異なるので、対象によって変わるものである。

- アクションピリオド … シーンにおいて、局面が変わる時点で排他的に区切られた時区間。イベントに基づく最小時区間
- シーン … 対象時空間において、大きく局面が変わる時点で排他的に区切られた時区間

4.3 サッカーデータベースのスキーマ設計

4.3.1 人間共存型空間データベースのスキーマ

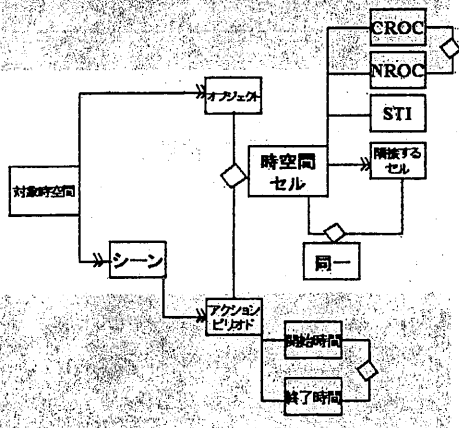


図 4: 人間共存型空間データベースのスキーマ

図 4に人間共存型時空間のスキーマを示す。実際に、人間共存型時空間に属する時空間のスキーマを設計する場合には、これを基本とし、必要があればエンティティや属性を追加してスキーマを設計すれば良い。

4.3.2 サッカーデータベースのスキーマ

ここでは、具体例としてサッカーの試合のスキーマ設計を試みている。図 5

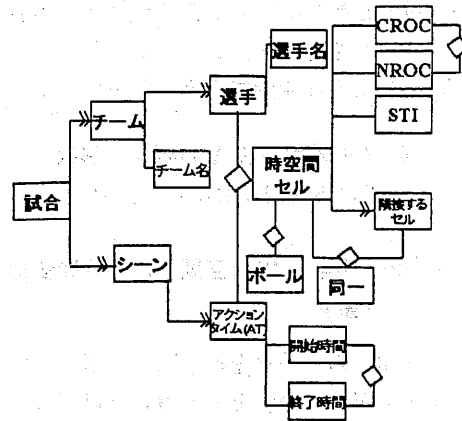


図 5: サッカーの試合のスキーマ

このスキーマで表される情報は次のようになる。

- 対象時空間に関する情報として、シーンやオブジェクトに関する情報が蓄積される。
- シーンに関する情報として、アクションピリオドが蓄積され、さらにアクションピリオドに関する情報として、開始時点と終了時点の情報が蓄積される。
- 時空間セルは質問とアクションの直積として表される時空間セル情報としては、現在の支配領域 (CROC: CURRENT REGION OF CONTROL) や次の時点における支配領域 (NROC: NEXT REGION OF CONTROL), その時空間セルの体積 (STI: SPACE TIME INTEGRATION) と隣接する時空間セルの情報も含めてエンティティとして蓄積される。

ここでのポイントはあらかじめ時空間セルの STI をデータベースに蓄積しておくところである。というのは、STI は支配領域の時間方向への積分演算の結果であり、その値をあらかじめ蓄積しておくことで検索の負担を減らしている。

4.4 サッカーデータベースの検索例

本節では、

設計したサッカー試合のデータベースを用いたイベント表示の例を示す。

4.4.1 検索例 1 (時空間セルを用いた検索)

設計したスキーマに基づき、「ある選手 A がフリーでボールをもらったシーンが見たい」という検索は以下のような手順で行うことができる。

- タイプ選手名 A の要素を選択し、それと続けている時空間セルの CROC が α 以上のシーン

4.4.2 検索例 2(時空間セルの時間変化を用いた検索)

「相手のマークを振り切ってボールを受けとったシーンが見たい」という検索は以下のような手順で行うことができる。

- ボールをもった選手の時空間セルの要素である NROC と CROC の差分を取った結果、正であれば該当する。

4.4.3 検索例 3(複数の時空間セルを用いた検索)

「ボールを持っている選手の相手のマークが集まったところでマークがうすくなった選手にパスが渡ったシーンが見たい」という検索は以下のような手順で行なわれている。

- ある選手がボールを持っている間の時空間セルの要素である NROC と CROC の差分を取った時に負の値を返して、かつ味方の選手の時空間セルの要素 NROC と CROC の差分を取った時に α 以上の値を返したときに味方にボールが渡ったシーンを抜き出す。

この検索結果として取り出されたシーンを実際にグラフィックスで表示した例が図 6,7,8 である。この図を見ると支配領域は片方の選手では徐々に広がっていて、片方の選手では徐々に狭くなっていることが分かる。

5 まとめ

本稿では、複数の人間が存在する共存型時空間を対象としたデータベースシステムの設計を行なった。

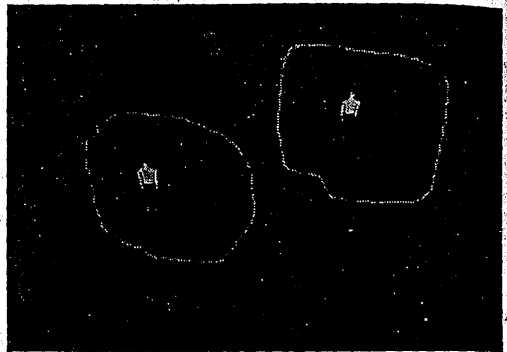


図 6: 結果 1

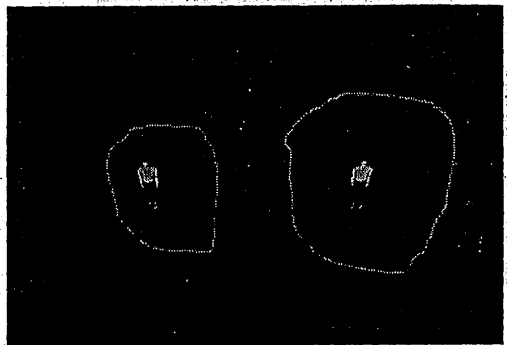


図 7: 結果 2

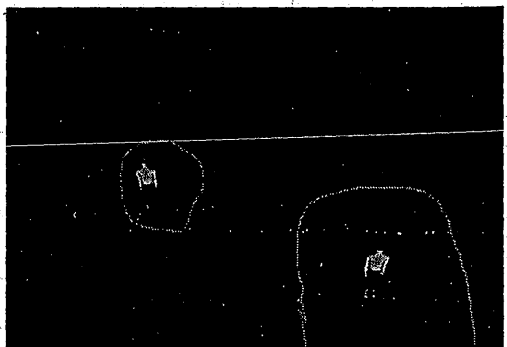


図 8: 結果 3

また、対象時空間における様々な局面を表現するための基本要素として時空間セルを提案し、時空間セルをデータベースに蓄積しておくことによって、多様な局面を直観的、かつ自由に検索するための手法を提案した。具体例として、サッカーの試合を挙げ、時空間セルの表現能力の検証を行った。この結果により、時空間セルを用いることによって、位置データを蓄積しておくだけでは検索することが困難であるような局面や協調イベントを直観的に検索できることを示した。

謝辞

本研究の一部は、文部科学省科学研究費補助金奨励研究(A)(課題番号 12780202)及び、文部科学省科学研究費補助金基盤研究(C)(2)(課題番号 12680338)による。

参考文献

- [1] 瀧 剛志, 長谷川純一: “チームスポーツにおける集団行動解析のための特徴量とその応用”, 電子情報学会論文誌 (D-II) Vol. J81-D-II, No.8 pp.1802-1811, 1998
- [2] H. Arisawa, H. Nagae, Y. Mochizuki, “Representation of Complex Objects in Semantic Data Model “AIS” and Implementation of Set Operators”, IEICE TRANSACTIONS, vol.E74, No.1, 1991.
- [3] H. Arisawa, T. Tomii, H. Yui, and H. Ishikawa: “Data Model and Architecture of Multimedia Database for Engineering Applications”, IEICE TRANS. INF. & SYST., vol E78-D, No.11, 1995.
- [4] Kiril Sarev, S. Imai, T. Tomii, H. Arisawa, Y. Tsuchitani: “Modeling and Detection of Events in Factory Work Simulation System”, Intelligent Manufacturing System September 22-24 pp.641-649 1999.
- [5] 富井尚志, 有澤博: “マルチメディアデータベースにおける映像モデリングと操作言語,” 信学論. Vol.J79-D-II. NO4. pp.520-530(1996).
- [6] Sunghoon Choi, Yongduck Sea, Hyunwoo Kim, Ki-Sang Hong: “Where are the ball and players?: Soccer Game Analysis with Color-based Tracking and Image Mosaick” Dept. of EE, Pohang University, of Science, and Technology Sam 31 Hyoja Done, Pohang, 790-784, Republic of Korea
- [7] Taeone Kim, Yongduck Sea, Ki-Sang Hong: “Physics-based 3D Position Analysis of a Soccer Ball from Monocular Image Sequences” Dept. of EE, POSTECH San 31 Hyoja Dong, Pohang, 790-784, Republic of Korea

XMLを用いたマルチメディア構造化文書STOICの意味表現 能力とその評価

越前谷健二 † 松田一章 †† 富井尚志 ††† 有澤博 †††

† 横浜国立大学院 環境情報学府

†† 横浜国立大学大学院工学研究科

††† 横浜国立大学大学院 環境情報研究院

〒 240-8501 横浜市保土ヶ谷区常盤台 79-7

E-mail: {echizenn, cross, tommy, arisawa}@arislab.dnj.ynu.ac.jp

あらまし

本稿では、我々が住んでいる実世界の意味情報までもを含めた情報を配信することを目的として、これを実現するために必要な新しい情報キャリアとして時空間オブジェクト情報キャリア STOIC を提案する。また STOIC は情報キャリアとして配信されるため、データ構造が人間にとっても分かりやすくなければならない。そこで構造化文書によって表現することを考え、XML を用いて STOIC を表現する。また、このようにして設計・記述された STOIC の有用性を示すために、CG を用いた表示システムを実装する。

Semantics expression of MultiMedia structured document STOIC with XML and its estimation

Kenji Echizennya Kazuaki Matsuda Takashi Tomii Hiroshi Arisawa

† Graduate School of Environment and Information Sciences, Yokohama National University

†† Division of Electrical and Computer Engineering, Faculty of Engineering,
Yokohama National University

††† Faculty of Environment and Information Sciences, Yokohama National University
79-7, Tokiwadai, Hodogaya-ku, Yokohama 240-8501 JAPAN

E-mail: {echizenn, cross, tommy, arisawa}@arislab.dnj.ynu.ac.jp

Abstract

In this paper, for the purpose of sending the information even including the semantic information on the real world where we live, we propose Spatio Temporal Object Information Carrier STOIC as the novel information carrier in order to realize sending the information even including the semantic information. Since STOIC is sent as information carrier, its data structure must be intelligible not only for computer but also for human. Therefore, we have considered expressing STOIC by the structured document, we decide to express STOIC using XML. Moreover, we construct the visualization system as CG in order to show its usefulness.

1 はじめに

我々が住んでいる3次元空間と、その空間中の「もの」とその「振舞い」によって表現される実世界の情報は非常に様々な「意味」を持っている。本稿ではこのような実世界の情報を、計算機上で配信することを目的とした、情報の表現手法を提案し、それを評価した。

このような実世界の情報を計算機上で扱う技術として、コンピュータグラフィックス (CG) があげられる。しかしCGは視覚的な情報を計算機を用いて表現することを目的としている技術なので、配信される情報というものは、その3次元空間の空間的な情報、「物体」の3次元的な形状、また、ある時点における位置変化などの、視覚的な情報のみとなってしまう、実世界を忠実に表現しているとはいえない。例えば、実世界ではある時点のオブジェクトAと別の時点のオブジェクトAはオブジェクトAという関係において、同一であるが、CGでは別々のオブジェクトとして表現してもよく、その同一性を保証しない。また、人間を表すオブジェクトAがある時点で物体Bをもちあげたことを表現するのに、実世界では、物体BはオブジェクトAの手に従属して動くという制約を満たし動いているにもかかわらず、CGでは、手と物体Bを別々に同じ方向へ動かし、視覚的には持っているように見えるが、そのデータには全くそのような意味的な情報は表現されていない。そこで我々は3次元空間と、その空間中の「物体」の3次元的な形状、また、その時点による位置変化などの視覚的に捉えることができるものだけでなく、その「物体」が持っている属性や、「物体」もしくはいくつかの「物体」の時間的な変化によって表現される概念的なものなどの、オブジェクトやオブジェクト間の意味情報までもを含めた実世界を表現する手法を提案する。以下このような実世界のことを時空間と呼ぶ。

また、我々はこの時空間の情報を配信することを目的としているので、これを記述する際に以下の2つの条件を満たしていなければならないと考えた。

- データ構造が人間にとっても理解しやすい
- 実世界の情報を忠実、また、一般的に記述してある

まず、これらのデータは人間によって扱われるもの

なので、人間が見た時に分かりやすくなければならない。そのため、この情報を構造化されたドキュメントによって記述することにした。ここでいうドキュメントというのは、バイナリーやアプリケーション依存の文字・数値の羅列ではなく、テキスト化された文書のことである。構造化することによって、その構造自体に意味を付加し、人間にも計算機にもその内容が把握しやすくなる。また、2つめとしてはだれがみてもおおよそ一緒である情報の表現によって、表示したりする環境、アプリケーションに依存しなくなると考えた。

以上のような考えに基づいて、本稿では対象とする一つのシーンをその内容までもを含めて表現し、伝えることができるキャリアを時空間オブジェクト情報キャリア (Spatio Temporal Information Carrier: STOIC) を提案し、STOICを用いたオブジェクトやオブジェクト間の意味情報までもを含めた実世界の情報を表現するシステムの構築を試みた。我々はこのSTOICをそのままデータベースに格納することを考えているのではない。データベースには、時空間を構成する要素がその素材に分けて様々な見方ができるように不偏的な形で記述されている [1][2]。STOICはそれらの素材を組み合わせて再構成される、ある一つのシーンを記述し、配信するものである。このSTOICの実現手法として、その枠組をXMLスキーマを用いて設計、実際のシーンをそれに従ったXML文書で記述した。XMLを使用した理由としては、XMLは自由に構造を定義できるマークアップ言語であり、その表現能力も高く、その構造やタグによっても意味を表現できるため、STOICの構造を定義し、記述する際に有効である。様々な業界でデータ、もしくは構造化文書を記述する際のデファクトスタンダードになりつつある。また、特別なソフトウェアが必要がないため、どんな計算機でもデータを見ることができる。さらに、様々な支援ツールやAPIが公開されているなどの理由があげられる。

本論文の構成は、以下のようになっている。第2章では、実世界の情報とは、第3章では、STOICの説明とXMLスキーマによる設計、また、その意味表現能力の評価、第4章では、STOICの視覚化システムの実装とその有用性、第5章では、まとめ、について述べる。

2 時空間情報のとらえかた

2.1 時空間情報の表現と伝達

我々の目的は様々な「意味」を持った実世界のあ
る時間的、空間的に閉じた一部分(シーン)を記述し
配信・通信することである。このシーンというもの
は3次元の空間であり、このような3次元的な空間
のデータを扱う既存の技術として、CGが挙げられ
る。CGは3次元空間中の物体の位置や形状、また
は、時間的に変化する位置などを、視覚的に表現す
るための技術で、その空間中で表現される「意味」
を表現することを目的としていない。例えば、ある
時点のオブジェクトAと別の時点のオブジェクトA
は実世界ではオブジェクトAであることで同一であ
るが、CGではこれらが同一のものであることを保
証しない。また、人間を表すオブジェクトAがある
時点で物体Bをもちあげたことを表現するのに、実
世界では、物体BはオブジェクトAの手に従属して
動くという制約を満たし動いているにもかかわらず、
CGでは、手と物体Bを別々に同じ方向へ動かし、
視覚的には持っているように見えるが、そこには
全くそのような意味的な情報は表現されていない。
我々が考えている、様々な「意味」までも含めて表
現するにはCGの表現能力は不十分である。

また、インターネット・モバイル通信の普及にと
もない、映像や3次元コンテンツの配信が盛んに行
なわれるようになった。配信するデータは不特定多
数のユーザであるので、要求される内容も様々であ
る。例えば、配信された情報をCGを用いて表示す
る場合、まずユーザの計算機環境の差によって同じ
「球」を表示するにも、ある人は非常に高性能な計算
機を使っているので、非常に詳細な球を表示したい
と要求し、また別の人は性能があまりよくない計算
機を使用しているので、計算機の演算能力に合わせ
た、詳細度をおとした球を表示させたいと要求する
ことも考えられる。このような場合、配信する側が
その要求に応えるために、同じシーンを表すいくつ
ものコンテンツを、その詳細度を変えて用意してお
くのは効率的ではない。さらに、ある人は高価なシ
ミュレーションソフト用のデータ形式で、別の人は
既存のCG表示用言語用のデータ形式が欲しいとい
うような、データ形式を指定した要求も考えられる。
このように、概念的に同じシーン、同じオブジェク

トなどを表現しているにもかかわらず、それぞれの
要求するデータ形式や、表示形式のコンテンツを用
意するというのもまた同様に効率的ではない。

以上のような問題点を解決するために、我々は、
伝えたい情報というものは環境やアプリケーション
によらず一つであって、その情報がどのようなデー
タ形式で処理されるかがアプリケーションによって
決まるべきだと考える。また、そのデータ構造は計
算機だけでなく、人間にとっても理解しやすくなけ
ればならないと考へ、実世界の情報を忠実に、また
一般的に表現するデータをテキストベースの構造化
文書によって記述することにした。

2.2 時空間情報の意味表現

前節では我々が表現しようとしている実世界の情
報は様々な意味を含んでおり、3次元空間の見かけ
だけの情報のデータを扱うCG技術では表現し切れ
ないことを述べた。このような問題を解決するため
に単にCGを拡張していくといった方法ではなく、
もう一度実世界というものを概念的にとらえ、新し
い情報記述について考える必要がある。

まず、我々はおおよそ実世界を表現するために次
のように考えた。まず、実世界というものをシー
ンに存在する「もの」とその時間的変化によって表
現する。ここで、シーンに存在するオブジェクトはそ
のシーンで存在しなくなるということではなく、必ず
存在し、また同一の「もの」である。シーンに存在
する「もの」はそのシーンで起こる事象に対して影
響を持たないオブジェクトと影響をもつオブジェク
トに分けて考える。前者を「場」、後者を「オブジェ
クト」と呼ぶ。この「オブジェクト」の時間的変化
を「変化」と呼び、この「変化」もしくは、「変化」
間のつながりによって表現されるものを「イベント」
と呼ぶ。実世界を表現するにはおおよそこれらの要
素が必要だと考えた。

これらの要素に対してさらに同じような意味を持
っているということ表現するために、タイプの同一
性について考える。例えば、Aという「オブジェク
ト」が人間である、と表現した場合、そのオブジェク
トAは人間であるという意味を持ち、手や足などを
持っているという情報を持つ。また、オブジェクト
Bを人間と表現した時、これらは同様の情報を持
っているはずである。このタイプという概念をもちい


```

<complexType name="human">
  <xsd:element name="HumanName"
    type="xsd:string"/>
  <xsd:element ref="HumanParts"/>
  <xsd:element ref="HumanStructure"/>
</complexType>
<xsd:element name="HumanParts">
  <xsd:complexType>
    <xsd:all>
      <xsd:element name="Hip" type="parts"/>
      <xsd:element name="Stomac" type="parts"/>
      <xsd:element name="LeftUpperLeg"
        type="parts"/>
      以下同様なパーツの定義なので省略
    </xsd:all>
  </xsd:complexType>
</xsd:element>
<xsd:complexType name="parts">
  <xsd:element ref="shape"/>
  <xsd:element ref="center"/>
</xsd:complexType>
<xsd:element name="HumanStructure">
  <xsd:complexType>
    <xsd:all>
      <xsd:element name="RightHip" type="joint"/>
      <xsd:element name="LeftHip" type="joint"/>
      以下同様なジョイントの定義なので省略
    </xsd:complexType>
  </xsd:all>
</xsd:element>
<xsd:complexType name="joint">
  <xsd:element ref="JointPosition"/>
  <xsd:element ref="connect"/>
</xsd:complexType>

```

図 2: XML スキーマの人間タイプの表現

間の構造を表す HumanName 要素、HumanParts 要素、HumanStructure 要素を持つ。HumanParts 要素は図 1 で示したパーツの集まりを要素として持つ。それらのパーツ要素は、その形状を表現するための Shape 要素と、その中心がどこにあるかを示すための center 要素を持つ。HumanStructure 要素はその子要素として、図 1 で示したジョイントの集まりを要素として持つ。それらのジョイント要素は、どの二つのパーツがつながっているかという参照と、そのパーツのつながっている位置を表す JointPosition 要素と、connect 要素をもつ。このように XML スキーマでタイプを定義することによって、例えばオブジェクト A とオブジェクト B は人間という意味で同一であるというタイプの同一性を示すことができる。

図 2 で記述した XML スキーマを元にして、XML インスタンスを記述することを考える。記述した例を図 3 に示す。

これは人間の構造・形状を、頭や胸などのパーツの集合と、それらをつなげるためのジョイントの集合によって表現する際に、ジョイントがどのパーツとどのパーツを結んでいるかを、パーツ集合中の要素への参照を用いることによって表現している。このように参照を用いることによって、例えば、オブジェクト A の右肩とつながっている「右上腕」とオブジェクト A の右肘につながっている「右上腕」が同一であるというインスタンスの同一性を示せる。

3.3 XML を用いた STOIC の記述能力の評価

これまでに時空間の情報を表現するためにはその視覚的な情報だけでなく、オブジェクトやオブジェクト間のつながりによって表現される意味情報が必要であることを述べた。そして、その枠組として STOIC を提案し、XML スキーマを用いて設計した。本節ではこの STOIC の記述能力の評価を行なう。

実世界の情報を表現するためにはオブジェクト同士の同一性や、オブジェクトの構造が持つ意味、シーンで起こった事象の意味、事象とオブジェクトの動きの関係が挙げられる。

STOIC では XML スキーマでのタイプの定義とインスタンス同士の参照を用いることによって、そのタイプの同一性、インスタンスレベルの同一性を表

```

<object>
  <HumanName>A</HumanName>
  <HumanParts>
    <Chest>
      <shape>
        <Box>
          <width>0.4</width>
          <height>0.3</height>
          <length>0.2</length>
        </Box>
        <center>0 0 0</center>
      </shape>
    <center>0 0 0</center>
    </Chest>
    <LeftUpperArm>
      <shape>
        <Cylinder>
          <radius>0.1</radius>
          <height>0.35</height>
        </Cylinder>
        <center>0 0 0</center>
      </shape>
    <center>0 0 0</center>
    </LeftUpperArm>
    以下同様にパーツのインスタンスなので省略
  </HumanParts>
  <HumanStructure>
    <LeftShoulder>
      <JointPosition>2 0.1 0</JointPosition>
      <connect>
        <locator href=
          "origin().ancestor(1,JointPosition)"/>
        <locator href=
          "origin().ancestor(1,Chest)"/>
      </connect>
      <JointPosition>-0.1 0.17 0</JointPosition>
      <connect>
        <locator href=
          "origin().ancestor(1,JointPosition)"/>
        <locator href=
          "origin().ancestor(1,LeftUpperArm)"/>
      </connect>
    </LeftShoulder>
    以下同様にジョイントのインスタンスなので省略
  </HumanStructure>
</object>

```

図 3: XML インスタンス

現できる。オブジェクトの構造が持つ意味やシーンで起こった事象の意味、その事象とオブジェクトの動きの関係についてはさらなる考察が必要だと考えられ、今後の課題となる。

4 STOIC と表示システムの実装

4.1 三次元グラフィックスと STOIC

本節では STOIC を用いて表現されたシーンを三次元コンピュータグラフィックスを用いて視覚化する実装システム述べる。

前章では STOIC は XML スキーマによって表現し、そのインスタンスであるシーンは XML インスタンスとして記述した。XML という構造化文書でシーンを記述することにより、テキストエディタでその XML インスタンスを見ることでシーンの概要を理解することはできる。しかしながら、実際に XML インスタンスに表現されているシーンは三次元空間の視覚的情報を含み、文字で書いてあるだけではその空間を直観的に把握しにくい。よって、シーンの視覚的情報がユーザーに直観的に把握できるように、三次元コンピュータグラフィックスを用いてそのシーンの三次元空間を構築することが必要となる。

今日、三次元コンピュータグラフィックスを利用するための API は多種多様である。また、そのような API を用いて実装されているアプリケーションも非常に多い。このような既存のアプリケーションのファイルフォーマットに変換するコンバータを用意する事も有用である。その一つとして、我々はすでに VRML への変換を試みている。しかしながら、既存のアプリケーションを利用するだけでは、STOIC に記述されている意味・内容はコンバータを作成する際に用いられることはあってもそのアプリケーションで利用することはできない。そのため、STOIC に記述されている情報を利用するためのアプリケーションの枠組と STOIC 用 API の構築を必要とする。本稿では以上のような事を前提としつつ、三次元コンピュータグラフィックス API を用いて STOIC の視覚化 API の構築と STOIC Viewer の設計を行う。今回、三次元コンピュータグラフィックス API として Java3D を用いた。

XML と Java は Web 指向、オープン性、ベンダ中立という同じ立場を取っていることにより、最近、

両技術を利用した設計・開発が多く見られる。Sun microsystems は XML と Java は補間技術であるとし、JAXP, JAXB, JAXM といった XML 用 API を多々提供している。

本稿では、このような背景から XML を用いて設計された STOIC 用のアプリケーション作成、API の構築を行なうに当たって Java 言語を選び、Java 上での三次元コンピュータグラフィックス API として Java3D を利用する。OpenGL の Java バインディングとして GL4Java[6], JSparrow[7], Magician 等があるが、今後、統一的なものとなるかに疑問があるため利用を控える。

4.2 Java 言語を用いた STOIC アプリケーションの構築

本節では Java 言語を用いた STOIC アプリケーションの構築について述べる。STOIC アプリケーションのシステム構成を図 4 に示す。このシステムを実現するための要素技術について順に説明する。

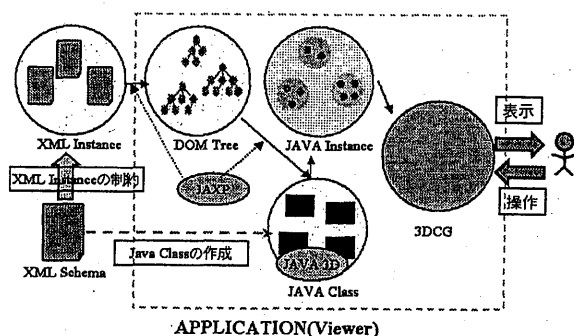


図 4: システム構成図

4.2.1 XML スキーマと Java クラス

STOIC を設計した XML スキーマを基に、その構造に沿って Java クラスを生成する。STOIC アプリケーションの設計はこの Java クラスを用いて行なわれる。

Java クラスの作成手順は以下のようになる。

1. 作成した XML スキーマを基に必要な Java クラスを想定し、そのクラスのプロパティを定義する。
2. XML スキーマで定義されるプロパティに accessor メソッド、mutator メソッドを与え、Java インターフェースを生成する。
3. インターフェース用のメンバ変数を定義するインプリメンテーション・クラスと、インターフェース内の各メソッドのためのコードを生成する。

主に階層構造表現における親子はクラスのメンバとして定義する。また、STOIC の階層構造の leaf には数値や文字列だけでなく、グラフィックオブジェクトが配置される場合がある。グラフィックオブジェクトが定義されている場合は、Java3D クラスのインスタンスを親のメンバとする。

4.2.2 JAXP と DOM API

XML をプログラムで処理する場合、DOM や SAX といった XML プログラミングの標準モデルが定義されている。また、Sun microsystems が提供する JAXP を用いることによって DOM や SAX をベンダー非依存的な方法で使おうとするときに障害となるベンダー固有の作業が可能となる。

今回のアプリケーション作成では JAXP を用いて DOM Tree を構築し、Tree を隅無く巡回し、Node に対する操作を記述していく。Node に対する操作は主に前項で述べた Java クラスのインスタンスを生成することである。

4.2.3 Java3D とシーングラフ

Java3D はシーングラフをベースにした三次元グラフィック API であり、アプリケーションプログラミングを行なうには、シーングラフと呼ばれるオブジェクト・モデルを構築していくことになる。主に、以下のようなノード (DOM Tree の Node とは別もの) の追加を行ないシーングラフを構築していく。

- TransformGroupNode

移動、回転、拡大縮小など座標変換操作を適用する

- Shape3DNode
物体の形状や頂点座標、色などの属性を管理する
- BehaviorNode
物体の振る舞いを制御する

また、ViewPlatformNodeにより視点の移動や回転を制御できる。

4.3 STOIC Viewer

前節の構築システムに基づき、STOIC Viewerの実装を行なった。今回のViewerでは、STOICインスタンスにおける物体の形状や構造のデータを使用し、シーングラフを構築し三次元グラフィックとして描画している。

5 まとめ

我々が住んでいる3次元空間と、その空間中の「もの」とその「振舞い」によって表現される実世界の情報は非常に様々な「意味」を持っていて、本稿では、そのような実世界の情報を計算機間でやりとりすることを目的としてこれを実現するためにSTOICを提案した。STOICは情報キャリアなので、人間にも分かりやすいように、構造化された文書で記述することを考え、構造化文書を記述する一手法であるXMLを用いてSTOICを表現した。また、我々が提案・定義したSTOICの意味表現能力に対する評価を行ない、最後にその有用性を示すために、実際にCGを用いた表示システムを実装した。

参考文献

- [1] H.Arisawa,H.Nagae,Y.Mochizuki:“Representation of Complex Objects in Semantic Data Model“AIS” and Implementation of Set Operators”,IEICE TRANSACTIONS,vol.E74,No.1,1991.
- [2] H.Arisawa,T.Torii,H.Yui,and H.Ishikawa: “Data Model and Architecture of

Multimedia Database for Engineering Applications”, IEICE TRANS.INF. & SYST.,vol E78-D,No.11,1995.

