第１１回　公開シンポジウム

人文科学とデータベース

国文学と非文字資料

2005 年 12 月 3 日

主 催　：第１１回公開シンポジウム実行委員会
共 催　：西日本国語国文学データベース研究会
協 賛　：情報処理学会関西支部
後 援　：人文系データベース協議会
　　アート・ドキュメンテーション学会
　　情報知識学会関西部会

目　　次

国文学と非文字資料

特別講演

・古典芸能研究におけるデジタルアーカイブの効用 ・・・・・・・・・・・・・・・・・・・・・・・・1
赤間　亮（立命館大学）

一般講演Ⅰ

座長　柴山　守（京都大学）

・「言語地図作成方法の変遷―GIS ソフトの導入と展開―」・・・・・・・・・・・・・・・・・・・・5
鳥谷善史・田原広史（大阪樟蔭女子大学）

・「立地分析のための分布地図表現－分布地図作成支援システムの構築－」・・・・・・・・・・１３
中島高司・加藤常員（大阪電気通信大学）

・「人文科学のための地理情報・資源共有化システム・プロジェクト」・・・・・・・・・・・・・２１
石川正敏（東京農工大学），川西陽一（京都大学），
奥村英史（ヒューマンオーク），原正一郎（国文学研究資料館），
桶谷猪久夫（大阪国際大学），貴志俊彦（島根県立大学），
村尾義和（島根県立大学），柴山　守（京都大学）

・「4D-GIS による遺跡空間データベースの構築と先史文化生態解析」・・・・・・・・・・・・・２９
津村宏臣（同志社大学）

一般講演Ⅱ

座長　田原広史（大阪樟蔭女子大学）

・「国書古典籍中の挿絵・絵本に描かれた実在キャラクター達の存在意義・・・・・・・・・・・３７
―情報学から文学論へのエチュードとして―」
相田　満（国文学研究資料館）

・「浮世絵と写真に基づく幕末から明治に至る女性の形態の定量的考察」・・・・・・・・・・・４７
森下あおい（滋賀県立大学），黒川隆夫（京都工芸繊維大学）

・「嵯峨本『伊勢物語』の木活字及び組版分析モデルに関する報告」・・・・・・・・・・・・・５５
津田光弘（イパレット／奈良女子大学），鈴木広光（奈良女子大学）

・「色彩のデジタル化の諸問題－日本文学作品画像データベースと色彩－」・・・・・・・・・６３
當山日出夫（花園大学）

一般講演Ⅲ

座長　及川昭文（総合研究大学院大学）

・「万葉集校本データベースから古字書データベースへ」・・・・・・・・・・・・・・・・・・・・７３
西端幸雄（大阪樟蔭女子大学），鈴木榮一（万葉情報システム調査会）

・「古今集データベースによる歌語の視覚化」・・・・・・・・・・・・・・・・・・・・・・・・８１
山元啓史（オーストラリア国立大学）

・「PDF ファイルを用いた歴史的資料のデジタル化とデータベースの試作」・・・・・・・・・・８９
竹内さおり・白川哲郎（大阪樟蔭女子大学）

・「階層に基づく遺物データベースシステム」・・・・・・・・・・・・・・・・・・・・・・・・９７
宝珍輝尚（大阪府立大学）

過去のプログラム

公開シンポジウム「人文科学とデータベース」1995 年プログラム・・・・・・・・・・・・・・１０５
第 2 回公開シンポジウム「人文科学とデータベース」・・・・・・・・・・・・・・・・・・・１０６
第 3 回公開シンポジウム「人文科学とデータベース」・・・・・・・・・・・・・・・・・・・１０７
第 4 回公開シンポジウム「人文科学とデータベース」・・・・・・・・・・・・・・・・・・・１０８
第 5 回公開シンポジウム「人文科学とデータベース」・・・・・・・・・・・・・・・・・・・１０９
第 6 回公開シンポジウム「人文科学とデータベース」・・・・・・・・・・・・・・・・・・・１１０
第 7 回公開シンポジウム「人文科学とデータベース」・・・・・・・・・・・・・・・・・・・１１１
第 8 回公開シンポジウム「人文科学とデータベース」・・・・・・・・・・・・・・・・・・・１１２
第 9 回公開シンポジウム「人文科学とデータベース」・・・・・・・・・・・・・・・・・・・１１３
第 10 回公開シンポジウム「人文科学とデータベース」・・・・・・・・・・・・・・・・・・・１１４

【プログラム】

受付　９：３０～

オープニング　９：５０～９：５５

午前の部　１０：００～１１：４０

（10:00～10:25）「言語地図作成方法の変遷―GIS ソフトの導入と展開―」
鳥谷善史・田原広史（大阪樟蔭女子大学）

（10:25～10:50）「立地分析のための分布地図表現－分布地図作成支援システムの構築 －」
中島高司・加藤常員（大阪電気通信大学）

（10:50～11:15）「人文科学のための地理情報・資源共有化システム・プロジェクト」
石川正敏（東京農工大学），川西陽一（京都大学），奥村英史（ヒューマンオーク），原正一郎（国文学研究資料館），
桶谷猪久夫（大阪国際大学），貴志俊彦（島根県立大学），
村尾義和（島根県立大学），柴山　守（京都大学）

（11:15～11:40）「4D-GIS による遺跡空間データベースの構築と先史文化生態解析」
津村宏臣（同志社大学）

昼　食　　１１：４０～１３：００

午後の部　１３：００～１７：４５

特別講演　１３：００～１３：４５

赤間　亮（立命館大学）「古典芸能研究におけるデジタルアーカイブの効用」

13:45～14:00　休憩

（14:00～14:25）「国書古典籍中の挿絵・絵本に描かれた実在キャラクター達の存在意義
―情報学から文学論へのエチュードとして―」
相田　満（国文学研究資料館）

（14:25～14:50）「浮世絵と写真に基づく幕末から明治に至る女性の形態の定量的考察」
森下あおい（滋賀県立大学），黒川隆夫（京都工芸繊維大学）

(14:50～15:15)「嵯峨本『伊勢物語』の木活字及び組版分析モデルに関する報告」
津田光弘（イパレット／奈良女子大学），鈴木広光（奈良女子大学）

(15:15～15:40)「色彩のデジタル化の諸問題－日本文学作品画像データベースと色彩－」
當山日出夫（花園大学）

15:40～16:00　コーヒー・ブレイク

(16:00～16:25)「万葉集校本データベースから古字書データベースへ」
西端幸雄（大阪樟蔭女子大学），鈴木榮一（万葉情報システム調査会）

(16:25～16:50)「古今集データベースによる歌語の視覚化」
山元啓史（オーストラリア国立大学）

(16:50～17:15)「PDF ファイルを用いた歴史的資料のデジタル化とデータベースの試作」
竹内さおり・白川哲郎（大阪樟蔭女子大学）

(17:15～17:40)「階層に基づく遺物データベースシステム」
宝珍輝尚（大阪府立大学）

クロージング　１７：４０～１７：４５

懇親会　１８：００～

古典芸能研究におけるデジタルアーカイブの効用

The utility of digital archiving for the study of traditional performing arts

赤間 亮

Ryo Akama

立命館大学 先端総合学術研究科，京都市北区等持院北町 56-1

Ritsumeikan University, 56-1, Tojiin Kita-machi, Kita-ku, Kyoto

あらまし： 日本の古典芸能の対象に行われている国内外のデジタルアーカイブ事業を紹介し、それらを踏まえて立命館大学アート・リサーチセンターを拠点に著者らが行っている研究を紹介する。ここでは、古典芸能のような無形文化の先端的なデジタルアーカイブ技術や、その応用方法を説明する。また、過去に生成された大量の上演資料を地道にアーカイブする必要性を述べ、それが上演情報をインデックスとすることで有機的に連動し、本来なら消えてしまう時間芸術である舞台上での上演内容を復元することになる点も指摘する。また、様々な機関が公開する目的の異なるデータベースを連動させ、一連の古典芸能アーカイブとして活用する手法についても紹介する。

Summary: In this paper, I will introduce some digital archives about Japanese traditional performing arts. Then I will show you our research activities doing in the Art Research Center of Ritsumeikan University. I will explain some cutting-edge technologies and applied technique. And I will make a point of showing that we should make steady efforts on the digitizing for a huge amount of materials of historical performing arts. The digital archives of traditional performing arts are working effectively by making index from the record of performance (Kabuki chronology), so it makes the results that we can restore contents of performance it was performed and lost everything at the same time in the past. And also, I will point out that some database outside made up with the different purpose from us can work effectively by making a database with a meta-data for traditional performing arts by ourselves.

キーワード: 能狂言，歌舞伎，データベース，演劇，モーションキャプチャ，容貌復元

Keywords: Noh-kyogen, Kabuki, database, theatre, motion capture, reconstruction of a face for historical celebrity

1，はじめに

デジタルアーカイブが、さまざまな分野で粛々と進められている。デジタルアーカイブという言葉が先端的で、その先には夢のような時代が訪れるという言説が先行していたのも、「懐かしい」ことと思える昨今である。したがって、“デジタルアーカイブが必要である”と力説するのが今回の目的ではない。むしろ、今回は、古典芸能研究の分野において、研究用コンテンツがどのようにデジタルアーカイブされているのかの現状を検証し、まさしく、デジタルアーカイブズを構築している私どものプロジェクトでは、それがどのように具体的な研究に役立てているのかを報告してみたい。

2，海外の状況

まず、海外の状況についての調査については、網羅できているわけではないが、いくつかの事例を報告したい。海外の日本演劇研究者が、ヴィジュアルな、あるいはマルチメディア型研究資料を求めていることは、予想していた。能楽や歌舞伎文楽は、海外の研究者にとって魅力的な対象であり、実に多くの研究者を数えることができる。彼らは、日本に留学中、文献資料はもちろんであるが、できるだけ舞台映像を手に入れて本国に帰る。本国の大学の講義で、学生に日本の古典芸能を紹介するのに、必要不可欠の教材であるからである。

2002年秋にロンドン大学では、日本学研究会の例会でスコットランド大学のHelen Parker 氏による発表があり、歌舞伎を教えるのにビデオがどれほどの効果があるかを力説された。その段階で、歌舞伎の舞台映像はほとんど販売されていなかったが、通常は、日本滞在中にNHKやCS放送の伝統文化放送(歌舞伎チャンネル)で流された映像をコピーして持帰っているようである。それも短期間での留学では十分な蓄積はできない。現在は、DVDによる販売が少しずつ進みはじめた。Parker 氏は、この発表を発展させて、2005年度、11th Conference of the EAJSにおいて、Border-crossing and multimedia resources in the study of traditional performing artsと題する発表をされているが、その要旨によると図像や映像が演劇研究において、とくに他ジャンルとの比較をする上で、いかに有効かを説いたものであったようだ。

本年10月11月と、大阪歴史博物館では、6月から9月まで大英博物館で開かれた展覧会の巡回展である「日英交流　大坂歌舞伎展」が開催されていた。12月から1月にかけては、さらに東京の演劇博物館に巡回する。この展覧会の実質的なディレクターは、ロンドン大学SOASのAndrew Gerstle 教授で、氏は5年の歳月をかけて、イギリス国内に存在する浮世絵を中心とした大阪歌舞伎の資料を悉皆調査され、調査と同時にデジタルカメラで資料をデジタルアーカイブされた。私は、2002年から2003年にかけてGersle教授をスーパーバイザーとして1年間、SOASに滞在し、共同でGerstle 教授の研究資料の画像データベース構築を担当した。これが現在、私の所属するアート・リサーチセンターのサーバー上で運用されており、今回の展覧会の研究・解説執筆の上で極めて有効に働いたと聞いている。

また、やはりロンドンのビクトリア&アルバート博物館では、38,000 枚という膨大な浮世絵を所蔵しているが、担当学芸員が15年も以前に目指した、「コンピュータを使った整理」が、私の研究テーマとして合致し、デジタルアーカイブとデータベース構築のコラボレーション研究が始まっている。すべての浮世絵のデジタル映像化とメタデータ付与の作業が3年後には完成する見込である。

3, 日本の古典芸能研究所アーカイブ

一方、日本では、古典芸能を研究対象にする研究所が、いくつも活動しているが、WEB上での資料公開に力をいれている研究所は比較的少ない。

最も先行しているのが、早稲田大学演劇博物館で、ここでは、所蔵する浮世絵(役者絵が中心)が約47000枚公開されている。これは、役者絵としては世界最大コレクションであるが、浮世絵公開データベースとしても世界最大である。また、義太夫の院本(義太夫節の台本)や、戦前の歌舞伎の舞台写真を見ることができる。神戸女子大は、民俗芸能にかかわる喜多コレクション(写真)を公開している。

ただし、浮世絵に関しては、国内の方が、海外よりも先行している模様で、演劇博物館以外にも、国会図書館、東京国立博物館、都立中央図書館などが大規模の所蔵品をWEB公開している。

なお、海外の浮世絵データベースとしては、サインフランシスコ美術館やライデン国立博物館などがあるが、十分なメタデータが付与されておらず、的確に所望する資料を見つけ出すまでのデータベースには到っていない。

4, GloPACプロジェクト

日本の古典芸能研究所におけるデジタルアーカイブが、このようにデジタルライブラリ的な段階に止まっており、演劇資料としての体系的なアーカイブに進む動きを見せていないのに対して、WEB大国たるアメリカでは、コーネル大学を中心とする演劇研究コンソーシアムが始まっている。The GLOBAL PERFORMING ARTS CONSORTIUMといい、GloPACと略されている。これは、組織や個人が演劇の研究や保存をするために、デジタル技術を駆使して、多言語対応、マルチメディア対応の資料に簡単にアクセスできるようにしようとする国際的なグループであり、コーネル大学、コロンビア大学、ワシントン大学、ハワイ大学、ビンガムトン大学(以上、米国)、シンガポール国立大学、法政大学などが参加している。

これは、演劇文化の盛んな各国の演劇資料を集積していこうとするWEB先進国アメリカならではの発想である。その中で、最も活発に活動しているのが日本演劇研究部門であり、海外の日本演劇研究者にとって、資料への容易なアクセスが如何に緊急の課題であるのかが理解できる。なお、データ蓄積のためのチームをこのプロジェクトでは、PARC(A Performing Arts

Resource Center)と読んでおり、この活動の意義を普及させる先導役プロトタイプとしてJPARC(Japanese Performing Arts Resource Center)の活動をWEB上でも確認することができる。
(http://www.glopac.org/gloparc/index.php)

こうした動きをみても、日本の伝統芸能資料は、現代の舞台映像や音声記録は別として、日本が世界最大の古典演劇資料の宝庫であることにも気付かされるのである。だが、問題点を指摘しておくと、本プロジェクトは、文系の研究者が中心であるが、それにデータベーステクノロジを組合せ、むしろこのグループが大量の古典芸能コンテンツを持っているのではなく、システムを提供するから各国のコンテンツを参加者側から提供し、しかも参加者が各自で書誌データを書込むという点である。日本やヨーロッパなどの歴史的な演劇資料を持つ、コンテンツホルダーにとっては、どのようなスタンスでこのプロジェクトに協力していくのか、難しい判断を要求される。

一方で、上記の参加大学の内、法政大学は、具体的には能楽研究所が積極的な関りを持っている。この研究所は、能楽分野専門研究所としては、最も歴史が古く、最大の機関である。しかしながら、上述、日本の演劇研究所の章でも全く触れなかったことで分かるとおり、デジタルライブラリとしての活動は、現状では全く見えてこない。(実際、デジタル撮影をすすめているとの報告が、GloPACのシンガポールでのカンファレンスで報告された。)

能楽研究所のような機関の場合、研究所員が完全に文系研究者であり、システム開発やデジタルアーカイブそのものを研究する必要はないと思われ、こうしたスタンスの機関にとっては、システムを提供してもらえるコンソーシアムに参加する意義は大きいと思われる。

5, 立命館大学アート・リサーチセンターにおける活動

立命館大学アート・リサーチセンター(以下、ARC)では、これまで、文部科学省の学術フロンティア推進事業による「無形文化財・時間芸術に関する総合的研究」、文部科学省オープン・リサーチ・センター整備事業による「デジタル時代のメディアと映像に関する総合的研究」、さらには、21世紀COEプログラムにより「京都エンターテインメント創成研究プログラム」により、いわゆる文理融合型の研究をとくに京都の大学として、日本文化に焦点をあてて進めてきている。その内、片山家能楽・京舞保存財団との提携による能楽のマルチメディア研究や、京都が発祥である歌舞伎の歴史に関するデータベースは、現在最先端の研究成果として知られている。

ARCは、7年前に設立された研究所であり、他の研究機関より後発であった。そのため、設立当初から、他の研究機関が持たない研究対象や手法を取り入れることで独自性を出すことに腐心してきており、古典芸能のデジタルアーカイブ研究としては、世界に類を見ない存在となっている。

ARCでは、すでに公開・非公開のデータベースが構築されており、とりわけ他機関の所蔵品のデジタルアーカイビングに協力し、多くのデータベースが立上っている。また、テキストや画像、動画といった既存のメディアを対象にしたデータベースだけでなく、新しい記録方法、鑑賞方法を開発するという技術開発も行っているところが特徴である。私は、デジタルアーカイブ研究を標榜するためには、技術開発と、アーカイブ構築の両面を併せ持つ必要があると主張するものであり、ARCの研究活動は、まさしくこの二面を併せて進んでいると言えよう。

1)舞台のアーカイブ(照明を含む))
2)役者の動きのアーカイブ
3)役者の衣裳のアーカイブ
4)役者の容貌や肉体のアーカイブ
5)舞台音楽・音声の感性的アーカイブ

以上は、現在、立命館大学21世紀COEの研究テーマとして、アーカイブ技術の開発・応用に関して実際に進行しているものである。1)については、現在までに南座の劇場内、さらには、国宝である西本願寺北能舞台が、きわめて高精細の3DCGが出来上っている。2)の役者の動きは、能役者、歌舞伎役者の他、日本舞踊や島原の太夫の動きなどが記録されている。短時間の作品であるが、これにキャラクターを載せたバーチャル能も制作されている。また、先述の大英博物館展では、会場の解説ビデオのなかで、歌舞伎の舞台実映像から、CGの演技となり、さらに展示されている浮世絵に変化していく作品を製作・展示した。これらの演技の記録データは、このように様々な応用ができ、舞台のCG上に容易に別々に収録した役者の演技を展開させることも可能である。

また、4)では、幕末の名優、8代目市川団十郎の容

貌を、残存する役者絵や、最古の隈取を元にして復元した。この復元技術を応用することで、ポートレートが残る過去の有名人については、立体的な容貌を復元させることのできるわけで、衣裳の現実感、音声の感性的な表現を含めて、現実の舞台や舞台をビデオ撮影した映像とは全く別の独立したメディアによる舞台芸術が実現できると考えている。音楽において、生演奏とレコードとが全く別物として鑑賞者を獲得しているのと同じ現象を、舞台芸術の世界でも実現しようというのが、私の構想である。

さて、こうした先端的な技術開発とともに、重要ものが、もう一方の過去に残されてきたさまざまな演劇資料のデジタルアーカイブである。ここでは、すでに標準的なデジタル化基準が定着し、安心して大量のデジタル化が行えるコンテンツ対象である必要が、まずある。そして、それらに対し、目立たず地道に粛々とアーカイブ作業を進めていくことが肝要であり、その地味な活動の成果が実は最大の研究成果を導き出すことを確認しておかねばならない。とりあつかわれるメディアは、

A)テキストアーカイブ(台本や音楽の電子テキスト)
B)画像アーカイブ(番付ポスター劇評・劇書などの幕外資料や、台本、興行日記などの幕内資料など、演劇資料の複製画像、)
C)音声アーカイブ(レコード・CD・ビデオからの音声など)
D)動画アーカイブ(舞台収録映像)
E)新メディア・データアーカイブ(3D、CG、モーションデータなどの新技術による記録データ)

などである。

Eについては、を新しい記録方式であり実際の上演を記録するまでには到っていないが、その他は、過去にある舞台上での上演があり、その製作・上演活動のなかで、様々な形態として生産され、残ったものである。一つの上演事実から生れ出た資料は、その後バラバラに拡散する運命にあり、以降それらは、物体として同じ場所に完全に集めることは不可能である。しかし、デジタルアーカイブ技術によってそれらの分散した資料が、再び一箇所(つまり上演情報のインデックス)に集められることが可能となった。例えば、上演年表データベースをインターフェイスとして、同じ時の上演資料を一括し、本来なら時間芸術として、消えてしまうはずの舞台を上演当時のテキストやイメージによって、再現することができるのである。

6, 外部データベースとその連携による効用

このような古典芸能のデジタルアーカイブに関しては、資料がさまざまな機関に分散しており、WEB 上で資料公開をしている機関においても、その目的や公開手法が微妙に異なるため、効率よく引出してくることが出来ない場合がある。例えば、国会図書館の錦絵データベースは、図書館の資料として整理され、メタデータが浮世絵に特有のデータ構造になっていないのはもちろんのこと、役者絵としての情報は付与されていない。そのため、1 万1千万もある浮世絵資料が歌舞伎史図像データとしては全く機能していない。この様な欠点を指摘しても、不特定多数の閲覧者のために公開している国会図書館には、歌舞伎史用にデータを追加することを期待すべくもない。

そこで、一つの方法として、公開されているシステムと連動する歌舞伎史研究用インデックスデータを外部に構築し、それを役者絵データベースと統合して、さらに上演年表データベースを連動させることで、国会図書館において公開されているデータベースへの役者絵としてのアクセス・エントランスを確保できるのである。このように古典芸能に関するデータベースは、それに関連する WEB 上の学術データベースの可能性をより大きくする効用がある。

7, 演者の意識改革

最後に、隠された効用を一つあげておく。新技術によるデジタルアーカイブには、かならず演者や所蔵者の理解が必要である。資料の収蔵管理者は、デジタルアーカイブという事業に対して比較的慎重に対応する必要があり、現実に資金的な事情も含めてなかなかアーカイブ事業が進まない。しかし、古典芸能の演者(役者)たちは、モーションキャプチャやマルチアングル撮影に対して比較的積極的な意識を持っている。それは、それらの記録することによる話題性にもよるだろうが、古典芸能の演者であるから保守的な意識が強いのではなく、古典であるからこそ、いかに新しさを保ち続けるかという課題に日常的に取組んでいるからであろう。

[注]なお、モーションキャプチャに関する研究は、本学情報理工学部八村広三郎氏、役者の容貌復原については、同じく田中弘美氏との共同研究である。

言語地図作成方法の変遷
－GISソフトの導入と展開－
A History of Drawing Techniques in Linguistic Atlas
- Application and Development of GIS Software-

鳥谷 善史 , 田原 広史
Yoshifumi Toritani , Hiroshi Tahara
大阪樟蔭女子大学 学芸学部, 東大阪市菱屋西 4- 2- 26
Osaka Shoin Women's University, 4-2-26, Hishiyanishi, Higashiosaka-shi, Osaka

あらまし:本稿では、方言学や言語地理学において作成されてきた言語地図の作成方法の変遷を情報処理技術の導入やデータベース作成の観点から整理し、現在導入されつつあるGISソフトを用いた方言地図データベース作成の現状や今後の可能性について報告する。また、言語地図に音声データを組み合わせた「声の言語地図」を取り上げ、マルチメディアを活用した地図データベースの現状と今後の展望についても言及する。

Summary: This paper illustrates a history of drawing techniques in linguistic atlases both in dialectology and linguistic geography with a reference to the application of the information technology and the development of linguistic databases. This paper will argue the current situation and the future directions of this area. It will also raise an 'acoustic linguistic atlas,' which combines acoustic data with the geographic data. Discussion will be made on the present status and future views in the geographical database along with multimedia.

キーワード:言語地図, GIS, 音声データ, マルチメディア
Keywords: linguistic atlas, GIS, acoustic data, multimedia

1. はじめに

言語地図とは、調査した地点の「ことば」に関する情報を地図上に表現したものである。一般的には調査地域の白地図を用意し、調査地点上に記号を配してその地点の「ことばの形」を表現する。ことばの分布域や勢力域、また、混交の様子等を解釈し、新古関係を解釈することにより、文献に現れなかったことばの歴史を再構築する。

ただし、ことばの分布には新古関係以外にも様々な言語外の社会的要因が働いている。昨今は、山脈・交通網・海路・藩境などとの相関を言語地図から読み取ったり、ことばの変異や変化と当該地域の社会との相関を読み取ったりしようとする社会言語学的研究も盛んになってきている。

3で改めて述べるが、これらの言語地図作成の手順は、調査地域を設定し、調査を行い、当該地域の白地図にハンコやシールを使って言語地図を作成するというものであった。そのため、結果の整理、記号の選択、押印等の作業には、かなりの時間と労力を費やすことになる。地図作成に際しては、まず「資料地図」、次に同類としてまとめられる変異形をまとめた「まとめ地図」、そして、最終的に研究者の視点にたった「解釈地図」ができあがる。こうして出来上がった「解釈地図」も、記号の選択が不適切で、解釈内容が上手く表現できていなかった場合は、再度記号の選択から行わなければならない。

このように、かなりの時間と労力を費やすことから、この分野における情報処理技術導入の機運は、他分野と比べてむしろ早かったと言える。しかし、ことばのデータが量的というよりもむしろ質的であること、また、併用語形の表し方の問題や、機器そのものの処理能力の低さ等、克服しなければならない問題が多かったため、実質的な導入は、Windowsの登場を待たなければならなかった。

現在ではGISという地理学を中心とする分野から展開されつつある、新しい情報処理技術の登場により、

地図作成に関わる言語地理学や社会言語学の分野でも新たな展開が期待されており、多くの研究者が努力している状況である。また、情報処理技術の進展により、音声データや画像データもコンピュータ上で活用できる状況となり、これらのデータを活用したマルチメディア地図とでも呼ぶべきものも登場しつつある。本稿では、言語地図作成の変遷と現在取り組まれつつある様々な試みを紹介し、現在の問題点や今後の可能性を検討する。

2. 言語地図の歴史

国際的レベルで言語地図の歴史を見た場合、徳川1993によれば、1879年、ドイツにおいてG. Wenkerの行った調査に始まるとされる。その後、言語地図の著名なものは、1897年から1901年に調査され1902年から1910年にかけて刊行されたJ.GilliéronとE.Edmontが作成した*Atlas linguistique de la France*(『フランス言語地図』)である。

日本における言語地図作成の歴史は古く、明治政府の文部省内におかれた国語調査委員会が1903年から1904年にかけて調査し、1905年に刊行された『音韻分布図』(地図29枚)と、1906年に刊行された『口語法分布図』(地図37枚)の66枚の地図がその端緒である。このように早い時期に言語地図を国家規模で作成した理由は、「国語ノ音韻組織ヲ調査スルコト」「方言ヲ調査シテ標準語ヲ選定スルコト」という目的があったためであり、近代化を急いだ当時の明治政府の言語政策観がうかがえる。

その後、しばらくの間、日本における言語地図にはほとんど見るべきものがなかった。実際には、明治末期に第二回目の調査が実施されていたが、そのデータは関東大震災によって失われてしまった。以降のこの方面の研究としては、戦後、国立国語研究所が刊行した『日本言語地図』(300枚)1～6集(1966～1974)を俟たなければならなかった。また、その後の大規模な言語地図集としては、やはり国立国語研究所が現在刊行中である、『方言文法全国地図』1～5集(1989～2002;6集までの予定)があげられる。

大規模なものを除き、各地で研究者が刊行している地図集は実際の所その数を正確に把握することはできないのであるが、「日本方言研究会」のホームページ(http://wwwsoc.nii.ac.jp/cdj/)に日本方言研究会に設けられた言語地図目録作成委員会の大西拓一郎氏が分担した調査対象(単行本)の一覧が公開されている。

このリストには約330冊の言語地図関連の単行本が掲載されているが、対象となった単行本は国立国語研究所図書館に所蔵する図書であること、最終更新日は2000.11.20であり、それ以降の分を含んでいないことを考え合わせると、実際に刊行されているものは400冊を下らないものと考えられる。

3. 言語地図作成方法の変遷

この章では、言語地図作成方法の変遷について述べる。まず、従来の手書きトレース法から始まり、続いてコンピュータを活用した言語地図作成の変遷にふれるが、コンピュータ利用については、さらに「汎用機を用いた試み」、「独自アプリケーションの開発」、「市販アプリケーションソフトの活用」という三つの視点に分けて説明する。

3. 1 手書きトレース法

第1章でも述べたとおり、言語地図作成のもっとも基本的な方法は、原図となるべき白地図を用意し、そこにオリジナルのハンコやシールを貼り付ける方法で作成されてきた。これを出版物にする場合は、印刷所でトレースをしてもらい刊行することとなる。この方法は、校正等を含め莫大な手間と時間及び費用がかかるのである。

この方法が手間・時間・費用の面で極めて効率が悪いのは明らかであるが、それにもまして、語形をハンコやシールの記号に置き換える際、かなりの経験と熟練を要すること、一度作成し、分布が表現できなかった場合は再度記号を別の記号に置き換えて地図化作業を行う必要があることも、手間・時間の面で大きな問題点であった。

これは、言語地図データは、基本的に質的データであるからである。最初に、調査目的や考察内容を決定し、これにそった調査項目や調査地点を選定し、調査目的や調査方法にあったインフォーマント(情報提供者)を選定する。そして調査を実施し、結果を集計して地図化する。つまり、言語地図作成は、基本的には仮説をたて、それを調査で検証し、地図上に明示していく知的作業である。単に地図上にデータをプロットしただけの「資料地図」では、目的の言語地図は完成しないのである。この辺りの事情から、コンピュータ化になじみにくかったのだと考えられる。

手書きトレース法による言語地図作成の手間や時間の部分を何とかして効率よくできないかという欲求は、極めて大きく、かなり早い時期からコンピュータを活用して地図を作成する試みがなされてきた。以下では、そのあたりの変遷について述べることにする。

3. 2 コンピュータを活用した言語地図作成

3. 2. 1 汎用機を用いた試み

言語地図をコンピュータで作成できないかとの考えを示したのは、管見によれば、山本 1965 及び徳川・山本 1967 である。

山本は、当時日立システム・エンジニアリング第一技術部に所属し、国立国語研究所が汎用機を導入した際の技術者である。一方、徳川は国立国語研究所地方言語研究室に在籍し、『日本言語地図』の作成を担当していた。

1948 年 12 月に国立国語研究所が設立された後、その設立の理念から、言語地図作成の要望に応えるかたちで、1955 年から『日本言語地図』の作成にとりかかる。調査は 1957 年から 1965 年まで 65 名の調査者により、北海道から沖縄までの 2,400 地点を 285 の調査項目で実施された。単純計算で延べ 65 万件に及ぶデータを 285 枚の地図にする作業がそこにはあったのである。

徳川・山本 1967 では、汎用機で言語地図を作成する試みの目的として、

1) 現行の手工業的言語地図作成行程に、電子計算機がどれだけ寄与できるか、その方向づけを行ない、限界を見極めて、新しいシステムの開発の資料とする。
2) 一般に分布図作成や言語の統計的処理に、電子計算機がいかに利用されうるか、それを認識するための資料とする。

といった二点をあげている。また、その結論として、

> 言語地図の作成について機械的作業の見通しがつき、今後さまざまな予測されないデータをもととして、高度のアルゴリズムまで高めて定式化することができれば、いっそう計算機らしい利用法が生まれるであろうことが明らかとなった。

と述べている。1967 年というかなり早い段階から地図化の試みがなされていたことが理解できるのである。

3. 2. 2独自アプリケーションの開発

① GLAPS (Generalized Linguistic Atlas Printing System)

徳川らの試みからほぼ 10 年が過ぎた頃に、後のGLAPSの端緒となる発表が、第 20 回日本方言研究会においてなされている(荻野 1975)。内容は、東大教育用計算機センターの MELCO-7700 を使って、プログラミングをおこない、「LINGUISTIC ATLAS OF SIZUKUISI」を作図する方法を述べたものである。この発表により、方言研究者自身がプログラミングをしてアプリケーションソフトを開発するといった時代に突入した。

GLAPSは、本来言語地図作成を目的に開発されたプログラムであるが、その後、「名義尺度データを解析する統計プログラム」、「各種アンケート調査の解析ツール」といった利用が中心となってきた。GLAPSの公開は、1981 年にユーザーズマニュアルの初版が発行されていることから、この頃からだと考えられる。また、1994 年に第2版が出版されていることから、一定の利用者が存在したこともうかがえる(荻野 1994)。

その後、荻野 1996 として科学研究費の報告書である『パソコン版GLAPSの開発』が出版された。この時期にハードウェアに大きな変革が起こり、研究の世界でもこれまでの大型汎用機(大型電子計算機)の時代からパーソナルコンピュータの時代へと変化したため、GLAPS自体も新たなハードウェアの変化に対応しようとしたのである。

GLAPSを使って作成された報告書や研究論文は多数見られることより、このソフトが言語研究の分野で重要な位置を占めてきたことがうかがえる。荻野 1994 の「付録2 GLAPSを用いた研究」には 110 件を超える研究成果があげられている。

② SEAL (System of Exhibition and Analysis of Linguistic Data)

SEALのユーザーズマニュアルがはじめて出版されたのは 1983 年のことである(福島 1983)。GLAPSのマニュアルが 1981 年に出ていることを考えれば、このソフトもコンピュータによる言語地図作成の初期から存在し、現在もバージョンアップを重ねている極めて息の長いソフトであると言える。ソフトの詳細については、福嶋 1994 に述べられている。

最新のマニュアルは福嶋・福嶋 2004 として出版されている。また、PDF 版が http://www.nicol.ac.jp/~fukusima/inet/index.html に公開されている。このアプリケーションソフトは、対話型である点だけでなく、バージョンアップのタイミングが時代をうまくとらえており、その点においても評価できる。

プログラミング言語は当初 N88BASIC で記述されていたようであるが、その後 OS にあわせたバージョンアップの際に、Visual BASIC に変更されている。現在の多くのアプリケーションソフトも、基本的には同じプログラム言語で記述されているのであるから、バージョンアップのタイミングやその拡張性を考えた場合、言語地図作成という分野では、今後ますます重要なソフトになる可能性がある。

GISとの関連で述べれば、最新バージョンの 7.0Jからは、いわゆるレイヤの機能を新たに加え、属性データの考え方を導入したことなどがあげられる。また、言語学の分野で使用する特殊なフォントであるIPA(the International Phonetic Alphabet:国際音声記号)を、凡例文字に使用できるようにしているなど、実質的な言語地図作成に対する細やかな配慮がなされている。

③ EGL(Editor for Geo-Linguistics)

EGLが最初に紹介されたのは、管見によると、1988年の第 15 回音声言語研究会においてである(前川 1988)。この付記に「EGLはまだ開発中のプログラムですが、試用を希望される方は筆者まで連絡頂ければ・・・」とあり、この頃から試用版の配布が始まったと考えられる。また同要旨に、

> EGLは 16 ビットパソコン PC-9801 で稼働する言語地図作成支援システムである。編集機能に優れた大型計算機上のGLAPSと地図の表現力に優れたパソコン上のSEALの特徴を両立させることを目的として開発された。EGLでは完全な対話型の操作によって、語形のまま入力されたデータの編集と言語地図の作成が可能である。地図の出力媒体は現在のところ CRT のみであるが、将来的にはプロッタ等を予定している。

とあることから、このソフトの開発のねらいは、GLAPSとSEALの長所を併せ持つものを作るという点にあったことがわかる。

その後、前川 1993 において、レーザープリンタを出力装置とすることで、かなり大規模でかつ短時間に美しい地図を印刷することができるようになったとの報告がある。しかし、より使いやすいソフトとするため、またOS の変更に対応するためには、「ソフトウェアのプラットホームを MS ウィンドウズなどに変更する必要がありますが、そのためにはプログラムを実質上すべて書き直さなければなりません。本業が忙しいこともあって、いまのところ二の足を踏んでいます。」とあり、結局、Windows 対応のものは作成されなかったようである。

最後に、このソフトの良い点を少しだけ述べておきたい。それは、併用語形に対する記号化も可能なこと、また、語形データの統合がソフト上でできることである。まさに、「言語地図作成支援プログラム」といえる存在であった。SEALについても、1999 年 version 5.5a 以降、併用をそれまで重ねて印字していたものから左右上下にずらして印字することが可能となっている。

④ その他のソフト

以上では、汎用性を持ち、公開され入手可能であり、かつ、言語地図作成を主たる目的として開発されたものを示した。これら以外にも、特定の調査を地図化するためのプログラミングは個別になされていた。

ここでは、以下の4つの研究を紹介しておく。

1) 柴田武氏が 1984 年頃から使っていた『糸魚川言語地図』を作成するためのソフト(柴田・井口 1992)
2) 高橋顕志氏が 1986 年に調査資料から結果を地図化するまでの一連の流れをシステム化し、後にSUGUDAS と命名したシステム(高橋 1992,2002a, 2002b、地図については以下のアドレスで公開中)
   http://home.hiroshima-u.ac.jp/hoogen/index.html
3) 田原広史が報告書を作成するにあたって、パソコンで山口県の方言地図を描くためにおこなったプログラミング(藤田・田原 1990)。
4) ダニエル・ロング氏が、1995 年頃から一般人の諸方言の使用領域に対する認識を統合し、方言使用領域意識を地図上に表示するために開発した「方言認知地図を処理するプログラム」(ロング,ダニエル 1990,1995)

3. 2. 3 市販アプリケーションソフトの活用

① ファイルメーカーpro と Excel による言語地図

筆者の一人である鳥谷は、表計算ソフト Excel とデータベースソフトファイルメーカーproを使い、『大阪府言語地図』を作成し公刊した(岸江・中井・鳥谷 2001)。

なお、この方法を最初に考え出したのは、筆者の一人である田原である。この方法による作成上の利点は、地図上に与えた記号を後から何度も変更し、より解釈を施した言語地図を作成できる点であった。作成方法の手順としては、あらかじめ画像データの地図を用意し、その地図をファイルメーカーのレイアウト上に貼り付け、貼り付けた地図上にフィールドを配置する。また、Excel 上では語形に対応した記号に置換しておき、最終的にフィールドに Excel の記号データをインポートすれば地図は完成する。

このインポートした記号データで思うように分布が現れない場合は再度、Excel 上で記号を置換してインポートを繰り返すのである。作成方法の詳細については、岸江・中井・鳥谷・石田 2000、中井・坂口 2000、鳥谷 2002、田原 2002 を参照されたい。

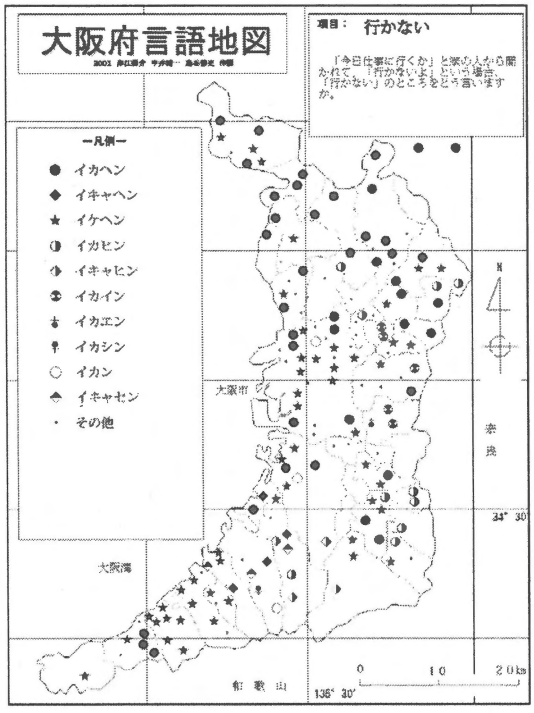
この時期の問題点としては、コンピュータに実装されている記号の種類が少なかったため、意図するような地図が表現できなかったところにあった。それに対し当時は、外字領域に新たに記号を作成することで対処した。その後、外字領域を使用するよりむしろ記号のフォントを作成し、研究者間で共有したりする方が簡便であることから、石田祐子氏により作成された「言語

朗」というフォントを使用するようになり、記号の種類の問題は解消された(岸江・木部・石田 2002)。

ただし、この作図方法は、あくまでも印刷をその主眼においている点である。この頃から、GIS ソフトへの関心が高まってはいたが、地図上での分析や他の相関データとの検証の方法までは及ばなかった。しかし、データを電子化していることから、今後 GIS へスムーズに移行できる可能性があるものと考えていた。

さらに、技術的な問題として、見栄えの良い調査地域の原図作成に際してはそれなりの習熟が必要なこと、ファイルメーカーproの操作に慣れる必要があること、などがあげられる。つまり、それぞれのアプリケーションソフトをそれなりに使いこなすスキルが求められるということである。

図1 ファイルメーカーpro と Excel による地図



② Adobe Illustrator による言語地図

大西 2002 にあるように、Adobe 社アプリケーションソフト Illustrator に新たに自動化ツールプログラムを組み込むことにより、国立国語研究所 2002『方言文法全国地図』5集が公刊された。4 集までは、手作業地図を原稿として、印刷所が書き直す方法(いわゆるトレース法)によって作成されてきた。3. 1で述べたようにこの方法は、確かに美しく見栄えの良い印刷物を提供してくれるが、校正の手間がかかる上、そのこと自体が価格にも反映してくる。まして地図化は手作業であり、多くの労力を必要とする。

