

第5回 公開シンポジウム

人文科学とデータベース

「データ」を読む・観る・解く

1999年12月18日

主 催

人文系データベース協議会

共 催

関西大学工業技術研究所

(知識アーカイブ研究グループ)

人文系データベース協議会

当公開シンポジウムの継続開催を目的として、16大学の会員からなる「人文系データベース協議会」が発足しました。今後年1回のペースで会員大学を巡回して開催します。ご関係各位のご支援をお願いいたします。

人文系データベース協議会議長
小沢一雅@大阪電気通信大学

会員一覧(敬称略50音順:1999.5.29)

出田 和久 (奈良女子大学)	高橋 晴子 (大阪樟蔭女子大学)
江澤 義典 (関西大学)	都司 達夫 (福井大学)
及川 昭文 (総合研究大学院大学)	中谷 広正 (静岡大学)
小沢 一雅 (大阪電気通信大学)	中村 敏枝 (大阪大学)
雄山 真弓 (関西学院大学)	八村 広三郎 (立命館大学)
加藤 常員 (大阪電気通信大学)	濱 裕光 (大阪市立大学)
川口 洋 (帝塚山大学)	深海 悟 (大阪工業大学)
黒川 隆夫 (京都工芸繊維大学)	宝珍 輝尚 (福井大学)
柴山 守 (大阪市立大学)	正木 久仁 (大阪教育大学)

目次

(招待講演)

1. 「顔の認知研究における顔データベースの利用」 1
関西大学総合情報学部 教授 加藤 隆 氏
2. 「Recent Research on Paleolithic Arts in Europe and the MultimediaDatabase」 11
Department of Historical Sciences, University of Cantabria,
Prof. Cesar Gonzalez Saintz

(一般講演)

3. 「PhotoVR 考古資料データベース『北スペインの旧石器洞窟美術』」 23
日本・スペイン産学共同プロジェクトの実現と諸問題
深沢 武雄(株式会社テクネ)
4. 「バーチャルリアリティによる遺跡探訪 - 3D Archaeo-Copter -」 35
中村 健、小沢 一雅(大阪電気通信大学)
5. 「照度差ステレオ法を用いた遺物の表裏形状の計測」 39
結城 宏和、宝珍 輝尚、都司 達夫(福井大学)
6. 「文学データベースのための文書の構造化と意味管理」 47
横田 一正、三宅 忠明、國島 丈生(岡山県立大学)、
劉 渤江(岡山理科大学)、田植 明子((株)リュービシステムサービス)
7. 「相対インデックス法を使った文構造分析」 59
雄山 真弓、岡田 孝、黒崎 茂樹(関西学院大学)
8. 「仮想電子辞書システムの設計と構築」 67
芳野 学、都司 達夫、宝珍 輝尚(福井大学)
- ・公開シンポジウム「人文科学とデータベース」1995プログラム 79
- ・第2回公開シンポジウム「人文科学とデータベース」プログラム 80
- ・第3回公開シンポジウム「人文科学とデータベース」プログラム 81
- ・第4回公開シンポジウム「人文科学とデータベース」プログラム 82
- ・第5回公開シンポジウム「人文科学とデータベース」プログラム 83

顔の認知研究における顔データベースの利用

Making Use of Face Database for Cognitive Research on Face Processing

加藤 隆
Takashi KATO

関西大学総合情報学部
大阪府 高槻市 霊仙寺町 2-1-1
Faculty of Informatics, Kansai University
2-1-1 Ryozenji-cho, Takatsuki-shi, Osaka 569-1095

ATR 人間情報通信研究所
京都府相楽郡精華町光台 2-2
ATR Human Information Processing Research Laboratories
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288

キーワード: 顔, 認知, 記憶, 顔データベース
Keywords: face, cognition, memory, face database

あらまし: 顔の認知研究は近年になって大きな進展を見せている。その重要な要因の一つとして、コンピュータ・グラフィックス技術の目覚ましい向上が挙げられる。すなわち、ワイヤーフレームモデルを用いた顔の変形操作やモーフィング技術を用いた顔の合成操作によって、それ以前には考えられなかった精巧で自然な顔刺激の作成が可能になり、様々な興味深い心理現象の発見と仮説検証を可能にしたことである。こうした合成顔を用いることは心理実験における的確な刺激統制という観点から望ましいことであるが、そのためには合成に用いる顔データベースの整備が不可欠である。本稿では、実験用の顔刺激の作成にモーフィング技術を用いた研究事例を取り上げ、顔の認知心理学的研究においては、実験で提示する顔がすべて実在のものである必要はないが非現実的な顔であってもならないこと、そして実験統制という観点からは仮説検証に関わる部分以外の刺激統制が要求されることを指摘し、顔データベースに対

する認知研究からの要求仕様について考察してみたい。さらに、研究目的から合成顔の使用が不適切な研究事例を取り上げ、モーフィング合成への適用も視野に入れた質の高い多量顔データベースへの期待を述べる。

Summary: This paper presents four examples of cognitive psychological studies on face processing to illustrate the nature of facial stimuli used in such experimental studies and to discuss “user requirements” for face databases that can be expected to facilitate cognitive research on face processing. The first two examples illustrate cases where research objectives necessitate the use of synthesized, rather than authentic, faces, whereas the third example shows a case in which availability of large database of authentic faces is essential to meet the research objective. The fourth example shows that well-controlled variations of the same individual faces would be needed to study the effects of visual changes in faces. The primary requirement for any face database suggested by these studies is that orig-

inal faces need be recorded under carefully and strictly controlled face and environmental conditions, regardless of whether authentic or synthesized faces are to be used. It is further suggested that in order to meet a wide variety of research objectives, large database of three-dimensional face models with good quality of color texture need to be constructed.

1 はじめに

顔は人間のコミュニケーションにとって極めて重要な情報の発信源である。したがって、顔についての興味は尽きず、顔に関する心理学的研究の歴史も古い。しかし、認知心理学においては、ごく最近までは決して活発な研究領域とはいえなかった。一つの理由は、実験における顔刺激の統制が困難であることから、厳密な実証的検証を重視する認知心理学者にとって、顔は正当な研究の対象として採用し難いものであったのであろう。そのため、ごく少数の研究者が顔の認知処理について継続的に研究を行う傍ら、顔の認知を主テーマとしないものの顔刺激を用いた実験の結果が時折発表されるという状況であった。しかしながら、英国の認知心理学者を中心に80年代から90年代にかけて顔の認知研究が徐々に盛り上がりを見せ、ここ数年は日本でも一種の顔ブームを呈するまでに変貌を遂げてきた。

この近年における顔研究の大きな進展にはコンピュータグラフィックス技術の目覚ましい向上が重要な要因の一つとして挙げられる。すなわち、ワイヤーフレームモデルを用いた顔の変形操作やモーフィング技術を用いた顔の合成操作によって、それ以前には考えられなかった精巧で自然な顔刺激の作成が可能になり、様々な興味深い心理現象の発見と仮説検証実験を可能にしたのである。こうした合成顔は心理実験における実験刺激的な統制という観点から極めて有用なものであるといえるが、同時に、その利便性と妥当性を高めるためには、合成に用いる顔データベースの整備が不可欠である。

本稿では、顔刺激の作成にモーフィング技術を用いた研究事例をいくつか取り上げ、顔の認知心理学的研究においては、実験で提示する顔がすべて実在のものである必要はないが、同時に非現実

的な合成顔であってはならないこと、また、実験統制という観点からは、仮説検証に関わる部分以外の厳密な刺激統制が要求されることを指摘し、顔データベースに対する認知研究からの要求仕様について考察してみる。さらに、研究の目的から、合成顔の使用が不適切な場合の研究事例を取り上げ、モーフィング合成への適用も視野に入れつつ、統制の取れた質の高い多量顔データベースの必要性を指摘し、今後の方向性について一つの示唆を提示してみる。

2 多量の顔刺激を必要とする場合

多数の提示刺激を要求する認知心理学的実験の一例として、顔の潜在学習に関する研究事例を紹介する。潜在学習とは、学習の意図や意識を必要とすることなく、変化する外界刺激の背後に存在する規則性について暗黙的な知識を獲得することを言う。こうした現象を実験的に検証するためには、同じ規則性を持つ数多くの刺激を提示する必要がある。同時に、比較対象となる統制条件においては、同じ規則性を持つ刺激と規則性を持たない刺激を提示する必要がある。したがって、実験全体で必要となる刺激の数がかなりのものとなる。

研究事例¹では、数的に限りのある顔データベースに対して、2つの顔の平均顔をすべての組み合わせについてモーフィング合成するという方法を適用し、実験刺激として使用できる顔の数を飛躍的に増大させている。この実験研究では、提示する顔がすべて実在のものである必要はないものの、人工的な線画ではなく実画像という生態学的妥当性を高めた刺激を用いるという目的からは、顔が非現実的なものであってはならないという制約があった。2つの実在の顔の平均顔を合成するという方法は、少ない数のオリジナルから十分な数の実験刺激を確保するという目的と、規則性を容易に操作するという目的に合致したものであり、同時に、合成された顔が妥当なバリエーションの範囲に納まることを保証しているともいえる。

¹ 研究事例1は[6]を一部改変して転載した。より詳しい研究報告は[4]にある。

2.1 研究事例1：顔の視覚的類似性に関する潜在学習

人間が一生を通じて数多くの顔と出会い、その過程で、典型的な顔とか特異な顔といったカテゴリを獲得するようになるのも、明示的な学習というよりは暗黙的な学習の結果であると考えられる。こうしたカテゴリの特徴は、カテゴリを規定する要因が単一のパラメータ値で限定されるのではなく、ある許容範囲を持つところにある。また、カテゴリの獲得が様々な視覚的ノイズ（規則に無関連な特徴における様々な変化）のもとで行われるところにも特徴がある。本実験の主目的は、こうした顔のカテゴリに関する知識が潜在学習 [11] によって獲得される暗黙的なものであることを示すことである。

そのためには、定義された規則性の有無によって分類された数多くの実面の顔を用意する必要がある。これについては、定義された規則の正事例（あるいは負事例）の顔を、それぞれ他の正事例（負事例）の顔と対にして生成された『対平均顔』を用いることで対応した。対平均顔を用いる長所は、少ないオリジナル顔から多数の顔を生成できることだけでなく、生成された対平均顔がすべて定義された規則に則っているため、少ないオリジナル顔を計測するだけで、生成された数多くの平均顔については計測する必要がないことである。ただし、対平均顔の集合には同一顔の半分が「オリジナル顔の数-1」個分出てくるため、後述のような統制が必要である。

実験では、始めに、顔の直後再認を偶発学習課題として用い、その後、未提示の正事例と負事例の対を提示して、どちらが再認課題で見た顔と類似の規則性を持っているかを判断させた。顔刺激の作成では、まず、オリジナル顔を正事例群と負事例群とに分け、それぞれの中でさらに、再認記憶課題用と規則性判断課題用とに分けた。そして、4通りの集合の中に含まれるオリジナル顔だけを用いて生成された対平均顔を当該の条件で用いた。つまり、再認記憶課題用の対平均顔に用いられたオリジナル顔と、類似性判断課題用の対平均顔に用いられたオリジナル顔には重複が一切ない。これは、類似性の判断が規則性に依らずに特定の顔の記憶に影響される可能性を除外するために行った統制である。

本実験では、目-鼻-口で囲まれる面積と眉-頬-顎で囲まれる面積の比を規則として採用した。規則性に比較的広い許容範囲を持たせるために、各オリジナル顔の持つ値が母集団の平均より大きいか小さいかで正/負事例に分けた。このような面積比は、各パーツの特徴およびパーツ相互の配置関係と相関を持つものであるが、問題は、そうしたパーツ（間）の様々なバリエーションにもかかわらず、面積比という規則性を学習できるかである。

2.1.1 方法

被験者. 男女学部生 72 名を各 36 名からなる 2 群に分けた。

実験刺激. 上述のような統制のもとに生成された対平均顔を実験刺激として用いた。刺激はすべて 236x236 ピクセルのフルカラー画像として提示した。正/負事例ともに、再認課題用として 435 個、類似性判断課題用として 21 個の対平均顔を、それぞれ 30 個のオリジナル顔、7 個のオリジナル顔から生成した。

手続き. 被験者は、始めに、直後再認課題を 48 試行与えられた。各試行において、8 個の顔が同時に 10 秒間提示され、その直後にテスト顔が提示された。被験者は、テスト顔が直前に提示された 8 個の顔のいずれかと同じかどうか判定するように求められた。実験条件では、再認課題で用いられた顔はすべて正事例であり、統制条件では、正事例と負事例が同数提示された。被験者は、次に、再認課題で用いられた顔には、すべてある規則性があったことを知らされた。そして、これから提示される顔の対のどちらが、それらの顔と同じ規則性を持っているか判断するように求められた。この類似性判断課題は計 10 試行与えられた。終了後、被験者は、再認課題で見た顔に共通する規則性について思いつくことを何でも自由に記述するように求められた。

2.1.2 結果と考察

直後再認におけるヒット率と虚報率は、それぞれ、実験条件では 0.56, 0.31, 統制条件では 0.53, 0.24 であった。虚報率についてのみ有意な傾向が見られた： $F(1, 70) = 3.45, 0.05 < p < 0.1$ 。類似性判断課題において正事例が選ばれた割合は、実

験条件では 68%、統制条件では 49%であり、この差は有意であった：『中のパーツ（目、鼻、口）が、顔全体に比べて小さすぎない』と答えている以外は、面積比に多少とも関係のある記述は実験／統制条件ともに皆無であった。

これらの結果は、面積比という、複雑であり、かつ、他の特徴要因と比べて顕著性に乏しい視覚的規則性について、実験条件の被験者が暗黙的な知識を獲得したことを示している。重要なのは、この視覚的規則性が単一のパラメータ値で規定された限定的なものではなく、かなり広い許容範囲を持った規則性であることである。しかも、カラー画像のリアルな顔が数多く提示されたため、肌の色合いなどを含めた数多くの要因における様々な変化がすべて学習におけるノイズとして作用していたことである。

2.2 顔データベースへの要求

実験的検証の目的に合致した的確なモーフィング合成を行うためには、検証対象に関係のない部分についての統制が重要である。上記の実験研究においては、髪型の異なる 2 人の顔を合成した時に、2 人の髪が重ならない部分に「ゴースト」が生じた。この実験研究においては、髪型は規則性の中に定義されておらず、むしろ、規則性の潜在学習におけるランダムな刺激変化として位置付けられていたため、特に問題にはならなかった。

しかし、顔の好感度などを評定させる場合などは、こうしたモーフィング合成の「副作用」は取り除いておく必要がある。頭の部分については、例えば、輪郭部分に沿って顔以外の部分を黒く塗りつぶすなどの方法でゴーストも消去することができる。しかし、顔の一部として提示しなければならない額部分については、そもそも顔データベース作成の段階で、モデル全員に髪を上げさせて撮影するという統制が必要である。

また、上記の実験研究を展開していく際に実際に経験したことであるが、連言形の規則を定義しようとする、100 人程度の顔データベースでは、正事例と負事例に十分な数の顔を確保することが困難になる。可能な組み合わせすべての対について平均顔を合成することによってオリジナルの数よりもはるかに多くの実験刺激を得ようとする方法も、組み合わせを作る元のオリジナルが少数で

あっては効果が望めない。やはり、多数の顔を集めたデータベースの整備が望まれるところである。

3 顔データの厳密な質的統制が要求される場合

ここでは、研究の目的から、モーフィング合成そのものが要求された研究事例²を紹介し、髪形等の外部特徴の統制以外に、撮影時の照明や化粧の有無、また画質そのものについても厳密な統制が必要となるケースを取り上げる。

この研究事例における興味は、顔の魅力度の判断に影響を与えるパーツ（部位）が特定できるかということであった。その実験的検証の方法は、魅力度の高い（あるいは低い）顔の部位を魅力度の低い（あるいは高い）顔の部位とモーフィング合成により入れ換え、合成された顔の魅力度を判定させるというものであった。

3.1 研究事例 2：顔の魅力度判断におけるパーツの魅力の影響

本研究は、顔の魅力の社会的意味合いについて検討するものではなく、魅力という感性情報の判断に対して、顔のどのような視覚情報がかかわっているかについて検討を加えるものである。特に、顔の魅力を決める特定のパーツが存在するのか、あるいは、顔の全体的配置がより重要であるのかという観点から実験的検証を試みようとしたものである。そのために、評価対象の顔のパーツを別の顔のパーツと交換するという操作を行うことにした。たとえば、魅力度の高い顔の目を魅力度の低い目と入れ換えたときに、他のパーツの操作に比べて魅力度が著しく低下すれば、目が他のパーツよりも魅力度に重要な役割を果たしているといえる。

しかし、単純に、魅力度の低い顔に魅力度の高い目や口を入れれば顔全体の魅力度が上がり、逆に、魅力度の高い顔に魅力度の低い目や口を入れれば顔全体の魅力度は下がるのだろうか。あるいは、個々のパーツの魅力は顔全体の魅力にそれほど影響を与えるものではなく、顔全体のバランスがむしろ魅力を決めるのであろうか。ここで

² 研究事例 2 は [9] を一部改変して転載した。

は、こうした点についても検討を加えるために、別のパーツを100%入れ換えるモーフィング操作と、オリジナルを50%残すモーフィング操作とを用いた。これは、50%混合のモーフィング操作の方が100%入れ換えるモーフィング操作よりも、ももとのバランスが保持されると考えたからである。

3.1.1 方法

被験者. 被験者は男女学生計56名である。被験者は以下に示す4つの提示リスト条件に男女比が平等になるように14人ずつ無作為に配置された。

実験刺激とデザイン. 評定の対象には20代女性32名の特定の表情を持たない正面顔を用いた。操作した顔のパーツは、眉を含む目の部分、鼻、口、内部特徴（眉、目、鼻、口）、耳を含まない輪郭、の5種類であった。予備実験で得られた評定結果をもとに、45人の顔モデルと各パーツをそれぞれ魅力度の高い順にランクづけた。

顔モデル（これをベースと呼ぶことにする）の魅力度、入れ換えるパーツの種類、パーツの魅力度、モーフィングの割合の4要因を検討するために、各パーツごとに、魅力度の高いベースに魅力度の高い/低いパーツを入れ、魅力度の低いベースに魅力度の高い/低いパーツを入れる操作を行った。このとき、ベースのパーツを別のパーツと完全に入れ換える条件と、ベースのパーツと別のパーツを50%混合する条件を用いた。

パーツの入れ換えにあたっては、当然のことながら、同じ人物のベースとパーツの組み合わせを用いても意味がない。ここでは、一つの提示リストでは同一のベースを繰り返し用いないこととし、上記の各操作条件で2種類のベースとパーツの組み合わせを用意することにした。その結果、モーフィング割合を被験者間要因とした計4種類の提示リストが作成された。上記の制約を満たすために、魅力度の高いベースは上位10位以内のものを、魅力度の低いベースは下位9位以内のものをを用いることとし、入れ換えパーツについては、魅力度の高いものと低いものとを、それぞれの上位5位以内と下位5位以内から無作為に選択して用いた。

実験手続き. 顔画像はすべて512x512ピクセルの約32,000色のカラー画像として17インチモニター上に提示した。被験者は、各顔画像を4秒間

提示され、その後、1～7（1=魅力的でない、7=魅力的である）の7段階で魅力度を評定するように求められた。被験者が評定値を入力すると、次の顔画像が表示され、また4秒後に評定を求められるというように評定が進められた。各リストの20枚の顔画像は被験者ごとにランダム化されて提示された。

3.1.2 結果と考察

詳細な結果の分析は原著[9]に譲るとして、ここでは結果の概要と全体的な傾向を指摘するにどめる。

全体的に、鼻と口は魅力度の高い顔の重要なパーツとは言えないという結果が示されている。それに対して、目と眉はかなり重要なパーツであることが示された。顔の輪郭は、魅力度の高い顔の場合には魅力度に影響するが、魅力度の低い顔の場合には魅力度に影響しないという結果が出ている。これは、顔の魅力度を高めるには、内部特徴の魅力度を上げる必要があるが、顔の輪郭だけでは魅力度の低下につながることを示唆している。このことは、魅力度の高い顔の内部特徴に大きな変化がなくても、たとえば、年齢とともに輪郭に変化が出てくれば、それが魅力度に影響を与えるということを示唆している。

今回の実験では、パーツを100%入れ換えるモーフィング条件とオリジナルを50%残すモーフィング条件とを用いて、顔全体のバランスが魅力度に影響を与えている可能性についても手がかりを得ようと試みた。魅力度の高い顔の場合は、オリジナルのパーツを50%残してパーツの入れ換えを行う方が魅力度の評定値が高いという傾向を示した。これは、ベースのももとの平均評定値と比較して考えると、むしろ、魅力度の低下が少なかったと考えるほうが妥当であろう。また、内部特徴をすべて入れ換える場合でも、50%混合のほうが魅力度が高いということが示されている。こうしたことは、パーツの魅力度さえ高ければ何でもよいということではなく、バランスがやはり重要であるということを示しているのではないだろうか。特に、目と眉の入れ換え条件で、オリジナルを50%残す方が魅力度評定値が高い。これは、目と眉の全体に占めるバランスの影響を示しているものかもしれない。目と眉は化粧

で見た目の大きさや太さを変えることができるわけであるが、こうした化粧は全体とのバランスを考えながら施されているものなのかもしれない。

3.2 顔データベースへの要求

上記の実験研究におけるパーツの入れ換え操作の特徴は、異なる個人間でパーツ部分の入れ換えを行っているということである。このため、2つの顔のテクスチャが顕著に異なる場合には、パーツの輪郭部分についてグラデーションをかけたとしても、入れ換えたパーツの部分が顔の他の部分は異なる印象を与えてしまう可能性が出てくる。

例えば、厚化粧した顔の目元に化粧を施していない顔の目元を入れてしまうと、目元だけが化粧落ちしたような感じになるということである。したがって、特に女性の顔データベースを構築する場合は、顔モデルの化粧の度合いについても統制をはかる必要がある。もちろん、地肌が色白や小麦色といったように肌の色が異なる顔の間でパーツの入れ換え操作を行う場合も同様の問題が生じるが、これは仕方のないことであろう。この場合は、そうした組み合わせを除いて実験刺激を作成することになる。

また、照明を含む撮影時の様々な条件の違いから生じる顔の肌の色合いの違いも問題になる。例えば、全体に青みがかかった顔と全体に赤みがかかった顔の間で部位の入れ換え操作を行えば、顔の一部だけが肌の色合いが異なる「奇妙」な合成顔を生成してしまうことになる。したがって、顔モデルの肌の表面という内部要因だけでなく、撮影条件という外部要因についても十分に統制された顔を集めてデータベースを構築する必要がある。

4 数多くの実在の顔を必要とする場合

ここまでは、研究の目的から、妥当な顔であれば合成顔でも構わない場合と、逆に、モーフィングによる合成顔の生成を必要とする場合を見てきた。ここでは、合成顔を使用することが研究目的からは不適切であり、あくまでも実在の顔を用いて検証すべき場合を取り上げる。ここで紹介する

研究³の目的は、顔の全体的印象という感性的判断に影響を与える特徴が特定できるかということであった。したがって、顔刺激を実験的に操作して検証するというアプローチではなく、実在の顔にそうした特徴があるかどうかについて、むしろ調査方法的に検証する必要があるのである⁴。

4.1 事例3：顔の印象を規定する要因について

本研究の目的は、人間が顔の印象という感性的な分類判断を行う際に、どのような部分特徴なり配置情報を用いているかを検討するところにある。ここで用いた実験手法の特徴は、被験者に怖そうな顔とか優しくそうな顔という印象に当てはまる顔を複数個選ばせるところにある。このような複数個の顔は、ある特徴要因については同じような大きさのパラメータ値を持ったり、別の特徴要因についてはばらつきの大きいパラメータ値を持ったりするかもしれない。しかし、被験者にとって重要な特徴要因については、被験者の妥協の程度が小さく、したがって、同じようなパラメータ値を持つ顔が集められるものと期待できる。つまり、被験者が集めた複数個の顔の各特徴要因について分散を求めれば、その大小の比較から、被験者が分類判断において（暗黙的に）重要視した要因が特定できるものと期待できる。

ここでの主な興味は、被験者が一般的にどのような具体的なパラメータ値を基準として用いたかではなく、被験者が一般的にどの特徴要因を重要視し、それぞれの選択の中で類似のパラメータ値を持つ顔を揃えたかということである。たとえば、被験者Aにとっては、幅の広い鼻が怖そうな顔の特徴であるかもしれないし、被験者Bにとっては、逆に鼻先の尖った顔が怖そうな顔であるかもしれない。しかし、鼻の先の形状が印象を判断する際に重要な特徴であるという点で共通しているならば、被験者が集めた複数個の顔の分散はそれぞれ小さいはずである。したがって、複数個の顔に見られる分散は、好みという個人差を超えて共通の重要な特徴要因を示唆する指標となり得るのである。

³ 研究事例3は[4]を一部改変して転載した。より詳しい報告は[6]にある

⁴ 顔刺激に対して同様の要求を持つ研究に[12]がある。

もちろん、各特徴要因のパラメータ値が持つ平均値を求めることで、被験者が抱いている印象の平均像が得られる。そして、2つの印象の間で平均値が有意に異なる特徴要因が特定できれば、それが印象判断を左右する特徴要因であると言えることができる。

4.1.1 方法

被験者. 40名の大学生を20名ずつの2群に分け、一方には、男性の怖そうな顔と優しそうな顔の検索を、もう一方には、女性のきつそうな顔と優しそうな顔の検索を課した。

実験刺激. 男性101名、女性102名の顔の実画像を用いた。肌の色合いや化粧の影響を抑えるために、カラーの原画像をグレースケール画像に変換して用いた。年齢は男女ともに20代と30代がほぼ同数である。なお、すべての顔について、規定の特徴ポイントの座標が計測されている。

実験手続き. 被験者は、コンピュータ画面に一度に10個表示される顔の中から検索課題に合致する顔を選ぶという基本操作を繰り返しながら、最終的に検索課題にもっとも合致する顔を8個集めるように指示された。検索は被験者のペースで、被験者自身の選択基準をもとに行われた。

4.1.2 結果と考察

検索結果の分析に先立って、分析の対象とする特徴や配置情報をパラメータとして定義し、各顔の計測データを基にパラメータ値を求めた。各顔がそれぞれわずかながら異なる距離から撮影されていることから、個々の顔に設定された基準値を100としたときの相対値として、各パラメータの値を求めた。基準値は、両目の中心を結んだ直線から口の中心に下ろした垂線の長さである。

各被験者の検索結果について2つの基本データを集めた。1つは、それぞれの個別特徴や配置特徴について、検索された8個の顔のパラメータ値の標準偏差を求めた。このような標準偏差は、被験者が選択した8個の顔に見られる反応の揺れを示すものである。反応の揺れを各特徴要因間で直接的に比較できるように、各顔のパラメータ値を母集団（男性101名、女性102名）における標準スコアに変換した上で、このような標準偏差を求

めた。

もう1つのデータは、検索された8個の顔のパラメータ値の平均値である。これは、異なるカテゴリに分類された顔が、特にどのような特徴要因において際立って異なるかを見るためである。

今回の分析に用いた特徴要因は、個別特徴（眉の傾き、目の横幅、顎の尖り具合など23種類）、位置関係特徴（左右眉頭の距離、眉と目の距離、口の端と顎先を結ぶ直線の傾きなど16種類）、面積特徴（両眉尻と両目尻で囲まれた面積、両目尻と口の両端で囲まれた面積など6種類）の3グループに分類し、それぞれの分類に対して2要因（顔の印象×特徴要因）の分散分析を行った。

分散分析において交互作用が有意であったことから、単純主効果の分析を基に、怖そうな顔と優しそうな顔の特徴について具体的に検討した。その結果、男性の怖そうな顔とは、目が細く、カーブの鋭い眉で、眉と眉の間隔が狭く、下唇が厚く、眉と鼻や口との間隔が短く、両眉頭と両目頭で囲まれる面積が小さいという像が浮かんできた。男性の優しそうな顔は、目が丸めで、眉が下がり気味、上唇は薄い、下唇が厚く、目と目の間隔は短く、眉と鼻や口との間隔は長いというものである。

女性のきつそうな顔は、データの比較ははっきりした像が出てきた。それによると、眉が上がっていて、目が細くて上がり気味、口が小さくて、顔が細く、内部特徴が縦方向に狭い範囲に収まっているが、眉と目の間隔は開いていて、両眉頭と両目頭で囲まれる面積が小さいという顔である。女性の優しそうな顔も特徴がはっきりしていて、眉が長くて下がり気味、目が大きく口も大きい、上唇は薄く、顔が大きくふくよかで、内部特徴が縦方向に広く配置されていて、鼻が長いというものである。

先に線画の顔で同じような主観的分類実験を行ったことがあるが、そのときの結論として、優しそうな顔は『顔のタイプ』として認識され、怖そうな顔は『顔の表情』として認識されているのではないかという考察を提示した。今回の実画像の実験結果でも同じような傾向が読み取れる。すなわち、標準偏差値データを見ると、優しそうな顔の場合に比べて、怖そうな（きつそうな）顔の場合には、分散が全体的に大きい。つまり、被験者が集めた怖そうな（きつそうな）顔は被験者内でバリエーションが大きく、一つのタイプに集約

されていないことを示唆している。ところが、表情に関係しそうな目と目の間隔については被験者内の揺れが特に小さくなっている。

4.2 顔データベースへの要求

顔データの質的統制という意味では、先の2つの場合と同様に、研究目的とは関係のない部分における統制が必要である。研究の目的が顔の全体的印象を決める内部特徴の同定にあるのであれば、髪型の違いや撮影条件による肌の色合いの違いなどは除去されなければならない。

しかし、こうした調査的研究の場合は、顔データの質的統制に加えて、結果の外的妥当性（一般性）を高めるために、より多くの実在の顔を集めたデータベースの使用が望まれる。研究結果が、データベースに収められた少数の顔に限定されたものではなく、上記の事例で言えば、20代から30代の男性あるいは女性の顔に共通した特徴であることを示したいからである。

5 同一の顔モデルの多様な顔が必要な場合

人の顔は多種多様な変化を見せる。表情として部位の形状や布置情報の変化を示したり、顔の向きによって異なる視覚情報を提示したりする。そうした変化の有り様が顔が豊かな情報源であることを保証しているといえる。そこで、同じ人物の顔でありながら、観察条件によって顔の認知がどのように異なるかについての興味湧いてくることになる。ここで紹介する研究⁵の目的は、同じ表情の顔でありながら、顔の見える向きによって観察者が受ける印象が異なるかを検証するものであった。そこで、(微妙ではあっても)表情という研究目的とは関係のない要因を統制するために、同じ人物の顔を同時に異なる角度から撮影した顔刺激を用いる必要があった。

⁵ 研究事例4は[10]を一部改変して転載した。

5.1 研究事例4：顔の印象形成における顔向きの影響

日本の伝統芸能である『能』では、同じ能面でありながら、その角度を変えることによって様々な「表情」を伝えようとする。人間の顔もこれと同様に、同じ表情であっても見る角度の変化によって印象が異なるのではないかと考え、観察者の視点と印象形成の関係について実証的検討を試みることにした。

本実験では、心の感情状態に関する印象として「たのしい-かなしい」の次元を用い、顔の形態的特徴に関する印象として「やさしそう-こわそう」の次元を用いた。これは、顔の観察角度の影響が評定する印象によって異なる可能性を見るためであり、そうした違いは一時的な印象と永続的な印象との間で起こりやすいと考えたからである。

なお、顔の動画刺激を用いた先行研究[3]によれば、普通の表情は、下方から撮影された場合よりも上方から撮影された場合に、より悲しく見えることが示されている。彼らの実験では、上下それぞれ40°という比較的大きな角度から撮影された顔が刺激として用いられている。本実験で用いた顔の撮影角度は、上下それぞれ20°と半分の大きさになっていて、こうした緩やかな角度の差による顔の見え方の違いが印象に影響を与えるかどうか注目された。また、本実験では、左斜め45°の上・水平・下から撮影した顔も用いており、視角の違いの影響をより広く検討するものとなっている⁶。

5.1.1 方法

被験者. 大学生男女計60名が後述の12の提示リスト条件に各5名ずつ無作為に配置された。

実験素材. 20代の男性24名の顔を6つの角度から撮影したカラーの顔写真を実験刺激として用いた。これらの顔はいずれも特定の表情を表すものではなく、眼鏡、口ひげ、アクセサリ等は一切つけていないものであった。6つの撮影角度は、正面上、正面水平、正面下、(顔の)左斜め上、左斜め水平、左斜め下であった。左斜めの角度は水

⁶ 著者たちが本研究を行った時点(1998年)では、この先行研究はすでに発表されていたが、残念ながらその存在を知ることなく実験を終了した。したがって、ここで述べられている先行研究との比較は、実験終了後の考察に基づくものである。

平方向に45度であり、上下の角度はそれぞれ垂直方向に約20度に統一されていた。また、垂直方向の3つの角度については3台のカメラで同時に撮影されていた。

実験デザインと手続き。1つの提示リストでは、同じ顔が一度しか現れないようにし、6つの角度条件がすべて同じ回数だけテストされるようにした。さらに、すべての顔が6つの角度条件で均等にテストされるように、合計6つの提示リストを作成した。次に、24名の顔を角度条件が等しくなるように2つのグループに無作為に分け、1つのグループの顔については「1. やさしそう～7. こわそう」の尺度で、残りのグループの顔については「1. たのしい～7. かなしい」の尺度で評定するように被験者に指示した。評定尺度の順序についてもカウンターバランスするために、最終的に12の提示リストを作成し、被験者間で均等に用いた。なお、提示リストの各グループ内の顔は被験者ごとにランダム化して提示した。

被験者は提示される顔の印象を与えられた尺度に基づいて評定するように指示された。各試行において、顔は2秒間提示され、続いて、評定尺度を示す画面が表示された。被験者がここで評定値をキーボードから入力すると次の顔が表示されるというように実験が進められた。始めの12の顔についての印象評定が終わると、評定尺度の変更を示す教示画面が現れ、被験者がここで定められたキーを押すと実験が再開された。なお、顔画像はすべて384x384ピクセルの約32,000色のカラー画像として17インチモニター上に提示された。

5.1.2 結果と考察

評定尺度ごとに、撮影角度を要因とする1元配置の分散分析を行ったところ、「やさしそう-こわそう」の尺度では有意な結果は得られなかった ($F(5, 295) = 1.98, MSe = 1.00, .05 < p < .1$) が、「たのしい-かなしい」の尺度では有意な結果が得られた ($F(5, 295) = 7.37, MSe = .78$)。「たのしい-かなしい」の評定値に対して Tukey-Kramer 法による多重比較を行ったところ、正面上から撮影した顔は、左斜め上から撮影した顔をのぞく他の4つの顔と有意に異なる評定値を示していることがわかった。また、左斜め上から撮影した顔の評定値は、正面水平から撮影した顔の評定値と有

意に異なることがわかった。

これらの結果は、「やさしそう-こわそう」という顔の常態的な印象については顔の角度の影響を受けないが、「たのしい-かなしい」という心の一時的な状態を示す印象については顔の角度による影響があることを示して興味深い。具体的には、相手の顔を上から見る場合、すなわち、顔が下を向いているように見える場合に、顔から受ける印象が「かなしい」方向に揺れることを示している。しかし、下から撮影された顔については、水平方向から撮影された顔に比べて、特に印象が「たのしい」方向に揺れたわけではなかった。つまり、類似の印象評定を用いた先行研究 [3] とは異なり、視角の影響は上から見る場合に限定されていた。この原因が実験刺激に対して用いられた撮影角度の大小の違いによるものなのかどうか現状ではわからないが、少なくとも上方から見た場合の視覚的变化に対して印象形成における感度が高いことが示唆される。

5.2 顔データベースへの要求

上記の実験では、上下3方向については同じ中間表情を同時にとらえていて、印象の差異は顔の角度の差異に一応帰属することができる。しかし、正面方向と左斜め方向の顔については、同じ中間表情が同時に撮影されているわけではなく、印象の違いが見られたとしても、これを撮影方向の違いだけに帰属することはできない。

もちろん、同じ人物が短い時間間隔で同じ表情条件で撮影されていることから、これらの中間表情に大きな差異はないものと期待できる。しかし、研究のそもそもの目的が、顔の微妙な視覚情報に対する人の敏感さを見るものである以上、たとえ僅かな差異であったとしても、その影響がないと安易に仮定することは自己矛盾を意味する。

6 おわりに

様々な人の顔や同じ人物の顔の変化に対する認知処理について実証的検討を加えるためには、ここまで述べてきたように、注意深く厳密に統制された顔刺激を用いる必要がある。しかし、将来の研究の方向性をすべて予測して顔データベースを

構築することは不可能であろう。例えば、眼鏡や口髭などをつけたときの印象の変化を調べようとすれば、そうした条件での顔も事前に同様の条件で撮影しておかなければならない。化粧の濃さの影響を見るとなれば、化粧の度合いを変えたいいくつかのバリエーションを撮影しておかなければならない。しかし、具体的にどのような化粧の濃さの度合いを条件として設定しておけば良いのか事前に予測することは困難である。また、顔の角度の影響を調べる研究に備えるにしても、はたして何度刻みで撮影しておけば良いのかという疑問がでてくる。

こうしたことを考えると、顔データベースとして整備できることには限界があることがわかる。今後は、基礎となる実在の顔をコンピュータグラフィックス技術を用いて変形操作するという方向を一層追及していく必要があるだろう。例えば、顔の角度の問題にしても、3次元レーザ・スキャナーでとらえた顔データを基にすれば、それを3次元空間上で自由自在に回転させることで、様々な角度からの見え方を検証することができる。こうした3次元顔データを実験刺激として用いた研究も既に行っているが [8, 1], データ収録時の顔モデルの負荷を含め技術的な課題があり、特にカラー・テクスチャ情報の質に不十分な点がある。今後、こうした課題が解決され、一つの顔データから様々な現実感のある顔刺激が合成できるようになれば、顔の認知研究もさらに発展を見ることであろう。

謝辞

本稿で紹介した研究事例は著者が客員研究員として ATR 人間情報通信研究所と共同で行ったものである。常に快適な研究環境を提供していただいた一瀬 裕社長、東倉洋一前社長、ならびに赤松 茂第二研究室長に深く感謝致します。

参考文献

[1] 蒲池みゆき・加藤 隆・赤松 茂 (1997). 顔の再認における視点依存性の左右非対称性— virtual view の観点から —. 電子情報通信学会論文誌 A, Vol. J80-A, No. 8, 1273–1278.