- [3] 松田一章,越前谷健二,富井尚志,有澤博:“XMLを用いたシーンの時空間オブジェクト情報の記述”,電子情報通信学会データ工学ワークショップ,2001
- [4] XML 1.0(Second Edition):
- [5] XML Schema:
Part0:Primer
Part1:Structures
Part2:Datatypes
- [6] GL4Java:
“<http://www.jausoft.com/gl4java/>”
- [7] JSparrow:
“<http://www.pfu.co.jp/jsparrow/index.html/>”

スキーマダイアグラムを用いたマルチメディアデータベース 検索手法

岸邊 佳世子 † 家富 誠敏 †† 富井 尚志 ††† 有澤 博 †††

† 横浜国立大学大学院 環境情報学府

†† 横浜国立大学 エコテクノロジーシステムラボラトリー

††† 横浜国立大学大学院 環境情報研究院

〒240-8501 横浜市保土ヶ谷区常盤台 79-7

E-mail: {kayo, eto, tommy, arisawa}@arislab.dnj.ynu.ac.jp

あらまし

本論文では、マルチメディアデータベースの利用者がスキーマやインスタンスの状況を直観的に把握できることを目的として、データベースを視覚化し、利用者が見たい・知りたいデータを自由にブラウジングすることができるインタフェースを提案する。マルチメディアデータベースでは複雑なスキーマに沿って大量のインスタンスが蓄積されている場合が多い。その際、スキーマ全体、インスタンスの全てを限られた画面上に視覚化することは物理的にも不可能であると同時に、必ずしも利用者の直観的なブラウジングを手助けすることにならない。特に、その大量のインスタンスの中から利用者が必要とするインスタンスを探し出すのは大変困難である。そこで本論文では、インタフェースを用いてマルチメディアデータベースを視覚化するとともに、そのスキーマおよびインスタンスのうち、必要とする部分を自由に抜き出して視覚化する手法として“ふるい”操作の提案を行う。

Diagram-based Querying for Multimedia Database

Kayoko KISHIBE †, Masatoshi IETOMI ††, Takashi TOMII †††, and Hiroshi ARISAWA †††

† Graduate School of Environment and Information Sciences, Yokohama National University

†† Ecotechnology System Laboratory, Yokohama National University

††† Faculty of Environment and Information Sciences, Yokohama National University

79-7, Tokiwadai, Hodogaya-ku, Yokohama 240-8501 JAPAN

E-mail: {kayo, eto, tommy, arisawa}@arislab.dnj.ynu.ac.jp

Abstract

In this paper, we propose the interface which the user can browse the visualized schema and instances, for the purpose of the user can grasp intuitively the schema and instances of multimedia database. In many cases, a lot of instances are accumulated along with the complicated schema in the multimedia database. In that case, it is physically impossible to visualize all the schema and instances on a screen, and a user becomes unable to perform browsing intuitively. It is very difficult especially to discover the instance which a user needs out of huge instances. Then, in this paper, while a multimedia database is visualized using an interface, as a method of the operation called “sift operation” that extracts freely the portion needed among the schema and an instance, is proposed.

1 はじめに

近年の計算機技術の発達により、文字、数値だけでなく、画像、映像、音声、3次元グラフィックスなどの大容量なマルチメディアデータを計算機上で取り扱うことが可能となってきた。それに伴い、マルチメディアデータを統括的に蓄積、管理、運用するためのデータベース、いわゆるマルチメディアデータベースの重要性が増してきている。

ところでマルチメディアデータベースに限らず、データベースを利用する際には、データベースの構造(スキーマ)や入力されている値(インスタンス)の状況を把握することが重要である。データベースの中にどのようなデータがどのように蓄積されているかを知ることによってデータベースを有効に活用することが可能になるからである。しかし、マルチメディアデータベースにおいては、映像、音声、画像などの様々なデータを、データを解釈して得られる多様なインデックス情報とともに蓄積することが求められ、そのためスキーマは一見複雑な構造となることが多い。このような一見すると複雑なデータベースのスキーマ、さらにインスタンスの状況を一般利用者が把握することは困難を伴うことが予想される。

これを解決するためには、マルチメディアデータベースのスキーマやインスタンスを視覚的に提示することで利用者の直観的理解を助ける手法が考えられる。特に画像や映像、グラフィックスのように視覚的なデータを扱うマルチメディアデータベースであればこそ、スキーマやインスタンスを視覚的表示することにより、より直観的な情報の把握が可能になると期待される。

このような考えに基づき我々は、拡張したERダイアグラムを用いてスキーマやインスタンスを概念構造そのままに視覚的表示するユーザインタフェースAVIS(Advanced Visualizer of Instance and Schema)を提案している。このAVISを用いることでデータベースを視覚化し、そのスキーマやインスタンスの状況を直観的に把握することが可能になる。

しかし、データベースには非常に大量のインスタンスが入っており、この全てを画面に表示することは、限られた画面上では物理的に困難であると同時に、利用者が自分の知りたい情報を知る上での妨げにもなってしまう。またスキーマに関しても、イン

スタンスと比較すればデータ数は多くはないが、必ずしもその全てを表示することは望ましくない。このとき、もし利用者が必要とする部分がデータベース全体ではなく、その一部分であるならば、利用者の要求する部分だけを抜き取り、表示することができれば有用であると考えられる。

本研究では、視覚的にデータベースの概念構造を表示し、それに対して操作できるインタフェースの設計とその実現手法について議論するとともに、本インタフェースを用いて、スキーマダイアグラム上に特に着目しているインスタンスを例として挙げた図を描くことにより、必要な部分のみを取り出して表示するという、データベースに対する“ふるい”という操作の手法を提案する。“ふるい”操作では、利用者には必要のない情報をデータベースから文字通りふるい落とすことで必要とするスキーマやインスタンスだけを画面に表示させる。これにより直観的な操作で複雑なスキーマや大量のインスタンスから利用者にとって必要な部分のみを取り出すことが可能になる。

2 視覚的インタフェースAVIS

マルチメディアデータベースは一見して複雑なスキーマを持ち、そこに大量のインスタンスが蓄積されている場合も数多い。そのため、データベース利用者がそのスキーマやインスタンスを見たい、知りたいときに、スキーマやインスタンスの状況を把握することは困難となる。そこでマルチメディアデータベース利用者が、スキーマやインスタンスの状況を直観的に把握できることを目的とし、マルチメディアデータベースを視覚化[1][2][3]し、その中から利用者が見たい・知りたいデータを自由にブラウジングできるインタフェースとして、我々はAVIS(Advanced Visualizer of Instance and Schema)[4][5]を提案している。このAVISの概念構成を図1に示す。

AVISを用いてデータベースのスキーマやインスタンスをブラウジングするためには、次のような機能が必要であると考えられる。

- データベースにおけるスキーマやインスタンスの直観的な視覚化
- データベースに蓄積されているデータから着

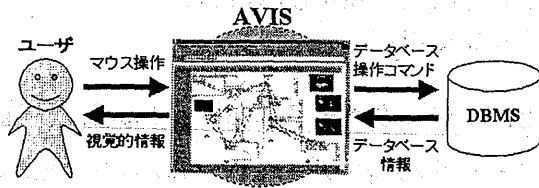


図 1: AVIS の概念構成

目している部分を抜き出して表示すること（ふるい操作）

- “ふるい” 操作によって抜き出した結果をブラウジングできること

まず本章において、視覚化およびブラウジング手法について述べ、次に次章において、本論文で新たに提案する“ふるい”操作について述べる。

2.1 データベースの直観的な視覚化とブラウジング

データベースを利用する上では、スキーマやインスタンスの状況を把握することが重要である。例えば、マルチメディアデータベースには、文字、映像、画像などのいろいろな値だけでなく、その解析情報が加えて蓄積される。そのため、そのスキーマはそれらデータ間のつながりが相互参照の入り組んだ一見複雑な構造となる。このようなマルチメディアデータベースのスキーマやインスタンスを視覚的にブラウジングすることでデータベースのスキーマ、およびインスタンスの複雑な関係を把握することが容易になると考えられる。

ところで、データベースの、特にスキーマの視覚化手法としては、既に意味データモデルである ER モデルに基づくスキーマのダイアグラム（図式的）表現が有用であることが知られている。我々はこの ER ダイアグラム [6] を拡張した表記法として、拡張 ER ダイアグラムを提案し、特にマルチメディアデータベースのデータベース化において有効であることを示してきた [7][8]。その視覚的インタフェースである AVIS ではこれを用いてスキーマの視覚化を行い、インスタンスもまたスキーマに重ねてダイアグラムで表現している。拡張 ER ダイアグラムで

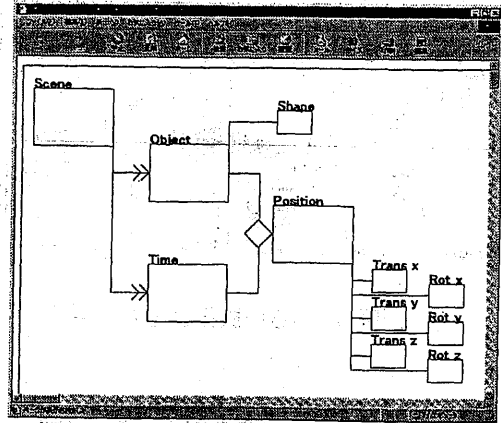


図 2: AVIS を用いたスキーマ表示例

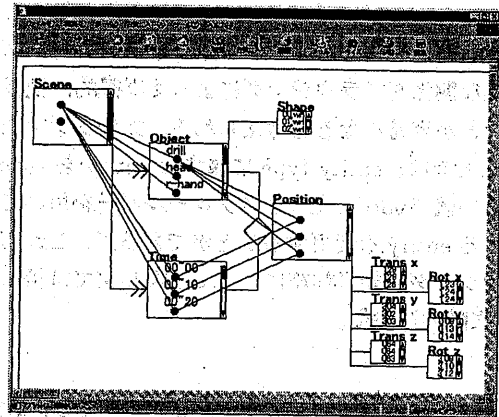


図 3: AVIS を用いたスキーマ・インスタンス表示例

は ER ダイアグラムと同様に entity-type(主体型) と relationship-type(関連型) という 2 つの概念によって対象世界をモデル化する。この拡張 ER ダイアグラムを用いたスキーマの視覚化例を図 2 に示す。

図 2 のスキーマ図では、ある時空間中の場 (Scene) に存在する物体 (Object) およびその場の時点 (Time)、またある特定の物体がある特定の時間に存在する位置 (Position) を表している。図 3 はさらにこのダイアグラムを用い、スキーマに重ねてインスタンスを表示した例である。図 3 では entity-type に属する entity が丸 (値を持つ場合は値を併記) で表され、それらの間のつながり (relationship) が線分で表現されている。

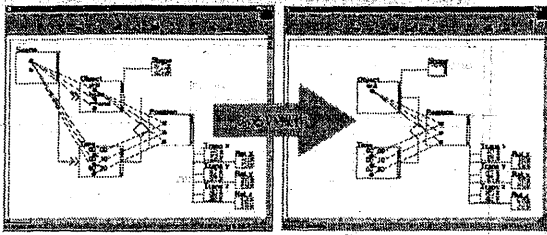


図 4: ふるい操作とその結果

我々が提案している視覚的インタフェース AVIS では、この拡張 ER ダイアグラムを用いてデータベースのスキーマ、およびインスタンスの視覚化を行っている。

このように、データベースの視覚化を行うことでデータベースのスキーマ、およびインスタンスの複雑な関係をブラウジングによって直観的に把握することが容易になると考えられる。ブラウジングを行う上では、entity-type に属する entity を entity が持つ値で sort したり、スクロールバーを用いて表示する entity を変更することができるが、これらブラウジング機能の詳細については本論文では扱わない(文献 [4] 参照)。

3 視覚的インタフェース AVIS を用いた“ふるい”操作

第2章では、我々の提案している視覚的インタフェース AVIS について述べた。このように視覚的インタフェースを用いてデータベースのスキーマやインスタンスを表示することで、直観的にそれらを見ることができ、データベースの中身をブラウジングする際に有用であると考えられる。

しかし、スキーマ全体、インスタンスの全てを画面上に視覚化することは物理的にも不可能であると同時に、必ずしも利用者の直観的なブラウジングを手助けすることにならない。特に、その大量のインスタンスの中から利用者が必要とするインスタンスを探し出すのは大変困難である。このことから利用者が必要とする部分のみを取り出して、ブラウジングしたいと考えることは自然な要求だと考えられる。

本章では、このように利用者が必要とする部分だ

けを取り出して視覚化する手法として、“ふるい”操作の提案を行う。

3.1 データベースにおける“ふるい”操作

マルチメディアデータベースにおける、複雑なスキーマ、および大量のインスタンスの中から、利用者が必要とする部分のみを取り出してブラウジングするためには、その部分のみをまるでもふるいにかけるかのように取り出してブラウジングする、という操作が考えられる。本論文では、このような操作を“ふるい”操作と呼ぶ。

“ふるい”操作は具体的には次のようにして行う。まず利用者が取り出したいデータを

- スキーマのどの部分に着目して“ふるい”操作を行うのかを記述する
- “ふるい”操作によって取り出したいインスタンスの条件を記述する

ことによつて示し、次に、実際にこの“ふるい”操作をデータベースシステムが実行することにより、その結果として取り出されたスキーマおよびインスタンスを視覚化して表示する。これによつて利用者がブラウジングを行う上で表示することが不必要と判断した部分を“ふるい”落として必要な部分だけを取り出して画面上に視覚化することができる。このふるい操作は、既に我々が提案しているマルチメディアデータベース検索言語である MMQL[9] における検索式の結合子で結ばれた要素操作のひとつ分に対応している。

3.2 “ふるい”操作のための基本要素

本節では、ふるい操作を記述するために必要な基本要素について述べる。

AVIS ではふるい操作を行う際、どのようなデータを取り出したいのかを視覚化されたスキーマダイアグラム上に図として描く。これによつて、

- スキーマダイアグラムと照らし合わせて直観的に取り出したいデータを記述することができる。
- 図として描くことによつてどのように“ふるい”を行うのかを直観的に把握することができる。

entity変数	☆	entityを表す変数 (root)
	□	entityを表す変数 (root以外)
つながり	☆ — □	2つのentity変数間の つながり (経路)
条件式	> 50	entity変数が 満たすべき条件
Label	\$label	entity変数の値を 参照する為のラベル

図 5: ふるい操作の基本要素

といった利点が考えられる。図としてふるい操作を記述する際の基本要素は図 5 のとおりである。この基本要素は 3.1 節で述べた 2 つの操作にそれぞれ対応しており、entity 変数とつながりという要素を用いて取り出したい(着目している)部分を記述し、条件式と Label を用いて取り出したい条件を記述する。

以下でこれらの要素について詳しく述べる。

3.2.1 ふるい操作における取り出したい部分の記述

AVIS を用いたふるい操作では、取り出したい部分を記述するために、以下の 2 つの基本要素を用いる。

1. entity 変数

ある特定の entity-type に属する任意の entity を表す変数のことをここでは entity 変数と呼ぶ。entity 変数は指定した entity-type 中を範囲として、それに属する entity 上を動きまわる変数である。entity 変数は図上では星型 (root を表す)、または四角形 (root 以外を表す) のいずれかで表される。(図 5)

- 2 つの entity 変数間のつながり (relationship)
entity を entity 変数で表現したとき、2 つの entity 変数間のつながりを図 5 のように 2 つの entity 変数間を線を用いてつなぐことで示

す。entity 変数間をつながりで示すことによって、それぞれの entity 変数が表す entity 同士が、実際につながりを持っていることを表す。

以上の要素を用いることで、ふるい操作において取り出したい部分を記述することが可能となると考えられる。

ところで、実際に取り出したい部分を記述する上では、通常、ある entity-type に主眼を置いてその entity-type から参照したい他の複数の entity-type へのつながりを考えることが多い。例えば図 2 のデータベースにおいて、物体 (Obj)、時点 (Time)、および位置 (Position) について考えたとき、『物体 (Object)』に主眼を置いた場合には、「あるシーン中に存在するある特定の物体だけに限ってその位置の時間的変化を追いたい」、また『時点』に主眼を置いた場合には、「ある特定の時点において、あるシーン中に存在している Object の位置を知りたい」というように明らかに観点が違ってくる。したがって、取り出したい部分を記述するには、主眼を置く場所を明確に記述する必要があると考えられる。

ここではこの主眼を置く entity-type の entity 変数を特に root 変数と呼び、図 5 のように、entity-type 中に星型を図示することで記述する。また、この root 以外の entity 変数は図 5 のように、entity-type 中に四角形を図示することで記述する。

この基本要素を用いた図 2 のデータベースにおけるふるい操作の例として、「名前が "drill" である物体 (Object) の形状 (Shape) と、その位置 (Position) の時間的変化を示す」というふるい操作を図 6 に、またその結果および結果のグラフィックス表示をそれぞれ図 7、図 8 に示す。

図 6 では、Position という entity-type 中に root 変数を、またこの entity-type 中の entity とのつながりを見たい entity-type 中にその他の entity 変数を記述することでこの操作における取り出したいデータの構造を表現している。

3.2.2 ふるい操作における条件の記述

AVIS を用いたふるい操作では、取り出したい条件を記述するために、以下の 2 つの基本要素を用いる。

1. 条件式

名前が“drill”のObjectを選択

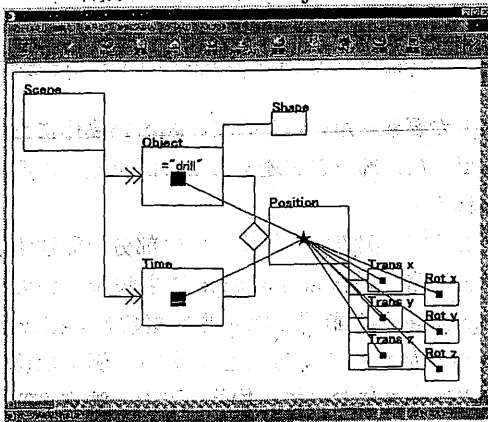


図 6: ふるい操作1の条件

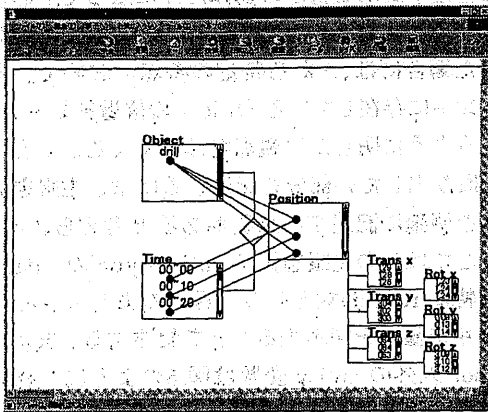


図 7: ふるい操作1の結果

大量のインスタンスの中から利用者にとって必要な部分がどの部分であるのかを条件および複数の条件の論理和 (OR)、論理積 (AND) を用いた条件式を記述することによって示す。このとき、各 entity-type 中の entity 変数は記述された条件を満たすかどうかチェックされ、条件を満たしているもの、およびその条件を満たしているものとのつながりを持っているものだけが取り出される。

2. Label

条件式を用いてデータベース中から取り出したい条件を記述する際、直接インタフェースに文字や数値のデータを入力することで、そ



図 8: ふるい操作1の結果のグラフィックス表示

の定数と比較するだけではなく、他の entity-type 中の entity の値と比較することも考えられる。このような場合、他の entity 変数が表す entity の値を Label を用いて参照する。このように entity 変数に Label をつけることによってこの entity 変数が表す entity の値を Label によって参照することができる。

以上の要素を用いることで、ふるい操作において取り出したい条件を記述することが可能となると考えられる。

このように条件でふるいをかけたインスタンスに対して、表示するかどうかを記述する際には、表示したい entity 変数を選択し、その選択された entity 変数を表す図形の色を黒く塗りつぶすことによって図示する。これによってデータベースにより指定されたスキーマおよびインスタンスだけが抜き出されて表示される。但し、もともとのスキーマには存在していた entity-type を表示する必要が無い場合には、つながりの両端の entity 変数が存在する entity-type 間に仮想的な relationship-type を生成してスキーマを表示する。

条件式を用いた例として図 6 のふるい操作では、取り出したい物体 (Object) として「名前が“drill”である」という条件を記述している。このように条件付けを行いたい entity 変数の上に条件式を記述することで、その entity 変数に対する条件を記述する。

3.2.3 仮想的な entity-type

ふるい操作を行う際、データベース中の entity や entity の列に対して演算を行うことで新しい entity を作成して見たいという要求が考えられる。

このような場合、3.2.2 節で述べた Label を用いて演算を行うことによって、もともとのスキーマには

各 Object における位置の x 座標の平均を計算

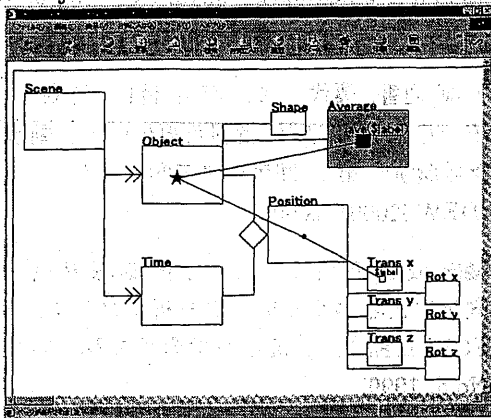


図 9: ふるい操作 2 の条件

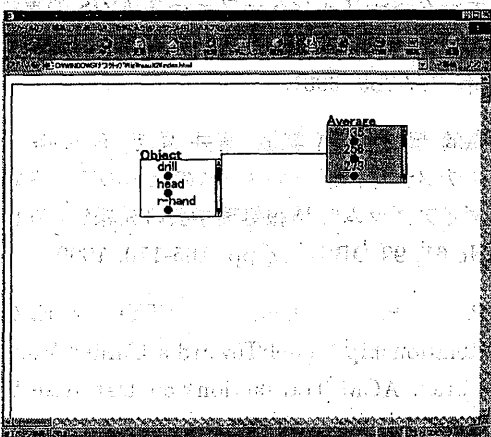


図 10: ふるい操作 2 の結果

存在しない entity-type および entity をまるでデータベース中に存在しているかのように仮想的に見ることができる。

このような仮想的な entity-type を作成するふるい操作の例として、図 2 のデータベースにおける「ある特定の物体 (Object) の位置 (Position) とその x 座標 (Trans x) の平均値を示す」というふるい操作を図 9 に、またこのふるい操作を行った結果を図 10 に示す。

図 9 ではもともとデータベースには存在しなかった Average という entity-type が新しく記述されている。このように仮想的な entity-type を記述することでその entity-type 中に新しく仮想的な entity

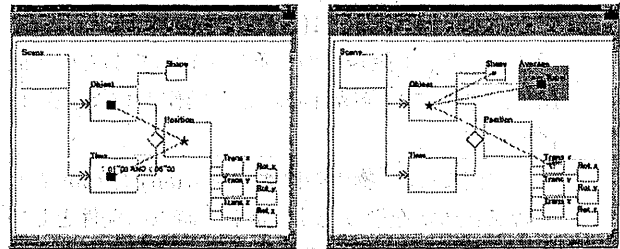


図 11: ふるいの組み合わせ (左:1 段階目 右:2 段階目)

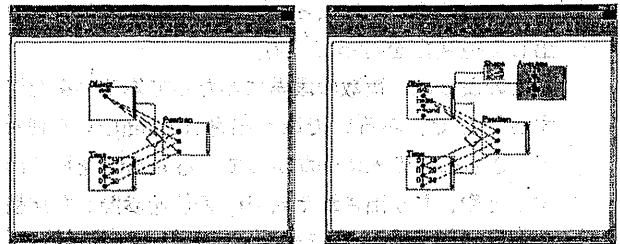


図 12: 図 11 の結果 (左:1 段階目の中間結果 右:最終結果)

を作成することが可能となる。

3.3 ふるい操作の組み合わせ

これらのふるい操作では基本操作を組み合わせることで着目した entity とその entity からつながっている entity の塊として取り出すことができる。しかし、

- つながっている entity のさらにその先につながっている entity を取り出したいとき
- ふるいで取り出した値、または生成された仮想 entity の値を参照してさらにふるいを行いたいとき

には、これだけでは行うことができない。そこで複数のふるい操作を組み合わせることで、前述のような複雑な操作を行うことを可能にした。

このようなふるいの組み合わせ例として、図 2 のデータベースにおける、「時点 (Time) が 1 秒から 5 秒までの間に存在する物体 (Object) について、その物体の形状、およびその位置の平均値を示す」と

いうふるい操作を図11(左:1段階目 右:2段階目)に示す。また、1段階目のふるいを行ったあとの中間結果、さらに2段階目のふるい操作の結果を図12に示す。

この例では、「物体の位置 (Position) を root として時点と形状を取り出す。但し、時点の条件として値が1秒~5秒の間のみ」という1段階目のふるいの結果に対して、さらに「1段階目に取り出された物体 (Object) を新たに root として、形状と平均 x 座標を取り出す」という2段階目のふるいを組み合わせかけてることによって、目的とする結果を取り出して視覚化を行っている。

このように、複数の段階にわたってふるいを記述することで、ふるいで取り出された entity を使って、さらにその先につながっている entity を取り出すことや、取り出された entity の値を参照した比較や演算を行うことができる。このとき、1段階目、2段階目、とふるい操作を手順的に記述することで操作を行う順番を明示的に指定することができる。このようにふるいを複数段にわたって記述することにより、MMQL[9]を用いた条件検索に相当する操作を行うことができる。

4 まとめ

本論文では、視覚的データベースインタフェース AVIS を用いたふるい操作について述べた。本手法により、データベースブラウジングを、視覚的・直観的な操作で行うことが可能になる。

本手法では、視覚的に表示したスキーマ上に直接、ふるい操作を記述することができる。また、そのふるい操作を行う際に、繰り返して複数段のふるい操作を行うことで、より複雑なふるい操作を順序を認識しながら行うことができる。これにより、一般利用者でも簡単な操作でマルチメディアデータベースのような複雑なスキーマ、および大量なインスタンスの中から必要な部分のみを取り出すことが可能になると考えられる。

参考文献

[1] Stuart K. Card, "Visualizing retrieved information: A Survey", IEEE Computer Graph-

ics and Applications, Vol.16, No.2, pp. 63-67, 1996

- [2] 古畑 理香, 藤代 一成, 市川 哲也, 竹島 由里子, "GADGET/IV: 情報可視化の半自動設計支援環境", 第11回データ工学ワークショップ (DEWS2000), 2000
- [3] 磯部 成二, 黒川 清, 塩原 寿子, 飯塚 哲也, "視覚化による多次元データ分析システム: INFORVISER", 情報処理学会論文誌, Vol.40, No.5, 1999
- [4] 家富 誠敏, 森寄 慎之, 富井 尚志, 有澤 博, "マルチメディアデータブラウジングのための視覚的データベースインタフェース AVIS の実装", 電子情報通信学会論文誌, VOL.J84-D-1 No.6, pp. 777-789, 2001.
- [5] 森崎 慎之, 家富 誠敏, 富井 尚志, 有澤 博, "マルチメディアデータベースのためのデータ定義ダイアグラム", 情報処理学会研究報告, Vol.99, No.61, 99-DBS-119, pp. 165-170, 1999.
- [6] P. P. Chen, "The Entity-Relationship Model: Toward a Unified View of Data", ACM Transactions on Database Systems, Vol.1, No.1, 1976.
- [7] H. Arisawa, H. Nagae, Y. Mochizuki, "Representation of Complex Objects in Semantic Data Model "AIS" and Implementation of Set Operators", IEICE TRANSACTIONS, vol.E74, No.1, 1991.
- [8] H. Arisawa, T. Tomii, H. Yui, and H. Ishikawa, "Data Model and Architecture of Multimedia Database for Engineering Applications", IEICE TRANS. INF. & SYST., vol E78-D, No.11, 1995.
- [9] 富井 尚志, 有澤 博, "マルチメディアデータベースにおける映像モデリングと操作言語", 電子情報通信学会論文誌, Vol.J79-D-II No.4, pp. 520-530, 1996.