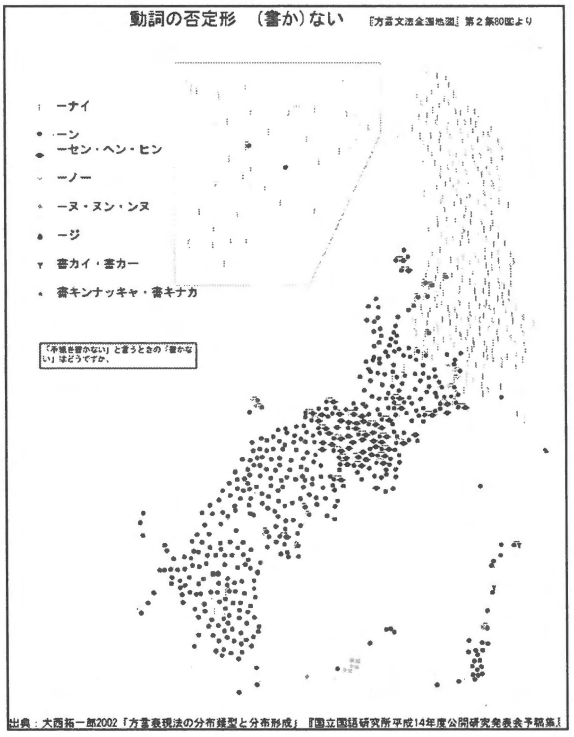
大西は Illustrator の利点として次の三点をあげている。

1) データがベクトルデータであるため、拡大や縮小をかけても曲線にギザギザが生じないこと
2)スウォッチと呼ばれる各ファイル内で記号や色などに固有のコードを与えて管理する機能があること
3) 印刷用プリンタとの相性が良いこと

この中で、1 と 3 は仕上がりの良さであり、2 は先の Excel やファイルメーカーproの方法で問題になった記号の種類の問題である。5 集という性格上、1〜4 集とあまり違うものは作ることができないが、その制約の中でおこなった最大限の電算化なのだろう。

データとプログラムの公開もインターネット上で行われている(http://www.kokken.go.jp/hogen)。これにより、研究内容に応じた新たな解釈地図の作成が可能となるが、そのためには Illustrator を所有しており、かつ、その操作に関する一定のスキルが習得できている必要がある。

図2 Illustrator による地図



出典:大西拓一郎 2002「方言表現法の分布類型と分布形成」『国立国語研究所平成 14 年度公開研究発表会予稿集』

4. GISソフトの導入と展開

筆者の一人である鳥谷は、前出の『大阪府言語地図』作成の段階で、GISソフトによる地図化を試みた。同書の共著者である、富山大学の中井精一氏の計らいにより、「MapInfo」を利用する機会を得たが、この頃の状況としては、実際にこのソフトを使用して『大阪府言語地図』を完成させるには、費用、技術の両面において未だ熟していなかった。

また、書籍形式での公刊を考えており、GISソフトそのものが印刷ということをあまり意識していないように感じられたことも、この段階でGISを導入しなかった大きな理由である。しかし、その後「富山大学人文学部GIS研究会」が結成され、鳥谷もメンバーとして具体的にGISソフトによる言語地図作成に取り組み始めた。これらの成果は鳥谷2003として公刊した。

このように研究は着実に進んでいるのではあるが、現状では誰もがGISソフトを使って簡単に言語地図を描くという状況ではない。方言学や言語地理学また社会言語学の分野において、これらのソフトを各研究者が同じような環境で持ち、かつ操作することは極めて困難であると考えている。

先ず言語地図作成におけるGIS導入の第一段階として、廉価でしかも汎用性のあるソフトが必要である。将来の一元的なデータベース化を考えた場合、専用のGISソフトが必要でもある。また、教育の場での使用なども考慮に入れるならば、操作が簡便で習得が比較的容易であり、統計データの地図化機能を重視したGISソフトを利用することが重要であろう。

そこで鳥谷は専用GISソフトと並行し、次に述べるMANDARAによる言語地図作成もおこなってきた。

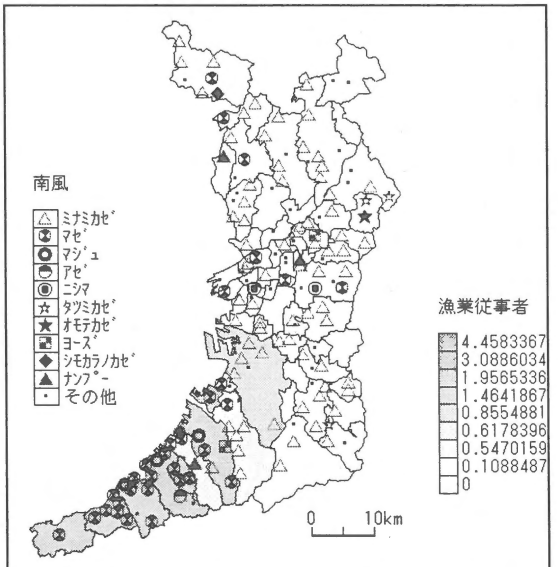
4. 1 MANDARAによる言語地図

『地理情報分析支援システムMANDARA』は、埼玉大学の谷謙二氏が開発したWindows対応のシェアウエアのGISアプリケーションソフトである(詳細はhttp://www.mandara-gis.net/を参照)。

このソフトを使用して先ず、『大阪府言語地図』と同じものを作成した。データをExcelで管理していたことから、MANDARAでの作図が簡単にできた。完成した地図にレイヤ機能を使い、鳥谷2001の論述内容を地図化してみた(図3)。この地図の解釈や意味を簡単に述べる。

一見、方言区画に対応して、和泉方言地域に南風を意味する方言形式「マゼ・マジ」が多いように感じられるが、1950年頃の漁業従事者割合の地図と重ねあわせてみると、この語形は、方言区画ではなく、実は漁業という職業の分布と密接に関係しているということが、視覚的・直感的に理解できるのである。

図3 MANDARAによる地図 レイヤ機能使用



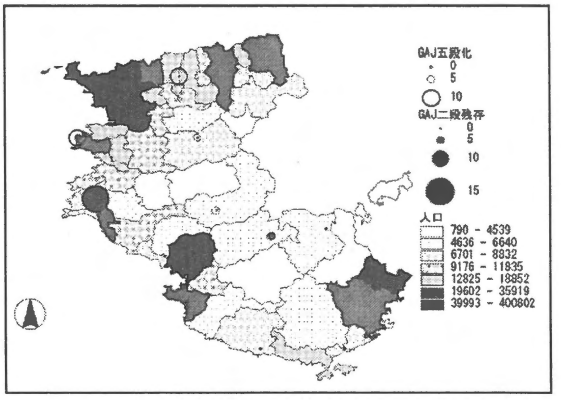
「マゼ」の分布と漁業従事者人口割合の相関図

4. 2 Arc GIS(View)を活用した言語地図作成

松丸2003では、Arc Viewを利用して、和歌山県における、動詞二段活用の残存状況と一段動詞のラ行五段化の発生率を地図化している。その際、前出の『方言文法全国地図』のデータを用い、発生地域のそれぞれの度数を地図上に示し、それらの状況と標高データや人口データを重ね合わせることにより、残存状況や発生率と標高や人口の関係を考察している(図4)。

結論として「(1)ラ行五段化については、人口が集中する和歌山市の周辺地域に見られること、(2)二段活用の残存については、その外側の比較的人口が少ない地域に見られること」を指摘している。

図4 Arc Viewによる地図 レイヤ機能使用



大西 2004a,2004b では、同じく『方言文法全国地図』のデータから言語地図を描き、ことばの分布と標高や人口との相関を考察している。ONISHI 2005 では、調査年代の異なる過去の言語地図を新たに GIS 上に再現し、重ね合わせることにより時間軸を加えた研究を行っている。

この試みは、2 でも述べたこれまで出版されてきた 400 以上はあるであろう、さまざまな言語地図のデータベース化を意味し、GIS 上で全ての言語地図が再現されるならば、20 世紀以降のことばの変化や変異の様子を通時的に示す事ができるであろう。そして、これらの膨大なデータからモデル化されることばの変化や変異から、ことばの分布に対するシミュレーションも可能になるのではないかと考えている。

5. マルチメディア言語地図

最後に、コンピュータを活用した新たな言語地図の試みについて述べたい。

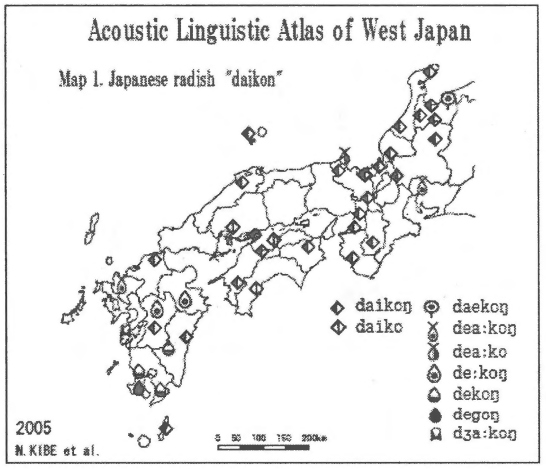
岸江・木部・石田 2002、岸江・木部・松永・福嶋 2002 では、コンピュータで作成した言語地図を、さらに Adobe 社の Acrobat を利用して pdf ファイルにすることにより、記号と実際の調査時に録音された音声データを同時に地図化する試みがなされている。

この地図は、木部 2003 において、南九州を対象に作成されてきたが、2003 年度以降、文部科学省より新たに科学研究費補助金を受けて、範囲を西日本に広げ、「西日本声の言語地図」として現在公開に向けての準備作業に入りつつある。

この地図の最大の特徴はことばの地理的分布と各地点の音声情報を同時に提示できる点である。これまでの研究では、ことばを研究するとはいっても、最終的にはその聞き取り結果は研究者個人の解釈内容であった。また、物理的に音声そのものを論文に添付することも不可能であった。木部・岸江・中井・松永・福島・鳥谷 2005 では、本研究に関わり研究者同士の個人差の他に、個人の中に「『地点ごとの聞き取り』と『項目ごとの聞き取り』の差」という興味深い報告をしている。

また、「これからの方言音声研究は、できるだけ音声データを公開し、研究者同士がデータを共有した上で、有意義な議論を行う必要がある」と述べている。これらの新たな試みは、データベースとしての意味合いだけではなく、学会における新たな研究スタイルを提示しつつある。

図5 「西日本声の言語地図」



6. おわりに

これまで、方言学や言語地理学及び社会言語学において取り組まれている、GIS を活用した言語地図作成や、pdf 形式を活用したマルチメディア言語地図データベースの取り組みと、これらの成果による新たな研究の可能性について述べてきた。

GIS については、本論の副題に「GIS ソフトの導入と展開」と付けたが、実際は未だ導入段階である。これは、ことばに関わる社会というものを、どのようなデータで明確に示すべきか、また、複合的に示すべきか否か、といったことに関して、われわれ研究者の考えが固まっていないということの証左であろう。

しかし同時に、ことばそのものが、極めて社会や文化に対して直接的に関係していることも、その展開を阻んでいると考えている。同時に、複合的に考える前に、一つひとつの事象とじっくり向き合うことが展開の始まりだと認識している。

次に、マルチメディア言語地図データベースについては、これまで、「ことば」を研究してきたわれわれにとって、その本質とでもいうべき音声を、耳というフィルターを通して文字化することにより成果を公表してきたわけであるが、捨て去られてきた(表現しきれなかった)部分も多かったと感じる。今、掛け値なしの「ことば」の研究が、情報処理技術の進展と共に始まろうとしている。

関連・参考文献

大西拓一郎(2002)「言語地図作成の電算化ー『方言文法全国地図』第五集を例にー」『日本語学』21-11,〔21-35〕,明治書院

大西拓一郎(2004a)「方言の東西境界と富山」『日本海沿岸の地域特性とことば』,〔241-266〕,桂書房

大西拓一郎(2004b)「地理情報誌システム(GIS)を利用した日本語研究」『日本語学』23-15,〔18-28〕,明治書院

ONISHI, Takuichiro (2005) 'Application of GIS technology to the studies in Japanese dialectology' Twelfth International Conference on Methods inDialectology (Université de Moncton), 〔1-10〕
荻野綱男(1975)「コンピュータによる方言地図の作成と解析」『日本方言研究会第20回発表原稿集』, 〔30-39〕, 日本方言研究会
荻野綱男(1994)『アンケート調査分析ソフトウェア GLAPSの使い方 第2版』,私家版
荻野綱男(1996)『パソコン版GLAPSの開発』,平成7年度科学研究費補助金総合研究(A)研究成果報告書
岸江信介・木部暢子・石田祐子(2002)「声の言語地図」『日本語学』, 21-11, 〔60-70〕,明治書院
岸江信介・木部暢子・松永修一・福島真司(2002)「南九州における音声言語地図の開発と作成の試み」『第74回日本方言研究会発表原稿集』,〔59-66〕, 日本方言研究会
岸江信介・中井精一・鳥谷善史(2001)『大阪府言語地図』,近畿方言研究会
岸江信介・中井精一・鳥谷善史・石田祐子(2000)「エクセルとファイルメーカープロを利用した言語地図の作製」『日本語研究センター報告』vol.8, 〔43-58〕, 大阪樟蔭女子大学日本語研究センター
木部暢子(2003)『南九州声の言語地図』平成13年度～14年度科学研究費補助金(基盤研究(C)(1))研究成果報告書
木部暢子・岸江信介・中井精一・松永修一・福島真司・鳥谷善史(2005)「方言音声研究の課題とその検討－「声の言語地図」からの提言－『第81回日本方言研究会発表原稿集』〔41-48〕,日本方言研究会
国立国語研究所(1966～1974)『日本言語地図1～6集』,大蔵省印刷局
国立国語研究所(1989～2002)『方言文法全国地図1～5集』,大蔵省印刷局(財務省印刷局)
柴田武・井口三重(1992)「コンピュータが描いた『糸魚川言語地図』をめぐって」『日本語学』11-5, 〔118-127〕, 明治書院
高橋顕志(1992)「『四国言語地図』を作る」『日本語学』11-9, 〔127-136〕, 明治書院
高橋顕志(2002a)「情報化時代の方言研究」『二一世紀の方言学』, 〔343-354〕, 日本方言研究会編,国書刊行会
高橋顕志(2002b)「ボーリング言語地図－グロットグラム(地域×年齢)分布図－」『日本語学』21-11, 〔48-59〕, 明治書院
田原広史(2002)「方言のデータベースとコンピュータ言語地図」『朝倉日本語講座10 方言』, 〔199-217〕, 朝倉書店
徳川宗賢(1993)「電子計算機の「言語地図」作成への適用」『方言地理学の展開』,〔377-389〕,ひつじ書房
徳川宗賢・山本武(1967)「電子計算機で言語地図を作る試み」計量国語学40号, 〔27-30〕,計量国語学会
鳥谷善史(2001)「語彙分布から見た大阪府の地域性－『大阪府言語地図』における風名語彙の分布から－」, 地域言語13, 〔37-50〕, 地域言語研究会
鳥谷善史(2002)「電子化言語地図の現状と問題点」『日本語研究センター報告』vol.10, 〔67-80〕, 大阪樟蔭女子大学日本語研究センター
鳥谷善史(2003)「社会言語学におけるGISの萌芽とその展開」『人文科学とGIS』, 〔29-43〕, 富山大学人文学部GIS研究会
中井精一・坂口直樹(2000)「データベースソフトによる富山県言語動態地図の作製について」『富山大学人文学部紀要』33, 〔59-70〕,富山大学人文学部
福島秩子(1983)『パソコンによる言語地図へのアプローチ SEALユーザーズマニュアル』自家版
福嶋秩子(1994)「言語地図作成システムSEAL」『日本語学』12-13, 〔106-113〕, 明治書院
福嶋秩子・福嶋祐介(2004)『パソコンによる言語地図の統合SEALユーザーズマニュアル第7版(SEAL version 7.0J for Windows) PDF版』
藤田勝良・田原広史(1990)『山口県域の方言動態小報告—通信調査の結果による概観—』私家版
前川喜久雄(1988)「PC-9801上の言語地図作成支援システム:EGL」『音声言語研究会第15回研究会発表原稿』,〔1-11〕,音声言語研究会
前川喜久雄(1993)「言語地図作成支援プログラム」『日本語学』12-4, 〔96-103〕, 明治書院
松丸真大(2003)「社会言語学におけるGISアプリケーションの活用と評価－ArcViewを用いた言語変化研究の試み－」『人文科学とGIS』, 〔44-52〕, 富山大学人文学部GIS研究会
山本武(1965)「日本言語地図の電子計算機による作成」国立国語研究所第一資料室月報
ロング,ダニエル(1990)「方言認知地図の書き方と読み方」『日本方言研究会第50回研究発表会原稿集』, 〔7-16〕, 日本方言研究会
ロング,ダニエル(1995)「第5章マルチメディア 2. 方言認知地図」『パソコン国語国文学』, 〔157-171〕, 啓文社

立地分析のための分布地図表現

－分布地図作成支援システムの構築－

Distribution Map Approach to Geographical Analyses

- Implementation of A Support System for Drawing Distribution Maps -

中島　高司・加藤　常員

Takashi Nakajima and Tsunekazu Kato
大阪電気通信大学大学院工学研究科，寝屋川市初町１８－８
Graduate School of Engineering, Osaka Electro-Communication University,
18-8 Hatuumachi, Neyagawa, Osaka 572-8530, Japan
nakajima@ktlab.osakac.ac.jp,kato@ktlab.osakac.ac.jp

あらまし：人文系の諸分野において基本的な研究活動のひとつとして分布地図の作成作業があげられる。人文系分野におけるデータの蓄積、管理はコンピュータの使用が日常になってきている。しかし、分布地図の作成に当たっては地図は電子化が進んでいるものの、安価で容易に利用できるアプリケーションソフトはないのが現状である。我々はこれまでに分布地図の作成支援を目的としたシステムの開発を進めてきた。本稿では構築した分布地図作成支援システムの開発した機能の概要について述べ、弥生集落遺跡のデータを用いて作成した分布地図による遺跡の立地分析の試みについて報告する。

Summary: Drawing a distribution map is one of basic activities for every study of the humanities. Computer is recently used to store or to manage data in the humanities. Today, since many kinds of maps have been digitalized, we can easily treat maps by computer, combining them with other types of data. In this paper, implementation of a support system for drawing distribution maps is presented and discussed. In addition, we present an archaeokogical study on a distribution map of a special type of sites, which is drawn by our system. Good understanding of geographical relation between the sites is obtaind.

キーワード：分布地図，GIS，支援システム，立地分析，可視範囲
Keyword: distribution Map, GIS, support system, geographical location, viewshed analysis

１．はじめに

近年コンピュータの機能性、操作性の向上、アプリケーションソフトの充実は、考古学、歴史学の研究者においても日常的に計算機を扱い、多くの研究者が独自なデータベースを容易に作成できる環境を生み出している。また、地理情報は電子化された多種多様なものが安価で提供されるようになった。分布地図の作成過程も従前の紙媒体上の作業から電子化された地図を利用した計算機上での操作となってきている。

コンピュータを利用した分布地図作成の方法のひとつとして GIS ソフトの利用が考えられるが、研究者が従前から蓄えてきたデータがそのまま利用できないといったことや使用できる地図が限られてしまうなど研究用のツールとしては使用しづらいのが現状である[1,2]。このため分布地図作成でもっとも労力を費やす対象点のプロット作業は、依然として一点一点を指定する操作が手作業で行われている。

本研究では研究者が従前から蓄えてきたデータ形式を極力変更することなく利用できる分布地図作成支援システムの開発を進めてきた[3]。本稿では本システムの現状の報告を行い、立地状況の分析に関する１，２の特化した処理機能を紹介する。開発した特化した処理は対象点周囲の可視性を判定するもので、従前に行った弥生時代の遺跡間の水平的な地理的分布分析の研究成果[4]を踏まえた活用事例を示し考察を行う。

２．既存データとその活用

紙媒体を基本とした人文系の資料も様々な局面で電子化が行われるようになり、多くの研究者がワープロや表集計ソフトを活用し独自のデータベースを作成している。提案する分布地図作成支援システムでは、個々に作成された既存データベースの活用を重要な視座に据えている。研究者が作成したデータを損なうことなく、負担や抵抗の少ない作業で既存データベースを活用するシステムをめざした。

（１）活用可能なデータベースの条件

既存データベースの活用にあたり、つぎのふたつの事項を条件とした。

ひとつは、ファイル形式は CSV 形式とした。これはファイル形式として一般的であり、多少異なっても研究者が使い慣れている Excel 等で簡単に変換が可能ためである。ふたつめには、緯度・経度がデータ項目に含まれている事が必須条件である。緯度・経度の項目の位置や表現形式は多岐にわたることが想定されるが項目の位置を本システム側で既存データ指定時に設定することで対応する。

（２）既存データベース活用に必要な機能

既存データベースを加工することなく活用するために以下の３つの機能が必要と考えられる。

① 既存データベースの情報保持機能

既存データファイルの緯度・経度のフィールド位置およびその表現形式を取得、指定する。得られた情報を参照ファイルとして保存する機能を設ける。

② 分布表示対象の複数指定機能

分布地図には、複数の対象群が異なった色やマークで示されることがしばしばある。ひとつまたは複数の既存データベースより対象のデータ(レコード)を選択し、１つのグループにする機能とそうしたグループを複数同時に扱える機能の設定が要求される。

③ 再編集可能な描画機能

一度作成した分布地図の再編集が必要となる場面は研究活動のなかで多々起こる。再編集の実現のためには、作成した分布地図の情報をファイル化し、保存する機能と保存されている情報から再描画する機能が必要となる。

３．分布地図作成支援システムの構成

分布地図作成支援システムは、図１に示す、既存ファイル設定モード、地図設定モード、分布地図編集モードおよび特殊処理モードの４つのモードを指針に構成する。

（１）既存ファイル設定モード

既存ファイル設定モードは、ファイル情報取

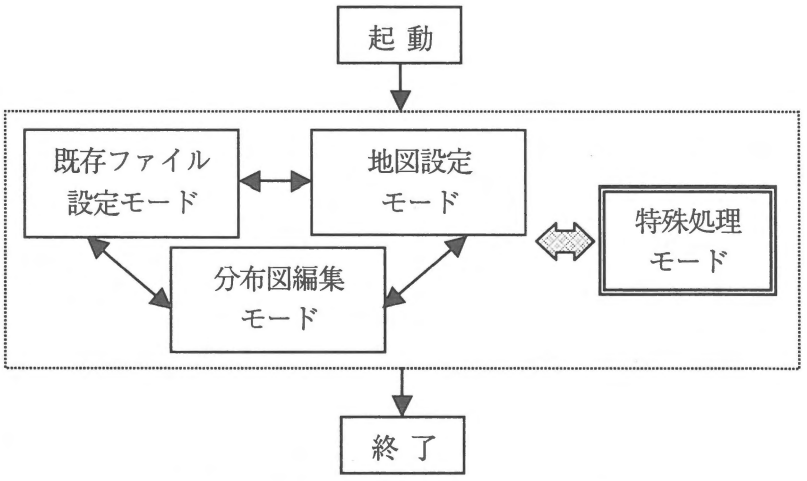


図１　分布地図作成支援システムの構成

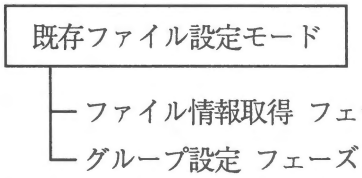


図2　既存ファイル設定モードの構成

得フェーズとグループ設定フェーズから成る(図2参照)。

① ファイル情報取得フェーズ

ファイル情報取得フェーズは、ユーザが独自に作成したデータファイルを取り込むモードである。画面上の [データファイルを開く]ボタンをクリックするとファイル選択画面ダイアログボックスが表示され、ファイルを選択する。すると自動的にExcelがコールされ、選択したファイルついての参照ファイルが既に作成済みであれば、即、グループ設定フェーズに移行する。また、なければ緯度、経度項目位置を順次、Excel画面上で「列」を指定することで取得し、緯度・経度の保存形式は、保存形式のサンプルを示したリストボックスで選択する。その内容は参照ファイルとして保存される。

② グループ設定フェーズ

このフェーズは、Excel 画面上で複数行を指定し、分布のひとつのグループとし、グループ名と表示色を指定する。ひとつのファイルに対し、複数のグループの設定を可能とし、グループ指定に当たっては、ソート等のExcelの機能を使用した結果に対して指定できる。指定後、描画ファイルとして保存される。

（2）地図設定モード

このモードには、地図種設定フェーズと、地図領域設定フェーズがある(図3参照)。

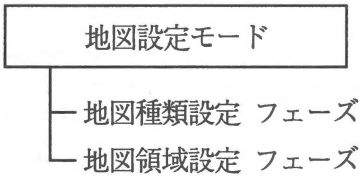


図3　地図設定設定モードの構成

① 地図種類設定フェーズ

分布地図の背景となる地図の種類をチェックボックスのチェックにより指定するフェーズである。指定できる地図は、海岸線、湖岸線、河川の3種類の水系データと、標高データをもとに一定標高幅ごとに色づけした彩色標高地図、等高線地図および行政界地図である。色彩標高地図に関しては基準色、広域、狭域および2色表示の四つの配色パターンを準備した。これらの地図は、任意に組み合わせることでき、多様な地図表示が可能となっている。複数の地図種をレイヤー構造にすることにより、地図種の選択操作、表示速度の向上を図る。

② 地図領域設定フェーズ

地図領域設定フェーズでは、地図の描画範囲を設定する。設定には3つの方法を準備した。ひとつめの方法は、日本全国の任意の範囲を選択する場合を想定し、縮小表示した全国図上でマウスドラッグにより範囲を指定する。ふたつめの方法は、左下緯度・左下経度、右上緯度・右上経度の値をテキストボックスに直接入力で指定する。3つめは、この設定操作に先立ち描画対象の分布情報(ファイル)が指定されているとき、指定されている緯度経度情報をもとに20万分の1地勢図の相当する範囲を設定する。

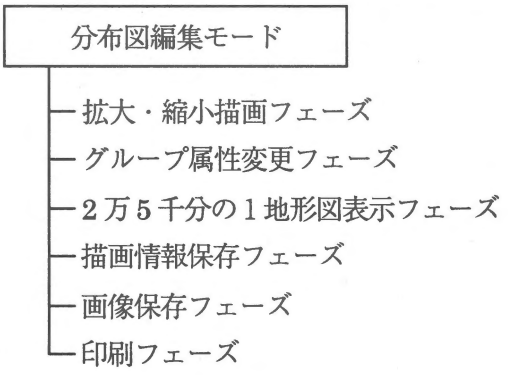


図4　地図設定設定モードの構成

（3）分布地図編集モード

分布地図編集モードでは、拡大・縮小描画フェーズ、グループ属性変更フェーズ、2万5千分の1地形図表示フェーズ、描画情報保存フェーズ、画像保存フェーズ、印刷フェーズの6つのフェーズから構成される(図4参照)。

① 拡大・縮小描画フェーズ

このフェーズは、表示している地図を拡大・縮小する機能がある。拡大では、描画範囲を狭め、狭めた範囲を拡大描画する。拡大描画には、表示範囲の中心を変えずに単純2分の1の領域を拡大描画する方法、指定した分布点を中心に

一定領域の広さで拡大描画する方法、マウスドラックの指定により任意の範囲を拡大描画する方法の３通り用意する。縮小は、描画範囲を広め、広めた範囲の縮小描画を行う。縮小描画では、描画されている２倍領域を描けるように縮小を行う。描画範囲内で表示される分布点は、自動的に制御される。縮小描画の場合、描画範囲が拡大されるが編集当初の分布対象点以上に増加することはない。

② グループ属性変更フェーズ

分布地図編集を進める上で、グループ名や表示色の変更が必要になることが考えられる。そうしたグループ属性の変更を行うフェーズを準備する。操作は、グループ名の上をマウスクリックすることで入力欄が表示される単純な方法を採用する。また分布点の色に関しても色変更ボタンをクリックする事で再指定可能とする。

③ 万５千分の１地形図表示フェーズ

表示されている分布地図に示されている分布点(マーカー)を含む２万５千分の１地形図画像が表示する。このとき、白地図上に表示されていたすべての分布点は、表示地形図上にも表示する。この地形図の表示は、分布点周囲の詳細な現地形の情報が確認でき、研究者に有益な付加情報になると考えられる。

④ 画像情報保存フェーズ

①から③の操作を適宜繰り返し、得られた分布地図を描画ファイルとして保存する処理である。保存では、再描画、再編集が可能となる情報をファイル化する。保存する項目は描画範囲(左下緯度・経度、右上緯度・経度)、グループの名前・色、各グループの分布点の緯度・経度である。

⑤ 画像保存フェーズ

作成した分布地図の他のアプリケーションでの２次利用を目的にビットマップ形式で画像として保存する。作成中のどの時点でも保存可能とする。

⑥ 印刷フェーズ

作成した分布地図をいくつかの付加情報(テキスト)とともにプリントアウトする機能を設定する。付加情報しては、分布地図領域の左下緯度・経度、右上緯度・経度、グループ数、各グループの名称、色、分布点数および描画ファイル名が印字される。

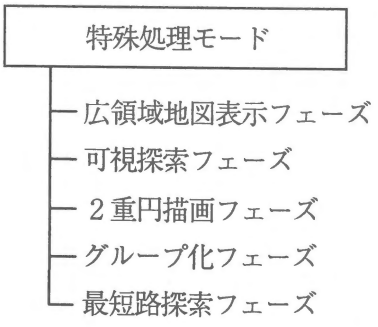


図５　特殊処理モードの構成

（４）特殊処理モード

特殊処理モードは、一般的、共通的な分布地図作成処理とは異なり、特定の主題に特化した処理群を設ける。本稿で挙げる処理群は、２種類遺跡群間の分布立地を分析するツールである(図５参照)。

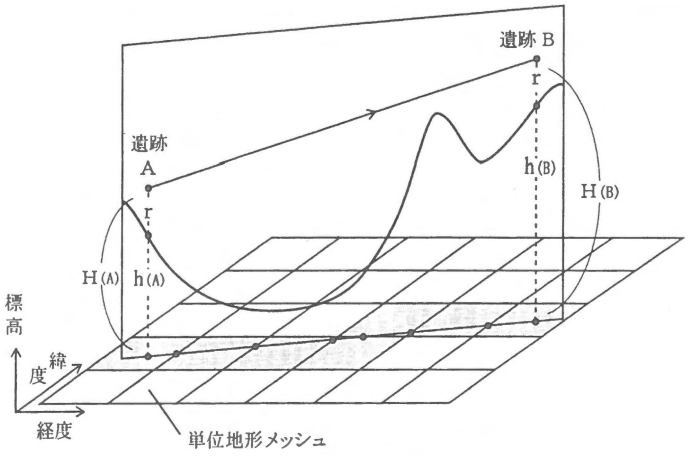


図６　可視性判定の概念図

① 広領域地図表示フェーズ

対象が広い領域に分布してい場合、限られた画面上の表示には、おのずと限界があると思われる。分析においては大縮尺で対象のすべてが地図上に表示される事が必要と考えられる。分布の観察に有効な広領域の地図画像を別フォームに生成しスクロールにより連続的な観察を可能にし全域の状況を把握する。表示された広領域地図上で任意の範囲をマウスドラックすると編集モードの拡大・縮小フ

ェーズと連動し、元のフォーム上に拡大または縮小して指定範囲の地図が描かれる。

② 可視探索フェーズ

任意の地点からの見通し関係を判定するする処理系である。基本は３次元地形データ(標高データ)によりコンピュータ内に生成した地形空間上で２点間の地形的障害の有無を判定する(図６参照)。中心地点(遺跡位置など)を定め、一定距離内のすべて地点(円領域)に対して行い、可視領域を探索する。標高データは、国土地理院が提供している数値地図 50m メッシュ(標高)を用いる。このデータは緯度・経度区画された約50m間隔の格子位置での標高値が10cm単位で与えられている。また、探索に当たっては可視性の判定を制御するパラメータとして可視距離と標高調整量を導入する。可視距離は定常的に見通せる距離の上限を定めるパラメータであり、拝呈を行う円領域の半径となる。一方、標高調整量は標高データの量子化誤差などいくつかの不確定要因を踏まえた上で、有効な可視可能性判定を行うためパラメータである。

このフェーズは対象すべてに対して一括でも一対象に対しても実行可能で、結果は①で述べた広領域表示、元のフォーム上のいずれにも表示できる。

③ グループ化フェーズ

近接する複数の対象を１グループにまとめ１対象にする処理である。一定の距離内につながる対象を１グループとしその一群の重心を代表点とする。データに偏りや時期的変遷を集約する場合などに有効と思われる。

④ 最短路探索フェーズ

対象間に適当な距離を設定し、その距離にもとづく最短路網を探索する。この処理は遺跡間の交流、親疎などの関係を検討するに有効と考えられる[5]。

⑤ ２重円描画フェーズ

注視する対象を中心に２つ円を各々設定した半径で描く機能である。極めて単純であるが対象間の位置関係などを示すのに有効な手法である[4]。

４．分布図作成支援システムの活用事例

３章で述べた各モードを実装した分布地図作成支援システムを構築した。ここでは弥生時代の２種類の異なった遺跡群(拠点集落遺跡および高地性集落遺跡)を採り上げ、地形的な立地関係の分析に活用した事例により、構築したシステムの操作を述べるとともに分析結果について検討、考察を行う。

４.１ 拠点集落遺跡と高地性集落遺跡

弥生時代の集落は水稲農耕と密接に関連した低台地上や沖積地の自然堤防上の微高地に営まれた。拠点集落遺跡とは、各地域の中核となる大規模で継続的に営まれた集落跡である[6]。拠点集落はその周囲に派生した多様な諸集落、諸施設を持つ。西日本では、弥生前期から後期に至る長期継続の拠点集落遺跡が多く検出されているが、東日本では極めて少ない。

一方、高地性集落遺跡は、弥生時代中期から後期にかけて西日本を中心に水稲農耕には不向きな標高１００ｍときには３００ｍを超える高所から検出される集落跡である。弥生時代中期から後期は、中国の史書が「倭国大乱」として伝える軍事的緊張があった時代にあたる。高地性集落遺跡は、こうした立地的特異性と時代的背景のもと、軍事用のノロシ通信施設であったとする考古学上の仮説がある[7]。

拠点集落と高地性集落は、地域ともに重なるが時期的には拠点集落が廃絶されると相前後して高地性集落が急増することが知られている。

４.２ 立地分析

拠点集落遺跡と高地性集落遺跡は、上述したように同時には営われなかったが、その平面的位置関係に関しては、先に行った研究により興味深い結果を得た。地形的制約が緩い領域では両者は 2.20km～2.75km 離れて存在することが示された。本稿ではこの結果を踏まえた両者の垂直的な分布分析の試みを分布図作成支援システムの活用事例として示す。

拠点集落遺跡データおよび高地性集落データは、異なったフォーマットの CSV 形式のファイルで各々３８１件と５８１件のレコードが登録されている。

(1) 既存ファイル情報取得操作事例

図７は、既存ファイル(拠点集落遺跡データファイルまたは高地性集落データファイル)を選択し、緯度・経度の項目(フィールド)位置を Excel の操作画面で指定、登録操作画面である。

緯度項目の列をマウスクリックするとExcel の列指定確認画面が表示されOK ボタンをクリックすることで登録シート(右フォーム)に登録され、登録シートの他の項目も設定し、設定終了ボタンをクリックすることで参照ファイルが作成される。

（2）クループ設定操作事例

図８は、分布対象のグループを設定する画面である。既存ファイルが展開されたExcel 画面と設定項目シートが表示される。設定項目シートのグループ番号および分布点の色を設定する。

次に必要な範囲(行)を Excel 上でマウスドラッグして選択し、Excel の行指定確認画面が表示されOK ボタンをクリックするとことで対象が選択される。保存ボタンをクリックすることで仮の描画ファイルの作成される。

（3）地図編集操作事例

図９は全国図表示枠より任意の範囲をマウスドラッグによって指定し、地図設定操作から彩色標高地図を選択し、描画した地図上に畿内に分布する高地性集落遺跡を５つのグループに分け描画した例である。図の右下シートには、それぞれ設定したグループ名と、分布点の数が表示さる。このグループ名の上をマウスクリックすると、名前変更欄が表示されるので、ここでグループ名の変更も可能となっている。また、地図の上をマウスドラッグすることで図９の地図内に四角枠の表示され、拡大ボタンをマウスクリックすると、選択範囲を拡大表示できる。

（4）２万５千分の１地形図表示操作

図８において任意の位置をクリックし、右中のタグシート操作の地形図呼び出しボタンをクリックするとクリック位置を中心とした２万５千分の１地形図が図１０のように表示される。表示範囲はクリックした近辺であるがスクロールバーを操作することで当該の地形図の全域を参照できる。

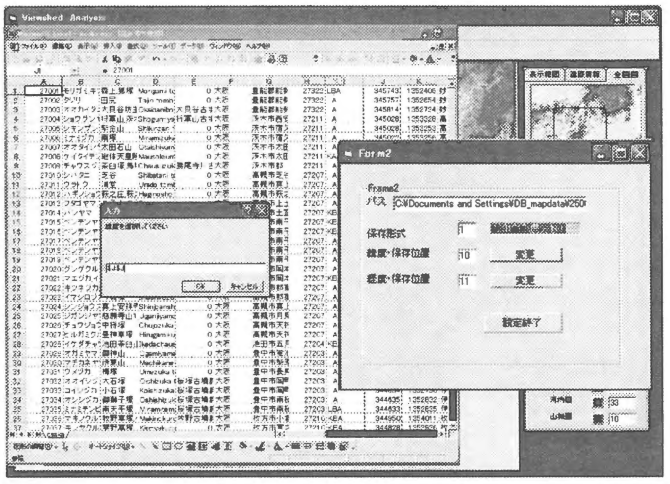


図７　既存ファイル情報取得操作事例

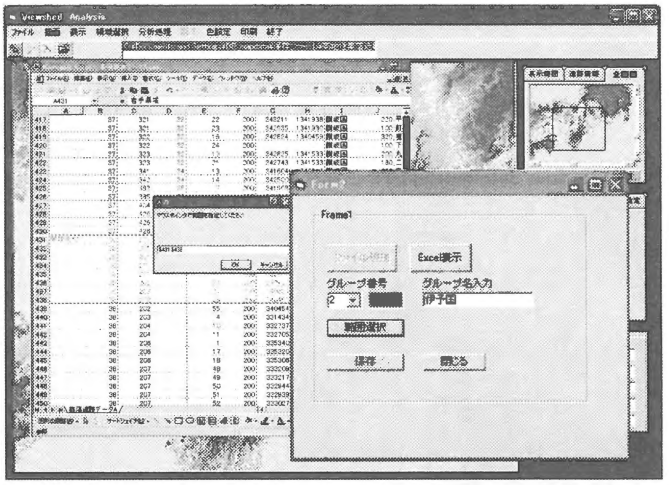


図８　グループ設定操作事例

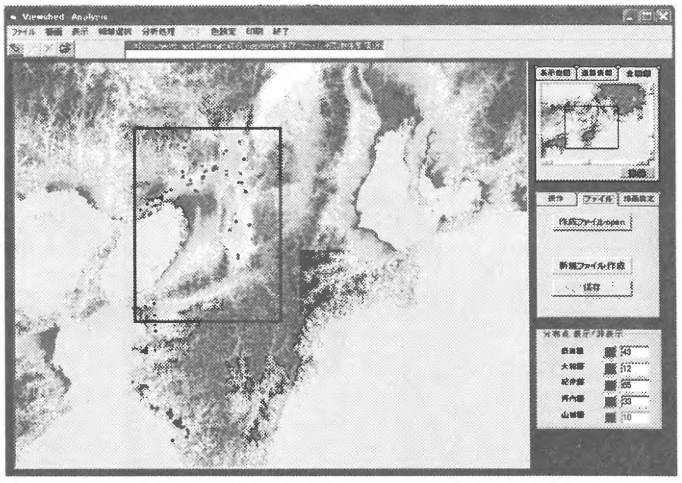


図９　地図編集操作事例

（5）特殊処理操作（立地分析処理）

メニューバーの特殊処理をマウスクリックすると特殊処理選択フォームが表れる。

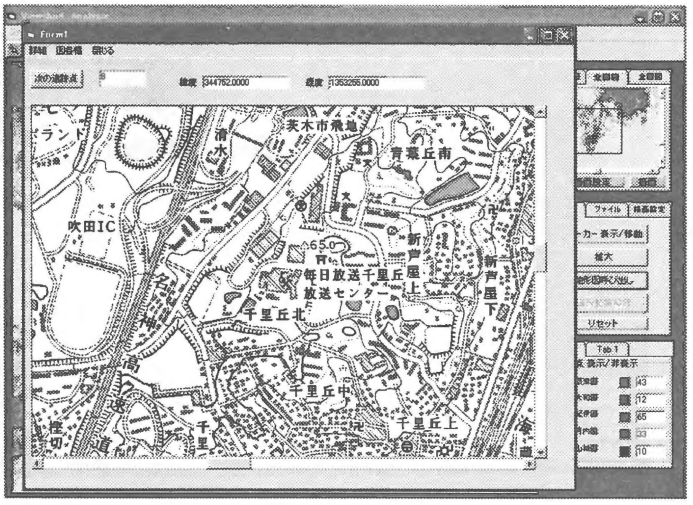


図１０　２万５千分の１地形図表示操作

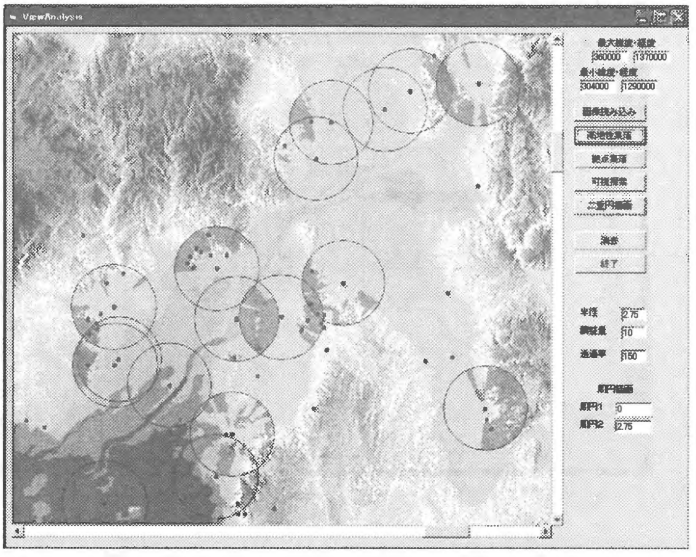


図１１　分析操作

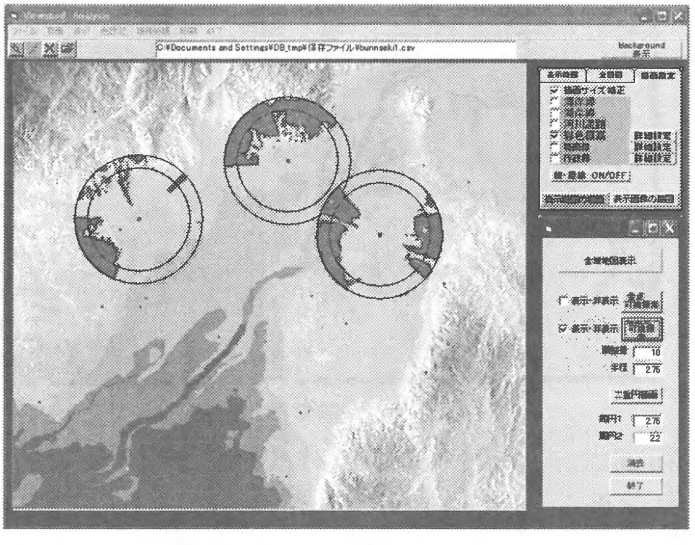


図１２　分析操作

① 広領域地図表示操作

特殊処理選択フォームから広領域地図表示選択すると広領域分析用フォームが表示される。表示されたフォーム上の画像読み込みボタンをマウスクリックすると広領域地図画像の読み込まれ、広領域地図が表示される(図１１参照)。スクロールバーにより連続的に全国土をスクロールすることができる。この時点で(1)および(2)で指定したファイルの対象データがグループごとに色分けされ分布点として表示されている。