- [2] Kappas, A., Hess, U., Barr, C. L., & Kleck, R. E. (1994). Angle of regard: The effect of vertical viewing angle on the perception of facial expressions. *Journal of Nonverbal Behavior*, 18, 263–280.
- [3] 加藤 隆 (1994). 人間の顔の特徴認識. 人工知能学会研究会資料 SIG-J-9402-1, pp. 1–8.
- [4] Kato, T. (1995). Implicit learning of complex visual regularity in faces. *ATR Technical Report TR-H-150*, ATR Human Information Processing Research Labs, Kyoto, Japan.
- [5] Kato, T., Oda, M., Yamaguchi, M. K., & Akamatsu, S. (1995). Facial features and configurations affecting impressions of faces. In Y. Anzai, K. Ogawa, & H. Mori (Eds.) *Symbiosis of human and artifact*, 485–508. Amsterdam, Holland: Elsevier Science B. V.
- [6] 加藤 隆 (1995). 顔の類似性の潜在学習について. 日本心理学会第 59 回大会論文集, p. 640.
- [7] 加藤 隆・山口真美・赤松 茂 (1996). 顔の性識別における特徴要因の顕著性と視点依存性. 電子情報通信学会ソサイエティ大会論文集, pp. 336–337.
- [8] 加藤 隆・阿磨大介・森岡久美子・赤松 茂 (1998). 顔の魅力度判断におけるパーツの魅力の影響, 信学技報, HIP97-53, pp. 17–22.
- [9] 加藤 隆・佐伯昌子・宅間三起・亀井美砂・向田 茂・赤松 茂 (1999). 微妙な顔情報に対する人の感受性について—類似性判断と印象形成の観点から—, 信学技報, HIP99-51, pp. 25–30.
- [10] Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.
- [11] 山口真美・加藤 隆・赤松 茂 (1996). 顔の感性情報と物理的特徴との関連について—年齢/性の情報を中心に—. 電子情報通信学会論文誌 A, Vol. J79-A, No. 2, 336–337.

Recent Research on Paleolithic Arts in Europe and the Multimedia Database

César González Sainz, Roberto Cacho Toca

Department of Historical Sciences. University of Cantabria.
Avda. Los Castros s/n. 39005. SANTANDER (Spain)
e-mail: gonzalec@ccaix3.unican.es / cachor@ccaix3.unican.es

Summary.

In this article the authors present the Multimedia PhotoVR Database, made by Texnai, Inc. (Tokyo) and the Department of Historical Sciences of the University of Cantabria (Spain) about the paleolithic art in northern Spain. For this purpose, it's made a short introduction to the modern knowledge about the European paleolithic art (35000-11500 BP), giving special attention to the last research trends and, in which way, the new techniques (computers, digital imaging, database, physics...) are now improving the knowledge about this artistic works. Finally, is made a short explanation about the Multimedia PhotoVR Database and in which way, these databases can improve, not only research and teaching, but also it can promote in the authorities and people the convenience of an adequate conservation and research of these artistic works.

Key Words: Multimedia Database, Paleolithic Art, Europe, North Spain, Research.

1. Introduction. The paleolithic European art.

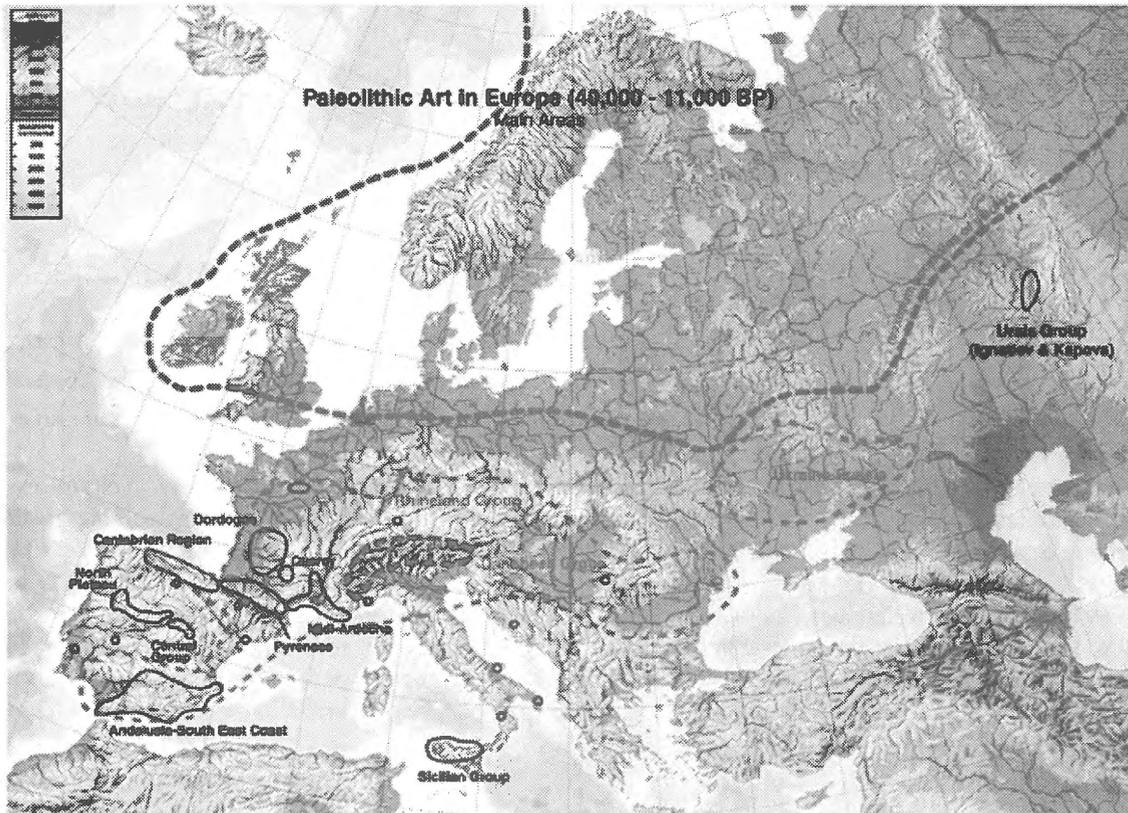
Between approximately 35000 and 11500 years BP, during the last glacial phases, the European continent saw to be born a first artistic cycle of surprising aesthetic achievements. The expressive force reached in the representation of a great variety of wild animals, with some very simple techniques, has been rarely reached in the history of the western art. We find this figurative art in caves, rock-shelters and sites in the open air, and at the same time, on very different objects of the daily life (pendants, spatulas, points of javelin, harpoons, perforated baton, estatuets or simple stone plates). The distribution all over the continent of those varieties is different: the cave art is concentrated in the SW (Spain, Portugal, France and Italy, mainly), while the portable art is much more extended.

This artistic cycle was developed by groups of hunters that lived in more open and cold territories than the current, where were using very various wild resources through the hunting, the fishing and the gathering. Such economic base demanded a relatively high mobility, of people, objects and ideas. This allowed the interaction to long distance, and permitted to fix a graphic style with many common elements all over Europe, in a moment when they were not stable roads, but only



[A goat engraved on a bone , Garma Cave]

migration routes of the herds of ungulates, and when the only one way of transmitting images to distance were the personal adornments, some little



objects that did not hinder the march and, indeed, the retina and the mind of the hunter-artist.

The development of this first figurative art is not only consequence of the greater intellectual capacities of our kind, *Homo sapiens sapiens*, extended at the beginning of the Upper Paleolithic (since 38000 BP in the Iberian peninsula, in the western extreme of Eurasia). But also of some forms of social organization of the groups of hunters more complex and flexible, extended through much more large geographic areas. This demanded the development of formulations to maintain the social cohesion, and for the transmission and exchange of information, more sophisticated than until that moment. The cave art, and the ceremonies to those which –sometimes - could be linked its accomplishment, was a part probably of this set of cohesion tools, for fixing and transmitting understable information for this people. This does not exclude its paper as artistic expression, and of personal and specially, collective affirmation.

The figurative art appears, therefore, within a novelties package that have been experienced by the human groups of the Upper Paleolithic. These novelties are well reflected in the archaeological record in referred what is to the technology (new materials –development of the instrumental on

bone, antler and ivory - and more complex transformation procedures, greater selection of the qualities to work, new technical supports, more various and specialized tools...), but that also affect to the subsistence (in many regions are observed important intensification processes) and that, without doubt, implied also more complex, versatile and differentiated regionally social structures. It is in that novelties context, and of now accelerated cultural change, in which is developed this new capacity –specific of our kind - of creating and reproducing graphic symbols.

2.The paleolithic art of the Cantabrian Region (Northern Spain).

The Cantabrian Region is one of the classic areas of the paleolithic art since the discovery of the paintings of the Altamira Cave between 1876 and 1879. It is a narrow corridor East-West orientated, in the north of the Iberian Peninsula, opened between Cantabrian Mountains, in the south, and the coast line in the north. It is a small region, with a length of 400 km. and a mean width of 40 km, communicated through the eastern extreme with the regions of the SW of France (Dordogne, Pyrenees), with which is verified an intensive interaction during the Upper Paleolithic. On the contrary, the glacier development in the southern

mountains closed the southward communication during great part of that period, specially in the central and western zones of the mountain chain. Nowadays the cantabrian region is organized in several administrative territories, that are from West to East: Asturias, Cantabria, Basque Country and the northernmost Navarra.

This region had an important human population during the Upper Paleolithic due to its good environmental conditions –because the nearness of the sea - and to the abundance of cynegetic, fishing, mollusks, vegetables, etc. resources, available for those groups of hunters-gatherers in some very variable ecological environments and located to short distance. This and the great quantity of caves in the territory allows to understand the abundance of well preserved archaeological deposits, as well as portable and rock art in those caves. The distribution of remains shows a certain concentration of the habitat in the coastal band –with a smoother weather, rich in resources and better communicated - and traffic zones in the inner valleys, between 0 and 200 meters over the sea level. Only exceptionally some caves can reach 500 m of altitude.

Since 1879 a hundred of decorated caves, of different size and importance, have been discovered here. Though only a few can resist a comparison in aesthetic terms with Altamira, there are very important sites like Peña Candamo, Tito Bustillo and Llonín, in Asturias or El Castillo, La Pasiega and La Garma, in Cantabria. These caves have several hundreds of painted and engraved representations with different styles and conventions which belong to successive human occupations along the Upper Paleolithic. Other many cavities shows works of chronology more concrete. Among those which reach greater showiness should cite La Lluera and Covaciella, in the west, Chufín, Las Monedas, Las Chimeneas and Covalanas, in the center, and Santimamiñe, Altxerri and Ekain, in the Basque Country (González Echeagaray and González Sainz, 1994). At the same time, the works of portable art have been multiplied in the archaeological sites of the Upper Paleolithic. The most important sites of portable art mobiliari are the caves of El Pendo, Tito Bustillo, Llonín, El Castillo, Altamira, El Juyo, El Pendo, El Valle, Bolinkoba, Urtiaga and Ekain (Barandiarán, 1995). In these and many more archaeological deposits has been able to study the development of the different phases of the Upper Paleolithic, between 38000 and 11500 years BP, with precision (an updated synthesis in Straus,

1992).

Within the unit that the paleolithic art presents in good part of Europe, the cantabrian region shows some distinctive features. Among them an relatively peculiar iconographic distribution, with many representations of the ungulates here more common: deers and specially hinds, horses, goats, auroch and bisons... and scarce of reindeers, mammoths and animals of especially cold environment, though they are present in different sites. The cantabrian caves show also a certain variety of abstract representations (square-shape signs, claviforms and others) specific of this region and different of the used in other areas of the SW of Europe.

The technical procedures are very similar to those of other European regions, though with some absences (bas-reliefs). Some techniques as like the dotted red outline, or the multiple line engravings, have an important role in the region during some periods. During the Upper Paleolithic was especially characteristic, in a first moment, the outside ensembles based on deep engravings (La Viña, La Lluera and Chufín), or with very simple painted motives (hands in negative of El Castillo, La Garma, Fuente del Salín). Afterwards, the sets of paintings in red with lines sometimes dotted in the Solutrean - 21.000 to 16.500 BP — (caves of Llonín and La Pasiega, Covalanas, Arco and Arenaza, etc.), in many cases with abstract square-shape signs. In Magdalenian (c. 16.500 to 11.500 BP) there is a great variety of expressive and technical resources, and this allows to the artist to make more realistic pictures. So, we find black paintings (caves of Candamo, Cullalvera, Las Monedas, Santimamiñe), or red paintings (specially in abstract signs called "claviform"), engravings (Tito Bustillo, Llonín, Venta de la Perra, Altxerri...), some so characteristic as the inner multiple line engraving used to reproduce hinds' heads, or the addition of several technical procedures, like polychromatic paintings (Tito Bustillo, Altamira, La Pasiega, El Castillo, Ekain...).

3. Recent trends in the investigation of the paleolithic art.

The investigation of the paleolithic cave art has been leaded along the 20th century, by the work of two french researchers: H. Breuil (1954) and A. Leroi-Gourhan (1965, 1983), whose works defined the guidelines of the research. After them, in the last twenty years, research is affected by two



[Negative hands , Garma Cave]

amazing phenomena:

1. The discovery of many decorated caves, especially in some regions of the Iberian Peninsula (in the North and also in Andalusia, in the interior Plateau and in Portugal), and in the classic french regions.
2. The application of the technological and data processing revolution to the documentation and research process of the paleolithic art.

In the Spanish case, the first of those factors is linked with the strong increases of the number of students in the Spanish university during the decade of 1970's. This permitted a multiplication of research centers as well as research and educational staff. At the same time, Spain has experimented social, political and cultural changes that has been accelerated from the last 1970s, in order to get the modernization and democratization of the country. In such context, the economic development of the last decades is supposing a strong impact on the territory (highways, dammings, buildings...) and, consequently, an important number of locations of new decorated

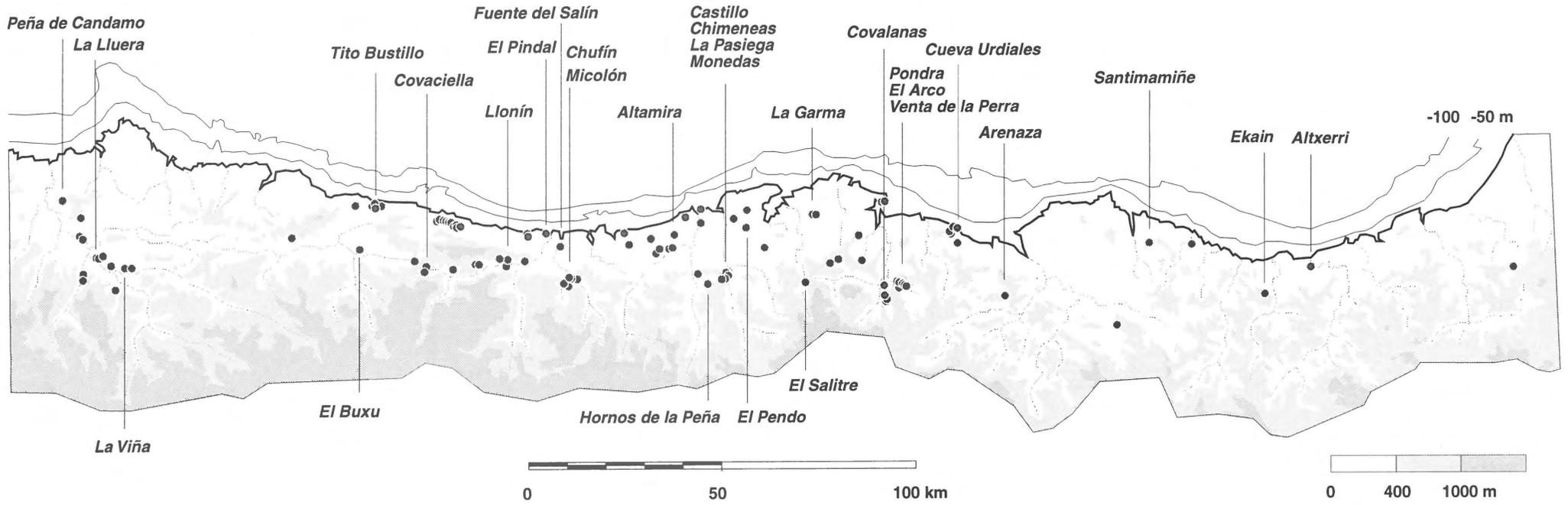
caves. These are of different type:

a) Is of great transcendence, among the recent novelties in western Europe, the documentation of paleolithic cave art ensembles located in the open air. They are stony outcrops in the riversides of some rivers, especially the Duero and its affluents, with very similar parietal compositions to those of the caves (sites of Mazouco, Domingo García, Siega Verde and the ensemble of Foz Côa, in Spain and Portugal) (Balbín et alii 1996, Zilhao et alii. 1997). Other sites on the open air are Piedras Blancas (Andalusia) and Fornols Haut (Eastern Pyrenees, France). These new ensembles have changed the way of understanding this first European art, before linked to the inner areas of the caves, dark and mysterious, and to magic/religious interpretations, that today are very restricted.

b) Among the discoveries of new cave ensembles there are two different situations:

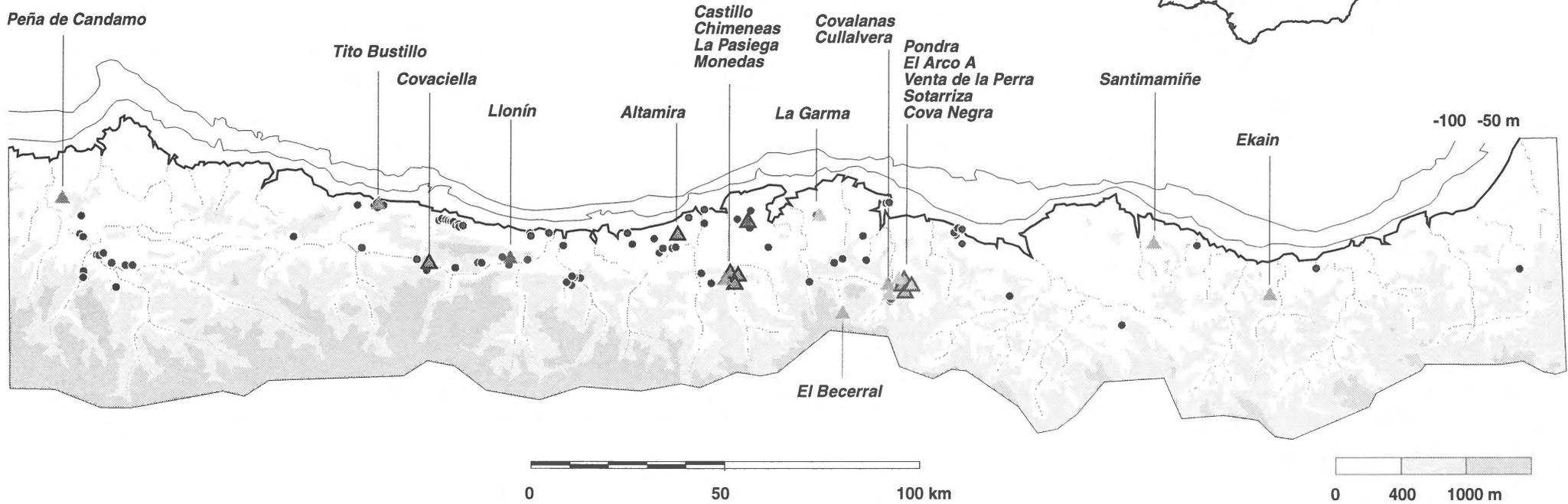
* Ensembles discovered several decades ago, where, a modern lighting, and a most adequate exploration permits to discover artistic evidences, often difficulty visible and not spectacular, or bad

Paleolithic Cave Art in Cantabrian Region

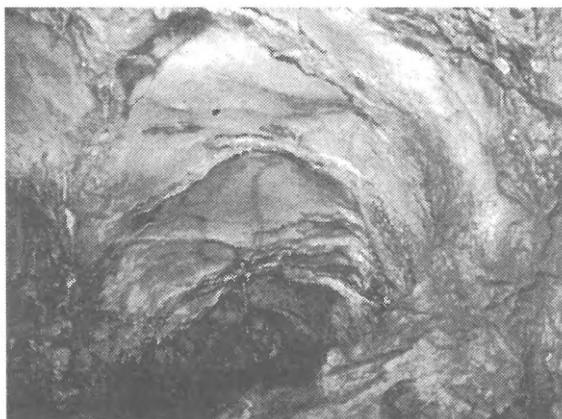


Cave Art in Cantabrian Region (Northern Spain)

Sites with Absolute Dates



- ▲▲ Charcoal Samples and Published Dates (^{14}C - AMS)
- ▲▲ Stalagmite Samples and Published Dates (TL)
- Other Caves With Paleolithic Art



[Red hinds in the ceiling ,Garma Cave]



[A black horse , Garma Cave]

preserved (caves of El Otero, Sovilla, Grande, Arco, Pondra, El Pendo, all in Cantabria, as well as other caves in other parts of the cantabrian region, Andalusia, France, etc.). These sites, generally small and with a few representations, are however very interesting to get a most real vision about the dimension and extent of the artistic phenomenon, in each region, during the Upper Paleolithic.

It is of special relevancy, in the cantabrian region, the wide series of outside ensembles located in rockshelters and caves of the basin of the Nalón river, in Asturias. These sites (La Viña, La Lluera and others) are subject of an important research project of the University of Oviedo, that includes the archaeological excavation of several deposits (Fortea, 1994).

* closed caves by natural processes, and discovered during the public works accomplishment (Covaciella, in Asturias), with sophisticated equipment of speleology (Chauvet, in L' Ardèche, France and La Garma, in Cantabria-) or of diving (cave Cosquer, partially sunken in the Mediterranean sea, near Marseilles, France) (Cosquer et al., 1995). In this case they are very well preserved sites, sometimes with wonderful pictures.

We are implied in the research of one of these sites, La Garma (Arias et al. 1999). The lower Gallery of this mountain is a cave whose entrance was closed by a cave-in at the end of Upper Paleolithic, so now this cave is accessible descending several abyss that join several caves to different height on the mountain. Throughout the 300 m. long of the lower Gallery, it can be found multiple remains of the last human occupations (between 14.000 and 13.600 BP according to the

available C14 dates), consistent in fauna remains, shell of mollusks, tools, charcoals, etc. Furthermore, there are an important number of residence structures in the same surface, including enclosures of stones – which marked the rooms limit and supported skin screens, or foliage - and different entity holes, piles of garbage, some fire... These occupation remains are distributed along a wide surface in the area close to the natural entry to the cave, and also in deeper places. Throughout the walls and ceilings of this gallery there are also a great number of representations of painted or engraved animals (bisons, horses, deers, megaceros, auroch, goats, carnivores...), abstract signs, hands drawn in negative, and multitude of spots of red color of random contours, or series of not- figurative engravings. Such representations correspond, according to the stylistic conventions, technical and other criteria, to different periods of the Upper Paleolithic, at least from the **Gravetiense** (26.000-21.500 ++BP) until central phases of the **Magdaleniense**, when that the cave was closed.

The main interest of the lower Gallery of La Garma –and something similar happens with other intact sites preserved until today - is the exceptional conservation of very assorted classes of remains of the human activity on the floor, and the link among these, or the residence structures well preserved in its perimeter, with the wall paintings. Thus for example, the distribution or rock art and other human remains in La Garma indicates that the engravings and paintings of some low ceilings were accomplished from inside the huts used by the paleolithic. The analysis of these evidences, yet in initial phase, seems to speak about some visits to the deepest regions of the cave more frequent that is traditionally considered, and about more complex and changing motivations of the parietal art along

the Upper Paleolithic.

Until here the main objects of our investigation at present. But, how it is being investigating, and with which ideas? Close to perspectives much more eclectic about interpretation, currently the effects of the technological revolution underway in the study of the paleolithic art are very important (Lorblanchet, 1995). The modern lighting procedures, instant photograph, photogrammetry, topography and informatic record of the spatial dispersion of remains, digital databases, digitizing and treatment of images with computer, chemical analysis of the pigments, experimental analysis of the technical work procedures etc., are changing each phase of the research process. Among the novelties has reached great importance the possibility of dating by C14, through the Accelerator Mass Spectrometer (AMS) some pictures painted with organic material, normally charcoal. At present we have several tens of dates in cavities of France and the cantabrian region (Valladas et al., 1992; Moure et al., 1997). Some of those results are making more precise the chronology traditionally considered, especially for the most ancient phases of the Upper Paleolithic, or pre-magdalenian, and specially, to calibrate the procedures of traditional dating and to decide what continues being valid. In the last two years, it has been begun to test, furthermore, the dating of travertine crusts associated with parietal compositions—superposed or underlayed to the representations—, through Thermoluminescence and other procedures as the Uranium/Thorium (U/Th). Many of the results of these new procedures, and of the research of the last decade, are exposed in a way short in different contributions to International Newsletter on Rock Art, edited in France by the Committee International d'Art Rupestre—CAR-ICOMOS— and other institutions.

The projects guided to interpret the meaning of the paleolithic art have tended to reflect, unflinching, the changes of mentality and of way of thinking and focusing the past that have been in our society (and that continue being produced at present). Today the prehistorian tend to doubt of global or simple explanations, valid for all the widest period in the different European regions (like the theories that link the representation with the magic of the hunt, with the expression of essential mythologies, or with rituals guided to the maintenance of the social cohesion, or to the transmission of information about the hunt), since these theories are not, necessarily, mutually excluding. At present time is tended to an analysis



[A panel with several red paintings: aurochs, irish elk and caprids, Garma Cave]

based on the comparison of the different affected regions, and of the different epochs of the Upper Paleolithic, as procedure to try of understanding something of the role played by the art, or its meaning. In the last analysis, it is considered to understand the artistic phenomenon inside the cultural system where it's generated—and it have sense -, or in relationship to other aspects of the system, archeologically registered and more easily interpretable (analysis of the spatial distribution in different levels, of the subsistence strategies and of the formulations of organization and mobility of the human groups).

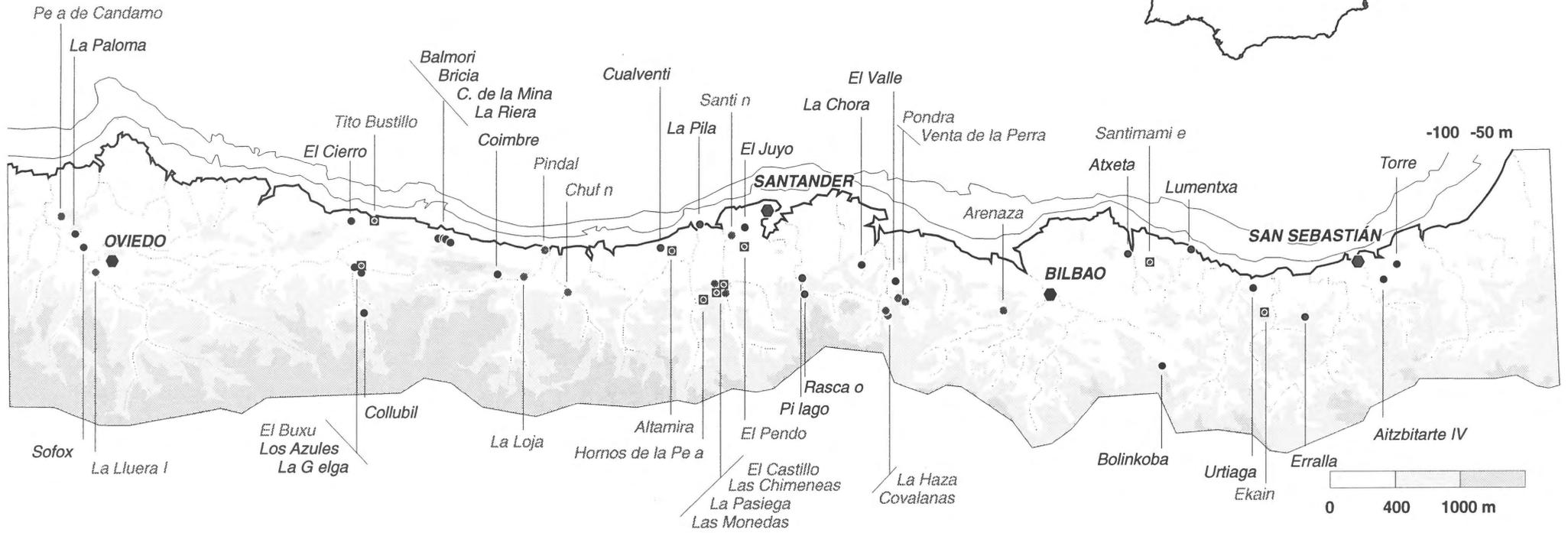
4. Multimedia PhotoVR Database on the cantabrian paleolithic art

In this context of the research the databases with digitized images have a great interest, because it permits a knowledge extraordinarily more immediate and, in many cases, accurate, for example, of the paleolithic art. And indeed it permits a diffusion all over the world of something before reserved to a few researchers of the implied countries (France, Spain and Italy mainly).

4.1. Accomplishment.

The multimedia database that we present now was accomplished by Texnai, Inc., of Tokyo, and directed by Takeo Fukazawa, cooperatively with the authors of this article, both researchers of the Department of Historical Sciences of the University of Cantabria (Spain). The accomplishment of this database had financial support of the Information Technology Promotion Association Association (IPA), an agency of the Ministry of International Trades and Industries of Japan. This database, will be available in several museums and research

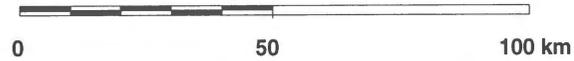
Multimedia PhotoVR Database
Paleolithic Art in Cantabrian Region (Northern Spain)



—19—

Cantabrian Region: Upper Paleolithic Sites Included in the Database

- Portable Art
- Cave Art
- ◻ Portable + Cave Art



centers of Japan (University of Tokyo, National Museum of Ethnology, in Osaka, and the International Reserch Center for Japanese Studies, in Kyoto). The spanish version is already available in the University of Cantabria, Museum of Archaeology and Ethnography of Oviedo, Regional Museum of Prehistory of Santander, Service of Historical Heritage of Biscay (Bilbao) and Sciences Society Aranzadi, of San Sebastián.

This work was organized and planned around the end of 1997. For the accomplishment, two teams of photographers and engineers of Texnai were working in the caves and museums of the north of Spain during the spring of 1998. They were accompanied and advised by a scientific team of research students of the Department of Historical Sciences, integrated in the project as scholarship holders. At that time, the main researchers defined the structure of three related databases: cave art, portable art and landscapes, giving content to a wide number of information normalized on each stony canvas or parietal representation, decorated object and landscape sights. Since June, 1998, and until February of 1999, we have been drafting the descriptions of each image, complimenting those fields in the University of Cantabria (Spain), and, at the same time, the graphic works were elaborated in Tokyo, where finally everything has been mounted and integrated in the database. The spanish version was ready in March of this same year.

4.2. Objectives.

We have attempted to make a multimedia product that facilitate a wide and updated introduction to the art of the hunters of the Cantabrian Upper Paleolithic, by using modern technics of virtual reproduction through an integrative and plural approach. In this way, the Cantabrian Region is shown as a geographical, natural and cultural unit during the Upper Paleolithic, with some inner changes that we have treated to specify and emphasize. On the other side, we intended to surpass the image of the regional paleolithic art reduced to Altamira and some other caves. Thus, besides Altamira, we include other 21 parietal ensembles, trying to offer a complete image of the artistic phenomenon, including large ensembles (Peña Candamo, Tito Bustillo, El Castillo, La Pasiega, Ekain...) and other little ones like La Loja, Pondra or Venta de la Perra. These little ensembles are very important to understand the real dimension of the paleolithic cave art.