HUMAN BODY/MOTION MODELING BASED ON THE "MEDIATOR" CONCEPT

Imai, S.*, Tomii, T.**, Arisawa, H. **

*Dept. of Comp. Sci., Gunma University, Japan

**Fact. of Env. & Info. Sci., Yokohama National University, Japan

In this paper, we propose a methodology for modeling of the human body and its motions in order to store and query them on a database. For storing all available data about human motions, we need a precise human body model as well as method for representing its motions. In a separate presentation, our group proposes bone-based human body model, coupled with a detailed computer graphics representation methodology [1]. Here we concentrate on the human motions modeling.

In a previous work, we have introduced the concept of mediator as a means for representing typical figures and structures in the human body, and its postures during some abstract motion [2]. Storage, retrieval, and replaying of the original motions all become possible through the mediator (Fig. 1).

Here, for the modeling of specific motions of human bodies in the real world, we define the following levels of a motion:

Real World Data level – A set of joint values for each of human body's joints. These values are extracted from "raw data" about the Real World Scene obtained by various devices, like cameras, scanners, etc.

Semantics level – Common motions or postures, as well as characteristic gestures. Semantic information is included as annotation into each posture's description.

Mediator level – Comprised of data about individual motions, gestures, and characteristic points, based on the common motions.

As a result, mediator-based modeling makes possible unified representation of individual human motions in the database and querying them.

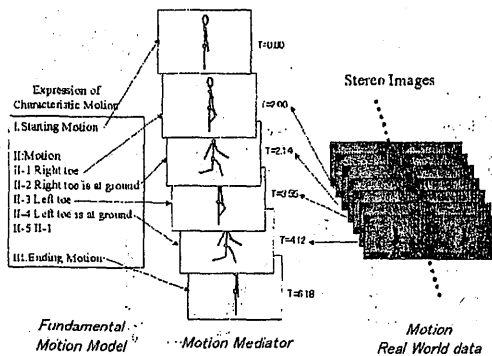


Fig.1: The Common Model, the Mediator and Real World Data of some motion

Reference:

Arisawa, H., Sato, T. Human Body motion simulation using bone-based human model and posture database, Premus2001. Arisawa, H., Imai, S. Mediator-based Modeling of factory workers and their motions in the framework of Info-Ergonomics, M. Takano, E. Arai and T. Arai eds., Human Friendly Mechatronics. Selected Papers of the International Conference on Machine Automation, 25-27, September 2000, Osaka, Japan Elsevier Science (to appear).

HUMAN-BODY MOTION SIMULATION USING BONE-BASED HUMAN MODEL AND POSTURE DATABASE

Arisawa, H.* Sato, T.**

*Fact. of Env. & Info. Sci., Yokohama National University, Japan

**Div. of Elec. & Comp. Eng., Yokohama National University, Japan

This paper presents a new methodology of human-body motion simulation with Bone-Based Human Model (BBHM) as a part of spatio-temporal analysis/modeling for human bodies and posture database.

Model-based simulation and load estimation are essential to the prevention of work-related musculoskeletal disorders and raise the need for precise models of human muscles and bones. Observable movements of the human body quite often involve muscle and bone movements that are by no means intuitive, thus making their recreation extremely difficult. However, in the medical science, data is available about the way larger-scale body movements translate into movements of part's bones. Several modeling approaches do exist, but they are highly specialized to the medical field and their use by non-specialists is not feasible.

In our approach an "action" of the real human body can be detected through video image analysis and mapped on a simplified human mock-up, called a "mediator". The mediator consists of minimum number of parts to describe the human structure and typical motions. Also, in order to achieve the precise motion simulation, we have developed a system that automatically derives the movements of precise skeleton model from a simple mediator, adjustable to the individual characteristics of humans.

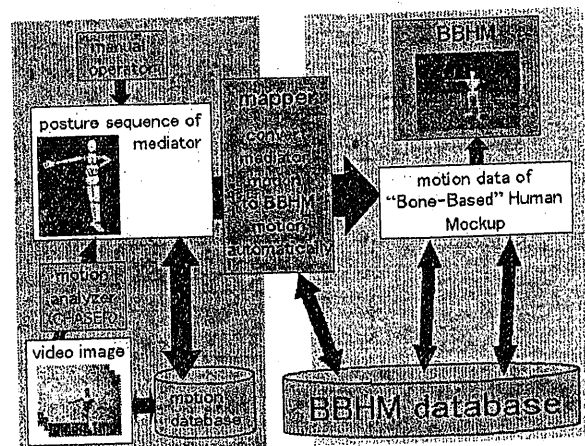
On the other hand, we have created a 3D graphical skeleton model by scanning human bones with 3D non-contact scanner. Based on medical knowledge, we have made the scalable model depending on the "real" human body.

Fig. 1 shows the total architecture of our system. With our system, it becomes possible to recreate postures of the skeleton model that faithfully represent real bone movements. Further work should concentrate on the implementation of muscle-related issues, load estimation based on the model, and detection of uncomfortable and strenuous postures, etc.

Fig.1 the total architecture of our BBHM database system

Reference

1. Hodgins. J, O'Brien. J. Perception of Human Motion With Different Geometric Models, IEEE Trans. On Visualization & Computer Graphics, 307-315, Vol. 4, No. 4, Oct.-Dec, 1998.



Motion Modeling and Simulation of Human Workers based on Info-Ergonomics Concept

Sayaka Imai¹, Takashi Tomii² and Hiroshi Arisawa²

¹ Department of Computer Science, Gunma University,
1-5-1 Tenjin-cho, Kiryu, Gunma 376-8515 JAPAN

E-mail:sayaka@cs.gunma-u.ac.jp

² Graduate School of Environment and Information Sciences,
Yokohama National University,

79-7 Tokiwadai, Hodogaya-ku, Yokohama, 240-8501 JAPAN

E-mail:tommy@ynu.ac.jp, arisawa@ynu.ac.jp

Abstract. This paper presents a new approach of modeling motion of factory workers and working simulation of their motion. Also a concept for Mediator-based human body/motion modeling for application to Info-Ergonomics is offered. In Manufacturing Process Design, many simulations are executed in order to improve efficiency of designer's work. But in many cases they are focusing on simulating and evaluating machines' performance. In the present paper we will propose a method of working simulation by using CG based on Info-Ergonomics. Also, we consider it as a possible way for unifying all the data used in various applications (CAD/CAM, etc) during the design process and evaluating all sub-systems in a virtual Factory.

1 Introduction

As a result of the significant advancement in computer technology many new applications of image and graphical data are brought to life. Those include the fields of Image processing and retrieval, Engineering Simulation, Tele-robotics, and so on. Computer graphics (CG) in particular has been widely used for a long time in CAD/CAM/CAE systems as a visualization tool. Recently some new applications supporting the designing process of a virtual factory and simulating its work are gaining popularity in areas like process planning and work scheduling. In addition to the CG, they involve many new other types of data, such as video and X-ray images, sensing information, etc. High complexity of those information causes a lot of difficulties to the researchers and developers. Here we would like to point out two problems we consider important for the future development of those systems.

First, as those applications use mainly application-oriented data, sharing data with other applications is quite difficult and very often —impossible. In our opinion, a flexible automation system requires data about a given product to be used in all applications.

Second, we think that eco-factory human factors are usually overlooked or underestimated. Human labor has an important role in the manufacturing process. It is examined from the viewpoint of industrial engineering, ergonomics, etc. and results are used

in product planning, working analysis, and work environment design. But aspects like comfort of the workers and optimal human-machine co-existence are not explored.

As a promising solution to the above problems we have proposed the use of Real World Database(RWDB) and Info-Ergonomics simulation[1]. RWDB is able to capture various types of data, namely characters, video images, 3D graphics, and shapes of 3D objects. We call all those *Real world data*. The data of all types are unified and stored in a Multimedia Database(MMDB).

On the other hand, we consider it necessary to focus on human-machine cooperation, especially for employees in the factories and analyze their work evaluation environment, and amenities against this background. We offer a model for analyzing employee's work evaluation and using the results in the manufacturing process design. We consider it possible to use Info-Ergonomics simulation data in human-machine cooperation simulation and in the production design, which would result in an integrated data model for manufacturing process applications. In order to realize this application, we consider that the human body/structure and motion modeling for the human Body/Motion database(for Working simulation system) is very important. We propose a methodology for modeling of the human body and its motions in order to store and query them on a database. For storing all available data about human motion, we need a precise human body model as well as method for representing its motions.

This paper presents an extended Info-Ergonomics system and Mediator-based modeling of workers in a factory, which includes the human body structure, work motion and work semantics.

2 Info-Ergonomics

2.1 Concept of Info-Ergonomics

Info-Ergonomics(Information Ergonomics) is the concept for computer aided system in the field of such as Ergonomics, Industrial Engineering, factory work design about factory workers and so on.

The technical bases of Info-Ergonomics are Real World Database System, Mediator based human body/motion modeling and graphic simulation system for human motions. Details of the concept of Info-Ergonomics is shown in [2].

2.2 Info-Ergonomics for Factory Workers' Motion Simulation

Info-Ergonomics is one of the research priorities in the integrated computer-aided manufacturing process design from the viewpoint of human factors. Recently modeling the machines in the factory, creating virtual machines by using CG, simulating their work and evaluating it are coming into use gradually [3].

But modeling human beings and creating "virtual employee" by CG still has very limited use, because of the human body's high complexity and the limits imposed by the computer techniques.

However, recently the development of computer technology enabled the modeling of complex machines based on kinematics and kinetics characteristics, analyzing machine's motions and represent virtual machine by means of CG in real time.

We focus on human body modeling along with the machines, and evaluating employee's working in the factory from human-machine co-operation and employee's point of view, so we propose Info-Ergonomics as the framework for CG simulation and evaluation of virtual employees' work.

The purpose of Info-Ergonomics can be defined as follows:

- Pursuing comfortable environment for employees (factory workers).
- Simulation and evaluation of employee's action with simplified human body.
- Modeling of skilled employees and storing their skills into Multimedia Database.

Info-Ergonomics provides the designers with modeling, evaluating, and visualizing tools for designing the optimized working environment for the factory employees. This is an important component of "virtual factory".

As the whole informations are schematized and integrated based on the RWDB concept, the user can extract arbitral part of them and reorganize multimedia data adapting to requirements of application.

The use of CG has significant advantages. For instance, Info-Ergonomics-based simulations allow more precise evaluation than real measurement because virtual employees can be made to do any more (even those causing pain, dangerous ones, etc) and, in addition, we can perform the evaluation itself at a much more detailed level visually.

And then, to replay the skilled employees' work in the Real World Database (RWDB) with CG, and to study from all angles is visual aids for beginners. In addition, we can keep safety and work space for employees, and design comfortable work environment.

3 The human Body/Motion modeling based on the "Mediator" Concept

3.1 Introduction of "Mediator"

In order to construct a human body/motion database, storing human body and motions, which are approximated to capturing data from real world by using 3D scanner and stereo video cameras, and spatiotemporal query method are needed. But surface data (polygons) captured by 3D scanner and spatial points sequence data analyzed from stereo video are very huge data, and are not able to display in real time. As a solution of this problem, we propose the "*Mediator*" concept about human body/structure and motion. In our previous research, we proposed the Shape Mediator[4][5]. In this paper we extended this concept to the human body and structure, and then we propose the human Motion Mediator. Each Mediator will define as follows.

- Shape/Structure Mediator — We suppose a human body as a 3D graphic model that is constructed from "parts"(head,arm,hand,...), which don't change its shape along time, and "joints". Each part is defined as collection of individual polygon surface and texture, and each joint is assigned Degree of Freedom and Range of Motion.
- Motion Mediator — We suppose the human's motions are sequence of postures, consequently we define the human's Motion Mediator as collection of characteristic internal postures, namely, relative time and joint angle(value) of each Shape/Structure Mediator in this time, in which characteristic internal postures of a motion of a human.

Shape/Structure Mediator is made individually. Therefore, Shape/Structure Mediator reflects difference of each human shape and range of motions. On the other hand, Motion Mediator is made for each motion of a human, and represent differences each motion of a human. By using this Mediator concept, you can represent the human body shape, structure and motion in database. Considering Mediator, expression of a characteristic by using minimum data and of common semantics between common data and Mediator.

In next subsection, we discuss about "Hub"(Common Model) and Mediator Creation based on Hub.

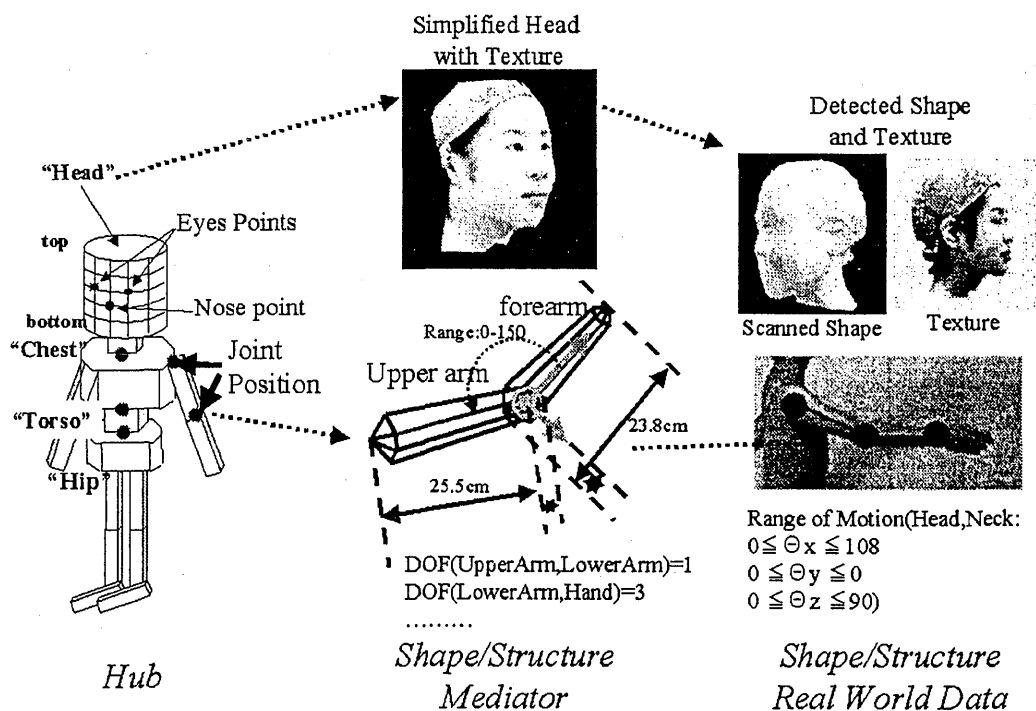


Fig. 1. Hub, Mediator and Real World Data about a shape and structure of the human

3.2 Creation of Mediator from a Hub

A Hub is the information for expressing the essence in common and standardizing about shape and structure of the human body and their motions. Definitions are given as follows.

- Shape and structure Hub — is expressed with the graphic object which has simplified the composition parts of a human body, and sticks the "semantic" information which the object has in the specific position of the surface. Furthermore, each part has "joint position" which defined connection between parts, in order to describe the structure of the whole human body.

- Motion Hub — is expressed with the intermediate posture along which it surely passes in case it appears in common with the motion is performed using the mutual position relation of the human body parts in shape and a structure Hub.

It can be said that the Hub shows the minimum “item” for describing the characteristics about shape, structure and motion of the human. The Degree of Freedom and the Range of Motion of the human are able to measure by carrying out easy motion for the human from a medical viewpoint. As opposed to the Hub, Mediator is data which transformed the Hub to the example a concrete individual exception and by motion, gives value. Next, the technique of creating Mediator from Shape/Structure Hub is follows. The Shape Mediator can be offered by transforming and carrying out fitting of the Hub using the detailed data obtained from the 3-dimensional scanner. Therefore Mediator can be constituted easily. Fig. 1 shows an association of Hub–Mediator–Real World Data about shape and structure. On the other hand, when Motion Mediator is created from a Motion Hub, it is important what intermediate postures are chosen as a Motion Hub. At a present stage, the designer has to design a intermediate posture for every motion. For example, a characteristic intermediate posture sequence which generally raise “Motion which tightens a screw with a drill” in factories can be show as follows.

Tightening a screw using a drill:

I: Starting Motion – The posture in which the person is touching the drill.

II: Motion

II-1: The posture in which the person located the drill at the tip of a screw.

II-2: The posture into which the person finished tightening a screw with a drill.

II-3: Next II-1, repeat sometimes.

III: Ending Motion – The posture to which the person turned the drill.

Generally, it is very difficult to extract human’s posture precisely from videos by using stereo video cameras etc. But, if it restricts to the specific postures specified by the Motion Hub, the analysis of postures (joint angles) can be carry out easily by means of detection of corresponding points of characteristic points on the human body. Therefore, the Motion Mediator can be create from the Motion Hub. Fig.2 shows a relation of Hub–Mediator–Real World data about a motion of tightening a screw using a drill of worker in factories. The posture sequence about Hub, Mediator and Real World Data shown in Fig.2 displays joint value data of motion for convenience as a shape of a posture.

4 Info-Ergonomics Database based on the human body/motion database

4.1 Objectives of Human body/motion Database

Objectives of Human body/motion database are follows.

1. Storing a huge number of information about human motion with their intention.
2. Expression each human motion of same human as different instances in the database.
3. Query information in the database from point of view aggregately and intensively, for example, queries which calculate average value of bend value in motions of 100 people.

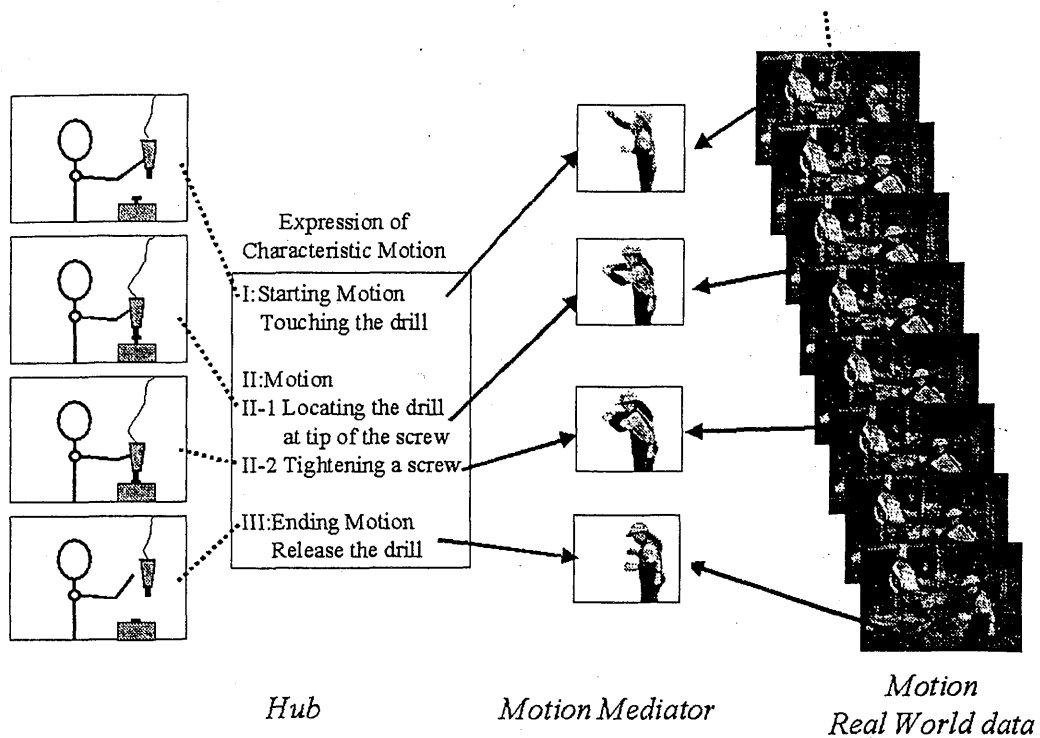


Fig. 2. Hub, Mediator and Real World Data about a working motion of the human

4. Spatiotemporal query based on common points between human body and motions

For the purpose of satisfying these objectives, especially 3. and 4., it is necessary to unify the character of the human body and motion in the database schema. And it is necessary to correspond between “type” and “instance” from the point of view of 2. and to express shape and motion by using a few data as a various type of index from the point of view of 3..

In the earlier our research, about the human body parts, we offered basic common information (Hub: for example, the human head can be approximate cylinder), and *Mediator* which fitted Basic Model to real body shape (for example, surface which captured by 3D Scanner). And we discussed its ability for spatiotemporal query [5]. But for the purpose of storing the human motions, the human structure and motion have to be modeled in database schema.

4.2 Realization of a database of body shape and motion based on Mediator

In order to accumulate motion of the human into a database, as we described previous section, the data belonging to each level of a common model, and a real world model must be able to be accumulated with the reference relation. The kinds of data are not only a character and a numerical value but a stereo image, and various things, such as polygon texture.

Moreover, query to a database is performed, combing the data accumulated freely and ad hoc query is performed according to user’s purpose. Meanings of Mediator

which was proposed in this paper, are not depended on data models for data expression. However, in order to realize the above query, the data model need to have the small unit of the data which defines as the mechanism with the simple reference relation of data, and is accumulated is desirable.

Then in this paper, in order to design human body, structure and motion database, we adopted functional data model "AIS(Associative Information Structure)"[6] as a schema description. Fig.3 shows database schema and its instances of the human body/motion.

For example, corresponding to operation with the real world, motion Mediator is defined, and accumulated as instances, the motion Mediator instance consists of the sequence of a intermediate posture, and the one intermediate posture is defined by the joint angle value for every joint. Thus, a complicated relation can also be directly expressed as relation of an instance level. Moreover, you can query to tracing this path freely. For example, a query, which is "Parts shape of the joint part both ends at the time of being in a certain motion", can be realized by tracing the following route, (1)-(2)-(3)-(4)-(5) in Fig.3.

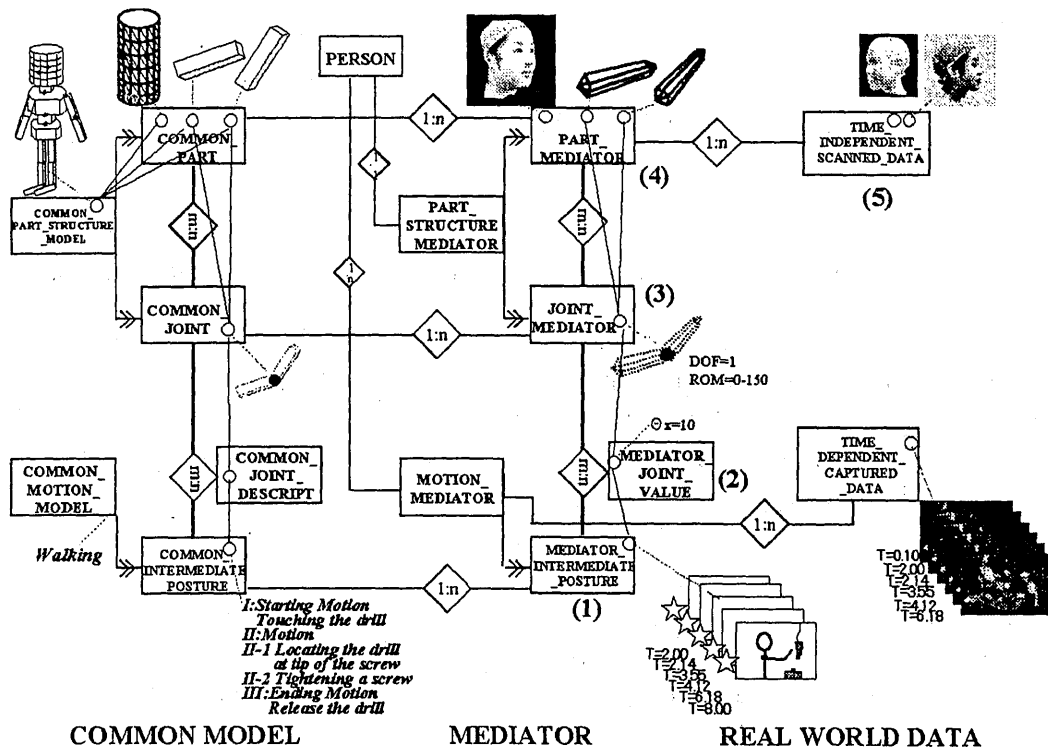


Fig. 3. Database Schema and its instances of the human body/motion

4.3 Consideration about flexible query of motion

In Fig.3, the instances of Hub, Mediator and real world data, which correspond to human shape, structure and motion, are accumulated separately. Moreover, as Shape/Structure

and motion of the human are independent each other, it is possible to combine a human body and motion by query variously. It is realizable not only the combination of motion of the human and the body of the human captured from real world, but to create the virtual motion animation as a query result. In short, 3-dimensional animation "Motion of the human B who imitated motion of the human A" can be create by querying "Motion of the human A" and "The body of the human B" and combining them. Mediator expresses parts shape and structure, and the individual feature of motion and has also described the difference between same parts shape and structure, and motion. Therefore, various query can be performed from Motion Mediator, for example, "An mean time of tightening a screw in motions of worker A", "An average of unit work time" and "Maximum of the distance which extend the worker A's right arm to take a drill".

Moreover, in Fig.3, The structure allows powerful and flexible data retrieval. As a most interesting type of query we consider the reconstructing of a new world using CG data from the database. Generally, CG Data includes shapes, polygons, textures, positions, lights, view points, motions and so on. Thus, 3-Dimensional CG data is huge and extremely complex. In the Info-Ergonomics Database, the query result is provided as a complex data object. And you can get working simulation data as a CG by reconstructing it.

5 Info-Ergonomics Simulation

5.1 Info-Ergonomics Simulation System

We have designed an Info-Ergonomics Simulation System based on RWDB and ENVISION(DELMI Corp.)[7], and are developing now a prototype system. Fig.4 shows the Architecture of Info-Ergonomics Simulation System. Basically working simulation is designed as a part of man-machine coexisting system, and then evaluated. All results can be accumulated into the Multimedia Database together with worker's data. Those data might be referenced by the Manufacturing Process designers when designing a new process or a factory. In short, we propose Info-Ergonomics as an approach supporting the work analysis and work design with knowledge and data.

6 Concluding Remarks

In the present paper we have reported the concept of Info-Ergonomics as a framework for Manufacturing Process Design. It supposes shared data usage between applications by means of *Real World Database* which can incorporate all the data through an integrated data model. Such approach would be beneficial to Manufacturing Process Design. Future work should focus on the detailed design of the *Real World Database System* and the further development of the simulation system.