② 可視探索操作

特殊処理選択メニューまたは広領域分析用フォームの可視探索ボタンをマウスクリックすると可視判定の基本となる可視探索ファイルを生成する。すでに可視探索ファイルが生成済みの場合は再生成しない。

可視探索ファイルの生成終了後(生成済み場合は即)、テキストボックスで指定した２つパラメータ、可視距離および標高調整量の値に応じた可視判定結果が図１１のように表示されている。パラメータ値を変えて再度、可視探索ボタンをマウスクリックするとそれら値に応じた結果が再描画される。図１１は各拠点集落遺跡を中心に円を探索し不可視領域を塗り潰している。また、この操作は可視探索ファイルが有る場合、一旦、可視領域を消去した後(消去ボタンをクッリク)に表示されている領域内の対象遺跡をマウスクリックするとその遺跡の分析結果のみの表示も可能である(図１２参照)。この画面で任意の領域をマウスドラックで指定すると、その領域を元のフォーム上に表示可能な大きさに拡大または縮小し表示する。この元のフォーム側においても先の可視探索が可能で詳細な関係が観測できる(図１２参照)。

③ ２重円描画操作

広領域分析用フォームで大小２つの円の半径を設定し、可視探索と同様にマウスクリックで遺跡をした後、フォーム上の２重円描画ボタンをクリックすることで、当該の遺跡位置に２重の円を描画する。

図１２は、詳細化した地図上で３遺跡をマウスクリックで指定し、先の研究結果にもとづき可視距離２.７５km、標高調整１０mで可視探索を行い、さらに内円半径２.２０ｋm、外円半径２.７５ｋmで描画したものである。

（６）印刷操作事例

メニューバーの印刷をマウスクリックすることで（１）から（５）の操作等を行って画面に表示されている結果を印刷する。図１３は、図１２の立地分析の結果を詳細化した彩色標高地図を印刷出力である。印刷ではＡ４用紙に分布地図と必要と思われる諸情報が印字さる。

５．立地分析結果の検討、考察

図１２は、拠点集落遺跡と高地性集落遺跡との立地関係を分析した結果図である。本分析を進めるに当たり、参考にした先行研究では両遺跡間水平距離２.５km前後を導いているが具体的な位置の地形状況については、「地形が影響する」との示唆に留まっている。図１２は、この示唆に対して具体的にどうような関係の所に所在するかを明確に示していると言える。すなわち、拠点集落の可視領域を図示した結果、可視領域と不可視領域の境界線上に高地性集落遺跡の分布点が見られ、拠点集落と高地性集落とは同時期に共存しないが、高地性集落と平地の集落の位置関係を推察するにあたり可視性がひとつの視点となると考察される。この両遺跡間の関係については専門分野から検討、評価が必要と考えている。

６．おわりに

開発を進めている分布地図作成支援システムおよび特殊処理機能について述べ、具体的なデータを用いた処理事例を示した。事例は可視領域分析で遺跡相互間の立地分析を試みた。興味深い結果が得られたと考えている。今後、専門的立場の方の意見を聞き、より有効な分析処理機能の付加を行っていく必要があると感じている。また、システム開発においては処理速度の向上、扱いやすいユーザインタフェースの構築等、改善していく予定である。

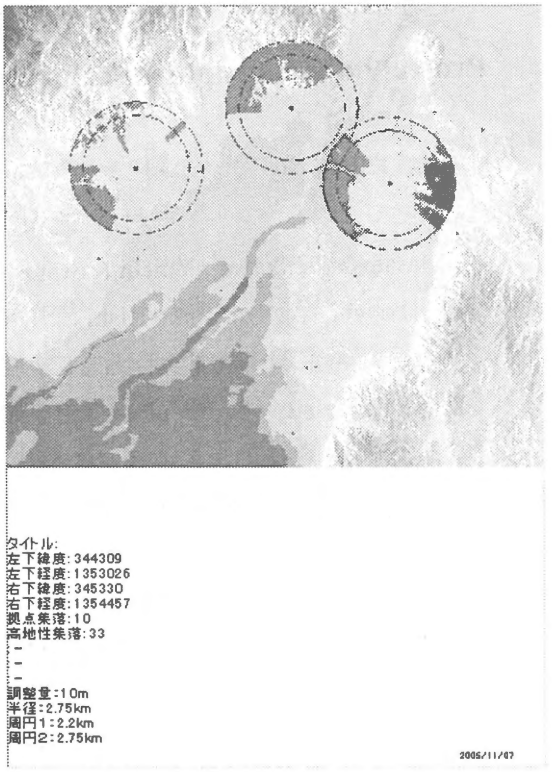


図１３　印刷結果表示

本研究を進めるに当たり、日頃ご討論、ご援助を賜る、大阪電気通信大学小澤一雅教授に深謝いたします。また、本研究の一部は、日本学術振興会科学研究費補助金（基盤研究（C）（２）No.１５５００１５９)によった。

参考文献

（１）杉本智彦：カシミール３Ｄ　パーフェクトマスター編、p.255、実業之日本社、東京(2003).
（２）㈱ゼンリン：ZPROFESSIONAL４取扱説明書、p.116、㈱ゼンリン、東京(2005).
（３）中島高司、加藤常員：分布地図作成支援システムの構築、情報処理学会研究報告,Vol.2005,No.51,pp.1-8(2005).
（４）Tsunekazu Kato and Kazumasa Ozawa: Distribution analysis of sites based on GIS, GIS Development,Vol.8,No.6,pp.22-25 (2004).
（５）加藤常員、小澤一雅：集落遺跡間の文物移動流のモデル化に関する一考察、情報処理学会論文誌、Vol.４０、No.３、pp.849-856(1999).
（６）酒井龍一：拠点集落と弥生社会,歴史発掘⑥,pp.118-139,講談社,東京(1997).
（７）森岡秀人：高地性集落性格論,論争・学説日本の考古学,4,雄山閣(1986).

人文科学のための地理情報・資源共有化システム・プロジェクト

Project on Geographical Information and Resource Sharing System for Humanities

石川 正敏[†], 川西 陽一[††], 奥村 英史[†††], 原 正一郎[††††], 桶谷 猪久夫[†††††], 貴志 俊彦[††††††], 村尾 義和[††††††], 柴山 守[†††††††]

Masatoshi Ishikawa[†], Yoichi Kawanishi[††], Hidefumi Okumura[†††], Shoichiro Hara[††††], Ikuo Oketani[†††††], Toshihiko Kishi[††††††], Yoshikazu Murao[††††††], Mamoru Shibayama[†††††††]

[†]東京農工大学大学院 工学教育部, 東京都小金井市中町 2-24-16

[†]Tokyo University of Agriculture and Technology, 2-24-16, Naka-cho, Koganei-shi, Tokyo

[††]京都大学大学院 アジア・アフリカ地域研究研究科, 京都市左京区吉田下阿達町 46

[††]Kyoto University, 46, Yoshida, Shimoadachi-cho, Sakyo, Kyoto

[†††]ヒューマン オーク, 京都府長岡京市奥海印寺太鼓山 22-113

[†††]Human ook, 22-113 Taikoyama, Okukainji, Nagaokakyo-city, Kyoto

[††††]国文学研究資料館 複合領域研究系, 東京都品川区豊町 1-16-10

[††††]National Institutes for the Humanities, 1-16-10 Yutaka-cho, Shinagawa-ku, Tokyo

[†††††]大阪国際大学 人間科学部, 大阪府守口市藤田町 6-21-57

[†††††]Osaka International University, 6-21-57, Fujita-cho, Moriguchi-shi, Osaka

[††††††]島根県立大学 総合政策学部, 島根県浜田市野原町 2433-2

[††††††]The University of Shimane, 2433-2, Nobara-cho, Hamada-shi, Shimane

[†††††††]京都大学 東南アジア研究所, 京都市左京区吉田下阿達町 46

[†††††††]Kyoto University, 46, Yoshida, Shimoadachi-cho, Sakyo, Kyoto

あらまし：本稿では，人文科学研究における時空間情報の共有基盤の構築を目的とした，HGIS 研究会 (WHGIS, Working group on Humanity Geographical Information Science) 研究会の各著者の研究成果について報告する．さらに，本研究会で構築を目指す地理情報共有システムについて検討する．所期の地理情報共有システムによって研究者は古代から現在まで空間情報をシームレスに閲覧し再利用することが可能となり，新たな知見に結びつくことが期待できる．

Summary : This paper reports the activities of our Working group on the Humanities Geographical Information Sciences (WHGIS). Our purpose is the construction of the common base for dealing with spatiotemporal resources on the Humanities. Furthermore, we suggest a Geographical Information Sharing System (GISS) for the Humanities. This GISS enables researchers to browse seamlessly spatiotemporal resources. Therefore we hope that the GISS is useful to discover new knowledge on the Humanities.

キーワード：時空間リソース, データベース, 地理情報システム, 情報共有, 地名辞書

Keywords : spatiotemporal resources, database, geographical Information system, information sharing, gazetteer,

1．はじめに

歴史学，考古学，地域研究のような人文科学研究分野では，研究対象と地域要因，時代背景との関りを把握することは非常に重要である．また，その関りに着目することで，新たな研究視点の創造に結びつくこともある．このような研究活動で利用される古地図や衛星写真，歴史史料などの資料の多くは，時間と位置に関する情報を含む．デ

ータの分類や偏りを利用者が理解し易い形で示すために情報の視覚化は有効であり，時間と位置を持つデータの視覚化には，地理情報システム(GIS)が広く利用される．GISより，古地図や衛星写真などの地理情報や，歴史史料などのさまざまな情報について，地図を媒介としてそれらの関連を視覚化することができる．さらに，様々な研究者が協調して研究するための効果的な環境を提供するためには，以下の項目に留意する必要がある[1]．(1) すべての事件，オブジェクトなどを地図に射影し，このとき3次元空間と時間を合わせた4次元空間として表現される．(2) インターネットを介した分散環境上で，研究者は，それぞれの研究テーマに関連した地理情報や関連情報を動的に選択できる．(3) GPSによる位置情報を含む注釈，写真などのデータベースへの登録と同時に，他の研究者による閲覧や編集が可能である．(4) 研究活動のための地名辞書，地図，衛星写真などのリソースの分析や再利用が可能である．

そこで，HGIS研究会(WHGIS，Working group on Humanities Geographical Information Science)は，これらの要求を満たす，新しい共同研究および情報共有環境の構築と実証を目指している．本稿では，各著者の研究成果について報告し，本研究会で構築を目指している地理情報システムの構想について述べる．

本稿の構成は，次の通りである．2章では，歴史史料や人文科学研究の書誌情報に関するデータベースである「北東アジア地域の社会科学のための資料・書誌情報データベース」について報告する．3章では，外字を含む歴史史料の電子化のための漢字フォントである「e漢字」のデータベースについて報告する．4章では，「古代・中世の日本の全地震史料」データベースにおけるGIS利用の効果について報告する．5章では，人文科学研究におけるGIS利用のために必要な地名辞書の構築について報告する．6章では，HGIS研究会で構築している地理情報共有システムについて述べ，7章で，これまでの議論を踏まえた人文科学のためのGISについて考察する．最後にまとめと今後の課題を述べる．

2．北東アジア地域の社会科学研究のための資料・書誌情報データベース

北東アジア・データベース研究会(研究代表，貴志俊彦，島根県立大学)では，様々な組織で個別に管理されている北東アジア地域についての資料を，横断的に活用できるように，書誌情報や年表情報をインターネット上で公開している．本章では，「北東アジア地域の社会学研究のための資料・書誌情報データベース(NEARDB)」の資料・書誌データベース[2]と戦前期東アジア絵はがきデータベース[3]について述べる．

2．1 資料・書誌データベース

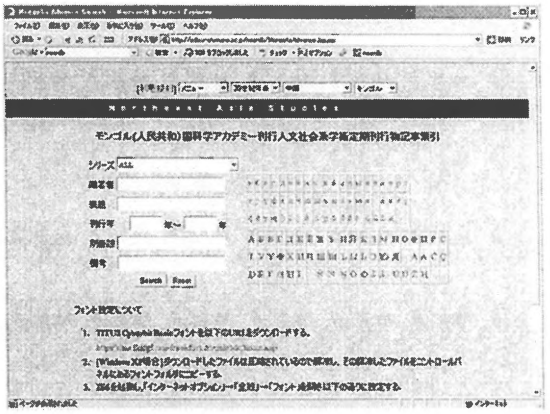


図 1 NEARDBの実行例 (モンゴル(人民共和)国科学アカデミー刊行人文社会系学術定期刊行物記事索引 詳細検索画面)

「北東アジア地域の社会学研究のための資料・書誌情報データベース」(NEARDB, http://gdb.u-shimane.ac.jp/neardb/)[2] は，中国，モンゴル，日本に関する資料の共有を支援するための多言語データベースである (図1). 本データベースは，2004年4月から島根県立大学で公開され，2005年9月までに約5000回の閲覧がある．本データベースは，以下のようなデータベース群からなる．

(1) 20世紀前半年表データベース
(2) スタンフォード大学フーヴァー研究所中国関係アーカイブ
(3) 『（北京特別市公署）市政公報』（1938年～1944年）件名データベース
(4) 上海租界工部局警務処文書（Shanghai Municipal Police Files）データベース
(5) 戦前期天津史文献目録データベース（邦文編）
(6) モンゴル(人民共和）国科学アカデミー刊行人文社会系学術定期刊行物記事索引

これらのデータベースに対して利用者は，単純なキーワードだけを用いた単純検索と，著者や表題のような各データベースのデータ構造に即した詳細検索が可能である．さらに，利用者が選択した複数のデータベースに対する横断検索も可能であり，新たな資料の関連の発見が期待される．

さらにNEARDBでは，一般的な日本語入力環境である，MS-IMEでは入力が困難なモンゴル語入力に対して JavaScript を用いた支援ツールを提供している．

2．2 戦前期東アジア絵はがきデータベース

図 2 戦前期東アジア絵はがきデータベース

過去，様々な地域で作成された絵葉書は，作成された当時の地域特色を知る上で重要な資料であるといえる．「戦前期東アジア絵はがきデータベース」は，戦前の日本，中国，韓国に関する絵葉書2134点を電子化し公開している(図2)[3]．このデータベースの検索には，キーワードに加えて地図や時間軸を用いた検索が可能である．このように地図や時間と結びつけて絵葉書を閲覧することによって，利用者は，当時の情景に関する分析が容易になると考えられる．

3．e漢字

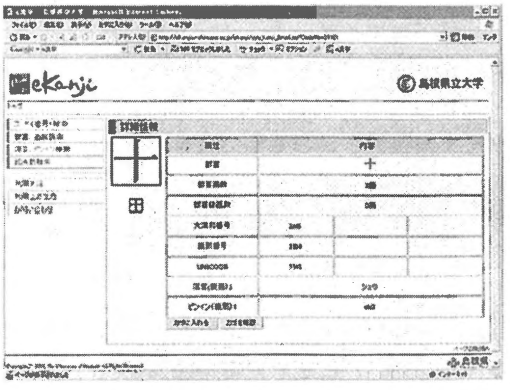


図 3 e漢字閲覧例

e漢字は，故勝村哲也名誉教授 (島根県立大学) の研究成果として作成された漢字フォント，および辞書コードやピンインなどの漢字属性データ集である．島根県立大学では，漢字文献の電子化支援を目的にe漢字をデータベース化し，インターネット上で公開している[4]．本データベースの実行例として，漢字「十」の閲覧例を図3に示す．e漢字は，諸橋大漢和辞典，京大康熙字典，Unicode2.0，中華字海に収録されている約96000字の漢字を収録している．本データベースでは，ビットマップデータであったe漢字フォントを24ポイント相当の大きさのGIF画像に変換して公開している．利用者は，フォント画像を自由にダウンロードしてWebページやMS-Word文書に貼り付けることができる．利用者が目的の漢字フォントを選択するために，本データベースでは，それぞれの辞書コードに基づいた検索と，部首や画数，読みなどの漢字属性に基づいた検索が可能である．

e漢字によって，外字を含む東アジア圏の歴史史料の多くの電子化が可能となり，効率的な資料の共有が実現できる．また，e漢字は，漢字に関する属性情報を集約しているため，漢字自体の研究に有益であると考えられる．

4．「古代・中世の日本の全地震史料」データベース

「古代・中世の日本の全地震史料」データベースは，歴史的な地震や災害に関する基礎資料と考えられている「増訂大日本地震史料」(武者，1941編，1943編）と「新収日本地震史料」（東京大学地震研究所編，1982年～1994年）を電子化し，データベース化したものである[5]．本データベースによって，情報量の多い基礎資料から必要な情報だけを容易に取り出せるため，効率的な地震分布や近年の地震との比較などが可能になる．

データベースの構築に当たり，基礎資料をXML文書として電子化した．また，基礎資料の電子化では，レイアウトや外字に関する情報も含む文書に関するすべての情報を電子化するためのDTDを定義した．本データベースは，検索に対して，レイアウトに関するXMLタグ情報に基づいて，元の基礎資料のレイアウトをWWW上で再現している．さらに，基礎資料にある重要な地震を対象として時空間データベースを構築した．ただし地震の観測点 (緯度経度) を基礎資料にある地名から推定した．図4は，Timemap[6]を用いて作成した時空間データベースを地図化した例である．このような地図化によって，歴史的な地震データの中から信頼性の高い，または低いデータを容易に選別ができる．

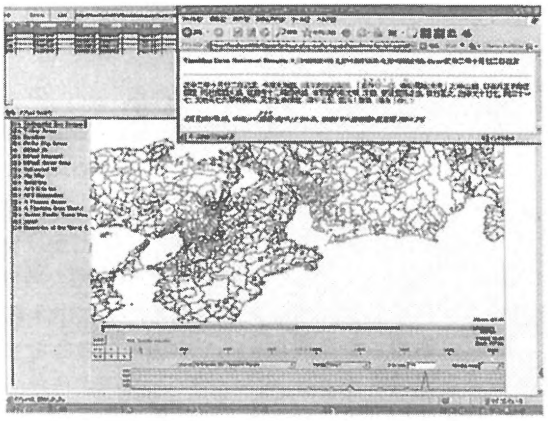


図 4 Timamap を用いた古代地震の表示例

このことから，地震学，火山学，歴史学など多くの分野の研究者の協力によって構築された本データベースの効率的な共有に GIS が有効であることを示していると考えられる．ただし，Timemap は，日本語や中国語などの多バイト文字で記述されたデータに対する多言語処理の機能がない．また，利用者による地理情報の編集機能も備えていない．そこで，今後，基礎資料のような時空間情報を人文科学における共同研究で効率的に利用するためのインタフェースなどが必要であると考えられる．

5．人文科学 GIS のための地名辞書の構築

人文科学において GIS の利用が普及しているが，国外に比べ人文科学研究における GIS 利用のためのツールが十分に整備されているとはいえない．そこで，本研究では，日本国内の歴史的な地名に関するデータベースを構築している[7]．本データベースは，大日本地名辞書(吉田東伍著，明治 33 年刊行)に記載されている 53676 ある地名の読みとローマ字表記および，地名に対応した緯度経度の座標値を入力している．今後は，大日本地名辞書にある地名と現代の地名の対応づけを行い，時間的推移を視野に入れた調査の効率化を目指す．さらにさらに本研究では，延喜式に記載されている 2861 の神社の名前とその位置情報との関係を示す辞書の構築を行っている．

このような地名辞書を構築することによって，年代や分野の異なる人文科学研究者間の地理情報共有や共同研究が可能になり，新たな知識の発見が期待できる．

6．地理情報共有システム プロジェクト

本章では，歴史学や考古学における共同研究の支援のために，HGIS 研究会が開発している地理情報共有システムについて述べる[8][9][10]．提案システムは，地図を媒介することで，地理情報や注釈，GPS データの共有と，時間と位置に基づいた歴史史料や研究論文の関係を効率的に示すことができる．

本章では，本プロジェクトのテストタイプシステムである地域に関する注釈を共有するためのシステム Gcollabo と，GPS データを含む注釈の追加システムについて報告する．最後に，本プロジェクトで開発している地理情報共有システムについて述べる．

6．1 Gcollabo

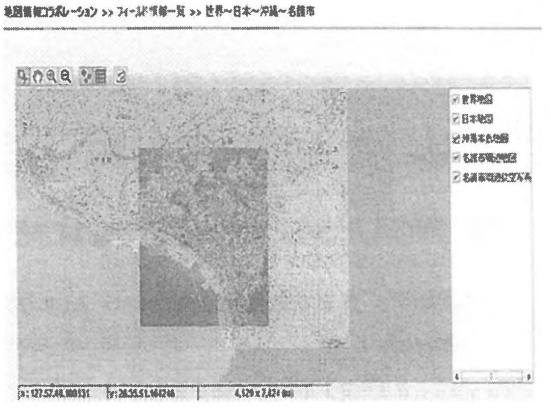


図 5 Gcollabo の実行例

Gcollabo は，地図や衛星写真などの地理情報と注釈の共有を支援するための Web アプリケーションである．図 5 は，沖縄県名護市の基本地図(世界白地図)に地図と衛星写真を重ねて表示している．地図上の点は，その地点に関する注釈が付けられていることを表す．利用者は，注釈を閲覧に加えて任意の地点に注釈を追加することができる．

Gcollabo では，地図や衛星写真などの画像と注釈を個別に管理し，それらの関係をフィールドという単位で管理する．つまり，利用者は，フィールド単位で地理情報を閲覧する．また，それぞれのデータ (フィールド，地図，注釈，フィールド-地図，フィールド-注釈) は，XML ファイルとして記述される．

Gcollabo の地図に対する操作には，地図の拡大縮小，平行移動，グリッドや注釈点の表示およびレイヤー表示の切り替えが可能である．ただし，レイヤーの順序を利用者が変更することができない．また，新規に追加された注釈は，フィールド全体の再表示後に閲覧できる．

Gcollaboは,簡易の地理情報閲覧システムであり,着目する地域について,利用者間での知識共有に有効であると考えられる.しかし,このシステムは,注釈をレイヤーとして扱わないため,地域や記入者などによる注釈の分類を示すことができない.また,レイヤーの追加や削除もフィールドなどのデータを手作業で修正しなければならないため,利用者は,このシステムで公開される地理情報や注釈を使った分析が困難であると考えられる.

6．2 GcollaboへのGPSデータの登録

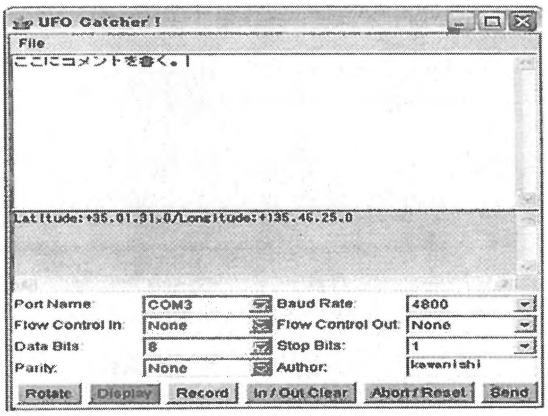


図 6 GPS データを含む注釈追加システム(クライアント実行例)

Gcollabo の応用として,GPS を用いて取得した位置データを含む注釈(以下,GPS 注釈と記す)の登録システムについて報告する.図 6 に GPS 注釈を Gcollabo に登録するためのクライアントソフトウェアの実行例である.本ソフトウェアは,ノート PC に接続した GPS デバイスから取得した位置データと注釈を無線 LAN でサーバソフトウェアに送信する.サーバソフトウェアは,Gcollabo で注釈を閲覧するために注釈ファイルの作成と,地図画像と注釈の関連を管理するファイルの更新をする.

このような Gcollabo の応用は,GPS と連帯した注釈を WWW 環境上でほぼリアルタイムに閲覧することができるので,効率的な意見交換が実現できると考えられる.

また,本サーバと Gcollabo は,ファイルを介して情報を共有するため互いに独立している.したがって,本サーバソフトウェアの出力されるファイル形式の修正も容易であるため,Google Maps などで閲覧するためのファイルも作成可能である.クライアントソフトウェアは,現在,ノート PC 上でだけ利用可能であるが,今後は PDA のようなモバイル機器での利用を目指す.

6．3 地理情報共有システムの開発

6．3．1 基本機能

前節までで述べてきた Gcollabo は,地図の閲覧と注釈の共有だけが可能であり,地図の追加や注釈の分類のような編集操作ができない.また,歴史学や考古学では対象地域の時間的な変化を分析が頻繁に行われるが,Gcollabo は地理情報の時間的な変化を表現することができない.そこで本研究では,次のような要求を満たす地理情報共有システムを開発する.

(1) 地理情報の時間的変化を閲覧する際の効率化
一般的な GIS では,地理情報の時間的な変化を表現するために年代ごとのレイヤーを作成しなければならない.したがって対象期間が長くなるほどレイヤーの数が増え,表示などの処理に時間を要すると考えられる.そこで本研究では,TimeMap などで利用されているタイムスライドバーを用いて,時間的な変化を容易に表現できるようにする.

(2) 地図,文献,GPS データを統合利用
歴史学や考古学などの研究では,地図や衛星写真に加えて,歴史史料や研究文献,GPS による位置データのような,地理情報として校正されていないデータも多数利用する.一般にはこれらの情報ごとに,利用されるソフトウェアが異なることが多い.そこで本研究では,地図を媒介として様々な情報の関連を明らかにするためのデータクリアリングハウスを構築する.

(3) 編集,保存機能
一般に Map Server, Global Map のような WWW 環境を利用した地理情報閲覧システムは,収集した情報について閲覧だけを許す.一方,研究活動では,過去にした収集した情報の追加や,不要な情報の削除のような編集によって情報を整理することが多い.したがって,本研究では,利用者に対して収集した地理情報に対する新たな地理情報の追加や不要な地理情報の削除のような編集操作を許す.さらに,編集結果をプロジェクトという単位で新たな地理情報として公開することを許し,利用者間の知識共有を支援する.

(4) 注釈の共有
本研究では,リアルタイムな知識共有の支援として,閲覧している地図に対する注釈編集機能を利用者に提供する.さらに,本システムでは,GPS を持つ端末によって追加される

注釈の閲覧および編集機能も持つ．これにより異なる地域にいる利用者間での，効率的な知識共有が可能になると考えられる．

(5) 多言語処理

アジア圏の資料は様々な言語で記述されており，それらの閲覧や編集を行うためには，多言語処理機能が必要である．さらに歴史史料には，Unicode のような標準的な文字コードにはない外字が含まれていることがある．したがって，開発する地理情報共有システムでは，外字の表示，入力，検索を含む外字処理機能が必要であると考えられる．

6．3．2 地理情報共有システムの構成

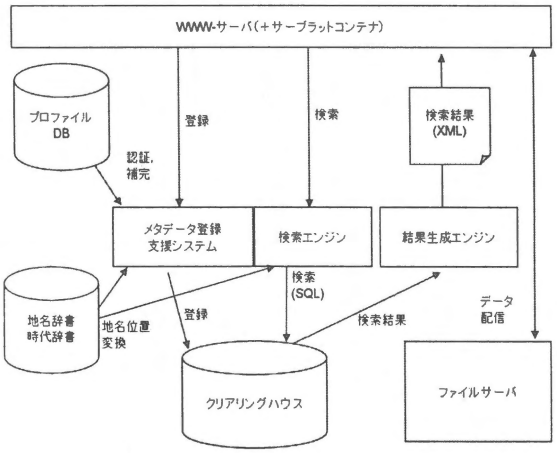


図 7 地理情報共有システム構成

本地理情報共有システムは，一般的なクライアント・サーバ形式のシステムであり，本節ではそれぞれのシステム構成について述べる．

図 7 は，サーバの構成を示している．以下にそれぞれの構成要素について述べる．

(1) WWW サーバ

http を利用してクライアントとの通信を処理する．また，検索エンジンなどは servlet を用いるので，servlet コンテナ機能ももつ．クライアントからの主な要求は，データ登録，検索，データ閲覧があり，クライアントから受け取った値を検索などの処理エンジンに渡される．

(2) クリアリングハウス

一般にクリアリングハウスは，地理情報であれば地理情報だけを対象とするような特定の種類の情報に関するメタデータだけを管理する．しかし，本研究で構築するクリアリングハウスは，情報の横断的利用を目的とするため，地理情報以外にも，歴史文献，GPS データ，プロジェクトのようなデータの種類に依存しないメタデータを管理する．さらに本クリアリングハウスでは，詳細情報を持つ Web ページへのリンクが書き込まれるような注釈データもあわせて管理する．本研究では，歴史的な地理情報に関するメタデータを記述するために，Dublin Core Metadata Elements Set Ver.1.1[11]に基づいて設計された ECAI メタデータ[12]を用いる．

(3) ファイルサーバ

一般にクリアリングハウスは，利用者の求める情報の所在地を明らかにするためのデータベースであり，データ本体はそのデータを作成した組織や個人が管理する．しかし，研究者個人でデータ配信を行うことは技術的，費用的に容易ではない．そこで，本研究で提案する地理情報共有システムでは，メタデータだけではなくデータ自身も管理する．ただし，データ本体は，内容ごとに属性の構成が異なるため，ここではデータベースで管理するのではなくデータ自身をそれぞれ 1 つのファイルとして管理する．また，データ自身の信頼性の保障とサーバのセキュリティの見地から，データの登録は，一定の基準を満たした者だけに登録を許可する．

(4) 地名辞書，時代辞書

これらの辞書は，歴史史料，研究論文，検索キーワードに含まれる地名や住所からの位置(緯度経度)の特定や，和暦のような各国の暦を統一した時間表現に変換するためデータベースである．CSV アドレスマッチングサービス[13]などの現在提供されているアドレスマッチングサービスの多くは，現在の住所を対象にしたものであり，過去の地名から位置を特定することは難しい．そこで，年代を考慮し，過去の住所にも対応した辞書の構築が必要であると考えられる．

(5) プロファイルデータベース

一般的な地理情報や歴史文献，GPS データのメタデータやデータ本体を登録するための認証や，ある GPS データ集合に対するタイトルや製作者などといった共通のメタデータの属性値を管理するデータベースである．

(6) メタデータ登録支援エンジン

このエンジンでは，まずプロファイルデータベースを参照して登録者の認証を行う．次に，

このエンジンはサーバが受信した地理情報や，歴史史料，GPS データのメタデータの形式の検査を行う．その際には，歴史史料に含まれる地名を位置に変換するなどのメタデータの補完処理もあわせて行う．

(7) 検索エンジン
利用者からの検索要求にしたがって，SQL を生成し，クリアリングハウスに問合せを行う．

(8) 結果生成エンジン
クリアリングハウスへの問合せ結果をクライアントでの表示処理をするための形式（XML 文書）に変換し，クライアントに送信する．

次にクライアントの基本的な機能について述べる．クライアントは，通信プロトコルとして http を利用したアプリケーションプログラムであり以下のような機能を持つ．

(a) 閲覧機能
地図の閲覧に必要な地図の拡大縮小，視点の平行移動，レイヤーの表示，非表示，描画順序の変更に加えて，地理情報の時間的な変化を閲覧するためにスクロールバーを模したタイムスライドバー操作や，各年代の地図を一定間隔で自動的に切り換えるようなアニメーション生成機能を持つ．

(b) 検索機能
クライアントでは検索時，キーワード以外に，地図上の範囲や，タイムスライドバー上の期間についても，検索条件として指定できる．また，キーワードによる検索では，メタデータの構造を無視した全文検索とメタデータの構造に即した詳細検索機能を持つ．詳細検索機能によって，人名と地名を区別した検索が可能になると考えられる．

(c) 編集機能
クライアントに表示される地理情報への編集操作として，注釈編集とレイヤー操作がある．注釈編集は，地図の任意の点に対する注釈の追加，削除操作である．レイヤー操作には，収集した地理情報の集合(レイヤー)からの不要なレイヤーの削除や，別の地理情報の追加などが含まれる．また，このような編集操作の結果は，プロジェクトとして保存されるだけではなく，新たな地理情報として公開される．このようなプロジェクト操作によって，本研究で構築する地理情報共有システムを用いた研究成果の公開や研究者間の共同研究支援が可能になる．

6．3．3 地理情報共有システムに関する考察

本地理情報システムは，注釈やある地域に関連する歴史史料などを地図上の点として表現する．このような関連付けの対象は，点として表現できるものだけではなく，川や道路のような線状のものや，山，国のような領域で表現されるものもある．さらに，歴史史料などにある位置情報や時間情報は，あいまいに記述されていることが多い．したがって，関連付けを行う領域は，任意の形状でありかつ，確立モデルなどを用いて領域の曖昧さを表現ができるようにする必要がある．

本地理情報システムは，閲覧，レイヤーの追加のような編集機能をもつが，面積や距離の計算のような空間分析機能は備えていない．しかし，研究活動において，人口密度や可視領域の計算のような空間分析は重要なツールであるといえる．そこで本システムにも，利用者の要求に合わせた空間分析機能を容易に追加するためのフレームワークが必要であると考えられる

7．考察

本プロジェクトには，コンテンツとして地名辞書，歴史的な地震を対象としたデータベース，絵葉書データベースのような位置情報を含むデータベースと，NEARDB のように地名だけを含むデータベースがある．しかし，それぞれのデータベースのスキーマは異なるため，利用者は横断的にデータベースを閲覧することができない．したがって，本地理情報共有システムのクリアリングハウスに登録するためには，メタデータ作成やスキーマ変換の支援が必要である．また，位置情報や時間情報を含むデータは，クリアリングハウスへのデータ登録だけではなく地名辞書や年代辞書のデータとしての利用も期待できる．本稿で述べたデータベース以外の電子図書館やデジタルアーカイブなどの既存のデータベースとの統合は，地理情報共有システムの利用の範囲を広げる上で重要である．ただし，データベースの統合で必要なスキーマ変換やデータ登録は，手作業で行われることが多い．したがって，効率的なデータベース統合のためのスキーマ変換規則などを記述するための GUI ベースの支援ツールの構築が必要である．

本研究で提案する地理情報共有システムは，日本国内に限らずアジアやアフリカなど，長期間かつ広範囲の地域を対象としている．つまり，メタデータなどの値は，多言語であり，かつ年代によ

って地域の名称が異なることがある．したがって，地名辞書や年代辞書は，地名，年代辞書以外に多言語辞書や考古学，歴史学に関するシソーラスやオントロジーの構築が必要であると考えられる．

8．おわりに

地域研究，歴史学，考古学を含む人文科学分野での研究では，歴史史料，学術論文，統計資料，現地調査の結果のような様々な情報を利用する．これらの情報の多くは各地に分散しているため，効率的に利用するには電子化と共有化は不可欠であると考えられる．また，それぞれの情報間の時間的空間的関係を認識することは，研究対象の多角的な考察や分析，正確な再現にあたり重要なことであると考えられる．そこで，本プロジェクトでは，人文科学で扱う時空間情報（地図，歴史史料，研究資料など)の共有に適した情報共有システムの構築を目指している．

また，本稿では，HGIS 研究会各会員の活動報告として，これまで構築したデータベースについて報告した．各データベースは，それぞれ独立しているが，今後は，相互利用のためにメタデータを構築し，本論文で述べた地理情報共有システムでの利用を目指す．提案システムよって，研究者はこれらのデータベースにある時空間情報を横断的に共有と再利用が可能となるので，人文科学研究における新たな知識の発見や視点の獲得が期待できる．

今後は，地理情報共有システムのプロトタイプの構築と評価に加え，ここのデータベースを横断的に扱うためのインタフェースの構築を目指す．

謝辞

本研究は，平成 17 年度日本学術振興会科学研究費補助金基盤研究 (S)「地域情報学の創出」(研究代表者 京都大学教授 柴山守，課題番号 17101008)との共同研究による．

文献

[1] Shibayama Mamoru：“Introduction to Project on Development of Resource Sharing System for Humanities GIS Applications”，PNC 2005 Annual Conference in Conjunction with PRDLA, ECAI, Oct. 31 - Nov. 3 2005, University of Hawaii USA.

[2] 石川正敏：“北東アジア地域の社会科学研究のための資料・書誌情報データベース(NEARDB)の構築”，北東アジア研究 第 8 号 pp.81-94，2005 年 1 月．

[3] 北東アジア・データベース研究会：“戦前期 東アジア絵はがきデータベース”，http://gsv.u-shimane.ac.jp/t-kishi/postcards/，島根県立大学，2004.

[4] 島根県立大学：“e 漢字”，http://ekanji.u-shimane.ac.jp/，島根県．

[5] Shoichiro Hara , Masatoshi Ishikawa , Masato Koyama, et al.：“GIS and Historical Documents - Application of GIS to historical earthquakes documents in Japan -”，PNC 2005 Annual Conference in Conjunction with PRDLA, ECAI, Oct. 31 - Nov. 3 2005, University of Hawaii USA.

[6] Ian Johnson, Andrew Wilson：“The TimeMap Project: Developing Time-Based GIS Display for Cultural Data”, Journal of GIS in Archaeology Vol 1, pp.125 - 134 , ESRI Inc., Redlands, 2003.

[7] Ikuo Oketani：“Construction of the Gazetteer of Japanese Place Names based on Humanities GIS”, PNC 2005 Annual Conference in Conjunction with PRDLA, ECAI, Oct. 31 - Nov. 3 2005, University of Hawaii USA.

[8] 石川正敏，川西陽一，奥村英史，原正一郎:“人文科学のための地理情報共有システムの構想”，情報考古学会第 20 回大会発表要旨，pp. 61 - 66，東京，2005 年 9 月 23-24 日．

[9] 石川正敏，川西陽一，奥村英史，原正一郎：“人文科学のための地理情報共有システムの設計”，情報処理学会技術報告書 2005-CH-68，pp. 49 – 56，2005 年 10 月 28 日，富山県．

[10] Masatoshi Ishikawa：“Development of Japanese GIS Tools for Humanities”，PNC 2005 Annual Conference in Conjunction with PRDLA, ECAI, Oct. 31 - Nov. 3 2005, University of Hawaii USA.

[11] Dublin Core Metadata Initiative：“Dublin Core Metadata Element Set, Version 1.1: Reference Description”，http://dublincore.org/documents/dces/, 2004.

[12] Electronic Cultural Atlas Initiative：“ECAI Metadata Clearinghouse 4.0 - Metadata Tags”, http://ecaimaps.berkeley.edu/clearinghouse/html/list_md_elements.php,1999.

[13] 相良 毅:“CSV アドレスマッチングサービス”，http://www.tkl.iis.u-tokyo.ac.jp/~sagara/geocode/，東京大学空間情報科学研究センター．

4D-GISによる遺跡空間データベースの構築と先史文化生態解析

Constructing Archaeological spatiotemporal information database with 4D-GIS through an analysis of the prehistoric cultural dynamics

津村 宏臣
Hiro'omi TSUMURA
同志社大学 文化情報学部, 京田辺市多々羅都谷 1-3
Doshisha University, 1-3, Miyakodani Tatara, Kyotanabe-shi, Kyoto

あらまし: 本研究の目的は, 考古学に時空間情報科学の理論的な考え方を導入することと、これを実際の遺跡発掘調査情報を基にした、実践的な集落論研究を通じて実践的研究を行うことである。対象は、東京・武蔵野台地東部の300㎢の範囲であり、縄文時代中期(約5400～4500年前)の集落が密集する地域である。本研究では、これまでの調査成果をふまえて、この地域に約2000人が居住していたことを想定し、それらの人々の文化生態的な動きを時空間的な解析から具体化することを目的とした。98遺跡(先史集落)の位置情報と出土遺物・所属時期などの属性情報をGISによりデータベース化し、以下の4つの時空間解析を実施した。1)多次元属性を参照できる集落分布図の作成、2)分布位相の特性を評価するVoronoi分割、3)集落間の視認関係を基盤としたネットワーク分析、4)文化属性データによる空間内挿(傾向面分析)と、放射性炭素年代測定値による(時間内挿)、を実施した。結果として、4D-GISが文化生態(動態)を描き出す技術として、きわめて有用であることを検証した。

Summary: The purpose of this study is to discuss the prehistoric settlement in the view of Geostatistical analytics, and to introduce the theoretical concepts of GIScience into Archaeology. The subject area is the eastern part of the MUSASHINO upland which has an area of about 300 square kilometers, at the time of middle Jomon period (cal BP. 5400-4500) with 98 archaeological sites, and the population in highest stage of development is estimated to be 2000. First, a database that has the mutual comparability in the space and site properties was built by using GIS. Then the following four analyses were applied. 1) Many digital site distribution maps, which can refer to miscellaneous properties, were made, 2) Since the geometrical character of its Diagrams symbolize the topological individuality, the Voronoi method (as interpolation) was applied, 3) The visual verification between sites was extracted by application of Viewshed Analysis for elucidation of inter-site spatial relations, and 4) Trend-surface analysis was enforced using point properties of archaeological cultural elements to reconstruct the spatially concrete images. Finally, the results were lined up along the "unconditional" time-series. We propose considering this approach as Four-dimensional "4D" GIS.