Furthermore, more than one hundred decorated objects are included in the database. These images were shot in several museums of the region and it allows to get an integrated vision of portable and cave art. Finally, it is offered also an image of the landscape environment where the caves are located. In this way the database integrates several hundreds of landscapes pictures throughout the cantabrian region, from the north of Navarra to the middle basin of the Nalón river (including panoramic images of 360°, for example, from the top of Peña de Candamo, the Macizo de Ardines, in Ribadesella, the Castillo mountain, in Puente Viesgo, Ekain mount, etc.).

4.3. The virtual reproduction.

Concerning the multimedia, the database includes conventional images (still picture and video) and other virtual reality files such as IPIX (spherical photographs) that allows to visualize a three-dimensional space in all directions and the sensation of being inside the cave, looking around and watching details of the decorated walls of the cave. Other kind of digital files are the QTVR objects, that allows to manipulate the paleolithic objects in the screen, and to examine its decoration or the techniques of manufacture with a great detail. The third kind of VR images is the QTVR panorama, which offers panoramic photos of 360° horizontally, that has been used for some landscapes. In round numbers this database includes about 200 IPIX files, 300 QTVR objects, 200 QTVR panorama and more than 1000 still pictures.

4.4. Possibilities of use.

The advantages of these techniques for the researchers and studios are not reduced to allow the access to caves - or to parts of caves —closed to the public, or very difficult to walk inside. Furthermore, it shows us the cave art in its ambient inside the cave. This is very important, since cave art don't use standard supports (like canvas or flat wall), but panels of different dimensions and directions in walls with irregularities and different conditions of dampness, hardness, etc. Sometimes these techniques of photographic virtual reality permit to appreciate and to understand the fitting of the paleolithic rock art to those conditions. Regarding portable art, usually people can see only one face of the object, through the crystal of a showcase in a museum. The QTVR objects of the database permit to manipulate in the screen the virtual object, and it

allows a very detailed analysis of the decoration in the better conditions (without quiver the hand and the magnifier, and choosing the better direction of the light in each moment...).

The organization of the material - different types of images and texts - in a database, permits to search and filter the information through several defined fields (including texts and logic fields), or combinations of these (representations, techniques, supports, museum, cave, , chronology, location, etc., for example: magdalenian black bison, in all the region or in a selected cave). Furthermore, different maps of the region, of each cave, stratigraphies, etc., are connected to the images, explanatory texts, traces of portable art, and available bibliography on each piece or each art ensemble. The ready texts are not limited to comment each image, but are included updated articles about each cave, the cantabrian region and other artistic areas of the European west, and more general approximations to the paleolithic artistic phenomenon. All this converts this database in an excellent support for teaching –that we are already using for teaching in the University of Cantabria- as well as for researching.

5. The role of the art of the paleolithic hunters in our civilization.

The works of cave art distributed along many caves of the north of Spain and other European regions, or the wonderful miniatures worked on great number of objects, offer today to our society a double permanent lesson:

- they are the live testimony of some beliefs and some forms of already disappeared life, primitive and, at the same time, very respectful with the territory and the environment, from where they got a wide diversity of wild resources, through the hunting, fishing and gathering.

- the spectacularity and artistic and technical mastery of many of the paleolithic representations, constitutes a moderation or regulatory factor, of our technified, complex and superb current civilization, that indicates us the way that the human nature procures to reach very high levels of artistic

expressivity and wisdom even in such simple and little developed cultural systems.

The decisive increase of the human action on the territory in the last decades makes increasingly necessary the vast documentation, and the development of a more active conservationist conscience. The diffusion of the knowledge that facilitate the multimedia databases can help also, as we believe, to promote such conscience and to sensitize to the authorities, and to the people, on the convenience of an adequate conservation and research of these manifestations.

Gratefulness.

We want to thanks to all the people who have worked in this project. Firstly, we thank to Mr. Takeo Fukazawa, president of Texnai, Inc. During more than one year we were in a permanent communication Spain-Japan through e-mail, trying to define the main lines of this project, beginning with an idea that, firstly, we saw as a *dreamt*. After defining the project we begun a hard bureaucracy works in Japan and Spain in order to get financial support and permissions for shooting inside the spanish caves and museums. Without the support of the IPA, of the Ministry of International Trade and Industries of Japan, and the administration of Asturias, Cantabria, Vizcaya and Guipuzcoa, in North Spain, we could not accomplish this project.

Afterwards, we could join a great japanese-spanish team of photographers, engineers and archaeologist who were working together in Spain during more than two month. They have a great responsibility in this project and we need to give thanks to everybody of them: Yuji Seki (Museum of Ethnology, Osaka), Norie Hiraide, Hiroaki Seki, Yuki Fukazawa, Shin Hamazaki, Peter Sun, Yoshinori Matsumoto (Texnai, Inc.), Koji Nagasaka (Nagasaka Video Service), and Nerea Gálvez, Ignacio Castanedo, María José Samperio and Gustavo Trueba (University of Cantabria). Finally, we want to thanks to Luis Teira the making of the ilustrations included in this article.

BIBLIOGRAPHY

Arias, P.; González Sainz, C.; Moure, A.; Ontañón, R. 1999. *La Garma. Un descenso al pasado*. Catálogo de la exposición. Gobierno de Cantabria y Universidad de Cantabria. Santander.

- Balbín, R. de; Alcolea, J.J.; Santonja, M. 1996. *Arte Rupestre Paleolítico al aire libre de la cuenca del Duero: Siega Verde y Foz Côa*. Fundación Rei Afonso Henriques, Zamora.
- Barandiarán, I. 1994. "Arte mueble del paleolítico cantábrico: una visión de síntesis en 1994". *Complutum* 5, pp.45-79.
- Breuil, H. 1952. *Quatre cents siècles d'art pariétal. Les cavernes ornées de l'âge du renne*. Centre d'études de de documentation préhistoriques, Montignac. (Reimpresión, Max Fourny, Paris 1974).
- Cacho Toca, R., Gálvez Lavín, N. (1999): Some applications of digital photography in the enhancement and reproduction of paleolithic rock paintings. En: BARCELÓ, J. A., BRIZ, I., VILA, A.: *New Techniques for Old Times CAA 98*. Proceeding of the 26th Conference, Barcelona, March 1998. BAR International Series, 757: 73-76
- Chauvet, J.M.; Brunel, E.; Hillaire, C.; Clottes, J. 1995. *La Grotte Chauvet à Vallon-Pont-d'Arc*. Editions du Seuil, Paris.
- Fortea Pérez, J. 1994. "Los "santuarios" exteriores en el Paleolítico cantábrico". *Complutum* 5 pp.203-220
- González Echegaray, J.; González Sainz, C. 1994. "Conjuntos rupestres paleolíticos de la cornisa cantábrica". *Complutum* 5, pp.21-43.
- González Sainz, C., Cacho Toca, R. y T. Fukazawa. "PhotoVR Multimedia Database: Paleolithic Art in North Spain". *International News on Rock Art*, en prensa.
- Leroi-Gourhan, A. 1965 (2ª ed: 1971). *Préhistoire de l'art occidental*. Lucien Mazenod, Paris
- Leroi-Gourhan, A. 1983. *Los primeros artistas de Europa. Introducción al arte parietal paleolítico*. Encuentro, Madrid.
- Moure, A.; González Sainz, C.; Bernaldo de Quirós, F.; Cabrera, V.; Valladas, H. 1997. Nouvelles dates absolues de pigments dans les cavernes cantabriques. *International News on Rock Art* 18, pp.26-29
- Lorblanchet, M. 1995. *Les grottes ornées de la Préhistoire. Nouveaux regards*. Errance, Paris.
- Straus, L.G. 1992. *Iberia before the Iberians.. The Stone Age Prehistory of Cantabrian Spain*. University of New Mexico Press, Albuquerque.
- Valladas, H.; Cachier, H.; Maurice, P.; Bernaldo de Quirós, F.; Cabrera Valdés, V.; Uzquiano, P.; Arnold, M. 1992. "Direct radiocarbon dates for prehistoric paintings at the Altamira, El Castillo and Niaux caves". *Nature* 357, pp.68-70.
- Zilhao, J. (coord.) 1997. *Arte Rupestre e Pré-Historia do Vale do Côa. Trabalhos de 1995-1996*. Ministério da Cultura. Lisboa.

PhotoVR考古資料データベース

『北スペインの旧石器洞窟美術』

日本・スペイン産学共同プロジェクトの実現と諸問題

深沢武雄

株式会社テクネ代表 / 東京都渋谷区宇田川町2-1、No.1008

e-mail: fukazawa@texnai.co.jp, url: http://www.texnai.co.jp

PhotoVR Archeological Image Database

Paleolithic Cave Arts in North Spain

On its co-project between a Japanese private company and a Spanish university

Takeo Fukazawa

Texnai, Inc. President / 2-1 Udagawa-cho, Shibuya-ku, Tokyo

概要：

人類の最も原初的な象形文化遺産のひとつとしてヨーロッパ後期旧石器時代の洞窟美術がある。中でもアルタミラやラスコーなどスペイン北部からピレネー山麓、フランス南西部にかけて集中する石灰岩洞窟に見られる彩色画や線刻画、ならびにその一帯から出土する装飾骨器など各種動産美術は、人類の象形文化あるいは精神文化の起源を探る上で極めて重要な意味を有している。本データベースは、かかる最も原初的な象形文化の一典型としてスペイン北部の旧石器洞窟美術をとりあげ、主としてバスク、カンタブリア、アストリアス地方に分布する壁画包含洞窟22所とその周辺景観ならびに各自治州考古学博物館収蔵の動産美術を、それぞれ対話型球面パノラマ、全周画像ムービーなど最新のPhotoVR生成技術によって情報化したもので、現地カンタブリア大学哲文学部歴史科学科ならびに株式会社テクネとの産学共同プロジェクトとして1997年9月から1999年3月までの約2年半余を要して実現したものである。本稿では、本データベースの概要、プロジェクト推進の経緯、PhotoVR技術について報告し、この種の人文系データベース構築に関わる社会的ならびに技術的諸問題について提起する。

[キーワード]

PhotoVR, QTVR, IPIX, 旧石器美術、洞窟壁画、動産美術、カンタブリア、考古資料データベース

Summary :

An image database using the latest PhotoVR technologies "Paleolithic Cave Arts in North Spain" was completed this year as a result of co-project between a Japanese private company and a Spanish state university spending two and half years since 1997. The database consists of three different kinds of sub-databases that includes a Cave art database, a Mobil art database and a Landscape database so to allow us to verify the relationship between those different categories, or to identify the atmospheres where Paleolithic peoples lived and created those arts. Where the PhotoVR means "Photographic Virtual Reality" and adding to thousands of still pictures, we shot about 200 spherical panoramas VR at 23 wall painted caves and its environs in Cantabrian regions to create the cave arts and the landscape databases, and more than 300 QTVR object movies for the Mobil art database at 5 archeological museums located in the same regions.

The total image record registered in this database exceeds about 1200 and when including the variations and additional images such as cave plans, Mobil arts traces, and stratifications, the total number of images exceeds 3000. Most images are those which are opened to the public for the first time in Japan by this database.

In this paper will be reported the process how this co-project could be realized as well as an overview of the database including the PhotoVR technologies. Paleolithic cave arts can be said to be one of the most important cultural heritages created by humanity as an origin of spiritual expression by images and symbols. People's more interests and studies on this field will be appreciated to be promoted by this database.

[Keywords]

PhotoVR, QTVR, IPIX, Paleolithic cave Arts, Rock arts, Mobil arts, Cantabria, Archeological Database

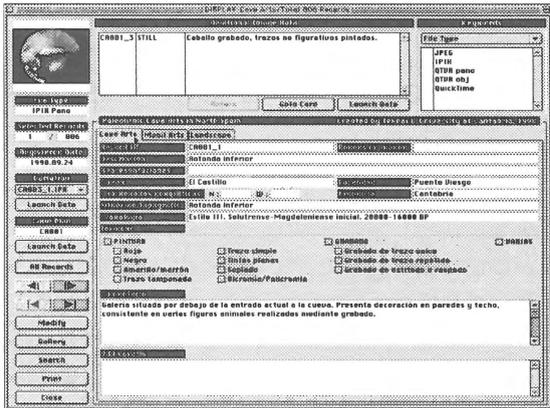
1. データベース概要

高度情報化社会の進展にともない、近年、文化財ならびに学術資料の情報化促進の気運が高まりつつある。情報産業の一翼を担う弊社としては、これを新市場創出のひとつの契機と捉え、主として考古資料の情報化とそのデータベース化に関する技術開発に取り組んできた。中でも象形資料の情報化については、従来の写真計測の発展型として対話型のPhotoVR生成技術[註1]に着目し、立体資料についてはその全周を対話的に観察可能なQTVRオブジェクトムービー[註2]、景観に関しては視点を取りまく空間を全て表示可能な各種球面パノラマ画像の実用化をめざしてきた[註3]。特にQTVRオブジェクトムービー生成技術については、データベース構築に伴う情報化効率の問題を解決するため、コンピュータ制御による自動全周撮影装置を考案し[註4]、縄文土器を素材とした長野県富士見町井戸尻考古館における撮影実験をかかわりに、東京大学総合研究資料館(現博物館)における同館収蔵学術資料の画像情報化ならびに国立民族学博物館収蔵のモンゴル民族資料の情報化などでその実用性が確認された。ヨーロッパ旧石器美術を対象とした本データベースのプロジェクトも、こうした一連の実証的技術開発の延長線上に発案されたもので、考古資料としての動産美術は桌上型全周撮影装置の格好な素材であり、また彩色画や線刻画を包含する旧石器洞窟空間は球面パノラマにとってはまたとない魅力的な素材とみなされた。

とはいえ、この種に学術資料の情報化は、単に技術的あるいは商業的な関心のみでは実現できない。そこには、対象分野に対する採算を超えた当事者の思い入れが肝要であり、この点についてもヨーロッパ旧石器美術は、人類が築き上げてきた形象文化の源流をなすものであり、イメージに携わる私どもにとっては最も興味をそそる素材のひとつであった。

こうして本プロジェクトは、ヨーロッパの後期旧石器美術に焦点をあて、その調査・研究の一拠点でもあるスペインのカンタブリア大学との共同プロジェクトとしてスタートした。

ヨーロッパ後期旧石器美術に関する一次資料としては、洞窟美術と動産美術が知られている。洞窟美術は、主として動物や人の彩色画や線刻画、あるいは幾何学的なシンボルなどで、洞窟の内奥壁面ある



〔洞窟美術 データウインドウ〕

いは開口部の岩肌を描かれている。動産美術は、裝飾骨器、裝飾石器、人像の彫刻など可搬性の遺物で壁画包含洞窟の内部またはその近傍から出土している。双方とも南は地中海沿岸からイベリア半島、北はピレネーを経て南フランスから東欧にかけて広く分布しているが、特に顕著なのはアルタミラ洞窟を擁する北スペインのカンタブリア地方とラスコー洞窟を擁するフランスのドルドーニュ地方であり、それぞれ180から200余の壁画包含洞窟が確認されている。本プロジェクトとしては、旧石器美術データベースの端緒を築く試みとして、まずはカンタブリア大学の地元である北スペインのカンタブリア地方、つまりギプスコア、ビスカヤ、カンタブリア、アストリアスの4自治州に範囲をしばり、そこに分布する主要な壁画包含洞窟23ヶ所ならびに各自治州考古博物館収蔵の動産美術から約150点を選んで資料撮影が開始された。撮影には、日本からは10名、カンタブリア大学からは5名が参加し、洞窟班と動産美術班の2チーム同時併行で延べ90日を要した。

既に触れた通り、本プロジェクトでは、洞窟内部にあつては壁画のしめる位置関係、あるいは旧石器人の行動環境をできるだけリアルに伝えるべく、魚

眼レンズを用いた球面パノラマ撮影を敢行した。また動産美術に関しては立体資料の全周をくまなく観察可能にすべく、小型の自動全周撮影装置を現地に運び、QTVRオブジェクトムービーの博物館撮影を行なった。この他に、通常のスチル写真とビデオ撮影が行われ、これらをデジタル処理してマルチメディア・データベースとして編集した。

データベース管理プログラムは、ACI社の4th Dimensionをカスタマイズし、使用目的を教育・研究・啓蒙と定めてスタンドアローン型でできるだけグラフィカルなユーザーインターフェースの実現を

めざした。構造としては互いにフィールド構成の異なる3つの画像データベース（洞窟壁画データベース、動産美術データベース、景観データベース）から成り、共通フィールドを媒介とした統合検索を可能とした。画像素材は、洞窟内外の球面パノラマ2000点、QTVRパノラマムービー2000点、高精細JPEG画像2000点、洞窟平面図22点、層位図30点、動産美術のQTVRオブジェクトムービー3000点、高精細JPG画像3000点。レコード総数は約1200となった〔註5〕。

[表1] 資料撮影場所一覧

地区	旧石器洞窟ならびにその周辺景観	動産美術収蔵博物館
アストリアス 自治州	テイト・ブスティーヨ 洞窟 エル・ピンダル洞窟 ラ・ロハ 洞窟 ラ・ジュエラI 洞窟 ラ・ペニャ・デ・カンダモ 洞窟 エル・ブシュー洞窟	オビエド考古学博物館
カンタブリア 自治州	エル・カステイーヨ 洞窟 ラス・モネダス洞窟 ラス・チムネアス洞窟 ラ・バシエガ洞窟 ラス・コバラナス洞窟 ラ・アーザ洞窟 チューフィン洞窟 サンティアン洞窟 エル・ペンド洞窟 ポンドラ洞窟 オルノス・デ・ラ・ペーニャ洞窟 アルタミラ洞窟	サンタンデル考古学博物館 アルタミラ洞窟研究博物館
ビスカヤ 自治州	サンティマミネ洞窟 ヴェンタ・デ・ラ・ペラ洞窟 アレナツァ洞窟 エカイン洞窟	ビルバオ考古民俗学博物館
ギプスコア 自治州	エカイン洞窟	サン・セバスチアン自然科学協会

[表2] 主要フィールド構成一覧

洞窟美術データベース	動産美術データベース	景観データベース
[内部データ] 参照画像 (ピクチャー) レコードID (テキスト) 名称 (テキスト) 画像表現 (テキスト) 最大長 (数値) 洞窟名 (テキスト) 地区名 (テキスト) 州名 (テキスト) 緯度軽度 (数値) 地形 (テキスト) 文化区分 (テキスト) 制作技法 彩色画 (論理 x 8) 線刻画 (論理 x 4) その他 (論理 x 1) 解説 (テキスト) 文献 (テキスト) ファイル形式 (テキスト) 登録年月日 (日付け) キーワード (論理 10 x 12) [外部データパス] 関連画像・形式・パス バリエーション画像パス 洞窟平面図パス *** [外部データ形式] JPEG 画像 (1024x768) IPIX Spherical (640x480) QuickTime Movie(320x240)	[内部データ] 参照画像 (ピクチャー) レコードID (テキスト) 名称 (テキスト) 画像表現 (テキスト) 寸法 (数値) 収蔵場所 (テキスト) 収蔵番号 (テキスト) 出土洞窟名 (テキスト) 地区名 (テキスト) 州名 (テキスト) 緯度軽度 (数値) 地形 (テキスト) 出土地層 (テキスト) 文化区分 (テキスト) 型式 (テキスト) 制作技法 (論理 x 8) 解説 (テキスト) 文献 (テキスト) ファイル形式 (テキスト) 登録年月日 (日付け) キーワード (論理 10 x 12) [外部データパス] 関連画像・形式・パス バリエーション画像パス トレース図パス *** [外部データ形式] JPEG 画像 (1024x768) QTVR Object(640x480)	[内部データ] 参照画像 (ピクチャー) レコードID (テキスト) 名称 (テキスト) 地区名 (テキスト) 州名 (テキスト) 近傍洞窟名 (テキスト) 撮影年月日 (日付け) 分類 (論理 x 10) 解説 (テキスト) 文献 (テキスト) ファイル形式 (テキスト) 登録年月日 (日付け) キーワード (論理 10 x 12) [外部データパス] 関連画像・形式・パス バリエーション画像パス *** [外部データ形式] JPEG 画像 (1024x768) QTVR Panorama(540 x 240) IPIX Spherical (640x480) QuickTime Movie(320x240)

[表3] ユーザーインターフェース機能一覧

検索ウインドウ	一般統合キーワード検索、テキスト検索
データウインドウ	データ表示、外部画像表示、印刷
ギャラリーウインドウ	ギャラリー表示 (画像形式、洞窟名、地域、ファイル名表示切換)
リストウインドウ	参照画像と主要フィールドのリスト表示・印刷
画像表示ウインドウ	JPEG、QTVR Object、QTVR Pano、IPIX、Quicktime Movie
地図検索ウインドウ	クリックブルマップキーワード検索
概説表示ウインドウ	旧石器美術概説、洞窟美術概説、動産美術概説、個別洞窟解説

2. 日・西産学共同プロジェクト実現の経緯

本プロジェクトの体制としての特色は、スペインにとっては第一級の文化遺産である旧石器洞窟美術を対象とし、その現地撮影からデータベース化までを、日本の、それも一民間小企業とスペインの州立大学が共同して実現した点にある。そこには予備調査、両者間の共同開発契約、現地4自治州にまたがる撮影許可申請、著作権問題、経費の工面などプロジェクトの開始以前に解決すべき課題が山ほどあるわけであるが、それらを発案からわずか6ヶ月という短期間で処理しえたことも特筆に値するといえる。また本プロジェクトを可能にしたのは、ひとえにスペインの大学、関係省庁、洞窟ならびに博物館関係者のあたたかい理解と協力によるものであるが、それを現実のものとした背景として特にインターネットの普及をあげておきたい。

■セビリアからの電子メール：

インターネットは、既に広く認められている通り、地球全体をひとつの巨大な情報通信システムとして政治的国境を超えた別次元の人間社会を構築しつつある。私どもとしても、その新世界の一員として、技術開発から商取引にいたるまで、インターネットは今や欠かせない存在となっているわけであるが、本プロジェクトの提携先であるカンタブリア大学とのコンタクトも私どものホームページと一通の電子メールが最初のきっかけとなった。

私どもが、既にQTVRに関する制作技術の実証テストを終え、その成果をホームページを介して公開しはじめた1997年の4月の半ば頃である。毎朝、いつも通りあけるメールボックスに、セビリアのホセ・マリアというフォトグラファーからQTVRパノラマムービーの撮影技術について問い合わせのメールが届いていた。米国アップル社の開設した情報交換のメールから私どものURLを知ったのがきっかけだった。以来、PhotoVRの技術動向に関するメール交換が暫く続き、その過程でスペインの洞窟美術を球面パノラマで撮影できればオモシロイということで意見が一致した。また、この際、ヨーロッパ旧石器美術全体のデータベース化まで実現できれば、学術的にも意味あるものとなるのではないかと夢が広がった。

■球面パノラマVR



[旧石器洞窟での球面パノラマ撮影]

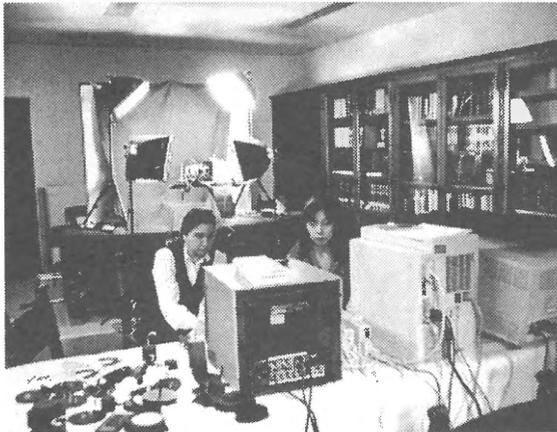
当時、魚眼レンズを使った球面パノラマとしては、前後2枚の魚眼写真から合成処理されるIPIX形式と3枚から合成されるSmoothMove形式のVR画像が知られていたが、暫くしてアンダルシアの鍾乳洞でテスト撮影されたSmoothMove形式の球面パノラマがホセ・マリアから送られてきた。光量が不足していたためかなり荒れた画像ではあったが、照明装置を工夫すればなんとかいけるとの感触をもった。残るは本物の旧石器洞窟の撮影である。そのためには専門の研究者の協力が不可欠であった。

■カンタブリア大学とのコンタクト

本プロジェクトの立役者となったカンタブリア大学のロベルト・カチョ研究員とコンタクトできたのは、それから1週間ほどたった6月中頃のことである。ホセ・マリアがスペイン中のホームページを検索し、スペイン北部で旧石器洞窟地帯の中心地でもあるカンタブリア州サンタンデルのとある大学のホームページが眼にとまったが、そのウェブ・マスターがロベルト・カチョであった。ロベルト・カチョがどんな研究者かは知るよしもないが、球面パノラマやQTVRを使った旧石器洞窟美術のデータベース化については間もなく理解を示し、とにかく他の研究者にも話をしてみようということになった。

■マルチメディアコンテンツ制作支援事業

国立民族学博物館特別展の企画でモンゴルでのパノラマ取材があり、その間、暫く音信は中断されたが、8月の末、マルチメディアコンテンツ振興協会(MMCA)から連絡が入り、6月にあらかじめ申請しておいた本プロジェクトについてヒアリングを受けることになった。その年は、通商産業省もマルチメディアコンテンツ産業の育成に本格的に乗り出

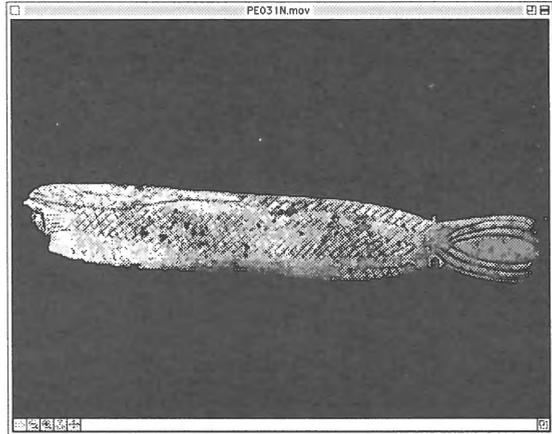


[考古博物館に於ける動産美術の QTVR 撮影]

し、情報処理振興協会（IPA）ならびにMMCAを介して、良質なコンテンツ制作のための資金的な支援事業を開始していたのである。この種のヒヤリングは、私どもにとっては、初めての経験ではあったが大した波乱もなく、カンタブリア大学との共同制作体制、撮影許可の可能性、権利関係、監修体制など基本的事項の確認が問われた程度であった。

■サンタンデル訪問

一方、サンタンデルでは、スペイン日本文化研究協会主催による日本文化シンポジウムが9月中旬に予定されており、ロベルト・カチョの友人で大学図書館長のアントニオ・サントスの計らいで、急遽、そこで日本における文化財情報化の現状ならびに本プロジェクトの主旨などについて私が話をすることになった。これが初のサンタンデル訪問となったが、到着した翌朝、それも講演に向う直前に東京からファックスが入り、マルチメディアコンテンツ制作支援事業採択の知らせを受け取った。これで資金調達にもめどがつき、がぜんプロジェクトは現実味をおびてきた。シンポジウムでは、お茶、生け花、折り紙などいわゆる日本の伝統文化に関する講演の続く中で、コンピュータを使った情報化の話はやや異質ではあったが、ことが地元の洞窟美術に関わるプロジェクトだけに一般市民に混じって考古専門の研究者の姿も見え、地元の新聞記者やテレビ局のカメラも入っての熱気のある講演となった。こうして夕方のテレビニュースでは、数回にわたって洞窟美術情報化の計画がオンエアされ、翌朝の新聞にも「カンタブリアの旧石器洞窟をPhotoVR撮影」と見出しがおどった。ロベルト・カチョの師事するセサル・ゴンザレス教授との初の会見も、こうした背景のもとで極めて順調に進み、現実的に約1年間でどれだけの数の洞窟美術がデータベース化できる



[動産美術の QTVR 表示]

かという具体的な計画にまで話が及んだ。この日、大学では学部長レベルでの共同開発に関する基本合意をとりつけ、年内には実施計画をまとめ、それを基に正式契約の文案を作成する運びとなった。

■洞窟撮影テスト

壁画包含洞窟内での魚眼撮影テストもサンタンデル訪問の欠かせない目的の一つだった。用意した機材は、等距離射影方式のNIKON8mmレンズとNikon F3、それに魚眼レンズ専用の回転雲台、三脚、ストロボ2個。テスト撮影は、まずカンタブリアとアストurias州境の山岳地帯にあるチューフィンという洞窟で行われることになった。

チューフィン洞窟は、峡谷の溪流をせき止めてできた細長いダムに対岸にあり、棒杭ひとつの木陰の船着き場から少し急斜面を上った石灰岩の岩場の底をえぐるように横長の裂け目を見せていた。開口部の岩棚では、馬や鹿の形を刻んだ線刻画がまず幾つか認められた。しかしながら、その輪郭はきわめて不鮮明で、洞窟壁画の撮影がそう容易でないことが思い知らされた。

洞窟は、いきなり天井の高さが60センチにも満たない横長の裂け目からはじまった。初めての者としては、天井の出っ張りに気を取られながらやっとの思いでガイドの尻を追った。約30メートルばかり暗闇をはっていったところで急に眼の前がひらけ、広々とした空間が僅かな懐中電灯の光に乳白色に浮かび上がった。洞窟の壁は全て鍾乳石でおおわれ、ガイドの指さす岩の窪みに、今度はかなり鮮明な、湾曲状に描かれた赤い斑点状のかたまりが見えてきた。岩棚をよじ登り、近づいてよくみると血の色に似たなんとも不思議な点列。私にとっては、この赤い点列が、肉眼で初めてみた旧石器時代の彩色

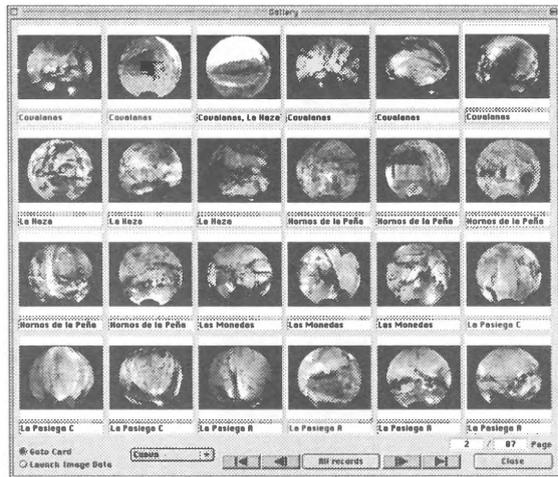
シンボルとなった。聞くところによれば、この種のシンボルはカンタブリア地方からフランスのドルドーニュ地方の洞窟にかけて広く分布しており、何らかのコミュニケーション手段として用いられた可能性が指摘されているが、その意味と目的については専門家の間でもまだ定説はない。

旧石器時代のシンボルについてはさておき、問題は、この種のシンボルをこうした洞窟内の暗闇でいかに球面パノラマ撮影が可能かということであった。前後2枚の魚眼写真から1個のスムーズな球面パノラマ画像を生成させるためには、2枚の境界面の露出が同一でなければならない。そのためには、ストロボを使った多重露光には無理があるし、光源は固定で、なるべく全周に光がまわるように配置されていなければならない。実際、後日、監督省庁から得られた洞窟撮影許可書には、壁画保存の観点から強力なストロボ撮影は不可。光効率90パーセント以上の蛍光性の低温光源で消費電力の総量が300W以下と定められた。状況からしてチューフィン洞窟でのテスト撮影は小型ストロボでのスチル撮影にとどめ、その翌日、別の場所のコバラナスという洞窟でとりあえずマルチストロボと多重露光の魚眼テストを行なった。結果は、やはり、固定光源を複数個、同時に用いなければならないということが確認できたただけであった。

■撮影許可申請

実施計画のための現地協議は、予定通り12月のクリスマスの時期に行われた。この間に、洞窟撮影としては、アストリアス、カンタブリア、ビスカヤ、ギプスコアの4自治州にまたがる22ヶ所が選定され、動産美術については各自治州の考古博物館とアルタミラ研究博物館に撮影許可を申請することになった。問題は、はたして4自治州がこのプロジェクトに理解を示すかどうかであった。洞窟は、各自治州の教育文化庁の管轄であり、中でもアルタミラ洞窟は、マドリッドのスペイン政府文化庁の直轄となっていた。本プロジェクトを実現させるためには、是が非でもこれら省庁の協力を得なければならない。私どもとしては、これら関係省庁の理解を促すため、各担当責任者には、既に、プロジェクトについてその趣意書を郵送しておいたが、第2回目の訪問では、直接、彼らと接触し、内々の合意をとりつけておく必要があった。

地元のカンタブリア州政府からは、ただちに全面協力の意向が伝えられたが、他州では、それほど簡単ではないことが徐々にわかってきた。旧石器洞窟

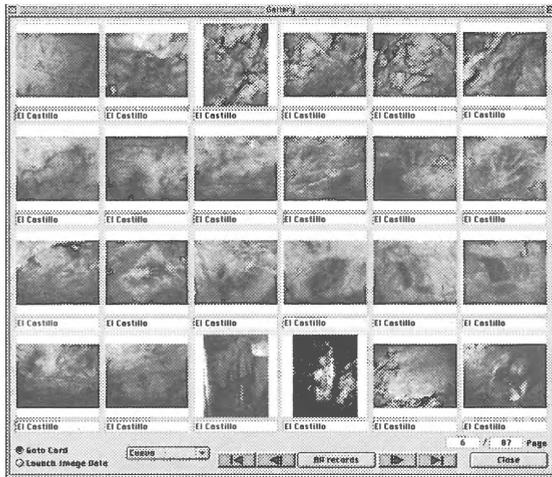


[球面パノラマギャラリー]

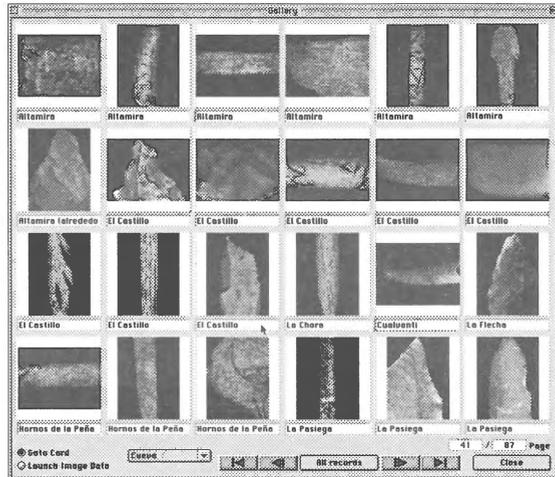
美術の情報化については、その意義を認める。しかし、問題は、自州の文化遺産の情報化を、何故、日本企業とカンタブリア大学主導で行わなければならないのかということであった。また研究者の間でもその必要性和遺跡保存の立場や管轄の問題から意見が分かれ、州政府としてのコンセンサスをまとめるためには相応の時間が必要のようであった。またアルタミラ洞窟の撮影に関しては、特に文化大臣の許可が必要であり、撮影許可申請に対しては、結局、最後まで正式な回答が得られなかった。