Acknowledgement

A part of this research is supported by the IMS-HUMACS project Japan. This work is also supported in part by the Grant-in-Aid for Encouragement of Young Scientists from the Ministry of Education, Culture, Sports, Science and Technology. (No.13780198)

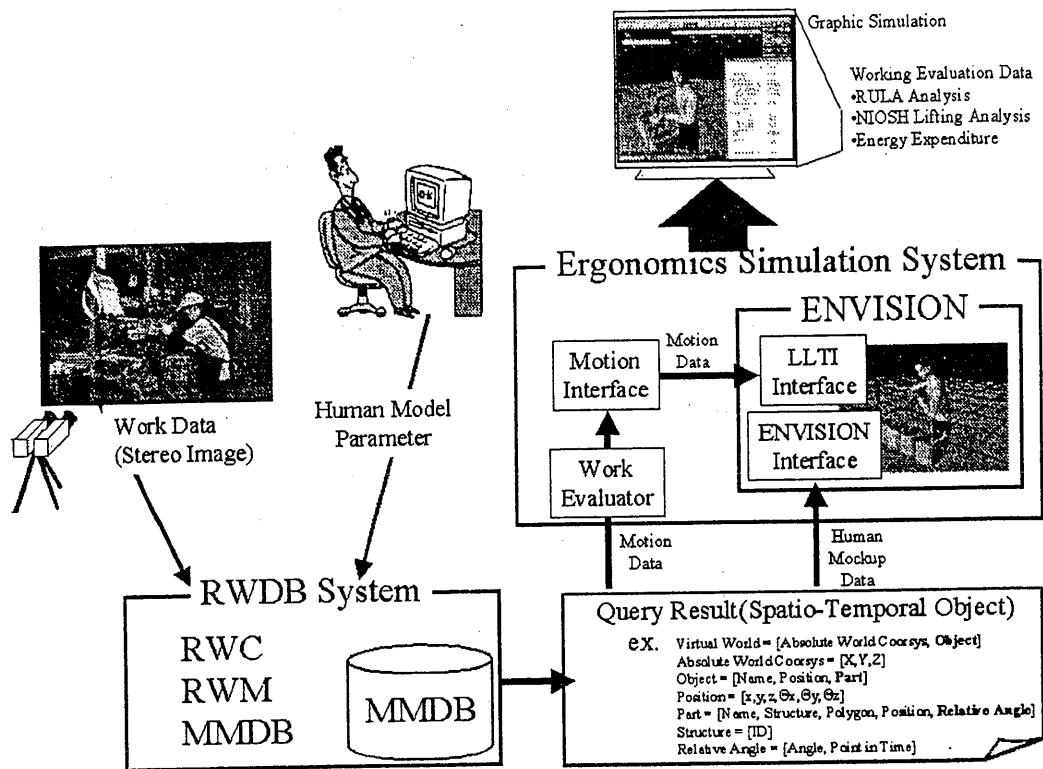


Fig. 4. Architecture of Info-Ergonomics Simulation System

References

- [1] S. Imai, K. Salev, T. Tomii and H. Arisawa: Modelling of Working Processes and Working Simulation based on Info-Ergonomics and Real World Database Concept. Proc. of 1998 Japan-U.S.A Symposium on Flexible Automation(JUSFA'98) (1998) 147-154
- [2] H. Arisawa, T. Sato and T. Tomii: Human-Body Motion Simulation Using Bone-based Human Model and Construction of Motion Database. International Workshop on Conceptual Modeling of Human Organizational Social Aspects of Manufacturing Activities HUMACS 2001(2001) (to appear)
- [3] K. Noro: Illustrated Ergonomics. JIS (1990)
- [4] H. Arisawa and S. Imai: Mediator-based modeling of factory workers and their motions in the framework of Info-Ergonomics. Human Friendly Mechatronics(ICMA 2000), E.Arai, T.Arai and M.Takano Eds., Elsevier Science B.V. (2001) 395-400
- [5] T. Tomii, S. Varga, S. Imai and H. Arisawa: Design of Video Scene Databases with Mapping to Virtual CG Space. Proc. of the 1999 IEEE International Conference on Multimedia Computing & Systems ICMCS'99 (1999) 741-746
- [6] H. Arisawa, T. Tomii, H.Yui, H. Ishikawa: Data Model and Architecture of Multimedia Database for Engineering Applications. IEICE Trans. Inf. & Syst., E78-D,11,(1995) 1362-1368
- [7] <http://www.delmia.com/>
- [8] <http://www.h-anim.org/>
- [9] <http://www.motionanalysis.com/>
- [10] T. Yabe, K. Tanaka: Similarity Retrieval of Human Motion As Multi-stream Time Series Data. International Symposium on Database Applications in Non-Traditional Environments (DANTE'99) (1999)

Human-Body Motion Simulation Using Bone-Based Human Model and Construction of Motion Database

Hiroshi Arisawa¹, Takako Sato² and Takashi Tomii¹

¹ Graduate School of Environment and Information Sciences,
Yokohama National University,
79-7, Tokiwadai, Hodogaya-ku, Kanagawa, Japan

² Graduate School of Engineering, Yokohama National University,
79-5, Tokiwadai, Hodogaya-ku, Kanagawa, Japan

Abstract. This paper presents motion simulation/evaluation system for factory workers in the framework of "Info-Ergonomics." One of the key technologies is CG simulation based on the precise human body mockup called "Bone-Based Human Model." Using BBHM, "real" motions of workers can be mapped for precise simulation. Another important issue is data and knowledge integration. For the purpose of schematizing such data and providing retrieval functions are discuss in an extended database system, "Real World Database."

1 Introduction

Motion Simulation of factory workers is one of the core technologies to achieve optimum Human-Machine Co-existing Environment in modern factory design. IMS - HUMACS Project has been pursuing such core technologies and their principles especially from the viewpoint of "Human" aspects[1] for about last 10 years. Among them the authors group proposed "Info-Ergonomics" as an integration of simulation/evaluation technologies of human workers' motion, in conjunction with database technologies which provides storing/retrieving functions for those knowledge[2]. We think that human centered factors are oftenly overlooked or underestimated although human labor has an important role in the manufacturing process design. It is usually examined from the viewpoint of Industrial Engineering, Ergonomics, etc. and results are used in product planning, working analysis, and work environment design. Info-Ergonomics is a promising solution to provide total environment of work simulation and evaluation. It consists of a number of component technologies like Human body modeling, Customizing human body model to specific workers, Motion description, Capturing motions from video images, Simulation, Storing time-dependent data into databases, and Spatio-temporal retrieval on databases. This paper outlines the Info-Ergonomics concept and technologies, and explains RWDB(Real World Database) as the back-ground technology. Then we will focus on the "precise" human body modeling, called Bone-Based Human Modeling. Total system architecture to realize Info-Ergonomics will be also discussed.

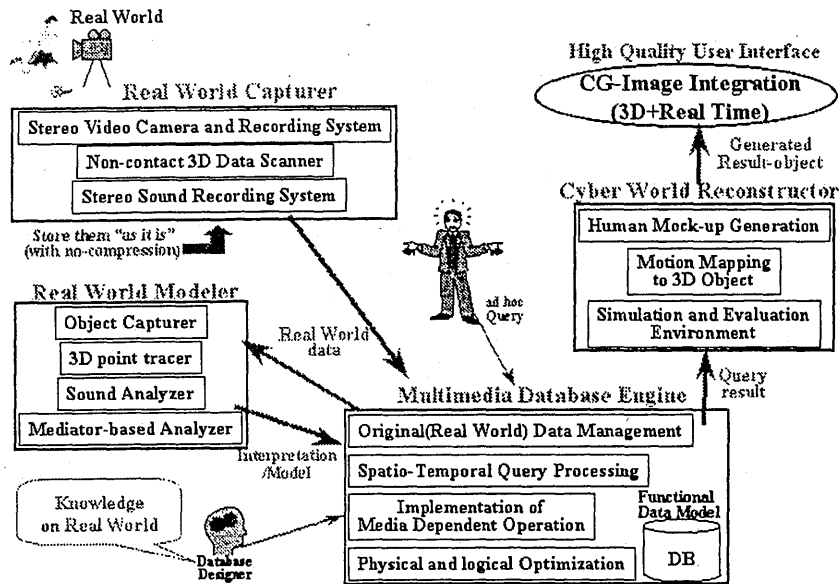


Fig. 1. Conceptual architecture of RWDB

2 Real World Database

To discuss the human body modeling and motion description, we must consider on the technical and the theoretical background of such data processing. For this purpose we proposed "Real World Database (RWDB)" as the Data processing platform. The objective of the RWDB is to provide a total environment for capturing, analyzing, storing, and retrieving physical or logical objects in the real world. Everything in the real world could be modeled through it and any type of data could be accumulated. For this purpose, RWDB must involve at least 4 components listed below. Conceptual architecture of RWDB is shown in Figure.1.

- Real World Capturer (RWC)

The objective of RWC is to capture the external form of objects in the real world. There exists various types of 3D or spatial information depending on capturing devices. The simplest one is a set of video cameras. We can get a sequence of frames from, for example, a left-eye camera and a right one simultaneously. Another type of input device is "3D Scanner" by which we can get a perfect surface (polygon) model for static objects. The practical solution is to get above two kind of informations from the real world and to combine two models into one in the database level.

- Real World Modeler (RWM)

RWM is a set of tools each of which analyzes original frame images and generate a new information. For example, the Outline Chaser[3] catches the outline of an object in a certain video frame, and then trace the outline in preceding and successive frames. Many algorithms are investigated and evaluated for range image

generation in Image Processing area [4]. All the results of analysis are stored into database preserving the correspondences to the original images.

– Multimedia Database (MMDB)

MMDB is a database which treats a variety of data types such as full texts, graphic drawing, bitmap images and image sequences. The features of such data are quite different from the conventional DBMS's ones, because some of them are continuous and might occupy much more space than traditional data types. As to data model, in order to integrate all types of data, introduction of simple primitives to describe real world entities and the associations between them are essential. Moreover, the query language of multimedia data handling must involve various types of media-dependent operations for retrieving and displaying. Especially, in RWDB, the result of a query creates a new "Cyberspace". A query may retrieve a certain unit work and project it to another human worker. The author proposed a total data model to describe the 2D or 3D data, and also presented query language MMQL for flexible retrieval [4].

– Cyber World Reconstructor (CWR)

The result of database consists of various types of data such as frame sequence and 3D graphics data. In order to visualize the query result, RWDB should provide a "player" of result world. CWR is, in this sense, an integrated system of 3D computer graphics and 3D video presentation. Unfortunately, the modeling method of objects in the field of 3D graphics and VR systems are quite different from the DB approach because the former is focusing on natural and smooth presentation of surfaces of objects and their motions, whereas the latter makes deep considerations on semantic aspects of these objects.

3 Info-Ergonomics Modeling

Based on the above RWDB Concept, we concentrate ourselves into modeling on human bodies and motions especially for factory workers. We offered Info-Ergonomics for storing Human Working Data and reconstructing it in CG Simulation.

3.1 Info-Ergonomics

Info-Ergonomics is a framework of information integration for conceptual design of Manufacturing Systems. Recently modeling the machines in the factory, creating virtual machines by using CG, simulating their work and evaluating it are coming into use gradually [5]. But modeling human beings and creating "virtual employee" by CG still has very limited use, because of the human body's high complexity and the limits imposed by the computer techniques. We focus on human body modeling and cooperative works with machines, especially from the viewpoint of comfortability, safety and efficiency.

Info-Ergonomics provides the designers with modeling, evaluating, and visualizing platform to design the optimized working environment for the factory employees. From the viewpoint of IMS-HUMACS Project, Info-Ergonomics is considered as a liaison between traditional simulation/evaluation technique and highly mental aspects. (See Figure 2.)

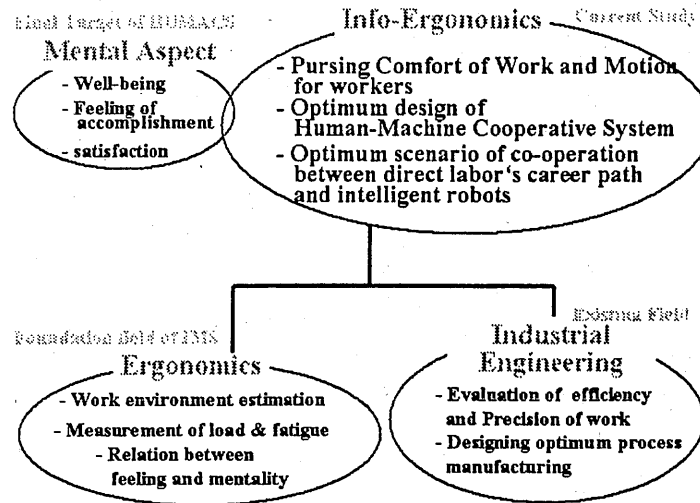


Fig. 2. Info-Ergonomics for Factory Workers

As the whole informations are schematized and accumulated based on the RWE concept, the user can extract arbitral part of them and reconstruct 3D objects and simulation data adapting to requirements of application.

The use of CG has significant advantages. For instance, Info-Ergonomics-based simulations allow more precise evaluation than real measurement because virtual employees can be made to do any more (even those causing pain, dangerous ones, etc) and in addition, we can perform the evaluation itself at a much more detailed level visually.

Replaying the skilled employees' work from Database to CG enables us to study body-motion from all angles intuitively. In addition, we can keep safety and work space for employees, and design comfortable work environment.

3.2 The Human Body and Motion Modeling

In order to store and retrieve Human motion data, we need an integrated datamode which can describe all kind of objects and materials about the human body, for use in all applications. Creating realistic model of the human body requires considerable amount of data because of object's complexity. Therefore from ergonomical point of view we must re-modeling the human body with small number of primitives. We call it "Simplified Human Body", which involves, for example, simplified head, arms, body and legs, connected by small number of joints each other.

The human body and motions can be described as follows.

1. Surface Model

Human body's figure can be described by a set of Polygon and Texture Models of body-parts, that is, each part of human body can be defined as CG surface model. An example of polygon model and texture image are shown in Figure.3(a)(b).

2. Structure Model

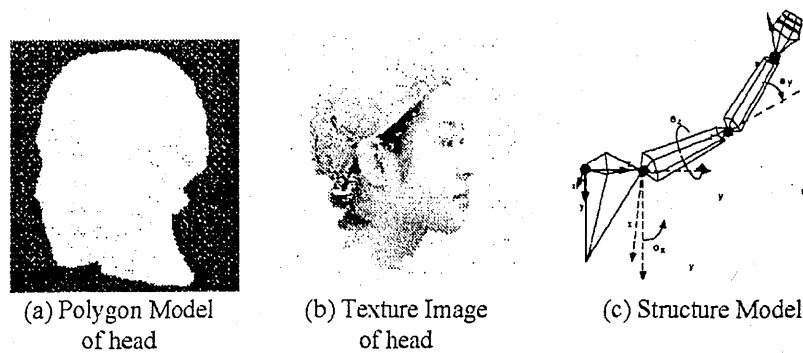


Fig. 3. Human Model

Human body consists of more than 200 bones, which are connected so that the body can move using the power of the muscles [6]. But mechanism is too complex to model it completely. We focus on movable joints and select 24 of them in order to create a simplified model of the human body, Simplified Human Model (SHM). Also we can define much more precise human body model (called BBHM) as discussing later. As for SHM, it can be regarded as hierarchical parts system, that is, each child component (usually a bone) of SHM has its own coordinate system, the origin of which means the connecting point (i.e. joint) to the parent component. An example of SHM connections between arm parts are shown in Figure.3(c).

3. Motion Model

For a certain employee's work, the action of each component of human body should be traced and modeled under the restriction of inter-joint structure. This Dynamic Model is defined on the Simplified Human Model discussed above. Each joint moves along with time. That is, the motion model of each workers (action) can be defined as the time sequence of joint-angles. On the other hand, it seems to be very difficult or impossible to define BBHM movement directly because of its complexity. Converting mechanisms from SHM motion to BBHM one will be discussed in the section 4

4 Simulation with Precise Human Mockup

4.1 Objectives of Precise Human Modeling

As discussed in the section 3, the objective of info-ergonomics is to explore comfort actions, etc. for the purpose of providing a platform that can be applied to all fields where human actions are handled (such as medical service, manufacturing, arts, and sports). Application of info-ergonomics to precise analysis of human beings at work sites such as factories will require a doll ("mockup") that not only has skin and hair, which is externally visible, but also has the internal structure of human body, especially bones, muscles, tendons, and others (flesh, blood vessels, nerves, etc.), together with freedom of joints. This research has therefore proposed and worked out a framework for the shape and structure of a detailed virtual human being that can be actually utilized on

the scene of medical treatment. When medical data is handled in info-ergonomics, there is a desire to visually simulate operations and rehabilitation in advance, using CG, with respect to a large number of humans, reflecting their individual differences. However, since data on shape, structure, and actions of humans in the real world amounts to huge volumes, it is necessary to schematize (represent) such data before simulation in order to create a database of human information. In addition, since individual human beings are different in height and arm length, we need a mockup that allows interactive changes to parameters for individual parts of human body. If the structure and actions of human body can be systematically stored on computer in this manner, it will become possible to analyze lumbago of workers and record movable range of joints and other actions to check progress in treatment.

Various human mockups have been proposed to represent human actions on computer. One of these mockups represents human actions mathematically, another represents each joint as a point, still another draws a line between one joint and another, and yet another mockup expresses each segment (each part of the body) as a rigid body. However, these mockups either have rough contours as a result of reduced amounts of information or fail to reproduce essentially smooth movement of human body because of an extremely limited number of joints. What is more, since these mockups cannot represent internal structure (bones, muscles, tendons, nerves, etc.), they can hardly meet the medical need for data as to which joints are under load or how the movements of bones are. Some of the mockups certainly represent bones and muscles of human body, but since being based on the analysis of only a part of human internal structure they are not instrumental in the understanding of the exterior (skin) of human body. Moreover these mockups do not permit changes to scales of individual parts so as to represent humans of different heights and arm lengths. As a result, a mockup that can meet our demands must be a 3D CG human mockup that has detailed contours, can faithfully reproduce the movement of skeleton, permits interactive editing of different body structures and movements of individual humans, and allows observation from all angles.

To this end, we first created a precision mockup based on human skeleton that can reflect different heights of individual humans and varied ranges of movements of their joints. By changing the angles of joints of this skeleton mockup, we simulated a series of actions of humans at work and practicing sports. Prototype of such precise human body is displayed in Fig. 4.

4.2 Bone-Based Human Model

In order to simulate and evaluate human motions, precise 3D human body model should have internal structure with the rigid bones and joints, and flexible surface structure required. We propose Bone-Based Human Model (BBHM) as one of such models.

BBHM is a 3-dimensional object which consists of the parts representing the bone shape of a human body, and an outer skin expressed by the flexible surface model. It has 155 parts, 76 joints, and 126 degrees of freedom. Parts are connected each other by movable joint which reflect the human body's motion. Almost all human body motion can be expressed precisely except a motion of ribs at the time of a breath etc. Using



Fig. 4. Bone-Based Human Model

such model, we can achieve precise load evaluation taking physique, joints' supple, and so on. into account.

4.3 Modeling and Schematization of Bone-Based Human Mockup

Based on the above discussion, a schema of BBHM, expressing the main point of human body structure is shown in Fig. 5. Also a "real" bone structure corresponding to Fig. 5 is displayed in Fig. 6.

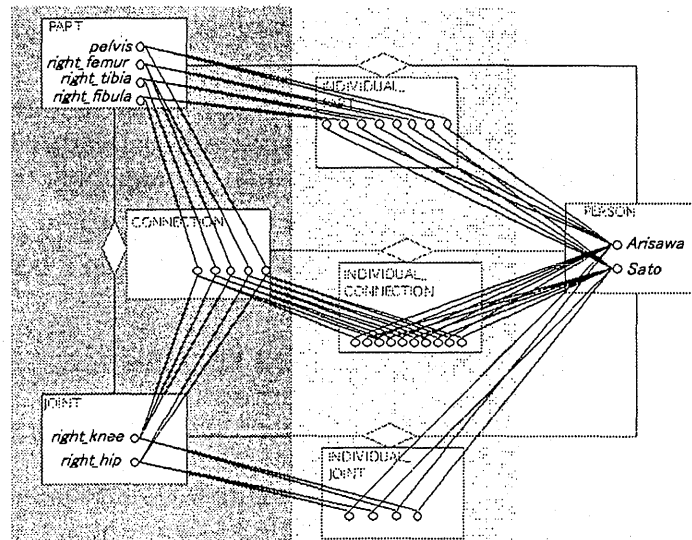


Fig. 5. expression of bone structure on database with instance examples

Bones and joints, which are independent sets of instances, are expressed with PART and JOINT, respectively. And these Cartesian aggregation, CONNECTION, express

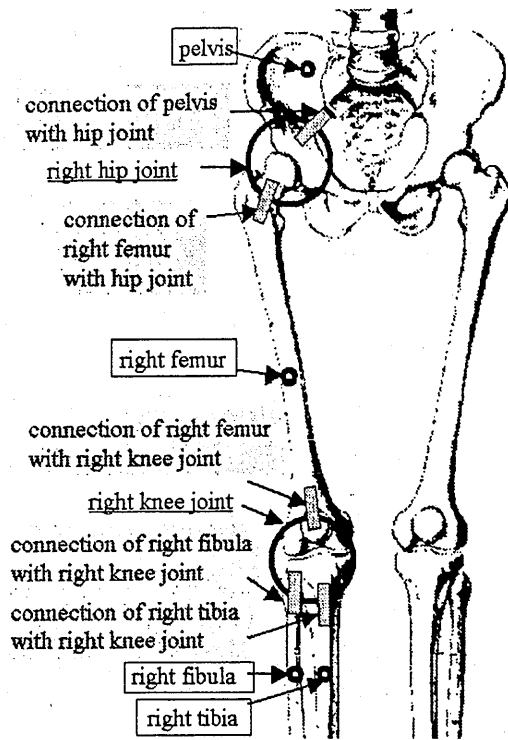


Fig. 6. bone structure of human-body

sets of connection. By relations among PART, JOINT, and CONNECTION, general human body structure can be defined.

For example, right_femur has connections between right_hip_joint or right_knee_joint. Right_knee_joint is constituted by three connections among right_femur, right_fibula and right_tibia.

Now, we can express the person's bone with Cartesian aggregation between human body parts (PART) and PERSON, which type name is INDIVIDUAL_PART. Characteristics of bone, for example, length, are attribute of INDIVIDUAL_PART. JOINT and CONNECTION are considered similarly.

As a result, we can schematize human body structure on a bone level.

4.4 Design and Realization of a Simulation System

To achieve human motion evaluation, we must embody three types of requirements:

1. precise simulations of actual motion in a certain situation
(ex. simulation of rehabilitation clients, etc.)
2. motion mapping from a person's one to another
(ex. an amateur dancer simulates a motion of a professional dancer, etc.)
3. extracting typical motions from a specific group of samples
(ex. workspace design for "aged" workers, etc.)

Based on the concept we proposed in the preceding section, we implemented BBHM simulation system by using engineering simulator ENVISION version 5.0 (der

corp., the U.S.A.[7]). The conceptual architecture and data flow of this system is shown in Fig. 7.

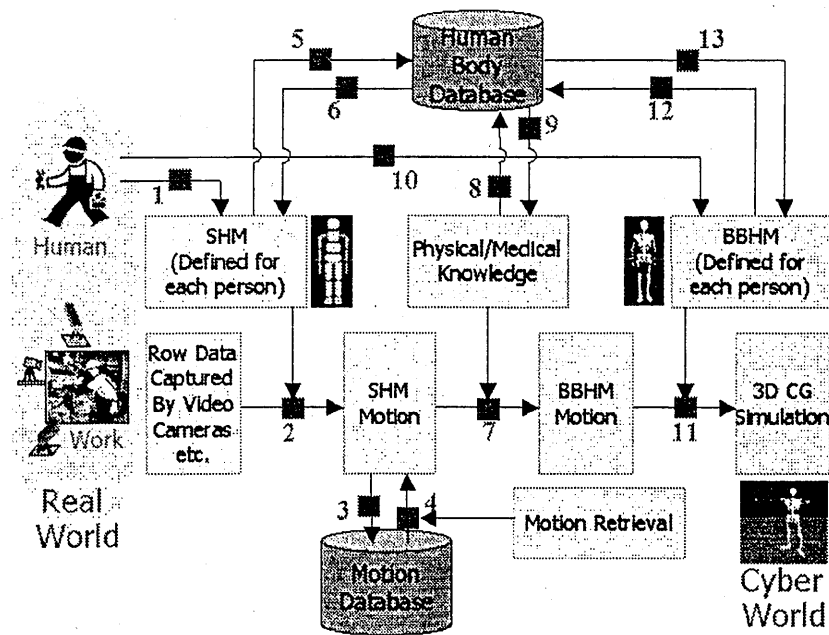


Fig. 7. Spacio-Temporal Analysis and Motion Modeling for Human Body

In this diagram, black rectangles denote “converters” or “modules”:

- 1,10 : measurement module for the human body
- 2 : motion capturer
- 3,5,8,12 : data loader (to database)
- 4,6,9,13 : database query processor
- 7 : mapper
- 11 : graphic simulator

We define human body motions on SHM firstly, then convert them to BBHM motions by using “mapper” and finally execute simulation/evaluation on the BBHM. Because motion data and mapper parameters are strictly schematized, we can various types of motion simulations on various types of human bodies.

We aim to evaluate both actual motion and imaginary motions. So we take two categories of motion definition:

1. obtain “real” motions from video images
2. define motions artificially by operating “3-D Mouse”

Every part's size, ranges of motions, and mapper parameters, are stored in the database, we can perform the simulation united with the individual. The default mapper parameters stand on medical knowledge provided by [8].

4.5 Motion Data Convert System and Motion Database

It is possible to express correctly the load state of each part of the human body by defining various motions to BBHM. But it is difficult to map the human motion to BBHM faithfully, because BBHM has a large number of degree of freedoms. In medicine, the ratio of each joint in a global motion of the human body was investigated. Therefore, we consider it is possible to map the human motion, which was evaluated, to BBHM correctly by using knowledge and technology based on medical knowledge. In our research, we propose "Mapper" algorithm as a method to map of the motion defined using SHM to BBHM by using dignity attachment parameters. The simplest method of calculating dignity attachment parameters from the ratio of the joint value(angle) of SHM and the method of calculating a joint value(angle) from a ratio of the joint value are follows.

$$P_{ij}(\%) = \frac{R_{ij}}{\sum_n R_{nj}} \times 100 \quad (1)$$

$$B_j = S_i \times P_{ij} \quad (2)$$

$$\left\{ \begin{array}{l} B_j : \text{BBHM's } j\text{th joint angle} \\ S_i : \text{SHM's } i\text{th joint angle} \\ P_{ij} : \text{dignity attachment parameter from SHM's } i\text{th joint to BBHM's } j\text{th joint} \\ R_{ij} : \text{rate at which SHM's } i\text{th joint angle effects BBHM's } j\text{th joint angle} \end{array} \right.$$

SHM Motions are stored in motion database. Motion schema is shown in Fig. 8.

4.6 Motion Simulation

As an example of motion evaluation on BBHM simulation system, we compared measured motion data with range of motion of BBHM. For example, we simulated the motion, which range of freedom is restricted by knee's angle (Shown in Figure 9.)

5 Conclusion

In this paper we described a motion analysis and modeling technique based on a precise human model called BBHM. We also described a motion definition technique, SHM and Mapper. As a result, detailed description about human body and various simulation of human body motion are possible with database.

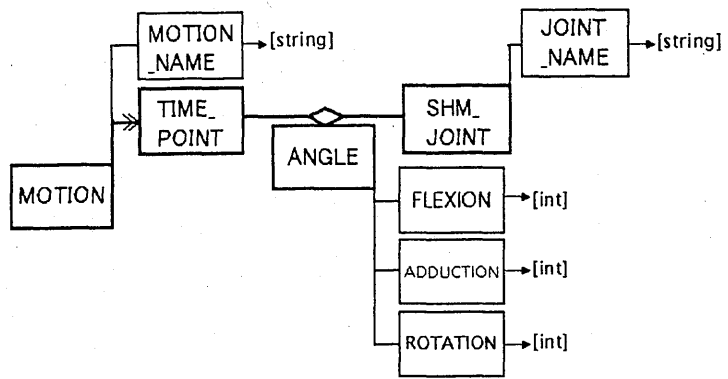


Fig. 8. Motion Schema

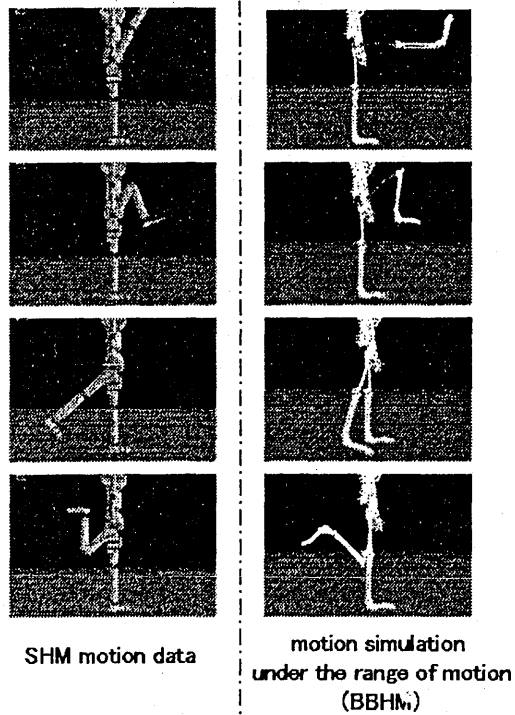


Fig. 9. evaluation of change of range of motion

Many problems are left for future. Implementation of components of human body other than skeleton model, motion analysis based on inverse kinematics and visualization of a load and a amount of fatigue about human body are remained as future problems. By modeling of the other components of the human body such as tendon and muscle, it becomes possible to evaluate human motions more precise and it becomes a lead to understand diseases and factors, which restrict degree of freedom of joint visually. In this paper, we proposed the human body model as rigid model. Modeling flexible surface of the human body is also open for future.