キーワード: 4次元GIS, セトルメントシステム, 時空間情報科学, 空間統計解析
Keywords: 4D-GIS, settlement system, spatiotemporal information science, geostatistics

1. はじめに

"なぜ遺跡がそこにそうあるのか"という問いかけは, これまで, 遺跡立地論や分布論として議論されてきた. 遺跡周辺の古環境評価や遺跡領域の復原から, 往時の遺跡空間とそこに住んだ人々の生活の姿が想像された. さらに, 個別の遺跡を総体的に捉えるセトルメントシステム論への議論の伸張は, 遺跡や遺物の多次元的な時空間的展開から先史社会を再構築する不可欠な視座であり, 多くの研究が実践的に先史社会の姿を描き出してきた.

しかし, 個別の遺跡や遺物の"そこにそうある"存在の背景には, 様々な自然・社会・文化的な背景が重層的に関わっていると考えるのが妥当で, "そこにそうある"ことが決定論的な前提となる限り, "なぜ"という問いに的確に答えることは難しい. 重層的な背景を確率論的に評価しない限り, その存在と時空間的なパターンに顕現した"複雑さ"は正しく理解できない.

近年, 時空間情報科学と地理情報システム(:以下 GIS)の発達により, 時空間情報を確率論的に評価する方法や技術が飛躍的に進展し, 日本考古学においても導入が試みられている(金田・津村・新納 2001). ここでは, この時空間情報科学の考

Fig.1　対象地域のDEMと等高線図・遺跡分布(ポリゴン)図・河川流路(ライン)図

え方とGIS技術を適用し，武蔵野台地東部の縄文時代中期集落の生態的特徴の解析を試みる．“なぜ遺跡がそこにそうあるのか”という問いに対し，多次元因子の“複雑さ”を評価し，仔細な時系列動態のなかからこれを理解する．

2．遺跡空間データベースと分布図作成

2-1　対象地域

具体的な検討対象として，武蔵野台地東部・縄文時代中期(cal B.P. 5400–4500)の集落遺跡をとりあげる．

武蔵野台地東部は，古くから細密な考古学的な調査が実施され，遺跡の存在が最もよく把握されている地域の1つである．遺跡分布については，これまで多くの研究者が，湧水や河川と集落との関係，密集する集落間関係，その相互の様相などについて検討してきた．空間事象を評価するうえで不可欠な時間情報についても，中期13期31細別の新地平編年(黒尾・小林・中山 1995)のような細かな時間軸が設定されている．

対象範囲は，国土地理院発行の 1:25000 地形図，「吉祥寺」・「東京西部」・「溝口」・「東京西南部」の図幅の範囲で，東西約 20000m、南北約 16000mをはかる．多摩川南岸域や石神井川北岸域を除き，縄文時代中期の集落遺跡は98遺跡確認されている．最も遺跡や住居址の数が多い中期後半の加曽利E3～E4式期にかけては，約2000人が居住していたと推察され(津村・小林・坂口ほか 2002)，先史社会としては比較的人口密度の高い地域の1つである．

2-2　遺跡空間データベースの構築

GIS を利用した遺跡空間データベースは，先述の4図幅を対象として作成した．西南隅(北緯 35° 35′ ，東経 139° 30′)を座標起点とし，東に向かってX軸，北に向かってY軸を設定し，北東隅(北緯 35° 45′ ，東経 139° 45′)までの範囲に東西約 20000m、南北約 16000mの座標メッシュを設定した．この範囲の 10mメッシュ DEM をラスターデータ化した等高線の内挿により作成し，対象98遺跡の範囲(ポリゴン)，河川流路(ライン)をデジタイズによってベクターデータとして作成した．また，98遺跡の属性データベースは，

①遺跡規模(住居址数/推定人口)
②存続期間(土器型式/暦年較正年代測定値)
③土器型式と組成比(ex.阿玉台・勝坂・加曽利…)
④炉の形態と組成比(ex.地床・埋甕・添石・石囲…)
⑤石器器種と組成比(ex.石鏃・石斧・石皿・石錘…)
⑥黒曜石原産地比(ex.神津島・信州・箱根…)

の6属性をデータ化し格納した．特に②は，碓井(碓井ほか 2000)で「不確実な性格(p.104)」のデータとされた遺跡の所属年代に，リニア(1次元)な時間情報としての放射性炭素年代測定と暦年較正年代値の導入を指向する．①については，便宜的ではあるが，同時機能の住居址20軒以上を大規模遺跡，5軒以上を中規模遺跡とし，それ以外を小規模遺跡とした．

2-3　遺跡空間データベースによる分布図の作成

Fig.1・2は，この遺跡空間データベースによって作成した遺跡分布図である．Fig.1のような，縄文中期というマクロな時期区分での遺跡分布をみると，従来の指摘のように，多くの遺跡が河川に沿って規則的に立地している状況が読みとれる．この傾向は，多摩川・野川・仙川・石神井川などで顕著で，遺跡間の距離が多少広いが，善福寺川や神田川でも看取できる．また，面的に展開する目黒川流域や，一見すると孤立的な遺

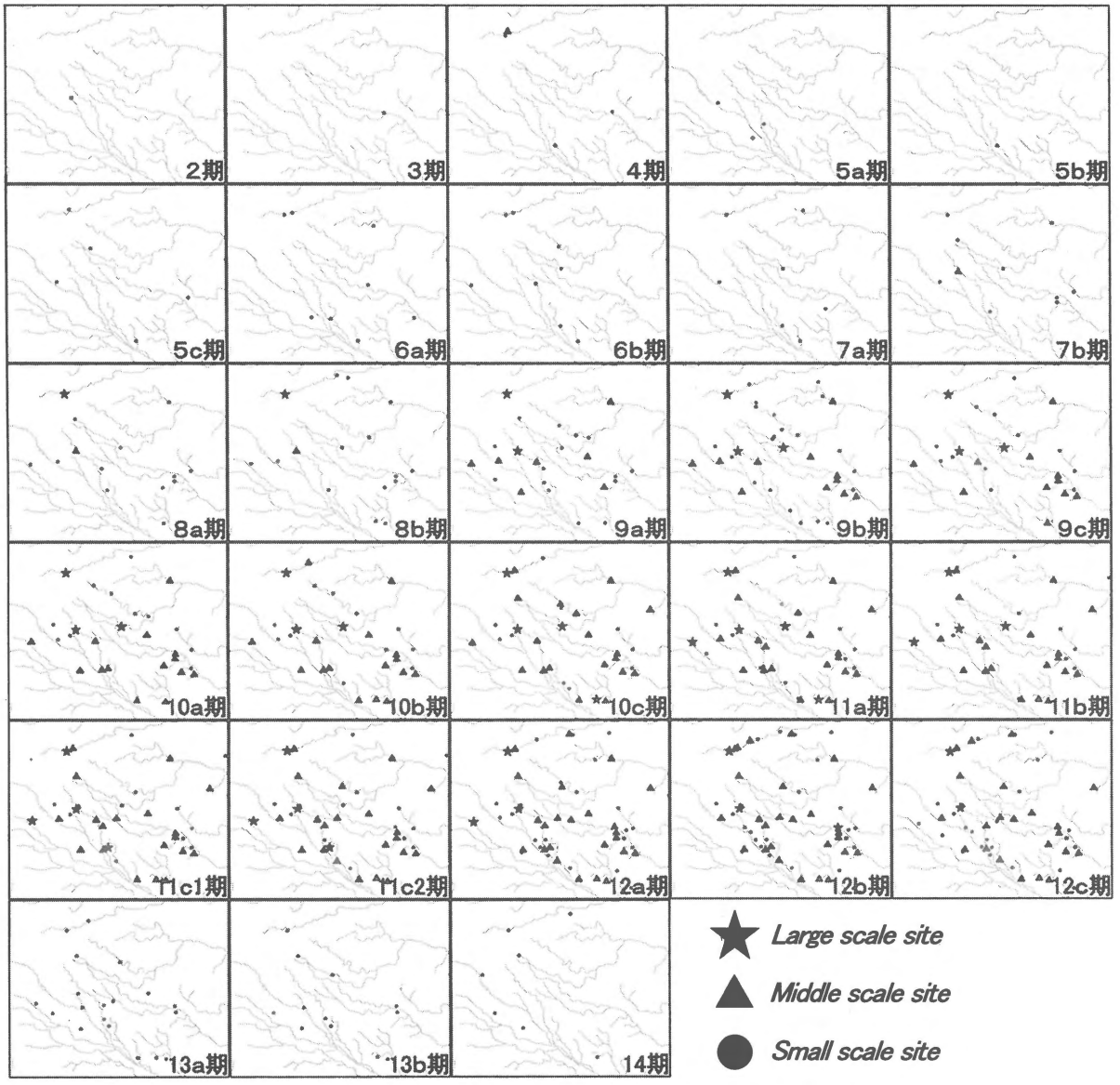


Fig. 2　対象地域の河川流路図と規模・土器型式時期別遺跡分布図

跡[三栄町(54)や代々木八幡(95)など]も，おおよそ支流や小さな沢に面した立地であることがDEMとの定量的比較で評価できる.

概略的に傾向をまとめると，①多くの遺跡は河川や沢などのいわゆる「水場」の近傍に立地し，②その中でも河川の合流部や屈・湾曲部に立地する遺跡[鳥森(12)，祖師谷大道北(48)，落合(56)，松ノ木(62)，高井戸東(65)，谷戸第二(66)，井荻三丁目(69)，堀北(76)，池淵(77)，武蔵高校裏(98)，永隆寺東方(94)など]がある程度認められる. ただ，③「水場」との空間的関係の希薄な分水界に立地する遺跡[南(3)，松原羽根木通(20)，東原(89)]も存在し，これは②の傾向を有する遺跡と地形的な属性を共有している蓋然性が高い，などが指摘できる.

3. 遺跡分布位相の評価 I　-ボロノイ分割-

3-1 ボロノイ分割の原理

分布位相を空間幾何的に評価する方法として，平面上に存在する複数の個体間に領域(多角形)を創出して対象間空間を評価する，ボロノイ分割による解析法を採用する.

このボロノイ分割は，同じ位置にない母点の集合を$P=\{p1, \cdots, pn\}(2 \leqq n < \infty)$とし，$d(p,pi)$を点pと母点piの間のユークリッド距離として，$V(pi)=\{p \mid d(p,pi) \leqq d(p,pj), \quad i \neq j, \ j=1, \cdots, n\}$で与えられる領域をpiの通常のボロノイ領域，$\{V(p1), \cdots, V(pn)\}$をPによって生成された通常のボロノイダイアグラムとする，と定義される.

ボロノイ分割は個々の領域をあわせたものが全領域になるモデルで，配置パターンによって領域形状が変化する. 隣接

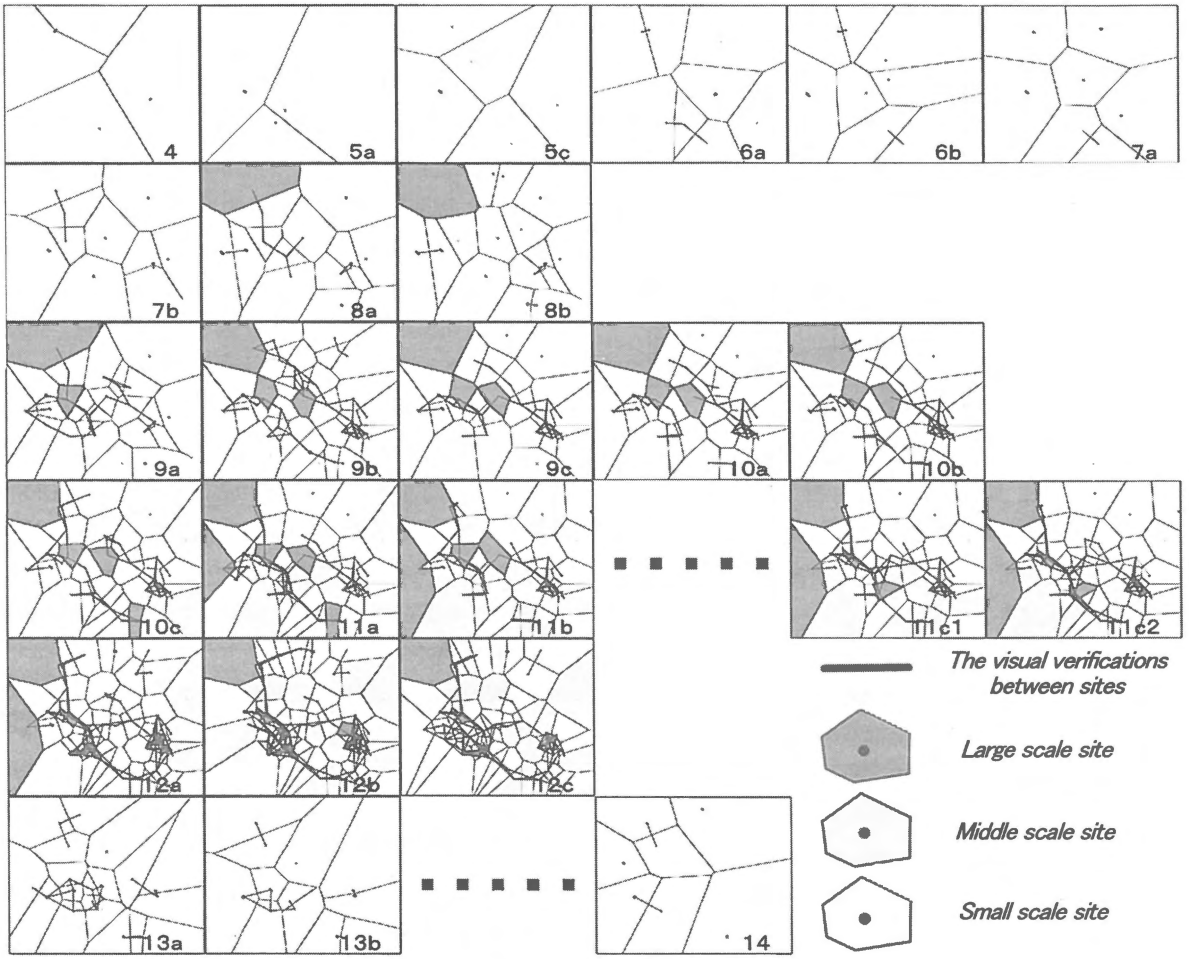


Fig.3 遺跡のボロノイ分割と遺跡間の視認ネットワーク

個体対の組合わさり方は，対象の位相を表現し，ボロノイ点の展開やボロノイ辺の辺長・辺数などは，評価の要素となる．六角形の領域が並接するモデルは，対象が均等配置の場合に現れ，逆に，領域が扁平に歪むモデルの場合，対象の配置が何らかの異次元属性に規定された場合に現れる．これにより，個々の領域の形状と隣接する領域との関係が明らかになれば，分布位相が評価できる．

3-2 ボロノイ分割の適用と時系列動態の抽出

細別時期別で遺跡が複数存在する時期に限り，ボロノイ分割を実施したのが Fig.3 である．領域の形状や配置に時系列での変遷が看取でき，先に視覚的に評価した縄文中期の“河川に沿う”背景とは異階層の別の背景の介在が示唆される．

遺跡の少ない6期でも，六角形状の領域は描出されておらず，分布を規制する何らかの因子が介在していることは明らかである．7期は，6期から8期への変遷過程のなかで理解できる．8期には，下野谷(86)や三鷹市立第五中学校(50)が大・中規模の遺跡として出現し，8b 期で対象地域にほぼ均密に遺跡が配置される．9a 期との様相の転換をみても，ここに時系列での画期が設定できるだろう．

9a 期では，遺跡数も増加し，領域面積や形状に多様性が現れる．目黒川周辺の明治薬科大(10)や松原羽根木通(20)の領域は六角形状を呈すが，仙川沿いの遺跡領域は短冊状や狭歪な不整形となる．9b 期でさらに様相が変化する．対象地域の北西隅から南東隅にかけて，六角形状の領域をもつ大・中規模遺跡が，相互に隣接個体対をとなりつつ線状に展開する．これに対し，小規模遺跡の領域の多くは不整形を呈している．この 9a 期から 9b 期への転換は，遺跡分布図の視覚的評価からは明らかにできない位相の変化である．また，善福寺川周辺では河川沿いの立地を反映して，10b 期まで短冊状の領域が連接するが，仙川など多摩川水系では，領域の様相からは看取できなくなる．これは，位相に顕現する因子に構造的な差があることを示唆する．

10c 期以降、善福寺川周辺で六角形状の領域が形成され始めたり，東原(89)の領域と短辺を接して短冊状の領域が放射状に配置されたり，奥沢台(5)や三鷹市立第五中学校(50)，大蔵(37)などの大規模遺跡の領域が狭歪で不整形な形状に変

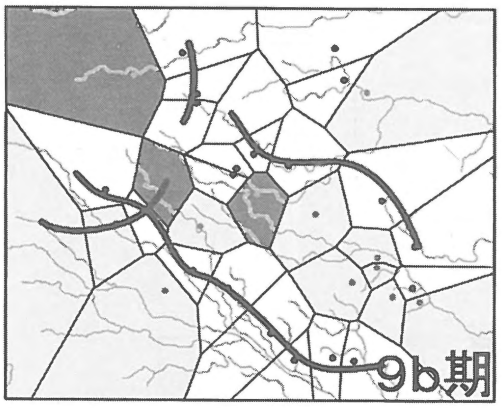


Fig.4 9b 期の遺跡分布とボロノイ領域

化しつつも，11c2 期までは同じ傾向での推移を読みとれる．また，目黒川周辺では，面積に違いはあるが，六角形状の領域が安定して展開している．以上から，10b 期と 10c 期の間に2つ目の画期を考えることができる．

12a 期に入ると様相が大きく変化する．まず，石神井川流域で長辺を接して並列する短冊状領域が顕著になり，東原を中心とした放射状の領域配置も傾向が強化される．こうした傾向は 12c 期まで継続する．13 期以降は，遺跡数の減少とともに領域面積が増加する．ただ，六角形状の領域はあまり看取できず，いずれも不整形な点は注意しておきたい．

3-3 集落分布の位相の評価

以上をまとめると，

①対象域全体で通時的には，ボロノイ領域が六角形を呈する均等な位相は，時期を問わず明確でない．

②時期が新しくなるにつれ，総体的にボロノイ領域は狭歪で不整形となる．

③同様に，遺跡規模とボロノイ領域の面積は対応せず，領域配列との規則性も看取できなくなる．

④配列(短冊状領域の横位並列や放射状配列)から，複数の背景が同じ階層で作用している．

などが指摘できる．

例として，9b 期について詳細に見ると(Fig.4)，遺跡は個別的には河川沿いに立地するものの，河川沿いに立地しない東原(89)遺跡のボロノイ領域を中心にした放射状配列や，多摩川下流域での短冊状領域の横位並列など，④の特徴が看取できる．また，遺跡規模との関係では，大・中規模遺跡の領域が六角形状を呈するのに対し，小規模遺跡の領域は相対的に狭歪で不整形である．これは，遺跡規模に応じて遺跡立地の背景が異なる空間コンプレックス(津村 2002a)を示唆する．具体的には，大・中規模遺跡は一定の間隔を保って均等に配置されるのに対し，小規模遺跡はその他の背景(例えば Fig.4 の曲線)で配置が決定している可能性が高い．

4．遺跡分布位相の評価Ⅱ －視認ネットワーク分析－

4-1 Viewshed Analysis と視認ネットワーク

対象地域の遺跡分布には，“河川に沿う”という自然的因子の他に，9b 期の大・中規模遺跡に顕現した均等配置のような，遺跡間関係を前提とした社会的因子の介在が明確となった．この社会的因子をより具体的に抽出するため，ここでは1つの方法として，遺跡間の視認関係について検討する．

この視認関係は，各細別時期別，各遺跡単位で GIS の Viewshed Analysis を実行し，眺望域に存在する遺跡を直線で結びネットワークとして抽出する．Viewshed Analysis は，特定の対象からの視界域を抽出する分析で，例えば，遺跡の範囲と視点の高さを指定すれば，対象範囲の DEM に視点の高さを加算し，周囲360° の範囲の仰角と伏角の計算からDEMで地形を考慮してその視界域を面的に抽出する．描出された視界域に存在する相互の遺跡を直線で結ぶよう設定し，これを“遺跡間の相互の視認関係”として扱った．こうして描出したラインを視認ネットワークと呼ぶ．

4-2 視認ネットワークの時系列動態

細別時期別に視認関係の有無を結線で描出したものがFig.3で，先のボロノイ分割の結果にオーバレイして表示した．ネットワークの形状が各時期で変化し，視認関係に顕現するような遺跡間関係に時系列で変遷があることが明確となった．

遺跡数の少ない6期までは，4期の下野谷(86)と下柳沢(85)をのぞいて視認関係はない．この傾向は 8a 期に変化があるものの基本的に 8b 期まで継続する．

本格的に遺跡が視認関係を持ち始めるのは 9a 期からで，その顕著な様相は 9b 期に定立する．9b 期から 10b 期には，対象範囲の北西隅から南東隅にかけて，南北に2つのネットワークが形成される，北側のネットワークは先述した善福寺川の河川沿いの遺跡立地と目黒川周辺の面的な展開とを，松原羽根木通(20)が中継して成立する．南側のネットワークは仙川の河川沿いの遺跡立地を顕現した様相と評価できる．この北側のネットワークの要となっている松原羽根木通は，分水界に立地する特異な立地の遺跡であり，この立地はネットワークを成立させる重要な空間であることが指摘できる．

10c 期は 11a 期に向かう漸移的な様相といえる．善福寺川流域では遺跡数の減少に伴ってネットワークが成立しなくなりつつある．また，御殿場(73)と東原(89)の存在で，対象地域の北西部を弧状に連結するネットワークが形成され始める．上記した松原羽根木通の例でもわかるように，これらの特異な立地はネットワークを成立させる要としての空間と評価できる．この傾向は 11c2 期まで継続するが，11a 期以降，神田川流域の下高井戸塚山(60)，烏山川流域の八幡山(52)や廻沢北(53)，仙川流域の釣鐘池北(47)が存在することで，河道方向とは全

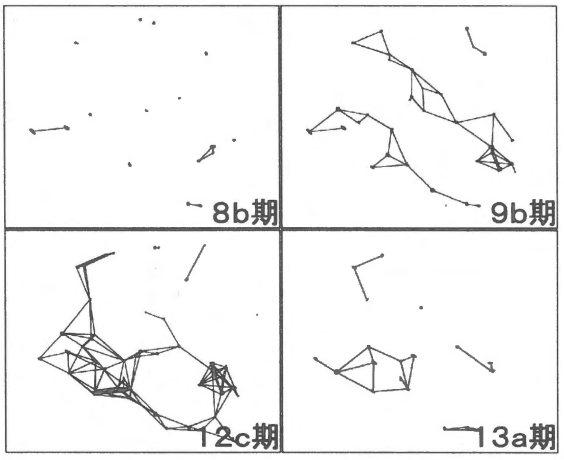


Fig.5 遺跡間視認ネットワーク形状の時系列動態

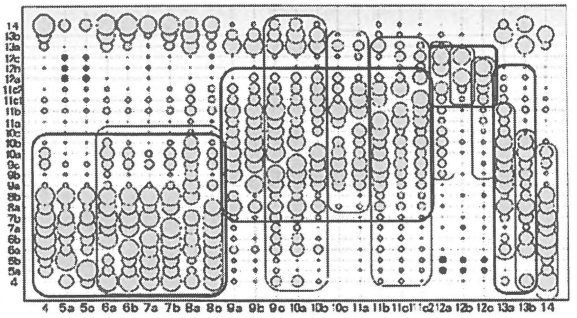


Fig.6 視認関係数の相関・分散共分散行列グラフ

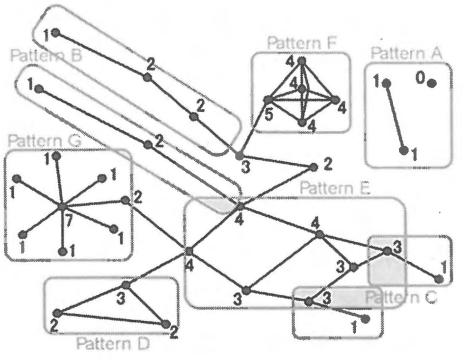


Fig.7 視認関係数とネットワーク形状のモデル

く無関係なネットワークが形成される. 特に, これらは松原羽根木通を中継して目黒川周辺の遺跡とネットワークを形成しており, 南北2つのネットワークを対象地域の中央で連結させる役割を担っている.

12a 期に出現する南(3)は、この南北のネットワークを対象地域の南東端で連結する要の立地である. 先述の松原羽根木通や東原と同様, 特異な立地条件の背景には, ネットワークの要としての機能が顕現している蓋然性が高い. また, この12期に石神井川流域での視認関係が強化している. 13a期には遺跡数が減少しネットワークも分断されるが, 野川流域や多摩川流域では 12c 期に形成された網状のネットワークが残存し, その他の遺跡も脆弱ながら視認関係を維持している. この傾向は13b期まで続くが, 14期にはネットワークは解体する.

4-3 視認ネットワークの評価

Fig.5は, 表記時期の視認ネットワークを図示したものである. まず特徴的なことは, 9b 期の視認ネットワークの形状が Fig.4 で抽出した曲線の形状と酷似することで, 9b 期の小規模遺跡の配置が小規模遺跡間あるいは大・中規模遺跡との視認関係によって規定されている蓋然性を示唆する. 全体的には, 遺跡数の増加に並行した視認ネットワークの複雑化が評価でき, 様相に時系列で動態があることもわかる. この複雑化についても, 8b 期のような"点"的な様相から 9b 期の"線"的な繋がり, 12c 期のような"面"的な拡がりへ, という方向性が看取できる.

また, 遺跡数が激減する 13a 期(19 遺跡)と増加以前の 8b 期(16 遺跡)の様相は興味深い. 同じ面積にほぼ同数の遺跡が河川に沿いつつ分布するが, 双方の視認ネットワークのあり方は対称的である. 遺跡の増え方と減り方の背景には異なった因子を評価する必要があることを示唆しており, 単純な人口増加や減少では何も説明したことにはならない.

次に, この視認ネットワークの形状と遺跡分布との関係を, 細別時期別での時系列動態とともに評価するため, 抽出された視認関係保有数とその組成について解析する。まず, 細別時期別の時系列での相関を明確にするため, 視認関係保有数の相関・分散共分散行列を抽出, グラフ化したのが Fig.6 である。相関の強さで円を大きくしている. これにより, 8b 期と 9a 期の間, 11c2 期と 12a 期の間, 12c 期と 13a 期の間にそれぞれ明確な画期が看取でき, その下位に 10b 期と 10c 期, 12b 期と 12c 期の間などにも変化を捉えられる. また, 視認関係保有数の出現頻度が, 間接的に視認ネットワークの形状を示唆する. 視認ネットワーク形状をモデル化(Fig.7)すると, 最頻値が1や2値の場合, ネットワークは"線"的であり, 4や5値の場合"面"的である. あわせて標準偏差を評価すれば, それらの組み合わせのパターンが抽出できる. 以上をまとめると,

①視認ネットワークとボロノイ領域の配列に相関があり, これも時系列での動態がある.

②視認ネットワークは遺跡の増加に並行して複雑化するが"点⇒線⇒面(網)"の方向性がある.

③遺跡分布の背景に, "視認関係の有無"に表出するような遺跡間関係が介在する可能性が高い.

などが評価できる. また, 河川との関係が希薄な立地が視認ネットワーク形成に深く関わっていたことを指摘したことも重要である. 往時の居住地選択の背景に, 視覚的な物質空間認知が関わっていた蓋然性が指摘できる.

5. 遺跡分布位相の評価Ⅲ -時-空間内挿分析-

5-1 傾向面分析とその原理

遺跡立地・分布の文化的背景を探るため，考古学的な文化要素によって確率論的に空間内挿する傾向面分析をおこなう．各遺跡で検出された土器型式(系統)・炉形態・石器器種組成・黒曜石原産地構成を対象に，その比を属性値として採用した(分類の詳細は津村・小林・建石ほか 2002 を参照)．傾向面は属性数と分布によって精度に差が出るため，細別時期別より資料数が確保できる土器型式による時期区分を採用した．ここでは，土器型式(系統)と炉形態について詳細を述べる．

傾向面は次元数によって，

LINEAR: $z = b0 + b1x + b2y$

QUADRATIC: $z = b0 + b1x + b2y + b3x^2 + b4xy + b5y^2$

CUBIC: $z = b0 + b1x + b2y + b3x^2 + b4xy + b5y^2 + b6x^3 + b7x^2y + b8xy^2 + b9y^3$

の多項式でもとめ，適合度と検定で判別した．以上により作成した傾向面が Fig.8である．概観すると，文化要素の資料特性に応じて傾向面の様相に違いがあることが読みとれる．

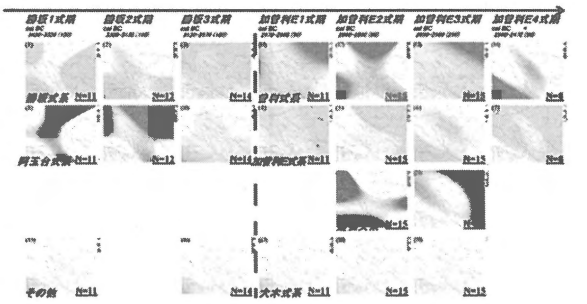
5-2 土器型式(系統)の傾向面と時系列動態(Fig.8-a)

縄文時代中期中葉と後葉に大別し，勝坂系土器群と曽利式系土器群を上段，阿玉台式系土器群と加曽利E式系土器群を中段，そのほか客体的に組成する曲隆線文系土器群や鳴神山系などの土器群と大木式系土器群を下段，一過性の連弧文系土器群を中下段に配した．凡例は濃色になるほど組成比が高くなるよう設定した．

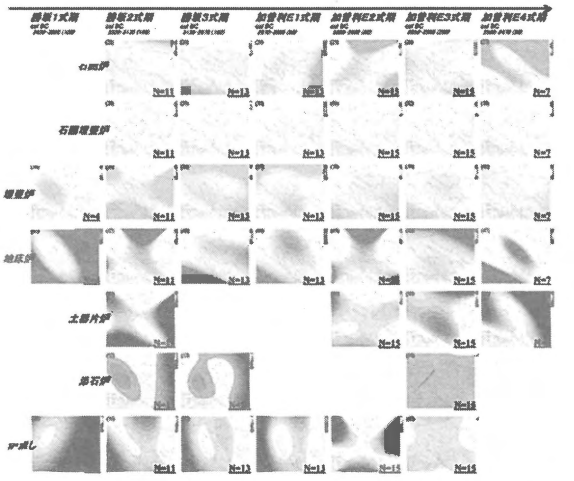
勝坂1～3式期では，勝坂式系と阿玉台式系が相反して展開し，段階的に勝坂式系が主体となる様相が看取できる．これらと展開を異にしたその他の型式が客体的に組成し，勝坂3式期には阿玉台式系と同様の展開を示す．興味深いのは，主要な2つの土器型式が排他的な関係で展開する点である．

土器文化が全体的に変化する加曽利E1式期には，加曽利E式系が勝坂式系の様相を保持したまま成立し，曽利式系や大木式系が客体的に組成する．それまでの土器型式の相反する様相は，この加曽利E1～2式期では看取できない．その後，加曽利E3式期の加曽利E式系と連弧文系，加曽利E4式期の加曽利E式系と曽利式系の間で再び顕現するが，主体としての加曽利E式系に客体的な土器系統が組成する様相と評価でき，中期中葉までとは構造的な相違がある．

8-a の各傾向面は，いずれも対象地域外に中心がある場合が多く，対象地域内に明瞭なまとまりや単位は見出せない．いずれの土器型式も対象地域以外に分布(≒文化)の核があり，この地域での展開は周縁的な様相の一部分を描出している可能性が高い．また，それぞれの時期間でのスムーズな推移は看取できず，いずれの時期も様相の転換が著しい．



8-a 土器型式(系統)



8-b 炉形態

Fig.8 文化要素の傾向面と時系列動態

5-3 炉形態の傾向面と時系列動態(Fig.8-b)

中期中葉の勝坂1式期では、埋甕炉と地床炉が排他的に相反して展開し，炉無しも地床炉と似た様相を呈す．土器型式での傾向面と異なり，入間・仙川上流域を中心とした小規模な分布の核が見出せる点は興味深い．勝坂2式期では，[石囲炉－石囲埋甕炉－地床炉]と[埋甕炉－土器片囲炉]が相似の様相を呈し，相反して展開する．また前段階での埋甕炉の核は，添石炉の核に置換する．勝坂3式期では，[石囲炉－地床炉]がやや相似の様相を呈し，埋甕炉と相反して展開する．前者は目黒・神田・善福寺川流域で，後者は多摩川流域で線状の単位を形成する．石囲埋甕炉は軌を異にし，烏山川中流域に緩やかな単位を形成する．いずれも前段階とは様相が著しく転換している．添石炉は、前段階と同様，入間・仙川上流域を中心に小規模な核を保持し，炉無しも変化はない．

中期後葉の加曽利E1式期では，[石囲炉－石囲埋甕炉]が相似となり，善福寺・神田・北沢・入間・仙川・野川の上流域を横断する核を形成する．また埋甕炉は多摩川中流域に，地床炉は善福寺川流域にそれぞれ核を形成し，3者が複合的に錯綜する．これは，中期中葉までの要素単位での相反した展開とは異なった状況である．また，[石囲炉－石囲埋甕炉]の展

開は，“河川沿い”という遺跡立地の背景では説明しえないことは明らかで，文化的情報を媒介する遺跡間関係が“河川に沿わない”因子を背景としていた蓋然性を示唆する．

加曽利E2式期には様相が転換し，それまで相似か複合的であった石囲炉と石囲埋甕炉が相反し始める．石囲炉は勝坂3式期と同様再び地床炉と，石囲埋甕炉は新しい要素としての土器片囲炉・炉無と相似となる．石囲炉は前段階の特徴をもつが，石囲埋甕炉は様相が転換する．また，それまで相反してきた埋甕炉と地床炉が複合的となる．

加曽利E3式期では，いずれの形態も空間的展開を異にし，対象地域で全くの錯綜的様相を呈す．石囲埋甕炉は多摩川流域に，埋甕炉は善福寺川中流域に，地床炉は多摩川中流域に，土器片囲炉は烏山川中流域に，それぞれ緩やかな単位を形成する．

加曽利E4式期では，中期中葉で看取された，排他的に相反した展開が再び顕著となる．いずれの形態も空間的に相似ではないが，石囲炉と石囲埋甕炉，地床炉と土器片囲炉が，それぞれ排他的に相反する．

以上をまとめると，

①対象地域内に分布の核が存在する．

②各時期別で様相が著しく転換する．

③中期中葉段階で看取できる相互・排他的な展開が，後葉段階で錯綜・複合的な展開となる．

などが評価できる．

土器型式の傾向面と比較すると，①のような異なった特徴と，②③のような共通する特徴が指摘できる．①については，炉は資料特性として不動であり，土器のようなモノとしての動的特性はない．これは，遺跡を残した当事者の文化的背景をより直接的に顕現する点で重要で，ここで評価した相似や相反などの空間的様相は，往時の人々の文化・社会的な距離を直接的に示唆する．また，②や③については，これが不動・可動の両特性の文化要素で共通して認められることから，この時系列動態は考古学的文化の実相として，文化・社会的な距離が遠く排他的な段階から，活発な交流を介した複合・重層的段階への推移と解釈できる．

5-4 傾向面の核と視認ネットワークの時空間的相関

Fig.9は，加曽利E1式期の炉形態の傾向面に同時期の視認ネットワークをオーバレイして表示したものである．各河川の上流域を南北にのびる線的な視認ネットワークが石囲炉と石囲埋甕炉と，北西から南東に向かってのびる2本のネットワークでは，北側で地床炉と，南側で埋甕炉とそれぞれ対応した展開が看取できる．先述したように，遺跡間の視認ネットワークの形成は，遺跡数増加に並行した社会の複合化過程と深く関係している．

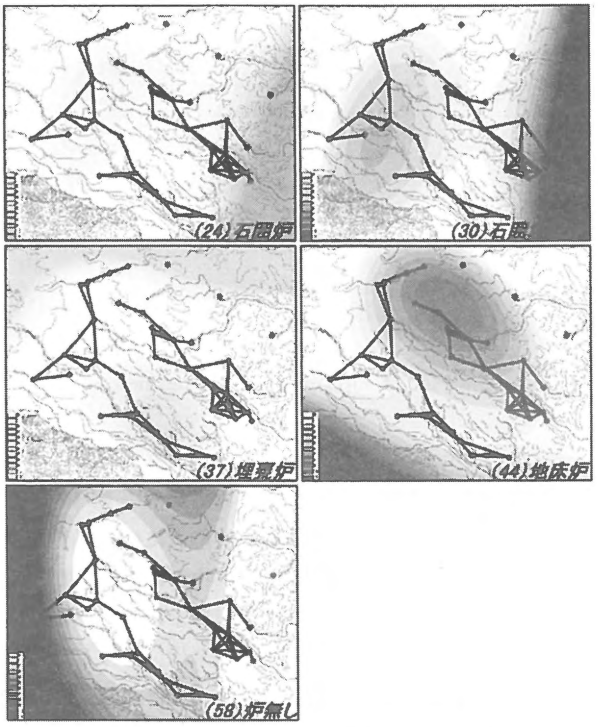


Fig.9 加曽利E1式期の炉形態の傾向面と視認ネットワーク

6. おわりに

本研究は，遺跡情報の処理・評価・解釈に際し，空間情報科学やGISがいかに適用し何を読むのかを，具体的な分析を通じて検討した．従来の時間事象による歴史叙述の枠を越え，空間事象による歴史叙述の方向性は示せたと考える．

引用・参考文献

碓井照子・森本晋・泉拓良・清水啓治・中雅明・松村憲治・高瀬啓司・鈴木繭子・水野行子・野田卓・関口靖之 2000「考古学データベースにおける考古学ﾒﾀﾃﾞｰﾀとｸﾘｱﾘﾝｸﾞﾊｳｽ」『地理情報システム学会講演論文集9』

金田明大・津村宏臣・新納泉 2001『考古学のための GIS 入門』古今書院

黒尾和久・小林謙一・中山真治 1995「多摩丘陵・武蔵野台地を中心とした縄文時代中期の時期設定」『縄文中期集落研究の新地平』

津村宏臣 2002a「空間錯体(Complex)の描出と遺跡間関係評価の方法」『動物考古学 18』

津村宏臣 2002b「地理情報システムの適用による“考古学的文化”概念の再考」『地理情報システム学会論文集 11』

津村宏臣・小林謙一・坂口隆・建石徹・西本豊弘 2002「縄文集落の生態論(2)」『動物考古学 18』

津村宏臣・小林謙一・建石徹・坂口隆・西本豊弘 2002 「縄文集落の生態論(3-1)」『動物考古学 19』

国書古典籍中の挿絵・絵本に描かれた実在キャラクター達の存在意義
—情報学から文学論へのエチュードとして—
About the significance of existences of the real existence characters drawn in the cut-in illustration in the book Japan classic
— Etude to construct literature theory based on information science —

相田 満
AIDA,Mitsuru
国文学研究資料館,東京都品川区豊町 1-16-10
National Institute of Japanese Literature,1-16-10 Yutaka-cho Shinagaea-ku Toyko

日本人が好んできた人物やキャラクターは誰か。国文学研究資料館に所蔵される古典籍刊本に収められる人物の画像を集積した『歴史人物画像データベース』は、そのような素朴な疑問に導かれて始まった。

このデータベースの出現頻度上位者には、①菅原道真 ②聖徳太子 ③曾我時致などの人物達が並ぶ。現代の我々からすれば、意外に思う人物が浮かび上がっているのだが、そもそも、昔の有名人はいったい誰で、各時代や時期では、どんな人物が好まれていたかという問いに、明確に応え得る指標はこれまで皆無に等しかった。しかし、このデータベースは、そうした問いへの答えとなる可能性もはらんでいる。

What is the person or the character that the Japanese had been liked? The data base from which it is started to be led to such a simple doubt is 'Historical character's Image-database. 'This data base is accumulation of the image of the person collected to the classical print books that NIJL（National Institute of Japanese Literature） owns.

The following people are listed by an appearance frequency high-ranking person of the data base. For instance,（1）SUGAWARA Michizane,（2）Prince SHOUTOKU,（3）SOGA Tokimune,and so on.We modern will often unexpectedly think of these results.To begin with, there was no standard by which who was most famous or who had been liked at ancients was shown up to now. However, there is a possibility of becoming an answer to such a question in this data base.