■日本・スペイン産学共同開発契約

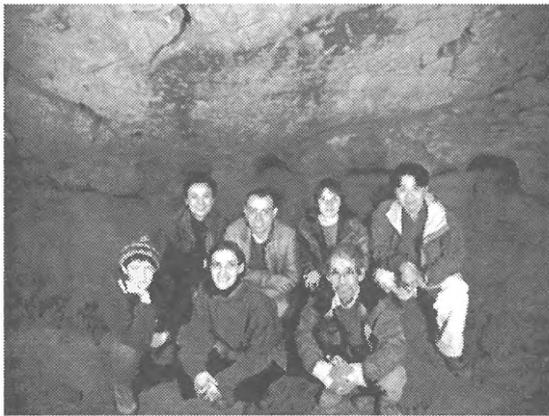
実施計画が確定し、それを前提にカンタブリア大学との正式な共同開発契約が成立したのは、撮影チームを現地へ送る1998年3月初旬のわずか1週間前になってからであった。カンタブリア大学にも制度としては産学共同開発の規定が用意されていたが、それはスペイン企業と主として理工学部に対するものであり、これまで経済効果には全く関係のなかった哲文学部にとっては産学共同開発自体が前例になく、ましてや日本の小企業が相手となれば、その書式を法的に整えるためにはかなりの人力と時間を要したらしい。ちなみに契約では、同年3月1日から9月までの契約期間、撮影場所と撮影目標点数、役務の範囲、大学に対する経済支援の金額、素材と成果物に関する共同著作権ならびに二次著作権の範囲などが明確に規定された。条件面については、概略、私どもから提示した通りではあったが、概して日本企業側に有利な文面となっており、成果物の著作権については互いに平等に留保されているものの、その行使についてはもっぱら私どもの裁量にまかされていることが眼についた。



[洞窟壁画パノラマギャラリー]



[動産美術ギャラリー]



[洞窟撮影班、ティトブスティージョ洞窟]

各自治州からの撮影許可は、更に遅れ、アストリアス州に関しては撮影直前の洞窟前の現場で契約が交わされた。撮影許可に関わる条件は、4州とも大体同じで、洞窟名と既に述べた照明機器の仕様、成果物の非営利目的の使用権に加え、最終成果物としてのデータベース・システムの寄贈に関する条項などがおりこまれていた。ビデオ撮影については、アストリアス州だけは不許可とされた。

■現地撮影開始

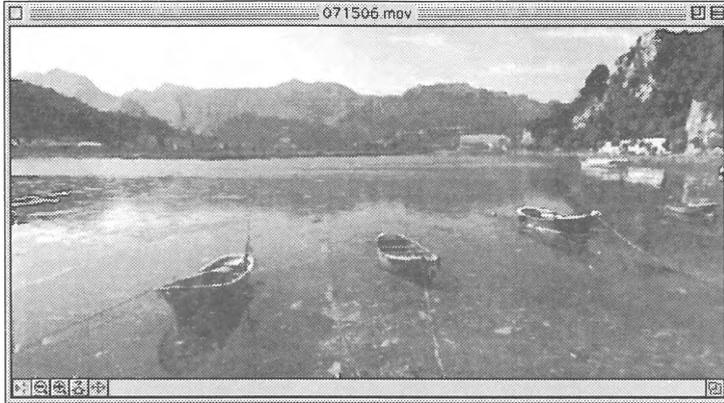
こうして現地撮影は、私どもにとっては、モンゴルの冬季撮影を終えてから間もない3月7日に日本を出発し、5月8日帰国まで60日間の予定で開始された。イベリア航空の協力もあり、デスクトップコンピューター1式、ラップトップコンピューター2式、小型全周撮影装置とその照明機器1式、洞窟撮影用の低温蛍光灯ランプ20個、高精細デジタルカメラ2式、電源ケーブル、電源トランス大小5式、魚眼レンズと専用ローテータを含む撮影機材2式、ビデオカメラ3式など必要な機材は全て日本から運

だ。撮影スタッフは、洞窟班と動産美術班の2班に分かれ同時に併行してスケジュールを消化した。洞窟班には、日本側からは4名、カンタブリア大学から2名、それに各州考古学局の専門家と公式ガイドが案内役をつとめた。動産美術のQTVRオブジェクト撮影には、日本側から2名ないし3名、カンタブリア大学から1名、これに博物館の研究者が補佐役として参加した。

現場では、遺跡の保全問題をめぐり、洞窟ガイドとの間に多少のいきちがいがあったものの、撮影そのものは概して順調に進んだ。ただし、洞窟撮影に関しては、現場の洞窟がたいてい山の中腹にあり、開口部までの機材運搬ならびに形状複雑な洞窟内部での照明機器の設定には相応の時間と労力を要した。また、生活様式の違いから、実際の撮影時間が予想外に制限され、60日間という日程は、移動や休日までを含めれば決して十分なものではなかった。そのために、洞窟周辺景観については、7月に再度、現地入りし、東はフランスのアトランティック・ピレネーから西のアストリアスまでを約30日かけて撮影した。また、その間にギブスコアでも第一級のエカイン洞窟の撮影許可があり、再度、洞窟内撮影が実現した。なお、アルタミラの撮影については、再三、マドリッド政府に問い合わせたが、高官が夏季休暇に入ったこともあり、応答さえ得られなかったのは既に述べた通りである。

■データベースの組み立て

この種の情報化作業で更に時間を要するのは、撮影の後処理である。90日間の撮影で得られた画像素材は、スチル写真が約4000点、球面パノラマ素材約220点、QTVRオブジェクトムービー素材



[QTVRパノラマ・ウインドウ]

300余点、QTVRパノラマ素材約200点、それにデジタルビデオ素材が合計48時間分にも達していた。マルチメディアコンテンツ制作支援事業では、その完成が10月末と規定されており、それまでの3ヶ月間に少なくとも仕様を満たす結果を提示しなければならない。作業は、急ピッチで進み、まずはおおまかに選別されたデジタル画像を全てスペイン側に送り、写真の同定作業とその書誌情報ならびに北スペイン旧石器洞窟美術に関する一般解説の執筆が開始された。また、日本側では、写真の選別とデジタル化から始まって、その画質補正から球面パノラマ、QTVRオブジェクトムービー、QTVRパノラマムービーへの変換、そしてデータベース管理プログラムのカスタマイズ作業が併行して行われた。完成した画像素材は、即、データベースに一括登録し、書誌情報については、スペイン側から逐次送信されてくるアクセス形式のテキストデータを翻訳と併行して変換・登録し、予定の10月半ばには、日・西版第1バージョンがほぼ完成の域に達した。

■現地発表

PhotoVR考古資料データベース『北スペインの旧石器洞窟美術』は、まがりなりにもこうして完成し、翌11月、サンタンデルで開催された国際マルチメディア・フェスティバルで発表する機会を得た。また翌年、1999年3月末から4月にはカナダブリア、アストリアス、ビスカヤ、ギプスコアの各省庁を再訪し、公式の完成発表とあらかじめ契約で規定

されたデータベース・システムの贈呈をもって、ひとまず本プロジェクトは完結した。今後は、その成果物をいかに世界に向けて公開し、また本プロジェクトをいかに将来にわたって発展させてゆけるかである。

■ネットワークで結ぶグローバル・データベース

近年、先史人類の残した形象文化の起源については、未だ推定の域を出ないとはいえ、時代的には中期旧石器時代にまで遡る可能性が問われ、また空間的にはヨーロッパから北アジアのみならずアメリカの南北両大陸を含む地球のほぼ全域にその証拠となる痕跡が確認されつつある。この種のデータベースが世界中で構築され、その多くがネットワークでひとつに統合された時、少なくとも先史人類の形象文化に関しては、また新たな地平が見えてくるかも知れない。

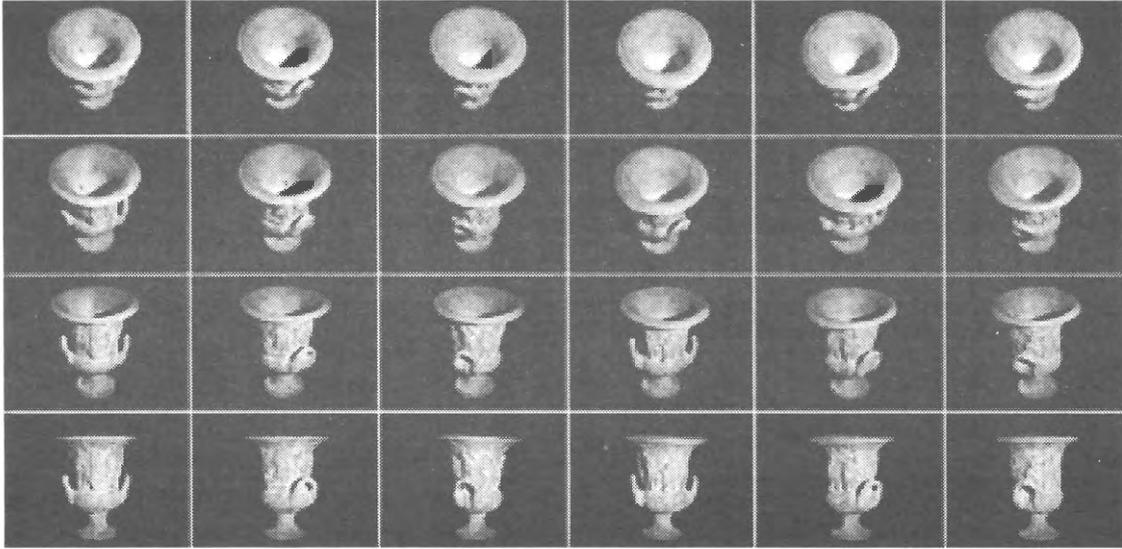
[註]

[1] 画像資料としてのPhotoVR

「仮想現実 (Virtual Reality)」とは、元来、「コンピュータが生成する3次元空間で一人の人間が没入することのできるもの」として、HMD (Head Mounted Display) やデータグローブ、データスーツなどVR入出力端末の開発で知られる米国VPL社の創始者ジャロン・ラニアー (Jaron Lanier) が1989年に提唱した言葉であった。以後、10年の間に、3次元画像のリアルタイム生成技術はソフトウェア、ハードウェア共に急速な進歩をとげ、これに、立体視、触感・力感のフィードバック技術、音

声合成認識技術が統合されて、現在、「仮想現実」は、科学技術シミュレーション、環境シミュレーション、エンタテインメント等に幅広く応用されつつある。

本データベースに関わる「PhotoVR」も、この延長線に開発された画像表示技術のひとつである。本来の「仮想現実」がコンピュータ内部に仮想的な3次元モデルを構築するのに対し、その「仮想現実」を2次元的な「写真」または他の方法で生成された2次元画像をデジタル処理することにより、対象とする景観や立体物を疑似3次元的に、しかも対話的に表出しようとするもので、一般にはQTVR、IPIX、



[QTVR オブジェクトムービーの画像マトリックス]

SmoothMovなどが知られている。

2次元画像による「PhotoVR」に対して3次元モデルによる仮想現実を「3D VR」と呼ぶことも可能であるが、そのどちらも対象物をよりリアルにコンピュータ上に再現しようとする目的には違いがない。ここでは、考古資料を対象物とした場合の双方の特徴を以下にまとめてみる。

[3D VR の特徴]

1) オブジェクト・インデペンダンス：対象物の有無、形状の如何にかかわらず、理論的には無限のモデリングが可能である。従って、考古資料の情報化としては、現存しない遺構・遺物の復元には極めて有効であるといえる。逆に、3Dモデリングには、極度の抽象化と仮に資料が現存する場合であっても、その3Dスキャニングには測定誤差と陰面補完にかかわる測定者の恣意が伴うため、2次資料としての価値が必ずしも保証できない欠点がある。

2) 数値変換と定量化：対象物の全てを3次元データとして保持するため、座標変換、投影変換を駆使することによって対象物を全ての視点から表出可能となる。また、同じ理由から寸法・容量など対象物の物理量について定量的な記録と算出が可能となる。但し、上と同じ理由により、2次資料としての価値は必ずしも保証できない。

[Photo VR の特徴]

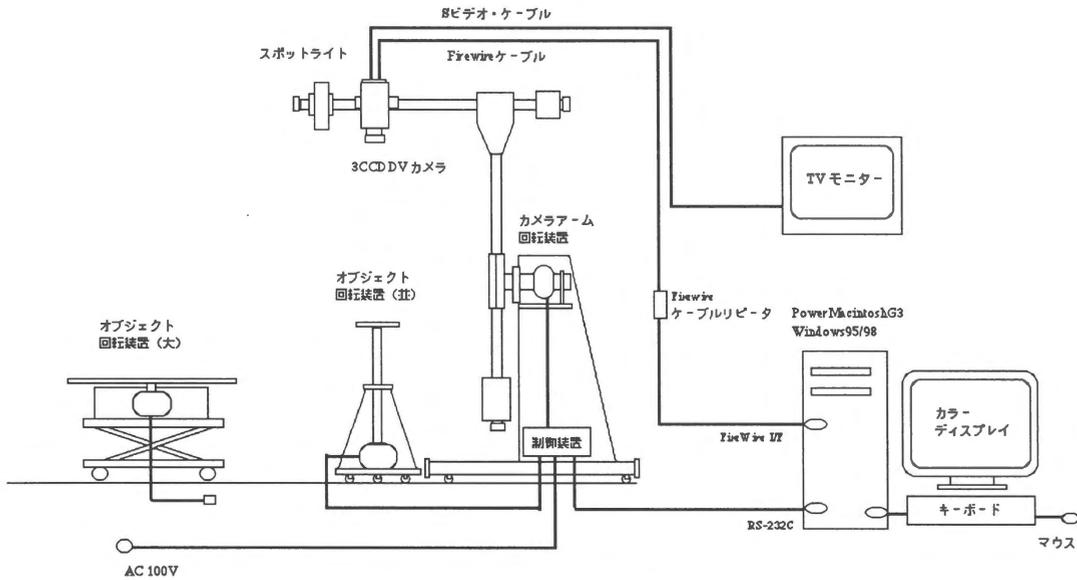
1) オブジェクト・デペンダンス：写真撮影可能な対象物に限定されるが、現存する遺構・遺物等に対しては従来の写真測量と同等の資料性を保持することができる。

2) 疑似3次元としての限界と可能性：対象物に関するデータは、2次元座標に投影された画素のみに限定されるため、3次元的な処理には決定的な限界がある。従って、遺跡空間のPhotoVRにあつては、3DVRに可能なウォークスルーや景観シュミレーションを実現することは不可能となる。

3) データ容量：QTVRオブジェクトムービーの場合は、対象物の静止画像を動きの数だけ含むためにデータ量が膨大となる。インターネットで送信する際には、このデータ容量がネックとなる場合が多い。ただし、将来は、回線の高速化によってデータ量の問題は解決されることが期待される。

[2] QTVR オブジェクト・ムービー

QTVRオブジェクト・ムービーは、米国アップル社の開発したQuickTime Movieの一形式で、通常のQuickTime Movieが時系列上の1次元表示であったのに対し、それを2次的に拡張したものである。複数枚の静止画像が縦系列と横系列のマトリックス状に配列され、その1個をポインターで指定して表示することができる。従って、立体物に関しては、その横方向の変化と縦方向の変化をそれぞれ連続的に撮影してQTVRオブジェクト・ムービーに組み込めば、マウスオペレーションによって対話的にその立体物をあらゆる角度から鑑賞可能となる。この場合、ひとつの立体物を滑らかに回転表示可能とするために必要な静止画像のフレーム数は、最低で縦横10度間隔、つまり水平360度、垂直90度のQTVRオブジェクト・ムービーの場合、36 x 10で360フレーム程度であることが経験的にわかっ



[自動全周撮影装置概念図]

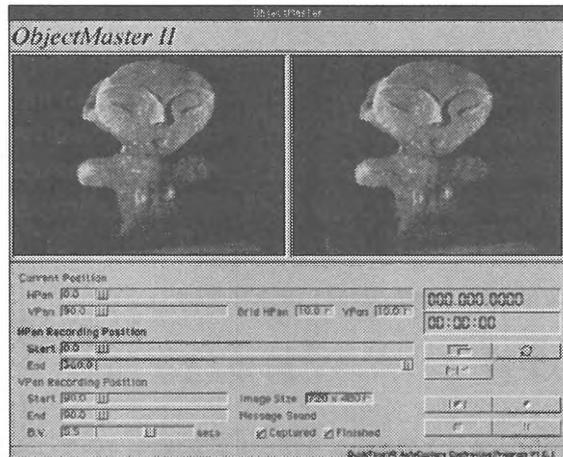
ている。撮影間隔をせばめればそれだけ滑らかな動きが期待できるが、それだけファイル容量が増えることにもなる。ちなみに640 x 480画素のフレームを360組み込んだ場合のファイル容量は、ビデオ圧縮して約25~30MBである。QTVRオブジェクト・ムービーは、通常のムービープレーヤで表示可能となっている。

[3] 球面パノラマ

等距離射影方式の180度魚眼レンズで撮影した複数の写真画像をビルド処理して得られるパノラマ画像で、専用のブラウザーを用いることにより、Macintosh 又はWindows上で撮影地点を中心とした全球面のパノラマ景観が対話的に鑑賞可能となる。カメラを中心として前後2枚撮影して合成するIPIX形式と3枚を合成するSmoothMov形式が知られている。両者とも球面の魚眼写真を正6面体の内側に展開し、その中心から矩形に射影して表示する方法をとっている。

[4] 自動全周撮影装置

QTVRオブジェクト・ムービーを画像データベースの素材として採用する場合、かかる全周画像をいかに効率よく滑らかに撮影できるかが問題となる。データベースというからには、通常、数百から数千の資料をできる限り短期間で撮影する必要があるからである。本プロジェクトで使用した自動全周撮影装置は、そうしたニーズのもとで立体物の全周撮影からデジタイズ・画質調整・圧縮・縮小・ムービー編集



[制御プログラムのユーザインターフェース]

までを一貫してパソコンで自動処理することを可能にしたものである。メカニズムとしては、被写体をのせる水平ローテーターとカメラ用の垂直ローテーターを2個のステッピング・モーターで駆動し、これらをコンピュータで同時制御して3次元自動撮影を実現している。

ソフトウェア機能としては、ローテータの2軸制御と共に、3CCDカメラからのFireWireデジタル転送、ムービーサイズを選択、画像圧縮の選択、画質調整、非圧縮PICT画像セットの格納、DV形式QuickTimeムービーへの変換機能などが組み込まれており、撮影条件を一回設定すれば、あとは自動的にQuickTimeムービーまたはPICTファイル画像セットが得られることになる。生成されたQuickTimeムービー

ビーは、アップル社の変換プログラムで縦横のフレーム数などを入力し、簡単にQTVRオブジェクト・ムービーが生成できる。

[5] QTVR パノラマ・ムービー

16mmから24mm程度の広角レンズを用い、カメラを中心に360度回転撮影して得られた12枚

から18枚の写真をステッチング結合処理し、これをQTVRパノラマ形式に変換したムービー画像である。通常のムービープレーヤーを用いることにより、Macintosh 又はWindows上で撮影地点を中心とした360度の円筒パノラマが簡単なマウス操作で対話的に鑑賞可能となる。

[参考文献]

深沢武雄, 自動3次元撮影装置の開発ならびにQTVR考古資料データベース作成の実際, 3D映像, Vol.11-No.3, 1997

Takeo Fukazawa, *Iconteque Project, The digital archiving project for cultural heritages and its publication using the latest PhotoVR and Multimedia technologies*, 4th Congress of Study on Japan, Santander, 1997

Takeo Fukazawa & Roberto Cacho Toca, *PhotoVR Iconteheque: Paleolithic Art in Cantabrian Region*, 26th Computer Applications in Archeology, 1998, Barcelona

Georges Bataille, *La Peinture Prehistorique Lascaux ou la Naissance de L'art*, Geneve, 1955
(出口裕弘訳『ラスコーの壁画』二見書房, 1975)

木村重信『美術の始原』新潮社, 1971

木村重信『ヴィーナス以前』中公新書, 1982

横山祐之『人類の起源を探る』朝日選書, 1987

横山祐之『芸術の起源を探る』朝日選書, 1992

John Wymer, *The Paleolithic Age*, 1995, New York (河合信和訳『世界旧石器時代概説』雄山閣, 1989)

Richard Rudgley, *Lost Civilizations of the Stone Age*, London, 1998 (安原和見訳『石器時代文明の驚異』河出書房新社, 1999)

George F. Carter, *Earlier than you think, A Personal View of Man in America*, Texas A&M University Press, 1980

バーチャルリアリティによる遺跡探訪

3D Archaeo-Copter

中村 健 小沢 一雅

Takeshi Nakamura and Kazumasa Ozawa

Osaka Electro-Communication University, Neyagawa, Osaka 572-8530, JAPAN

Abstract. An experimental system for virtual travelling over an archaeological region has been implemented. The central idea for building this system is simulation of 3D travelling over a region by taking a helicopter. Nara Basin has been employed as an example of archaeological region.

Keywords: 3D, Virtual reality
, Archaeological site

1. Introduction

Computer technical innovation dramatically changes visualization of archaeological monuments. A typical way of archaeological visualization has been presented by animated pictures generated by computer graphical techniques. A number of papers on such type of visualization have already been published[1-3]. The authors have also been engaged in making animated pictures of Japanese ancient scenery. During our working process, a special type of modeller named ASM has been developed by which some animated pictures of scenery of ancient villages have been created[4].

In recent years, another technique has been available for archaeological visualization: Since computing power and memory space of computer has grown quickly, real-time generation of pictures can practically be carried out by a personal computer. This may bring about a new revolution in archaeological visualization. Namely, the virtual reality techniques (or realistic virtual imaging) are now available. An immediate application of such virtual reality techniques is given by the so-called "walk-through" an ancient place [5].

In archaeology, geographical relations between a find and other finds, between a find and its site, and between a site and other sites are seriously considered for understanding the ancient society. To handle such relations in a new type of archaeological visualization, a fairly wide range of region should be covered, including a number of archaeological sites. The most macroscopic view of the region will be given by bird-looking from a very high view-point. If this was possible, we would be able to understand immediately inter-site geographical relations. On the other hand, it will also be desirable that when we wish to make

a precise inspection of a site, we can look it from a lower view point or can "walk through" the site.

This story has reminded us of travelling the region by taking a helicopter. This paper presents an experimental system to simulate such type of virtual travelling over an archaeological region by using the virtual reality techniques. We call the system 3D Archaeo-Copter.

2. 3D Archaeo-Copter

The system has been designed according to the following typical steps to visit a site by taking a helicopter:

- (1) Looking at the entire region from a very high view-point, determine a site to visit,
- (2) Circling over the site at a lower altitude, reconfirm the site,
- (3) Landing at a point near the site, confirm the site with eyes,
- (4) Make a precise inspection by walking through the site.

The system acts in terms of transition between the four modes corresponding to the above (1)-(4) as is shown in Figure 1. The first mode is termed HA (High altitude) mode, which is initially presented for a user of this system. The display image in HA mode is a distribution map of a selected kind of sites superimposed on the 3D digital terrain picture (See Figure 2). A user can move anywhere on this map as if controlling a helicopter. Specifying a small area of the region, he can switch the present HA mode to LA (Low altitude) mode in which he can make bird-looking inspections. Display images in this mode are given by high-speed 3D imaging based on the 3D digital terrain data (See Figure 3). The next is G (Ground) mode. Here, a user can obtain text information of the site; e.g. name, address, archaeological significance or recommended walking routes of the site (See Figure 4). The final mode is WT (Walk-through) mode in which he can freely walk around in the

site, visiting pit-dwellings or high-floured houses. Figure 5 presents an example of frame picture generated in a walk-through process.

After precise inspection of a site, if trying to visit another site, you have to back to the HA mode. Then you can visit another site in the same way as described above. Figure 6 illustrates the conceptual system structure of 3D Archaeo-Copter. The three sets of data play key roles in performing functions of the system. Object models include all the surface models of objects to be visualized; e.g. Keyhole tomb mounds, houses, fences and trees. Site data is a database of Keyhole tombs in Nara Basin. 3D digital terrain data is employed to generate the display images in HA and LA modes. Especially, object models and the terrain data need so large amount of memory that space and time to be travelled have been limited by hardware. The present system runs on Windows 95/98 through API called DirectX.

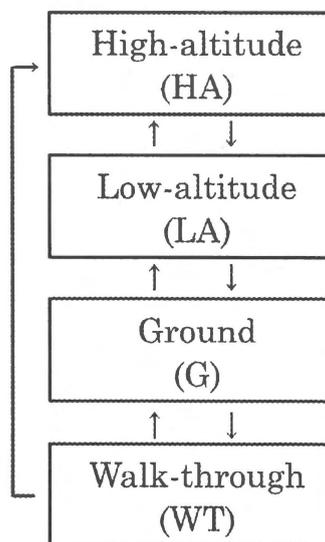


Figure 1
Transition diagram of the four modes.

3. Region

The region which the present system covers is Nara Basin where the Japanese ancient regime begun in the 4th century. In later periods, ancient capitals had

Figure 6 Conceptual system structure of 3D Achaeo-Copter

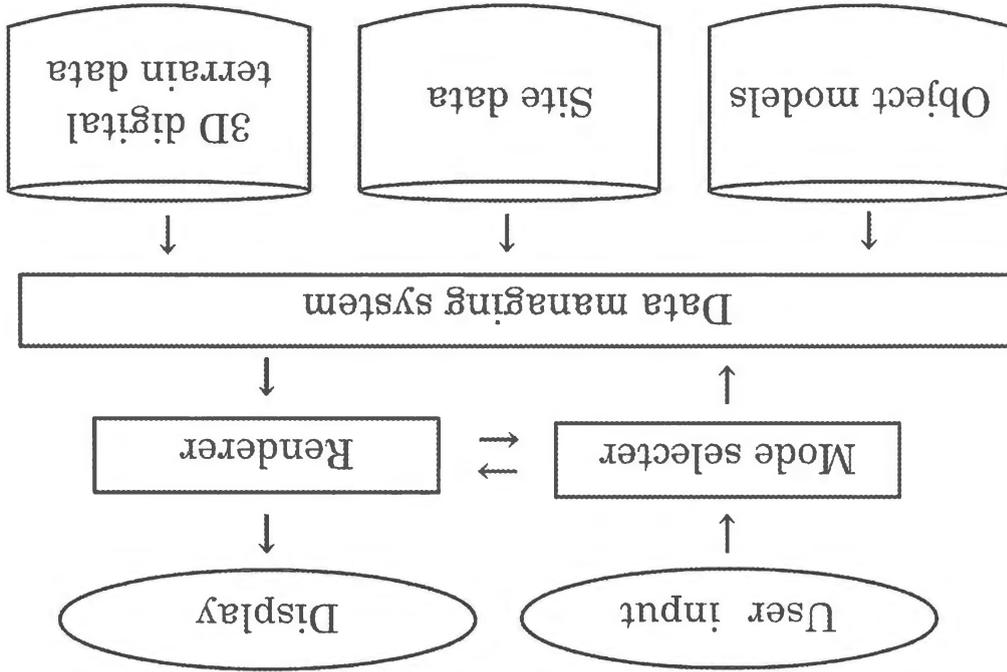


Figure 5 Walk-through

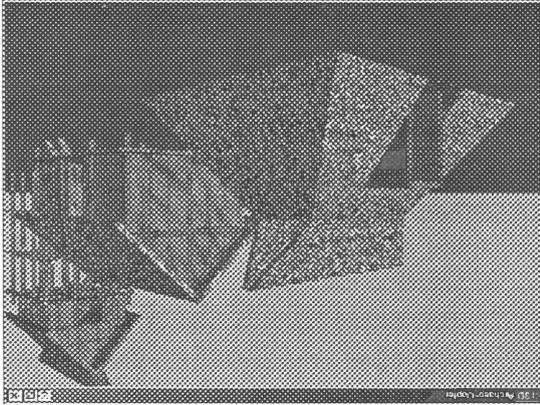


Figure 3 Low-altitude



Figure 4 Ground

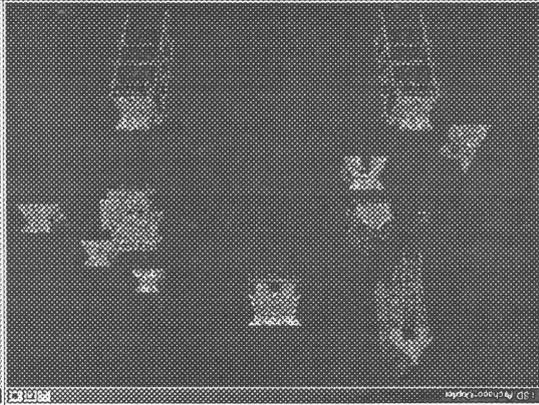
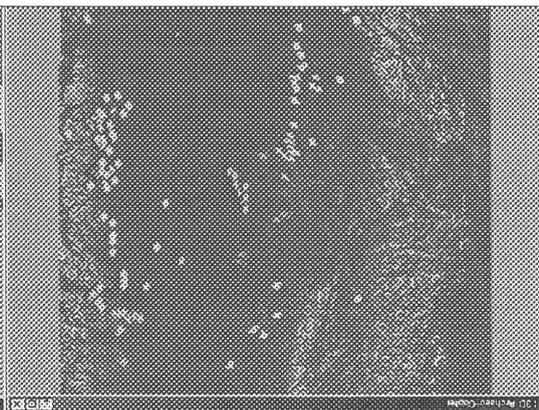


Figure 2 High-altitude



continued to be built within this region until it moved to Kyoto. Because of its long history, Nara Basin is widely recognized as one of the most important regions with rich archaeological monuments.

The 3D digital terrain data provided by the government has been employed for generating the distribution map in HA mode. The minimum unit of the terrain data is a 50m by 50m square area.

The sites handled by the present experimental system are limited to a small set of monuments; i.e. the Keyhole tombs and their neighbored villages. Namely, the system can treat simply Nara Basin in the Ancient Tomb Period. In fact, the authors have been engaged in building a database of Keyhole tombs since 1980. This is a reason why we have employed the set of Keyhole tombs as an example set of sites for 3D Archaeo-Copter.

4. Conclusion

This paper presents the conceptual system structure and functions of 3D Archaeo-Copter. Virtual travelling by the present system has been limited to Nara Basin only in the Ancient Tomb Period. As stated previously, this region includes a rich amount of archaeological monuments stored up during a long range of time. One of our future tasks is to extend the time range so that a user could visit more sites other than the Ancient Tomb Period. When this was taken place, another option to select a period of interest would be required in HA mode. Secondly, it will be a future work to extend the region to be handled. Such extensions obviously need a huge memory space beyond the present situation of computer hardware. Technical innovation is, however, so rapid that both extensions in terms of time and region look to be taken place very soon.

References

- [1] K. Ozawa, Reconstruction of Japanese ancient tombs and village. Proc. of CAA92. ISBN: 87 7288 1127. Aarhus University Press, Aarhus, 1993, pp.415-423.
- [2] S. Morimoto & M. Motonaka, Reconstruction of the 8th-century Imperial Palace of the Heijo capital at Nara in Japan, Proc. of CAA92. ISBN: 87 7288 1127. Aarhus University Press, Aarhus, 1993, pp.425-427.
- [3] P. Boland and C. Johnson, Archaeology as computer visualisation : 'Virtual Tours' of Dudley Castle c.1550.. ISBN: 0 86159 114 3. British Museum Occasional Paper No.114, 1996, pp.227-233.
- [4] K. Ozawa, ASM : An ancient scenery modeller. ISBN: 0 86159 114 3. British Museum Occasional Paper No.114, 1996, pp. 109-118.
- [5] R. Kadobayashi and K. Mase, The VisTA-walk system. Abstract of CAA98, Barcelona, 1998, pp.57-58.

照度差ステレオ法を用いた遺物の表裏形状の計測

Measurement of the front and back form of a relic by using the Photometric Stereo Method

結城宏和*、宝珍輝尚*、都司達夫*

Hirokazu YUUKI*, Teruhisa HOCHIN*, Tatu TSUJI*

* 福井大学 工学部 情報工学科

〒910-8507 福井市文京3丁目9-1

*Department of Information Science, Faculty of Engineering, Fukui University
3-9-1, Bunkyo, Fukui-shi, Fukui 910-8507 Japan

あらまし：カメラ2台を使用して遺物の表面と裏面を同時に計測し、照度差ステレオ法で遺物破片の形状を求めた結果を報告する。

Summary : This paper describes the results of the measurements of the front and back forms of relics. Pictures of two sides of relics are taken with two cameras. The front and back forms are obtained from these pictures by using the Photometric Stereo Method.