Acknowledgement

A part of this research is supported by the IMS-HUMACS project Japan. Also, work is partly supported by a Grant-in-Aid for Scientific Research (No. 12558026)

References

1. Results of Research on Organizational Aspects of Human-Machine Coexistence Systems(HUMACS) in Manufacturing Systems (Summary Report). Manufacturing Science Technology Center, IMS Promotion Center (2001)
2. Arisawa, H., Imai, S.: Mediator-Based Modeling of Factory Workers and Their Motion in the Framework of Info-Ergonomics Human Friendly Mechatronics (Selected Paper ICMA2000). 395-400 (2001)
3. Michihiko Hayashi, Takashi Tomii, Hiroshi Arisawa: A Form Acquisition of Subjects in Image Databases. IPSJ SIG Notes Vol.97, No.7, 97-DBS-111-13 (1997)
4. H. Arisawa, T. Tomii, H.Yui, H. Ishikawa: Data Model and Architecture of Multimedia Database for Engineering Applications. IEICE Trans. Inf. & Syst. , Vol.E78-D No.11 (November, 1995)
5. Kageyuu Noro: Illustrated Ergonomics. JIS (1990)
6. R. Nakamura, H. Saitoh: Foundmental Kinematics. Ishiyakushuppan Cop. (1995)
7. <http://www.delmia.com/solutions/html/ergonomics.htm>
8. I. A. KAPANDJI: Physiologie Articulaire. Maloine S. A. Editeur (1980)

複数オブジェクト共存空間上の “局面”モデリングとサッカー戦術への適用

坂木 和 則[†] 四十物 裕司^{††} 鬼 東 郷[†]
富井 尚志^{†††} 有澤 博^{†††}

本稿では、実世界などの時空間の中で、特に、複数のオブジェクトが共存する空間中で生じる事象のデータベース化を目的とする。このような時空間では、ある時点あるいは明確に分離されるある時区間における情勢、状況、つまり“局面”を検索できることが非常に重要であることから、“局面”を検索する手法を提案する。

その“局面”検索手法として、現実世界から自動的に取得できる特徴量を局面テンプレートとして蓄積する。検索は、検索テンプレートを用いて大量の局面テンプレートとのマッチングを行い、該当した局面テンプレートを取り出す、テンプレートマッチング型検索を行う。この検索法の妥当性を検証するために、サッカーを具体例として取り上げた。

“Phase” Modeling and Application of Soccer Tactics on Coexisting Objects Space

KAZUNORI SAKAKI,[†] YUJI AIMONO,^{††} GO ONITSUKA,[†]
TAKASHI TOMII^{†††} and HIROSHI ARISAWA^{†††}

We aim at realizing the database which accumulates the behavior in spatiotemporal space, such as real world, especially in space where two or more object coexist. We propose the technique to retrieve “Phase”, because it is very important that a situation or a circumstances, that is “phase”, can be retrieved in spatiotemporal space.

For, “phase” query method, we accumulate quantity of feature as “phase” template which can get automatically from real world. When we want to query, we match a lot of “phase” template for query template, and pick up “phase” template which satisfies the condition. In order to verify the validity of our “query method”, we took up soccer as an example.

1. はじめに

時空間における物体の3次元的な振る舞いを蓄積し、多様な検索を行なうことができるデータベースを時空間データベースと呼ぶ。特に現実世界で起こる様々な事象は、時間的な要素と空間的な要素が複雑に結び付いて成り立っているためそのモデル化は非常に興味深い課題である。

当研究室では、現実世界を様々なマルチメディアデータを用いることでできるだけ忠実に記録し、さまざまな検索をかけた結果を現実感豊かに表現できるデータベースとしてリアルワールドデータベース (RWD-B: Real World DataBase) システムを提案している¹⁾。これが実現されれば、スポーツにおける人体の動きの解析、医学におけるリハビリテーション計画、生産技術分野での作業者と工作機械の協調作業の解析・調査など、幅広い応用分野での利用が考えられる。

ところで一般に、複数オブジェクトが共存する空間における事象・事象に対する人の解釈や検索したいと考える内容は多様であることから、検索者にとって自由度の高い検索である ad-hoc 検索、特に時空間での出来事を用いた検索に応えられる柔軟性、適応性が必要である。現在、複数のオブジェクトが共存する空間から最も簡単に抽出でき、かつ有用なデータとして個々のオブジェクトの位置が挙げられるが、これをそのま

[†] 横浜国立大学大学院 環境情報学府

Graduate School of Environment and Information Sciences, Yokohama National University

^{††} 横浜国立大学大学院 工学研究科 電子情報工学専攻

Division of Electrical and Computer Engineering, Faculty of Engineering, Yokohama National University

^{†††} 横浜国立大学大学院 環境情報研究院

Faculty of Environment and Information Sciences, Yokohama National University

まデータベースに蓄積するだけでは、これらの値をインデックスとした、非常に簡単な検索しか行うことができない。特に、対象とする時空間に複数のオブジェクトが存在するような場合では、物体相互の位置関係やそれらの協調動作を考慮した検索を行うことが必要になるだろう。

そこで本稿では、複数のオブジェクトが共存する空間において、検索者が最も興味のある事柄がいわゆる“局面”であることに注目した。“局面”を情報表現のプリミティブとして使用することにより、同じ“局面”を探しだすという検索が実現できないかと考えた。しかし、ここで問題となるのは、この“局面”の要素の記述法である。というのも、そのプリミティブは現実世界から実際に抽出でき、かつ、それが様々な“局面”を表すことが出来なければならないからである。この問題を解決するために、プリミティブを1枚の局面テンプレートとして記述し、検索に関しては各テンプレートをマッチングさせて行なう、テンプレートマッチング型検索手法を提案する。また、このようなモデル化の有用性を示すためにサッカーの戦術に基づく質問処理に対してその妥当性を検証する。

以下、2章では複数オブジェクトが共存する空間のモデリングについて述べ、3章では、その適用例としてサッカーを取り上げ、サッカーDBにおける“局面”検索手法について述べる。4章では、その実現性について検証し、5章でまとめを述べる。

2. 複数オブジェクト共存空間のモデリング

2.1 オブジェクト空間の特徴

現実世界で起こる事象を3次元時空間に置き換えると、次の2つに分けられる。

- (1) あるオブジェクトの形状や構造の成り立ちに注目し、その動きによって導き出される事象
- (2) 個々のオブジェクトの大局的な動きをもとに導き出される事象

(1)では、いわば個々のオブジェクトの局所的な動作に注目していることが言え、例えば、個々の人間の各部が動くことによって生じる出来事(歩く、投げるなど)が挙げられる。このような出来事に関するモデリング手法として当研究室では、Mediator概念の提案を行ってきた²⁾。他に、単一のオブジェクトの空間的な移動に対する検索手法として軌跡を用いる手法などが挙げられる³⁾。

これに対して(2)では、広い空間におけるそのオブジェクトの動きに注目していると言え、例えば、チームスポーツや囲碁、将棋などで起こる出来事をさす。

この種の事象はオブジェクトが協調作業を行いつつながら時間が進むため、前者とは全く異なるアプローチが必要となり、3次元時空間を考える上で非常に興味深い時空間領域である。

このような空間における研究は特に盛んであり、大きく分けると次のアプローチに分類できる。

- 群集の可視化を行うことにより、あらゆる評価を行う研究⁴⁾⁵⁾
- 群集における協調行動を機械学習やマルチエージェントなどを用いることによりシミュレートする研究⁶⁾⁷⁾
- 群集で行われている事象を検索により取り出す研究⁸⁾

文献^{4)および5)}では、群集における個々のオブジェクトの位置関係を考慮し、そのオブジェクトが確保する領域を考えることにより、評価を行っている。また、文献⁷⁾では個々のオブジェクトをエージェントとしてとらえ、協調行動のシミュレーションなどを行なっている。しかし、これらの研究では、その事例における評価やシミュレーションは行なえるが、過去にあった事例との比較や、似通った行動に対する集約的な見方ができない。さらに文献⁸⁾では、個々のオブジェクトの軌跡を用いて検索することにより、見たいシーンをみることができる。しかし、これでは直観的に検索を行なうことができるが、位置の制約を受けることや、時区間の設定の仕方が難しい。そこで本研究では、“局面”の概念を導入することにより、複数オブジェクトが共存する空間におけるモデリングを行う。

2.2 “局面”概念の導入

我々が観察するリアルワールドの時空間においては、いろいろなオブジェクト—例えば人や車—が、時間の流れに沿って形や位置を変えており、それがまた様々な「状況」や「事象」を創出している。我々がリアルワールドデータベースに対して期待する機能とは、こうした「ある意味を持った状況や事象が起きたある時点(または時区間)のある領域」を検索の結果として取り出し、コンピュータ上に再現することであると考えられる。我々はこれを一般的に“局面”と呼ぶことにし、以下では 実際の局面の記述法と検索方法について議論を進めよう。ところで、このような“局面”の概念は、時空間的なアプリケーションにおいては非常によく使われ、かつ有用である。例えば、ボール競技においてはそのボールの保持者が変わるたびに、全く異なった“局面”に変わり、戦術が組み立てられている。また広い意味で言えば、囲碁や将棋なども各駒が2次元的に配置することにより“局面”を構成し

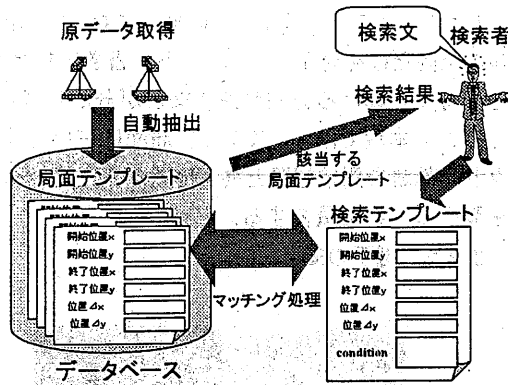


図1 局面検索の概念図

一般	特徴量				
		特徴量2	特徴量3	特徴量4	特徴量5
開始時点 情報	物体1				
	物体2				
	物体3				
終了時点 情報	物体1				
	物体2				
	物体3				
差分情報	物体1				
	物体2				
	物体3				

図2 局面テンプレート例

ている。

この“局面”はいくつかの段階に分けて考えることができる。また、しかもその“局面”はその時点のみを見れば判断できる“局面”と、逆にある時点ごとの一連の流れを見ないと判断できない“局面”がある。そこで、ある1時点で“局面”が決定できるような場面を、1ステップ局面と呼ぶことにすると、始めと終りの2時点で決定できる場面は、2ステップ局面、始めと終りとその中間の3時点で決定できる場面は、3ステップ局面と分けて考えることができる。

以降本稿では、多ステップ局面の基礎となることから、2ステップ局面について考えていく。

2.3 局面テンプレートと局面検索

以上に述べた“局面”の概念を具体的に記述し、またデータベースに蓄積されている“局面”の検索を行なうために、以降では局面テンプレートと検索テンプレートを提案する。その概念を図1に示す。

まず“局面”を記述するために、局面テンプレートを導入する。局面テンプレートには、局面を検索するために必要な特徴量とその差分が書かれている。例えば、2ステップ局面における局面テンプレートでは、局面の開始時点と終了時点における各インスタンスごとの特徴量の値とその差分が書かれる(図2)。

一方、大量の局面テンプレートを検索するために検

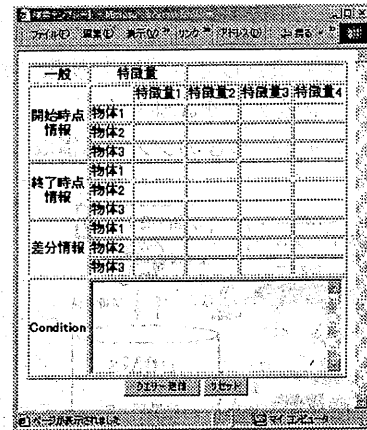


図3 検索テンプレート例

索テンプレートを用いる(図3)。検索者は、この検索テンプレートに検索したい局面の条件をインスタンスごとに書き込む。各インスタンスには定数だけでなく、変数を書き込み、conditionの欄には他の欄で指定した値の関係条件を書くことができる。

しかしこの手法の問題点として、各局面におけるインスタンスすべてに変数や定数を書き込むことは困難であることが挙げられる。そこで検索テンプレートにはすべての項目を埋める必要はなく、書かれない項目はDon't careとして扱うことにすることで、限定したい条件のみを検索テンプレートに書き込むことが可能となる。

検索はこの検索テンプレートに書かれた値と、データベースに蓄積された局面テンプレートをマッチングさせることにより行う。そして検索テンプレートに書かれた条件に該当した局面テンプレートを検索結果として検索者に返す。

これにより時間に依存しない“局面”を対象とした検索ができる。

3. 局面検索のサッカーDBへの適用

前章で提案した局面テンプレートを用いた局面検索が実際に有効であるのかを証明するための1つの方法として、ある決まったアプリケーションフィールドを想定して適用してみることが考えられる。

そのアプリケーション例としては、様々なものが考えられるが、

- 様々な見方により、局面の切り替わりが分かれること
- 同時点に、多数の局面が起こりうること
- その局面がある特定の場所のみ起こるとは限らないこと

などの理由により、局面検索を目的としてデータ

ベース化する対象としては非常に難しいと考えられるサッカーを対象として、“局面”検索における局面テンプレートの有効性を評価する。

3.1 サッカー DB のシステム構成

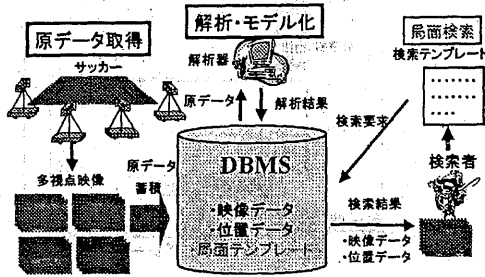


図4 サッカーシステムの概念図

本稿では、局面テンプレートの有効性を示すために、サッカーシステムのプロトタイプを試作した(図4)。

サッカーは主に屋外で行われ、また対象となる空間が広いことから、複数台のカメラを用いて多視点からの映像を取得する。その映像から解析を行うことより、選手やボールの位置を抽出し、映像とともにDBに蓄積する。その位置データをもとに、各局面の時区間の区切りを検出し、局面テンプレートを自動的に作成する。検索者は、検索テンプレートに検索内容を記述し、それをもとに検索することにより当てはまる局面テンプレートを検索結果として表示する。

3.2 局面テンプレートの構成要素

サッカーにおける局面を表す特徴量として最も単純なものは選手やボールの位置情報を逐次蓄積することである。しかし、複数のオブジェクトが共存している空間において、それらの動きを表すには単なる位置データだけでは表しきれない。

複数のオブジェクトが共存している空間において、局面を表す特徴量として必要となるものは、各オブジェクトの位置だけではなく、そのオブジェクト間の相対位置関係やオブジェクトの軌跡などが重要となる⁸⁾。

そこでサッカーにおける局面テンプレートの特徴量として、オブジェクト位置の他にボールの軌跡と選手の支配領域を用いる。

3.2.1 ボールの軌跡

ボール競技においては、ボールを中心に出来事が進み、そのボールの動きにより選手達のとりうる行動が変わってしまう。つまりそのボールの軌跡には、その時点における選手が行おうとしている戦術に関する情報が含まれていると考えることができる。この様なボールの動きに関する特徴量としては、今後のボールの方向や速さなどが特に重要になる。

しかし、そのようなボールの動作に対して、単にボールの位置だけを局面テンプレートに書き込むだけでは、その時点における方向や速さなどが分かりづらい。

そこで本稿では、2ステップ局面においてオブジェクトの軌跡を開始位置と終了位置から正規化ベクトルを算出し、特徴量とする。

3.2.2 選手の支配領域

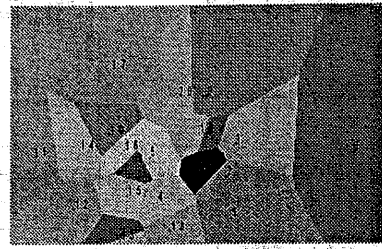


図5 支配領域

複数のオブジェクトが存在する共存空間において、人間は必ず同じ時空間に存在する他の人間との兼ね合いで動いている。特にサッカーの様なチームスポーツでは、各選手はチームの勝利のために相手選手や味方選手の動きに連動して動いている。

このように、オブジェクトはある種の支配領域を持っており、その支配領域の変化がその時空間における様々な情報を反映していると考えられる。

そこでその選手の支配領域の変化を捉えるために、複数のオブジェクトが共存する空間とは対象時空間に存在する複数の選手の支配領域の総和であると捉え、ある時点における空間をそれぞれの選手の支配領域に分割する。

ここで言う選手の支配領域とは誰よりも早く到達できる範囲のことであり、各個人の空間的な位置関係、および、個人の運動能力によって決定される。一般に、数理的に空間中にいくつかの点が与えられたとき、各点の支配領域はポロノイ領域で表現できる。このポロノイ領域を拡張することで、集団で移動する個人の支配領域を定式化する手法が確立されている⁴⁾。そこで本稿では、その手法を用いて選手の支配領域のモデル化を行い、支配領域の例を図5に示した。

この支配領域では、その支配領域の面積がどのように変化するかということが特に重要となる。そこで、この2つの特徴量を蓄積する。

3.3 サッカーにおける局面モデリング

3.3.1 時間軸の区切り方

サッカーの局面テンプレートにおいて、重要となるのが時区間の切り方である。その時区間が撮影された

映像から自動的に抽出でき、かつ、その時区間に検索したい“局面”が表せなければならない。そこで本稿では、その時区間がある選手がボールをもつ時点と、ボールを放す時点で区切ることにした。これは、サッカーにおいて興味のある場面の多くがボールを起点としていることが多いためである。

3.3.2 サッカーにおける局面テンプレート例

局面テンプレートを図6に示す。局面テンプレートには、局面に関するアトリビュート情報が記述されている。この各行は各選手やボールにおける、ある局面の開始時点や終了時点の情報とその差分情報が書かれている。

この局面テンプレートでは、ある試合における1局面を表す必要がある。そのために、その局面テンプレートにはその試合を一意に識別するための値として、試合が行なわれた日、場所、対戦したチーム名、テンプレートの開始時刻、終了時刻が必要となる。

また、サッカーにおいては選手とボールに関する記述を行う。

選手の指定の仕方には様々な指定法が行えることが必要である。そこで各選手ごとにチーム別、選手名、ポジション、チーム内での各選手の順位を蓄積する。ポジションは大きく分けてフォワード(FW)、ミッドフィルダー(MF)、ディフェンダー(DF)、ゴールキーパー(GK)に大別できるためこれを用いる。また、サッカーでは各選手が作るラインの概念が重要となる。そこでチームごとの選手の順位が各チームの自ゴールから近い順番に番号を付け蓄積する。これにより、チーム内における各選手の相対的な位置を基にした指定が行える。

選手の指定を行った後に、それらの条件を指定する。その条件には、単なる位置座標だけではあらゆる局面を表すことは困難である。まずサッカーはボールを扱う球技であるため、その局面においてどの選手がボールを保持しているのかを蓄積しておく必要がある。また、ボールを保持している選手には必ずといってよいほど敵の選手がプレッシャーをかけてパスコースを防ぐ動きを行う。具体的には、敵選手はボールを保持している選手に近づくことによりその後ろへのパスをさせないようにする。この動作は、サッカーの試合の中で選手達が意識的に行う基本動作の1つといえる。従って各局面において、この基本動作を行っている選手を特別に蓄積することにより複雑な局面を表すことが可能になる。また前節で述べた支配領域の大きさの変化を蓄積する。

この他に、ある特定の領域で行なわれている局面を

一般情報		2001/08/21										
場所		横浜国立大学記念体育館										
試合		Norty / Forza										
局面情報		200108210032										
開始時刻		00:00:52:03										
終了時刻		00:00:53:15										
開始と終了の差分時刻		00:00:01:12										
選手に関する情報	チーム	選手ID	選手名	ポジション	順位(自差)	位置座標x	位置座標y	支配領域の大きさ(20cm)	ボール保持	パスを出せる選手	エリアID	
												開始時点
選手に関する情報	A	N_1	竹島直人	DF	2	68	138	1211	x	x	8	
		N_2	佐川直樹	MF	3	113	191	1426	x	x	7	
		N_3	中村洋介	GK	1	74	16	2784	x	x	2	
		N_4	岡本博之	MF	4	47	221	881	x	x	8	
		N_5	川口雄也	FW	5	24	225	328	x	x	8	
	B	F_1	松田一幸	GK	1	69	267	800	x	o	8	
		F_2	中野真由	DF	2	21	247	450	o	x	1	
		F_3	上原崇宏	MF	4	16	150	822	x	x	6	
		F_4	志村和也	DF	3	21	170	488	x	o	8	
		F_5	伊藤直之	FW	5	102	144	1410	x	x	4	
	A	N_1	竹島直人	DF	2	74	136	1612	x	x	7	
		N_2	佐川直樹	MF	3	107	190	1185	x	x	7	
		N_3	中村洋介	GK	1	74	13	2312	x	x	2	
		N_4	岡本博之	MF	4	79	213	981	x	x	8	
		N_5	川口雄也	FW	5	8	234	608	x	x	8	
B	F_1	松田一幸	GK	1	69	267	1910	x	o	4		
	F_2	中野真由	DF	2	2	244	271	o	x	8		
	F_3	上原崇宏	MF	4	29	141	1628	x	x	6		
	F_4	志村和也	DF	3	1	142	458	x	x	6		
	F_5	伊藤直之	FW	5	118	120	1443	x	x	4		
ボールに関する情報	開始時点		終了時点		差分		ボールの軌跡		方向ベクトルx		方向ベクトルy	
	開始時点		終了時点		差分		ボールの軌跡		方向ベクトルx		方向ベクトルy	
	開始時点		終了時点		差分		ボールの軌跡		方向ベクトルx		方向ベクトルy	
	開始時点		終了時点		差分		ボールの軌跡		方向ベクトルx		方向ベクトルy	
	開始時点		終了時点		差分		ボールの軌跡		方向ベクトルx		方向ベクトルy	
	開始時点		終了時点		差分		ボールの軌跡		方向ベクトルx		方向ベクトルy	
	開始時点		終了時点		差分		ボールの軌跡		方向ベクトルx		方向ベクトルy	
	開始時点		終了時点		差分		ボールの軌跡		方向ベクトルx		方向ベクトルy	
	開始時点		終了時点		差分		ボールの軌跡		方向ベクトルx		方向ベクトルy	
	開始時点		終了時点		差分		ボールの軌跡		方向ベクトルx		方向ベクトルy	

図6 サッカーにおける局面テンプレート

表すためには、エリアの概念が必要である。そこでサッカーフィールドを9分割することにより、特定のエリアの指定を行なう。

ボールに関する条件の指定には、局面の開始時点と終了時点における位置情報だけでなく、その方向や軌道が重要となる。そこで、ボールの開始時点と終了時点の位置座標からボールのベクトルを算出し、正規化したものをボールの方向として蓄積する。また、ボール競技においては前節で述べた通り、軌跡が必要である。そこでボールの方向ベクトルと実際のボールの軌跡との最大の差を蓄積する。

これにより検索したい選手・ボールの指定ができ、その条件を自由に行なうことができる。

4. 実現性に関する検討

前章までにおいて提案した局面テンプレートを用いた“局面”検索に関して、最大の問題となるのは次の2点についてである。

- (1) 局面テンプレートの作成法
局面テンプレートはイベントごとに大量に作成されるため、これを手動で作成していく事はほ

ば不可能である。このため、自動的または半自動的にテンプレートが作成される必要がある。

(2) 局面テンプレートの表現能力

その局面テンプレートが検索者の意図した検索をほぼ満たす必要がある。

この2点についての妥当性について、サッカーを例に検証した。

4.1 局面テンプレートの自動作成

自動的に局面テンプレートを作成することの妥当性を検証するために、本研究では実際に室内で行われたフットサルの試合を複数台のカメラを用いて撮影した。

これにより得られた画像群に対して、まず、オブジェクトの領域を抽出する必要がある。オブジェクトの抽出法には、大きく分けると次の2つに分けられる。1つ目の手法はあらかじめ作成したオブジェクトの典型的なピクセルマップを用意し、それと取り込まれた画像とのマッチングを行うことにより、オブジェクト領域を抽出する手法^{9), 10)}である。もう1つの手法が、オブジェクトの色に注目し、その色分布を頼りに画像からオブジェクトの領域を抽出する手法¹¹⁾である。前者のテンプレートを用いる手法では、小さなオブジェクト領域の抽出には良いが、より大きなオブジェクト領域の抽出となると用意しなければならないテンプレートの数が多くなってしまふ。そこで本稿では後者の手法を用いた。これによりオブジェクトの領域が抽出できるため、その領域から位置座標を算出する。

つぎに、これを各時点について行った後に各オブジェクトの追跡を行う。この追跡結果から得られる座標群を基に局面の開始と終了時点をも算出する。

この作業の中で、特に自動化が必要とされる要素技術としては、以下に挙げるものがある。

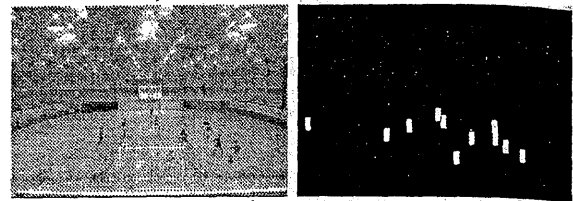
- フィールドにおけるオブジェクトの位置データの自動取得
- オブジェクトの自動追跡
- 局面の開始と終了時点の自動取得

以降、これらの内、最も困難であるオブジェクトの位置データの自動取得法については、実際に抽出システムを作成した。この結果(適用例)を図7に示す。

図7に示すような抽出を映像1000フレームに対し行った結果、再現率79.0%、適合率91.3%であった。これによりオブジェクトの位置データをほぼ自動的に取得できることが検証できたと言える。

4.2 局面テンプレートの表現能力

今回作成したサッカーの局面テンプレートが、検索者の意図した検索をほぼ満たしているのかを検討する。そのためにまず、サッカーの試合中で実際に用いられ



(a)オブジェクト抽出画像 (b)オブジェクト抽出結果

図7 自動抽出結果

図8 スイッチにおける検索テンプレート

ているサッカー用語から、特に選手やボールの動きに関する用語を集めた¹²⁾。この用語を大きく大別すると、パス、シュート、オフェンス、ディフェンス、ドリブルに分けることができる。それぞれの用語について調べた結果、パス23用語、シュート3用語、オフェンス6用語、ディフェンス10用語、ドリブル4用語の全47用語に分類できる。それらの用語で示されている動きが、実際に今回作成したテンプレートで表現できるのかを調べた。ここで、表現できるとは1つの局面テンプレートにその用語の意図するシーンの特徴量が全て表せることを意味する。

この結果を表1~3に示す。ただし、○はその用語の意図する場面を検索テンプレートの変数を変更することにより表現することができる。×は全く表現できないことを表す。

これらの検索テンプレート例として、スイッチにお

ける検索テンプレートを図8に示す。スイッチは、2人の選手が互いに逆方向に交差しながらボールを受け取ることを言う。そこで検索テンプレートでは、まず2人の同じチームの選手をa,bとし、開始時点と終了時点においてボールの保持を交互に切り替える。また開始時点においてパスを出せる候補としてボールを持っていない選手を指定する。そして、互いに交差することの条件として、選手aと選手bの位置差分情報 $(X_a, Y_a), (X_b, Y_b)$ の内積が負であることをConditionに書き込むことで検索できる。