キーワード:画像データベース,キャラクター,日本古典文学,絵本
Keywords: Image-database,Character,Japanese classics literature,Picture books

1.デジタルアーカイヴにおける「古典」
1.1.記憶の中の文化—放送コンテンツ—

「古典」とは何か。この問いは、デジタルアーカイヴとして、後世に残し得るべきコンテンツとは何かという問題とも共通する。

たとえば、放送コンテンツを例にあげて考えてみよう。

かつてテレビ放送業界では、ビデオテープの方が貴重なため、生放送の記録が残されないのはもとより、録画番組でさえも、ほとんどが製作される端から消磁されていた。そのため、1970年代までのNHK大河ドラマや連続人形劇などの放送番組では、総集編しか保存されていなかった。稿者が幼少期に視聴したテレビ人形劇"[1]なども、そのほとんどは記憶の中に埋没してしまっている。

しかし、現在は、発信メディアの多チャンネル化と、情報機器の性能向上によって、安価に潤沢なサイバースペースを確保することが可能となった。そのため、コンテンツをとにかく多く保有することの方が重視されるようになり、蓄積・記録する対象を選別・吟味する意識はやや希薄になりつつある。

ユーザ側でも、すでに民生機はアナログ地上放送の全チャンネルの同時録画が可能になりつつあることからもわかるように、情報コンテンツは、「選別してから入力・蓄積する」という方針から、まずは「データを貯めておいてから選別する」という方向へと大きく変わりつつある。そのような状況の変化は、デジタルアーカイヴ作成対象についての価値観や選別意識、特に「古典」観というものについても、変容を生じさせつつあるようである。

1.2.古典とは何か

近年は「評価・選別」という基準を度外視して、徹底的なデジタル化、アーカイヴ化が果たされるべきだという意見が聞かれるようになった。

「可能な限り何でもかんでも保存すべき」という選別基準は、一見無定見な見識にも思われる。しかし、長い歴史の中で価値観はどのようにも推移するものである。たとえば、それまで一見無価値にも思えたものが、新たなパラダイムを獲得することによって、俄然宝石のような輝きを発し始めることは、これまでにも頻発しているものである。

かつて、科学研究補助費・特定領域研究「古典学の再構築」(1999.04 ～ 2003.3)にちなんだ『文学』(岩波書店,2000.7)で、「転換期の古典」というミレニアムを意識した特集が組まれたことがあった。

特集冒頭には、「古典とは何か」ということを考える、東野治之氏[歴史学(上代)]と小峯和明氏[日本文学(中世・説話文学)]の 2 本の論文が配されたが、論中に提示された古典観は全く対照的なものであった。

まず、東野氏は、

> 古典とは、その書物が単に古いというだけでなく、時代をこえて読みつがれ、生活の規範とされたものでなければならないとされたのは、吉川幸次郎氏であった。その意味で、日本の『古事記』や『日本書紀』を古典というのは当たらないとも言われたという。これはまことにもっともな定義である。『古事記』は、本居宣長が注目するまで、ほとんど忘れられた書であったし、『日本書紀』も、その神代巻のみが神道家に尊重されたものの、それ以外はやはり近世に入るまで、かえりみられることがなかった。この点、日本の伝統社会で真に古典と呼べるのは、漢籍、とりわけ経書であったといってよかろう。
>
> (東野「日本古代の『春秋』受容」,文学 1-4,pp1-10)

と述べ、対して小峯氏は、

> いったい、古典はいつから古典になったのか、先験的な与えられた聖典意識がおのずといくつかの古典を特権化し、その認識にもとづいて研究も進められ、学会も組織されてきた。
>
> (小峯,,説話の輪郭—説話学の階梯・その揺籃期をめぐる—,文学 1-4,pp11-20)

という認識に立つ。古典とは、作り意味づけられるものであるという前提に立つのである。これは、日本文学自体が中国的価値観とは異なる独自性を主張する長い営みの上に成り立ったものであるという事情を汲んでの見解である。

2.日本人の和漢

2.1.日本人の劣等意識

1878 年(明治 11)、第 3 回パリ万国博覧会で、最も勝利をおさめたのは日本美術(ジャポニズム)だったといわれている。しかし、当時の日本人は、自国の作品をくだらないと思う日本人の劣等意識は払拭できていなかったらしい。このことについて、フランスにおける第 1 期の大蒐集家の一人に数えられるエドモン・ド・ゴンクール(Edmond de GONCOURT)は、次のように書いた。

> 〔日本人には、日本の芸術〈特に日本陶器〉は中国のものであったという〕この真実がしみこんでいて、中国を彼らの芸術の故国、美しきものの到来する処、彼らにとってのギリシアと考えることが習慣になっている。ある日、成島〔内務官僚の成島謙吉〕に、お国の人々の間にも蒐集家はいますか、と私が尋ねると『確かに』と答えた。次いで彼は、自国の芸術にたいする翻訳できない軽蔑のニュアンスをこめて付け加えた。『でも、わが国の蒐集家たちは、中国の作品しか持っていません』と……」
>
> (Edomond de GONCOURT,*La maison d'un artiste,Paris,*1881[*2])

シャン・ド・マルス会場の、各種目の展示場や各国の陳列場が立ち並ぶ中、日本館はどぎつい装飾の中国館の隣にあり、簡潔な建築はその気品を際だたせることになった。実に、日本館の成功は、中国館の隣に位置した幸運のせいだともいわれているのであった。

全体、近代以前の日本文化は「漢」の世界を至上とする価値観に覆われていたといっても過言ではない。たとえば、仮名書きの書き物に対して、奥書・序文などに「幼童」「童蒙」のためという類の表現が多用されるのは[*3]、そうした意識の現れといえよう。漢字文化の世界は、異国のものというよりも渾然一体となって根底意識に染みついていた。そのような環境化、「和」的世界が「漢」に優越して評価されたことは、相当に画期的なことであったのである。しかし、それらが心性として「漢」的価値観と

完全に切り離された「和的」価値観の優位性を獲得するには、近代に入ってからもまだ、相当の年月を要したようである。

2.2.日本のキャラクターの定着度

日本における「和」と「漢」のありようを考える際に、もう一つ気にとめておくべき現象がある。

たとえば、諺や、故事成語・詩句に使用される表現に、日本人の名前が現れることは、日本文化に「和」的世界が定着することと、関連するといえよう。ところが、意外にも人口に膾炙する和漢の秀句が選ばれた 11 世紀頃成立の『和漢朗詠集』には、日本人名は登場しない。平安時代に編纂された、諺・金言集と呼ばれる類の『世俗諺文』『明文抄』『口遊』などにおいても同様である。

古典時代において、人の口を突いてでる慣用句に日本人が意識されることは、ずいぶん後代になってからと予想されるのである。その意味において、『平家物語』冒頭などは、日本人名が慣用的に定着した現象の発露としての極初期作品とも言い得よう[4]。

3.日本の昔の有名人

3.1.何をどうやって調べるか

前近代日本を中心に、日本人で有名だった人には誰がいたのだろうか。そして、その結果がわかれば、現代の我々の意識とどのように異っているかということも明確になろう。そのための方法・尺度として、どのようなものが有効だろうか。

悉皆的な調査方法としては、日本人の名前が、文化的著作物の中にどれだけ出現するかというランキングをはかることがまず考えられよう。今でいえば、Yahoo[1]などの検索エンジンに出現する人名のヒットランキングを、長期的なスパンで観察するという方法である。

この方法は、網羅的であるという点で、かなりの信頼性が期待されよう。しかし、どれだけのサンプルあれば有効であるかが予想しづらい。

[1]Yahoo!カテゴリ

後者の場合は、過去に遡っての調査ができないという欠点がある。しかし、将来的には文化動向の定点観測を継続的に行うことによって、最も有効なツールとなりるだろう。(ただし、その実現のためには、まず事前に、どのような人名がリストアップ可能かという事前のデータを整えておく必要がある。)

このように、尺度となり得るものには、いろいろなものが考えられよう。たとえば、ほかにも人形や、彫刻、絵画など、文字に現れたものだけに限らない、さまざまなジャンルのメディアにおいて、どのような現れ方をするかなど、さまざまなアプローチが可能である。

およそ、文化動向の定点観測を行うにふさわしい対象の特徴としては、

①多数の中から精選された結果が集約されたもの

②その時々で、人口に膾炙していることがはっきりしていること

などがあげられよう。これらの条件が満たされれば、有効性が保証されるであろう。

3.2.叢伝

和漢古典籍には、その人物の生き様や特徴の観点から類似する人物を集成した列伝を 1 タイトルの書物にまとめた作品が少なくない。「(人物) 叢伝」と呼ばれるジャンルに分類される。

叢伝は、基本的に多くの書物から取材された人物の伝記、奇事、逸事をなどをさらに選別・抽出した形で作られる。その意味で、その編纂意識には、その編纂当時において、「この人物だけは記録したい」といった類の意識が働いていたと思われる。その意味において、有名人という観点での取材源として最適な特徴を持つが、しかし、その総体は必ずしも判然としてはいない。『国書総目録 補訂版』や『古典籍総合目録』にもとづくそれぞれのデータベース化においても、書籍の主題分類は未整備の状態にある[5]。

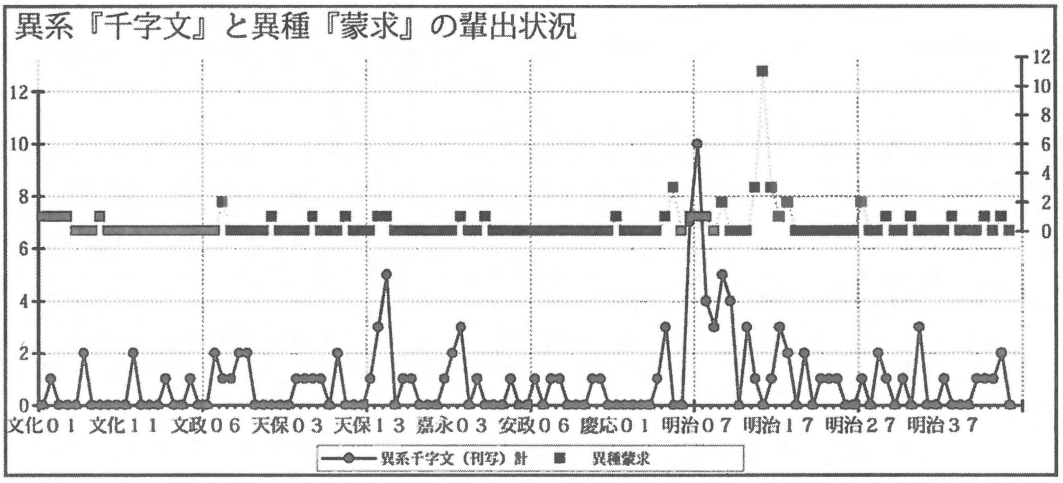
この類の書物は、中世以前のものは多くない。たとえば試みに、『本朝書籍目録』(編者未詳・鎌倉末期成立) 中から叢伝に含め得る典籍を抽出してみても、そのタイトルは決して多くないことが確認される[2]。しかも、佚書となったものも少なくない。

[2]本朝書籍目錄中に見える「叢伝」的典籍（和田英松『本朝書籍目錄考證』より）			
1.神事	神別記	10巻	佚書。新井白石『紳書』に「神別記に、功徳ありし人々をまつりし事をしるせしなり、今は此書なきなり、」
	和漢春秋	不明	大外記師弘撰。佚書。
16.人々伝（人之伝[刊本]・人伝[神宮文庫本]）	儒伝	3巻	佚書。
	摂関	2巻	忠仁公[良房]～安祥院[吉忠](天正九上巻奥書)
	大臣	1巻	佚書。
	大将	1巻	佚書。
	本朝神仙伝	1巻	大江匡房撰。現存。
	菅家	1巻	群書類従本系現存。
	日本儒林	1巻	菅原大江二家、その他の伝。佚書。
	故賢	1巻	佚書。
18.雑々	高名録	1巻	江帥抄。『河海抄』などに引用あり。佚書。
	名所抄	1巻	佚書。地名・居所に関するものだが、中国の『皇覧』の例もあるように、その地に関する人物に関する記述もあった可能性も高い。
	居宅抄	1巻	
19.雑抄	『打聞』『古事談』『続古事談』など		

ほかにも、『○○蒙求』のような唐の李瀚撰『蒙求』の体裁に倣った続撰本（国書で約70種存在『○○蒙求』の書名が多く約500名を掲載）や、『○○二十四孝』など、幼学書の伝統を引き継いで、その編纂主旨を書名に反映させた漢文で書かれた叢伝の類もある。[3]に幕末以前に輩出された、日本人を扱う『蒙求』型人物叢伝の一覧を示したが、この類の書については、ほぼその総量が判明しており、その数は決して多くはないが、幕末から急増する輩出状況を見せている[4]。

また、叢伝の中には、絵の配されたものもある。『蒙求』型人物叢伝の多くが漢文により記されたものが多いのに対して、絵入人物伝は、仮名書きのものも多い。この類の典籍類は、現代でいえば漫画・絵本のように、初等教養段階か低年齢の読者を想定して作られたものが多い。そのため、たとえば「百人一首」の続撰本のように、文字だけで構成された人物総伝よりも厳選され、絞り込まれた人物を扱うものが多いという特徴を持つ。

3.3.絵と人物伝がセットになったもの



[4]唐の李瀚撰『蒙求』の体裁に倣った続撰本の幕末期の輩出状況〔転載元：幕末・明治期の「蒙求」,相田満,国際日本文学研究集会会議録第18回(1994),国文学研究資料館,pp109-133(25)〕

[3]日本人を扱う『蒙求』型人物叢伝の輩出状況(幕末以前)						
参考:相田,『蒙求』型類書の世界,相田　満/和漢比較文学会編『和漢比較文学の諸問題』8/和漢比較文学叢書/(汲古書院)/pp109-128(20) 参考:相田,長崎浩斎(健)・東条琴台(耕)の『千字文』『蒙求』目録—付・日比谷加賀文庫蔵『補訂浩斎所蔵千文蒙求二種書目』翻刻—,相田　満/中野猛:編『説話と伝承と略縁起』94/新典社研究叢書/(新典社)/pp222-244(23) 参考:相田,幕末・明治期の「蒙求」,相田　満/国際日本文学研究集会会議録第18回(1994)/(国文学研究資料館)/pp109-133(25) ※日本人を扱わないため、除外したものもある。たとえば、『蒙求続貂』(恩田仲任)・『医林蒙求』(樋口好古)など。 ※人物を扱わない、入門・初学書としての『蒙求』は除外した。たとえば、『俳諧蒙求』(堀麦水)・『吹寄蒙求』(銅脈先生・偏屈道人)・『国字蒙求』(伊藤有隣[全節])。						
扶桑蒙求	不明	永済	1239	暦仁2	著者存命	尊経閣文庫蔵『書札類稿』所収の元禄13年(1700)の加賀宰相(綱紀)より水戸の増田寿得にあてた書状中に、この書名がある。『聖徳太子玉林抄』巻1、巻1、『上宮太子拾遺記』第7などに逸文があり、『玉林抄』巻19の逸文は「大津初詩　素盞創歌」との対標題が記される。彰考館蔵に『扶桑蒙求私註』があるが、これは巻六のみの零本で、しかも見出しが四字句標題を採らず、釈氏のみを扱う。その意味で、『扶桑蒙求』と「私註」との関連にはやや微妙なものがあるが、無関係とも言い切れない。『和談抄』と合冊で、彰考館蔵。
伝法蒙求	3巻2冊		1642	寛永19	刊	駒沢大蔵。
釈書蒙求		祖寛	1673	延宝元	序	『元亨釈書』より取材。国会図書館に祖本の1冊本があるほか、延宝4年版1冊本が高野山大・東大・東北大などに、元禄10年(1697)版の5巻4冊本が大谷大・大正大・龍谷大などに蔵せられる。
本朝蒙求	3巻3冊	菅亨(中徹)	1679	延宝7	序	貞享3年(1686)刊。文庫などに蔵せられる。巌谷小波編『大語園』の取材資料にもなる。明治10年(1877)の刊本もあるように、息長く享受された。
桑華蒙求	3巻3冊	木下公定(葵峯)	1710	宝永7	序	外題「新撰自註桑華蒙求」。日本と中国の故事を対にする。無刊記本のほか、天保15年(1839)木活字本、福田宇中箋注の明治15年(1882)刊本などがある。引用書目録を付すものもあるが、これは後人の附会。岸鳳質『扶桑蒙求』は、この書の日本の故事の部分から多く採る。
医学蒙求	3巻3冊	伊東見龍	1742	寛保2	序	寛保3年刊(1743)。首巻は標題を陰刻して独立させる。歴代の医人列伝。
蒙求拾遺	3巻3冊	大江(桂)広保	1749	寛延2	序	宝暦2年(1752)刊。
医家蒙求	不明		1754	宝暦4	輩出	『江戸時代書林出版書籍目録集成』収載の『宝暦四年書籍目録』記載。
平安蒙求	不明		1772	明和9	輩出	『江戸時代書林出版書籍目録集成』収載の『明和九年刊書籍目録』記載。
俳諧和漢蒙求	1巻1冊	千秋館大虎	1780	安永9	序	人物名二字を対比しているが、四字標題ではない。東京国立博物館、香川大蔵。
日本蒙求	3巻	恩田仲任	1780	安永9	前後	巻頭「隷事」、外題「日本蒙求」。日本の故事より取材する。内閣文庫に自筆本を蔵するほか、無窮会図書館織田文庫などに蔵せられる。
扶桑蒙求	3巻	岸鳳質(根岸嶸谷)	1818	文政元	序	天保14年(1843)刊。明治4年(1871)の後刷もある。日本の故事から取材するが、木下葵峯『桑華蒙求』から日本に関連する故事を節略したものが目立つ。
大東蒙求	1巻	岸鳳質(根岸嶸谷)	1818			山岸徳平論文による。写本。未見。
扶桑禅蒙求	3巻	岸鳳質(根岸嶸谷)	1818			『補訂浩斎所蔵千字文蒙求二種書目』記載。
扶桑釈氏蒙求	3巻	岸鳳質(根岸嶸谷)	1818			『補訂浩斎所蔵千字文蒙求二種書目』記載。
国史蒙求標題	3巻4冊	衣笠孝郷	1825	文政8	序	京都大学蔵。

[5]Web公開版「歴史人物画像データベース」

そこで、こうした特質に着目して、古今の人物群から取材された典籍で、絵入りのものを対象として、集中的なさらに調査を行うならば、「日本人が知っていた本当に有名な人物は誰か」ということを証するための有効な指標が求められるという道筋は成り立ちそうである。

「歴史人物画像データベース」は、こうして構想されたものであった。

4.歴史人物画像データベース

4.1.概要

「歴史人物画像データベース」は、古典籍に描かれた挿絵の人物画像（主に日本人）を集積したものである。取材源は国文学研究資料館所蔵本にかかり、明治以前の国書古典籍刊本に収められた人物画像を切り集め、それぞれの人物が各典籍でどのように描かれてきたかを比較対照できるように配慮している。

2005年度中公開を進めているものとして、

①高精細画像による Web 公開版[5]（648 名 898 画像収録[9 作品より取材]）
http://base1.nijl.ac.jp/~rekijin/syouzou.html

②検索システム付 DVD 版[6]（3,168 名分 4,740 画像収録[59 作品より取材]）

の2種類を予定している。①の高精細画像版は、TIFF 形式の画像を ZOOMA 形式フォーマット（ドリームテクノロジ社）に変換することにより、画像の高速表示を実現し、画像の無断引用・無断転載を防ぐためのプロテクトをかけたものである。

②のデータベースは DVD-ROM に格納されたスタンドアロン環境で動作するが、異体字、異称、暦日検索など、高次の検索を可能としたシステムが同梱するほか、人物に関する詳細情報として、『芳賀人名辞典』『地下家伝』をデータベース化したデータを搭載する。また、表示に使用される画像は Jpeg 劣化画像を使用する。

4.2.取材典拠作品の特徴

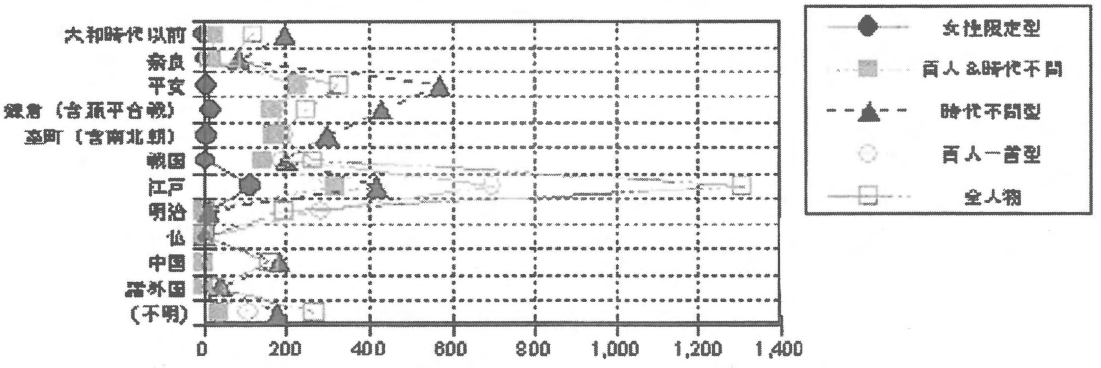
検索システム付 DVD 版に使用した 59 作品の内訳は次にあげるとおり。

英雄百人一首・百人一首図絵・若鶴百人一首・武家百人一首・役者三十六家選・粋興奇人伝・続英雄百人一首・近世名婦百人撰・列女百人一首・本朝孝子伝・秀雅百人一首・万家人名録・畸人百人一首・義烈百人一首・絵入赤穂四十七士詩伝・徳川十五代記・武勇魁図会・本朝百将伝・狂歌百人一首・新編歌俳百人撰・大日本国開闢由来記・本朝列仙伝・前賢故実・先賢遺芳・先哲像伝初輯・八幡太郎一代記・利運談・宮本武勇伝・岩倉具視公之実伝・明治英銘百詠撰・現今英名百首・絵本威武貴山・絵本武者備考・明治中興雲台図録・明治中興凌煙図録・源義経一代記・絵本武蔵鐙・和洋奇人伝・今人名誉百首・徳川東国武勇伝・神君長篠大合戦・玉石雑誌・絵本武勇伝・豊臣昇雲録・好古事彙・源平盛衰記図会・絵本故事談・絵本忠経・絵本徒然草・絵本堪忍記・円光大師御一代記・（弘法大師行状記）・武稽百人一首・俳人百家撰・近世報国百人一首・訓蒙皇国史略・（集古十種）・都絵馬鑑・絵本倭比事・三国七高僧伝図会・神風遺談〔計 59 作品〕

[6]検索システム付DVD版

[7]DVD版59作品の典籍の特質

種別	時代不問型	時代限定型	百人一首型	女性限定型
書名例	英雄百人一首 百人一首図絵 若鶴百人一首 武家百人一首：他	近世名婦百人撰 万家人名録 絵入赤穂四十七士詩伝 徳川十五代記：他	英雄百人一首 百人一首図絵 若鶴百人一首 武家百人一首：他	近世名婦百人撰 列女百人一首
件数	32/59	25/59	19/59	2/59
比率	52.5 %	41.8 %	31.1 %	3.3 %



[8]　時代別・典籍の特徴別の出現数比較

これらの典籍の成立年は、刊年の判明するものだけに限れば 1686 年(貞享 3)～ 1893 年(明治 23)成立のもので占められ、不明のものもすべて江戸時代以降にかかる。

これら 59 作品を、それぞれ収録される人物の集まり方から 4 つのタイプに分けて分類したものが[7]で、さらにそこから、登場人物の時代別に分類し分析したものが[8]で、いずれも、扱われる人物の時代は、江戸期・平安・鎌倉・室町の順に多いということがわかる。

また、さまざまな時代から取材された人物で構成された典籍（時代不問型：▲印）には、平安・鎌倉の人物が繰り返し現れることが多く、結果的には、そうした人物が「古典的な有名人」として、人々の記憶に残されていくことがわかるのである。

5.1.集計

では、「日本の古典的な有名人」としては、どのような人が意識されてきたのだろうか。

調査を進めるに際して、まず DVD 版を作成する際に使用した 59 作品の中から、分析対象とするにふさわしい性格を持つ典籍を抽出した。

具体的には、女性だけを集めたものとか、百人一首型や収録人物の時代が限定されている作品などは、収載傾向に偏りがあるため、分析の対象外とすることとした。その結果、さまざまな時代から集められた人物が集約された「時代不問型」の絵入叢伝が分析対象となり、31 作品、延べ 1,748 名が、分析作業の対象となった。

そして、その結果、出現頻度上位 29 位以内に入った人物と、調査対象作品の内訳は、[9]で示すとおりとなった。

5.日本の昔の有名人は誰か

[9]時代不問型典籍中に現れる人物群

1/菅原道真(17)
2/聖徳太子(15)
3/曾我時致(14)
4/源義家(13)
5/楠木正成(13)
6/足利尊氏(13)
7/西行(12)
8/武内宿祢(11)
9/熊谷直実(11)
10/源義経(11)
11/柿本人麻呂(11)
12/法然(11)
13/源頼朝(9)
14/豊臣秀吉(9)
15/平清盛(8)
16/楠木正行(8)
17/日本武尊(8)
18/藤原鎌足(8)
19/源満仲(7)
20/源頼政(7)
21/梶原景季(7)
22/曾我祐成(7)
23/新田義貞(7)
24/小野篁(7)
25/在原業平(7)
26/足利義満(7)
27/武蔵坊弁慶(7)
28/坂上田村麻呂(7)
29/源信(7)
※[( )内は出現頻度を示す件数]

【調査対象と した作品】

英雄百人一首,百人一首図絵,武家百人一首,粋興奇人伝,続英雄百人一首,列女百人一首,本朝孝子伝,畸人百人一首,義烈百人一首,武勇魁図会,本朝百将伝,狂歌百人一首,新編歌俳百人撰,本朝列仙伝,前賢故実,先賢遺芳,先哲像伝初輯,絵本威武貴山,絵本武者備考,利運談,絵本武蔵鐙,絵本堪忍記,絵本忠経,絵本倭比事,訓蒙皇国史略,三国七高僧伝図会,集古十種,玉石雑誌,都絵馬鑑,和洋奇人伝,絵本故事〔計 31 作品〕

[10]Web公開版のGoogle人気ランキング

1/織田信長[245000]
2/豊臣秀吉[217000]
3/聖徳太子[182000]
4/源頼朝[106000]
5/源義経[105000]
6/紫式部[104000]
7/法然[83700]
8/菅原道真[76800]
9/伊達政宗[73000]
10/日本武尊[62800]
11/平清盛[52200]
12/道因法師[51300]
13/足利尊氏[42000]
14/清少納言[39000]
15/小野小町[37700]
16/天智天皇[37600]
17/三遊亭円朝[31300]
18/平将門[31000]
19/持統天皇[28500]
20/和泉式部[27200]
21/松本幸四郎[27000]
(※含現代人)
22/紀貫之[24200]
23/西行法師[23400]
24/大伴家持[23200]
25/藤原定家[20800]
26/武蔵坊弁慶[22700] [6]
27/在原業平[21400]
28/柿本人麻呂[21300]
29/足利義満[20900]
30/柴田勝家[19300]

上位にあがった人物の内訳を見ると、中世以来盛んに行われるようになった信仰の対象と深い関わりがあるものが目立つ。

こうした評価法によって浮かび上がる人物群に共通することは、さまざまな物語伝承を包含するだけでなく、中には信仰の対象となるに至るほどに強烈な個性を持っていることである。

まず、1 位の菅原道真は天神信仰、2 位の聖徳太子は太子信仰、4 位の源義家は八幡信仰という具合にである。また、3 位の曾我時致は曾我兄弟の兄五郎の方で、曾我兄弟の話は、若くして非業の死を遂げた曽我兄弟への同情が、やがて鎮魂から信仰に変わり、伝説化するとともに、説経の題材として室町期以降、正月に決まって扱われるのみならず、歌舞伎においても、いつしか「吉例」として「曽我物」が正月公演の祝儀物として定番化するに至っている。

そして、そうした個性が魅力となり、その時々の人々の好尚にかなうように、再解釈が繰り返されてきたということであろう。

一方、日本の紙幣の絵柄に使われた人物のランキングでは、1 位に聖徳太子[7 回]、2 位に菅原道真と和気清麻呂が並び[6 回]、4 位・武内宿禰[5 回]、5 位・藤原鎌足[4 回]、6 位・神功皇后[3 回]などと続き、菅原道真や聖徳太子が上位者に並ぶ点で共通する[7]。

このようなことを考えあわせると、この一覧が日本人の心性に深く根ざしたものであることが首肯されてくるのである。

6.現代人との意識差

これらの人物群が、現代人の感覚とどのように異なるのだろうか。そのことを明らかにするためには、これまでは、アンケートを主とした社会調査・教育調査が必要であった。しかし、近年では Web 検索エンジンのヒット件数を利用した調査でも効力を発揮するだろう。

参考までに、高精細画像による Web 公開版[5](648 名 898 画像収録[9 作品より取材])を使用して Google 検索にて件数を求め、その件数順に再配列したものが[10]の一覧である。

ここで、留意しておかねばならないのは、時代不問型の総集では、同時代の人物はどうしても上りづらくなるという傾向があるということである。

加えて江戸時代には、時の為政者を直接に描いたりすることが憚られたため、上記調査には上りづらいこともある。たとえば、東照権現に祀られた徳川家康は 261,000 件にも上るが、それがデータベースに上がらないのは、近世期の刊行物において図像化が憚られた事情による。こうした人物画像が多数上るようになったのは、明治期以後の刊行物にかかることである。

7.情報ナビゲーションの必要性

昔の有名人はいったい誰で、各時代や時期では、どんな人物が好まれていたかという問いに、明確に応え得る指標はこれまで皆無に等しかった。現代ならば、アンケートなどを通しての標本調査なども考えられようが、そのためにかかる手間も費用も莫大なものとなる。ところが、最近ではこのように Web 検索エンジンのヒット数という、新たなツールを手に入れることができ、これまで不可能だったことが、発想次第で興味深いデータを入手することができるようになっているのである。しかし、それとても過

去に遡っての調査は不可能である。

このような指標は、学術面においては、その当時に通底する哲学・イデオロギーを把握するという、文学・歴史・民俗・教育学などの分野で意義ある情報となるだろう。また、「古典的有名人」の人気とその言説とには相関関係があるところから、文学的研究の指針としても有効で、さらにはエンターテイメント産業におけるコンテンツの掘り起こしや、地域振興の新たな素材ともなり得るものといえる。

デジタルアーカイブスが陸続と発信される現在、その情報量は激増の一途をたどっているが、その現状は、ただ紙から電子データへと媒体の形態を変えただけのものにとどまったにすぎず、ユーザ側はかえって膨大な情報の選択肢に押しつぶされかねない状況にある。

潤沢なデジタル資源から、いかに情報を読み解き、説明するかという、データマイニングをサポートするための枠組みは、ユーザからの発案を待つよりも、発信者側のコンセプトに依存する。特に初等・中等教育においては古典学に関わる素養の減衰が、ますます顕著となってきている現在、想定されるべきユーザ像自体が見直されるべき時にさしかかっていることが痛感されてならないのである。

8.おわりに―〝人物〟から〝キャラクター〟へ―

最後に、本論において、標題に「キャラクター」をあげながら、本文中のほとんどでは「人物」という用語を使用してきたことに関して述べておこう。

印刻を意味するギリシア語に由来するキャラクター(character)は、一般的に性格・性質という意味を持つが、そこから派生して、今では狭義の演劇・小説・漫画・アニメなどの登場人物（の役柄）を指す概念が一般的になっている。そして、今やその言葉は、あらゆる創作活動の原点と認識され、現代のクリエータ達にとって作品の出来を左右するものは、キャラクターの成否如何にかかるという発想も常識と化し、キャラクター商品を始め、作品世界を飛び越えた愛好者と、巨大な経済的価値を生み出すものへと変貌し、日本から世界に発信される主要な文化コンテンツともなっている。

キャラクターの重要性を打ち出したものに、小池一夫氏の主導になるキャラクター造形学科（大阪芸術大学）の試みがあるが、文学、とりわけ日本古典文学の世界においても、享受の実態に則してみれば、さまざまな作品はキャラクターに従属するといっても過言ではない。ところが、不思議なことに、これまでの日本の文学研究、とりわけ古典分野の研究においては、こうした現象を受け止める素地、すなわちキャラクターを主体とし、積極的に評価する研究体系が構築されることはなかった。また、研究論文も皆無に近い状態にある。

その最大の理由としては、文学以外のメディアに現れるキャラクターの評価軸・分析のための観点が定まっていないことがあげられよう。絵画・演劇・伝承・信仰など、さまざまな場で共有・変容して現れても、それまで時代・媒体（メディア）などに立脚して思考されることの多かった研究の場では、対応しきれなかったためである。

ある典籍が、内容とは別次元で、尊崇され、特別な扱いを受けるに至る現象を聖典（カノン）化と呼ぶが、人物造形の生成と享受の様相は、まさに現代的な意味におけるキャラクターに異ならず、それを扱う研究主題は、ある意味で、聖典の対極に位置するものかもしれない。文学史に対するキャラクター史というものも、本来ならとうに構想されるべきであったろう。

今や日本人の生活は、ずいぶんと古典や伝統と縁遠いものになっている。しかし、それに対して中国・台湾などでは、伝統的キャラクターと新渡の日本発のサブカルチャに由来するキャラクターとが混在して享受される風景が、あちこちで見られる。キャラクターを共有するということは、世代や地域をつなぐことのできる共通の話題や教養を共有することであるといっても過言ではない。

そのような状況下、本稿で述べたようなデータベースの試行版が、「人物　画像」の掛け合わせによるGoogle検索で2位を、「人物　画像　データベース」の掛け合わせ検索では1位にランクされている[8]。そうした実績を勘案するにつけ、今後このデータベースが果たす役割は、ますます重みを増していくことだろう。

9.補記

本論は国文学研究資料館文学形成系における研究プロジェクト「古典形成の基盤としての中世資料の研究」による研究成果の一部である。記して謝意を伝えたい。

*1 そうした中、海外吹き替え版が発見されたり、『ひょっこりひょうたん島』のように、当時小学生だった伊藤悟氏が５年間の放送すべてをノートに詳細に記録、残してあったテープと照らし合わせてリバイバル作品が復活したという、きわめて希有かつドラマチックな例もある。「伊藤悟のひょっこりひょうたん島漂流」（http://www009.upp.so-net.ne.jp/hyoutanjima/index.html）

*2 木々康子,林忠正とその時代―世紀末のパリと日本美術―,第二章 万国博覧会と日本美術,一八七八年パリ万国博覧会,pp29-37,筑摩書房,1987

*3 たとえば、名所図会のはしりである『都名所図会』凡例など。

*4『平家物語』冒頭部分は以下のとおり。

『祇園精舎の鐘の声、諸行無常の響きあり。沙羅双樹の花の色、盛者必衰の理を彰す。奢れる人も久しからず、只春の夜の夢の如し。猛き者も終には亡びぬ、偏に風の前の塵に同じ。

遠く異朝を訪らへば、秦の趙高、漢の王まう、梁のしゅうい、唐の禄山、これらは皆旧主先皇の政にも従わず、楽しみを究め、諫めをも思い入れず、天下の乱れんことを悟らずして、民間の憂ふるところを知らざっしかば、久しからずして、亡じにし者どもなり。

近く本朝を伺ふに、承平の将門、天慶の純友、康和の義親、近く本朝を窺ふに、承平の将門、天慶の純友、康和の義親、平治の信頼、此等はおごれる心もたけき事も、皆とりどりにこそありしかども、まぢかくは六波羅の入道前太政大臣平朝臣清盛公と申し人のありさま、伝い承るこそ心も詞も及ばれね。』

*5 相田満,和漢古典学のオントロジの資源化のために―『国書総目録』の分類について―,ナオ・デ・ラ・チーナ（東アジアの出版文化ニューズレター）8,(文部科学省科学研究費特定領域研究「東アジアの出版文化」（東北大学東北アジア研究センター内）),pp105-112(8),2004

*6「弁慶」のみの場合は、214000 を数えたが、店舗名・商品名等が多く混じったため、数字としては採りあげなかった。

*7http://www.77bank.co.jp/museum/okane/02.htm

*8 相田満,「歴史人物画像データベース」構築奮闘記,情報知識学会誌 15-2(May 2005),pp7-14(8),2005

浮世絵と写真に基づく幕末から明治に至る女性の形態の定量的考察

Quantitative Study on Body Shapes of Women from the Edo Period through the Meiji Era Based on Ukiyo-e and Photographs

森下あおい*　　黒川隆夫**
Aoi morishita　　Takao kurokawa
*滋賀県立大学 人間文化学部　彦根市八坂町
University of Siga Prefecture,　Hassaka-cho, Hikone,
**京都工芸繊維大学大学院　工芸科学研究科　京都市左京区松ヶ崎
Graduate School of Science and Technology, Kyoto Institute of Technology
Matsugasaki, Sakyo-ku, Kyoto

集団での正確な人体計測が行われるまでの過去の人の形態を把握するには，絵画や写真は重要な手掛かりとなる．本稿では，明治期に撮影された写真を対象として，そこから推定される女性の人体寸法を計測し，そのデータベースの構築と人体形態の定量的化を試みた．得られた計測値を江戸後期，幕末～明治期において浮世絵に描かれた女性の形態の計測値，および現代女性の人体計測値と比較し，その推移と形態の特徴について考察を行った．その結果，幕末から明治期の浮世絵には現代人に近い特徴が描かれていることなどを指摘した．

Paintings and photographs offer vital clues towards estimating the body shapes of people in bygone eras when no valid measurements could be performed in order to determine their torso parameters in both a precise and collective manner. In this paper, attempts were made to construct a database and quantify the body parameters of women who appear in photographs taken in the Meiji Period, by estimating and measuring their body shapes and sizes from this data source. The values obtained were subsequently compared with the torso shape measurements of the female models depicted by Ukiyo-e paintings in the latter half of the Edo Period and in the period from its last days through the Meiji Era; also with the measured data on the body shapes of women in the modern days. The comparison results were then analyzed, in order to clearly elucidate temporal changes in the torso shapes of women of these different times, as well as their respective body characteristics. Based on the results of this investigation, the author points out that while the Ukiyo-e models from the end of the Edo Era through the Meiji Period were painted as figures characteristically similar in some portions to the body shapes of modern women.

キーワード：浮世絵，写真，人体形態，定量的分析．
Key words: Ukiyo-e, Photograph, Body shape, Quantitative analysis.