キーワード：遺物、照度差ステレオ法、破片

Keywords: relic, photometric stereo method, broken piece

1 はじめに

一乗谷朝倉氏遺跡では、毎年多くの遺物が発掘されている。その数は約200万個にもものぼる。発掘された遺物は、完全な形で発掘されているものもあるが、多くは破片の状態で発掘されている。これらの破片をデータベース化することで、従来人手に頼っていた集計作業を自動化できることや、様々な破片から計算機を用いて新たな仮説を導出することができると考えられ、破片のデータベース化が必要となってきた。また、これらの遺物に対して形状、紋様、色、重さなどの特徴を用いて遺物の検索を行いたいという要求や、遺物の形状からもとの物体を推定したいという要求がある。しかし、微小で大量の複雑な破片の形状計測には時間と経費をかけられ

ないという制約がある。

我々は、現在、微小で大量の雑多な破片の形状計測に関する検討を行なっている。これまでに、照度差ステレオ法を用いて遺物の3次元形状の測定を球体を用いて行った[1,2]。また、現在までに、球体並びに実際の遺物として白磁の陶器の破片の形状測定を試み、計測する物体ごとに照度の強さを変化させることによって形状の計測を行なえることを明かにしてきた[2]。しかし、形状計測にあたり、破片を上方から撮影した画像のみを用いているため、裏面の形状は計測できていなかった。

ここで、金谷らは、表面と裏面を別途計測し、表面と裏面の対応点を検出することによってこれらの2面の貼り合わせを提案している[6]。しかし、この方法では、例えば円筒状の破片に対しては対応点を検出できないので適用できないという問題や、測定に手間がかかりすぎるという問題がある。また、レーザー光を用いた測定で、物体またはカメラを回転させる方法が考えられるが、これも微小で大量の破片を測定するには手間がかかりすぎる。これらの点から望ましい方法は、表面を計測した状態で裏面を計測することである。このために、アクリル板やガラス板といった透面な板に破片をのせて計測可能かについて検討を行い、これまでの実験で、照度差ステレオ法を用いた計測方法において、アクリル板を用いても計測できることを明かにしてきた[3]。

そこで、本論文では、照度差ステレオ法を用いて、破片の表面と裏面を同時に計測する方法について述べる。また、製作した計測装置を用いて破片の形状を求めた結

果を示す。

以下、2では、測定対象とする遺物について述べ、3で照度差ステレオ法について説明する。4で作成した計測装置について述べ、5で計測手順を述べ、6で本装置を用いた測定結果について述べる。最後に7でまとめを述べる。

2 対象とする遺物データ

本研究では、福井県にある越前一乗谷朝倉氏遺跡より発掘された遺物を対象とする。

発掘された遺物は約200万個あり、遺物のなかには発掘後すぐに茶碗や皿と識別され選別されるものもあるが、なかには破片の形状が小さすぎて識別することができないものもある。復元対象の遺物はこの識別できない破片である。

対象となる破片の例を図1に示す。図1は約43mm×41mm程度の小さな遺物である。

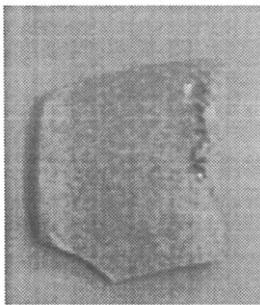


図1: 対象とする遺物

3 照度差ステレオ法について

3.1 明るさからの3次元形状の推定

一般に、物体面に照射された光エネルギーは一部吸収されるものの、大部分は反射され、これが面の明るさとして観測される。このとき、面の明るさを決定する要因として、面に照射されている光量、面の光反射特性、面の向きなどが考えられる。

図2は、一般によく知られた2種類の光反射特性であり、左が鏡面反射、右が拡散反射である。通常の物体面の光反射はこの2つの反射成分が混じりあっている。鏡面反射では、光の特定の方向にしか反射されないの

その方向から観測しない限り面の明るさを測定することはできない。しかし、もしその反射光が観測されたならば、観測方向と光源の向きから面の向きを推定することができる。

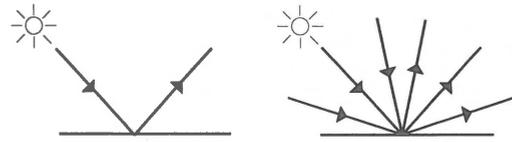


図2: 鏡面反射(左)と拡散反射(右)

完全拡散反射面の明るさは、どの方向から見ても同じ明るさとして観測される。つまり、カメラをどの方向に置いても測定した明るさは同じである。図3の右図のように面に対して低い角度から照射されている場合は、単位面積あたりの光量が少ないので暗くなる。図3の左図のように、明るさが最大となるのは、面の法線方向から照らされているときである。光源の方向と面の法線方向との角度を θ とすれば、完全拡散反射面の明るさ I_p は $\cos\theta$ に比例する。すなわち、 I を点 p での入射光量として $I_p = I \cos\theta$ となる。

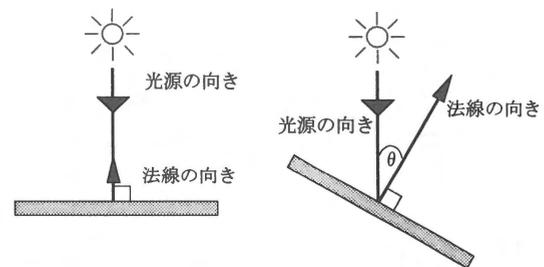


図3: 光源と平面の法線の向き

ここで、図4のようにカメラ中心の座標系を導入し、 $z = z(x, y)$ 、 x 方向の勾配を $p = \frac{\partial z}{\partial x}$ 、 y 方向の勾配を $q = \frac{\partial z}{\partial y}$ とすると面の法線ベクトルは $(p, q, 1)$ で表せる。光源方向も勾配ベクトルを使って $(p_s, q_s, 1)$ と表せる。このときの $\cos\theta$ はこの2つのベクトルを使って、式(1)のように表せる。

$$\cos\theta = \frac{pp_s + qq_s + 1}{\sqrt{p^2 + q^2 + 1}\sqrt{p_s^2 + q_s^2 + 1}} \quad (1)$$

すなわち、面の明るさは面の勾配 (p, q) の関数として表すことができる。 (p, q) 平面を勾配空間と呼んでいる。

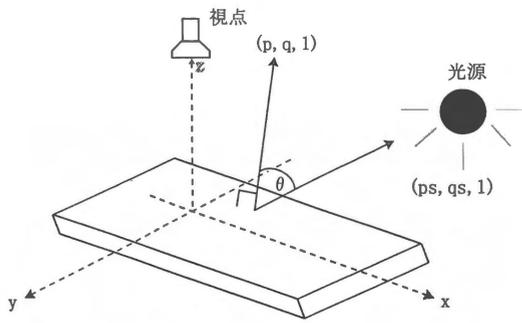


図 4: 光源と面の向き

明るさと面の向きとの関係は反射率分布図 $R(p, q)$ と呼ばれ、最大値が 1 となるように正規化する。したがって、完全拡散反射面の反射分布図は 以下のように式 (1) そのものである。

$$R(p, q) = \frac{pp_s + qq_s + 1}{\sqrt{p^2 + q^2 + 1}\sqrt{p_s^2 + q_s^2 + 1}} \quad (2)$$

この反射率分布図を勾配空間上に明るさを高さとし、等高線を描くと図 5 のようになる。もし、光源がカメラの方向と一致していると $(p_s, q_s) = (0, 0)$ であり、反射率分布図は同心円を描く。

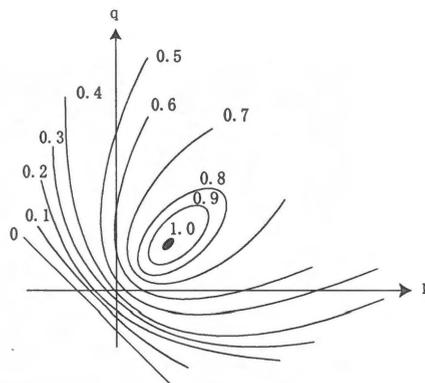


図 5: 反射率分布図

明るさが最大となる場所は、面の法線方向と光源の方向が一致しているところであり、逆に一番暗い 0 となる場所は勾配空間上の直線である。(式 3)

$$0 = pp_s + qq_s + 1 \quad (3)$$

この直線の外側は光が届かず反射率分布図が定義されない。これは図 6 のように求める点が物体の影になるところである。

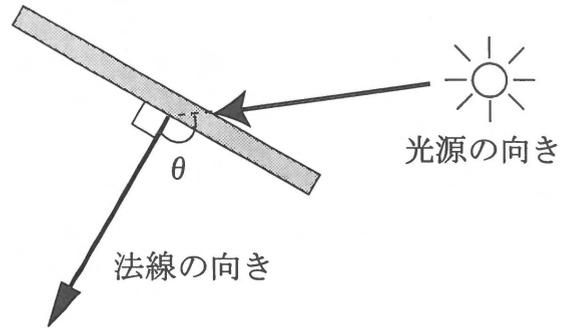


図 6: 光源と平面の向き (光があたらず)

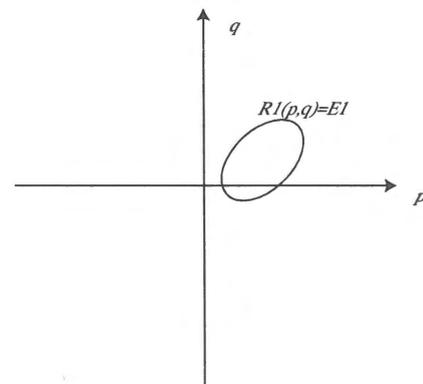


図 7: R_1 のグラフ

3.2 照度差ステレオ法

カメラと対象物体の位置は固定しておき、最初に置かれた光源下の明るさを E_1 とすると照度方程式 $R_1(p, q) = E_1(x_0, y_0)$ の解は図 7 のような楕円になる。次に光源の位置を変えて得られた照度方程式を $R_2(p, q) = E_2(x_0, y_0)$ とすると 2 つのグラフは 2 点で交わる。交わったところの (p, q) の値が求める面の傾きある。さらに光源の位置

を変えて得られた照度方程式をたてると2つの解のうち
のどちらが正しいか求めることができる。(図 8)

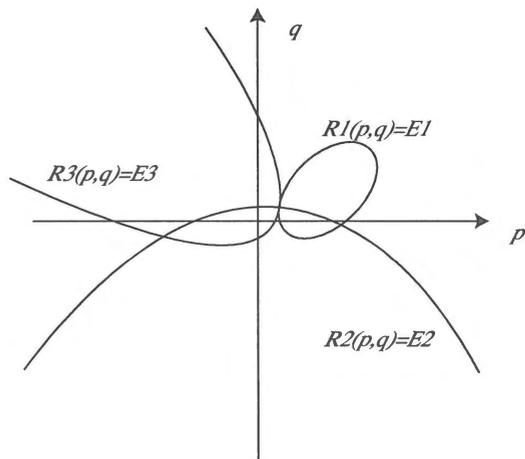


図 8: R_1 , R_2 , R_3 による面の傾きの決定

3つの光源を用いて、1つ1つの画素に対して求めた
 p, q の値は、X-Z軸、Y-Z軸の傾きであり、それを利用
してZ軸の値を求めることによって立体を表すことが
できる。

3.2.1 特徴

照度差ステレオ法は、物体とカメラの位置は固定し、
異なる複数の光源を利用して画像を撮る方法である。し
たがって、カメラは一台しか使用しなくてもよく、ステ
レオ画像処理である両眼ステレオ法等で起こる探索点の
問題が起こらない。カメラを1台しか使用しないのでコ
ストが安いという利点もある。

しかし、照度差ステレオ法は、物体の面の勾配を求め
るのに使用されるものであり、3次元形状を推定する用
途には用いられないのが一般である。

4 計測装置

製作した計測装置の全体を図 9 に示す。この計測
装置全体の大きさは、幅 450mm、奥行き 450mm、高
さ 141.5mm である。この計測装置はフレームに木材を
使用して頑強なものにし、また、外部からの環境光の侵
入を防ぐために計測装置全体を暗幕で囲っている。上部
からデジタルカメラを使用して撮影し、下部からは、ビ
デオカメラを用いて撮影するように製作されている。

計測装置内部写真を図 10 に示す。また、計測装置の
内部のモデル図を図 11 に示す。

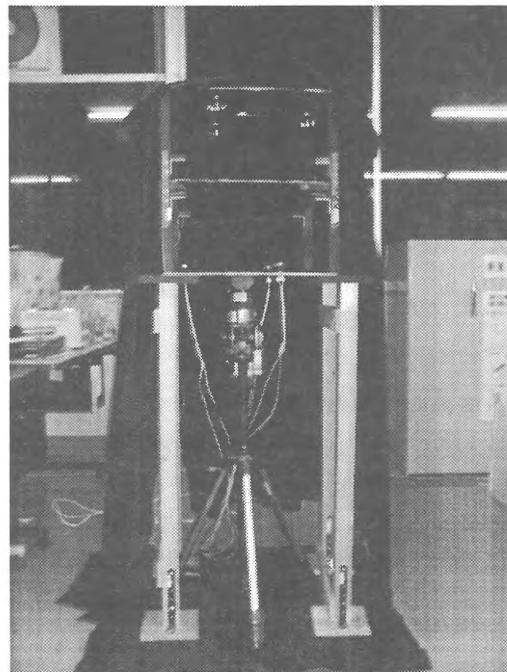


図 9: 計測装置

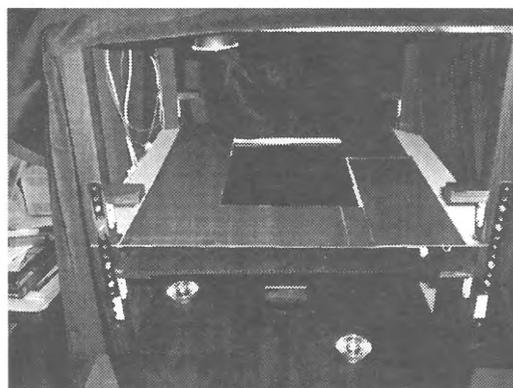


図 10: 計測装置内部

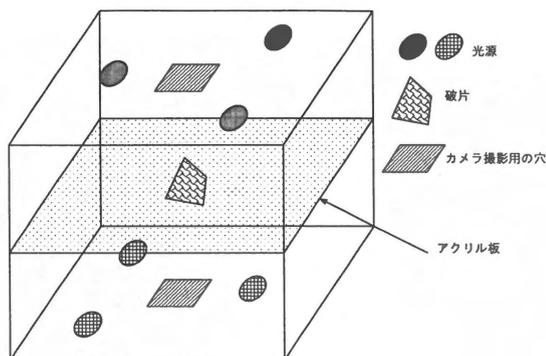


図 11: 計測装置内部のモデル図

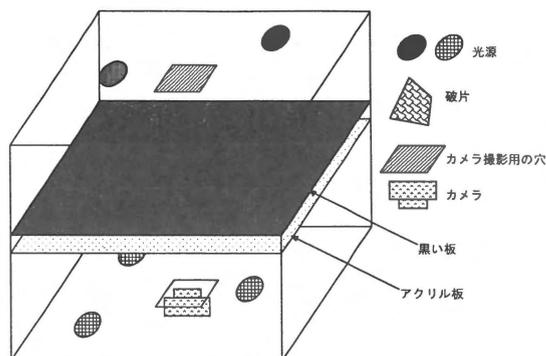


図 13: 計測図 (裏)

この計測装置内部は、中央を厚さ 3 mm のアクリル板で仕切っており、このアクリル板の上に計測する遺物を置く。計測装置の上部、下部には、それぞれの光源となる電球が 3 つずつ設置してあり、自作のスイッチで ON、OFF を切替えることができる。電球の照度は、電源の電圧、電流を操作することによって変更が可能である。実際に計測を行うときには、アクリル板を通して、カメラや電球が破片と一緒に写らないようにするために、表面を計測するときは、アクリル板とビデオカメラの間に黒い板を挟んで撮影し (図 12)、裏面を計測するときは、アクリル板とデジタルカメラの間に黒い板を挟んで計測する (図 13)。なお、ビデオカメラの出力は、直接パソコンに入力されている。

これは、表面からみた物体との位置関係と裏面からみた物体の位置関係も同じである。

表 1: 計測装置の光源の角度

	光源 1	光源 2	光源 3
ϕ	60	180	300
θ	30	30	30

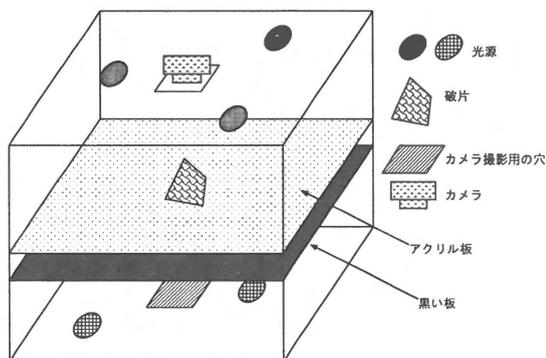


図 12: 計測図 (表)

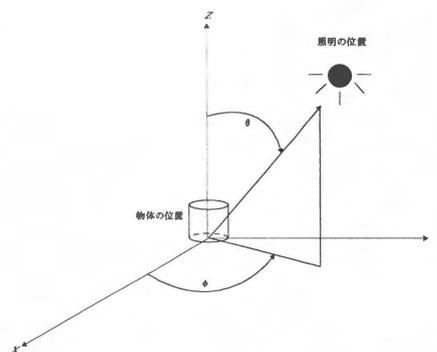


図 14: 電球とカメラのモデル図

3 光源とも破片を置く中央の位置から均等な距離と角度に位置するように固定してある。

5 計測方法

電球とカメラと物体の位置関係を図 14 に示す。光源となる電球の位置関係を示す θ と ϕ の値を、表 1 に示す。

計測を行ったときの照度は 110 ルクス。電圧は 6.5 V、電流は 2.8 A である。この照度の値は、これまでの

実験結果より、最適と思われる値である。

5.1 手順

撮影手順を以下に示す。

1. 計測対象となる遺物をアクリル板の上に置く。
2. 裏面を撮影するビデオカメラが、アクリル板を通して写らないように黒い板を挟んでおく。
3. 3つの光源それぞれについて、表面を撮影する。
4. 次に裏面を撮影する前に黒い板を挟む場所を上部のデジタルカメラが写らない場所に入れ換える。
5. 裏面を撮影する。

以上を繰り返して遺物の撮影を行う。カメラと遺物は、動かす必要がない。

6 計測

6.1 計測対象の遺物

計測に用いた遺物の表面画像と裏面画像をそれぞれ 図 15 と 図 16 に示す。

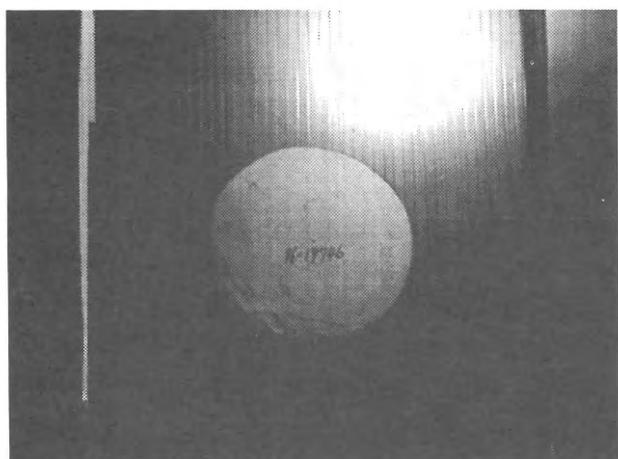


図 15: 表面画像

この遺物は、縦 77 mm × 横 88 mm × 高さ 17 mm の素焼きの皿である。

6.2 計測結果

表面画像からの形状の計測結果を 図 17 に示す。

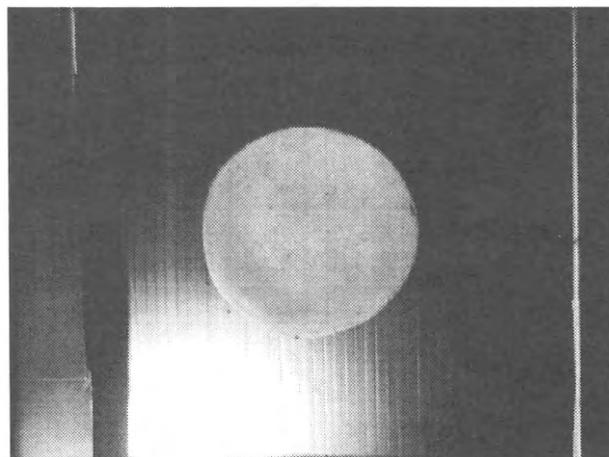


図 16: 裏画像

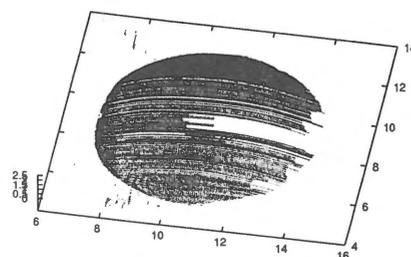


図 17: 表面計測結果

また、裏面画像からの形状の計測結果を 図 18 に示す。さらに、表面、裏面の結果を張り合わせた結果を 図 19 から 図 21 に示す。

図 19 は斜め上方向から見たものであり、図 20 は横方向から見たものであり、図 21 は中央で切った断面図である。図 21 の下の線が、裏面の計測から得た結果であり、上の線が表面の計測から得られた結果である。

6.3 考察

図 17 の中央の計測できない部分は、遺物番号による影響である。図 21 より、計測された素焼きは、厚さ 15 mm といえる。図 19 から 図 21 より、素焼きのおおまかな形状を求めることができたといえる。

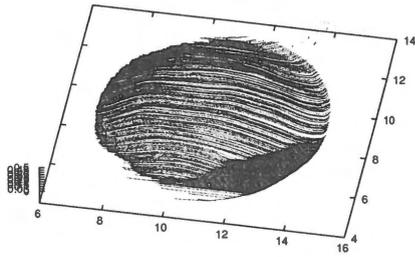


図 18: 裏面計測結果

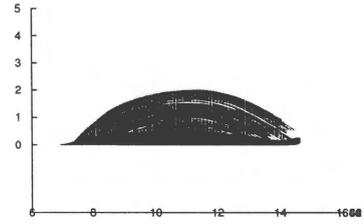


図 20: 表裏面計測結果 (側面)

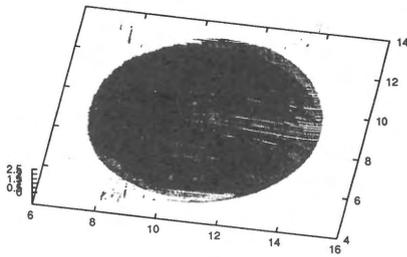


図 19: 表裏面計測結果 (上方)

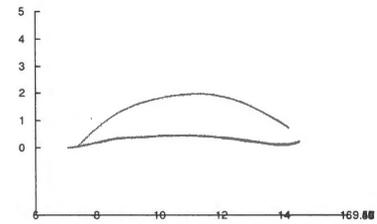


図 21: 表裏面計測結果 (断面)

計測する遺物の表面は、凹形よりも凸形の方が正確に求めることができるようである。すなわち、この例の場合裏面画像から得た形状の方が誤差が大きくなっている。

カメラと遺物は、固定してあるにもかかわらず、図 17 と図 18 結果より、同じ大きさとなっていない。これは、表面は、デジタルカメラ、裏面はビデオカメラで撮影したためにレンズの歪みや倍率の違いによって生じたものと考えられる。

7 まとめ

遺物の 3 次元形状を求めることを目的として表面、裏面同時計測可能な装置を製作した。形状計測には、照度差ステレオ法を用いている。計測の結果、おおまかな形状を求めることは可能だが、撮影に用いた 2 台のカメラ

が異なっていたため、レンズの歪みや倍率の為に、正確に表面と裏面の張り合わせができなかった。

今後は、同じカメラ 2 台を使用した表面と裏面の形状計測、ならびに、より多くの遺物破片に対しての計測と評価が課題である。

謝辞

データの収集・分類作業や考古学におけるデータ管理についての議論などでお世話になった福井県立一条谷朝倉氏遺跡資料館の水村伸行氏に深く感謝致します。

参考文献

- [1] 結城宏和、寶珍輝尚、都司達夫. “照度差ステレオ法を用いた遺物の3次元形状の推定”. 日本情報考古学会第5回大会 発表要旨, pp.65-70,1998.
- [2] 結城宏和、寶珍輝尚、都司達夫. “遺物破片の計測－照度差ステレオ計測装置の製作と計測－”. 第4回公開シンポジウム 人文科学とデータベース, pp.61-70,1998.
- [3] 結城宏和、寶珍輝尚、都司達夫. “破片の裏面形状測定のための基礎検討”. じんもんこもん(-)1999 人文科学とコンピュータシンポジウム, pp.87-8,1999.
- [4] 池内克史. “反射率地図に基づき、二次元濃淡画像より三次元形状を再構成する2手法”. 電子通信学会論文誌”, 842-849(1982).
- [5] Berthold K.P.Horn and Robert W.Sjoberg.
“Calculating the reflectance map”. APPLIED OPTICS, Vol18, No11, pp.1170-1179(1979)
- [6] 金谷一朗、金本優子、陳謙、千原國宏. “遺物破片の形状計測の一手法”. 日本情報考古学会第5回大会. pp71-6,1998.

文学データベースのための文書の構造化と意味管理
Structured Documents and Semantic Objects for Literature
Databases

横田 一正 三宅 忠明 國島 丈生 劉 渤江* 田槇 明子†

Kazumasa Yokota, Tadaaki Miyake, Takeo Kunishima, Bojiang Liu, Akiko Tamaki

岡山県立大学 情報工学部

〒719-1197 岡山県総社市窪木 111

Faculty of Computer Science and System Engineering

Okayama Prefectural University

Soja, Okayama, 719-1197

キーワード: 文学データベース、構造化文書、XML、意味オブジェクト

Keywords: literature database, structured document, XML, semantic object.

あらまし: 電子化された文学作品をデータベースに蓄積し、情報検索や統計処理のツールを使って文学的考察を行うことは多く行なわれており、文学研究に大きく貢献してきている。しかし、さらに深く文学研究を進めるためには、構造的あるいは内容的関連を表すことが重要で、作品内の特定の部分を抽出したり、それらの間、さらには複数の作品間にリンク付けを行う必要がある。そのために、文書管理技術によって文学作品を構造化し、さまざまな意味情報を扱い、さらに文学研究の支援ツールを整備したシステムの研究開発を行っている。本稿では、ケルト文学の『デアドラ伝説』を対象に、XML による構造化文書の生成、総索引表(コンコーダンス)等の作成、構造化文書間のリンク付けなどをサポートしたプロトタイプシステムについて報告し、XML で記述できない利用者のオブジェクトの扱い、それらを含む管理機構について議論する。

Summary: There are many works for literature databases, in which electronic documents are stored and information retrieval and statistics tools are supported. However, for advanced literature research, it is important that structural or semanti-

cal similarities are represented. That is, any part in a document is specified as an object and linked to another one in the same or another document. For the objectives, we have been developing a system based on some document management techniques, in which literature documents are structured, user-specified objects are defined, and various support tools for literature research are provided. In this paper, we report a prototype system for a literature database which includes varieties of a Celtic story, *The Dierdire Legend*. The system structures documents based on XML, generates various concordances, and supports link tools among multiple documents, including personal memos. Further, in this paper, we discuss representation and management of user-specified objects, which cannot be described in XML, and their management.

1 はじめに

『デアドラ伝説』[15]は8世紀頃から伝わるケルト(アイルランド)文学の代表的伝承物語であるが、口承によって伝えられたため数多くの変種が存在している。岡山県立大学はその収集において世界的に見て最大規模のひとつであり、現在100編近くを所蔵している。岡山県立大学では上記のような分析・研究のために、1997年からそれらの電子化を行なっているが、当初は、それは出版を目的としたもの

*岡山理科大学総合情報学部

†(株)リョービシステムサービス

で、文学研究のためのデータベース化や支援システムの構築を目的としてこなかった。従来の文学データベースの研究では、電子化された文書を対象に、情報検索や統計処理技術によって、文学研究のための有用な情報を引き出すことがほとんどであり、データベースシステムの観点からのデータベース化は考慮されなかった。

比較文学の分野では複数の作品における類似性や語彙の変化に着目し、作品の系統樹を作成することが大きな課題となっている。とくに『デアドラ伝説』のように変種の多いものにとっては重要である。さらに、変種間やそれらの語彙間の差異を比較文学的に分析することによって、文化的あるいは政治的な背景を考察することができ、より深い文学的考察を行うことができる。このため、作品間の構造的類似性や内容的類似性を可能にするためには、構造化された文学データベースの構築と、それに基づく支援システムがより進んだ文学研究のために必要不可欠となってくる。そこで、これらをサポートした『デアドラ伝説』のデータベース化を行っている。

電子化文書を構造化するためには XML (eXtended Markup Language)[2][24] が一般的になってきており、本研究でも XML に基づくアプローチを採用している。従来構造化は文書の再利用を目的としたものが多いが、本研究で XML を採用した理由は、それがデファクトスタンダードになりつつあること以外に、以下のようなものである。

- 文書スキーマの定義

文書に対する検索機能が必要になると考えられ、文書の意味的構造が明確になっていることが重要である。言い換えれば、文書スキーマが明確に定義できることが必要である。XML では、DTD (Document Type Definition) によって独自の文書型定義を行うことができる。

- リンク機能

XML では、XLink[12], XPointer[13] という2つの規格によって、強力なリンク機能を提供している。具体的には、文書中の任意の粒度のオブジェクトをアンカーとできる、3つ以上のアンカーに対してリンクを張ることができる、文脈外リンク (out-of-line link)¹によって元の文書に手を加えずにリンクを張ることができる、などの機能を有用と考えている。

¹XML に関する技術用語の日本語訳は [24] によった。

- インターネットでの技術的優位性

HTML と同様、WWW での利用を考慮して作られた規格であり、Netscape Navigator や Internet Explorer などの主要な WWW ブラウザでの対応が期待できる。また、オフィスアプリケーションでも将来の XML 対応を公表しているものもあり、文書フォーマットとしてデファクト・スタンダードとなりうるものが期待できる。

- レイアウト情報の定義

見栄えにこだわった WWW ページが氾濫している現状から分かるように、電子文書といえども、それを扱うのが人間である以上、レイアウト情報を保持する機能は必要になる。XML では、CSS (Cascading Style Sheet)[23, 1], XSL[3] などのレイアウト定義言語を持ち、レイアウト情報を定義することができる。

XML で特定するオブジェクトが (表題、著者、章、段落など) が他の文書とのリンクの単位となると共に、検索等のデータ操作の単位ともなる。

本稿では、まず2節でこれまでに実装したプロトタイプシステムの概要を説明し、3節から7節まではその中の個々のサブシステムについて報告する。8節でプロトタイプシステムの経験から今後の拡張点について、とくに XML の構文的制約で記述できないオブジェクトについて議論し、まとめる。

2 プロトタイプシステム

電子化されたテキスト文書は構造を持たないフラットファイルであり、そのままでは表層的な分析しか行なうことができない。そこで、深層的な分析が行なえるよう、個々の文献に対して基本的に、以下に示す4種類の構成要素を持たせている。

- テキスト文書

個々の物語を分析や加工処理を施さず、単に電子化したフラットファイル。

- 構造化文書

テキスト文書に構文的な概念を持たすために章、段落、文などに構造化したものの。構造化は XML (eXtensible Markup Language) に用いられるようなタグを挿入することで行う。この構造化情報により文章の位置を特定することができる。

- 内容記述
物語のストーリーを個々の場面に分割し時系列で結び、トップダウンで記述したものの。これは抽象化のレベルにより複数の記述がありうる。
- メモ文書
原テキストに対応させてユーザ独自のメモを記述したもの。

内容記述もメモもXMLで記述されるので、共通のリンク機能によって、関連付けることができる。つまり、リンク機能をサポートした問合せ機能は、元の文書、内容記述文書、メモ文書（個人文書）を関連付けながら検索することが可能になる。

上記の4種類の文書をもつ文学データベースのためのプロトタイプシステムは、図1に示すように、構造化システム、索引抽出システム、検索システム、内容記述システム、メモ記述システムの5つのサブシステムから構成される。ここでは内容記述システム以外の4つについて説明する。

3 構造化システム

プロトタイプシステムの構造化システムでは、最小構成要素を文書の文単位として構造化を行なうこととした。文書を構造化するための構造識別子は、表1に示す7つである。

表 1: 各構成要素の構造識別子

構造識別子	構成要素
{title}	表題 (著者等も含む)
{chapter}	章名
{section}	節
{paragraph}	段落
{sentence}	文 (韻文も含む)
{stanza}	韻文
{source}	出典

一般の構造化文書は再利用を主たる目的としたものが多く、構造化の粒度はもっと大きなものがとられるが、文書の構文的条件を考慮した検索やリンクの粒度等のために、文単位としてある。また『デアドラ伝説』は散文だけではなく、韻文や戯曲、それらの混在したものなど、さまざまな形で書かれているが、構造化するためには、構文の文法記述が必要で、プロトタイプシステムではその単純化のために、散文と韻文を含む散文にとどめることにした。

このシステムにより生成される構造化文書は、図2のようにXMLに基づいたタグを構造

識別子としてテキスト文書の各構成要素の前後に挿入することで構文的概念を付加したものである。

```

<section number="1">
<paragraph number="1">
<sentence number="1">
What caused the exile of the sons of Uisliu?
</sentence>
<sentence number="2">
It is soon told.
</sentence>
</paragraph>
<paragraph number="2">
<sentence number="3">
The men of Ulster were drinking in the house
of Conchobor's
storyteller, Fedlimid mac Dail.
</sentence>
<sentence number="4">
Fedlimid's wife was overseeing everything and
looking after them all.
</sentence>
.....