ここで検索の際に用いる特徴量別に分類すると、パス、シュートは主にボールの軌道に関する検索、オフエンス、ディフェンス、ドリブルは主に選手の動きに関する検索といえる。

まず、ボールの軌道に関する検索では、今回の時間区間の切り方がボールに注目したこともあり、表現力が高いといえる。しかし、シュートに関する用語は、シュート自体の定義が曖昧であるため、その用語が示す全ての事例を表すことがもともと不可能であり、低い表現力となった。

また、主に選手の動きに関する検索では、その選手の動きがボールとは無関係に行なわれることがほとんどであるため、ボールの軌道に関する検索よりは低い表現力であった。個々の結果を見るとまず、オフエンスやディフェンスに関する用語では、おおよそそのような守備、または攻撃と言ったものを表す用語が多いため、2ステップの局面では表しにくいという結果となった。しかし、ドリブルに関する用語は、位置や支配領域の差分がはっきりしやすいなどの理由により、テンプレートへの適用が容易であったと言える。

5. まとめ

本稿では、複数の人間が共存する空間を対象とした検索システムの設計を行なった。また、そのような空間においては様々な“局面”に関する検索が重要であることを述べ、“局面”検索に対して有効な手法として局面テンプレートをを用いた検索法を提案した。

この際に、最も問題となるのは、局面テンプレートの自動抽出とその表現能力である。そこで具体例として、フットサルの試合を撮影し、局面テンプレートの自動抽出を行った。また、これによって得られた局面テンプレートの妥当性の検証を行った。この結果により、局面テンプレートをを用いることによって、位置データを蓄積しておくだけでは検索することが困難であるような局面や協調イベントを直観的に検索できることを示した。

謝辞 本研究の一部は、文部科学省科学研究費補助金奨励研究(A)(課題番号12780202)及び、文部科学省科学研究費補助金基盤研究(C)(2)(課題番号12680338)による。

参考文献

- 1) 有澤博: “リアルワールドデータベースとその実現技術1-3”, bit, Vol.28, No.9-11, 1996
- 2) 今井さやか, 富井尚志, 有澤博: “Mediator概念に基づく身体形状・動作のモデル化とデータベース設計”, 情報処理学会論文誌: データベース, Vol.41, No.SIG1(TOD 5), pp.100-108, Feb.2000.
- 3) Martin Erwig, Markus Schneider: “Query-By-Trace: Visual Predicate Specification in Spatio-Temporal Databases”, IFIP 2.6 5th Working Conference on Visual Database Systems, pp.199-218, 2000
- 4) 瀧剛志, 長谷川純一: “チームスポーツにおける集団行動解析のための特徴量とその応用”, 電子情報学会論文誌(D-II) Vol. J81-D-II, No.8 pp.1802-1811, 1998
- 5) 高柳英明, 佐野友紀, 渡辺仁史: “群集交差流動における歩行領域確保に関する研究-歩行領域モデルを用いた解析-”, 日本建築学会計画系論文集, 第549号, 2001年11月
- 6) <http://www.robocup.org/>
- 7) 原章, 長尾智晴: “マルチエージェントの協調のための最適なチーム構成の自動獲得”, 信学技法[人工知能と知識処理], Vol.98, No.296, AI98-36, pp.23-30, 1998
- 8) 矢島史, 中西吉洋, 田中克巳: “動きの直接指定と時間関連指定に基づく移動体映像検索”, 情報処理学会 DBS 研究会技術報告, Vol.2001, No.71 pp.169-176, 2001
- 9) Sunghoon Choi, Yongduck Sea, Hyunwoo Kim, Ki-Sang Hong: “Where are the ball and players?: Soccer Game Analysis with Color-based Tracking and Image Mosaick” Dept. of EE, Pohang University, of Science, and Technology Sam 31 Hyoja Done, Pohang, 790-784, Republic of Korea
- 10) 越後富夫: “視覚的な手がかりによる動画像検索” 情報処理, Vol.41, No.4 pp.351-356, 2000
- 11) 苗村昌秀: “動画像から被写体を精度よく切り出す方法~サッカーシーンへの適用例を中心に~” NHK 技術だより, No.106, pp.8-13, 1998
- 12) http://www.city.kashiwa.chiba.jp/reisoru/yougo/idx_reiyougo.htm
- 13) 四十物裕司, 坂木和則, 鬼束郷, 富井尚志, 有澤博: “時空間MMDBにおけるサッカーの戦術とad-hoc検索”, 情報処理学会 DBS 研究会技術報告, Vol.2001, No.70 pp.71-77, 2001

表1 サッカー用語パス系1

ワンタッチ	ボールを止めないでファーストタッチでパス・シュートをする	○	
アーリクロス	ディフェンスラインとゴールキーパーの間を狙って、浅い位置から早めに入れるクロス	○	
アウトスイング	ゴールから遠ざかる方向にスワープするキック	○	
インスイング	ゴールに近づく方向にスワープするキック	○	
くさび	前方にいる選手の足元への縦パス	×	
クロス	ウイングサイドからゴール前の地域に出すパス	ボールを受けとった領域が2の場合	○
		そうでない場合	×
スイッチ	2人の選手が互いに逆方向に交差しながらボールを受けること	○	
スクエアパス	フィールドを横切るパス	○	
スルーパス	ディフェンダーの間を通してディフェンスラインの裏側へ通り抜ける味方へ通すパス	×	
スワープキック	ボールに回転をかけて曲げるキック	○	
フリック	ボールを少しかすらせ、ボールを後方へ送る技術	○	
ブルバック	ゴール近くまで切れ込んでから後方へ戻すクロス	ボールを受けとった領域が2の場合	○
		そうでない場合	×

表2 サッカー用語パス系2

ポストプレー	前線にいるトップの選手にボールを当てて、そこからさらに攻撃を組み立てていく方法	×
リバースパス	ある方向にドリブルし、それによって逆方向にできたスペースに出すパス	○
ワンツース	味方選手に短くパスを出し、自分は前へと走る。そこへ先ほどの選手からすぐにパスが出てくること	×
パス	ボールを相手に渡すこと	○
壁パス	味方にパスして、またすぐパスを受け取ること	○
サイドチェンジ	攻撃の時、ボールを逆サイドへ動かすこと	○
スローイン	タッチラインからの試合の再開方法	○
ゴールキック	シュートが相手ディフェンダーに触れずにゴール・ラインを割ったときは相手ボールとなり、GKが蹴るキックのこと	○
コーナーキック	ボールがディフェンス(GK含む)の選手に触れてゴールラインから外に出たとき、オフェンス側に与えられるキックのこと	○
バックパス	タッチラインに平行に攻撃側と逆の方向に行われるパス	○
フィード	フォワードに出す縦パス	○

表3 サッカー用語シュート系

シュート	ゴールに向かってボールを打つこと	×
ゴール	ボールがゴールに入ったこと	○
オウンゴール	味方のゴールにボールを入れてしまうこと	○

表4 サッカー用語ドリブル系

ランウイズザボール	ボールを持って移動する技術。少ないタッチでボールを運ぶ	○	
スクリーン	ボールと相手の間に自分の身体を入れて、ボールを奪われないようにする技術	ボールをもらってからはなすまで相手を背負っている場合	○
		そうでない場合	×
ドリブル	ボールを1人で進めていく方法	○	
キープ	ボールを保持していること	○	

表5 サッカー用語オフェンス系

クロスオーバー	2人の選手がお互いに逆方向に走りながら交差する動き	ボールが関わっている場合	○
		そうでない場合	×
ブルアウェイ	ディフェンダーから消える動きのこと		×
オーバーラップ	後ろの選手が前まで出てきて攻撃すること	オーバーラップした瞬間にパスがきた場合	○
		そうでない場合	×
ダイナゴラルラン	コート中央からサイドあるいは、その斜めに走る動き	ボールのやりとりが行なわれる場合	○
		そうでない場合	×
デコイ	ディフェンダーを引きつけるためのおとりの動き		×
トライアングル	ボールを持っている選手を中心に三角形を保ち、サポートできる体制を作ろうとする動き		○

表6 サッカー用語ディフェンス系

チェック	一旦逆方向へ2,3歩ダッシュしてから急に方向を変えてダッシュする動き		×
カバー	味方選手が攻撃者に突破されたときに、他のディフェンダーがすぐ守りに行くこと		×
ジョッキー	ボールを持っている選手に対して、ある方向に追い込むようにプレッシャーをかけること		○
ディフレクティング	ボールが何かに当たってコースを変えること		○
フィスティング	ゴールキーパーがこぶしを使ってボールを大きくはじき返す技術		○
フォアチェック	できるだけ相手のゴールの近くでボールを奪い、攻撃につなげようとする戦術		○
マーク	ボールを奪うために相手のプレーヤーのポジションに合わせて体勢をとること		×
ラインディフェンス	ディフェンダーが相互にチャレンジとカバーとの関係を保ちながら行う戦術		×
マンマーク	決まった選手を守備側が一人でマークすること		×
インターセプト	相手のパスやドリブルを予測しボールを奪うこと	ルーズボールを取った場合	×
		そうでない場合	×

大規模マルチメディアを対象とした スクラップアンドビルド型DBMSの構築

田中俊顕[†] 中家路也^{††} 松田一章^{††} 家富誠敏^{†††} 富井尚志^{††††} 有澤博^{††††}

[†] 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

^{††} 横浜国立大学大学院 工学研究科 電子情報工学専攻

^{†††} 横浜国立大学 エコテクノロジーシステムラボラトリー

^{††††} 横浜国立大学大学院 環境情報研究院

〒 240-8501 横浜市保土ヶ谷区常盤台 79-7

E-mail : {toshi,nakaie,cross,eto,tommy,arisawa}@arislab.dnj.ynu.ac.jp

あらまし

我々の研究室では、フレキシブルなシステムのオンライン拡張性を考慮したDBMSとして、共有メモリを介して協調作業を行うマルチエージェントシステムに基づいたスクラップアンドビルド型DBMSを提案している。本論文では、同システムにおいて、共有メモリに対し複数のエージェントが同時書き込みを行った場合の、効率のよい排他制御を実現する機構を設計し、実機を用いて評価を行った。

Construction of Scrap-and-Build DBMS for Large Scale Multimedia

Toshiaki TANAKA[†], Michiya NAKAIE^{††}, Kazuaki MATSUDA^{††}, Masatoshi IETOMI^{†††},
Takashi TOMII^{††††}, and Hiroshi ARISAWA^{††††}

[†]Graduate School of Environment and Information Sciences, Yokohama National University

^{††}Division of Electrical and Computer Engineering, Faculty of Engineering,
Yokohama National University

^{†††}Ecotechnology System Laboratory, Yokohama National University

^{††††}Faculty of Environment and Information Sciences, Yokohama National University

79-7, Tokiwadai, Hodogaya-ku, Yokohama 240-8501 JAPAN

E-mail : {toshi,nakaie,cross,eto,tommy,arisawa}@arislab.dnj.ynu.ac.jp

Abstract

We propose Scrap-and-Build DBMS based on the multi-agent system which does cooperation work through a shared memory as DBMS in consideration of the on-line extendibility of a flexible system. In this system, the mechanism in which efficient exclusion control when two or more agents perform simultaneous writing to a shared memory was realized was designed, and this paper estimated using the system.

1 はじめに

近年、マルチメディア DB の研究が盛んに行われている。マルチメディア DB の応用分野は様々であり産業、医療、スポーツなどへの幅広い利用が期待される。それらの利用の中には、一度構築したデータベースに対して必要に応じてデータベース及びシステムの規模を拡大しつつ何年何十年と長期間の運用が求められるケースも考えられる。例えば、サッカーや野球などのスポーツでプロリーグ戦の試合を(古いデータを消さずに)数十年にわたって、全てデータベースに蓄積する場合などがそうである。しかし、データベースを長期間運用していくとなると、時代の変化と共にユーザの検索内容が変わっていくことや、時間の経過と共に蓄えられるデータ量も徐々に増え、長い年月をかけて非常に膨大な量になることが考えられる。だが、大規模なマルチメディア DB では、データベース設計の初期段階で将来を見越したシステムのハードウェア構成やスキーマ、検索オペレーションを決めることは難しい。

それゆえ、このような時々刻々変化するデータベースに対しては、その時々々のデータ量、計算機の性能、ユーザのニーズ、マルチメディア情報処理技術に応じてハードウェア構成、スキーマ、検索オペレーションを必要に応じて変更することが重要となる。しかし、ただ変更を行えるだけでは不十分で、実用面を考えれば上で挙げた様々な変更を利用者へのサービスを維持しながら、すなわちシステム全体を停止することなく行なえる必要がある。

この考えに基づき、我々の研究室では、データフォーマットやデータに対するオペレータの変更、ハードウェア構成の変更を柔軟にかつ、DBMS を停止させることなく行うシステムとして、スクラップアンドビルド型 DBMS [4] を提案している。同システムはマルチエージェントシステムをベースとし、すべてのエージェントが共有メモリを介して互いの情報を参照し協調作業を行っていくシステムで、データに対する様々なオペレータをそれぞれ独立したエージェントとして扱い、エージェントの追加、削除を自由に行うことで、システムの柔軟な拡張を可能にしている。

このような複数のエージェントが共有メモリを参照し、協調作業を行う場合、複数のエージェントが同時に同じデータに変更を行う可能性がある。そのような場合どちらかのデータが勝手に上書きされ、無効となるといったことが考えられ、排他制御の機構が重要となる。

そこで本研究では共有メモリにおいて、複数のエージェントによる同時書き込み時の、効率のよい排他制御を実現する機構を提案する。これにより、共有メモリ内のデータの整合性が保たれ、複数エージェントによる協調作業が可能となる。以下、2章でスクラップアンドビルド型 DBMS の概念、3章で同システムの実現手法について述べ、4章で大規模共有メモリにおける排他制御の方法を提案する。5章では、提案する方法を並列計算機上に実装し、評価を行い、最後にまとめを述べる。

2 スクラップアンドビルド型 DBMS の概念

本章ではスクラップアンドビルド型 DBMS の概念について述べる。ソフトウェアアーキテクチャの観点から見た、スクラップアンドビルドの意味は、「システム全体が独立性の高い構成部品の集まりとしてできていた場合、その中のいらぬ要素を自由に外せて、新たに必要になった要素を自由に付け加えることができるもの」とであると言える。この考え方を DBMS に適応したものが我々の提案するスクラップアンドビルド型 DBMS の概念である。概念図を図 1 に示す。

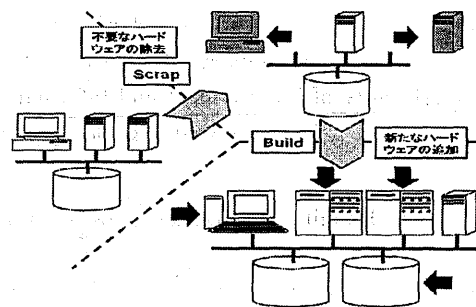


図 1: スクラップアンドビルド型システム

特に、大規模なマルチメディア DB では、長期間の運用の中で変化するデータ量、計算機の性能、ユーザのニーズ、マルチメディア情報処理技術に応じてハードウェア構成、オペレータを柔軟に拡張する必要があることから、スクラップアンドビルド型 DBMS においてスクラップ (削除) 及びビルド (追加) することが必要なものは大きく分けて次の 2 つであると考えられる。

- 計算機資源のスクラップアンドビルド

大規模なマルチメディア DB では、データベース運用開始後も時間の経過と共に蓄積されるデータが増大していくことから、システムのハードウェア構成をデータベース設計の初期段階で決めることは難しく、運用開始後のハードウェア構成の拡張が必要となる。

- オペレータのスクラップアンドビルド

ユーザの新たなニーズに対応するためや、マルチメディア情報処理技術の発展により可能となる新たな検索をサポートするため、必要に応じて既存のオペレータを拡張 (追加、削除) することが必要となる。

また、以上を行う場合、システムのオンライン拡張性を保持する必要がある。オンライン拡張性とは「部分的な拡張、修正、追加があったとしても、他の部分の稼働を停止することなく、かつ、それらを変更しなくてよいこと」[1]と定義されている。データベース運用中のシステムの変更や修正は、システム全体を停止することなく行うことが望ましく、スクラップやビルドをシステム全体を停止することなく行うことができれば、データベースの運用方針やシステムの変更を自由に行うことが可能になると考えられる。

関連研究として、拡張可能 DBMS である、Earth[3] や COMMON[2] が挙げられる。COMMON では、ダイナミックリンクライブラリを利用した動的な機能拡張を行っているが、ハードウェア構成の拡張にまでは触れていない。また、Earth では、システム構成の異なる複数の拡張可能 DBMS を用意し、個々のシステム・アーキテクチャに適応させている。しかし、運用中にシステム・アーキテクチャが変更される場合の対応については考えられてはいない。

これに対し、スクラップアンドビルド型 DBMS では、DBMS の機能のみを拡張するのではなく、システムの各ハードウェア構成要素までも、柔軟な拡張を行うことを可能にしている。これにより、より多様な DBMS の拡張が可能となると考えられる。

3 スクラップアンドビルド型 DBMS の実現手法

2章で述べたような自由な拡張を実現するため、マルチエージェントシステム [5] をベースとした DBMS を構成した。本システムのアーキテクチャを図 2 に示す。検索の処理は、まずユーザからの検索文をコ

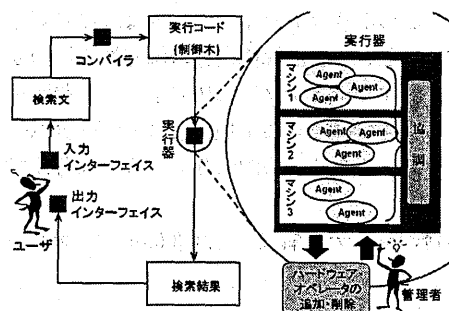


図 2: スクラップアンドビルド型 DBMS のアーキテクチャ

ンパイルして実行コードに落とし、それを実行器のプリミティブであるエージェントが協調しながら検索結果を生成する。そして最後に最終検索結果をユーザに渡すという流れである。実行器によって実行される個々の演算は、当然 DBMS がよって立つところのデータモデルやデータベースの物理構造、そして提供すべき ADT オペレータ等によって異なるが、ここでは関係 DB などによく用いられ、かつ我々が開発中の DBMS プロトタイプとの相性も良いことから、「集合 (群) へのポインタを引数としてもらい、数値演算や集約演算、画像の加工などの演算を行い、集合を変更し、結果についての値を返す」という形の操作が同時並行的に実行されることを仮定している。実行器を構成するエージェントは、大きく、DB アクセスエージェント、メディア依存処理エージェントに分けられ、それぞれに様々

なエージェントが存在する。DB アクセスエージェントは、データベースにアクセスするエージェントで、R-tree やハッシュなど、参照方法の違いにより複数のエージェントが存在する。メディア依存処理エージェントは、各種メディア依存処理を行うエージェントで四則演算、画像の合成のエージェントなどが挙げられる。

これより、我々が提案、実装するスクラップアンドビルド型 DBMS の構造と演算処理の実現手法について詳細に述べていく。マルチメディア DBMS の実現手法にはいろいろな方式が考えられるが、ここではエンティティ(あるいはオブジェクト) 集合と、集合間の構造・リンク情報のすべてをメモリ空間上におき、メディアデータのみをディスクなどの外部記憶上に置く、いわゆるオンメモリデータベースを想定する。このようなアーキテクチャは現在必ずしも一般的ではないが、計算機資源の価格性能比の変化によって、マルチメディア DB のように、参照構造と原データを分離できる中規模以下のデータベースにあっては、むしろ一般的になると我々は考えている。

この場合のシステム構成図を図 3 に示す。図 3 の

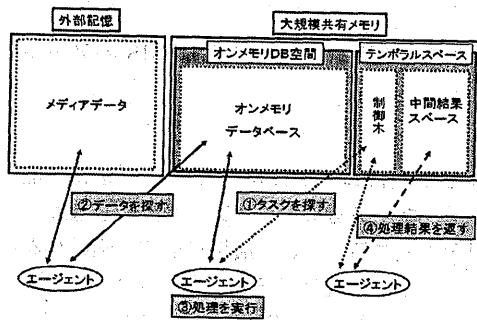


図 3: スクラップアンドビルド型 DBMS の構成

ように本システムでは、大規模共有メモリ上にオンメモリ DB 空間と、エージェントが一時的に作業を行うスペースであるテンポラルスペースを構築し、そこを各エージェントが参照することで、大規模共有メモリを介した協調作業を行う。以下、システムの各構成要素、エージェントについて詳しく述べる。

3.1 大規模共有メモリ

大規模共有メモリは、大きくオンメモリ DB 空間と、エージェントのテンポラルスペースである制御木及び中間結果スペースから構成される。それぞれの空間の役割を以下に述べる。

3.1.1 オンメモリ DB 空間

本システムではメディアデータを除くデータベースを大規模共有メモリ上にオンメモリで構築している。これがオンメモリ DB 空間である。エージェントのデータベースへのアクセスはこのオンメモリ DB 空間に対して行い、必要に応じて外部記憶からメディアデータを取り出す。

3.1.2 制御木

ここでは、マルチメディア DB に対する検索をストリームデータ(映像、音声など)に対するストリーム処理と考え、検索手順をストリーム処理の流れを示す有向グラフで記述したものを制御木と呼ぶ。制御木の例を図 4 に示す。図 4 では、小さい四

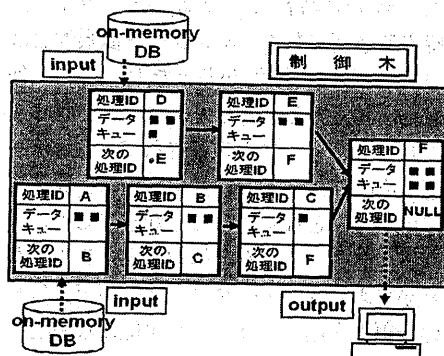


図 4: 制御木の例

角が、処理されるデータのデータ ID とそのデータの処理結果に付加される ID(返り値 ID) の組を、各ノードがそれぞれ処理の種類を、矢印が処理の流れを表している。各ノードは、自分の処理 ID と、次にデータが流れる先の処理 ID(次処理 ID)、処理対象のデータの情報を管理するためのデータキューを保持している。

各ノード間にはデータ ID のみが流れ、それぞれのノードでは、流れてきたデータ ID をデータキューに登録する際、そのデータ ID に対する返り値 ID を一意に決め、データ ID と返り値 ID を組みにし、エントリとしてデータキューに登録する。また、エージェントの制御木に対するアクセスは処理 ID を用いて行われる。

3.1.3 中間結果スペース

中間結果スペースとは、検索処理の途中で生じる中間結果が書かれる空間である。エージェントは、処理結果を返り値 ID と共に中間結果スペースに書き込み、ここに書かれたデータにはデータ ID を用いてアクセスする。なお、中間結果スペースと制御木は、検索の度に生まれる一時的なものである。

3.2 エージェント

検索時におけるエージェントの処理の流れを示す。

1. エージェントは、処理 ID をもとに制御木を参照し、該当するノードのデータキューにキューがあるか確認する。(read アクセス)
2. あれば、先頭のキューに対する処理を行うために予約を入れ、予約が認められればキューを取り出す。(write アクセス)
3. データ ID を指定することで、対象となるデータをオンメモリ DB 空間又は中間結果スペースから読み出し、処理を行う。
4. 処理が終われば、処理結果を返り値 ID と共に中間結果スペースに書き込む。
5. また、制御木に対しては、次処理 ID を持つノードにアクセスし、返り値 ID を渡す。(write アクセス)

以上の繰り返しで検索処理が行われる。このうち、1,2,5 は制御木に対するアクセスである。

3.3 並列計算機でのシステム構築

エージェントが大規模共有メモリを介した協調作業を行うためには、大規模共有メモリに対する高速

な読み書きのアクセスが可能であることが求められる。

一方、並列計算機の中にはこのような共有メモリ機構をハードウェアレベルでサポートしているものもある。並列計算機は、必ずしも計算機資源の拡張性に優れているとは言い難いが、将来ネットワークが高速化されれば、計算機資源の拡張を柔軟に行える、ネットワーク分散したコンピュータクラスタにも容易に移行できることから、今回、本システムを実現するものとして並列計算機上でシステム構築を行った。並列計算機上でのシステム構成を図 5 に示す。図 5 のように並列計算機の各プロセッサをまたいだ形で大規模共有メモリが存在し、各プロセッサにエージェントが分散して配置されている。

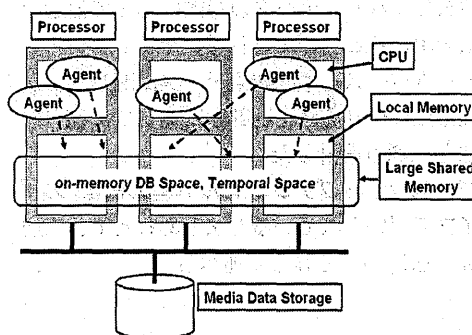


図 5: 並列計算機上でのシステムの構成

本章で述べてきた考え方で DBMS を構築するにあたって、議論すべき点は多々あるが、中でも「複数プロセッサにまたがる大規模共有メモリと複数プロセッサ上に存在するエージェントからの対等で効率の良いアクセスをどう実現したらよいか」は重要な課題である。そこで本論文では、以下この点に絞り、大規模共有メモリの構成技法や並列計算機による評価について述べていく。

4 大規模共有メモリにおける排他制御

3.1 節で述べたように、制御木と中間結果スペース、オンメモリ DB 空間では、その役割の違いから参照の頻度やそこに書き込まれるデータ量に違いが

ある。特に制御木に関してはその特徴として、他の二つと比べデータ量が圧倒的に小さいことや多くのエージェントから参照され、頻繁に更新が繰り返されることが挙げられる。

その際同時に複数のエージェントが制御木の同一の部分に変更を加えることも考えられ、特に制御木のような頻繁に更新されるデータの場合、同時に変更要求が起きたときの排他制御機構が重要となる。

4.1 制御木の管理方法

制御木の排他制御を効率良く行うことができれば、たとえハードウェア的にサポートされる共有メモリを用いたとしても、物理的に異なるプロセッサのメモリ上に、制御木を無秩序に分散させ管理することは望ましくなく、データの特徴に応じて最適な管理方法を用いることが望まれる。そこで本システムではデータベースと中間結果データに関しては通常の管理方法に任せるが、制御木に関しては独自の管理方法を二通り設計した。

以下で、スクラップアンドビルド型 DBMS のプロトタイプで用いられている方法とあわせて、三通りの管理方法について詳しく述べる。

4.2 server-client type

第一の方法が server-client type である。server-client type の概念図を図 6 に示す。この方法は、た

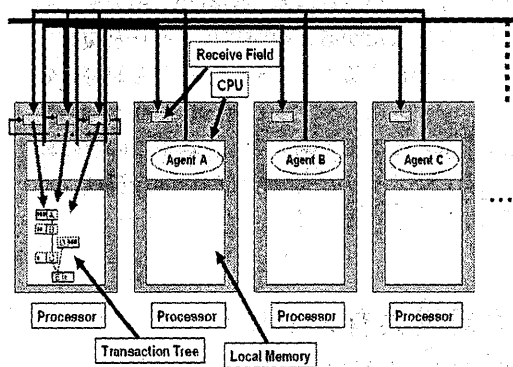


図 6: server-client type

だ一つ server となるプロセッサのローカルメモリ上で制御木を管理し、エージェントの制御木に対するアクセスは、server との通信を通して行うというものである。これは、スクラップアンドビルド型 DBMS のプロトタイプでも用いられている方法で、この場合、制御木は一ヶ所で一括管理するため、制御木に対する排他制御を容易に行うことができる。また、エージェントの把握もこの server のみが行うため、client 側のプロセッサでは、エージェントのオペレータ処理のみを行えばよい。

しかし、server 側で制御木やエージェントの管理を行うことや、制御木にアクセスしようとする client からの通信が集中することから、server となるプロセッサに負荷が集中し、制御木に対する処理に時間がかかってしまうことが問題となる。また、何らかの障害により、server 機能がダウンしてしまうと、システム全体が機能しなくなるといった問題がある。

4.3 server-client type 改良型

4.1 節で述べた管理方法を改良し、処理時間の短縮を図ったものが、server-client type 改良型である。server-client type 改良型の概念図を図 7 に示す。こ