1．はじめに

筆者らは人体の計測データのない江戸時代，および明治時代の日本人成人女性の体形と衣服着衣形態について，美術作品，服飾資料，写真資料をもとに定量的に分析している．

明治時代以前の人の体形については，人類学研究で人骨を用いた種々の推定[1)]が行われているが，女性の体形については標本数も少なく，多くは解明されていない．明治時代になると，文部省が1900年から行った体格調査のデータがあるが，対象が子供であり計測項目も身長，体重，胸囲，座高に限られている．従っ

て，正確な成人女性の体形は，衣服サイズ設計のために戦後漸く始められた全国規模での人体計測におけるデータによって把握することになる．

一方，江戸時代の絵画作品や明治時代の写真には女性の体形に関する資料が数多く残されており，それらを資料として活用し定量化することでデータベース化が可能である．

筆者らはこれまでに，江戸時代に浮世絵に描かれた女性の人体形態と小袖着衣形態の計測を行い定量化した[2) 3)]．小袖着衣形態の定量分析の結果では，浮世絵に描かれた小袖着衣形態は，当時の小袖の形態寸法をほぼ反映した描き方が行われており，浮世絵の着衣形態はかなり写実的と見なせることを確認した．また1600 年～1900 年までの期間に描かれた浮世絵の人体形態を時系列で定量分析した結果，形態特徴が顕著に表出する人体部位やその時期を明らかにした．さらに江戸時代を代表する作者が描いた女性の人体形態の特徴から，同時期の作者や時期が隣接する作者間の類似性についても明らかにした[4)]．

本稿では明治時代に撮影された写真から，女性の形態を推測して描出し，その人体形態の計測を行い定量化を試みた．そして得られたデータを筆者らがさきに算出した，江戸時代～明治時代の浮世絵の計測データ，1951 年，1994 年の現代人女性の人体計測値と比較することで，幕末から明治期の女性の形態の特徴を考察した．

2. 資料および方法

2. 1 資料

資料は，明治時代に撮影された写真 30 枚[5) 6) 7)]で，いずれも着物を着衣した成人と考えられる女性の立ち姿である．比較のために用いた浮世絵資料は，筆者らがさきに計測した江戸時代～明治時代に描かれた浮世絵 76 枚から推定し人体形態を定量化した人体計測データである．人体計測の実測データは，1951 年の柳沢のデータ[8)]，1994 年の人間生活工学センターによるデータ[9)]を使用した．

2. 2 方法

写真の立位女性の人体形態を推定，描出するために，写真の着衣形態を慎重に観察した．それらは頭部や頚部，手，足など，露出している部位と着衣に完全に覆われた部位に分けられる．そこでまず，首や手足など人体が露出して形態の判明している箇所や，着衣が人体と密着に触れシルエットを表出している箇所を手掛かりに，内在ずる人体形態を推定し裸体を描出した．胴部から腰部の着衣によって形態が把握し難い箇所は，着物の着付けの観点から帯の位置に照らして適切と考えられる位置を推定し形態を描いた．この裸体の描出に際しては，着物の着衣形態と人体形態との関係に精通している画家の協力，指導を得た．図 1 はこのようにして写真から描いた人体形態の例である．

写真に写されている女性は何等かのポーズをとっているため，サイズの計測やプロポーションの算出が困難である．これらの困難を解決するために，本研究では女性の舞踊家をモデルとして 30 作品から抽出したすべてのポーズを再現してもらい，それらを人類学的計測の標準立位姿勢とともにデジタルカメラで撮影した．撮影に先立って，モデルのウエストの位置にベルトを着用させるとともに，肩峰点，頚前点，頚側点，乳頭点，前腋点，後腋点，臍点，腸骨稜点，脚付け根の位置で大腿部中点の位置，膝線の位置，内果点，外果点にマークをつけて後の計測のための基準点とした．モデルの正面からの直立状態と，ポーズ状態はカメラの距離などの撮影条件を一定に保った．

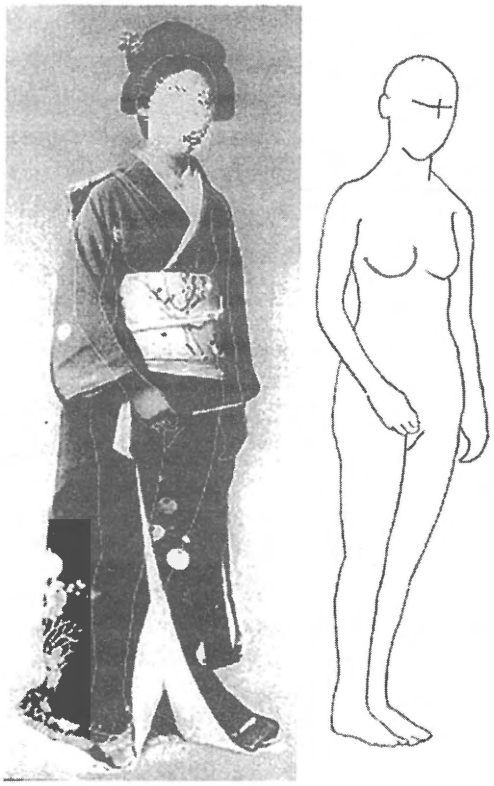


図 1　写真から描出した人体形態の例
写真資料[6)]（ポーラ文化研究所）

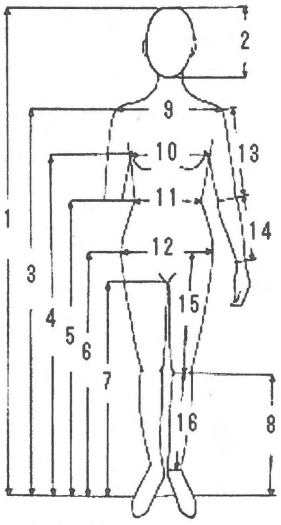


図２　計測箇所

表１　計測項目

NO	計測項目	
1	身長	高径
2	全頭高	
3	肩峰高	
4	乳頭高	
5	ウエスト高	
6	ヒップ高	
7	股下高	
8	膝関節高	
9	肩峰幅	幅径
10	胸部横径	
11	ウエスト幅	
12	ヒップ幅	
13	上腕長	四肢長
14	前腕長	
15	大腿長	
16	下腿長	

表１，図２に計測項目と計測箇所を示す．計測項目は，人体形状を把握するために有用であり，写真の着衣形態から計測可能な16項目である．計測は撮影した写真上で基準点および人体形態のシルエットを基に定規を用いて行った．

次にモデルによるポーズと正面立位で同一サイズをそれぞれ a_{ij} および b_{ij} とし（1）によって定義した．

$$P_{ij} = b_{ij} / a_{ij} \qquad (1)$$

ここで i（i=,...,30)は写真の作品番号，また j(j=,...,16)は計測項目の番号である．これらの比率 P_{ij} はポーズを取ることによって生じたモデルの人物像の各部位ごとの直立時の描写比を意味する．

また写真から描出した人体に関しても，計測用基準点の位置を推定し，モデルの写真と同様に各寸法 a'_{ij} を計測した．これらの計測値と(1)を用いて写真の人物が正面立位と仮定した場合の人体寸法 b'_{ij} を次の式を用いて算出した．

$$b'_{ij} = P_{ij}\, a_{ij} \qquad (2)$$

以下ではすべての人体寸法を身長 b'_{i1} に対するパーセンテージ，すなわち対身長示数 I_{ij} で表示することにする．

$$I_{ij} = (b'_{ij} / b'_{i1}) \times 100 \quad (j = 2, \ldots, 16) \qquad (3)$$

データベースには，制作年代に加え，実寸，示数の項目を設定し，データ取得とともにデータベースに入力した

3. 結果

3. 1　主成分分析による明治期写真の女性の形態

表２に計測項目について分散共分散行列に基く主成分分析を行った結果の各主成分の因子負荷量を示す．

主成分の解釈は，累積寄与率が66.95%の第４主成分まで行った．第１主成分では膝関節高が大きく正に負荷し，その他ヒップ高，股下高，下腿長，前腕長でも比較的大きく正に負荷している．一方，乳頭高，肩峰高は負に大きい．このことから第１主成分は人体下部の高さと四肢末端部の長さ成分と考える．第２主成分では高径項目がすべて正で大きく負荷し，特に上部で大きい．また胸部横径，肩峰幅，全頭高，および上腕長，下腿長が負に負荷している．これは人体上部の高さと幅を示す成分と考える．第３主成分では幅径項目がすべて正に大きく負荷し，乳頭高，前腕長，上腕長が負にやや大きい．これは体幹上部の幅の広さと上肢の長さの成分と考

表２　各主成分の因子負荷量

	第1主成分	第2主成分	第3主成分	第4主成分	第5主成分
全頭高	0.0406	−0.0633	−0.0159	−0.0794	0.1469
肩峰高	−0.3210	0.5276	0.1440	0.0693	0.0731
乳頭高示数	−0.4143	0.5080	−0.3543	−0.3490	−0.1057
ウエスト高	0.0900	0.2248	0.1924	−0.2518	0.1323
ヒップ高	0.2604	0.3995	0.1669	0.5644	−0.0798
股下高	0.2368	0.3204	0.1213	0.0073	0.0794
膝関節高	0.5078	0.2161	0.1463	−0.3267	−0.2041
肩峰幅	−0.0420	−0.0776	0.4555	−0.2903	−0.0650
胸部横径	−0.0467	−0.1084	0.4672	−0.1071	−0.0391
ウエスト幅	−0.0057	0.0212	0.2380	−0.1520	0.0335
ヒップ幅	−0.2931	0.0734	0.4306	0.0476	0.4800
上腕長	−0.1594	−0.0145	−0.1335	−0.0242	0.3096
前腕長	0.3422	0.0104	−0.2188	−0.0640	0.7462
大腿長	−0.1830	−0.0506	0.1110	0.4935	−0.0069
下腿長	0.2608	0.2712	−0.0817	0.0956	−0.0135
固有値	8.1545	6.1940	4.4382	3.6269	3.3687
寄与率	24.36%	18.50%	13.26%	10.83%	10.06%
累積寄与率	24.36%	42.86%	56.12%	66.95%	77.01%

える．第 4 主成分ではヒップ高と大腿長で正に大きい一方，幅径項目，乳頭高，ウエスト高で負に大きい．これは腰部の高さと体幹の幅の成分と考える．

このようにて明治期写真における女性の人体計測値を主成分分析した結果，第 1，2 主成分からは各々高さを示す因子，第 3，4 主成分からは体幹部の幅と四肢の長さを示す因子が示された．

3．2　写真および浮世絵による幕末～明治期の形態

表 3 に 15 項目の計測項目の平均値，標準偏差を，図 3 にはこれらの平均値を図示した．図 4～6 は平均値の推移を示す．比較のため，江戸後期，幕末～明治期の浮世絵から算出した人体形態データと，1951 年と 1994 年の人体計測のデータを合わせて示した．なお 1951 年の計測では股下高は計測されていないため，データには股下高が欠損している．また 1994 年の資料ではヒップ高は計測されていない為，ここには 1995 年のワコール人間科学研究所による人体計測データ [10] を使用した．

全頭高は，明治期写真の値が最大で，次いで後期浮世絵，1951 年の人体計測値が大きい．明治期写真と幕末～明治期の浮世絵の値には大きく差があるが，明治期写真以降，1994 年人体計測値へは小さくなる傾向にある．　高径は，1994 年の人体計測値と明治期写真の計測値では，肩峰高，バスト高の値にほとんど差がないが，ウエスト高，股下高で明治期写真の値が小さい．明治期写真と幕末～明治期浮世絵では，肩峰高，バスト高にほとんど差はなく，ウエスト高，股下高，膝関節高での差も小さい．ウエスト高と膝関節高は，時代とともに大きくなる傾向である．　幅径は，1994 年人体計測値と比較して明治期写真は肩峰幅は小さいが，その他の幅径はすべて大きい．幕末～明治期浮世絵の計測値も，1994 年人体計測値と比較して肩峰幅が小さい．肩峰幅，胸部横径，ヒップ幅は時代とともに大きく推移している．四肢長では，明治期写真がどの項目も最も値が小さく，幕末～明治の浮世絵は最も長い．

表 3　各計測値の平均値と標準偏差

項目		種別	後期浮世絵 (1800～1850)	幕末～明治期浮世絵 (1850～1900)	明治期写真 (1880 ～1900頃)	人体計測 (1951)	人体計測 (1994)
高径	全頭高	平均値	14.81	13.06	15.17	14.75	14.01
		SD	1.21	0.41	0.66	0.53	11.60
	肩峰高	平均値	81.21	81.47	80.06	80.00	80.97
		SD	1.63	1.29	1.75		50.00
	乳頭高	平均値	70.93	71.63	71.06	70.84	71.18
		SD	1.85	1.66	2.07	1.17	48.00
	ウエスト高	平均値	57.78	59.27	59.85	61.20	61.25
		SD	1.26	0.81	1.49	0.86	41.00
	ヒップ高	平均値	48.28	50.37	48.35	49.01	48.10
		SD	2.24	0.75	1.78	1.04	
	股下高	平均値	42.16	44.32	42.26		45.15
		SD	1.48	0.86	1.27		36.80
	膝関節高	平均値	22.70	24.32	25.97	24.87	26.31
		SD	1.53	0.25	1.81	0.69	21.80
幅径	肩峰幅	平均値	22.45	19.49	20.63	22.34	22.00
		SD	2.81	1.63	1.30	0.80	21.80
	胸部横径	平均値	16.19	16.92	18.29	17.17	17.16
		SD	1.34	1.22	1.29	0.96	18.10
	ウエスト幅	平均値	15.94	15.46	15.77	15.05	14.37
		SD	1.04	1.03	1.02	0.92	19.20
	ヒップ幅	平均値	20.67	17.48	21.13	20.14	20.97
		SD	1.19	1.22	1.67	0.76	19.60
四肢長	上腕長	平均値	18.94	18.95	17.07	18.66	18.53
		SD	1.84	1.78	1.03	0.58	17.00
	前腕長	平均値	13.81	15.17	12.95	13.37	14.70
		SD	1.56	1.32	1.82	0.53	17.30
	大腿長	平均値	25.61	25.77	22.50	25.98	26.45
		SD	0.79	0.80	1.36	0.90	28.60
	下腿長	平均値	18.78	19.85	18.87	20.53	21.02
		SD	0.83	0.71	1.41	0.79	21.30

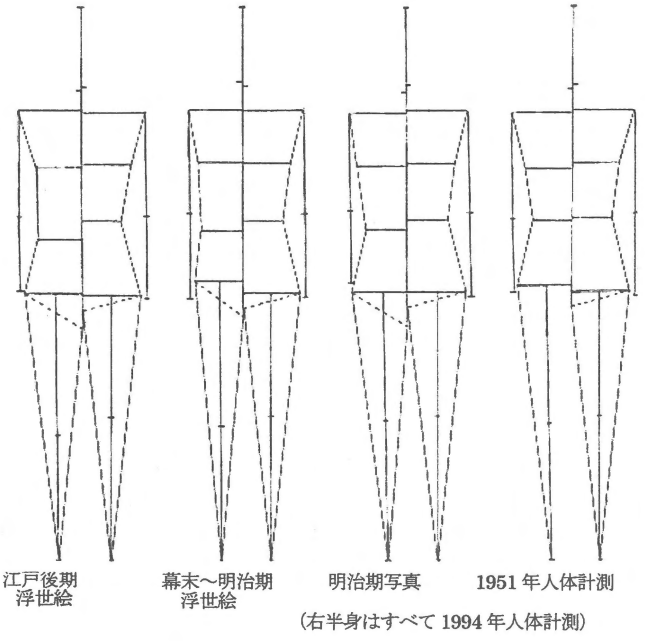


図3　計測平均値に基づく人体図

また1994年人体計測値と比較すると，明治期写真は前腕長が極めて短く，幕末～明治期の浮世絵は逆に前腕長が長い．

シルエットの特徴は，明治期写真の胸部から胴部の脇のラインの形状は1951年，1994年の体形に比較的近い．幕末～明治期浮世絵は脚部の長さと細さが特徴的で抑揚が少ないシルエットである．後期浮世絵は，肩峰幅の大きさとウエスト高が低さなど，抑揚の大きいシルエットである．またすべての人体形態で，ウエスト高の差が顕著にシルエットになって表れている．

4. 考察

明治期写真による主成分分析では，今回資料とした明治期写真の女性の形態は，第1主成分では人体下部の高さ，および四肢末端部の長さ，第2主成分では人体上部の高さ，第3主成分では体幹上部幅の広さや上肢の長さが示されたことから，高さに関する項目が多く説明された．また幅径で上部の幅径を表す因子である．

明治写真の女性の形態を江戸後期，幕末～明治期の浮世絵人体計測のデータ，1951年と1994年の人体計測のデータと比較すると，写真の女性の形態は，肩幅は狭く，胸部ヒップ幅が大きく上肢の短い体形であり，体幹部全体のシルエットに比較的抑揚が表れている。

写真を観察すると，着物の胸部からウエストにかけての着付け方は，胸元はゆったりとしている。しかしウエストでは帯の部分でやや絞った着付けになっている．このことは明治期写真の形態が胸部，

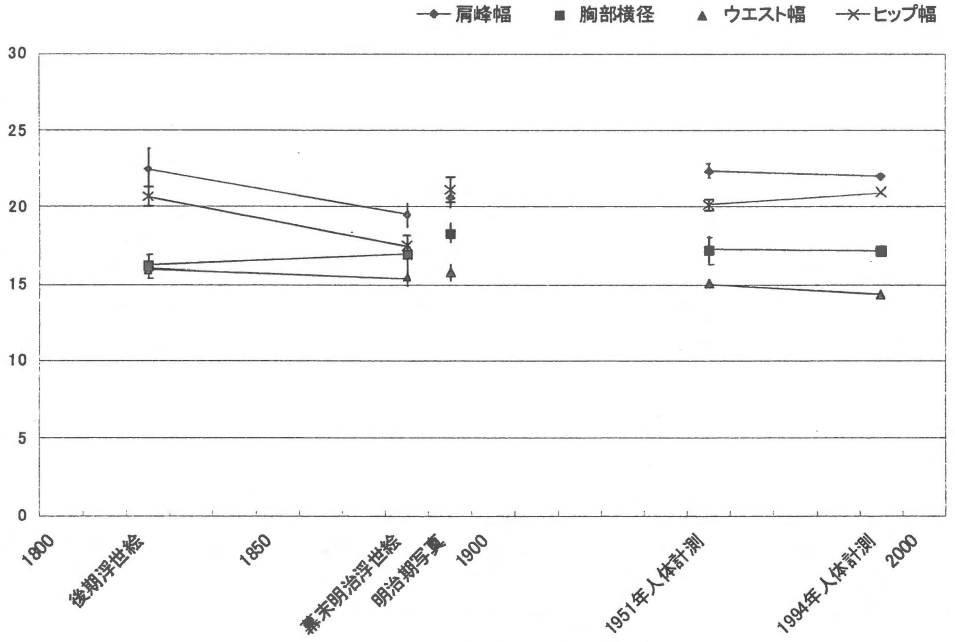


図4　幅径の推移

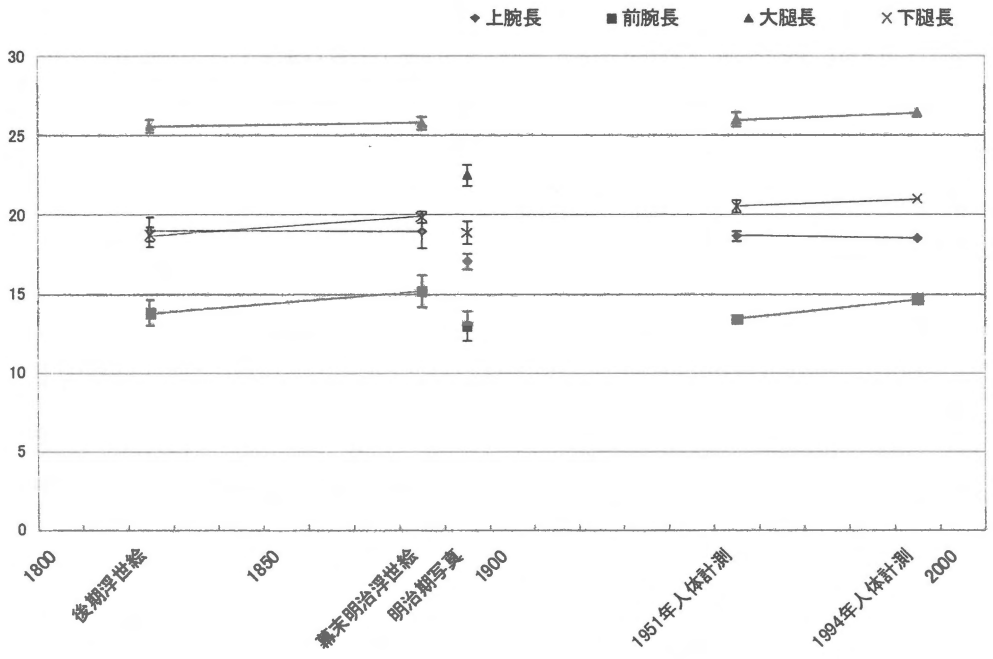


図5　四肢長の推移

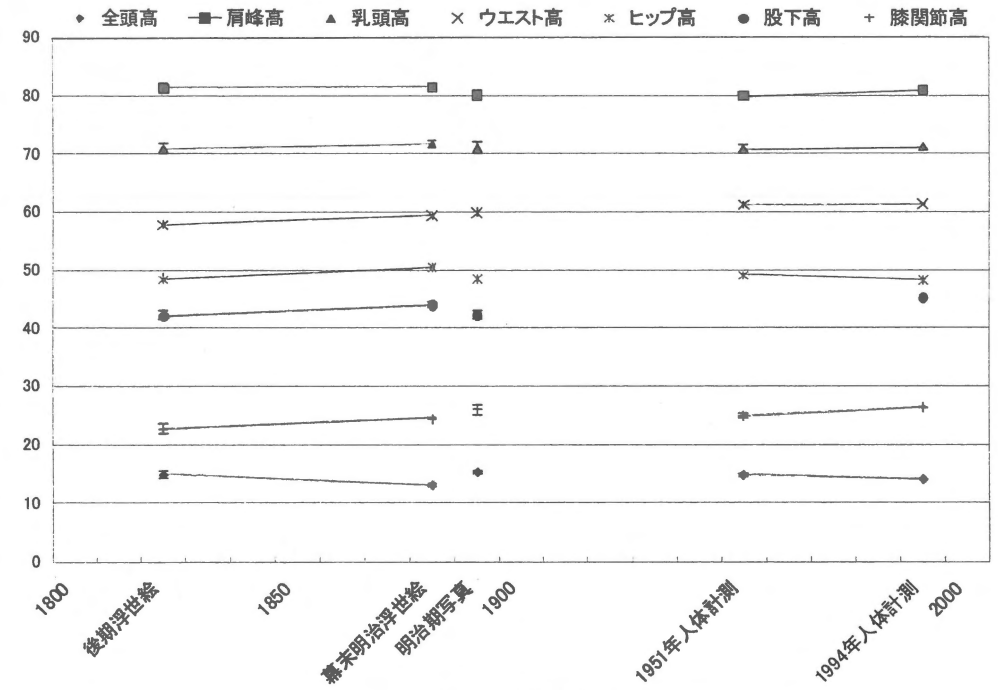


図6　高径の推移

ヒップの幅が強調された要因と考える.

全頭高の明治期の資料としてベルツによる1880年の人体計測データがある[11]. これは17歳～27歳の日本人女性10名について82項目にわたりを計測したデ

ータから平均値を算出すると，その全頭高は 15.51 であった．また 1905 年に発行された「人体画法」に記述されている女性モデルの全頭高は 14.92 とされている．このことから今回算出した明治期写真の 15.17 は，同時期の人体計測値をほぼ反映していると考える．

幕末から明治期の浮世絵の計測値は，全頭高が小さく細身で四肢が極めて長い．この体形は明治期写真と幅径や四肢長で大きく特徴が異なるが，体幹部の高径項目の値では幕末から明治期の浮世絵と明治期写真は比較的似ている．一方,幕末から明治期の浮世絵を後期浮世絵とでは，体幹部の幅径のバランスがかなり異なる．これまでの分析で,江戸時代を通して浮世絵の人体形態は，後期浮世絵が示すようなウエストの位置の低さに特徴があることを確認しているが，幕末から明治期の浮世絵ではウエストの位置は比較的高く，明治期写真や 1951 年のウエスト高に近い．このことは，幕末から明治期の浮世絵は江戸期浮世絵とは異なる形態特徴を持つと言える．

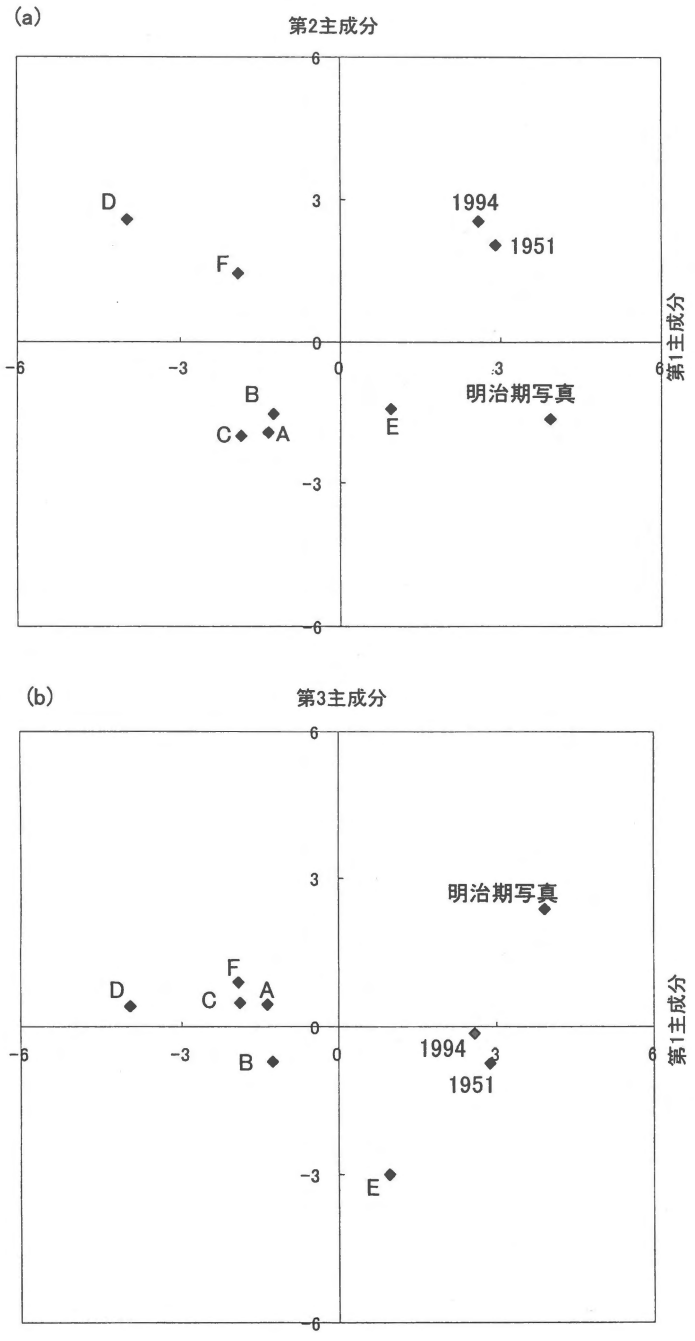


図 7　第 1～3 主成分空間における各年代グループの位置

次に江戸時代～現代までを通した人体の形態の特徴とそれぞれの年代の類似性を見る．

江戸時代の浮世絵のデータは，1600 年から 1900 年の間に描かれた浮世絵 76 枚を筆者らが 50 年ごとに古い時代から A～F に分類し定量化した年代グループである．

まず，A～F の江戸時代の浮世絵，明治期写真，1951 年，1994 年の計測値の主成分分析を行い，累積寄与率が 78．99%の第 3 主成分までの各主成分の解釈を行った．第 1 主成分は肩峰幅，バスト幅，ヒップ幅，全頭高で正に大きく負荷し，高径項目，上肢が負で大きく，第 1 主成分は痩身度を表す因子と考えた．第 2 主成分は，高径項目と下肢長で正に大きく負荷し，体幹下部の幅径が負に大きいことから，下部の高さと幅の因子，第 3 主成分は膝関節高，ウエスト高で正に負荷し，大腿長，上腕長で負に大きいことから四肢の上部の長さと末端を除く胴部高さの因子と考えた．

これらの各年代グループの主成分得点を空間にプロットすると図 7 (a) (b)のようになる．

(a)の第 1 主成分，第 2 主成分から，1994 年と 1951 年人体計測値は，体幹部の幅が広く高径が大きい体形で互いに近い．明治期写真は幅径が大きく高径が小さい為，1951 年や 1994 年の人体計測値との距離はやや離れている．F の幕末～明治期の浮世絵も高径は大きいが，幅径は小さい為，1951 年，1994 年の人体計測値，および明治期写真とはやや離れた距離にある．しかしその他の A～D の江戸期浮世絵の位置

表4　第1～3主成分空間における年代グループ間の距離

	A(1600～1650)	B(1650～1700)	C(1700～1750)	D(1750～1800)	E(1800～1850)	F(1850～1900)	明治期写真	人体計測(1951)	人体計測(1994)
A(1600～1650)	0.0000								
B(1650～1700)	1.6348	0.0000							
C(1700～1750)	0.5816	2.2164	0.0000						
D(1750～1800)	7.1457	7.9181	6.7203	0.0000					
E(1800～1850)	6.2872	4.6524	6.8688	12.3097	0.0000				
F(1850～1900)	4.3739	5.2343	3.9113	3.6768	9.6259	0.0000			
明治期写真	7.4956	8.4106	8.0400	14.0959	8.5904	9.8393	0.0000		
人体計測(1951)	9.4241	7.7893	10.0057	8.5356	7.6447	7.0630	7.8567	0.0000	
人体計測(1994)	9.0322	8.5018	9.6138	7.1747	8.4208	6.6711	8.0496	1.3609	0.0000

と比較すると，明治期写真，Fの幕末～明治期の浮世絵，Eの後期浮世絵は1951年，1994年の人体計測値と近い距離にある．

(b) の第1，第3主成分の空間においても1951年と1994年人体計測値に対して，明治期写真，Fの幕末～明治期の浮世絵，Eの後期浮世絵が比較的近い空間にある．

このように (a)，(b)から江戸時代中期～後期，幕末から明治期の浮世絵と写真は，主成分空間で現代人の形態に比較的近接している．以上のことをより詳細に見るために，3次元主成分空間（第1～3主成分）における各年代グループ間の距離を求め，その結果を表4に示した．この表から，1951年，1994年の人体計測値とFの幕末～明治期の浮世絵のユークリッド距離が極めて近いことが判明した．

5. おわりに

本稿では，明治期写真に写された女性の人体形態を推定して描出し人体各部位の寸法やプロポーションを算出するという手法によって人体寸法のデータベースの構築を試みた．

このように絵画や写真資料の定量的分析から得られる当時の人体寸法や体形に関する情報は，人体計測値のなかった時代の体形を知る手掛かりとなると考えている．

今後は，より広い資料によって過去の時代の体形と着衣形態を客観的な数値として把握し，日本人女性の体形と着衣の変遷過程を明らかにするとともに，衣服と体形との長時間にわたる相互作用の観点から浮かびあがる人間の感性や，着衣に表された文化および歴史的作用の意味を吟味することが課題である．

謝辞

本研究にご協力をいただきました谷山文衛氏，舞踊家の孔ちはる氏に感謝する．本研究の一部は文部科学省科学研究費補助金（No.17652082)の助成を受けた．

引用文献

1) 平本（1972）縄文時代から現代に至る関東地方人身長の時代的変化，人類誌，vol. 80. No3, 221-236.
2) 森下，黒川（2003）浮世絵を通してみた江戸時代女性の人体表現について，人文科学とデータベース論文集，51－69
3) 森下（2004）浮世絵から見た江戸時代女性の着衣形態の定量化，意匠学会大会論文集，10－11
4) 森下，黒川（2005）浮世絵に描かれた江戸期女性の人形状の定量的考察，ヒューマンインターフェイス学会第9回ノンバーバルインターフェイス研究会講演予稿集，1－8
5) 小沢（2001）幕末・明治の美人図鑑，世界文化社
6) 幕末・明治美人帖愛蔵版（2002），新人物往来社
7) 小沢（1985）日本写真全集1，小学館
8) 柳沢(1961) 日本人の成人女子の身体比例について，人類誌，vol. 69. No2, 55-65.
9) 日本人の人体計測データ集（1996）人間工学生活センター．
10) 鈴木（2000）日本人のからだ健康・身体データ集，朝倉書店．
11) E. BAERZ (1883) DIE KORPERLICHEN EIGENSCHAFTEN DER JAPANER，ドイツ東亜自然民族協会誌，85，35－103.
12) 川崎（1905）人体画法，隆文館．

嵯峨本『伊勢物語』の木活字及び組版分析モデルに関する報告

An Analysis Model of Japanese Wooden Type and Composition; "Ise-Monogatari, Saga-Press"

津田 光弘
Mitsuhiro TSUDA
イパレット／(奈良女子大学大学院 非常勤), 〒560-0041 大阪府豊中市清風荘 2-8-30
Ipallet, 2-8-30 Seifu-sou, Toyonaka, Osaka 560-0041, JAPAN

鈴木 広光
Hiromitsu SUZUKI
奈良女子大学 文学部, 〒630-8506 奈良市北魚屋西町
Nara Women's University, Kitauoya-nishi-machi, Nara 630-8506, JAPAN

あらまし:
嵯峨本「伊勢物語」の組版推定を目的とし、デジタル化した印字面から全ての活字の数値化と同異識別を行った。分析モデルには、冊子から印字面、行、各印字に至るまでの階層的なアノテーションを用いた。各解析の段階で専用のツールを作成して効率よく分析が行えるようにした。「アノテーション＝印字」という単位で、連綿体を含むすべての印字ボディのサイズを測定し、得られた幅、長さ、面積、出現順の数値と、画像そのものの目視比較を併用して予想される活字への分類を行った。以上によって、活字・印字の関係データベースを作成し、活字の大きさ、組版の構造を推定し、江戸初期の木活字による印刷工程の一端をうかがい知ることができた。

Summary:
In order to estimate composition of "Ise-Monogartari, Saga-Press", we digitized the print and measured size of all printing bodies from these images, and identified them. For this analysis model, it used structured annotations that have consisted of volumes, pages, lines and each printing. At each stage of the analysis process, it applied an original application in order to raise efficiency. By defining a unit of "annotation = printing type", we measured width and height of all printing body including continuous type, and classified them used values such as width, length, area and an appearance order together with the pattern in their images. By the above, we made a type-printing database, and estimated a standard size of a printing type and a structure of composition, and were able to know a printing technique by wooden types in the Edo era early days.

キーワード: 画像, 木活字, アノテーション, 測定, 分類, データベース
Keywords: image, wooden type, annotation, measurement, grouping, database

1. はじめに

嵯峨本は角倉素庵の開版による一連の印刷本をいうが、その中でも、慶長十三年から十五年にかけて何度も繰り返し制作された「伊勢物語」はその代表的なものといえる。この本は、一字および二字、三字等の連綿体の木活字を組み合わせて刷られているが、異版が数多く存在することが、これまで多くの研究によって知られている。ところで、印刷の側からみると、木活字や組版用器具が現存していないなどの諸事情から、活字制作から組版、印刷にいたる一連の工程についてはいまだ多くの事柄が解明されていない。

そこで報告者らは、近畿大学のご厚意により、同大学中央図書館所蔵の『伊勢物語』(小汀文庫旧蔵)を高精度のデジタル画像におさめ、コンピュータ上で同書における版面の精査および印出字の調査を行い、活字コマの総数および活字の使用状況、活字コマの寸法規格などの算出を行った。[1]

本発表では、上記のデジタル画像を用いて印字面から活字の数値化と同異識別を行うデータ分析モデルを報告すると共に、分析プロセスに沿って用いたツールを紹介する。

2. 調査資料について

近畿大学中央図書館所蔵『伊勢物語』は川瀬一馬氏の分類にいう第一種(慶長十三年初刊)の(イ)版に属するものである。上下2冊本であり、本文は上巻51丁、下巻63丁から成る。鶯色、薄紅色などの色変わり料紙も用い、紙質は丁によって僅かではあるが異なる。半丁は9行で構成される。各行は単字活字に換算して18字詰になる。また、和歌は2字下げ(1字下げ箇所あり)で組まれている。

3. 分析プロセス

分析は次のようなプロセスに拠った。

(1) 事前調査
(2) 撮影
(3) デジタル画像の精度調査
　印字面の精度、(撮影の繰り返し精度)、
　解像度と誤差の推定
(4) 前処理
　印字面グリッドの作成と印字面の回転補正
　印字測定用アノテーションの構成
　繰り返し精度の確認
(5) 印字測長
　アノテーションの編集による印字の測長(縦・横)、
　付帯情報のテキスト入力
(6) 活字単位のグルーピング
　長さ、幅、面積等を参考とした目視による分類
(7) データの解析
　活字数算出、活字辞書(データベース)の作成、
　活字組み版モデルの推定

以下で、撮影以降のプロセスを詳細に説明してゆく。

4. 撮影

貴重書の撮影であると共に、測定のための精度を意識した撮影を行う必要がある。資料の保全はもとより、資料とカメラとの正確な正対、文字の判読性、撮影ひずみ、画像の階調、繰り返し精度の確保などを配慮して撮影を行った。撮影は、堀内カラー・アーカイブサポートセンターが行った。

撮影には繰り返し精度を確保できるデジタルカメラを使用した。デジタルカメラは2004年時点で最適な選定を行った。印字計測のために資料を可能な限り精細に撮影することを考慮し、1600万画素のデジタルカメラバックを選定した。

資料に対しカメラが正確に正対するようセッティングを行い、俯瞰撮影を行った。資料には透け防止のため間紙を挿入し、無反射ガラスが装着されたブックホルダーに密着するように資料をセットし撮影した。表1に主な撮影機材と撮影条件を記録する。

撮影によって得られた画像データは、後の画像処理での損失を考慮してRAWデータおよび、階調16bit、色空間はAdobe RGBのTIFFフォーマットで保存した。

以上によって、全階調情報を保つ計測用画像データであると同時に、デジタルアーカイブの目的に沿う画像データを得た。

表1. 主な撮影機材と撮影条件

カメラ	マミヤ 645AF
デジタルカメラバック	コダックDCSプロバック645
レンズ	マミヤ AF150mm f3.5
照明	コメットストロボ CB2400IV
その他の機材	ｽﾀｼﾞｵｽﾀﾝﾄﾞ, ﾌﾞｯｸﾎﾙﾀﾞｰ, 精密方眼, 他
撮影方法	俯瞰正対撮影
撮影距離	約 160cm
絞り	f16

5. デジタル画像の精度

撮影したデジタル画像は縦横 4072 画素であり、その中央部、面積比 30%の範囲に充分なマージンを伴って、印字面が半丁ごとに写っている。各画像の精度を確認するため、被写体の周囲に精密方眼のメッシュを配置している。この精密方眼から撮影の繰り返し精度も確認することができる。印字測定に先立ち、得られたデジタル画像の今回の目的への適合性を検証する。

(1) 印字面の面内精度(位置依存性)
(2) 繰り返し精度

の2つを行い、誤差の範囲を確認した。

このうち、(1)については市販の画像アプリケーション Adobe Photoshop 7.0 を使い計測した。(2)については、次プロセスで回転をかけた後に、iPalletnexus(イパレットネクサス)を用い、ベクトル・アノテーションを作成して行った。(1)と(2)の間で解像度の変更はない。

先ず、精密方眼を印字面に配置した確認用の撮影画像を用いて印字面内の位置精度を求めた。この精密方眼は、あらかじめ JIS 1 級の定規で測定を行い、長さ精度が JIS 1級と同程度であることを確認している。個々の印字の測定を行うため、面に依存する歪がなく均一であることが必要である。印字面の9箇所で精密方眼の単位長さあたりの画素数を求めた。その結果、単位長さ 127.0mm に対して 1552±3pixel であった。この数値の範囲で面の傾斜を認めたが、対象とする印字の大きさは2倍格で約 25mm であり、面による影響はこの場合約 1.2pixel となる。1倍格を合わせ全印字の約 93%がこの範囲に収まる。一方、ポインティング・デバイスによる位置の繰り返しの誤差は片側最大 1〜2pixel であったため、面による影響はこれに比べて大きいとは言えず、正対俯瞰による撮影は測定要求を満たすと判断した。

また、繰り返し精度については、被写体である印字面の外側に写してある精密方眼の単位長さあたりの画素数から間接的に求めた。測定用の画像を作成した段階で、全撮影画像で測定、集計を行い、単位長さ 40mm で 4pixel 以内であることを確認した。これはポインティング・デバイスの偶然誤差を含んだ値であり、繰り返し精度も測定要求を満たしていると言える。

以上から、撮影した画像の平均解像度を 310.3dpi(有効桁数は 4 桁で±0.6dpi の誤差を含む)とした。これは 100% 表示時に約 12 画素で 1mm という精細さであり、資料現物を肉眼で見ながら細心の注意を以って測定する場合と同等の品質が、コンピュータの画面上で得られる。最終的に、面内精度と繰り返し精度を考慮し、最大の印字4倍格に対して、誤差は最大で±4pixel であると見積もった。印字ボディはその内側にあり、ほとんどの印字は測定時の系統および偶然誤差の合計が±0.3mm よりも小さな値に収まると考えた。

ただし、以上のように数値を追ったが、実際には印刷時の刷りによる影響や印字のつぶれ、かすれなどが見積もることができない相違として存在するために、仮に同じ活字から印刷されたとしても、ここで算出した値を超える場合もある。この精度は、分析の最後のプロセスでの活字サイズ算出時の誤差の見積り、また、測定手順を検証する参考値として記録する。

また、(2)と同時に各画像の階調精度も確認した。これは、測長時に印字の周辺濃度に違いがあると微々とはいえポインティング操作の誤差要因になる可能性があると考えたためである。結果は測定誤差を含み約1%の変動に留まっていることを確認し、階調の繰り返し精度に関しても問題がないと判断した。

6. 前処理

解像度と精度が得られたので、次に印字計測のための準備を行う。測定方法としては、各印字の外周に長方形を用いてその縦横の画素数をコンピュータ画面上で得る。印字数は上下冊で当初の見積りでは1万数千字存在し、一般の画像アプリケーションでは作業上効率が良くない。そのため、測定ツールには画像上でアノテーションを編集できる iPalletnexus を用いるが、オリジナルでは測定表示機能が無いため、アノテーションの位置と大きさを表示する「数値表示ダイアログ」や「グリッド編集」、「等倍サムネイル」等の新機能を追加した。iPalletnexus では画像上のアノテーション・データは個々に XML ファイルに保存するが、主要な項目は同時に検索用の RDB ファイルに格納する。アノテーションを編集操作しながら自動的に測定データが蓄積されてゆく。[2]

先ず、印字面全体の画像をスムーズに原寸大で閲覧できるようにした。大きなディスプレイに拠らずとも一般的なコンピュータ環境で閲覧できるように、分割画像形式に変換し、1画像すなわち半丁単位でXMLファイルを作成しそれらの構造化を行って、基礎となる閲覧環境を作った。前処理の段階自体は、次のプロセスに分かれる。

(1) 印字面グリッドの作成
(2) 印字面の回転補正
(3) 測定用画像の作成
(4) 印字測定用アノテーション群の構成
(5) 繰り返し制度の調査(5 で説明)

前処理工程での最も大きな課題に、回転補正がある。今回の測定方法を設計するに際し、画像に対する印字面の平行をどのようにして得るかという問題が顕在化した。印字面の各行が本の小口に対し正確に平行となっている保証はない。実際に画像を見ても、印字面が紙に対して明らかに角度を持っている丁もある。角度のずれは、本自体に起因するもの、撮影の際に生じたセッティングに起因するものの2種類があり、2重の回転がかかっていることになる。逆に、この問題を予測していたので、回転補正を撮影後の画像処理で行うことが必然であったため、本の外形を撮影時のセッティングの厳密な基準にする必要がなく、撮影時間の短縮につながった。

しかし幸いなことに、嵯峨本『伊勢物語』の半丁は9行、各行が単字活字換算で18字という性質がある。この性質を利用して、回転角度とサイズを共に編集可能なグリッド・アノテーションを使って半丁での印字面角度を測定した。図1に表示するように、印字面の上端と下端、左右および各行がいずれもグリッド内に収まるようにグリッドを配置した。行内で左右に印字がぶれている場合もあるため、推測をしやすくするために補助線を用いたが、この補助線は後のインテルの伏線ともなる。このグリッド・アノテーションの四隅の座標を記録し、その数値を利用して各画像の補正用回転角度を算出した。このプロセスには、精度の高い位置決めを行う必要性から1600×1200の高解像度モニタを用いている。回転とトリミングの画像処理には Photoshop 7.0 を用いた。回転に伴う誤差は、無視できる範囲である。

補正後の画像を再度、分割画像形式に変換して最終的な測定用の画像を作成した。この段階で後の作業での操作性向上を意図して、色空間を sRGB、8bit の極高品質な JPEG 画像としている。