```

図 2: 構造化文書

文書からの各構成要素の切り出しは、lexにより字句解析ルーチンを生成することにより行なっている。切り出し規則は、文書内の空白や改行、ピリオドなどのインデント情報を元に行なっている。ただし、対象データとして扱っているデアドラ伝説には、文献によって韻文が含まれるなど文書ごとに論理構造が異なっているため、構造化のための文法記述はある程度文書に依存した形となっている。このシステムで使用しているlexは、入力文字列が2つ以上の規則に適合したり、規則に適合する文字列がいくとおりも存在するような仕様を含むファイルにも字句解析ルーチンを生成する。このとき、字句解析ルーチンは

- 適合する文字数が最大になる規則を優先する
- 適合する文字数が同じ場合には、仕様書ファイルの先に書かれている規則を優先する

という2つの暗黙の選択基準のいずれか1つを採用することにより、トークンの切り出しを行なっている。

このような性質を持つ字句解析ルーチンでは、入れ子状態にある節、段落、文、韻文の規則を一度に定義した場合、字句解析ルーチンは前者の選択基準を採用してしまい、結果として文のみをトークンとして切り出すことになるた

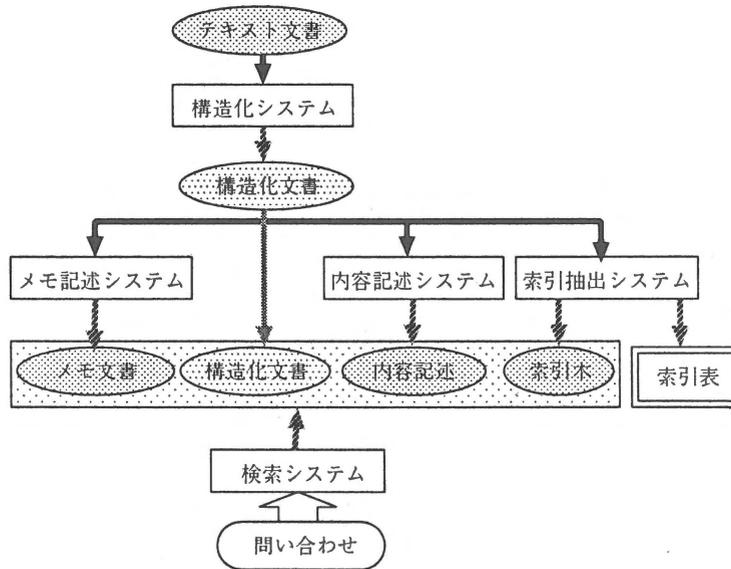


図 1: 文学データベースシステム

め、トークンの切り出しは以下に示す4段階に分け、テキスト文書の構造化を行なう。

1. 電子化されたテキスト文書を入力文字列として、表題、章名、韻文、出典とその他の5つのトークンに切り出し、その他以外のトークンの前後にタグを挿入する。
2. 第1段階で生成されたデータを入力文字列として、節と第1段階でタグの付けられた表題、章名、出典をトークンとして切り出し、切り出された節の前後にタグを挿入する。このとき、節には第1段階でタグの付けられた韻文が含まれていることもあるということを規則として書いておく。
3. 第2段階で生成されたデータを入力文字列として、表題、章名、節のタグ、出典と段落をトークンとして切り出し、切り出された段落の前後にタグを挿入する。ここで、段落に韻文が含まれる可能性があることを規則として書いておく。
4. 第3段階で生成されたデータを入力文字列として、表題、章名、出典、節と段落のタグと文に切り出し、切り出された文の前後にタグを挿入する。ここで、文に韻文が含まれる可能性があることを規則として書いておく。

図3に構造化システムのインタフェースを示す。

この構造化システムで生成される構造化文書には一文毎にタグが付けられているので、必要

な情報の存在する位置の論理的な特定を行うには良いのだが、フラットファイルとしてみた時に構造識別子という付加情報が含まれており、図2のように構造化文書だけを見て文章を理解するのはなかなか難しい。そこで、構造化文書から構造識別子を抜きとり成形し、図4のようにテキスト文書として表示するブラウザの構築も作成している。これにより、構造化文書一つでテキスト文書の役割も果たすようになる。

4 メモ記述システム

文書が紙に書かれている場合、文書中にユーザが独自のメモを書き込むということがあるが、これに近い機能を電子文書上でも行なえるようにするシステムがメモ記述システムである。

メモ記述システムでは、先に述べた構造化文書ブラウザを利用して、構造化文書からでもタグを取り除いたテキスト文書からでも、メモを付加したい任意の文字列を指定することが可能である。指定した文字列に対してメモ文書を作成し元文書とリンクを張ることでメモを付加していく。

任意の文字列をメモを付加するオブジェクト(メモオブジェクト)として指定する方法として以下の2つが考えられる。

- メモオブジェクトとしたい任意の文字列にタグを付けることで構造化文書を変更する
- 構造化文書から指定したメモオブジェクトの位置情報を抽出する

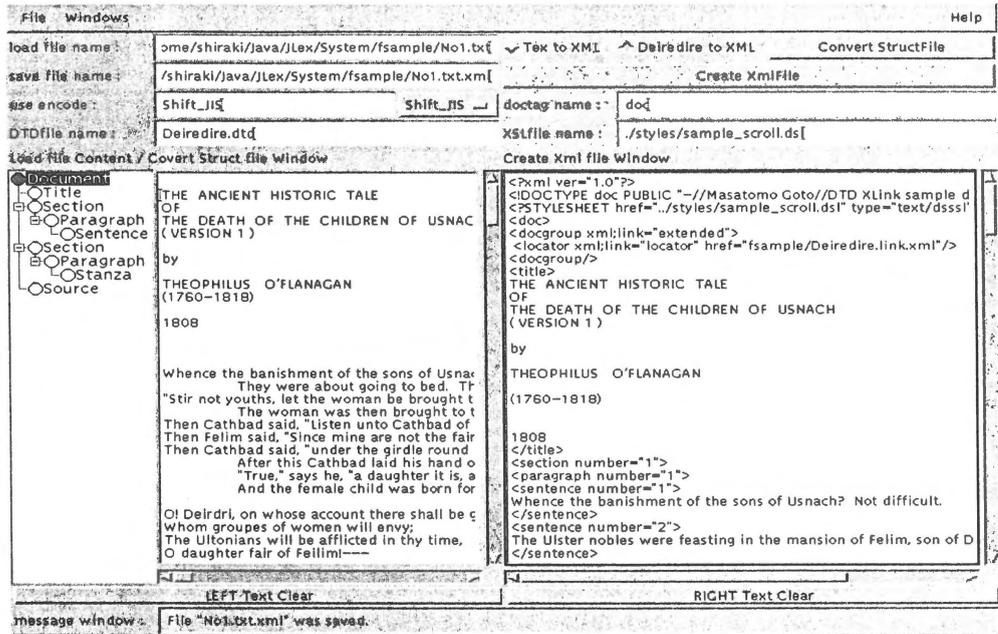


図 3: 構造化システムの実行例

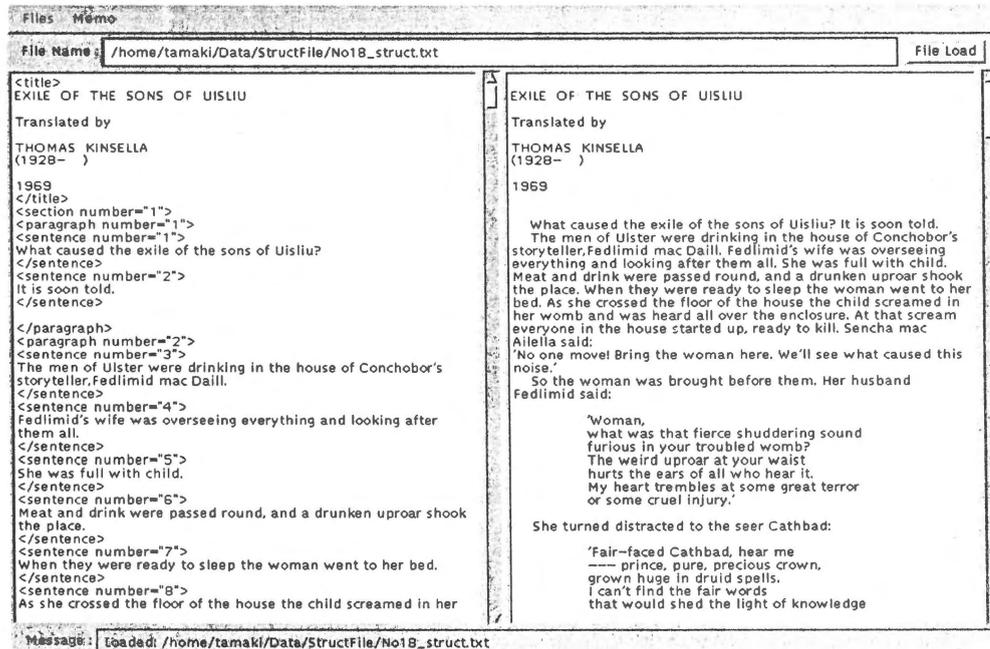


図 4: 構造化文書ブラウザ

前者のタグ挿入形式でメモオブジェクトを指定した場合、構造化文書にタグが直接書き込まれるため構造化文書が変更される。また、メモオブジェクトとして指定した任意の文字列が既存のメモオブジェクトと重なるとき、メモオブジェクトが入れ子状態になっていれば良いのだが、少しずれた形で指定されている場合、メモオブジェクトを図5のように2つに分割してタグを付けるなどの処理が必要となる。ただし、

× `<para>...<kw>...</para>`
 `<para>...</kw>...</para>`
○ `<para>...<kw>...</kw></para>`
 `<para><kw>...</kw>...</para>`

図 5: アンカーの分割

これはシステム側の処理上の問題であり、ユーザには関係のないことであるから、実際に表示させる時には分割されたメモオブジェクトを、あたかも一つのものであるかのように表示しなければならない。メモオブジェクトの前側に挿入される開始タグには、各メモオブジェクトに対するメモIDとユーザIDが情報として付加される。

後者の位置情報によるメモオブジェクト指定方式では、構造化文書からメモオブジェクトの存在する位置情報を見るだけなので、構造化文書が変更されることはないが、ユーザが文書に直接変更を加えた場合、メモオブジェクトの存在位置を指定しているため、位置情報にも変更を加えなければならなくなる可能性がある。本研究で処理対象データとして扱っているデアドラ文書は、読み込み専用データであるため、構造化文書を参考文献として位置情報を取り出す形式を採用することとした。

実際に文書にメモを付加するには、構造化文書ブラウザ上で任意の文字列を選択するか、もしくはメモ記述エディタに直接メモオブジェクトを指定する。ただし、構造化文書のタグは構成要素の判別のために付加された構造識別子なので、現在の実装では、それをメモオブジェクトとすることはできない。また、メモ記述エディタからメモオブジェクトを指定した場合、指定文字列が構造化文書内に複数箇所ある場合はユーザに確認を取る。指定がなければ最初に現れた文字列の位置がメモオブジェクトの位置情報として記録される。また、メモは基本的に私的なものであるから、各ユーザが付加したメモ文書のみを表示させるようにする必要がある。既存のメモ文書は特定のウィンドウより開いたり、必要でなくなったメモの削除を行なう

ことができる。

5 リンクエディタ

複数の構造化文書を関連付けるため、構造化システムの一部としてリンクエディタがある。これは、文脈外リンクを用いることにより、元文書に手を加えることなく管理することを可能にした。文脈外リンクを用いた理由は、必ずしも文書が書き込み可能ではないからである。リンクの定義がどの異種文書からも参照できる必要があるため、図6のような拡張リンクの定義のみを集めたXML文書(リンク定義文書)を文書群ごとに一つだけ、文書群中の各文書から参照するという方法をとった。この方法を用いると、すでに存在する文書群内のリンクの定義の変更はリンク定義文書に対して行えばよく、元文書に手を加えなくてもよい。

内容的に関連のある文書をひとつのまとまりとして定義することができる。つまり一般的に、XMLの構造化文書を含むハイパーテキスト文書の場合、たとえ内容が同一であっても設計方法は多様で、図7のように結果の文書集合の同定は構造的には困難である。

そこで同一内容の文書集合を文書群として定義している。つまり文書群とは、内容に対応した論理的な単位であって、図8のようにひとつの文書が複数の文書群から共有されても、入れ子構造になっていても構わない。この文書群は、文書単位のリンク付けで十分で、操作の基本的な単位となる。

6 索引抽出システム

索引抽出システムでは、文献中で使用されている語彙の使用例や使用頻度を知ることによって比較文学研究の促進を図ったり、検索の高速化や高度化を図るためのデータを作成する。このシステムは、文書より使用されているすべての単語を切り出すキーワード抽出システムと総索引表(concordance)を生成する索引表作成システムの2つのサブシステムから構成される。

実際に比較文学研究において索引を利用しようとする場合、抽出語や不要語はユーザや使用目的によって異なる。また、文献によって、使われている語が同じとは限らず、すべての不要語をキーワード抽出システム内で自動的に削除してしまうというのは難しい。そこで、単語の切り出しは文書の題名、章名、出典を除く構成要素を対象に、前置詞や冠詞等も区別なくすべての単語を切り出し、切り出された単語から外

```

<docgroup xml:link="extended">
<location xml:link="locator" href="main.xml" />
<location xml:link="locator" href="www.xml" />
<location xml:link="locator" href="ohp.xml" />
</docgroup/>
<relation xml:link="extended">
<point xml:link="locator" href="main.xml#root().child(1, section)" />
<point xml:link="locator" href="www.xml#ID(section_2)" />
<point xml:link="locator" href="ohp.xml#span(ID(p3),ID(p5))" />
</relation/>

```

図 6: 拡張リンクによる関連付けの例

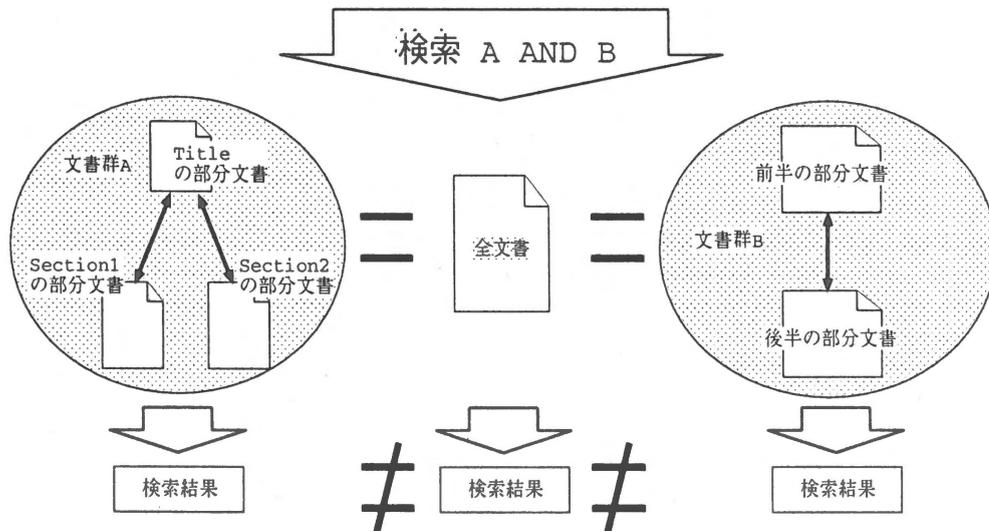


図 7: 同一文書の検索結果の相違

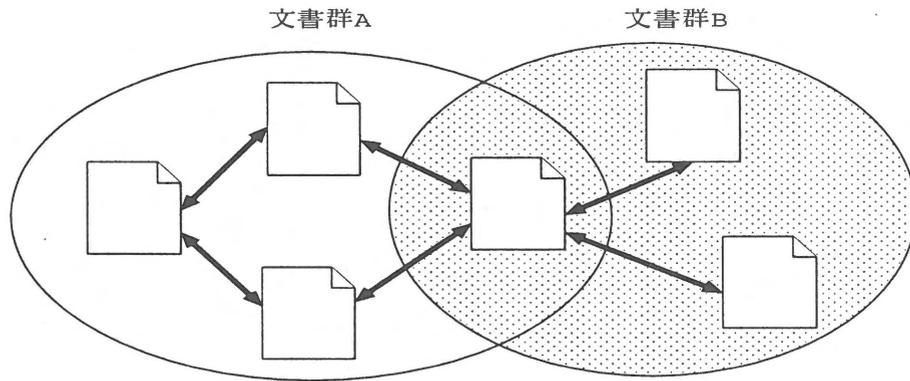


図 8: 文書関連・文書群の概要

部プログラム（索引表作成システムの前処理）によって目的に応じた語を抽出したり、不要語をユーザが削除することで修正キーワードファイルを作成することとした。内容記述のためには固有名詞の切り出しも行う [5]。なお、単語の切り出しは先の構造化システムでも使用した lex を使用し行なっている。

索引表作成システムでは、基本的にキーワード抽出システムにより作成された修正キーワードファイルを用いて総索引表を作成する。生成される索引表には、キーワードとその前後の文脈を表示させる索引表と検索・内容記述システムにおける処理対象データとして使用する索引表の 2 種類がある。前者の索引表は、一般に使用される KWIC 索引と KWOC 索引の両者の特徴をいかしたもので、出現頻度表の生成も行なわれる。これは図 9 に示すように、出現頻度とキーワードを使用頻度が高い順に並べて変えたのもで、これにより、どのキーワードがもっとも多く使用されているかが視覚的にわかるようになる。頻度表のキーワードをクリックすることにより、それに対応する KWIC 索引を表示できる。ここで、ロケーション情報とは一番右端にある“(2,30,132,20)”のような情報のことで、表示されている数字はキーワードが存在する節、段落、文、韻文中にキーワードが存在する場合は韻文の番号で構成される。このロケーションは最初は文単位で作成されているので、それを操作することで、パラグラフ単位や節単位などで表示することも可能である。

さらに、異なる文書から抽出された出現頻度表ををマージすることによって文書間の比較を容易に行なうことができるようになる。実際にマージを行なった例を図 10 に示す。

これによって『デアドラ伝説の』のさまざまな変種を比較することにより、文学的考察を深

Keyword	Not	Wol
Ailleen	11	0
Ainli	0	2
Alba	0	5
Albain	6	0
Allil	0	2
Although	0	1
Anell	0	1
Ardan	11	2
Ardan	1	0
Art	8	0
Assassinated	0	1
Atchy	0	1
Atha	2	0
Ballyshanon	0	1
Being	5	2
Beloved	1	0
Bewall	0	1
Binedar	0	1
Boinne	6	0
Boyb	6	0
Break	0	1
Cathbad	0	6
Caught	0	1
Clann	2	0
Close	2	0
Colam	17	0
Conachar	49	0
Conor	0	29
Cormac	0	2
Cruaidh	6	0
Cruitire	16	0
Cuilicnn	6	0
Dall	0	1
Dear	0	3

図 10: マージした出現頻度の表示

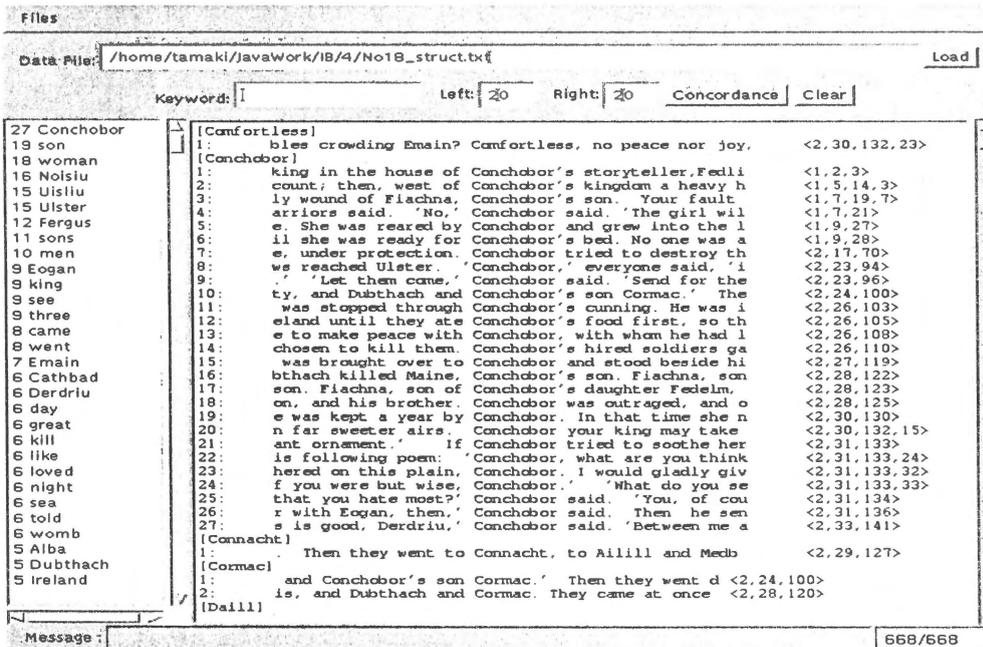


図 9: 出現頻度の表示

めることもできる [16, 17]。

7 検索システム

検索処理はまず、索引抽出システムから出力されたキーワード表から、必要となるキーワードを抽出し、索引木 (B+ 木) を生成している。従来の情報検索システムと異なるのは検索結果である。

単独の文書を対象にした結果得られるのは文書の部分構造の集合である。XML で構造化された文書は木構造になっており、ロケーションは木の葉 (文) を指している。B+ 木はこのノードアドレスを指しており、ロケーションの階層を操作することによって、同一文章中、同一パラグラフ中などの構文的条件を検索条件に付加できる。したがって検索結果の構造のレベルは検索条件によって異なっている。基本的な操作は以下のようになっている。

- 縮退: Reduce Loc by i

$$Loc = [l_1, l_2, \dots, l_n] \Rightarrow [l_1, l_2, \dots, l_{n-i}]$$

- ロケーション検索:

Select Loc where (tag) が C を満足する

{ Loc | D の Loc' の (tag) が

(C から得られた) C' を満足し、

Loc は縮退操作によって得られ、

C を満足する }

- 文書検索: Select Doc where ...

Reduce ロケーション検索の結果

by 文書識別子

文書群は文書の集合として定義されるので、文書群に対する検索では、

- OR 条件では得られた文書を含む文書群の探索
- AND 条件では、AND を除去して文書群定義に含まれた文書を検索した結果、再度 AND の評価をおこなう

の 2 種類の操作が必要となる。基本操作は以下の 4 種類である。

- Select 文書群 G where C

$$\{G | G \text{ の } D \text{ が } C \text{ を満足する}\}$$

- Select 文書群 G where $C_1 \wedge C_2$

$$\{G | C_1 \text{ を満足する } D_1 \text{ と } C_2 \text{ を満足する } D_2 \text{ を含む } G\}$$

- Select 文書群 G where $C_1 \vee C_2$

- Select 文書群 G where $\neg C$

文書群は一般的には検索条件を満足する極小集合の定義としている。

4 節で述べたメモを含む検索も、この文書群としての検索となる。さらに 3 節で述べたよう

に、『デアドラ伝説』は構造化文書やメモ文書の他に、抽象階層をなす内容記述の文書を持っている。この内容から元文書の構造を引き出すのもこの文書群検索がおこなう。

このように文書群の操作はさまざまな内容を持っている。つまり

- デアドラのようにさまざまな変種を統合すること
- 内容的に関連がある複数のデータ形式の電子化文書を論理的に統合する
- ハイパーテキストの設計の多様性に対して、内容的なまとまりを定義する
- 電子化文書の更新履歴をまとめて管理する
- メモや内容記述のような付加情報をユーザごとに統合する。

などである。

8 構造化文書の拡張

プロトタイプシステムでは多くの機能の積み残しがある。それらは以下のものである。

- 文書の再帰構造
- 構造の細分化
- 意味オブジェクトの拡張
- ラッパーとタグのフィルタリング
- データベースエンジンとの結合

構造化文書は一般的に再帰構造になる。これはロケーション情報を木にもつ索引木と、構文条件をロケーション情報で操作する縮退操作に影響する。そのためにロケーション情報に構文情報(タグ情報)を付加するよう拡張することで解決できる。ただし問題が系統的に解消できたとしても、「同一文章中にキーワード A と B を含む」という検索条件には多義的となるので、ユーザはそれを知らなければならない。

現在の構造化は文レベルであるが、さらに自然言語処理技術を使って詳細化することも検討している。一般的に『デアドラ伝説』のように長文の物語の内容記述は困難であるので、現在の内容記述は、場面の時系列として物語を構成し、各場面に登場人物や場所を付加するに留めている。場面は抽象化できるので抽象階層としても扱っている。因果関係など人が記述しても困難なものも多いが、意味内容を反映したオブ

ジェクトを指定できれば、より深い内容検索を行うことができる。

構造化文書にメモを付加したり、内容記述文書等の対応をとることはプロトタイプシステムでも行っているが、それは元文書が更新されないという強い仮定に基づいていた。さらに一般的に意味内容を考えたリンク付けを行うためには、XML による構文的オブジェクトと意味オブジェクトの有機的なリンク付けが必要である。またいったん意味オブジェクトを導入すると、意味オブジェクトに対する索引木や検索が必要となる。そのために現在、意味オブジェクトの記述言語として、演繹オブジェクト指向パラダイムに基づいた知識表現言語 QUIK [?, 19, 25, 10] を予定している。現在記述実験を行っており、両オブジェクト間の意味管理やリンク管理の機能要件を明確化している。意味オブジェクトは、電子化文書の個別化や情報統合、異種文書管理で重要な役割を果たす [7, 8, 10]。これらの問題を解決するため、図 11 のような統合管理システムを構想し、研究を進めている。

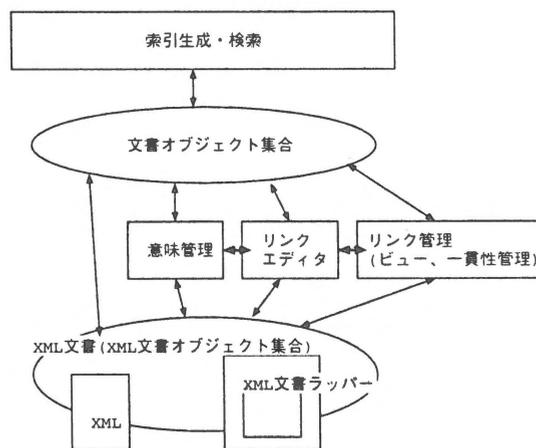


図 11: 統合管理システムの概要

現在は対象とする文書の種類が少ないが、『デアドラ伝説』関係の解説や付属資料も付加するとすれば、さまざまなデータ形式を対象としなければならない。そのためには元のデータ形式を XML 形式に変換するラッパーの位置付けが重くなってくる。そのときは、変換を可逆性を保証するためのタグのフィルタリングを考えている。これはプロトタイプシステムでの構造化文書ブラウザの一般化でもある。

今回のプロトタイプシステムではデータベースエンジンとの結合まではできなかった。現在実験を開始しているのは、ObjectStore PSE Pro (for Java) である。システムはすべて

Java で実装されているので、Java オブジェクトとして蓄積し、管理する予定である。意味オブジェクトは QUIK オブジェクトとなるので、QUIK と Java の連携機能も現在実装中である。

本論文では、ケルト文学の『デアドラ伝説』を対象にした文学データベースのプロトタイプシステムについて述べ、その拡張として、さらに文学研究を進めるための課題について議論した。本研究の特徴は以下のようにまとめることができる。

- XML による構造化文書を対象にした文学データベースを指向している。
- リンク機能により、関連文書や文書群にたいする操作を実現した。
- 意味オブジェクトを導入することにより、電子化文書の管理機構を拡張している。
- 文学研究の要件を反映したツールを整備しつつある。

このプロトタイプシステムの経験から、いくつかの新しいシステムを進行中である。本稿の文学データベースの文学研究向けの強化、民間説話データベースの構築、マルチメディア情報を含む異種文書管理システム、電子化文書の効果的提示システム（デジタルライブラリ、デジタルミュージアム、マルチメディアブック、戯曲の視覚化）、デジタルテーマパーク、QUIK の改良、などであり、現在平行して研究開発を進めている。

謝辞

さまざまな議論を頂く岡山県立大学横田研究室の皆様へ感謝します。なお、本研究の一部は文部省科学研究費特定領域研究(A)「高度データベース」および基盤研究(C)による。

参考文献

- [1] Bert Bos and Hakon Wium Lie and Chris Lilley and Ian Jacobs: "Cascading Style Sheets, level 2", *W3C Recommendation REC-CSS2-19980512* (May 1998)
available from <http://www.w3.org/>
- [2] Tim Bray and Jean Paoli and C. M. Sperberg-McQueen: "Extensible Markup Language (XML) 1.0", *W3C Recommendation REC-xml-19980210* (Feb, 1998)
available from <http://www.w3.org/>
- [3] James Clark and Stephen Deach: "Extensible Stylesheet Language (XSL)", *W3C Working Draft WD-xsl-19981216* (Dec, 1998)
available from <http://www.w3.org/>
- [4] 八村広三郎: "人文科学とデータベース", 「情報処理」 Vol.38 No.5 通巻 387 号, 5 月, (1997).
- [5] 本行弘明、池口仁誠、三宅忠明、横田一正: "分散環境での文学データベースの内容検索", 第 55 回情報処理学会全国大会, 九州, 9 月, (1997).
- [6] 本行弘明、白木善隆、田植明子、国島丈生、横田一正: "文学データベースのためのプロトタイプシステムの実装", 第 57 回情報処理学会全国大会, 名古屋, 10 月, (1998).
- [7] 国島丈生、横田一正、白木善隆、劉渤江: "XML リンク機能による異種文書の統合方式", 情報処理学会データベースシステム研究会資料 vol.117, No.7, 1 月 (1999).
- [8] Takeo Kunishima, Kazumasa Yokota, Bojiang Liu, and Tadaaki Miyake: "Towards Integrated Management of Heterogeneous Documents", *Cooperative Databases and Applications '99*, pp.39-51, Springer, Sep., 1999
- [9] Bojiang Liu, Kazumasa Yokota, and Nobutaka Ogata: Specific Features of the QUIK Mediator System, *IEICE Transaction on Information and System*, Vol. E82-D, No.1, pp.180-188 (1999).
- [10] 劉渤江、横田一正、国島丈生、三宅忠明: "異種文書の統合のための意味とリンクの管理機構", 電子情報通信学会データ工学ワークショップ, 5B-2, pp.1-8, 鹿児島, 1999 年 3 月.
- [11] W. Li, and Y. Wu: "Query Relaxation by Structure for Web Document Retrieval with Progressive Processing," *Proc. Advanced Database Systems Symp.*, pp.19-25 (1998).
- [12] Eve Maler and Steve DeRose: "XML Linking Language (XLink)", *W3C Working Draft WD-xlink-19980303* (Mar, 1998), available from <http://www.w3.org/>

- [13] Eve Maler and Steve DeRose: XML Pointer Language (XPointer), *W3C Working Draft WD-xptr-19980303* (Mar, 1998), available from <http://www.w3.org/>
- [14] Dennis R. McCarthy and Umeshwar Dayal: "The Architecture of An Active Data Base Management System," *Proc. ACM SIGMOD International Conference on Management of Data*, pp. 215-244 (June, 1989).
- [15] Tadaaki Miyake: *Select Version of DEIRDRE*, University Education Press, May, (1998).
- [16] 三宅忠明、横田一正、國島丈生、田槇明子: "ケルトの悲恋ロマンス「デアドラ」における資料の言語分析", 人文科学とコンピュータ 99 シンポジウム, pp.103-104, 大阪 (1999).
- [17] 三宅忠明、横田一正: "情報工学的アプローチによる文学研究," 英語青年, Feb. (2000).
- [18] 村田真. XML 入門. 日本経済新聞社, Jan. (1998).
- [19] Toshihiro Nishioka, Kazumasa Yokota, Chie Takahashi, and Satoshi Tojo: "Constructing a Legal Knowledge-base with Partial Information", *Proc. ECAI'94 Workshop on Artificial Normative Reasoning* pp.40-55, Amsterdam (Aug., 1994)
- [20] 大田友一、横田一正、西田豊明、佐藤哲司: 情報の共有と統合、岩波講座マルチメディア情報学、7 卷、岩波書店, 12 月 (1999).
- [21] 白木善隆、本行弘明、田槇明子、国島丈生、横田一正: "異種文書データの統合管理システムの構想", 電気・情報関連学会中国支部連合大会, 岡山, 10 月 (1998).
- [22] 田槇明子, 本行弘明, 白木善隆, 國島丈生, 横田一正: "文学データベースのための索引・検索機能の拡張", 電気・情報関連学会中国支部連合大会, 岡山, 10 月 (1998).
- [23] Hakon Wium Lie and Bert Bos: "Cascading Style Sheets, level 1", *W3C Recommendation REC-CSS1-961217* (Dec, 1996) available from <http://www.w3.org/>
- [24] XML/SGML サロン: 標準 XML 完全解説, 技術評論社, 5 月 (1998).
- [25] 横田一正: "演繹オブジェクト指向データベース Quixote の法的推論への応用", 人工知能学会誌, vol.10, no.1, pp.24-30 (1995).
- [26] Kazumasa Yokota, Yutaka Banjou, Takashi Kuroda, and Takeo Kunishima: "Extensions of Query Processing Facilities in Mediator Systems," *Proc. Int. Workshop on Knowledge Representation Meets Databases (KRDB'97)*, p.17.1-8 (Aug., 1997).

相対インデックス法を使った文構造分析

Data Mining of Sentence Structures using Relative Indexing of Vertices

雄山真弓、岡田孝、黒崎茂樹

Mayumi OYAMA, Takashi OKADA, Shigeki KUROSAKI

関西学院大学情報メディア教育センター

Center for Information & Media Studies, Kwansei Gakuin University
〒662-8501 兵庫県西宮市上ヶ原, Nishinomiya, Hyogo, 〒662-8501

キーワード: データマイニング、相対インデックス、構文解析木、知識発見

Keywords: data mining, relative indexing, syntactic parse tree, knowledge discovery.

あらまし

データマイニングは、大規模データベースから新しい知見の発見、これまで予期できなかったパターンの導出、新しいルールの導出を行う方法である。コーパスは大規模データベースである。最近ではテキストマイニング技術も急速に進歩してきた。データマイニングの手法をコーパスからの知識発見に適用すれば、良い結果が得られるであろう。しかし、構文解析木で表現されたコーパスを分析するには、構造を扱う処理を行わなければならない。本論文は、構文解析木の表記法について論じる。次に、相対インデックス法を使って、構造を持つデータからデータマイニングを行う方法について解説する。最後に、英語のコーパスデータを使い、自動詞の *think* と他動詞の *think* を含む文の構文構造上の特徴について、実際の分析により得られた結果を報告する。

Summary

Data mining deals with the discovery of hidden knowledge, unexpected patterns and new rules from large databases. Corporuses are large databases and text-mining techniques are making rapid progress. If we apply the technique of data mining to discover knowledge in corporuses, we will get good results. But, to analyze corporuses with syntactic parse trees we have to treat their structures. In this paper, we show the inscription of syntactic parse tree first and we propose the data mining methods using the relative indexing of vertices for knowledge discover from corporuses. We show the distinctive structural features found between an intransitive verb "think" and a transitive verb "think".