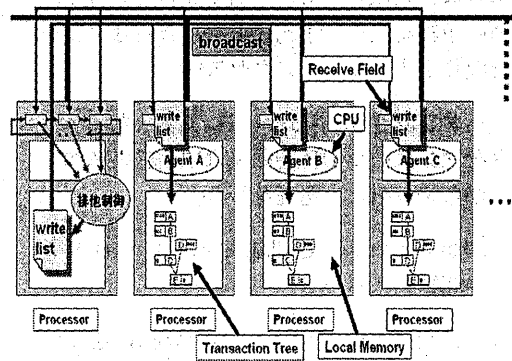


図 7: server-client type 改良型

の方法の、server-client type との大きな違いは、制御木を client 側の各プロセッサで、一つずつ保持・管理することである。

クライアント側で制御木を保持することにより、制御木に対する read 処理は、それぞれのローカル

メモリにある制御木を参照すればよく、処理時間の大幅な短縮が見込める。一方、write 処理は、各プロセッサ上にある制御木の、プロセス間の整合性を保つため、まず各エージェントは書き込み命令を一度 server に送り、そこで同一データに対する書き込み命令があった場合、命令の到着順による排他制御を行い、有効な書き込み命令のみを server が client にブロードキャストする。そして各 client で、送られてきた書き込み命令を制御木に対して行うという方法をとる。この方法により、各プロセッサにある複数の制御木を同一の状態に保つことができる。

この改良により、read、write 処理時間の大幅な短縮が見込めるが、依然 server 機能の停止が、システム全体の停止につながるという問題は残る。また、各 client に制御木が置かれる為、4.1 節の方法と比べ、制御木の管理に使用されるメモリの領域が何倍にもなってしまうことや、client 側の負荷の増加が欠点として挙げられる。

4.4 broadcast type

4.2, 4.3 節で述べた方法に比べ、各プロセッサの自立性の高い管理方式が broadcast type である。broadcast type の概念図を図 8 に示す。この方法

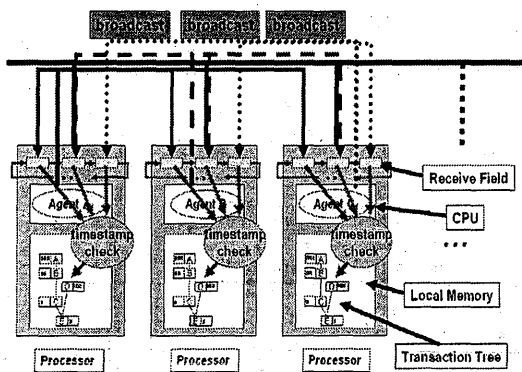


図 8: broadcast type

は、4.2、4.3 節の方法と異なり、全てのプロセッサが対等の関係にある。各プロセッサはそれぞれのローカルメモリに制御木を持ち、管理する。read アクセスは自メモリの制御木に対して行い、write ア

クセスは全てのプロセッサにブロードキャストを行うことで実現する。また、server となるプロセッサが存在しないため、排他制御やエージェントの管理もそれぞれのプロセッサで行う必要があるが、プロセッサごとに write 処理命令の到着順が異なる可能性があるため、到着順による排他制御では、各制御木を同一の状態に保持できない。そこで、broadcast type では、タイムスタンプを用い、同一データへの書き換え命令があった場合、それぞれの命令のタイムスタンプをチェックし、タイムスタンプの小さい命令を有効にするといった排他制御を行い、各制御木の同期をとる。

新しくエージェントを登録する際は、エージェントの ID、処理名、引数と戻り値のデータ型といった情報を全てのプロセッサにブロードキャストし、到着した情報を各プロセッサ上で管理する。検索文をコンパイルするエージェントは、自メモリ上に管理されるエージェントの情報を参照し、コンパイルの際に、検索に必要なエージェントが全て揃っているかチェックする。

4.2、4.3 節の方法では、server 機能を持つプロセッサが落ちてしまうことが、システム全体の停止につながっていたが、broadcast type では、たとえ一つのプロセッサが落ちたとしても、そのプロセッサで動いていたエージェントの機能が停止するだけで、システム全体に与える影響は小さい。

broadcast type は、複雑な排他制御を行うことや、ブロードキャスト通信が同時に多発した場合の通信のオーバーヘッドを考えると、4.2 節と比べ、制御木に対する処理に時間がかかることが予測される。しかし、システムの信頼性においては、三つの方法の中で最も優れていることから、スクラップアンドビルド型 DBMS における制御木の管理方法としては、broadcast type が最も適していると考えられる。そこで、broadcast type の妥当性を検証するため次章で評価を行う。

5 評価

4 章で述べた制御木の管理方法をそれぞれ、並列計算機アーキテクチャ上で構築し、制御木に対する read 処理にかかる時間、write 処理にかかる時間を測定し、性能評価を行った。

なお、今回の測定に用いた並列計算機は、HITACHI SR2201(プロセッサ数32、メモリ512MB×9、256MB×23、計10.5GB)で、OSはHI-UX/MPP、通信方式はリモートDMAである。

5.1 実験結果

各実験はそれぞれの処理を10回ずつ測定し、その処理にかかった最小時間と最大時間を除き、平均をとったものである。また、writeの命令を4種類用意し、同一データに対する同時書き込みが、頻繁に起こる環境を設定した。なお、実験結果の図の横軸“同時アクセス数”は同じタイミングで起こる制御木に対する処理要求の数を表し、縦軸“time[us]”は同時に起こる処理要求を全て行うためにかかる時間[us]を表している。

5.1.1 制御木に対する read 処理時間

各管理方法の read 処理時間を図9に示す。図9か

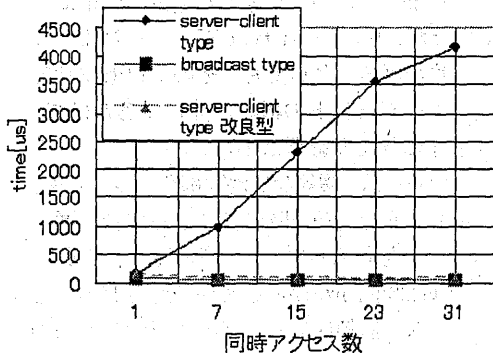


図9: 同時アクセス数と read 処理時間の関係

ら、server-client type の read 処理時間は同時アクセス数に比例して増加していることがわかる。これは、read 処理の際も、server との通信を行うことや server 側で全ての read 処理を行うことが要因として挙げられる。一方、4.3、4.4節で示した方法のように、自分のローカルメモリにアクセスする場合、同時アクセス数によらず、また他のプロセッサとの通信を行わないことから、高速にアクセスできるこ

とがわかる。

5.1.2 制御木に対する write 処理時間

制御木に対する write 処理時間は、大きく3つに分類できる。

- プロセス間通信にかかる時間
- 排他制御にかかる時間
- 制御木にアクセスして書き込む時間

各管理方法の全 write 処理時間を図10に示す。また、write 処理時間の内訳のうち、同時アクセス数と通信時間の関係を図11に、同時アクセス数と排他制御にかかる時間の関係を図12に示す。

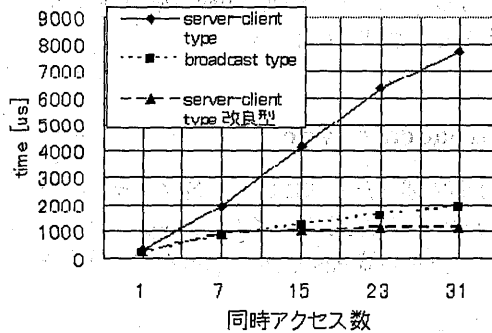


図10: 同時アクセス数と write 処理時間の関係

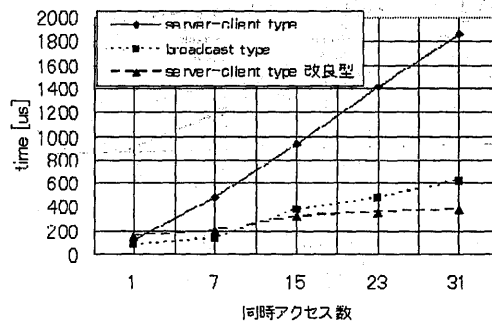


図11: 同時アクセス数と通信時間の関係

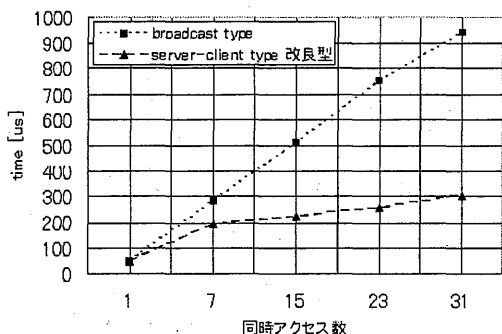


図 12: 同時アクセス数と排他制御時間の関係

図 10 から、他の二つと比べ信頼性の高い broadcast type は、write 処理による性能評価において、同時アクセス数が 7 以下で最も、また 15 以下でも、server-client type 改良型と同程度に、処理時間が短いことがわかる。

図 11 から、同時アクセス数が 7 以下では、broadcast type が最もプロセス間通信によるオーバーヘッドが小さいということがわかる。また、同時アクセス数が増加してブロードキャストが同時に多数発生する場合でも、server-client type 改良型とそれほど差が開いていないことがわかる。

排他制御の方法の違いによる両者の差が図 12 よりわかる。今回は、同一データに対する同時書き込みが頻繁に起こる環境を設定したので、同時アクセス数が増加するほど両者の差が開いているが、同一データに対しての書き込みがそれほど起こらない環境では、タイムスタンプまでチェックする必要がないため、同時アクセス数が増加しても、両者の差はそれほど開かないと予想される。

6 まとめ

本稿では、大規模マルチメディアデータを対象としたスクラップアンドビルド型 DBMS における大規模共有メモリの実現手法の提案を行った。特に、制御木の管理方法に重点を置き、二つの管理方法を設計し、従来までの方法とあわせて三つの方法を並列計算機上に構築し、実験による比較・評価を行った。

4 章で述べた管理方法の中で最も信頼性の高い

broadcast type をスクラップアンドビルド型 DBMS に適した管理方法として挙げたが、write 処理についても同時アクセス数が 7 以下において、最も処理時間が短いことや、同時アクセス数が増加しても server-client type 改良型と大きく差が開かないことから、一定の実現性、妥当性を示したといえる。

今回は、制御木に対する処理時間について、同時アクセス数の関係を検証したが、処理時間には、同時にアクセスされる命令の中にどれくらいの割合で、同一データに対する命令があるかということも、関係してくると予測されるので、今後検証していきたい。

参考文献

- [1] 森欣司: “自立分散システム [I]~[V・完]”, 電子情報通信学会誌, vol.84, No.6, pp.403-408, No.7, pp.484-490, No.8, pp.611-617, No.9, pp.663-669, No.10, pp.734-740, 2001.
- [2] 増永浩二、宝珍輝尚、都司達夫: “拡張可能 DBMS における部品の管理と呼び出しの一方法”, 情報処理学会論文誌, Vol.40, No.SIG6(TOD3), pp.152-161, 1999.
- [3] 早田宏、渡辺美樹、田中圭、山崎伸宏: “オブジェクト指向データベース・エンジン Erath: キャッシュ共有のための設計空間”, 情報処理学会論文誌, Vol.39, No.7, pp.2240-2249, 1998.
- [4] 河信司、市間健太郎、高橋賢一郎、家富誠敏、有澤博: “スクラップアンドビルド型マルチメディア DBMS の構成方式”, 情報処理学会研究報告, 情報処理学会, Vol.98-DBS-116(2), pp.273-280, 1998.
- [5] 村上国男: “マルチエージェントシステムとその応用”, 電子情報通信学会誌, vol.78, No.6, pp.570-577, 1995.

文章セグメントの単一化による多文書自動要約

浅野 秀胤† 田村 直良††

† 横浜国立大学 大学院 工学研究科 電子情報工学専攻

†† 横浜国立大学 大学院 環境情報研究院

{nocchi,tam}@tamlab.eis.ynu.ac.jp

1 はじめに

本研究では複数の文章を一文章にまとめることを目標とし、文章の事件展開構造化、内容が重複する文の単一化、パス選択による要約文生成からなる要約モデルを提唱し、論じる。

電子化された文書が爆発的に増加している中で、必要な情報を素早く、効率的に入手するために自然言語処理技術に対する要求が高まっている。特に新聞記事やニュース記事のような出来事について書かれた文章に注目してみると、一つの出来事に関する記述が複数の文章にまたがることが多い。また、そのような場合には単に出来事の続きが記述されているだけでなく、視点を変えた文章なども存在する。さらに、複数の文書にわたった新聞記事はまとめて読まれることは考慮されていないので、内容が重複する部分が存在する。この結果、出来事全体の把握が難しくなっている。以上のような特徴を持つ出来事に関する文章に関しては、冗長性の低減を考慮した複数文書要約の実現により、効率的な情報の取得が可能になり、有効であると考えられる。

このような新聞記事について、小倉 [1] は Allen の時空間モデルを用い、文章に含まれる文すべての時間関係を求めることで文章中の文を時系列順に整列する手法を提案し、この手法の有効性を示している。

そこで本研究でも、文章をまとめる観点として時間を用いて複数文書自動要約を実現する手法を提案する。文章中出现する時間表現を手がかりに、事件展開構造を生成する。ここから、様々な観点でパスを選択する事によって、多様な要約文を生成することができる。

2 文章の事件展開構造化

2.1 事件展開構造の定義

文章中出现する時間情報に注目してみると、明らかに時間的な前後関係がわかるものと、わからないものがある。前後関係がわからない場合は、文中中出现する時間情報が異なっているが、比較でき

る部分まで同一である場合である。詳細さが異なる時間情報は、どちらが先に起きた出来事かについては曖昧であるため、単純に比較することは意味がない。また、時間情報と文の記述している出来事について調べてみると、この二つには関連性があることがわかる。一般に、時間情報が詳細なほど文の記述も詳細であり、逆に大まかな時間情報しかない文では出来事を大まかに述べている場合がほとんどである。このことから、時間情報の詳細さが異なる文は並列に扱う事が妥当であると考えられる。

以上より、文章の時間的な関係は図1の様に、部分的に並列になっていると考えられる。そこで、これを事件展開構造として定義する。事件展開構造とは、時間情報を元に事件記事を構造化したものである。

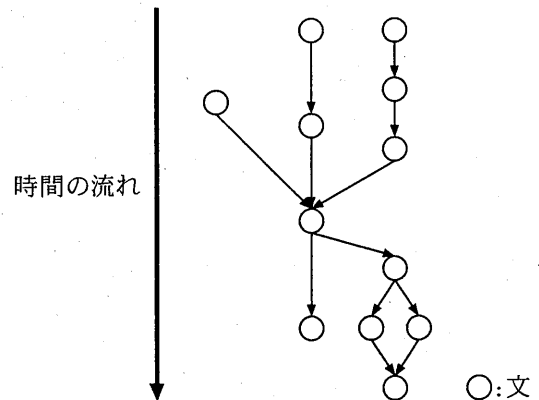


図 1: 文章と時間の関係

定義 1 事件展開構造とは、文をノードとし、文が表す事象の時間的な前後関係をリンクとして文章を構造化したものである。ノードの時間情報に前後関係があるものは時系列順に整列し、時間情報が同一、もしくは前後関係が不確かなものは並列に並べることにより構造を生成する。

2.2 時間セグメント

時間表現やテンス・アスペクトなどといった時間情報に注目して文章を調べると、文章内では文の時

間情報に関連性がある文は連続して出現していることがわかる。これを時間セグメントと定義する。

定義 2 時間セグメントとは、文章中の時間的な境界から次の時間的な境界の間に出現している文の集合である。そのため、同じセグメントに属する文は何らかの時間的な関連性を持って起きた事象について記述されている。

ここで、時間情報は何らかの時点を明示的に示している。そのため、時間情報がある文の直前で時間の連続性が途切れている。よって、時間セグメントに時間情報は多くとも一つしか存在しない。

時間セグメントの種類は大きく分けて以下の2種類ある。

- 現在のセグメント

時間情報を持たず、時制によって現在の出来事や状態について書かれた文と判別できる文が集まったセグメントである。このセグメントでは文章が書かれた時点での状況や、行われていることが記述されている。

- 過去のセグメント

過去の出来事について書かれた文が集まっているセグメントである。このセグメントには過去の出来事や、そのときの状態を表す文が含まれている。そのため、このセグメントには時間情報がある文、もしくはアスペクトによって過去と判別できる文が属している。このため、過去のセグメントには時間情報があるものとなないものが存在する。

文章中には時間セグメントの配置順がそこに書かれている時間情報と無関係に出現している。そのため、時間セグメントを時系列順に整理することにより、文の時間順整理が実現できると考えられる。

2.3 時間セグメント生成

ある文の前が時間セグメントの区切りとなるのは以下の場合である。

1. 文中に時間表現が出現した場合
2. 過去の文の次に現在の文が出現した場合
3. 現在の文の次に過去の文が出現した場合
4. 段落が変わった場合

1の場合、時間表現によって文の関連性が途切れるため、新しいセグメントとする。2,3については、過去の出来事についての記述か、文章が書かれた時の状態や出来事であるかが変化している部分で

ある。そのため、これも新しいセグメントとする。4の段落が変わった場合であるが、このような時筆者が意識して文の区切りとしている。そのため、時間的な関連性も途切れている。よって、ここでもセグメントを改める。

また、以下の場合例外として新たなセグメントを生成しない。

- 文頭に接続詞がある場合
- 段落の最初の文に時間情報がなく、次の文に時間情報がある場合

次に、時間セグメントが持つ時間情報を定める。

1. 記事が書かれた時間

現在のセグメントには、記事が書かれたときどのような状況であるかが記述されているので、このような場合には記事が書かれた時間を時間セグメントの時間情報とする。

2. 時間表現などを元にした時間情報

時間情報を持った文が出現する場合、文の時間情報をセグメントの時間情報とする。

3. 知識を元にした時間情報

個々の事件にもよるが、事件の大まかな流れはあまり変化せず、一定である場合が多い。そのため、一般的な事件の流れと変わらない部分については、時間情報が省略される場合がある。そこで、一般的な出来事の成り立ちを知識として定義する。図2は一般的な事件の流れを表している知識の例である。これらについて表1のような手がかり語をそれぞれ用意し、用いることで時間情報を補完する。

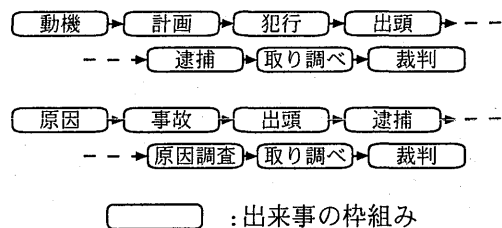


図 2: 一般的な事件の流れの例

このような知識を用いてもまだ時間情報がないセグメントが存在する。このようなセグメントは、図3のような場合が考えられる。

以下、それぞれの場合について述べる。

(1)については、過去のセグメント a と c の間に挿入的に現在のセグメントが表出している場合である。この場合、セグメント a と c は本来一つのセ

計画	標的にした
犯行	誘拐した 殺害した
出頭	出頭した 自首した
逮捕	逮捕
取り調べ	押収した

表 1: 出来事と手がかり語の例

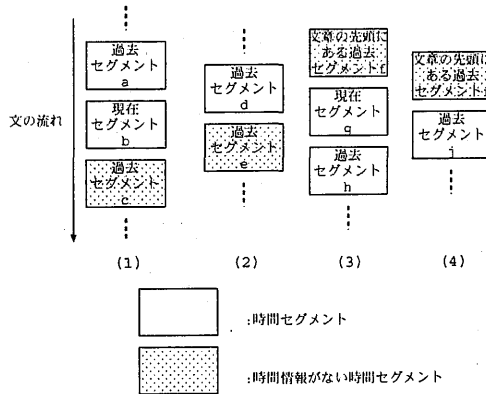


図 3: 時間情報がないセグメントが出現する場合

グメントであったと考えられる。しかし、セグメント a の最後に記述されている事象について、現在の状況などを補足している時などにこのような状態になる。

(2) については、段落が切り替わった場合などによりセグメントが分割されたが、本来は同じセグメントであった場合である。この場合、文章構成上新たな段落となっているが、時間的な流れは直前のセグメントを継承している。

(3) と (4) は (1) と (2) がそれぞれ文章の先頭であった場合である。

よって、以上のように時間情報がないセグメントは、本来直前の過去セグメントと同一であったものが、文章の構成上分割されたために発生していると考えられる。よって、このようなセグメントは以下のルールで統合する。

- 文章中の過去のセグメントに時間情報がなかった場合、そのセグメントの前に出現した過去のセグメントに統合する。
- 文章の最初に出現した過去のセグメントに時間情報がなかった場合、その次に出現する過去のセグメントに統合する。

2.4 事件展開構造の生成

事件展開構造は、時間セグメント A、B を以下のルールで整理することにより生成できる。

- A が持つ時間情報より B が持つ時間情報が明らかに後であるならば、A → B の順にセグメントを整理する。
- A と B の時間情報が全く同じであるか、もしくは細かさが異なる時間情報のうち比較できる部分までは同一であるならば、並列に並べる。
- A についての前後関係を知識を用いて調べた場合、次に起きる出来事のうちもっとも最初に出現するセグメント B の直前にセグメントをおく。
- A が持つ時間情報が時点であり、B が持つ時間情報時区間の時に、両方の時間情報が完全に一致する場合がある。時区間は時間情報の時点から、もしくは時間情報の時点までの出来事である。そこで、時間情報の時間からの場合は時点の時間情報より後の出来事とし、時間情報の時間までの場合は時点の時間情報より前の出来事であるとする。

3 内容が同一な文の単一化

関連文書の一つにまとめた時、内容が重複する文がある。このような文を単一化することにより、冗長性の低減する。文の内容は格フレームを用いる。格フレームで表現された文について、すべてのスロットが同一か、言い換えられた形であれば単一化していく。

また、事件展開構造における単一化された文の扱いを決定する必要がある。ここで、文が属している時間セグメントが、時間的に関連がある文の集合であることを考えると、文がどちらか片方にのみ出現するようにすると文の流れが途切れてしまう。よって、双方の時間セグメントに文が存在していることが妥当である。そこで、単一化された文にはリンクを生成し、同一の文であることを明示しておく。

4 パス選択による要約文生成

前節までで生成した事件展開構造の例を図 4 に示す。この構造から、リンクをたどりながらノードを選択していくことで、時間に沿った文章を生成することができる。本節では、このようなパス選択による要約文の生成について述べる。

要約文を生成する時、どのような条件でパスを選ぶかを決定する必要がある。ここで用いる条件により、生成される要約文が変化する。これは、条件の選択と同時に要約文の観点を選択しているためである。例えば、最短パスを選択すると、もっとも要約率が高く、出来事を大まかに述べた要約が生成され

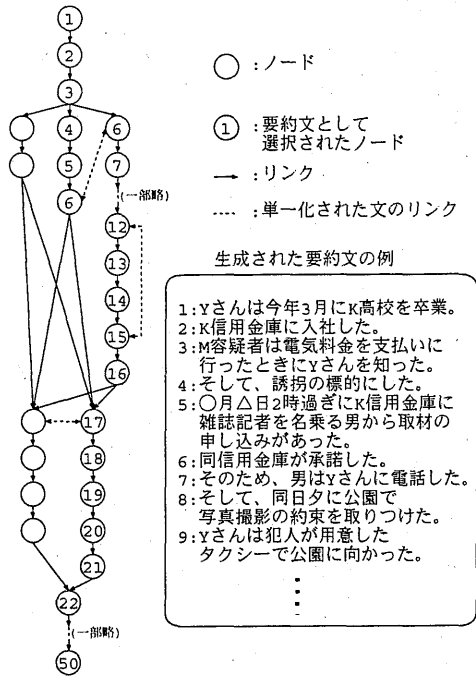


図 4: 事件展開構造の例

る。また、最長パスを選択すると要約率は低いが、出来事を細かく述べた要約を生成することができる。そこで、これらの要約を生成し、評価を行う。

	事件 1	事件 2
全文数	59	24
生成された要約文の数	14	1
最短要約文数 (要約率)	37(62.7%)	24(100%)
最長要約文数 (要約率)	50(84.7%)	24(100%)

表 2: 本手法での要約率

本研究では実験データとして、1993年毎日新聞のコーパスから2文書にわたる2つの事件について人手で事件展開構造を生成し、パス選択を行うことによって要約文を生成した。表2はそれぞれについて最短パスを求めた要約と最長パスを求めた要約の要約率を示したものである。

ここで、事件2については時間的に並列な部分がなく、文の単一化も起らなかったためにすべての文が要約として用いられたため、要約率が100%になっている。

これらの3つの要約文それぞれについて以下のようなアンケートを行った。この結果を表3に示す。

1. 文の意味的に文章の繋がりが誤っている部分があるか
2. 事件1の最短要約と最長要約のどちらが詳細

な要約であると感じたか

3. 出来事全体を時間を追って把握できたと感じたか

		事件 1		事件 2
		最短要約	最長要約	
被験者 1	繋がりの誤り数	0	1	0
	出来事を把握できたか	○	○	○
被験者 2	繋がりの誤り数	0	1	0
	出来事を把握できたか	○	○	○
被験者 3	繋がりの誤り数	0	1	1
	出来事を把握できたか	○	○	○
被験者 4	繋がりの誤り数	0	1	0
	出来事を把握できたか	○	○	○

表 3: 要約文についての調査

それぞれの要約文の平均誤り率は、事件1の最短要約で0%、最長要約で2.1%、事件2の要約で1.4%であった。被験者が誤りとした文は被験者により異なり、あまり一致が見られなかった。これは、要約文の表層的な整形を行わなかったため、文の意味の取り方が被験者によって異なっているためだと考えられる。

また、すべての被験者が最短要約より最長要約の方が詳細な要約であると回答した。このことから、要約文の観点とパス選択の条件付けに関わりがあると考えられる。

さらに、すべての被験者が出来事を把握できたと感じている。よって、この手法は適切な要約文の生成に有効な手段であると考えられる。

5 まとめと今後の展望

複数文書を時間を観点としてまとめ、事件展開構造の概念を元にパス選択による要約文の生成を行う手法について提案、実装し、評価を行った。

今後の展望としては、要約文生成の際に表層的な整形を行うことで、よりわかりやすい要約を生成することができる。また、本手法で定義した事件展開構造を、時間情報だけでなく主題や文章中の登場人物の視点などで生成することで、事件記事以外の文書についても要約を生成できると考えられる。

参考文献

- [1] 小倉牧人, 田村直良. 文間の時間制約モデルと事象の時系列化への応用に関する研究. 情報処理学会自然言語処理研究会研究報告 NL-140, 2000.