次に、各画像に改めてグリッド・アノテーションを配置して各印字がそのうちに収まるように位置とサイズを調整した。回転はかけない。このグリッド・アノテーションと印字の区切りを基にして、印字アノテーションを作成した。図2は前工程で準備した最終的なアノテーションの配置を示す。

iPalletnexus では通常、自動的に重複のないアノテーションIDが割り振られるが、各印字のアドレスを理解しやすくするため、丁、表・裏、行、印字の位置を基にして IDをつけ、文字枠数や文字、前後関係などの付帯的な情報もメタデータとして盛り込んでデータを構成した。また、2倍格に3字連彫、3倍格に2字連彫などの区別もつくようにした。とくに「に」「し」については1倍格から3倍格までである。和歌については、1字あるいは2字下げとなっている。以上のような特徴を全て記載できるように、データ構造を設計した。半丁のデータ構造は図3のように階層構造になっている。

図1. グリッドの回転による補正

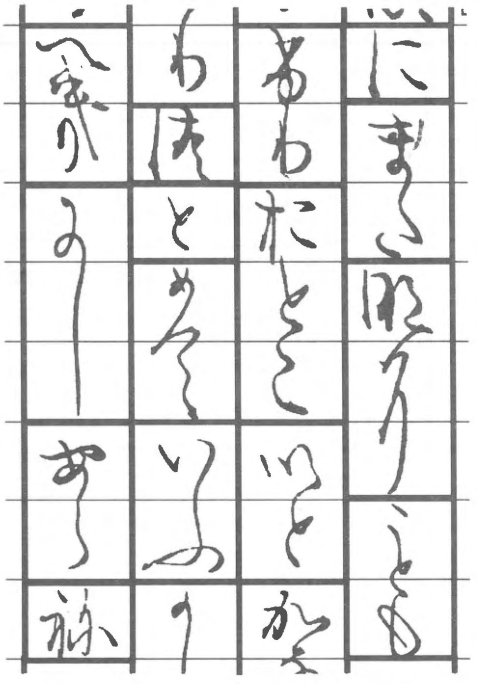


図2. 前処理プロセスでの最終印字面
（印字アノテーションは強調している）

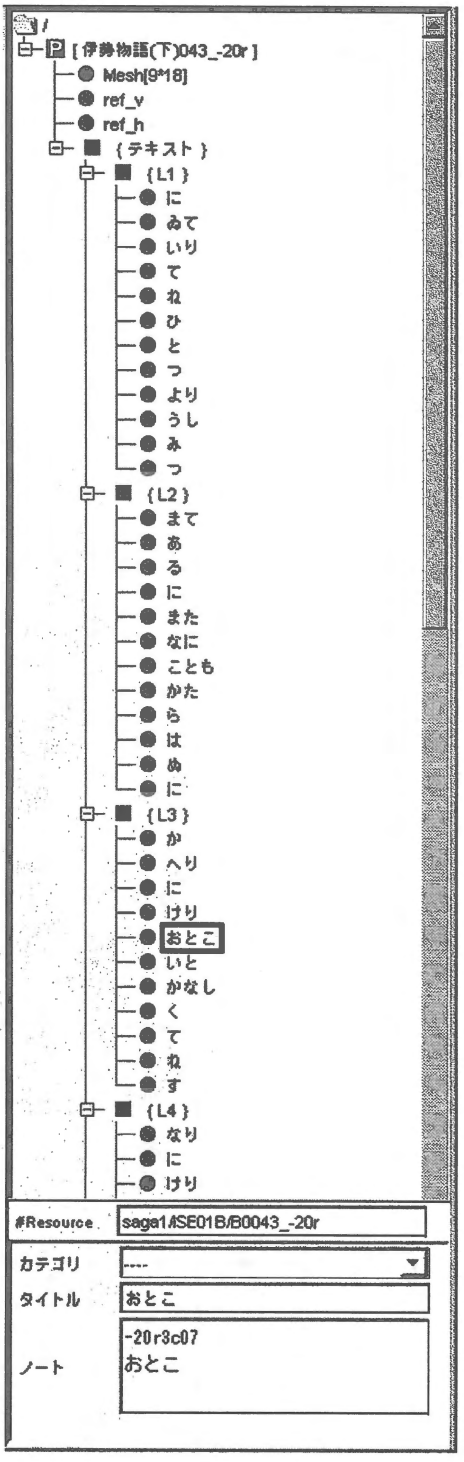


図 3. データの構造
(iPalletnexus のアノテーション・ツリー)

[説明]「下 20 丁オ」の構造化例
テキスト階層下に行(L1～L9)を配置し、その下に各印字のアノテーションを配置する。アノテーションのタイトルに読みを設定している。

7. 印字測長

この分析プロセスでは、全1万数千印字について仮配置した長方形のアノテーションを手作業によって編集し、その横幅と縦の長さを記録した。この作業を繰り返して、図4に示すようなデータを各印字面で得た。アノテーションを再編集すると同時に元画像と同一の解像度の印字画像(以降、切り出し画像と呼ぶ)が作成される。境界判断のため、切り出し画像の周囲には 10 画素のマージンを設けて作成してある。この切り出し画像は、後プロセスの分類で使用する。測長の際に、ポインティング・デバイスやアノテーションの座標を確認できるように数値表示ダイアログを追加した。図 5 は、カスタマイズを行った iPalletnexus のユーザ・インタフェースである。

精度検証の際に記述したように、測定は注意深く行っても、ポインティングの始点、終点で片方 1～2pixel 程度の偶然誤差を生じる。判断の系統誤差や精度を合わせて考慮すれば、合計数ピクセル程度の誤差が生じる。しかし、それ以上に印刷の際に生じた差異や、印字のかすれ、つぶれという影響は大きい。これらの影響を含んでいる測定値のみを用いては、同じ活字による印字であるとは必ずしも判断できない。従って、ここで行った測長は、必ずしも絶対的な長さを求めることが目的ではなく、長さという数量を用いて次の分析プロセスである分類を効率化することを目的としている。

図 4. アノテーションの編集結果

8. 印字単位のグルーピング

この分析プロセスの目的は、印字アノテーションを同異識別して活字のグループを得ることである。測長したデータを基に活字のグループ化を行う。グループは、活字の倍格数、読み、活字の順に構造設計し、この構造を反映する活字 ID を定義した。活字の特徴をあらわす数値として、印字の横幅と縦の長さを先の測長プロセスで求めた。既に記したように、これらの数値は刷りの性質上、必ずしも絶対な判断基準ではない。また、活字がそれぞれ全く固有の形をしているとも限らない。しかし、これらの数値は数多くの印字を識別する際にその有効な支援情報となるはずである。従って、これらの数値を利用すると共に、画像の目視確認による判断を併用する分類ツールを Macromedia Flash で制作した。(図 6)

このツールには、長さ、幅、それらの面積、出現順序などで切り出し画像をソートする機能を設けた。また、図 6-1 のように、長さと幅で切り出し画像を自動レイアウトする機能なども設けた。印字は必ずどれかひとつのグループに属するという分類原則に基づいてデータベースのテーブルを設計しプログラムを作成した。印字に対応するアノテーションと活字のグループをそれぞれリレーショナル・データベースのテーブルとして設け、このテーブルからグループごとのデータを得て切り出し画像を表示する。分類の際には、切り出し画像の透過比較や、周辺の印字との関係を見るために切り出し元の画像を呼び出しての表示など、複数の判断を用いて総合的に分類できるようにした。それでも実際には、同一活字であることを推測することが難しい場合もあり、「か」「こ」「し」のように印字数が多くかつ特徴点が少ない場合には、判断に迷うこともあった。

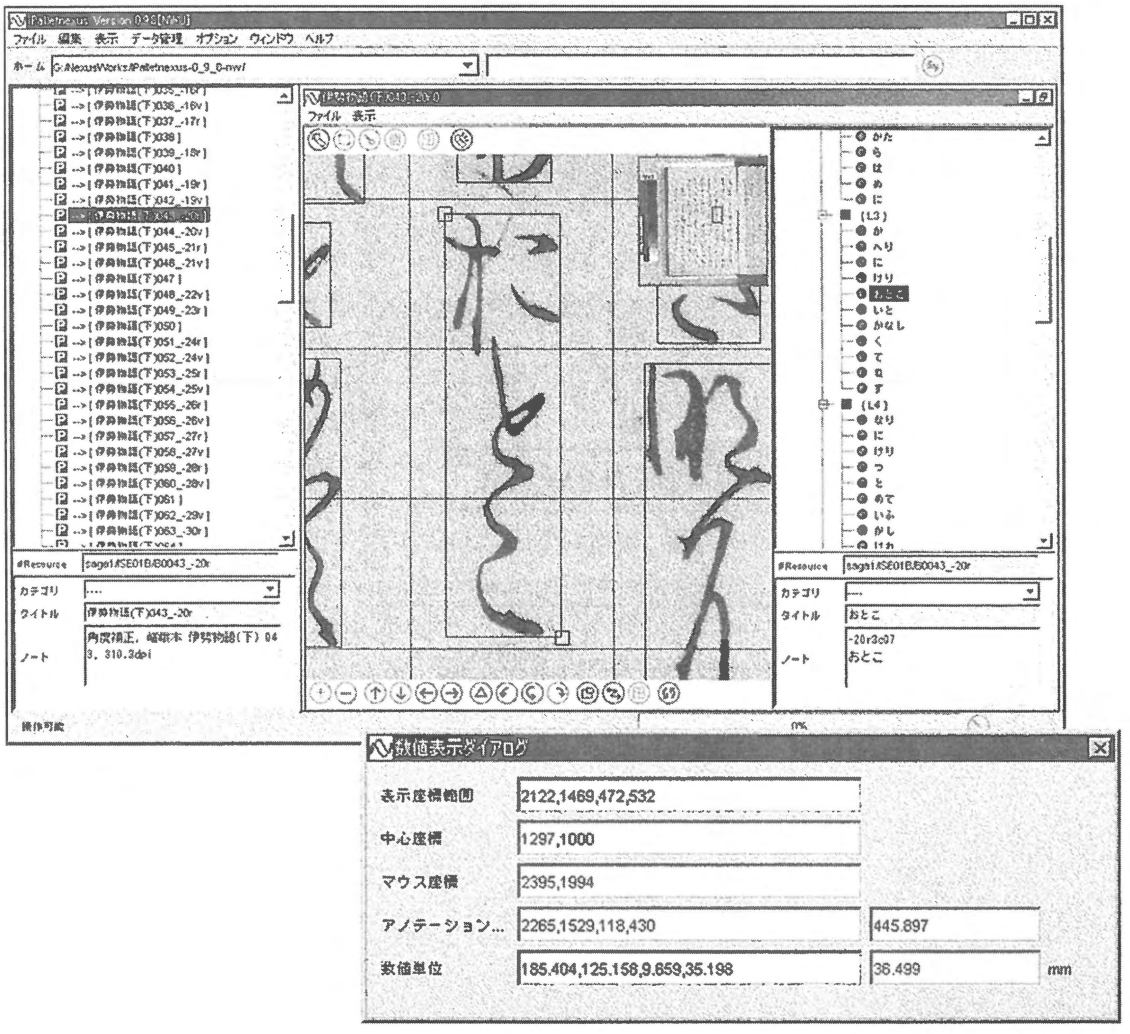


図 5. 測長ツール(カスタマイズした iPalletnexus の全体(上)と数値表示ダイアログ(下))

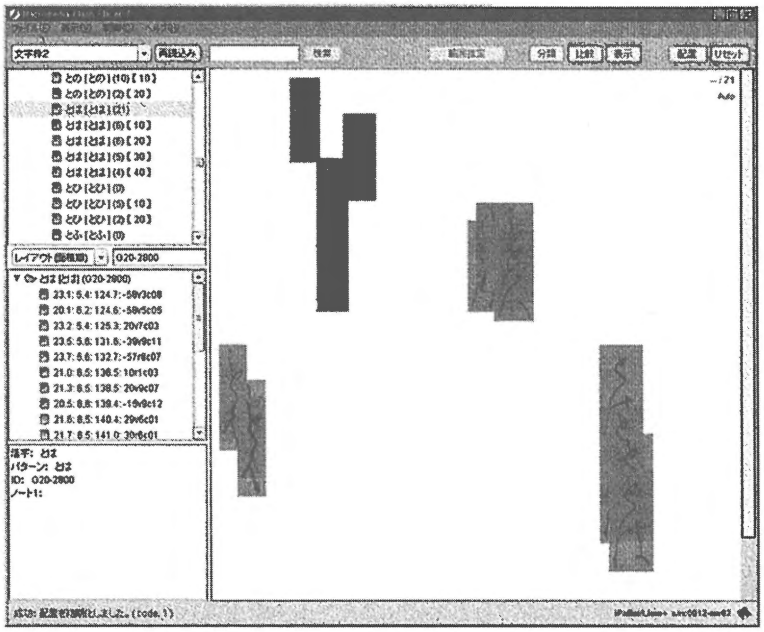


図 6-1. 分類ツール(レイアウト)

[説明]

2倍格「とは」をレイアウト・モードで自動分類した例。分類の軸を、印字の長さと幅にしている。

この事例では4つの印字に明確に分かれている。幅、長さ共に 1mm 以上の差異がある。実際は、このような事例ばかりではない。

レイアウトはポインティングデバイスで変更可能であり、必要に応じて保存できるようにした。

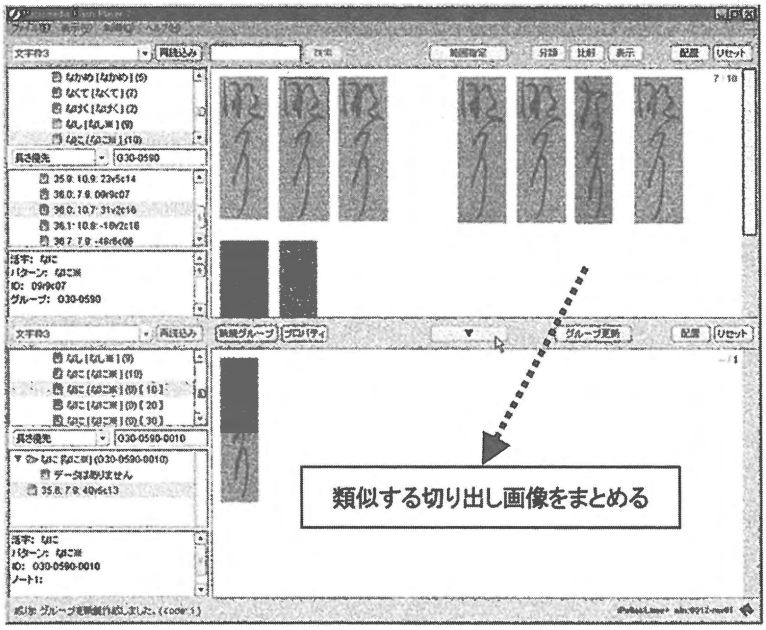


図 6-2. 分類ツール(グルーピング)

[説明]

3倍格「なに」を分類している例。

印字が少ない場合には、1字ずつ判断して分類するほうが確実である場合も多い。新しいグループを作成し、仮の印字グループから活字グループへ移動する。切り出し画像を選択後に、移動ボタンで決定するとデータベースのグループ ID が書き換わる。

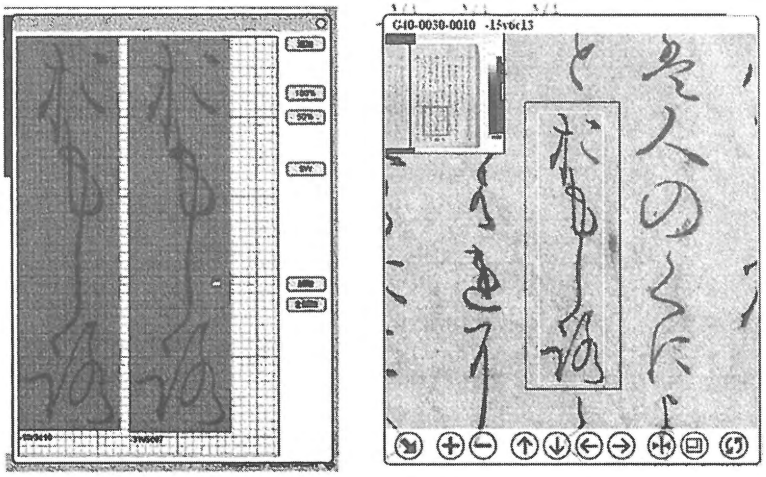


図 6-3. 分類ツール
(比較表示(左)、画像表示ツール(右))

[説明]

4倍格「おもしろ」の例。類似性の判断や、印字の前後関係から活字の分類を判断するためのツール類。切り出し画像から呼び出す。

画像表示ツールは元になる画像をアノテーションと共に表示する。iPallet/Lime(ライム)の応用。

9. 分析結果と組版の推測

以上のプロセスによって、1倍格から4倍格(規格外を1活字含む)の印字総数 15726、活字数 2123 という結果を得た。ただし、極めて分類困難な「し」などがあり、今後の検討次第でまだ若干数の変更もあり得ることをおことわりしておく。

また、測長結果から、活字ボディのサイズは縦12.6mm、横 14.2mm の単字活字を基本(全格)とし、縦寸法の2倍、3倍、4倍の整数倍格で規格化されていると推測した。この活字の幅の推測には、最大の印字幅とインテルの写り込みが役に立った。インテルは幅 2.6mm と推定した。これらのデータをもとに、インテルと活字、クワタで構成される組版を推測した。図7にビジュアル化した仮想組版を表示する。ただし、推定した活字ボディのサイズはあくまで計算上の推定値であり誤差も含む。実物のサイズにも若干の大小はあったはずである。したがって、組まれた活字がぐらつかないように字間にごく薄い込め物が使用されたと考えられるが、その形状を示す証拠が版面上に存在せず、積極的に図示したような形状であったと主張するわけではない。

なお、印字を測定し分類する過程で興味深い結果も得たが、それらについては、本報告の主旨ではないため、省略させていただく。

10. おわりに

現存する貴重な印刷本の画像から遡って、江戸時代の初期における木活字の組み版を得る、という一種の「逆問題」を解く分析モデルを構築し結果を得た。

分析上、アノテーションや情報技術を積極的に利用したが、その利用は主として大量のデータの記録や管理、理解しやすい表示などの支援のために用いた。一方で経験や直感、微妙な判断を伴う部分については従来からの手作業や目視に拠って行った。このような役割分けもまた、人文科学における情報技術活用のひとつの方法であると考える。

最後に、複数ツール間のデータ連携や操作インタフェースがまだ充分にこなれていないため予想外な労力も要した点は反省を込めて記録し今後の課題としたい。

謝辞

貴重な文献の撮影、利用に際してご高配を賜りました近畿大学と同大学中央図書館、研究にあたりさまざまな面でご指導をいただきました森上修氏、そして、嵯峨本『伊勢物語』の書誌情報をご教示くださいました高木浩明氏に感謝申し上げます。また、精度の高いデジタル撮影を行っていただいた堀内カラー・アーカイブサポートセンターの川瀬敏雄氏をはじめとするセンター員の皆様に感謝をします。なお、本発表は平成 16～17 年度文部科学省科学研究費(特定領域研究「江戸のモノづくり」)研究課題「嵯峨本の印刷技法の解明とビジュアル的復元による仮想組版の試み」(研究代表者 鈴木)による研究成果の一部である。

参考文献

[1] 鈴木 広光:「嵯峨本『伊勢物語』の活字と組版」, 平成17年度 日本近世文学会 秋季大会, 2005.

[2] 津田 光弘, 馬場 章, 川瀬 敏雄, 肥田 康:「絵地図画像の研究支援ツールの開発」, 人文科学とコンピュータシンポジウム, pp.241-244, 2002.

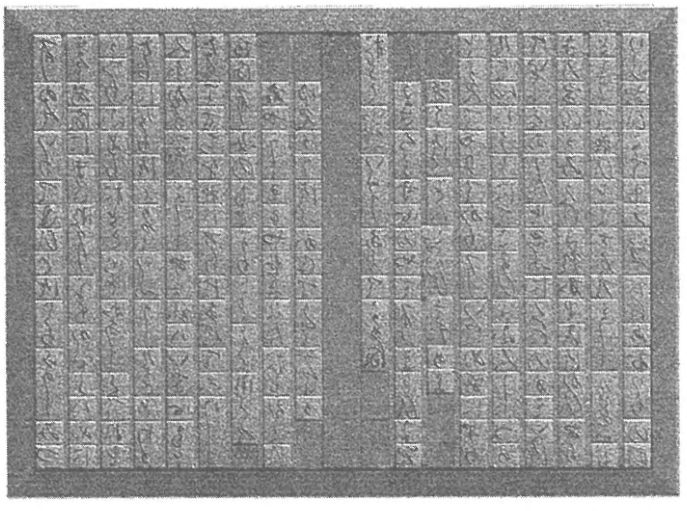


図 7. 推定した仮想組版(2D ビジュアル)

色彩のデジタル化の諸問題

－日本文学作品画像データベースと色彩－

The problem of the digitalization of the color

— The Image-database and the Color of the Japanese literature work —

當山　日出夫

TOUYAMA Hideo

htoym@kcn.ne.jp

花園大学文学部（非常勤講師）

HANAZONO UNIVERSITY (DOCENT)

〒606-8456　京都市中京区西ノ京壺ノ内町　8-1

8-1,Nishinokyou-Tsubonouchi-cho,Nakagyou-ku,Kyouto-shi

あらまし：日本文学研究において、画像データベースが重要である。そして、色彩が重要な学術的意味を持つようになってきている。現在、コンピュータで画像データ処理が、容易に出来るようになった。しかし、画像の色彩の同一性を、保証することは困難である。今後は、人文学研究者にも、色彩学の基礎的知識が、リベラルアーツの一部として、必要となってきている。コンピュータのディスプレイ、スキャナ、デジタルカメラ、画像処理ソフト、などの特性と問題点を知っておかなければならない。

Summary:In the Japanese literature research, the image database is important. Then, it becomes a element with important color. At present, it became possible to do image data processing easily at the computer. However, it is difficult to guarantee the identity of the color of the image. Now, the literally study person, too, needs the basic knowledge of the chromatics. The characteristic of the display, the scanner, the digital camera, the image processing software of the computer and so on must be known beforehand.

キーワード：画像データベース　色彩学　スキャナ　RGB　CMYK

Keywords : Image database　Color science　Scanner　RGB　CMYK

【1】はじめに ― 本稿の目的 ―

人文学研究の分野において、テキスト・音声など、各種資料のデジタル化、データベース作成が進行している。本稿では、画像データに焦点をしぼって、論ずることとする。

今日、画像データの作成・利用は、非常に容易になってきている。コンピュータの性能の向上、処理ソフトの発達、デジタルカメラ(デジカメ)やスキャナの低価格化・高機能化などに起因することは言うまでもない。

ところが、筆者の見るところ、画像データの作成・利用には、ある陥穽があるように思われてならない。例えば、ディスプレイで見ている色彩の同一性は、いったい何によって保証されているのか、共有できる色彩のモノサシとしては何がつかえるのか、そもそもコンピュータのディスプレイで見える色は、人間の眼による色彩の認識とどう違うのか……このような基礎的な問題点について、はたして利用者の多くが考えてみたことがあるだろうか、ということである。

以下、現時点での一般的な利用技術の範囲内において、画像データの色彩をめぐる諸問題について、思うところを述べてみたい。

ただ、これは、現代色彩科学の観点からは、ごく初歩的な問題にすぎない。が、ここで結論を先取りして言えば、この程度の基礎知識をふまえないで、学術的なカラー画像データベースなど論じても無意味なのである。筆者は、リベラルアーツ(人文学的基礎教養)としてのデジタル色彩学が、今後きわめて重要になってくることを指摘したい。これが、本稿の意図である。

【2】学術資料における色彩の重要性

例えば、ある古典籍の古写本があるとする。それを翻刻してテキスト化することは学術的に価値がある。翻刻には、必ず研究者の解釈が介在する。故に、従来、古典籍の翻刻という一見単純な作業も、研究業績として評価の対象とされてきた。では、画像データの場合はどうであろうか。

日本文学関係に限って例をあげれば、中古～中世を中心に数多く残されている絵巻物が貴重である。その他、奈良絵本なども、描かれた絵が重要な意味を持っている。近世になって版本がおこなわれるようになってからでも、丹緑本など彩色した書物の存在がある。また、絵をふくまないとしても、どのような紙(料紙)が用いられたが貴重である本もある。石山切などの古筆あるいは嵯峨本(古活字)などが、あげられよう。

日本語史の観点からは、訓点資料がある。墨筆の漢文本文に加点された朱点・白点・角筆など、色彩を抜きにして研究できない。資料によっては、時期を別にして加点された朱点を、色彩の区別から判別する必要がある。さらには、その漢籍・漢訳仏典等などでは、どのような紙(料紙)に筆写されたのかによって、その由来・来歴を考察することもある。つまり、紙の色や文字の色そのものが重要な学術情報なのである。

現在、文学・歴史研究における画像資料の重要性は、ますます大きくなってきている。従来の代表的な研究としては、『絵巻で読む中世』(五味文彦)(注1)、『絵巻物に見る日本庶民生活誌』(宮本常一)(注2)などがよく知られているところである。

研究目的・分野によって、多種多様な画像データが存在することになるが、色彩の観点からは、次のように整理できる。

(1). とりあえず文字が読めればよい。絵として何が描かれているか判別できればよい。
(2). 色彩そのものに学術的意味がある。

問題は、(2)のレベル利用において、現在の画像データは、その利用に耐えるであろうか、ということである。また、結論的には、古典籍の翻刻と同様、その画像データ化にも、研究者の解釈と責任がともなう、と筆者は考える。

【3】色があればいいのか

現在では、フォトレタッチソフトが普及しており、極言すれば、とにかく色さえついていれば、後は画像処理技術でどうにかなる、とまで言える状況かもしれない。場合によっては、モノクロ画像をカラー化することさえ可能である。だが、ここに重大な問題がある。それは、基準となる色彩が決まっていないことである。あるいは、少なくとも、基準となる色彩のモノサシが、研究者(ここでは色彩科学の専門家ではなく一般の人文学系研究者)同士の間で、共有されていないことである。

(後述するが、一般に使われている『Kodak Color Separation Guide and Gray Scale』は、色彩のモノサシとしては、不適である。)

インターネット上で、ある画像データを見つけたとして、「今、私が、ディスプレイで見ている画像の色は、本当の色なんだろうか？ Photoshopはパソコンに入っているから、画像データとして取りこんでしまえば、色の調整はできるのだけれど、はたしてどの色で見れば、それが本物の色といえるんだろうか？」この疑問に、どうすれば答えることができるのか、である。

ここで、誤解のないように結論を先に確認しておきたい。筆者が言いたいのは、色彩の忠実な再現(確かにこれが理想的な目標ではあるが)ではない。人文学系研究者に求められるのは、オリジナル(原本・実物)の色彩と、ディスプレイ画像との距離感覚である、ということである。あるいは、デジタル画像・写真(銀塩フィルム)・印刷(カラーオフセット)・ディスプレイの表示などにおいて、色彩に、それぞれどれほどのバイアスがかかるかの自覚的知識である、と言ってもよい。

【4】色彩の出力の問題

論点の整理としては、先に、色彩の出力、つまり、ディスプレイでどのように色彩が見えるのか、を考えておく。

(1). sRGB・CMYK・xy色度図・ガモット

現在の色彩科学では、人間の可視光領域の色彩のすべてを、xy色度図上に表現可能である。その成立のプロセスについては、金子隆芳(1988)、池田光男・芦澤昌子(1992)、大山正(1994)、などで、分かりやすく解説されている。

常識的な知識としては、光の三原色は、赤・緑・青、いわゆるRGBであり、色(絵の具)の三原色は、シアン(青)・マゼンタ(赤)・イエロー(黄色)、とされている。が、これはあくまでも通念の範囲であり、厳密には、次のようになる。

[1] RGB

RGBと言っても実は一種類ではない。一般のコンピュータディスプレイで採用されているのは、sRGBカラースペース、という(現実には、これがWindowsパソコンの標準)。この他にも、AdobeRGB、AppleRGBなど、表現できる領域の異なる各種のRGBカラースペースがあり、実際に利用さている。つまり、一般のWindowsパソコンによるRGB値(各0～255)では、絶対的な色の指定は出来ないのである。これは、あくまでもWindowsパソコンに限って有効ということである。

しかも、このsRGBは、xy色度図に比して、表現できるカラースペースが狭い。要するに、人間が見ることの出来る色彩を、Windowsパソコンのディスプレイ上で100%再現することは、原理的に不可能なのである。

[2] CMYK

印刷の三原色、シアン・マゼンタ・イエローについてはどうであろうか。これについても、各種の問題がある。

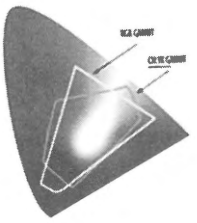
第一に、シアン(C)・マゼンタ(M)・イエロー(Y)だけでは、現実の印刷は不可能であり、さらに黒(K)が、必要とされる。(注3)

第二に、CMYKで表現できるカラースペースと、sRGBとでは異なる。ほとんど重複してはいるが、一方では表現できない色彩が相互に存在する。簡単に言えば、仮に理想的にスキャナやデジカメで画像データを得たとしても、対象が印刷物である場合、ディスプレイ上には、そのすべての色彩を表現することは不可能ということであり、逆に、ディスプレイで見えている画像を、忠実にプリントすることは不可能、ということになる。(注4)

第三に、このCMYKもsRGBと同様、xy色度図に比較すると、表現できる領域は狭い。結果的には、コンピュータのディスプレイの画像データにせよ、書物としてのカラー印刷にせよ、人間が見ている可視光の色彩を完全に表現して見るということは、出来ないのである。この現象をガモットという。

画像データの作成、色彩のデジタル化においは、まずこの点を確実に把握しておかねばならない。〈図表(1)ガモット〉に、xy色度図と、RGBおよびCMYKの表現領域の一般的な違いをしめす(Adobe社のTechnical GidesのHPより)。(注5)

図表(1)ガモット



(2).ディスプレイの調整

デスクトップパソコンで、ディスプレイが独立している場合、ある程度以上の上級機種になれば、ディスプレイ自体に調整機能がある。標準的には、Windowsマシンの場合、標準は、ガンマ値(ディスプレイの明るさの程度)は、2.2(Macでは1.8)。色温度(光の色味、太陽光や蛍光灯などの色の違い、K=ケルビンで表示する)は、6500K、である。(注6)

ところが、ノートパソコンでは、この調整機能の無い機種が多い。しかも、現在では、パソコンは、DVDやiTUNEのために存在すると言っても過言ではない状況にある。つまり、動画像再生用にカスタマイズしてあり、静止画像データをsRGBで忠実に表示することを、目標としてはいない。

(3).表示ソフトでちがう

同じ画像データでも、使用するソフトによって、同じに見えるとは限らない。画像処理(レタッチ)ソフトの代表といえば、Photoshopであるが、これで見える画像は、やや明るくなることが知られている。(注7)これを、逆にいえば、Photoshop(Windows版)で見た画像を、一般のWindows標準の画像表示(当然、インターネットのためのExplorerもふくまれる)で見ると、暗くなってしまうのである。

例として、古代の「くれなゐ(紅)」を擬似的にsRGBで再現すると、「240・50・90」になる(あえて、きりのいい数値にした)。この色彩サンプル画像を、Phtoshopで表示した状態で、『色々の色』(注8)で計測してみると、「217・72・112」になる。しかし、これをPhotoshopカラーピッカーで計測すれば、同じ、「240・50・90」の計測値をしめす。花子フォトレタッチ3(ジャストシステム)などで表示して、『色々の色』で計測すると、最初に意図した「240・50・90」をしめす。

つまり、インターネット上の画像データを、Explorerやで見るときと、それを、自分のパソコンのPhotpshopに取りこんで見るときとでは、画像データが内部的に持つ情報としてのsRGB値は同じであっても、ディスプレイに見える色が違うのである。

(4).その他の不確定要因

同一機種のディスプレイでも、同じに見える保証はない。筆者は、ここ数年、立命館大学で情報処理の授業を非常勤で担当している。パソコン教室には、各自のパソコンの他に、教師用のパソコン画面を見せるためのディスプレイが、2名に1台の割合で中央に設置してある。教師のパソコン画面やPowerPointなど見せるわけであるが、どう見えているか教室内を観察して見て回ると、それぞれに微妙に色彩が違うのが実際である。

第一に、ディスプレイ自体、やはり機械である以上、個体差がある。

第二に、機械である以上、経年劣化は避けられない。

第三に、照明と角度。当然ことだが、部屋の照明と見る角度によって、ディスプレイの見え方も影響をうける。

以上を総合すれば、Windowsに限っても、現在のパソコンのディスプレイで一貫して同一の色彩が表示されることを、保証することは不可能に近い、ということになる。それが、まして、インターネットを通じての画像データの表示となれば、利用者がどのような環境で見ているか……想像を絶した世界とでも言えよう(これは、極言すればのことであって、実際には、ある程度のゆれの範囲内にあることは確かではあるが。)

【5】色彩の入力の問題

画像データの入力について考えてみる。筆者は、ここしばらく、色覚異常について勉強している。特に、最近のカラー印刷された辞書(英和辞書・古語辞書など)が、色覚異常の人には、どのように見えているのか、色覚シミュレーションソフトを使って、その問題点を考察するこころみをおこなっ

ている。(注9)

この過程で、まず問題となったのは、スキャナで辞書の印刷された色を採取し測定することが、はたして可能であるかどうか、という点である。単に、紙に印刷されたものをスキャンするだけのことなのであるが、実際には、実に多様な不確定なパラメータあるいはノイズがからんでくる。

(1). 機種の選択

現在、主な市販のスキャナといえば、メーカはCANONとEPSONになるが、それぞれに採取する画像データの色彩は異なる。また、同じメーカであっても、機種が異なれば、機能の差がある。同じ機種であっても、個体差のあることが予想される。

(2). 同一機種でも常に同じではない

まったく同一の機械であっても、機械である以上、作動時の誤差が生じる。また、電源の影響も無視はできない(注10)。同じ画像をスキャンするにしても、時によって色彩が違ってくるのである。もちろん、事前のキャリブレーションは、不可欠である。

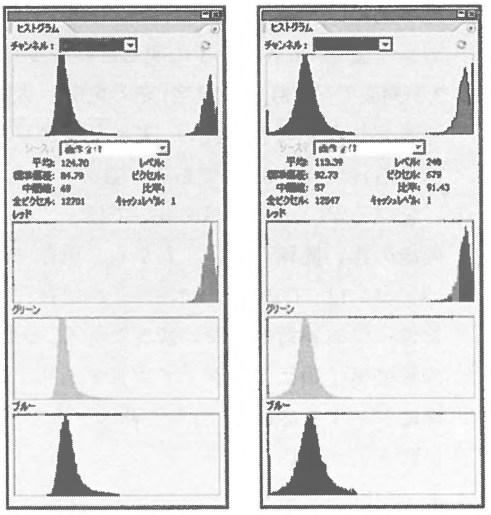
(3). 色彩の表現・構成要素には種々の方式があるが、色相・彩度・明度、による方法がある。Photoshopのカラーピッカーでいえば「HSB」、MS-Officeでいえば「HLS」が該当する。スキャナは、色彩のこれらの要素について、またRGBの各色について、同じように反応するわけではない。それぞれに固有の癖があるし、さらには、メーカによる意図的なバイアスがかけてある。

(4). CANONにせよEPSONにせよ、自動のカラー調整機能がついている。EPSONであれば「自動露出」であるし、CANONであれば、「自動色調整」である。その他、「アンシャープネスフィルタ」(EPSON)「輪郭強調」(CANON)など、多種多様な自動調整機能がある。実際に使ってみると、確かに便利である。よりはっきりしたメリハリのある画像データが得られる。しかし、これは、逆に言えば、色彩をある方向に強調することによって、よりそれらしく見せている、つまり、色彩を変化させてしまっていることになる。利用にあたっては、どのように色彩が変化するかを確認しておく必要がある。〈図表(2)自動調整機能の影響〉に、『初級クラウン英和辞典』の「A」の例をしめす(1200dpi)。

ただ、ここで重要なのは、これらの機能を使用しないことにあるのではない。場合によっては、使ってもよい。だが、使うか使わないかは、研究者が、対象とする資料の特性と利用目的に応じて、学術的責任のもとに判断しなければならないことである。この点が、重要なのである。

図表(2)自動調整機能の影響

自動色調整オフ 輪郭強調オフ	自動色調整オン 輪郭強調オン



(5). 画像の精度と保存方式

簡単に言えば、高精細な画像データであればメモリを大量に必要とするし、処理速度もおそくなる。低精度で、かつ圧縮(JPEGなど)した画像データであれば、軽くあつかえる。色彩の点についてのみ述べるならば、スキャン精度は、色彩にはさほど影響を与えない。理想的に均一に塗りつぶされた印刷面のスキャン画像データとしては、300dpiでも1200dpiでも、同じになる。まったくの均一色であれば、ヒストグラムは縦の一直線になる。しかし、現実には、印刷むらや、紙の裏面の文字の影響などが大きく作用するので、これらを観察するためには、より高精細なスキャンが必要になる。

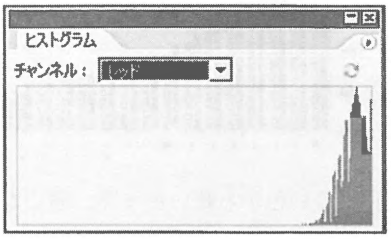
画像データの保存方式としては、なるべく圧縮は避けた方がよい。少なくとも、JPEGは色彩が変わってしまう可能性(危険性)がある。「色彩の保存」という観点からは、避け

るのが賢明である。少なくとも、最初のオリジナルのデータを保存するときの１回に限定すべきであろう。

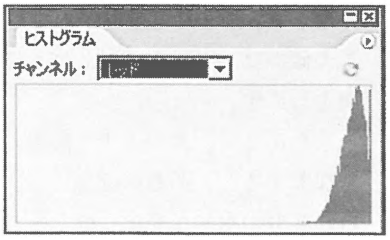
図表(3)(4)(5)に、画像データとそのRのヒストグラムを示す。(5)は、1200dpiの画像(3)を低圧縮(高画質)で、５回のJPEG圧縮を経たものである。全体の色彩の計測値としては、ヒストグラムの平均値(あるいは中間値)を採用するため、さほどの影響はない。しかし、(4)と比較すると、微細な色調のディテールが失われていることが見てとれる。

図表(3)初級クラウン「d」

図表(4)BMPヒストグラム



図表(5)JPEGヒストグラム



スキャナで印刷物を画像データ化するだけでも、少なくとも以上の各パラメータについては、確認をとらなければいけない。どういうバイアスのかかった画像データを使うことになるのか、自分自身できちんと把握しておかなければならないのである。

これが、デジカメ画像や、従来のフィルムやプリントされた写真のスキャンになると、印刷物のスキャン以上に、さらに多くの不確定要因がからんでくる。

(6). 撮影の光源。蛍光灯のもとの色彩と、太陽の自然光による色彩とでは、違う。ただ、人間の眼による色彩の認識は、それを自動補正するので、同じに見えてしまう(注11)。しかし、機械であるカメラやフィルムは、このような融通がきかない。この場合、問題なのは、光の明るさではなく、色温度である。

(7). フィルムの種類。旧来の銀塩写真における昼光用フィルム(太陽光のもとでの撮影用)であっても、メーカによって(コダックとフジとで)違う。同じメーカであっても、フィルムの種類(エクタクロームとコダクローム)で違う。同じフィルムであっても、製造のロットナンバー(乳剤番号)によって発色が異なる(ちなみにプロの写真家は、フィルムのロットナンバーまで管理する)。もちろん、その後、現像・プリントによる発色の差は言うまでもない。

(8). レンズによる色の違い。カメラ用のレンズには、必ずなにがしかの色収差がある。光の波長による屈折率の違いで(つまり、プリズムの原理)、すべての光が一箇所に集まらない。レンズによって、色がついてしまうのである。

(9). デジタルカメラであれば、当然ながら、そのメーカや機種固有の発色の癖があるはず、と考えておかなければならない。同一機種でも、個体差を配慮しなければならない。

画像データをあつかう以上、少なくとも上記の、不確定なパラメータ、場合によってはノイズを、把握しておく必要がある。これは、色彩を正確に再現する、ということではなく、どのようなバイアスがかかった画像データであるのか、そのバイアスの方向と程度を知っておくことである。

従来ならば、これは、まさにプロの写真家や印刷業にたずさわる者の専門知識に属することがらであった。一般の利用者(人文学系研究者)は、それを、とりあえず信じていればよかった。あるいは、影印複製とはこの程度のものと、あきらめていればよかった。しかし、コンピュータの技術の進歩、スキャナやデジカメの普及は、それではすまない状況を生み出してきている。画像データの色彩においても、研究者としての学術的判断と、それにともなう責任が生ずる、と筆者は考える。

【6】スキャナ画像の具体的問題

以下、筆者の実験した辞書のスキャナ画像データの諸問題について述べ、筆者の考えたプロセスを明らかにしておきたい。実験の対象としては、一般的に業界標準とでもいうべき、『Kodak Color Separation Guide and Gray Scale』〈図表(5)Kodak Gray Scale〉をもちいる。

(1). 機種による違い

筆者が使用しているのは次の2機種である。

EPSON GT7400U

CANON LiDE500F

それぞれで、スキャナの自動補正機能をすべてオフにした状態でのRGB値測定の結果は次のようになる。150dpi、保存形式はBMP。花子フォトレタッチ3(ジャストシステム)で表示。測定は、前述『色々の色』による(ディザパターン8ピクセルで各色の中央部を計測)。

EPSON GT7400U	シアン	69・160・243
	イエロー	254・239・56
	マゼンタ	234・53・127
CANON LiDE500F	シアン	55・161・234
	イエロー	254・246・0
	マゼンタ	248・39・93