1. はじめに

本研究の目的は、構造をもつデータを対象としたデータマイニングツールの開発とその応用を行うことにある[1][2]。言葉や文は構造を持つデータである。我々は、複雑な表現をしようとすればするほど、文のなかで使われる単語の係り受け構造が複雑になる。構文解析木（以下、構文木）はこの係り受け関係の複雑さを括弧などによって表現したリスト形式のデータである。近年では、日本語や英語の膨大な文例を集めた大規模データベースの蓄積が行われている。特に、日本電子化辞書研究所で作成されたEDRコーパスは、英語や日本語の文例と共に、文を構成する単語の品詞やその共起関係、文の係り受け構造を表現する構文木などを含むコーパスデータである [3]。構文木のような、構造をもつ大量のデータは、テーブル型のデータベースでは蓄積できないため、構造そのものを分析する場合、一般の統計分析の手法では扱うことが難しくなった。

ここで述べる構造のあるデータの分析とは、単に構文木全体のパターンを分類する方法ではなく、文のなかに含まれる単語の使い方等で異なるグループに分けられたデータから、それぞれのグループが持つ特徴的部分パターンを探索することを意味する。例えば、英語の文例から動詞（例えば think）を1つ選択しそれが自動詞として使われる場合と他動詞として使われる場合で、その動詞の置かれた位置を視点として、全ての周辺構造とその属性を調べて、自動詞としての使い方と他動詞としての使い方に違いや共通点等が見いだせるかどうかを探索し、新しい知識を発見するものである。

本論文では、最初に構文木の構造表記法について考察した後、データマイニングでよく使われる各種技法毎に、構造情報の取り扱いの可能性を評価する。さらに、構文木の構造表記法として提案した相対節点インデックス化を英語コーパスに適用し、構造、単語、品詞を組み合わせたデータを作成する。最後に、構文木に含まれる品詞とその構文木中の位置関係で、使われる頻度の多いルールを得た結果を報告する。

2. 構造の表記法

構文木は順序づけられた有向木であり、順序付きでなくまた閉サイクルを含む化学グラフなどと比較すると、多様な表記法が可能である。通常のコーパスでは、括弧で表現したリスト構造で表されており、これが一

般的な表記法であろう。図1(a)に示した“My son has a red pen”の構文木のリスト表記例を次に示す。

My son has a red pen

```
(S (NP (pron "My") (n "son"))
  (VP (v "has")
      (NP (a "a")
          (NP (adj "red")
              (n "pen")))))
```

なお、各節点の先頭に斜体で付した記号は通常の文法での概念である。

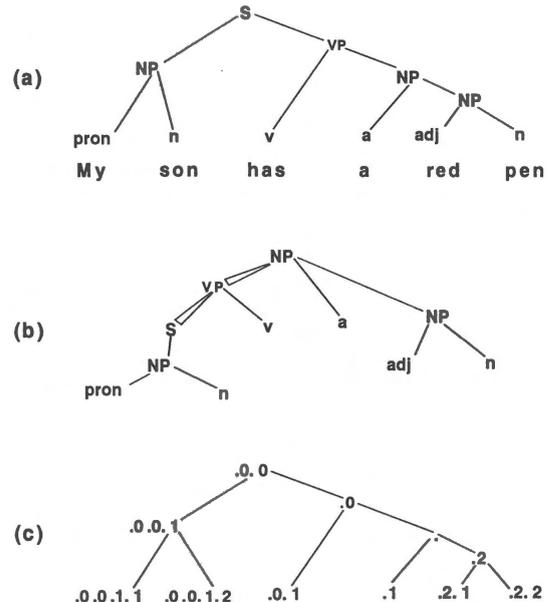


図1. 木構造表現法

ここでは、構文木中に特定の視点となる節点を定めた場合に、可能となる2種の表記法について考察する。上記で“a red pen”の名詞句を視点として定めた場合を例とする。

(A) 視点を摘んで持ち上げ振り回すと考えれば、視点が根節点となるように変換され、図1(b)のようになる。なお、元来の親節点は左端になるように配置し、それを明示するため、くさび形の記号で示している。実際にこの木をどの様に符号化するかは各種の方法が考えられ、考え方だけを取り入れて、変換せずに元の木のままで扱うこともできる。この木は、各節点の深さが視点からの距離に対応しており、視点を含む部分木を探索する際に便利である。この技法を逆転木化と名付けることとする。

(B) 図1(c)は、図1(a)に示す木のすべての節点に、視点を原点としてユニークな番号付けを行った例である。すべての節点は、原点からその節点に至るユニークな経路を有する。そこで、経路中で親節点への移動には .0 を、

構文木	単語	品詞	A	B
and	and	CONJ	[4].1	[4]-3
—	—	BLNK	[4].2	[4]-2
theatre	theatre	NOUN	[3].1.1	[3]-3-2
s	s	SUF	[3].1.2	[3]-3-1
—	—	BLNK	[3].2	[3]-2
think	think	VT		
—	—	BLNK	[1].2	[1]-1
they	they	PRON	[2].2	[2]-3
—	—	BLNK	[2].3	[2]-2
have	have	VT	[2].4.1.1	[2]-1-5-2
—	—	BLNK	[2].4.1.2	[2]-1-5-1
problem	problem	NOUN	[2].4.2.1.1	[2]-1-4-2-2
s	s	SUF	[2].4.2.1.2	[2]-1-4-2-1
—	—	BLNK	[2].4.2.2	[2]-1-4-1
with	with	PREP	[2].4.3	[2]-1-3
—	—	BLNK	[2].4.4	[2]-1-2
journalist	journalist	NOUN	[2].4.5.1	[2]-1-1-2
s	s	SUF	[2].4.5.2	[2]-1-1-1
.	.	PUNC	[5].2	[5]-1

図2. 構文木構造と単語の品詞および相対節点インデックス表現AとB (視点: think)

付し、子節点への移動には、左側から数えた子節点番号を付すと図に示した番号付けができる。この技法を、相対節点インデックス化と名付けることにする。

この様にして得られた番号は、視点から各節点への相対的な位置情報を表しており、節点付随の情報と組み合わせることにより、構造と節点内情報の相関を探索する上で重要な手がかりを与えてくれる。なお、子節点の番号付けは右側から数える方法もあり、最適な番号付けは課題に応じて選択すべきものであろう。

図2は、英文 [And theatres think they have problems with journalists.] について構文木の係り受け構造と記号を含む単語とその品詞、think を視点とした相対節点インデックスの値を記号化して表現している。節点にぶら下がる子供の数を左側から数えた表現をAで、右側から数えた表現をBに示している。[4]は視点(think)から数えて4番目の親の位置 0.0.0.0. を表している。Aの表現の[3].1.1は、視点thinkから3つ目上の親に付いた子供中で、左側から数えて1番目の子供のそのまた左側から1番目の子供を意味する。Bの表現では[3]-3-2はthinkから3つ上の親について子供中で、右側から数えて3番目の子供に移りそのまた右から数えて2番目の子供を意味する。A、Bの表

現は、共に相対節点インデックス法を用いている。分析では、両方のデータを使うことによって詳しい情報が得られる。本研究で使用した相対節点インデックス化データは、変換プログラム EDRitem を LISP で作成し、EDR英語コーパスデータを入力データとして作成した。

3. データマイニング技法と構造情報の取り扱い

現在データマイニングに利用されている主要な技法は以下のようなものであろう。

1. 多変量解析
2. 決定木およびルールの導出 [4]
3. 帰納論理プログラミング [5]
4. Graph based induction [6]
5. ニューラルネットワーク
6. クラスタリング
7. 相関ルール探索 [7] [8]

これら技法は、その内容として教師付き学習と教師なし学習の何れかに限定されているものと、その双方を含むものがある。また、構文木のトポロジカルな形

状だけに興味のあるケースは少なく、実際に使われている語まで掘り下げて解析する場合には、属性値の種類が非常に多くなることを考慮する必要がある。属性値数を制限して解析する場合には、意味論にまで踏み込んで概念階層化等の方法により、何らかの符号化を行う必要があると考えられる。以下、各技法について、構造情報取り扱いへの拡張可能性を簡単に評価しておく。

3.1 多変量解析

多変量解析の方法は、本来連続数値情報の解析から発展してきた。これまでの多くの研究成果もあり、信頼性のある方法論である。従って、内容についての解説は不要であろう。

この方法では、構造情報を直接取り扱うことは、原理的に無理がある。しかし、特定の語句に視点を定めた上で解析する場合、以下のような方法が考えられる。

- a. 相対節点インデックス化により得られた節点番号を属性とする。
- b. 個々の文における当該位置の語句の内容を属性値とする。

こうすれば、通常の表形式のデータがそろふこととなり、各種の解析が可能である。この考え方は、多変量解析に限らず他の技法でも適用できる。なお、属性値数が多いことと missing value が頻発することに留意し、それに対応できる技法を選択すべきである。

3.2 決定木およびルールの導出

決定木は教師付き帰納学習で多用される簡単な方法であり、計算時間も早い。しかし、前項で述べた相対節点インデックス化による方法は、原理的には適用できるが、属性値の種類が多く、現実的には適用困難であろう。

この方法の適用対象を逆転木に拡張し、浅い節点から順に識別力の高い変数を選択して、決定木を構成する方法は、我々がすでに発表した[9]。しかしながら、決定木は元来、識別に用いる変数を greedy に探索するため、識別力のない節点を飛び越して、その節点に繋がる深い節点に存在する良い知識を得ることは難しい。機械的に判別すればよいと言う立場ではなく、興味ある知識を得ると言う立場からは、この方法は制御が困難であろう。

3.3 帰納論理プログラミング

発見概念を記述する述語を自動的に生成するもつと

も柔軟な方法であるが、大規模コーパスへの適用は計算時間の点から当面困難と予想される。

3.4 Graph based induction

この方法によると、連結された特徴的な部分グラフを高速に探索することができる。また、教師付き学習の場合にも拡張されている。属性値数を制限できれば、今後有効に活用できると期待される。

3.5 ニューラルネットワーク

逆伝搬型、あるいは Kohonen 型の何れのネットワークも、基本的に構造入力に困難である。しかも、結果として得られるネットワークの解析が困難であるため、相対節点インデックス化により、何らかの自動識別を行う場合以外は適用範囲に限られるであろう。

3.6 クラスタリング

この方法のベースには、ユークリッド空間であれ、記号空間であれ、基本的には事例や生成されたクラスター間の距離概念が必要である。相対節点インデックス化を行い、また属性値の一致数により距離を定義する可能性が考えられる。適当な概念階層化により合理的な距離定義が得られるならば、興味ある結果の得られる可能性がある。

3.7 相関ルール探索

元来マーケティング調査を行うため、顧客が購買するアイテム間の相関を調べる目的で提案された方法である。最小サポート数の概念を設定することにより、大容量のデータベースの解析を可能としている。全く異なった分野からのものではあるが、対象が表ではなくアイテムを要素とするリスト構造であったため、時系列的な購買行動の解析、要因・結果分析などに拡張されている。さらに、扱うアイテム群の概念階層化も実用化されている。

また、元来教師なしの学習を対象としていたが、出力されるルールの中からクラス記述を帰結部に有するルールのみを解析することで、教師付き学習への適用も試みられている。さらに筆者らは最近、この方法の体系的な枠組みを一新することにより、カスケードモデルによる教師付きの学習法を提案している[10]。4.ではカスケードモデルを使った結果も報告する。

これらの発展経過を見ると、相関ルール探索は構文木からのデータマイニングに今後大きな役割を果たす可能性がある。以下、いくつかの具体的な方法論を提

案してみよう。

1. 相対節点インデックス化により得られた節点番号と節点属性をアイテムと考え、通常の相関ルール探索を行う。元来多種多様な商品を対象としたものであるため、属性値の概念階層化などを考慮する必要がなく実行可能である。その結果から、特定の位置関係にある複数の語間の特徴的な相関関係を得ることが期待できる。
2. 上記と同様の処理を、クラス属性を加えてカスケードモデルにより解析すれば、識別に有効な語間の相関関係を得ることが期待できる。
3. アイテムのリスト構造を部分木に拡張することにより、個別節点を抽出するだけでなく、特徴的な部分木全体を獲得することが期待できる。かなりのシステム開発が必要となろうが、大きな成果が期待できる方法論である。

4 相対節点インデックス化を使った構造データの分析

4.1 入力データについて

EDR英語コーパス(約16万文例)から、動詞 think の自動詞(182文例)と他動詞(300文例)を含む文を全て検索し利用した。検索・変換プログラム EDRitem で、データの形式はデータ番号、単語、相対節点インデックス番号、品詞から構成される。相対節点インデックス記号は枝の数を右側から数える場合(A)と左から数える場合(B)がある。以下にデータ構成を示す。

番号	単語	相対節点 インデックス記号	品詞
EC00002661	_	[3]-2	BLNK
EC00002661	a	[4]-7-5-3	ART
EC00002661	_	[4]-7-5-2	BLNK
EC00002661	spokesman	[4]-7-5-1-2	NOUN
EC00002661	_	[4]-7-5-1-1	BLNK
EC00002661	for	[4]-7-4	PREP

EDRコーパスは文中のスペースについても構造を与えている。_ はスペースを意味する。品詞の種類は記号も含めて

ADJ, ADV, ART, AUX, BE, BLNK, CONJ,
DEMO, INDEF, ITJ, NOUN, NUM, PREP,

PRON, PTCL, PUNC, SUF, SYM, VI, VT, WH,
UNIT, VERB
の23種類がEDRコーパスデータで定義されている。

4.2 分析の方法と結果

データマイニング手法を用いて分析を行うには、対象とするデータについて、データクリーニングを行っておくことは分析のために重要である。データ分析用ソフトウェア JMP を使ってデータクリーニングの手作業と分布及びクロス集計をおこなった。データクリーニングの方針として、相対節点インデックス記号や品詞の数が対象としているデータに対して極度に少ないものは省いた。データマイニングを行う対象データは以下の通りである。

think を含む文から取り出した、文番号、構造・品詞・単語の組み合わせのアイテム数

自動詞の場合：3606件(文例：182)

他動詞の場合：6607件(文例：300)

分析はインテリジェント・マイナーの相関関係探索手法を用いた。分析に用いたパラメータ値と導出されたルール総数、最小確信度を以下に示す。

	自動詞	他動詞
・最小サポート値	9.5%	21.5%
・最小確信度	100%	100%
・ルール総数	10,693	9,331
・品詞が関係するルール	51	65
・think 周辺の特徴的 構造パターンの数	9	5

最小サポート値は、導出されたルールのトランザクションに対する割合をしめす。アイテム X,Y の関係 $X \rightarrow Y$ というルールがある場合、ルールの確信度は次のように計算できる。

$$\text{確信度}(X \rightarrow Y) = \frac{\text{サポート係数}(XUY)}{\text{サポート係数}(X)}$$

品詞と構造が関係するルールは、自動詞 51 件、他動詞 65 件であるが、そのルールから構造上同じ内容のものをまとめると、自動詞 think の周辺構造の特徴的なパターン 9 件、他動詞 think の周辺構造の特徴的なパターン 5 件となった。それぞれのパターンを図 3、図 4 に示す。また、各ルールのサポート係数、文例数と、文例、出典を表 1、表 2 に示す。

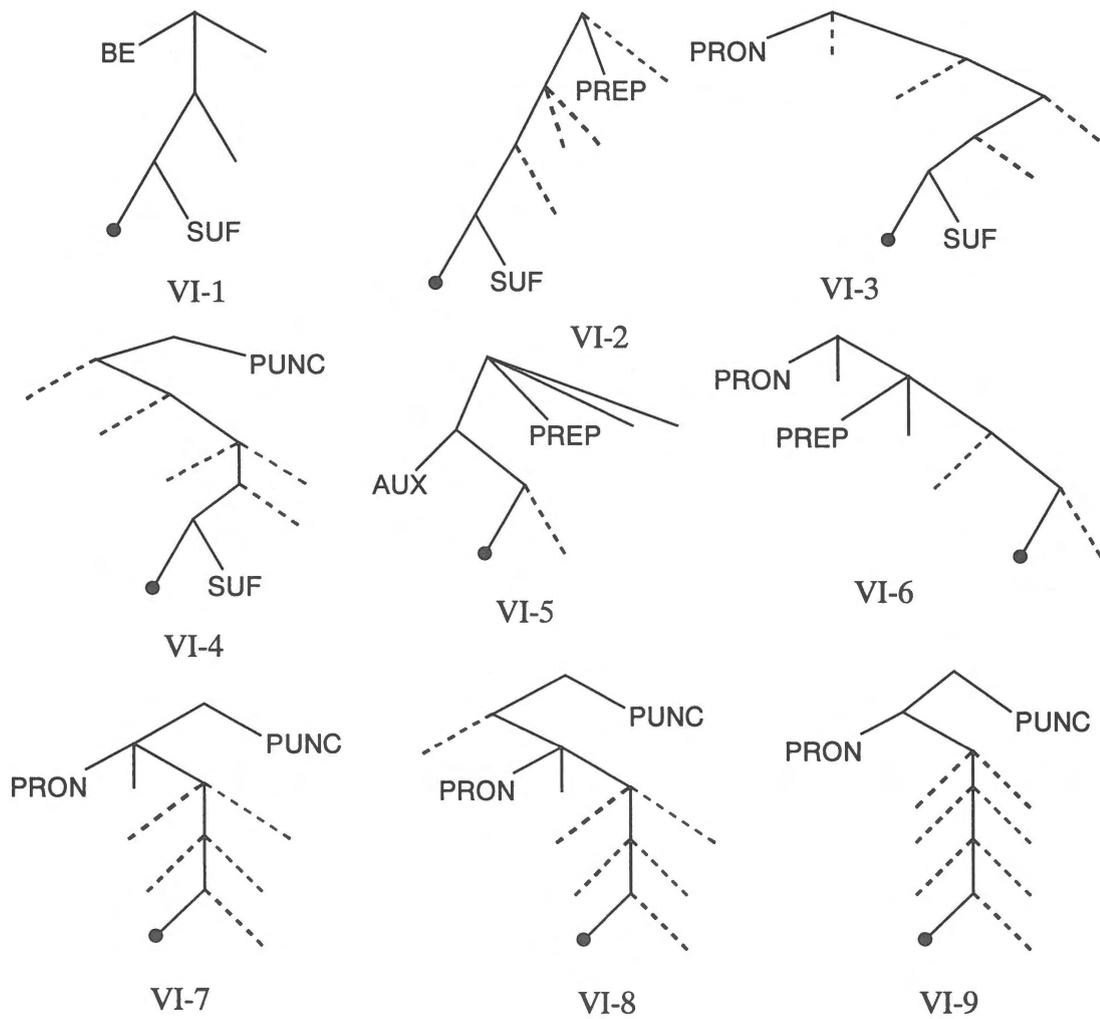


図3. 自動詞 think の周辺構造パターン (黒丸は think の位置)

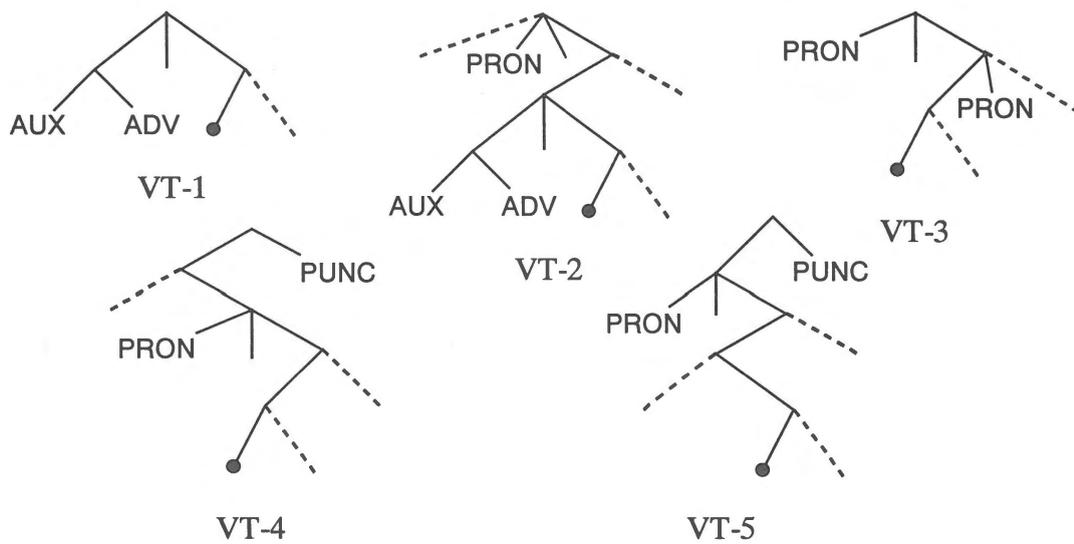


図4. 他動詞 think の周辺構造パターン (黒丸は think の位置)

表 1. 自動詞の場合のルールとサポート係数(S),文例 (確信度=100%) [出典]

VI-1 (SUF-BE) (S=9.890%, 文例数=18) [The Japan Economic Journal] Clearly, Nakasone was thinking of a similar role to Saionji's, observers say.
VI-2 (SUF-PREP)(S=9.890%, 文例数=18) [The Japan Times] But the postal part has held firm, and this is what everybody thinks of as the ministry's real work.
VI-3 (SUF-PRON)(S=9.890%, 文例数=18) [旺文社表現辞典] Are you also thinking about a link with the CAM system?
VI-4 (SUF-PUNC) (S=9.890%, 文例数=18) [旺文社表現辞典] Do you always relate to your customers thinking only about profits?
VI-5 (AUX-PREP)(S=10.989%, 文例数=20) [The Japan Times] But before we can think too much about spring, we have Christmas to enjoy.
VI-6 (PERP-PRON)(S=10.440%, 文例数=19) [The Independent] But how long was it before I was doing so without even thinking?
VI-7 (PRON-PUNC)(S=12.637%, 文例数=23) [旺文社表現辞典] Can't you think of a more efficient way to get it done?
VI-8 (PRON-PUNC)(S=9.890%, 文例数=18) [The Japan Times] Although such an argument seems plausible on the surface, I don't think it gets to the point.
VI-9 (PRON-PUNC)(S=10.440%, 文例数=19) [旺文社表現辞典] He obviously thinks he's a great comedian.

表 2. 他動詞の場合のルールとサポート係数文例 (確信度=100%) (出典)

VT-1 (AUX-ADV)(S=23.667%, 文例数=71) [The Japan Economic Journal] And in the end, I don't think we should have any doubt.
VT-2 (AUX-ADV-PRON)(S=22.000%, 文例数=66) [The Japan Economic Journal] But I don't think funds will move so widely.
VT-3 (PRON-PRON)(S=22.000%, 文例数=66) [The Independent] All those chaps have African forebears, I think you'll find.
VT-4 (PRON-PUNC)(S=24.667%, 文例数=74) [The Japan Times] A spokesman for the city said it cannot disclose at this time what price it thinks is acceptable.
VT-5 (PRON-PUNC)(S=30.000%, 文例数=90) [The Japan Times] Besides, I think, Americans like the novel because there is a lot of action in the story.

4.3 カスケードモデルによる分析

カスケードモデルのプログラム DISCAS を使って分析した結果を以下に示す。分析に用いたデータはインテリジェント・マイナーと同じである。パラメータは以下の通りである。

Min-support:	5%
Threshold:	0.04
自動詞のルール	16件
他動詞のルール	14件

カスケードモデルのルールの特徴は、{ [4]-3PRON: no} のように [4]-3の位置にPRONが存在しないことを示す構造パターンを導出する。ルール例を以下に示す。

相対インデックスの右側からの表示と左側からの表示を全て含むデータを用いた結果である。

() 内はルールがあてはまる文例数。

- R1: IF [[3].2 PREP: y] THEN Verb = VI (50)
R2: IF [[2]-3-1ADV: y] on [[3].1 PRON: n]
THEN Verb = VT (76)
R3: IF [[2].2 PREP: y] THEN Verb=VI (47)
R4: IF [[2].2 PRON: y] on [[4]-3PRON:n]
THEN Verb = VT (73)
R5: IF [[4].2 PREP: y] THEN Verb = VI (39)
R6: IF [[3]-3PRON: y] on [[4].1 PRON: n]
THEN Verb = VT (143)
R7: IF [[3]-3PRON: n] on [[4]-3PRON: n]
THEN Verb = VI (171)
R8: IF [[3]-3PREP: y] THEN Verb = VI (56)
R9: IF [[3]-3PRON: y] THEN Verb = VT (175)
R10: IF [[2].1.1AUX: y] on [[3]-3PRON: n]
THEN Verb = VT (63)
R11: IF [[2].1.2ADV: y] on [[3]-3PRON: n]
THEN Verb = VT (57)
R12: IF [[2].2PRON: y] THEN Verb = VT (74)
R13: IF [[3]-3PRN: n] on [[4].1PRON: n]
THEN Verb = VI (172)

5. まとめ

構文木データに、相対節点インデックス化を使って構造を含むデータの分析を行った。データマイニングの基本であるデータクリーニングをJMPのソート機能やエディター機能、グラフ機能を有効に使って比較的簡単に行うことができた。言語データの分析には、

人間とコンピュータとのインタラクティブな作業も大切であることを実感した。今後の研究としては、think以外の動詞について分析をおこなう予定である。特に、あらゆる表現を難しい単語を使わなくても表現できるといわれている基礎動詞 (be, do, have, see, take, get, put, keep, say, go, come, make, let) の13種について自動詞と他動詞の関係を調査する。さらにこれまで、行ってきた日本語のコーパスについても同様な方法で研究をおこなう。また、視点とした枝位置を基準とした番号付けによるインデックス化も試みる予定である。

参考文献

- [1] 雄山、岡田、李：“構文解析木を対象とするデータ解析法の研究(1)－方法論についての考察－”，重点領域研究、シンポジウム「人文科学における数量的分析」 pp.71-78 (1996)
- [2] 雄山、岡田、李：“構文解析木を対象とするデータ解析法の研究－構文木からの知識発見システム SYKD の開発と応用－”，情報処理学会研究報告 98-CH-38, pp.69-77 (1998)
- [3] EDR:EDR 電子化辞書 1.5 判仕様説明書, 日本電子化辞書研究所(1996)
- [4] Quinlan J.R.: “C4.5: Programs for Machine Learning”, Morgan Kaufmann, (1993); 古川訳: “AIによるデータ解析”, トップラン (1995).
- [5] 国藤編: “小特集 帰納論理プログラミング”, 人工知能学会誌, Vol.12, No.5 (1997)
- [6] 吉田、元田: “逐次ペア拡張に基づく帰納推論”, 人工知能学会誌, Vol.12, pp.58-67 (1997).
- [7] Agrawal, A. et. al.: “Database Mining: A Performance Perspective”, IEEE Trans. on Knowledge and Data Engineering, Vol. 5, No.6, pp.914-925 (1993).
- [8] Agrawal, A. et. al.: “Fast Algorithms for Mining Association Rules”, Proc. VLDB, pp.487-499 (1994)
- [9] Li, G. et.al.: “Knowledge Discovery from Syntactic Trees”, 1996年度人工知能学会全国大会 (第10回), 08-01, 東京 (1996).
- [10] Okada, T.: “Rule Induction in Cascade Model Based on Sum of Squares Decomposition” Principles of Data Mining and Knowledge Discovery, pp.468-475, Lecture Notes in Artificial Intelligence 1704, Springer (1999).

仮想電子辞書システムの設計と構築

Design and Construction of Virtual Electronic Dictionaries

芳野 学†, 都司 達夫‡, 宝珍 輝尚‡

Yoshino MANABU†, Tsuji TATSUO‡, Houchin TERUHISA‡

†福井大学 工学研究科 ‡福井大学工学部

Fukui University, 3-9-1 Bunkyo, Fukui-Shi 910, Japan

キーワード: 電子辞書, データベース, ビュー

Keywords: electronic dictionary, database, view

あらまし: 母体とする複数の辞書をその論理構造にしたがって構成要素に分解し, 項目結合可能性の制限の下にユーザが必要な項目のみを自由に組み合わせて新たに定義できる「仮想電子辞書」の提案とその設計・実装について述べる. これらの仮想電子辞書の基本概念の説明とともに, 定義構文, 検索構文を提案し, その実装方式を説明する.

Summary: In this paper, we describe design and implementation of virtual electronic dictionary system. In our system, one or more dictionaries are decomposed into their components according to their logical structures and some of them are corrected and combined logically into a virtual dictionary under the connectivity condition. Along with the description of the fundamental ideas underlying our virtual dictionaries, defining and retrieving primitives of virtual dictionaries are introduced and their implementation is explained.

1 はじめに

記憶メディアの大容量化と低廉化およびマルチメディア技術の進展にともなう, 各種電子辞書の出版は盛んに行われている. 特に, 画像や写真, 音声を取り込んだ「マルチメディア辞書」の発展は著しい. 本来, 辞書にはきわめて豊富で多角的な情報が内包されており [1], これらの情報は「辞書の文法」にしたがって, 構造的に

記述されている. 手で引く場合にはこの構造の強制に縛られており, 辞書が潜在的に保有している豊富で多角的な情報を手軽に利用することは不可能である. 電子辞書の存在意義はこのような構造の強制を解放し, ユーザの望む情報をユーザの望む配列で提示することができることであると考えられる. これには, 望む情報を特定するために, 辞書の文法にしたがった構造化を行う必要がある. 電子辞書を含めた電子化出版物の構造化には, たとえば EPWING[8]などの規格が存在しており, それにしたがって, 電子辞書も出回り始めている. また, [2]では辞書のハイパーテキスト化の方式が提案されている. ところが, 辞書の論理構造 (スキーマ) を前面に押し出し, ユーザレベルで既存辞書からニーズにかなった, 仮想的な辞書を定義し利用するための枠組みはあまり, 見当たらない.

そこで, 本報告では, 上記のような目的に沿った「仮想電子辞書」の設計と構築について述べる. 仮想電子辞書として, 一種類の辞書内で有用な情報のみを再配列して得られる「部分辞書」と複数の辞書内の有用な情報を「結合対」を介して結合して得られる「結合辞書」を提案する. さらに, これらの仮想辞書に関連する諸概念を提案するとともに, 仮想電子辞書の定義・操作言語を提案し, その実装方式を説明する.

2 仮想辞書の設計

市販電子辞書における相互検索機能を使えば, 複数辞書にまたがった場合の検索操作性は向上するものの, 個々の辞書を意識して, 切り替えるといった操作が必要であり, 煩わしい. ここでは, 個々の辞書の実体とは別個にユーザが定

義する仮想辞書を設計する。仮想辞書は説明項目の実体を持たずに、辞書実体のどの部分を使用するかで定義情報のみ有する。ここでの仮想辞書の考え方は、データベースシステムにおけるビュー機能に相当している。各実体辞書の文法構造を表現する辞書スキーマ群に対するビュースキーマとして仮想辞書のスキーマをユーザが定義できる。仮想辞書に対する検索要求は辞書実体に対する3で述べる検索ライブラリ中の検索関数呼び出しにより、実行される。以降では特に「仮想辞書」と断らない限り、単に辞書といった場合は実体辞書を指すこととする。なお、本研究で用いた辞書の元データは三省堂(株)から提供を受けた、新クラウン和英辞典[5]、ニューセンチュリー英和辞典[6]であり、Websterの英語シソーラス[7]を参考にした。

2.1 結合対

「仮想辞書の自動生成」を行う場合には、項目同士の結合操作がポイントとなる。例えば、英和辞書から、「反意語仮想辞書」を作成する場合、英和辞書の見出し説明中に現れる「反意語」が結合対となり、(同じ)英和辞書の見出しと結合してその反意語の詳しい情報を与える仮想辞書が定義される。なお、辞書の見出しは必ず、結合対となり得る。ここでは、種々の辞書の辞書構造と各項目のデータの分類を掲げて、仮想辞書生成時の結合対について考察する。

(1) 和英辞書

和英辞書の論理構造を表1に示す。[]付きの項目は辞書の部分構造を決定するための項目であり抽象項目と呼ぶこととする。また、[]なしの項目は辞書に記載されているデータであり、具体項目と呼ぶ。

それぞれの具体項目のデータは、結合可能性の観点から表2のように分類される。「見出し」以外の項目には記号(1byte)も含まれている。ここでは、「英訳の例文」「用例の例文」「関連語の例文」をまとめて「例文」としている。また、用例や例文は(和単語列、英単語列)の対として与えられるが、ここでは例えば、用例の場合、和単語列成分は「用例和」、英単語列成分は「用例英」として別個の項目としている。

なお、以降の結合可能性分類表における右端の欄の“-”, “○”, “△”, “×”は、“結合は無意

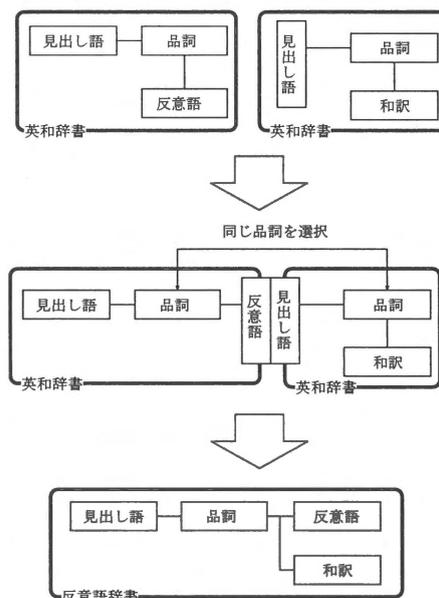


図1: 結合対と結合条件

味”, “結合可能”, “データに処理を施すことで結合可能”, “結合は困難”をそれぞれ表している。見出しは例えば国語辞書など、仮名和単語や片仮名和単語を見出しとする辞書とそのまま結合可能である。“△”のついたものは、適当な構文解析や整形によって、有意な単語に分解することにより、他の結合対と結合可能である。「漢字訳」は、見出しとほとんど1対1の関係にある。しかし、たとえば「いたみ」のように「痛み」と「傷み」のような同音異義語があるが、和単語を結合対とする他の辞書と結合可能である。「英訳」は多くの場合単一英単語であり、英単語を見出しとする他の辞書と結合可能である。英単語列の場合には、英語辞書の成句と結合したり、名詞+名詞や形容詞+名詞のような複合語の場合には最後の名詞が有意であり、結合対となり得る。「関連語和」の場合には、見出しを含む複合語である場合が多く、たとえば、国語辞書の参考語と結合可能であり、より詳しく、その日本語の意味を知ることができる。また、用例や例文はいずれも、単語列として他の説明項目と全く一致することはほとんどない。ただし、構文解析によりその中の名詞や動詞について他と結合することは可能である。

対を介して合わさっただけであり、実用にはならない。仮想辞書を実用的に使用するためには、ユーザ自身が仮想辞書を自分のニーズにしたがって、きめ細かく定義できる必要がある。このために、ここでは、本辞書システムにおける、仮想辞書定義の定義要素を以下に述べる。以降の説明においては例として、英和辞書と英語のみの英語シソーラス辞書を使って、和訳付き英語シソーラスの仮想辞書を作成する場合を取り上げる。

(1) 辞書構造の定義

複数の辞書を結合しても、それらの辞書の対応する項目がすべて必要であるわけではない。各辞書の項目を選択的に取り込み、それを可能な配置により再編成することにより、仮想辞書の構造を定義できる。和訳付き英語シソーラス辞書の場合、例えば、

和訳付き英語シソーラス辞書(見出し、
見出し項目(品詞、品詞項目(
類義語、類義語項目(語義))、
反意語、反意語項目(語義))))

のようにその構造を入れ子で指定する。

(2) 母体となる辞書群の指定

結合の対象となる実体辞書または仮想辞書を1つ以上指定する。1つの場合には、結果として得られる辞書を仮想部分辞書といい、2つ以上の時は仮想結合辞書という。和訳付き英語シソーラスの仮想辞書の場合には、母体辞書群は英和辞書、英語シソーラス辞書である。

(3) 母体辞書との対応付けの定義

(1)で定義した仮想辞書の実体項目が実体辞書のどの項目に関連づけられるかを定義する。実体辞書の論理構造(スキーマ)にしたがって、見出しから始まって辞書項目に至るパスを実体辞書の項目として指定する。たとえば、図2の英和辞書スキーマにおいて、“見出し項目.品詞項目.成句項目.成句”とすれば、以下の(5)の選択条件を満足するすべての成句を指示することになる。また、“見出し項目.品詞項目.成句項目”とすれば、選択条件を満足するすべての成句について、その「重要度」、「成句(の綴り)」、「意味」、「用例」を指示する。和訳付き英語シソーラス辞書の場合、例えば、

見出し = ET. 見出し、

見出し項目.品詞 = ET. 見出し項目.品詞、
見出し項目.品詞項目.類義語 =

ET. 見出し項目.品詞項目.語義項目.類義語、
見出し項目.品詞項目.類義語項目.語義 =

EJ. 見出し項目.品詞項目.語義、

見出し項目.品詞項目.反意語 =

ET. 見出し項目.品詞項目.語義項目.反意語、
見出し項目.品詞項目.反意語項目.語義 =

EJ. 見出し項目.品詞項目.語義

となる。ここで、ETは英語シソーラス辞書、EJは英和辞書を表す。

(4) 結合対の指定

2つの母体辞書を D_1 、 D_2 とする。 D_2 の項目 $e_1...e_n$ が結合可能であり、 D_1 の見出しと結合する時、“ D_2 は項目 $e_1...e_n$ を介して D_1 と結合可能である”といい、 $\langle D_1, D_2.e_1...e_n \rangle$ と表記してこれを結合対という。和訳付き英語シソーラス辞書の場合には、結合対は

$\langle EJ, ET. 見出し項目.品詞項目.語義項目.類義語 \rangle$

$\langle EJ, ET. 見出し項目.品詞項目.語義項目.反意語 \rangle$

となる。結合対による2つの辞書の結合は関係データベースにおける結合演算に相当する。

(5) 選択条件の指定

和訳付き英語シソーラスの仮想辞書の場合、基本的には、英和辞書も英語シソーラス辞書も品詞ごとに分けられているが、互いに異なる品詞を結合しても、有意な情報とはいえないし、組み合わせの数も膨大になる。このような場合には、「選択条件」として、“ET. 見出し項目.品詞項目.品詞 = EJ. 見出し項目.品詞項目.品詞”というようにして結合の条件を指定する。他に、動詞ばかりの仮想辞書を作成する場合には、“ET. 見出し項目.品詞項目.品詞=動詞 and EJ. 見出し項目.品詞項目.品詞=動詞”の指定になる。

(6) 見出し指定と見出しフィルタ

見出しの指定を行う。見出しとして複数の辞書項目を指定してもよい。見出しには既定として、高速アクセスのための索引が付与される。通常、結合条件の選択演算により、仮想辞書のレコード集合は選択的に定義されるが、見出し集合に働くフィルタによっても選択することができる。見出しフィルタとして、正規表現が利用できる。また、逆引き辞書[4]のように、見出しの読みを逆順にして、並べ替えることもできる。