負データが極端に少ない訓練事例を用いる OCR 誤認識検出

田中 大輔† 田村 直良††

† 横浜国立大学 教育人間科学部 マルチメディア文化課程

†† 横浜国立大学大学院 環境情報研究院

{dtanaka,tam}@tamlab.eis.ynu.ac.jp

1 はじめに

本稿では、OCR システムによって認識された文章中の誤りを機械学習により検出する手法について論じる。

近年、官公庁における情報公開や、新聞記事データベースに見られるように、膨大な文書をインターネットで閲覧できるようになってきている。これらを実現するために、紙媒体の文書を情報検索に利用できるテキストデータに電子化することが要求されている。一般に文書をテキストデータに電子化するには、イメージスキャナにより紙媒体の文書を画像化し、認識する文字領域を選択し、OCR によって文字認識を行う。

現在、市販されている OCR エンジン、ノイズの少ない原稿では 99% 前後の認識精度を実現している。しかし、一般的には、誤り文字が含まれていることが気にならない認識精度は 99.95% 以上が必要とされている。そのため、誤認識を訂正する何らかの後処理が必要である。後処理は人手に依る作業が大部分を占め負担が大きい。そこで、これら人手による作業の補助となるシステムが望まれる。

今までの研究では、誤認識の訂正を行っているものが数多い。自動的に誤認識を訂正することは、正しい文字までをも改悪する弊害を持ち合わせている。そのため、99.95% 以上の精度を満たすことができていない。竹内ら [6] は共起情報を用いて検出を行い、精度の高い文字の修正を行っている。共起情報は誤り文字の含まれない文書から文字 trigram モデルとして作成している。そして、OCR 出力した文字がその文字 tri-gram モデルに含まれていなければ誤りであるとしている。しかし、誤り文字列は常に、日本語として存在しない文字列であるとは限らない。つまり、誤りのない文章の trigram モデルに合致する OCR 誤認識文字を含む文字列が存在する。よって、そういった文字 trigram モデルに含まれる誤り文字は検出される可能性が低いと考えられる。また、久光ら [3] は OCR の文字候補付きの出力結果と、形態素解析を用いて後処理の効率化を図っている。この手法では OCR 認識エンジンの精度に依るところが大きい。

これらの他に OCR 誤認識文字検出及び訂正に関するさまざまな研究が為されているが、多くは 95% 以下の OCR 認識精度を前提にしている。本研究では、99% 程度の精度の認識結果に対して誤認識文字を発見して、99.95% 以上の精度の文章を得ることを目標としている。手法として機械学習によるが、このような前提では、正解（正データ）に対して認識誤り（負データ）の数が極端に少なく、十分な学習効果が得られない。本研究では、重回帰分析、C4.5 を基準にクラスの分布に極端に偏りがある訓練事例の場合にも高精度の分類器を構成することを試みる。

2 誤認識文字発見手法の概要

本研究では OCR 認識された文章を入力とし、文字ごとにくつかのパラメタを抽出した事例データを用い機械学習することにより、誤認識を自動的に検出するシステムを実現する。手法の概要は以下の 4 つのフェーズからなる。

1. OCR 認識文字のパラメタを得るために、統計的確率を算出をする¹。
2. 訓練データから事例データを作成する。
3. 誤認識文字判定対象である OCR 出力文字のパラメタ抽出をする。
4. 判定モデルを作るための最適な事例データの抽出と AdaBoost[2] を用いた機械学習を行い、判定モデルを導出する。

3 事例データの作成

3.1 パラメタの設定

本研究では、機械学習で用いるパラメタとして以下の 8 種類を設定する。

¹衆議院ホームページ [4] より国会会議録約 9000 万文字のコーパスを得た。

1. 形態素コスト・前後の接続コスト

単語の構造的正当性の指標として、形態素コストと接続コストを形態素解析プログラム「茶筌」[5]を使用して求めている。形態素解析プログラムは、ある特定の位置からはじまるすべての可能な形態素を辞書引きによって得る。そして、辞書引きによって得られた個々の形態素に対して、その直前の位置に存在するすべての形態素との接続可能性のチェック、および、コストの計算を行なう。

誤り文字が含まれる形態素は、辞書に該当する形態素が存在しない、または、接続の可能性が低い等が考えられる。つまり、形態素コスト・接続コスト共に他と比較してコストが、高いと推定される。

2. 画数

複雑な文字ほど、誤認識が起りやすい可能性がある。そこで、画数データを用いる。各文字に対してそれぞれの画数を割り当てる。

3. 文頭であるか

パラメタを算出しようとする文字が文頭の場合、前との接続コストを算出することができない。そこで、文頭と文中を区別するパラメタで代用する。文頭である形態素には1を、そうでないときは0を割り当てる。

4. 形態素の長さ

OCR 誤認識文字が含まれる文は、形態素が短く区切られて形態素解析される場合がある。そこで、形態素の長さをパラメタとして用いる。

5. 文字 bi-gram 確率

OCR 誤認識を含む文字列は日本語の文として意味を為さない事があると思われる。つまり、OCR 誤認識を含む連続した2つの文字の共起確率は低いと考えられる。そこで、本研究では N-gram モデルを用いる。しかし、tri-gram モデルでは、文字の組み合わせが膨大になり、計算量もかかりことから、bi-gram モデルを採用する。

あらかじめ適切な文字列から構成されたコーパス [4] より、文字 bi-gram 確率を算出しておく。求めた文字 bi-gram 確率より、認識文字の文字 bi-gram 確率を検索し割り当てる。

6. 品詞 tri-gram 確率

OCR 誤認識を含む文字列は日本語的に意味を為さない事がある。つまり、OCR 誤認識を含む連続した3つの形態素の品詞共起確率

は低いと考えられる。そこで、本研究では N-gram モデルを用いる。品詞の場合、組み合わせ数が少ないため tri-gram モデルを採用する。

あらかじめ適切な文字列から構成されたコーパス [4] より、品詞 tri-gram 確率を算出しておく。求めた品詞 tri-gram 確率より、認識文字の品詞 tri-gram 確率を検索し割り当てる。

7. 文字誤り易さ確率

OCR 文字認識においては、文字ごとに間違いやすさに偏りがある。この偏りは、大量の OCR 誤認識を調査すれば、文字ごとの誤り易さを算出することができる。しかし、手作業で行わなければならない、多大な労力を要する。そこで、同じ文章で画像解像度 400dpi の画像と画像解像度 200dpi の画像との OCR 結果比較し、差が出た文字を誤りやすい文字とみなす。文字単位で差分が出た文字の数をカウントし差分の出た確率をパラメタとして用いる。本論文では、この確率を文字誤り易さ確率と呼ぶ。

3.2 スムージング

本研究では、N-gram モデルを用いているパラメタがある。しかし、統計的に求めたパラメタにはゼロ頻度問題が存在する。そこで、ゼロ頻度問題に対処するために、スムージングを用いる。

● ワン・カウント法 (One-Count Method)

本研究では、スムージング手法としてワン・カウント法を用いる。ワン・カウント法は、モデルの次数、学習データ量にかかわらず優れた性能を示すとされている [1]。

4 分類器の生成

OCR 誤認識文字の検出では、OCR 出力を「正」と「負」の2クラスに分類する。ここでは、線形モデルを作ることのできる重回帰分析を用いる。

4.1 ブースティング

本研究で使用する訓練データは、クラスの数に極端な偏りがある。弱判定器にしきい値の調整が容易な重回帰分析を用いた AdaBoost アルゴリズム [2] により高精度の分類器を作成する。 $D_t(i)$ はラウンド t で求められた事例 i の重みである。

各ラウンド t の弱判定器は、事例ベクトル $(X_{i1}, X_{i2}, \dots, X_{im-1})$ とその重みの積

$$D_t(i) \cdot \hat{Y}_i = D_t(i) \cdot (b_0 + \sum_{j=1}^{m-1} b_j X_{ij}) \quad (1)$$

と、判定クラスの重みの積

$$D_t(i) \cdot Y_i \quad (2)$$

から残差に関する最小2乗法により最適な b_0, b_1, \dots, b_m を求める。

4.2 事例データの削除

使用する事例データは、クラスの分布に極端な偏りがあり、的確なブースティングを行うのは困難である。そこで、事例データ中の割合が多い正データに限り削除する。削除は、訓練データとして用いるデータをクローズドな環境で重回帰分析を行い、正解を誤判定した事例を削除し、正データ数を減らす。本研究ではこの手法により、正データを10分の1まで削減した。

4.3 再現率・適合率

本研究はOCR誤認識文字を検出することを主眼を置いているため、「負」クラスのみに着目した再現率・適合率・F値を求める。

以下のように定義した再現率、適合率を求め、検討を行なう。本研究では、OCR誤認識文字検出を目的としているため、誤認識文字をどれだけ検出できたかを表す再現率を重視する。このため $\gamma = 3$ とする。

$$\text{再現率 } R = \frac{\text{「誤」判定が正しかった文字数}}{\text{実験データ中の誤り文字数}} \quad (3)$$

$$\text{適合率 } P = \frac{\text{「誤」判定が正しかった文字数}}{\text{「誤」と判定した文字数}} \quad (4)$$

$$F = \frac{(1 + \gamma^2) \times P \times R}{\gamma^2 \times P + R} \quad (5)$$

5 実験結果と評価

5.1 実験データについて

誤認識文字判定モジュールで行う訓練と評価のために、国会議事録55536文字からなるOCR認識文字の特徴ベクトルを抽出し、正誤のクラスに分類したものを作成する。「正」、「負」のクラスの分類に際しては、OCR認識対象文書となった国会議事録原本のコピーを参照し、手作業で行う。目標とする99.95%を達成するための再現率 R を以下の式に示し、表1に訓練データの内訳を示す。なお、OCR出力の文字認識制度を f で表す。

$$\text{目標を達成するための再現率 } R = \frac{99.95 - f}{1 - f} \quad (6)$$

よって、本研究で目標とする99.95%の精度を実現するには、76.37%以上の再現率が必要となる。

総文字数	正	誤	OCR認識精度
55536	55419	117	99.789%

表1: 実験データの内訳

5.2 実験結果

実験結果と評価、検討を示す。

1. オリジナルデータを用いた重回帰分析の結果
OCR出力から事例データを作成し、そのデータをそのまま用いて機械学習を行った結果を2に示す。

No.	R	P	F
1	0.452	0.072	0.30%
2	0.667	0.092	0.41%
3	0.528	0.207	0.46%

表2: オリジナルデータを用いた重回帰分析の結果

2. 重回帰分析を用いて正データを削除したデータを用いた結果

クローズドな環境で重回帰分析を行い、そこでの誤判定データを削除したデータを用いた重回帰分析の結果を表3に示す。

No.	R	P	F
1	0.484	0.143	0.39%
2	0.576	0.373	0.55%
3	0.811	0.160	0.58%

表3: 重回帰分析を用いて正データを削除したデータを用いた結果

3. ブースティングを行ったデータを用いた結果

クローズドな環境で誤判定データを削除したデータに、ブースティングを行ったデータを用いた重回帰分析の結果を表4に示す。

No.	R	P	F
1	0.355	0.224	0.34%
2	0.848	0.222	0.66%
3	0.811	0.167	0.59%

表 4: ブースティングを行ったデータを用いた結果

4. ランダムに正データを削除したものを用いた結果

比較対象として、ランダムに正データを削除したものを用いた重回帰分析の結果を表 5 に示す。

No.	R	P	F
1	0.452	0.126	0.36%
2	0.667	0.112	0.45%
3	0.623	0.213	0.52%

表 5: ランダムに正データを削除したものを用いた結果

5. C4.5 を用いた結果

比較対象として、分類器に C4.5 を用いた結果を表 6 に示す。

No.	R	P	F
1	0.642	0.290	0.57%
2	0.612	0.578	0.61%
3	0.500	0.226	0.45%

表 6: C4.5 を用いた結果

ブースティングを用いた手法では、表 4 で示すように、No.1 以外のテストデータに関して、他の手法よりも良好な結果が得られた。しかし、No.1 のテストデータは、重回帰分析を用いて正データを削除したデータを用いた結果よりも、ブースティングを行った結果悪化した。これは、ブースティングの特性として、基本的に弱分類器には 50% 以上の分類精度が求められているからであると考えられる。

6 おわりに

本研究では、文字ごとに文の構造、意味、OCR システムによる誤認識文字の偏りに関する特徴ベクトルを抽出し、OCR 出力から事例データを作成した。次に、弱判定器として重回帰分析を用いた

AdaBoost アルゴリズムにより、クラスの分布に極端に偏りがある訓練事例の場合にも高精度の分類器を構成することを試みた。

本研究では、人手による後処理の補助となるようなシステムの構築を目標としている。方針として、OCR 誤認識文字をできるだけ検出することを主眼を置き、OCR 誤認識文字を検出するものである。OCR 誤認識文字の検出には重回帰分析と機械学習システム C4.5 を用い、OCR 誤認識文字を検出することで実現した。

機械学習を行う際、形態素コスト、接続コスト、文字誤り易さ確率等、合計 8 種のパラメタを用いた。また、3 分割交差検定によりシステムの評価を行った。訓練データに関しては、現在の OCR システムに合致させるために、実際の OCR 出力を用意した。評価においては、OCR 誤認識文字をできるだけ検出することを主眼を置いているため、再現率を重視した。その結果、AdaBoost アルゴリズムを用いた機械学習では目標としている OCR 誤認識文字を訂正した後の誤り文字含有率 99.95% を達成した。

参考文献

- [1] S. Chen and J. Goodman. An empirical study of smoothing techniques for language modeling. *Proceedings of the Thirty-Fourth Annual Meeting of the Association for Computational Linguistics, Morgan Kaufmann Publishers*, pp. 310–318, 1996.
- [2] Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *European Conference on Computational Learning Theory*, pp. 23–37, 1995.
- [3] 久光徹, 丸川勝美, 島好博, 藤澤浩道, 新田義彦. ocr 誤認識後処理の効率について. 情報処理学会自然言語処理研究会研究報告 NL-104, pp. 17–24, 1994.
- [4] 衆議院. <http://www.shugiin.go.jp>.
- [5] 松本裕治, 北内啓, 山下達雄, 平野善隆, 松田寛, 浅原正幸. 日本語形態素解析システム「茶筌」version2.0 使用説明書 第二版. NAIST Technical Report, 奈良先端科学技術大学院大学 松本研究室, 1999.
- [6] 竹内孔一, 松本祐治. 共起情報と統計的形態素解析による ocr 誤り訂正. 情報処理学会自然言語処理研究会研究報告 NL-121, pp. 17–24, 1997.

主題構造解析による新聞記事からの気象情報の抽出と応用

大村 高史† 田村 直良††

† 横浜国立大学大学院 工学研究科 電子情報工学専攻

†† 横浜国立大学 環境情報研究院

{takashi,tam}@tamlab.eis.ynu.ac.jp

1 はじめに

近年、電子化された文書は増加の一途をたどっており、WWWの普及と相まって、これらの文書を手軽に入手することも可能になってきた。しかしながら、膨大な文書の中から欲しい情報を探したい、あるいは文書の集合を分析して傾向をつかみたいといった、利用者のさまざまな要求に十分対応できる情報アクセス手段はまだ乏しいのが現状である。

現在、我々は容易に入手可能であり、多方面に渡り事実の提示や物事の解説が多く含まれる新聞記事を情報源として、気象庁が発信している情報からでは得ることができない、社会生活に密着した気象に関わる知識を獲得することを目標として研究を行っている。気象情報は我々の生活と非常に深い関係を持っており、このシステムの実現は、我々の生活にとって非常に高い利益をもたらすと考えられる。その一端として、本研究では、文章から知識発見するために必要なデータベースの構築を目指し、データベースに蓄積するための事例の抽出方法について検討する。

事例の抽出は、基本的には新聞記事から気象に関する語句表現と、その気象の事象が発生した地域の名称を抽出することである。これは、それぞれ単独に固有名詞の抽出をするのでは全く意味がなく、気象情報とそれが発生した場所の組として情報を抽出する必要がある。単独の固有名詞の抽出に関する研究は盛んに行われているが、このように複数の語句を関連付けて抽出する研究はまだあまり注目されていない。

本研究で取り扱う2つの語句は、同一文内に定型的に出現するものではなく、その上、必ずしも同一文内に含まれているとも限らない。この場合、表層上のパターンマッチによる抽出手法は非常に困難であり、文章の内容を理解した上での抽出が必要不可欠である。そこで本研究では、文章中の話題の展開を把握することにより、語句の関連性を重視した抽出手法を提案する。

2 主題構造解析

文章は文の累積したものであるが、単なる集合体ではなく、時間的・線条的な累加・連続として成立する流れであり、文の継起的連続というところに、文章の特性がある [2]。

本研究では、1文にはその文において中心の話題となる語句(主題)が存在すると仮定する。そして、文間の主題の連鎖関係を解析することで、文章全体の話題の展開を把握することができると考えている。この解析を以降では主題構造解析と呼ぶ。以下に、主題構造解析について説明する。

2.1 主題の抽出

主題構造解析をするために、トピックと主題の抽出を行う。トピックと主題の定義を以下に示す。

- トピック：文章は、その全体の話題を表すようなトピックを持っていると仮定する。本研究では、新聞記事のトピックをそのタイトル中に出現する「名詞句」及び「名詞」と定義する。
- 主題と題述 [2]：各文は、主題構造を持つと仮定し、主題と題述とから構成されているとする。具体的には、係助詞「は」の直前に出現する「名詞句」もしくは「名詞」を主題と定義し、文の主題以外の残りの「名詞句」及び「名詞」を、題述と定義する。

主題を抽出する際、上記の定義によって主題が抽出されない文に関しては、明確な主題を持たない、前文の主題を維持している文であるとして、前文との連鎖関係を重視することにする。

また、主題となりうる語句が複数存在する場合は、1文中でより文末に近い語句を修飾している主題を抽出する。ただし、同じ語句を修飾している場合は、より文末の近くにあるものを抽出する。

2.2 主題の連鎖関係の種類

記事中の各文間が、下記の条件の、5種類の連鎖関係のうちで少なくとも1つを満たすものとし、何

らかの結束性を持っているとする。

- A主題維持：直前の文の主題と同一か、基準以上の類似度のある主題を持つ場合。
- B主題変化：直前の文の題述のいずれかと同一か、基準以上の類似度のある主題を持つ場合。
- C主題回復：最も近い主題変化の直前の主題と同一か、基準以上の類似度のある主題を持つ場合。
- Dトピックの導入：文章のトピックと同一か、基準以上の類似度のある主題を持つ場合。
- E主題派生：上記のいずれにも該当しない場合。この場合、直前の文やトピックとは関連性の低い文となる。

ただし、基準以上の類似度とは、一方の語句が他方の部分文字列になっている場合とする。

2.3 主題の連鎖関係の決定

主題の連鎖関係を決定するルールを示す。

- ルール1：原則として、結束関係の強さは $A > B > C > D > E$ とし、可能な限り結束性の高い連鎖を採用する。ただし、主題が存在しない場合は、前述した通り前文との連鎖関係を重視し、「主題の維持」を採用する。
- ルール2：第1文に関しては、前文が存在しないため、「トピックの導入」が採用できる場合にのみそれを採用する。それ以外の場合は、連鎖関係は採用しない。
- ルール3：ルール1にもとづき、連鎖関係を決定した結果、D「トピックの導入」とE「主題の派生」のどちらかになった場合、連鎖関係の修正処理を試みる。これは、主題を採用する際にあらかじめ抽出しておいた主題候補(主題を修飾する名詞句及び名詞)をもとに行う。主題候補を主題と仮定し、再度ルール1にもとづき連鎖関係を決定し、その結果がA～Cの連鎖関係となった場合にのみ、この主題と連鎖関係を採用する。ただし、主題候補が複数存在する場合は、主題の直前に存在するものから順に修正処理を試みる。

3 気象情報の抽出

気象に関する事項の抽出のために気象語句のタグ付けを行い、その後タグ付けされた記事を対象に気象に関して記事の分類を行う。

3.1 気象語句のタグ付け

本研究では、気象語句を、分類語彙表 [3] に掲載されている気象に関連の深い項目 (1.514 天災、1.5150 気象、1.5151 風、1.5152 雲、1.5153 雨・雪、1.5154 天気、1.5155 波・潮) に含まれる名詞とする。動詞を含む他品詞に関しては、気象を示すものを分類することが難しく、又高い精度が望めないことが予想されるため省略する。

タグ付け手法としては、前述の気象語句をリストとして用意しておき、記事毎に表層上のパターンマッチング手法により気象語句にタグ付けを行う。パターンマッチングでは、文章の任意の位置からマッチングを行うと、誤ってタグ付けを行う確率が高くなるため、形態素解析ツール茶筌 [4] の結果で得られる形態素の境界をもとに、完全一致法によってマッチングを行う。気象語句が、二つの形態素にまたがる場合に関しては、その区間による完全一致でマッチングを行う。

3.2 気象記事の分類

気象語句のタグ付けを行った記事を対象に、気象と関連の深い記事と気象と関連の浅い記事を機械学習システム C 4.5 [1] を用いて分類する。気象記事をデータとして蓄積する際に分類を行い、蓄積するデータに分類項目を付け加えることによって、データ利用者が用途によって利用しやすくなることが期待される。

4 地名の抽出

4.1 地名データベース

電子化された郵便番号簿をもとに地名を階層化し、PERL の GDBM を用いて地名データベースを作成する。階層化することで、地名を単に名称だけでなく包含関係も考慮した系統的な情報の蓄積を実現することができる。

気象に関する記事をデータとして蓄積することを考慮すると、せまい地域に関してデータを蓄積することは、あまり有効ではない。そこで、第1階層：「国名」、第2階層：「地方名」、第3階層：「都道府県名」、第4階層：「市区郡名」として階層化し、第4階層よりせまい地域に関しては階層化しない。

4.2 地名のタグ付け

気象語句のタグ付けと同様に形態素解析ツール茶筌の結果で得られる境界をもとに、地名データベースに存在する地名とのマッチングを行う。また、タグ付けされた後の処理を効率化させるために、階層毎に異なるタグを使用する。

4.3 地名の抽出

タグ付けされた地名には、気象語句と全く関連の無い地名も多く含んでおり、気象語句と関連性の高い地名のみを抽出する必要がある。

手法としては、まずあらかじめ新聞記事に対して主題構造解析を行い、その結果をもとに、地名を抽出するにあたって関連性が高いと考えられる文の集合を抽出する。この文集合を以下では、related text とする。

- related text の抽出方法

related text の抽出方法としては、気象語句と地名の関連性を重視し、以下の条件をもとに抽出を行う。ただし、ここではある文Tとその前文との連鎖関係によって条件を設定している。また、本研究では、離れた文に出現する気象語句と関連した地名をいかにして抽出するか、いかに気象語句と関連の低い地名を削除するかということを目的としている。そのため、1文中で気象語句と地名の両方が出現している場合は、その2つの語句は関連性が高いということを前提としている。

1. 前文との連鎖関係が「主題の維持」の場合、文Tと前文は同一の内容に関して述べられていると考えられ、文Tと前文の中で出現する気象語句と地名は関連性が高いと考えられるため、文Tと前文は related text であるとする。
2. 前文との連鎖関係が「主題の変化」の場合、前文の題述である語句、つまり文Tの主題である語句が気象語句である場合にのみ、文Tと前文を related text とする。これは前文の内容に気象語句が出現し、その気象語句に関して文Tが述べられていると考えられるためである。
3. 連鎖関係が「主題の回復」の場合、文Tと主題の回復元である文の関係は、主題の維持の場合と同様であるため、related text であるとする。
4. 連鎖関係が「トピックの導入」の場合、タイトル中に出現する語句を主題としている文Tとタイトルの内容は非常に密接したものであると考えられるため、同様に文Tとタイトルは related text とする。
5. 前文との連鎖関係が「主題の派生」の場合、文Tは他のいかなる文とも related text を構成しない。

主題構造解析をもとにした related text の抽出例を図1で示す。

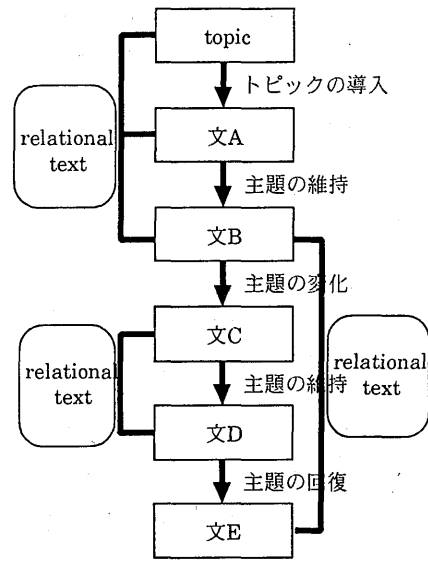


図 1: related text の抽出例

- 地名の抽出方法

作成された related text から地名を抽出する。この作成された related text は、主題構造解析の結果をもとに得られたものであり、この中に出現する気象語句と地名は関連性が高いと考えられる文の集合である。そのため、related text 内に気象語句と地名の両方が出現した場合は、その地名を全て抽出する。

4.4 地名のあいまいさの解消

本研究における第4階層の地名には、いくつかつづりが等しいものが存在する。そのようなあいまいな地名に対して、地名の同定を行う。手法を以下に示す。

地名の抽出において文章を左から右へ走査する際、都道府県名が出現すれば配列にスタックする。もし文章中の地名にあいまいさがなければ、この配列を参照せずに同定できる。あいまいな地名が出現した時には、それと同定すべき地名候補の上位階層地名がスタックに格納された地名と比較され、等しい地名が選択される。

5 実験と評価

地名の抽出の際、最も重要なことは、主題構造解析の結果を用いることによって気象語句と関連性の高い地名を正確に抽出できているかどうかである。これに関して、「A: 本論文で提案したシステムによって地名を抽出する方式」と「B: 1記事中に気象語句と地名が共起した場合に限り、その地名

を気象語句と関連のある地名として抽出する方式」の2つの方式を比較して検討する。その結果を表1で示す。ただし、実験では93年度日本経済新聞からランダムに選出した56記事に対して地名の抽出を行い、その結果に対して検討を行う。また、評価では再現率、適合率、F値を求める。F値($0 < F < 1$)とは、再現率Rと適合率Pのトータルでの精度を表す指標である。表1の結果から、B方

抽出手法	再現率	適合率	F 値
A 方式	36.8	58.3	0.451
B 方式	68.3	50.0	0.568

表 1: 実験の評価 1

式では、離れた文で気象語句とそれと関連性の高い地名が出現した場合、地名を抽出することは不可能である。そのため、この手法の再現率は極めて低い値になっている。つまり、このことから記事中には気象語句とは離れているが、関連性の高い地名が多く存在することが分かる。

またB方式では、A方式と比較すると再現率は良いが、適合率はやや低くなっている。これは、気象語句とは離れた文に存在する、気象語句と関連性の高い地名は抽出することができているが、その反面、関連性の低い地名に対しても抽出を行ってしまっているためであると考えられる。このような場合の多くは、気象の意味として用いられていない語句に気象のタグ付けを行ってしまっている場合である。

ここで実験として、本システムによって離れた語句を抽出する際の抽出精度について、地名と対応する気象が全ての気象語句とする場合と、気象の意味として用いられている気象語句に限った場合とに分けて、それぞれに対して本システムによって地名を抽出し、その結果を表2に示し比較検討する。ただし、気象の意味として用いられているかどうかについては人手に判断によるものである。

地名抽出の対象	正解率
全ての気象語句	44.1
気象の意味として用いられている気象語句	73.3

表 2: 実験の評価 2

表2の結果から、気象語句と関連のある離れた地名を抽出する際、気象語句が気象の意味として用いられている場合については、約7割の正解率が得られた。しかし、気象語句が気象の意味として用いられていない場合が多くあるために全体としてはあまり良い結果が得られなかった。このことから、主題構造解析自体は有効な手段であるが、格構造解析等により気象語句が気象の意味として用いられている

かどうかを解析する手段が新たに必要であることが分かった。

図2に地名の抽出例を示す。

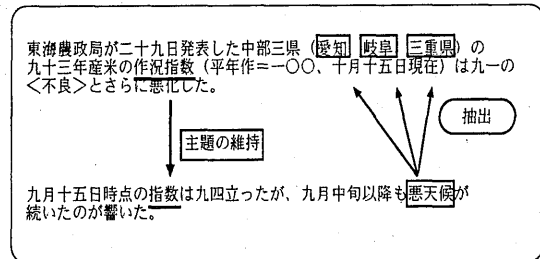


図 2: 地名の抽出例

6 まとめと今後の展望

本研究では、新聞記事を対象に主題構造解析をすることによって話題の展開を把握することにより、気象語句と関連性の高い地名を抽出する手法を試みた。その結果、パターンマッチング手法や1文中の共起情報からの抽出ではなすことができなかった、離れた文からの関連語句の抽出をすることが可能となった。今後は、蓄積されたデータをもとに、テキストマイニングを行い、傾向から利用者が気付かない新たな知識を発見できるシステムを構築することが大きな目標である。

また、主題構造解析における1文の単文化や主題の抽出精度向上、気象語句の抽出における抽出精度向上など解決しなければならない問題が残されている。

参考文献

- [1] J.R. キンラン. AIによるデータ解析. トップラン, 1985.
- [2] M. A. K.Halliday. An introduction to functional grammar second edition. くろしお出版, 2001.
- [3] 国立国語研究所. 分類語彙表. 秀英出版, 1993.
- [4] 松本裕治, 北内啓, 山下達雄, 平野善隆, 松田寛, 浅原正幸. 日本語形態素解析システム「茶筌」version2.0 使用説明書 第二版. NAIST Technical Report, 奈良先端科学技術大学院大学 松本研究室, 1999.

以下の頁は著作権者の許諾を得ていないため、公表できません。

抜刷 4 ~ 抜刷 6

抜刷 8 ~ 抜刷 8

抜刷 13 ~ 抜刷 13

抜刷 28 ~ 抜刷 28

抜刷 ~ 抜刷

抜刷 ~ 抜刷