2.3 仮想辞書の定義構文

仮想辞書には、仮想部分辞書と仮想結合辞書がある。いずれの場合にも「辞書構造の定義」、「母体辞書の指定」と「母体辞書との対応付けの定義」は必須である。

(1) 仮想部分辞書

仮想部分辞書では辞書の結合は行われず、単一辞書について、ユーザの必要とする項目のみを再編成する。数ある辞書項目をすべて検索対象とはせずに、使用目的に沿って選んだ辞書項目のみをよりコンパクトにまとめた辞書を単一辞書から抽出することが目的である。たとえば、英和辞書の中から、成句のみを抽出して成句辞書を作成したい場合には、

```
DEFINE 成句辞書 (見出し, 見出し項目
                (品詞, 成句))
FROM    EJ
ASSIGN
    見出し = EJ. 見出し,
    品詞 = EJ. 見出し項目. 品詞,
    成句 = EJ. 見出し項目. 品詞項目.
            語義項目. 成句項目. 成句
WITH ENTRIES 見出し
```

となる。DEFINE 句では、仮想辞書名の下に辞書項目とその構造が定義される。実体辞書では「品詞項目」の下に3つの項目があり、「語義項目」には5つの項目があったが、成句のみを選ぶことにより、これら2つの抽象項目はなくなり、構造は浅くなっている。FROM 句は母体辞書の指定であり、AS以降は母体辞書による定義内容である。ASSIGN 句は仮想辞書の各辞書項目が実体辞書のどの項目に関連づけられるかを定義する。WITH ENTRIES 句は仮想辞書の見出しを含む1以上の複数の辞書項目を指定する。物理的にはこれらを複合キーとして、索引が生成される。

見出しに対する索引は、実体辞書の索引をそのまま使用できないこともないが、実体辞書のすべての見出しが成句項目を持っているわけではない。したがって、仮想辞書の見出しに対してフィルター検索をかけた場合、無駄な探索を行ってしまうので、仮想辞書用に新たに別途作成する。

(2) グループ핑

次に、英作文を支援する仮想辞書の構成を考える。英文作成において、実際にもっとも役に立つ情報は用例であろう [3]。例えば次のように、和英辞書から例文、用例、関連語を集めて、用例対訳辞書 A を定義する。

```
DEFINE 用例対訳辞書 A (見出し,
                       用例項目 (用例和, 用例英))
FROM    JE
ASSIGN
    見出し = JE. 見出し,
    (用例和, 用例英) =
    (<JE. 見出し項目. 英訳項目. 対訳項目. 例文和,
     JE. 見出し項目. 英訳項目. 対訳項目. 用例和,
     JE. 見出し項目. 英訳項目. 対訳項目. 例文和,
     JE. 見出し項目. 英訳項目. 関連語. 関連語和,
     JE. 見出し項目. 英訳項目. 関連語. 例文和>,
     <JE. 見出し項目. 英訳項目. 対訳項目. 例文英,
     ..... (用例和 と同様) >)
WITH ENTRIES 見出し
```

ASSIGN 句の「用例和」は和英辞書の5つの和文の辞書項目を1つにまとめた新たな単一項目である。「用例英」も同様の新たな単一項目である。このような操作を辞書定義レベルのグルーピング (GROUPING) という。このときの問題は、「用例和」と「用例英」の対応のさせ方である。一般に2つのグルーピング A (a_1, a_2, \dots, a_n) および B (b_1, b_2, \dots, b_n) があり、仮想辞書の DEFINE 句において (A, B) と並置され、母体辞書において (a_i, b_i) ($i = 1, 2, \dots, n$) と並んでいるとき、(A, B) の項目間は (a_1, b_1) (a_2, b_2) ... (a_n, b_n) と対応づけられる。2つのグルーピングの項目間がこのように対応づけられることを ASSING 句において、() で囲むことによって表す。

ところで、用例は英和辞書にも記載されており、その仮想辞書、“用例対訳辞書 B (見出し, 用例項目 (用例英, 用例和))” が用例対訳辞書 A と同様にして定義できる。

(3) 仮想結合辞書

和訳付き英語シソーラス辞書の定義をまとめると、

```
DEFINE 和訳付き英語シソーラス辞書 (見出し,
                                       見出し項目 (品詞, 品詞項目 (
                                       類義語, 類義語項目 (語義)) ),
```

(反意語, 反意語項目 (語義))))))

FROM ET, EJ

ASSIGN

見出し = ET. 見出し,

見出し項目. 品詞 = ET. 見出し項目. 品詞,

見出し項目. 品詞項目. 類義語 =

ET. 見出し項目. 品詞項目. 語義項目. 類義語,

見出し項目. 品詞項目. 類義語項目. 語義 =

EJ. 見出し項目. 品詞項目. 語義,

見出し項目. 品詞項目. 反意語 =

ET. 見出し項目. 品詞項目. 語義項目. 反意語,

見出し項目. 品詞項目. 反意語項目. 語義 =

EJ. 見出し項目. 品詞項目. 語義

WITH ENTRIES 見出し, 見出し項目. 品詞

CONNECTED BY

<EJ, ET. 見出し項目. 品詞項目. 語義項目. 類義語>

<EJ, ET. 見出し項目. 品詞項目. 語義項目. 反意語>

WHERE

EJ. 見出し項目. 品詞 = ET. 見出し項目. 品詞

となる. 結合対の指定には, CONNECTED BY 句を導入した. また, WITH ENTRIES 句により, ET の見出しと品詞の対が指定されており, この対をキーとして索引が生成される. 仮想結合辞書は, 複数の辞書を結合して得られるが, 必ず, FROM 句において複数の辞書と, CONNECTED BY 句において結合対を指定する必要がある.

(4) 見出しなしの仮想辞書

既述の仮想辞書には必ず見出しが定義されていた. ユーザは見出しを指定することにより, 関連する辞書項目を直接検索することができる. ところが, 場合によっては, 見出しにはよらず, 定義されたデータを”縦断的”に検索したい場合がある. たとえば, 2.3 で述べた用例対訳辞書は便利であるが, 調べたい言葉の見出しに関するものだけしか知ることができない. たとえば, ある言葉の用例の和英対訳を辞書全体にわたって検索できれば, 手引きでは調べ尽くせない有用な例文を数多く得ることができる. この仮想辞書は既定義の仮想辞書である用例対訳辞書 A を母体辞書として, 次のように定義できる.

DEFINE 用例対訳辞書 and 英 (用例和, 用例英)

FROM 用例対訳辞書 A

ASSIGN

用例和 = 用例対訳辞書 A. 用例項目. 用例和,

用例英 = 用例対訳辞書 A. 用例項目. 用例英

このとき, 用例対訳辞書 A の定義における用例和と用例英のグルーピングとそれらの対応関係は用例対訳辞書 and 英においても有効であることに注意されたい.

見出しなしの仮想辞書はいわゆる全文検索機能に対応するが, 検索の対象を仮想辞書で定義された項目のみに限定することが可能であり, 効率的な検索が行える.

(5) 辞書の縦列結合

見出しなしの複数の仮想辞書を縦続接続することができる. 縦列結合する仮想辞書は構造等価でなければならない. 2つの辞書が構造等価であるとは, DEFINE 句で指定される (木) 構造が同一であること, および, 具体項目のデータ型が一致していることである.

(4) の例では, 和英辞書を母体とする用例対訳辞書 A のみが検索の対象となっているが, 英和辞書を母体とする (2) の用例対訳辞書 B の用例も併せて検索したい. 辞書を縦列結合してあらたな仮想辞書を作り出す構文は以下の通りである.

DEFINE 用例対訳辞書 and 英 (用例和, 用例英)

FROM 用例対訳辞書 A, 用例対訳辞書 B

ASSIGN 用例和 =

[用例対訳辞書 A. 用例項目. 用例和 |

用例対訳辞書 B. 用例項目. 用例和]

用例英 =

[用例対訳辞書 A. 用例項目. 用例英 |

用例対訳辞書 B. 用例項目. 用例英]

[… | … | …] は縦列結合を表す. 縦列結合は結合対による結合によらず, したがって, 結合対の指定は不要であること, やはり, 母体辞書におけるグルーピングとそれらの対応関係が保存されることに注意されたい.

(6) 実体辞書定義文

仮想辞書と同様に実体辞書についても実体辞書定義文を同様に定める必要がある. これは, 定義構造を仮想辞書のものと同じにし, 操作上区別なしに扱えるようにするためである. これにより, たとえば, 母体辞書定義として, 仮想辞書と実体辞書を混在させることができる. 実体

辞書定義文は実体辞書の構造を定義するとともに、それぞれの項目に対応する辞書の実データを検索するライブラリ関数のアドレスを関連づけする。項目の検索にあたってはこの関数を呼び出せばよい。

```
DEFINE 英語シソーラス辞書 (見出し語,  
    見出し項目 (品詞, 品詞項目 (語義,  
        語義項目 (用例, 類義語, 関連語,  
            成句, 反対語, 反意語))))
```

ASSIGN

```
見出し語 = thesaurus_key(),  
見出し項目.品詞 = thesaurus_part(),  
見出し項目.品詞項目.語義  
    = thesaurus_sense(),  
見出し項目.品詞項目.語義項目.用例  
    = thesaurus_example(),  
見出し項目.品詞項目.語義項目.類義語  
    = thesaurus_synonym(),  
見出し項目.品詞項目.語義項目.関連語  
    = thesaurus_relate(),  
見出し項目.品詞項目.語義項目.成句  
    = thesaurus_idiom(),  
見出し項目.品詞項目.語義項目.反対語  
    = thesaurus_contrast(),  
見出し項目.品詞項目.語義項目.反意語  
    = thesaurus_antonym()
```

2.4 仮想辞書の検索

仮想辞書の検索は FIND 文により行う。仮想辞書の DEFINE 文の構造定義中の項目に具体値を代入して、条件付けする。具体値の指定には、正規表現が使用できるほか、論理和 | も使用できる。例えば、和訳付き英語シソーラス辞書の場合、

```
FIND 和訳付き英語シソーラス辞書 ("*take",  
    見出し項目 ("動詞", 品詞項目 (  
        (類義語, 類義語項目 (語義), #)))
```

とすれば、"take"で終わる動詞の英単語の(類義語, 類義語項目 (語義))がリストアップされる。#はその項目以下の項目がすべてマスクされ、検索条件にも検索対象にもならないことを示す。以下はこの検索文による検索結果である。

見出し語: test

品詞: 動詞

類義語 (1): check

和訳 (1): 1. 阻止する, 防ぐ, 抑制する
2. テストする, 検査する, 調べる
3. 預ける 4. 王手をかける

和訳 (2): 1. 一致する, 符合する
2. 検査する
3. 小切手を振り出す
4. [キングに] 王手をかける

類義語 (2): example

和訳 (1): 1. 調べる, 検査する, 審査する
2. 試験する
3. 診察する, 検診する
4. 尋問する

類義語 (3): prove

和訳 (1): 1. 説明する, 立証する
2. 試す, 試験する

和訳 (2): 1. 判明する, わかる

類義語 (4): demonstration

和訳 (1): 1. 実証する, 明らかにする
2. 宣伝する, 説明する
3. [感情などを] 顔に出す

和訳 (2): 1. デモする

類義後 (5): try

和訳 (1): 1. 試す, 使ってみる
2. 試みる, やってみる
3. しようと努める, 努力する
4. 試練を与える, 無理させる
5. 審判を行う, 裁判する, 審理する
和訳 (2): 1. 試みる, やってみる, 努力する

また、

```
FIND 和訳付き英語シソーラス辞書 ("*take",  
    見出し項目 ("動詞", 品詞項目))
```

の「品詞項目」のように、抽象項目のままにしておくと下位の項目がすべて検索対象になる。

見出しなしの仮想辞書の検索も同様に FIND 文により行うことができる。例えば、

```
FIND 用例対訳辞書和英 ("*考え*", 用例英)
```

は用例和に“考え”を含むような用例の和英対訳表を作る。用例対訳辞書と和英は2つの仮想辞

書を母体としており，これらは”対応づけられたグルーピング”を含む。したがって，グルーピングされた各項目ごとに対応関係が存在するので，和文に対して英文が一意に定まる。

3 仮想辞書の実体化

仮想辞書は実体辞書に対するビューであり，その検索は次の理由から一般的にオーバーヘッドが大きい。

(a) 実体辞書にあった抽象項目が仮想辞書定義において，消失した場合でも，実際の検索は実体辞書を用いるために，抽象項目を介して行う必要がある。

(b) 仮想結合辞書の結合処理コストが非常に高い。検索時に毎回結合処理を行うことは実用的でない。

仮想辞書の定義と実体辞書の辞書項目検索ライブラリルーチンを用いて，実体辞書と同様にその物理データ構造を二次記憶中に作り出す操作を「仮想辞書の実体化」という。これにより，上記のオーバーヘッドは回避できる。

なお，2.3(1)で述べたように，仮想辞書の見出しに対する索引だけは，仮想辞書定義時にすでに作成されており，実体化しなくても使用できる。また，(b)に対しては，結合によって連結される相手辞書のレコードの参照情報が結合索引として，仮想結合辞書定義時に作成される。これらの対策により，項目データも含めて完全に実体化しなくても，見出し指定検索の場合には，実用に耐えうるパフォーマンスは得られる。ただし，見出しなしの仮想辞書の場合には，定義項目の全検索が必要であり，それ用に最適化した辞書データの内部データ構造を自動生成することが必要となる。実体化は MATERIALIZE 文で行う。

MATERIALIZE 用例対訳辞書中英

4 仮想辞書システム

4.1 モジュール構成

仮想辞書システムは(1)ユーザからの仮想辞書定義や，仮想辞書検索の指示を受け，それぞれを DEFINE 文，FIND 文に変換する GUI，(2)DEFINE 文，FIND 文を構文解析し，対応する内部データ構造を1次記憶中に作成するパーザ，(3)

内部データ構造を参照しながら仮想辞書操作を行うインタプリタ，(4)母体辞書 DB，母体辞書定義文 DB 仮想辞書定義文 DB，(5)GUIからの要求にしたがって実際に母体辞書に指定された実体辞書や仮想辞書の検索を行うライブラリ，(6)仮想辞書実体化サブシステム，などからなる。

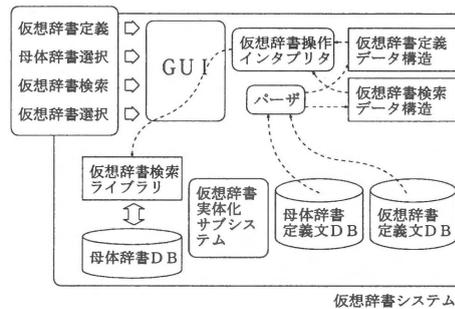


図 2: 仮想辞書システム

(1) 仮想辞書の定義

仮想辞書の定義を行なう場合，ユーザは仮想辞書定義文を作成する。仮想辞書定義文を受け取った仮想辞書定義文パーザはこれを解析し，1次記憶中に仮想辞書定義内部データ構造に変換する。

(2) 仮想辞書の検索

(1)と同様にユーザは仮想辞書検索文を作成する。定義文と同様に仮想辞書検索パーザにより，仮想辞書検索内部データ構造に変換され仮想辞書インタプリタに渡される。検索データを受け取ったインタプリタは主メモリ上で被検索辞書の辞書定義データを探して，必要な検索ルーチンを決定する。決定したらインタプリタは必要な検索ルーチンを母体辞書サーバに指示し，母体辞書サーバはその指示にしたがって母体辞書から検索を行なう。検索結果はインタプリタを通して，GUIなどから出力される。

4.2 GUI

前節で述べた言語ベースのインターフェースでは項目指定のためのパス表現が多くてかつ長たらしいので，仮想辞書はやはり使いにくい。仮想辞書の定義や操作には GUI ベースの分かりやすいインターフェースを提供することが必要で

ある。特に本システムでは辞書スキーマを前面に押し出す必要がある。すなわち、仮想辞書の定義にあたっては、母体辞書のスキーマをグラフィカルに提示し、定義に必要な項目をその中から切り出し、可能な範囲で望みどおり組み合わせる機能を提供することにより、新辞書のスキーマ構築を支援する必要がある。

本研究では、このような GUI を CGI を用いて実装することで、Web 上で仮想辞書の定義、検索を可能とした。

GUI は大きく分けて、仮想辞書定義部と仮想辞書検索部とに分けられる。

(1) 仮想辞書定義部 (付録 (1) 参照)

仮想辞書定義部では、最初に仮想辞書の辞書構造を作成する。

辞書構造の作成は、

1. 仮想辞書中のどこに項目を追加するかを選択。
2. 母体辞書を選択。
3. 母体辞書中の辞書項目を選択。
4. 追加された項目の名前を変更する。

という手順で行なわれる。辞書構造を決定すれば自動的に母体辞書との対応付けも行なわれる。辞書構造が決定されたら次は仮想辞書の見出し、結合対、結合条件を指定する。完成した仮想辞書定義文は仮想辞書定義文 DB に登録され、検索時に利用される。また、他の仮想辞書の母体辞書として利用することができる。

(2) 仮想辞書検索部 (付録 (2) 参照)

仮想辞書検索部では、最初に検索を行なう仮想辞書あるいは実体辞書を選択する。次に選択した辞書の辞書構造を参照しながら検索対象と検索条件を指定する。検索条件は任意の項目を選択するだけで自動的に作成される。検索は作成された仮想辞書検索文に従って行なわれるが、一度検索した検索文中の検索条件のパラメータを変更しながら再検索を行なうこともできる。

5 おわりに

仮想電子辞書の設計について述べた。本設計に基づき現在実装中である。既存の実体辞書の問題点として、記述が不統一でスキーマに沿った構造化ができない項目説明があったり、項目として切り出す時の終端子として決めたマークが

必ずしも一定でなかったり、スキーマとしてこちらが想定していない項目が出現したりする例外が出現し、構造化できないものがあつた。可能な限り手作業で対応したが、なお十分ではない。これらのいくつかは、言葉が持つ多様性を十分に説明するためには、ある程度不可避であると思われる。電子辞書の規格化 (たとえば, [8]) がこのあたりも含めて十分に検討され、標準化されたならば、逆に辞書の編集段階から、規格準拠を考慮でき、統一のとれた電子辞書を構築できる。本研究における仮想電子辞書の場合には、このような標準化された実体辞書が母体辞書として使用できるような環境が望まれる。この点に関して、現在、我々は XML(eXtensible Markup Language) を用いて、仮想辞書の定義、検索、辞書情報の記述を行なうための取り扱いを検討中である。付録 (2)(3) に、仮想辞書定義のための DTD とそれに基づく ”和訳付き英語シソーラス辞書” の XML による定義を示す。

謝辞

和英辞書および英和辞書の元データは三省堂(株)より提供していただいた。謝意を表す。

文献

- [1] 吉田 将: 辞書構築における諸問題, 情報処理, 27, 8, pp.933-939, 1986.
- [2] 内藤, 山下, 松山, 柵木: オンライン辞書のハイパーテキスト化手法, 情報処理学会論文誌, 34, 2, pp.320-329, 1993.
- [3] 武田明子, 古郡延治: 例文をもとにした英文書作成支援システム, 情報処理学会論文誌, 35, 1, pp.53-61, 1994.
- [4] 岩波書店辞典編集部: 逆引き広辞苑, 岩波書店.
- [5] 山田和男 編: 新クラウン和英辞典, 三省堂.
- [6] 木原, 福村, 芦川 編: ニューセンチュリー英和辞典, 三省堂.
- [7] NeXT Computer Inc.: Chap.5 in " NeXT Applications", 1990.
- [8] EPWING コンソシアム: <http://www.epwing.or.jp>
- [9] <http://www.m-w.com/cgi-bin/netdict>

付 録

(1) GUIによる仮想辞書の定義



(2) GUIによる仮想辞書の検索



(3) 仮想辞書定義のための DTD

```
<?xml encoding="EUC-JP">

<!-- 仮想辞書の宣言 -->
<!ELEMENT Define ( Head, From, Body ) >

<!-- 仮想辞書のヘッダ情報 -->
<!ELEMENT Head ( DicName, Locate, Author, Update, Note ) >
<!ELEMENT DicName ( #PCDATA ) >          -- 辞書名 --

<!-- 仮想辞書の所在 -->
<!ELEMENT Locate ( #PCDATA ) >          -- 辞書の存在する URL --

<!-- 作成者 -->
<!ELEMENT Author ( #PCDATA ) >          -- 作成者名 --
<!ATTLIST Author
      Mail          #IMPLIED          -- 作成者のメールアドレス --
>

<!-- 最終更新日 -->
<!ELEMENT Update ( #PCDATA ) >          -- 更新日 --

<!-- 備考 -->
<!ELEMENT Note ( #PCDATA ) >           -- 備考 --

<!-- 母体辞書情報 -->
<!ELEMENT From ( ParentDic* ) >
<!ELEMENT ParentDic ( #PCDATA ) >      -- 母体辞書名 --
<!ATTLIST ParentDic
      Alias          #IMPLIED          -- 母体辞書の別名 --
      Locate         #REQUIRED         -- 母体辞書の存在する URL --
>

<!-- 仮想辞書の構造および項目情報 -->
<!ELEMENT Body ( Struct, Assign, Connect, Index ) >
<!ELEMENT Struct ( ( Concrete | Abstract | Assign | Connect )* ) >
<!ELEMENT Concrete >
<!ATTLIST Concrete
      Name          #REQUIRED          -- 具体項目の名前 --
>
<!ELEMENT Abstract ( ( Concrete | Abstract | Assign )* ) >
<!ATTLIST Abstract
      Name          #REQUIRED          -- 抽象項目の名前 --
>

<!-- 辞書項目の関連情報 -->
<!ELEMENT Assign ( ( Relate, Connect, Group ) | Entity* ) >
<!ATTLIST Assign
      Name          #REQUIRED          -- 関連情報を付加する辞書項目名 --
      Type ( NUM, PENUM, UC, PRON, WORD, AWORD, SENT, ASENT, ETC ) "WORD"
      -- データ型に関する定義 --
>

<!ELEMENT Relate ( #PCDATA ) >          -- 対応する母体辞書項目パス --
<!ELEMENT Index ( Key* ) >
<!ELEMENT Key ( #PCDATA ) >            -- 索引を付加する項目パス --
<!ELEMENT Connect ( Entry* ) >
<!ELEMENT Entry ( #PCDATA ) >          -- 結合辞書の見出し語に代入する項目 --
<!ATTLIST Entry
      Name          #REQUIRED          -- 結合を行なう母体辞書 --
>
```

```

<!ELEMENT Where ( Expr ) >
<!ELEMENT Expr ( #PCDATA ) >
<!ATTLIST Expr
      Name          #REQUIRED      -- 条件を付加する項目パス --
      Op            ( EQ, NEQ ) "EQ" -- 演算子 --
>
<!ELEMENT Entity ( #PCDATA ) >      -- 実体 --

```

(4)XML を用いた仮想辞書定義文

```

<Define>
  <Head>
    <DicName>和訳付き英語シソーラス辞書</DicName>
    <Locate>http://pear/~yoshino/kenkyu/jet.xml</Locate>
    <Author Mail="yoshino@pear">Manabu Yoshino</Author>
    <Update>1999 11/8</Update>
  </Head>
  <From>
    <ParentDic Alias="ET" Locate="http://pear/~yoshino/kenkyu/et.xml">英語シソーラス辞書
  </ParentDic>
    <ParentDic Alias="EJ" Locate="http://pear/~yoshino/kenkyu/ej.xml">英和辞書</ParentDic>
  </From>
  <Body>
    <Struct>
      <Concrete Name="見出し語"/>
      <Abstract Name="見出し語項目">
        <Concrete Name="品詞"/>
        <Abstract Name="品詞項目">
          <Concrete Name="類義語"/>
          <Abstract Name="類義語項目">
            <Concrete Name="類義語和訳"/>
          </Abstract>
          <Concrete Name="反意語"/>
          <Abstract Name="反意語項目">
            <Concrete Name="反意語和訳"/>
          </Abstract>
        </Abstract>
      </Abstract>
    </Struct>
    <Assign>
      <Relate Name="見出し語">ET. 見出し語</Relate>
      <Relate Name="品詞" Type="PENUM">ET. 見出し語項目. 品詞</Relate>
      <Relate Name="類義語">ET. 見出し語項目. 品詞項目. 語義項目. 類義語</Relate>
      <Relate Name="類義語和訳">EJ. 見出し語項目. 品詞項目. 語義</Relate>
      <Relate Name="反意語">ET. 見出し語項目. 品詞項目. 語義項目. 反意語</Relate>
      <Relate Name="反意語和訳">EJ. 見出し語項目. 品詞項目. 語義</Relate>
    </Assign>
    <Index>
      <Key>見出し語</Key>
      <Key>見出し語項目. 品詞</Key>
    </Index>
    <Connect>
      <Entry Name="EJ">ET. 見出し語項目. 品詞. 類義語</Entry>
      <Entry Name="EJ">ET. 見出し語項目. 品詞. 反意語</Entry>
    </Connect>
    <Where>
      <Expr Name="EJ. 見出し語項目. 品詞">ET. 見出し語項目. 品詞</Expr>
    </Where>
  </Body>
</Define>

```

公開シンポジウム「人文科学とデータベース」プログラム

日時:平成7年12月25日(月)、26日(火)
場所:大阪電気通信大学B310教室(AV教室)

25日(月曜日)午前10時から午後5時まで

開会挨拶:重点領域・領域代表者:
及川 昭文(総合研究大学院大学)
シンポジウム実行委員長:
小沢 一雅(大阪電気通信大学)

特別講演:古地震データと活断層
講演者:寒川 旭
通産省工業技術院地質調査所大阪地域地質センター
地域地質研究官(理学博士)

1. IntelligentPadシステムを用いた歴史学研究支援データベースの構築
赤石美奈、中谷広正、伊東幸宏、阿部圭一、田村貞雄
(静岡大学)
2. 4次元歴史空間システムにおける地理情報処理について
小林努、加藤常員、小沢一雅
(大阪電気通信大学)
3. 視点に依存する属性付け機構を持つ木簡研究支援システム
ー構造進化型データベースの概念ー
森下淳也(姫路獨協大学)、上島紳一(関西大学)、大月一弘(神戸大学)
4. 古典籍とJIS漢字
富山日出夫(花園大学)
5. 手書き文字時系列筆跡パタンの一解析と今後の計画
東山孝生、山中由紀子、澤田伸一、中川正樹
(東京農工大学)
6. 絵画DBとイメージ検索 ー浮世絵の線画表現とデータ圧縮効果ー
濱裕光、志賀直人(大阪市立大学)
7. 画像データベースの自然言語インタフェースについて
伊東幸宏、中谷広正(静岡大学)
8. 多視点距離データを用いた3次元形状モデリング
横矢直和、増田健(奈良先端科学技術大学院大学)

懇親会

26日(火曜日)午前10時から午後3時まで

9. ハイパーメディア・コーパスの構築と言語教育への応用について
上村隆一(福岡工業大学)
10. 「歌物語」語彙の数量的分析と研究
西端幸雄(大阪樟蔭女子大学)
11. 高次辞書データベースのための語彙知識自動獲得システム
亀田弘之(東京工科大学)、藤崎博也(東京理科大学)
12. 社会調査結果の視覚化データベース
吉田光雄(大阪大学)
13. 「間」に関するデータベースの構築
中村敏枝(大阪大学)
14. 方言音声データベースの作成と利用に関する研究
田原広史、江川清、杉藤美代子、板橋秀一
(大阪樟蔭女子大学)

クロージング

第2回 公開シンポジウム「人文科学とデータベース」プログラム

日時：平成8年12月21日（土）午前10時より午後4時まで
場所：大阪電気通信大学（寝屋川キャンパス）B310大教室

午前：招待講演

- 「三浦梅園の主著「玄語」のデータベース化と解説の試み
～江戸時代のハイパーテキストを読み解く～」
赤星 哲也，日本文理大学 北林 達也，三浦梅園研究所
- 「江戸図データベースの作成と今後の課題」
黒川 隆夫（京都工芸繊維大学）

午後：一般講演

- 「Shape from motion を応用した什器類の立体データ作成」
中島 重義・岡本 次郎・濱 裕光・細川 省一
（大阪市立大学工学部）
- 「古地図に描かれた内容のデータベース化の試み」
出田 和久・正木 久仁・小方 登・山近 博義
（奈良女子大学文学部）
- 「考古学のためのデータベースシステム」
宝珍 輝尚・中田 充・白井治彦・都司 達夫
（福井大学 工学部 情報工学科）
- 「インターネット・イントラネットにまたがる
分散型 図書館目録データベースの構築と運用」
芝 勝徳（神戸市外国語大学）
- 「音楽における印象語検索システムの開発とその有用性」
原田 章・吉田 光雄（大阪大学人間科学部）
- 「方言認知地図プログラムと統計処理地図」
ダニエル・ロング（大阪樟蔭女子大学）

午後：特別講演

- 「科研費のしくみと申請のポイント」
及川 昭文（総合大学院大学教授／(元)文部省学術調査官）

午前11時より午後4時まで：
♪インターネット・カフェ「じんもん」

第3回 公開シンポジウム「人文科学とデータベース」プログラム

日時：1997年12月20日（土曜日）午前10時より
場所：大阪電気通信大学（寝屋川キャンパス）B310大教室

午前：招待講演 10:00～12:00
「邪馬台国大和説を科学する」

- 「倭人社会と卑弥呼の王権」
岡山大学名誉教授 吉田 晶
- 「前方後円墳成立の歴史的意義
－「邪馬台国時代」における首長と農民層の共同幻想－」
奈良女子大学教授 広瀬 和雄

昼休み（12:00～13:10）

午後の部 一般講演（1） 13:10～14:25

- 「歴史学研究支援システムの構築」
三浦崇，伊東幸宏，小西達裕，田村貞雄（静岡大学），
赤石美奈（北海道大学），中谷広正，阿部圭一（静岡大学）
- 「地理情報を利用した遺物データベースシステムについて」
宝珍 輝尚，都司 達夫（福井大学），河合 秀夫（大阪電気通信大学）
- 「木簡研究支援データベースシステム
－知見と仮説に基づく再構造化－」
森下淳也，大月一弘（神戸大学），上島紳一（関西大学），
大庭脩（皇学館大学），杉山武司（姫路獨協大学）

休憩（14:25～14:50）

一般講演（2） 14:50～16:05

- 「短編推理小説の論理構造の分析」
西島恵介，神山文子，藤田米春（大分大学）
- 「形状分析ツールの開発とその応用
－浮世絵に描かれた役者の同定と分類－」
モハメド・アラミン・ブイヤン，阿古弥寿章，濱裕光（大阪市立大学），
松平進（甲南女子大教授）
- ＜技術紹介＞「3次元形状入力へのおさそい
－人文科学の道具として－」
濱 裕光（大阪市立大学）

休憩（16:05～16:30）

特別講演 16:30～17:30

- 「マルチメディア時代の縄文文化」
国立民族学博物館教授
第4研究部長 小山 修三

クロージング 17:40

懇親会 18:00～

午前11時より午後4時まで：

♪インターネット・カフェ「じんもん」

第4回 公開シンポジウム「人文科学とデータベース」プログラム

日時：1998年12月19日（土曜日）

場所：大阪電気通信大学（寝屋川キャンパス）

午前：招待講演 10:00～12:00

- ★「邪馬台国九州説を科学する ―数理歴史学の立場から―」
産能大学教授 安本 美典 氏

昼休み（12:00～13:10）

午後の部一般講演（1） 13:10～14:25

- ★「中期インドアリアン聖典のデータベース」
逢坂 雄美（仙台電波工業高等専門学校）

- ★「高地性集落遺跡データベースからみた弥生時代の情報通信」
加藤 常員（大阪電気通信大学）

- ★「データベース倫理について」
江澤 義典（関西大学）

休憩（14:25～14:50）

一般講演（2） 14:50～16:05

- ★「階層構造グラフによるデータモデルの適用例：木簡データベース」
杉山武司（姫路獨協大学）、森下淳也、大月一弘（神戸大学）
上島紳一（関西大学）

- ★「遺物破片の計測 ―照度差ステレオ計測装置の製作と計測―」
結城宏和、宝珍輝尚、都司達夫（福井大学）、
河合秀夫（大阪電気通信大学）

- ★<技術紹介>「地場工芸品立体展示システム―人文科学の道具として―」
橋本 隆之（インテックシステム研究所）

休憩（16:05～16:15）

特別講演 16:15～17:15

- ★「選挙研究とデータベース」
関西大学教授 三宅 一郎 氏

クロージング 17:15

懇親会 18:00～

第5回 公開シンポジウム「人文科学とデータベース」プログラム

■日時:1999年12月18日(土曜日)午前10時より午後4時まで

■場所:関西大学総合情報学部(高槻キャンパス内)

■主催:人文系データベース協議会

■共催:関西大学工業技術研究所(知識アーカイブ研究グループ)

■午前の部:招待講演 10:00~12:00

☆「顔の認知研究における顔データベースの利用」

講師:関西大学総合情報学部 教授 加藤 隆 氏

☆「Recent Research on Paleolithic Arts in Europe and the MultimediaDatabase」

講師: Department of Historical Sciences, University of Cantabria, Prof. Cesar Gonzalez Saintz

講演通訳:天理大学国際文化学部助教授 箭内 匡 氏

昼休み(12:00~13:00)

■午後の部:一般講演(1)13:00~14:15

☆「PhotoVR 考古資料データベース『北スペインの旧石器洞窟美術』

日本・スペイン産学共同プロジェクトの実現と諸問題」

深沢 武雄(株式会社テクネ)

☆「バーチャルリアリティによる遺跡探訪 - 3D Archaeo-Copter -」

中村 健、小沢 一雅(大阪電気通信大学)

☆「照度差ステレオ法を用いた遺物の表裏形状の計測」

結城 宏和、宝珍 輝尚、都司 達夫(福井大学)

休憩(14:15~14:40)

■午後の部:一般講演(2)14:40~15:55

☆「文学データベースのための文書の構造化と意味管理」

横田 一正、三宅 忠明、國島 丈生(岡山県立大学)、

劉 渤江(岡山理科大学)、田植 明子((株)リュービシステムサービス)

☆「相対インデックス法を使った文構造分析」

雄山 真弓、岡田 孝、黒崎 茂樹(関西学院大学)

☆「仮想電子辞書システムの設計と構築」

芳野 学、都司 達夫、宝珍 輝尚(福井大学)

■クロージング 16:00

■懇親会 16:30~

主催:人文系データベース協議会

共催:関西大学工業技術研究所(知識アーカイブ研究グループ)

代表者 江澤 義典(関西大学 総合情報学部)

分担者 飯田 記子(関西大学 総合情報学部)

冬木 正彦(関西大学 工学部)

西本 秀樹(関西大学 総合情報学部)

榎原 博之(関西大学 工学部)

堀井 康史(関西大学 総合情報学部)

公開シンポジウム「人文科学とデータベース」1999

発行日 1999年12月18日

発行所 公開シンポジウム「人文科学とデータベース」実行委員会

〒569-1095 大阪府高槻市霊仙寺町2-1-1

関西大学総合情報学部 江澤研究室内

電話: 0726(90)2450

FAX: 0726(90)2450

印刷・製本 株式会社 三協印刷社

大阪市西成区天下茶屋1-28-16

電話:06(6651)5147