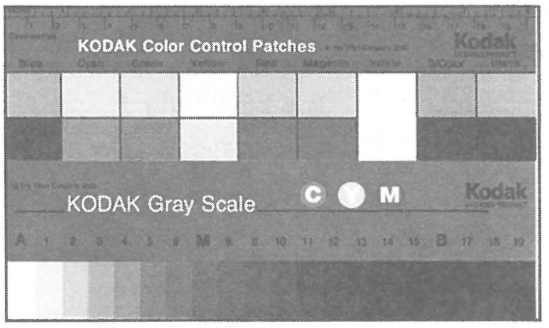
シアンはあまり差が出ていないが、イエローとマゼンタでは、かなり測定値が異なる。実は、どちらが正しいのか、これだけでは分からない。『Kodak Color Separation Guide and Gray Scale』には、それをスキャンした時のRGB値やLab値が示されていないからである。なぜならば、これは、あくまでも写真撮影において、写しこんだものと同じ色彩を、印画紙やポジフィルム上で再現・確認することを意図した製品だからである。

辞書という厚みのある本のスキャンということで、あつかいやすさの観点から筆者は、CANONを使うことにした。この機種は、蓋を完全に平らに開くことが出来るので、辞書などのスキャンには都合がよい。

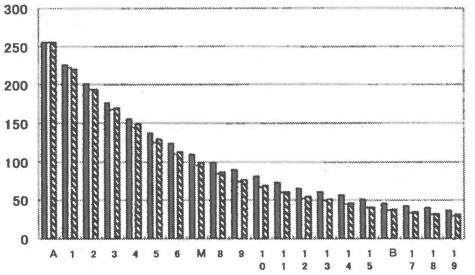
(2). CANON LiDE500Fの特性

では、「CANON LiDE500F」というスキャナは、どのような特性(色彩のバイアス)を持っているのか、確かめなければならない。今度は、Kodakの『Gray Scale』をスキャンしてみる。同様に、自動補正機能をオフにしてカラーモードでスキャンした結果の、各段階のRGB値をグラフ化すると、〈図表(6)GrayScale測定値〉のようになる。これの測定は、PhotoshopCS・CS2のカラーピッカーによる。

図表(5)Kodak Gray Scale



図表(6)GrayScale測定値



[1]. 明るい色から暗い色まで、同じようには反応していない。明るい箇所には敏感に反応するが、暗い箇所では鈍くなる。ある程度以上暗くなると、ほとんど差異が出なくなる。なお、この傾向は、EPSONで試してもほぼ同様の傾向をしめす。

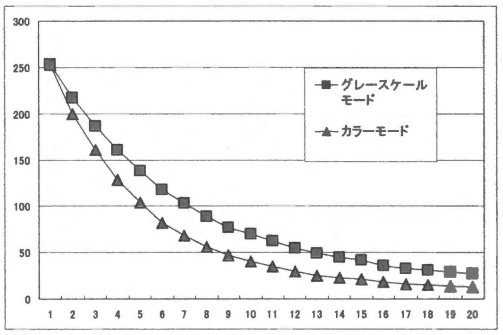
したがって、スキャナで辞書の画像データを採取すると、明るい部分、つまり、用紙の白い部分の微妙な変化には、きわめて敏感に反応する。裏側に印刷してある文字の色が用紙を透かしてかすかに見えるが、それまで忠実に取りこんでしまう。もし、仮にこのスキャナで古写本などを直接スキャンするとすれば、料紙の色合いが非常に強調されることになるし、また、裏側の文字まではっきりと読み取れることになる。状況によっては、これがノイズになることもあり得る。

[2]. 『Gray Scale』は、基本的に無彩色のはずであるが、RGB値にばらつきがある(色がついている)。肉眼で『Gray Scale』を目視して見て

も、かすかではあるが色がついていることが認識できる。これは、『Gray Scale』自体の本来の目的が、白黒写真フィルム(銀塩写真)を対象として、その白(ハイライト)から黒(シャドー)までの階調を整えるためのものであることに起因すると、筆者は理解している。言い換えれば、本来スキャナのテスト用ではないのである。『Gray Scale』の説明書を読んでも、20段階で等間隔に明度差のある旨は記してあるが、色彩(色相・彩度)についてはまったく言及していない。

[3]. そのことの確認のため、スキャナを「グレースケール」の設定に変えて、同じ『Gary Scale』をスキャンしてみる。それを、先にカラーでスキャンした画像データをPhotoshopの機能で無彩色化(イメージ → モード → グレースケール)したものと比較してみる。各色均一の数値にはなるが、各段階のＲＧＢ値の測定ができる。結果は、〈図表(7)グレースケールとカラーモード〉のようになる。スキャナを「グレースケール」で使用した方が、ゆるやかなカーブを描く。

図表(7)グレースケールとカラーモード



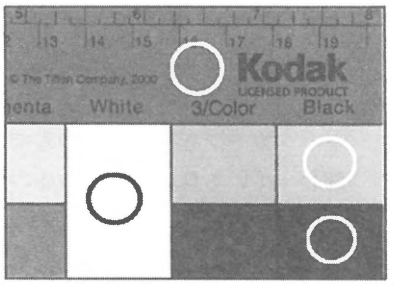
これから言えることは、例えば、墨色による版本・写本などの画像データを、モノクロ画像で見るということを前提とするならば、同じデータサイズ(メモリ)であるならば、最初から、データをスキャンする段階で、グレースケールモードによる画像データ化をしておく方がよい、その方が、より豊かな階調を得られる、ということである。

画像データについて論じるとき、その精度(dpiやpixel数・画素数)をよく問題にする。だが、画像の情報量として重要なのは、画素数や精度だけではなく、本質は階調や色彩の豊かさであることを忘れてはならない。カラーだから情報(学術的情報)が豊富であるということはないのである。

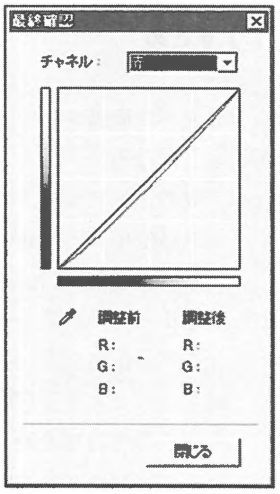
[4]. RGBのバランス

上記の『Gray Scale』の画像データを見る限りでも、かなりRが強いことがみてとれる。実際に、画像データをディスプレイで表示(sRGBモード)してみても、実物と比べて赤みがかっていることが観察される。これは、たまたま筆者の所有しているものの固有の特性かもしれないが、やはり、なにがしかの補正の必要がある。しかし、ここで『Gray Scale』は使えない。

図表(8)Color Control Patches部分図



図表(9)補正トーンカーブ



前述のごとく、これはあくまでも白黒フィルム用であると、筆者は理解している。RGBのバランスを整えるのには、『Gray Scale』の方ではなく、『Color Control Patches』の方をつかうことにした。こちらの方は、Kodakが無彩色であることを指定している箇所がある。〈図表(8)Color Control Patches部分図〉の○印の部分である。

それぞれ、18%反射・25%のグレー・ブラ

ック・ホワイト、と説明書には記されている。なお、18%反射というのは、写真の世界でいう標準反射率である。反射光式露出計では、撮影にあたって18%反射板を測定することが基準であるし、入射光式露出計ではこの測定値を示す。また、現実には、『Kodak Color Separation Guide and Gray Scale』においては、これらの箇所しか、RGB値が保証されているところは無いのである。(注12)

結果としては、スキャナのトーンカーブ補正を、Rを12ポイントずらすことで、どうにかバランスをとることが出来た〈図表(9)補正トーンカーブ〉。このとき、基準としては、RGB値を測定して、揺れの範囲が５以内であることと、Lab値の測定で、a*とb*の値が、0±1の範囲内であること、とした。

[5]. 自動補正機能はオフ

トーンカーブを補正するのみで、スキャナに備わっている自動の調整機能の類は、すべてオフで使用した。前述のごとく、自動調整機能で、どのように色彩が変化するか、RGBヒストグラムにおいて確認はしてある。

【7】まとめ

以上に述べた画像データの問題点を要約すると、次のようになる。

(1). 一般のコンピュータであつかえるカラースペース(sRGB)と、印刷(CMYK)、人間の見る色彩(xy色度図)には、ガモットの制約がある。

(2). 画像データの入力(スキャナ・デジカメ)、保存形式、出力(ディスプレイ)には、それぞれ固有のバイアスや制限がある。また、利用目的に応じた設定変更が必要である。

(3). 『Kodak Color Separation Guide and Gray Scale』は、デジタル画像データの汎用的な色彩の基準としては使えない。

(4). 同一画像データであっても、Photoshopで見るとのと、IEで見るのとでは、ディスプレイに表示される色彩が異なる。

リベラルアーツ(人文学的基礎教養)としてのデジタル色彩学が、以上のような問題意識のもとに構築されなければならない。画像データの色彩にも、学術的判断と責任がともなうのである。

【8】おわりに ― 色彩の基準と距離感覚 ―

我々は、日常、色彩にあふれた生活をおくっている。しかし、身の回りにあるもので、確実に、これは完全な黒である、sRGBで完全な赤(R255)である、というものが存在するだろうか？　一般的には、『Kodak Color Separation Guide and Gray Scale』がある。しかし、これが汎用的な色彩の基準とならないことは、述べたとおりである。(注13)

純粋に技術的な問題としては、ＴＶやインターネットで見る画像の色彩については、何の保証も無いのが実際である。定まっているのはNTSCやsRGBの規格であり、それを実現しようとする技術の努力だけであるといってもよいであろう。

だが、にもかかわらず、日常的に色彩名を使用して、コミュニケーションが成立している。さらに、人間の色彩の認識では、光源(色温度)が変わっても、同じ色を同じ色として見るようになっている(メタリタリズム)。これに、言語文化における色彩の認識と構造が加わる。

人文学の研究資料の画像データ化という点に限っていうならば、現代色彩科学の成果を積極的に利用すべきである。コンピュータで色彩が簡単にあつかえるようになり、手軽に、スキャナ画像やデジカメ画像を利用できるようになっている。また、フォトレタッチソフトも、廉価で高機能なものが多数ある。誰でもが自由に色彩をあつかえるようになった。だからこそ、コンピュータで使える色彩と、人間が見る色彩の世界との違いについて、正しい理解が必要になってきている。

重要なのは、実物との距離感覚である。古典籍を見慣れた研究者ならば、白黒の影印本であっても、およそ墨筆と朱筆の区別はつく。また、カラーの複製本でも、実物との違いを自然に感知できる。カラー画像データについても同じである。実物との違いを、どこまで直感的に看取できるかが、問われる。そのためには、より多くの実物(原本)に接するより他に道はない。コンピュータによる画像データベースが容易に利用できるようになればなるほど、実物を見ることの重要性が増してきている、このように筆者は認識する次第である。

以上、画像データをめぐる色彩の同一性の保証という観点から、いくつかの問題点を考察してきた。これは、あくまでも一般論としてであって、

特定の画像データベースを批判する意図はないことを、最後に強調しておきたい。筆者は、その作成者の努力に多大の敬意を表するとともに、それが今後より有効に利用されるためには、人文学研究者にも現代色彩科学の基礎的知識が不可欠であることを、主張したい。専門知識としての色彩科学ではなく、リベラルアーツ(人文学的基礎教養)としてのデジタル色彩学である。

[参考文献]
(1)千葉憲昭(2004)、『カメラ・常識のウソ・マコト』、講談社
(2)千々岩英彰(2001)、『色彩学概説』、東京大学出版会
(3)池田光男・芦澤昌子(1992)、『どうして色は見えるのか―色彩の科学と色覚―』、平凡社
(4)池田光男(2003)、『色彩工学の基礎』、朝倉書店
(5)池田光男・池田幾子(1995)、『目の老いを考える』、平凡社
(6)池田光男(1988)、『眼は何を見ているか―視覚系の情報処理―』、平凡社
(7)金子隆芳(1990)、『色彩の心理学』、岩波書店
(8)金子隆芳(1995)、『色の科学―その心理と物理―』、朝倉書店
(9)金子良二(2005)、『コンピュータで知る 色彩と画像表現』、九天社
(10)金子隆芳(1988)、『色彩の科学』、岩波書店
(11)川添泰宏(1996)、『色彩の基礎―芸術と科学―』、美術出版社
(12)栗野輪美(他)(2001)、『天空からの虹色の便り マルチメディア 宇宙スペクトル博物館〈可視光編〉』、裳華房
(13)村上元彦(1995)、『どうしてものが見えるのか』、岩波書店
(14)南雲治嘉(2003)、『デジタル色彩表現』、グラフィック社
(15)(財)日本色彩研究所編(2004)、『デジタル色彩マニュアル』、クレオ
(16)太田登(2001)、『色彩工学(第2版)』、東京電機大学出版局
(17)大山正(1994)、『色彩心理学入門―ニュートンとゲーテの流れを追って―』、中央公論新社
(18)世良歩(2005)、『カラーマネジメントガイド』、アスキー
(19)鈴木孝夫(1990)、『日本語と外国語』、岩波書店
(20)當山日出夫(2005)、「色覚異常者にカラー印刷の辞書はどう見えるか―スキャナ画像データの利用をめぐる諸問題―」、『情報処理学会研究報告(CH-68)』、情報処理学会
(21)内田恵二(1998)、『色覚のメカニズム―色を見る仕組み―(色彩科学選書 4)』、朝倉書店
(22)矢部國俊(2004)、『標準DTPデザイン講座 Photoshop』、翔泳社

[参考HP]
(1)http://www.hikarun.com/w/『色々の色』(中原ひかる)
(2)http://konicaminolta.jp/entertainment/colorknowledge/index.html『色色雑学』(コニカミノルタ)
(3)http://www.adobe.co.jp/support/techguides/color/colormodels/main.html『Technial Guides』(Adobe)

(注1) 五味文彦(2005)、『絵巻で読む中世』(ちくま学芸文庫)、筑摩書房、原本は1994年、筑摩書房(筑摩新書)
(注2) 宮本常一(1981)、『絵巻物に見る日本庶民生活誌』(中公新書)、中央公論新社
(注3) 一般のパソコン用カラープリンタでも、CMYの他にK(黒)のインクを必要とする。
(注4) プロの印刷・出版業の世界では、異なるカラースペース間のカラーマネージメントが、最重要の課題である。
(注5) 参考HP(3)
(注6) ちなみに、筆者の使用しているモニタは、NANAOのEIZO FlexScanM170 であるが、sRGBモードでは6500K、文書作成用のテキストモードでは5000Kになる。
(注7) 参考文献(1) pp.149-50
(注8) 参考HP(1)
(注9) 参考文献(20)、2005年10月28日、於高岡万葉歴史館
なお、次の論文は、すでに入稿済みである。
當山日出夫(2006a)、「英語辞書におけるユニバーサルデザイン―初級英和辞書を色覚異常の観点から検証する―」、『福祉と人間科学』(第16号)、花園大学社会福祉学会
當山日出夫(2006b)、「全訳古語辞典のカラー印刷と色覚異常―ユニバーサルデザインの視点から―」、『日本語辞書学の構築』(仮題)、おうふう
(注10) ディスプレイでも、FlexScanM170の説明には、調整は電源を入れてから20分以上経過しておこなう旨、注意が記されている。つまり、それだけ時間がたたないと、機械として安定しないのである。また、マニアックなオーディオの世界では、「いい音」のためには、電源の工事から始めるのが常識である。それほど電気を使う機械は、電源の安定性に依存する。
(注11) 色彩科学の用語では、メタメリズムという。
(注12) あくまでも、絶対的なRGB値ではない。そのバランス(つまり、グレーである、色がついていないこと)を、意味するだけである。
(注13) ただ、高価格ではあるが、写真の世界の標準として、マクベスカラーチェッカーがある。これは、各色についてRGB値が指示してある。また、『Kodak Color Separation Guide and Gray Scale』は、紙の印刷物であるため、光にさらされることによる退色・劣化や、使用によるよごれの付着等は、避けられない。そのため、説明書には、頻繁に利用する場合は、耐用期間は1ヶ月と記してある。

万葉集校本データベースから古字書データベースへ

From Manyoshu variorum database to the old dictionary database

西端　幸雄
Yukio Nishihata
大阪樟蔭女子大学　学芸学部，東大阪市菱屋西　4-2-26
Osakashoin Wemen's University, 4-2-26, Hishiyanishi, Higashiosaka-shi, Osaka

鈴木　榮一
Eiichi Suzuki
万葉情報システム調査会
Investigation Committee for information system about "Manyoshu"

あらまし:万葉集の主立った古写本・版本の画像データを対比できる「万葉集校本データベース」の紹介と、現在構想中の古字書データベースの紹介、さらに、将来、それぞれをどのように関連付けていくかを説明する。

Summary: We introduce "Manyoshu variorum database" which can compare image data of main copied books of old time and printed books of Manyoshu now introduce the old dictionary database elaborating a plan again and explain how We connect those databases in the future.

キーワード:万葉集，校本，データベース，古字書
Keywords: Manyoshu , variorum , database, old dictionary

1. はじめに

国文学の諸領域の中で、上代文学の分野は、これまで電子メディアに最も縁が薄かった。

そのため、1997 年に、有志が集まり、「万葉情報システム調査会」を結成した。この会の結成の趣旨は、上代文学の分野における先学の諸研究から最新の研究成果までをデータベース化し、それらを幅広く効率的に検索するシステムの実現を目指すというところにある。

その構想したところについては、すでに、シンポジウム「人文科学における数量的分析 IV」及び「第 11 回文献情報のデータベースとその応用に関する研究会」(1999.3:総合研究大学院大学)と「第 14 回 西日本国語国文学データベース研究会(DB-West)」(1999.6:大阪樟蔭女子大学)において発表した。

そして、その構想の一端は、本「万葉集校本データベース」の公開と『CD-ROM版萬葉集』(2002:塙書房)『萬葉集索引』(2003:塙書房)の刊行という形で実現させたのであるが、さらに、文字データをはじめ画像データなどをも含めたデータベース構築等といった複数のアイデアの実現化へと向かっている。その一つの活動が、本稿第５項以下で紹介している「古字書データベース」である。

また、将来的には、これらのデータベースをより充実させ、さらに公開することにより、様々な研究目的への用途が見込まれ、上代文学の研究が飛躍的に発展するだろうと確信している。

2. 万葉集研究の現状

ところで、古代日本の文化と日本人の精神生活を窺うことのできる古典として、また、当時の事物や社会の様相を研究するために欠かすことのできない第一級資料として、古来、その重要性が広く江湖に認められている『万葉集』は、原本が現存しないため、その原文復元は大変困難な作業であるといえよう。

古典の基本的研究は本文の吟味にはじまり、その本文の吟味は諸本の対校によるといわれる。この対校作業は古典研究の基礎であるが、古写本・版本の形でしか残っていない『万葉集』においては殊にそれが重要である。

そして、その目的を実現すべく、かつて、1912年文部省文芸委員会において、国家の事業として『万葉集』諸本の対校と定本編纂をする事となり、1924年には『校本萬葉集』(現在、新増補版『校本萬葉集』(岩波書店)全21巻が刊行されている)が刊行された。さらに、また、我が国の古典研究における単語索引の先駆けとして、『万葉集』に使用された単語の全てが検索できる『萬葉集總索引』(正宗敦夫)が1929〜31年に刊行されたのである。

しかしながら、現在の古典研究はこれらの事業の上に立ってさらに深化してきており、本文研究においても詳細を究め、『校本萬葉集』だけでは十分とはいえず、また『校本萬葉集』自体、非常に見にくく(図１参照)、また、誤りが散見でき、各種古写本・版本を直接見て対校すべき必要が生まれている。とはいえ、現実的には、『万葉集』の主立った伝本だけでも、30点を優に超えており、それぞれを刊行された書籍の上で相互に対校するといった作業は煩雑を極め、多大な時間と労力を要し、容易ではない作業となる。

3. 万葉集校本データベースの紹介

そこで、「万葉集校本データベース」では、使用の許可が得られた諸伝本について、そのすべてを画像データとして取り込み、１首単位に切り出し、同一歌番号の同一句の画像データを並列させ、画像を左右に移動させることにより、各伝本間の異同の有無を容易に対校できるようにした。

このデータベースの完成により、検索したい歌番号を入力すると、瞬時に諸伝本の画像データに直接あたって対校でき、これまで書籍の上で行っていた対校作業に比べ、簡便に、また、より一層精緻な本文校訂・付訓が可能となり、古写本が伝写されてきた過程において発生したであろう誤写・誤脱・誤読の特徴が明らかになり、本文校訂・付訓の作業を能率的に、かつ正確に行うことができるようになるものと確信している。また、これまで誤りとされていた本文や付訓を訂正することが可能となり、さらに、難訓歌とされている歌の理解も進めることができる。

この「万葉集校本データベース」のデータ構成は、以下の通りである。

Ⅰ画像データベース　本データベースの根幹を成すデータであるが、寛永版本（使用許可済）の画像データを基軸とし、歌単位（異伝含め約6000首）で切り出し、句毎に整理し、他の古写本画像データ（『万葉集』の古写本『桂本』、『藍紙本』、『元暦校本』、『広瀬本』、『西本願寺』、『京都大学本』等、約12種類）と、句毎にそれぞれの画像を並列させたものである(図２参照)。

Ⅱ翻刻本文データベース　鎌倉時代の仙覚に始まって、室町・江戸時代を経て今日に至るまでの『万葉集』の注釈書における歌本文の翻字データである。手写された『万葉集』の古写本を、注釈を行った人々がどのように読解していたかを知るための重要なデータである。ここで、採用している注釈書は、古注釈としては、『萬葉集註釈』『萬葉代匠記』『萬葉集童蒙抄』『萬葉考』『萬葉集略解』『萬葉集槻の木』『萬葉集攷証』『萬葉集古義』をはじめ約20種類の注釈書、また、明治以降の注釈書としては、『万葉集註疏』(近藤芳樹:1868頃)に始まり、現代の『万葉集釈注』(伊藤博)に至る約20種類の注釈書を取り上げている。

Ⅲ訓み本文データベース　Ⅱの翻刻本文データと対をなすデータで、『万葉集』の注釈書において、万葉仮名、または漢文形式で書かれたりしている歌本文を、和語として訓んだ結果をテキストデータ化したものである。という点で、これも、注釈を行った人々がどのように読解していたかを知るための重要なデータである。

Ⅳ注釈データベース　仙覚以降の主立った注釈書に掲載されている各歌毎の注釈データを一覧できるものである。

Ⅴ口訳データベース　Ⅳの注釈データベースと

対をなすデータで、各注釈書において、Ⅳの注釈に基づいて各歌の意味をどのように解釈しているかが一覧できるデータである。口訳データとしては、当然のことではあるが、明治以降の、『万葉集註疏』(近藤芳樹:1868頃)に始まり、現代の『万葉集釈注』(伊藤博)に至る注釈書を網羅している。

本データベースは、1999年度から、「寛永版本」の歌単位画像データ、各種古写本の歌単位画像データ、各種注釈書の歌単位テキストデータの入力作業を行い、基本歌テキストデータとリレーションさせ、2003年度に完成させ、万葉情報システム調査会が運営管理しているホームページ (http://www.manyou.gr.jp/SMAN_1/) でデータを公開している。

また、近い将来、DVD-ROM版でのデータ提供も予定している。

4. 万葉集校本データベースの利用

次に、その「万葉集校本データベース」の利用の一端を紹介する。

既に述べたように、奈良時代に成立した『万葉集』の原本は失われているが、平安中期以降の古写本・版本が多数残っている。ところが、長い間伝写が繰り返されてきたため、それらの中には、本文の乱れ(誤写・誤解)が散見される。そのため、中には、今日でも訓みが確定しない歌もあるほどである。

今、提示するのは、大和三山が妻争いをしたという伝説を背景として歌われた有名な歌である。

その末句において、古写本・版本間での本文の乱れが激しいのであるが、その点を、本「万葉集校本データベース」を利用して、各古写本・版本間の句データを比較すれば、歌本文の校訂が行えるという点を述べておく。

かぐやまは　うねびををしと　みみなしと　あひあらそひき
高山波　雲根火雄男志等　耳梨與　相諍競伎
かみよより　かくにあるらし　いにしへも　しかにあれこそ
神代従　如此尓有良之　古昔母　然尓有許曽
うつせみも　つまを　あらそふらしき
虚蝉毛　嬬乎　相 挌良思吉(13)

その異同の状態を、まず、『校本萬葉集』によって示すと、図１のようになるのだが、異同の状態の状態を知る手掛かりにはなるが、略号が使われていたりして、各古写本・版本間の比較を行う場合には、かなり使い慣れないと分かりにくい面がある。

一方、本「万葉集校本データベース」で、当該句の異同を調べると、図２のように表示され、各古写本・版本を直接に見ながら比較ができる点で、正確な対校作業ができる。

ところで、今、問題として取り上げる末句を翻字すると、以下のようになる。

［寛永版本］　相 格(アヒウツ)良(ラ)思(シ)吉(キ)
左傍訓「アラ」

［元暦校本］　相 挌(アヒテカク)良(ラ) 思(オモヒ) 吉(キ)

［広瀬本］　相 挌(アラソフ)良(ラ)思(シ)吉(キ)

［紀州本］　相 挌(アヒウテヲシ) 良(トヲ) 思(モキヤ) 吉
右傍訓「アヒテカタ」「オモヒ」を消す。左傍訓「アヒテウチ？？オモヒキ」を消す。

［神宮文庫本］　相 格(アヒウツ)良(ラ)思(シ)吉(キ)
左傍訓「アヒウチヨシト」

［西本願寺本］　相 挌(アヒウツ)良(ラ)思(シ)吉(キ)
左傍訓「アヒテカクラオモヒキ」

［京都大学本］　相 挌(アヒウツ)良(ラ)思(シ)吉(キ)
左傍訓「ミルニテカクオモヒ」

［陽明本］　相挌(アイ)良(ラ)思(キ)吉
左傍訓「ミルニテカク　オモヒ」

以上のように、各古写本とも、本文の表記と訓みが一定していない。

まず、本文の表記としては、２文字目の用字について、寛永版本と神宮文庫本が「格」字としているのに対して、その他の古写本は「挌」としているという揺れがあるが、『説文通訓定聲』に「格、叚借為挌」とあり、両字は、相通する。その点で、当該箇所の表記は、「挌」でも「格」でもよいということになるが、平安期の成った元暦校本や類聚古集が「挌」字を用いている点で、元もと「挌」字であったとしてよいと思われる。

また、この句の訓みであるが、単純に二分するなら「相挌」の箇所を「あひうつ」と訓むか「あらそふ」と訓むかの問題となるのだが、『説文解字』に「挌、撃也」とあり、『広韻』に「挌、闘

也」とある。ということは、可能性として、「うつ」も「あらそふ」も成り立つということになる。その点で、漢字の表す意味から訓みが確定できないことになる。そこで、この 13 番歌全体の解釈を行ってみると、「香具山は畝傍山を取られるのが惜しいと、耳梨山と争いあった。神代から、このようであるらしい。昔もそうだったからこそ、今の世の人も、妻を……」(新日本古典文学大系による)ということになる。つまり、「昔もそのようであった(いにしへもしかにあれこそ)」だから、「今の世の人も(うつせみも)」そうであるということになると、4句目の「相諍競伎(あひあらそひき)」を受けた句が、当該の末句であるということになる。以上の点からして、この末句の訓みは、「あらそふらしき」とする方がよいということになる。

以上に説明したように、本「万葉集校本データベース」そのものが、本文の校訂の答えを提供はしてくれないが、まず、問題点の発掘と、その問題解決の糸口を提供してくれるものであると位置づけられよう。

5. 古字書データベース

以上に紹介してきた「万葉集校本データベース」と関連するデータとして、我々が最もその必要性を感じていたのが、我が国の平安時代から鎌倉時代初期に編纂された古字書データベースである。

その必要性の背景には、万葉集の時代、人々が漢字理解を行う場合、中国から渡来した字書を元にしていたということがある。ただ、その当時の日本に渡来していただろう中国字書は、『切韻』にしろ『玉篇』にしろ、完本の形では現存していない。そのため、それら中国字書を抄録したり、抜粋して利用している我が国の古字書を見る必要がある。

しかし、今日、我が国の古字書の整備は、本文の影印や漢字・和訓索引が刊行されているという点では、進んでいるとも言える。しかし、特に古字書というものは、過去の字書や文献を引用することによって編集されたり、増補されてきた。その点では、複数の古字書を縦断的に、また、網羅的に見る必要があるのだが、そうした機能を備えた資料は、『古字書綜合索引』(長島豊太郎編:1957)が唯一である。その内容は、8種類の彼我の古字書の見出し漢字索引である点では、上記の条件を備えていると言える。ただ、本書における見出し漢字の表記が手書きである点で、漢字についての相当な知識を有するものしか使えないという問題と、すでに絶版になっていて、入手が困難であるという問題を抱えている。

また、古字書も、他の古文献と同様、原本は存在せず、写本・版本の形でしか伝わっていない。そうすると、それらが伝写されてきた過程において発生したであろう誤写・誤脱・誤読も当然残存している。ところが、現在のところ、それらの誤りは部分的に指摘されてはいるが、統括的な指摘はまだ成されていないのが現状である。そうした場合、先に紹介した「万葉集校本データベース」と同様に、各古字書の同一見出し漢字の本文を並列的に並べることで、相互比較が容易にできるのではないか、さらに、多くの古字書を網羅的に見ることによって、『万葉集』をはじめとする上代文献の読解が進むのではないかというところから、本「古字書データベース」作成を計画したわけである。

この計画では、『新撰字鏡』『篆隷万象名義』『類聚名義抄』『和名類聚抄』等のデータベース化を図る予定である。

6.『新撰字鏡』のデータベース化

そこでまず、初年度には、平安時代の寛平4年(892)に草案本三巻が成り、昌泰年中(898 ～ 901)に十二巻に増補された『新撰字鏡』をデータベース化することにした。

この『新撰字鏡』を選択した理由は、まず、我が国の漢和辞書としては、現存最古のものであるという点と、その成立が、中国の『一切経音義』『切韻』『玉篇』の三本に、すでに現存しない我が国の古字書等を併せて編集したものであるという点と、さらに、後世の『世尊寺本字鏡』『天理図書館蔵観智院本類聚名義抄』に影響を与えているという点で、最初にデータベース化する意義のある資料であると判断したからである。

現存する『新撰字鏡』は、天治元年(1124)に法隆寺一切経の一部として書写された十二巻本系の天治本と、享和3年(1803)の序を有し、十二巻本系から和訓の注を有する項目だけを抜き出した享和本、享和本の校異に用いられた抄録本の群書類

従本の、三種類があるが、そのうちデータベース化するのは、『天治本新撰字鏡』である。

この『天治本新撰字鏡』は、所収見出し漢字数、約 21,300 字で、そのうちに、約 3,000 の和訓が存する。字書としての体裁は、基本的には部首分類体をとっているが、それぞれの漢字は、今日の漢和辞典のように、画数順に並べるなどといったように、整然と整理はされていない。

その和訓を検索するための索引は、『新撰字鏡国語索引』(京都大学文学部国語国文学研究室:1958)があるが、見出し漢字を検索するための索引としては、前述の『古字書綜合索引』があるだけで、『新撰字鏡』を漢字辞書として利用しづらいのが現状である。

そこで、『天治本新撰字鏡』として刊行されているうち、本文の状態がよい、大槻文彦によって編纂され、1916 年に六合館から出版されたものを使用して、全巻をスキャニングし、さらに、見出し漢字と注文を切り出し、画像データベース化している。

現在、この『天治本新撰字鏡』データベースをテストデータとして、検索方法の検討、および、画像の表示方法などの検討を行っているが、特に、漢字の検索方法については、ＪＩＳ外字が多いため、多角的な検索を可能とするように努めている。

その漢字検索方法のパターンは、以下の通りである。

①大漢和番号検索
②自由入力検索(ＪＩＳ内字のみ)
③複合検索(１つ以上の条件で検索可能)
- 部首
- 漢字音
- 画数

①の大漢和番号検索は、『諸橋大漢和辞典』(大修館書店)に付されている通し番号を入力して、当該漢字を検索する方法である。②の自由入力検索とは、パソコンが通常扱える漢字(ＪＩＳに登録されている漢字)を入力して、検索する方法である。また、③の複合検索とは、検索したい漢字の部首か漢字音か総画数が分かっている場合、それらのうち、１つ以上を条件として検索する方法である。

これらの内、どれかの方法で検索し、該当する漢字がない場合には、「漢字が存在しません」と表示をし、該当する漢字が存在した場合には、それらの漢字すべてを表示できるようにしている。(『新撰字鏡』においては、同一漢字が複数回、見出し漢字になっているケースは少ない)

今は、図３に、③の複合検索を行った例を挙げておく。

部首として「日」、漢字音として「ケイ」、画数として「8」を与えて、[検索する]をクリックすると、図４の検索結果画面に切り替わる。

現在は、『天治本新撰字鏡』のデータしか取り込んでいないが、将来は、『篆隷万象名義』『類聚名義抄』『和名類聚抄』等のデータベース化も計画しているので、図４に表示されている画像の横に並列して、他の古字書のデータも並ぶことになる。

7. 古字書データベースの利用

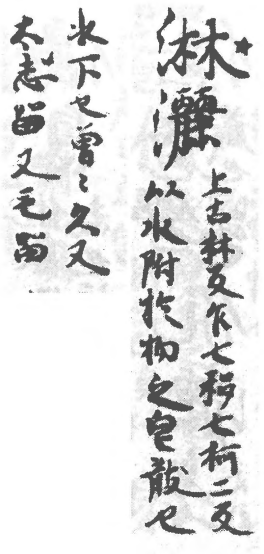
ここでは、現在計画している「古字書データベース」にどのような利用価値があるのか、その一端を紹介することとする。ただ、「古字書データベース」は、今のところ、前項で説明した『天治本新撰字鏡』しかデータベース化していないため、今は、その『天治本新撰字鏡』が『万葉集』の訓みと意味解釈に重要な働きをするという点を紹介するにとどめる。

石灑　岸の浦廻に　寄する波　辺に来寄らばか
言の繁けむ(1388)

この歌の初句に用いられている「灑」字は、『万葉集』中、孤例である。その点で、この文字の訓みと意味は、『万葉集』だけを見ていても確定できないのである。

そこで、古字書にこの文字を求めると、『天治本新撰字鏡』に「淋灑」という見出しで掲載されている(下に掲げた例参照)。それによると、注文に記載の和訓「曾々久」から訓みは、ソソクであることが分かり、さらに、意味は、義注「散也」「水下也」からして、水などが飛散するように激しく流れ下る様を表すことが分かる。よって、当該の初句の訓みは、「いはそそく」となり、上の歌の初句から３句までの解釈は、「岩を激しく洗

うように　岸の浦辺に　寄せてくる波のように」となる。



8. 古字書データベースの将来

前述のように、現在は、『天治本新撰字鏡』しかデータベース化していないが、将来は、『篆隷万象名義』『類聚名義抄』『和名類聚抄』等のデータベース化を計画している。

このように複数の古字書をデータベース化する理由は、既に触れたことではあるが、古字書が過去の字書や文献を引用することによって編集されたり、増補されてきたものであることから、まず、それら過去の字書や文献をどのように利用しているかのという字書生成の過程を知ることができるという点がある。

また、古字書が伝写される過程で、誤写・誤読といった誤りが発生していることが多々あるのだが、近い関係にある古字書相互を比校することによって、その誤りを正すことも可能となる。

さらに、古字書に引用されている過去の字書や文献の中には、部分的に残っていたり、現存しないものもあるのだが、そうした、佚文を収集することによって、現存しない資料の復元が可能にもなる。

このような利用方法を考えると、それぞれの古字書を書籍の形で並べて比較するより、本「古字書データベース」が構想している同一見出し漢字の本文を同一画面上に並列的に並べるということの方が、きわめて有用性が高いと言えるのだろう。

加えて、上記の、我が国において成立した古字書群のデータベース化と平行して、それらの古字書に多大な影響を与えた、中国の古字書についても、データベース化を図りたく考えている。

それに、奈良時代から平安時代にかけての国語・国文学系の総合データベースを目指す端緒として、先に紹介した「万葉集校本データベース」とこの「古字書データベース」とをリンクさせることにより、「万葉集校本データベース」の各画像内の文字をクリックすれば、「古字書データベース」内の画像データを表示できるようにする予定である。これによって、『万葉集』の表記文字の訓みの検討がより一層簡便に、また、厳密に行えるようになるものと確信している。

図 1

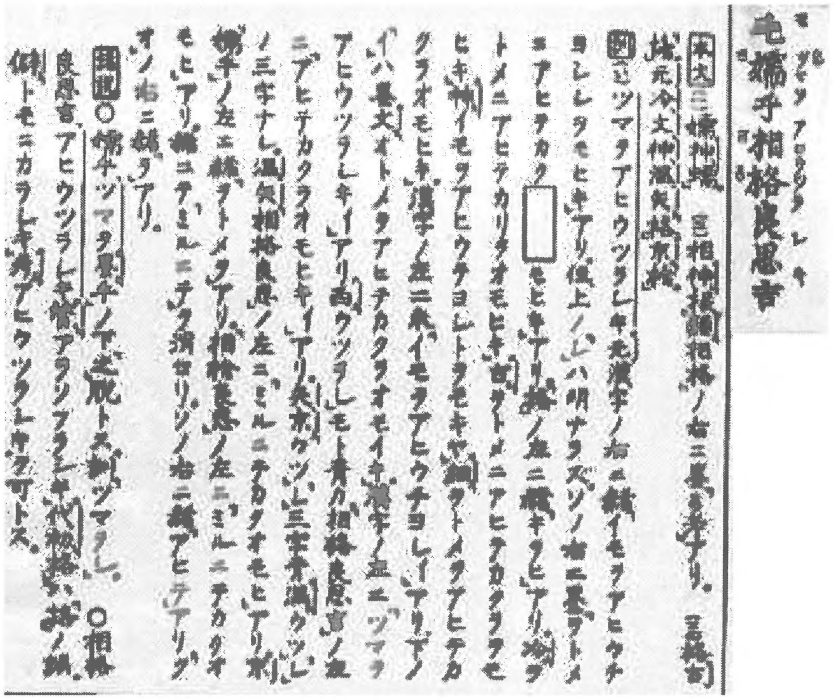

手相格良思吉

図 2

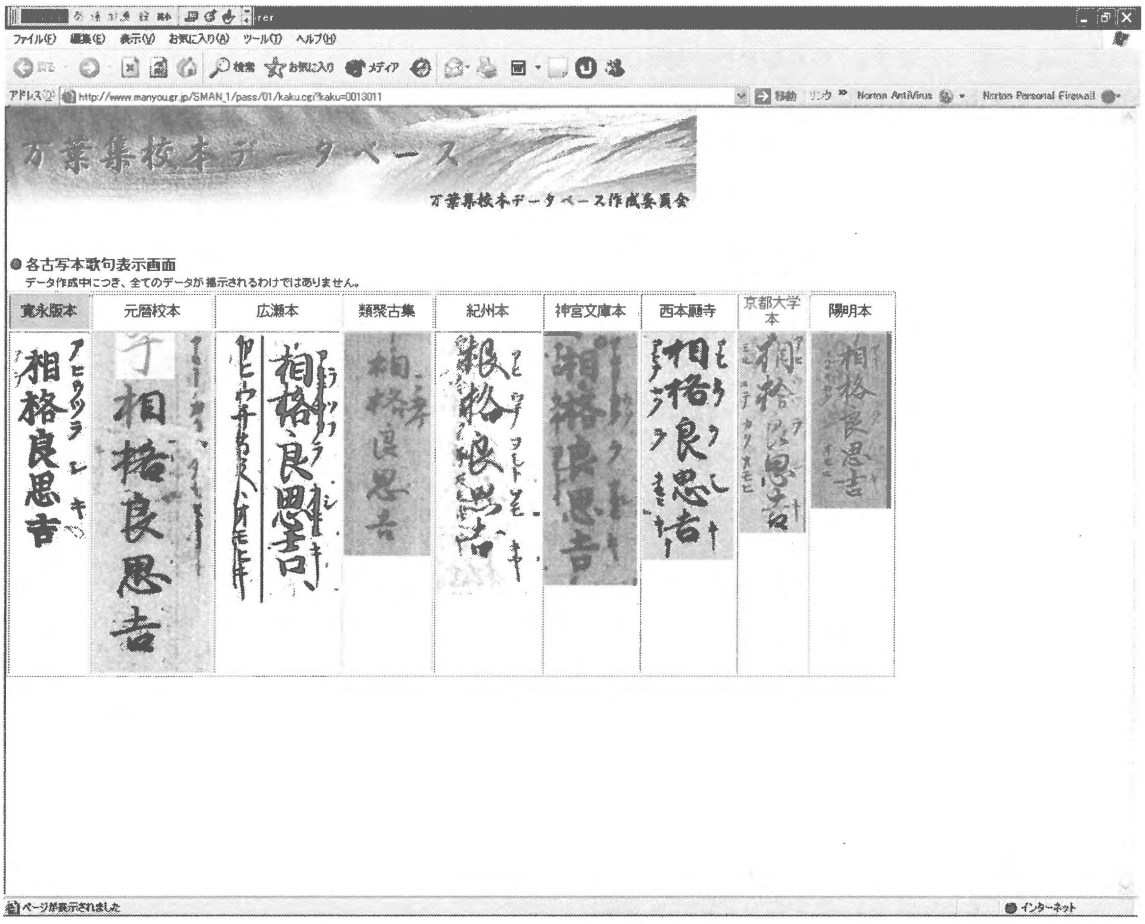

万葉集校本データベース
万葉集校本データベース作成委員会
各古写本歌句表示画面
データ作成中につき、全てのデータが掲示されるわけではありません。
寛永版本
元暦校本
広瀬本
類聚古集
紀州本
神宮文庫本
西本願寺
京都大学本
陽明本

図 3

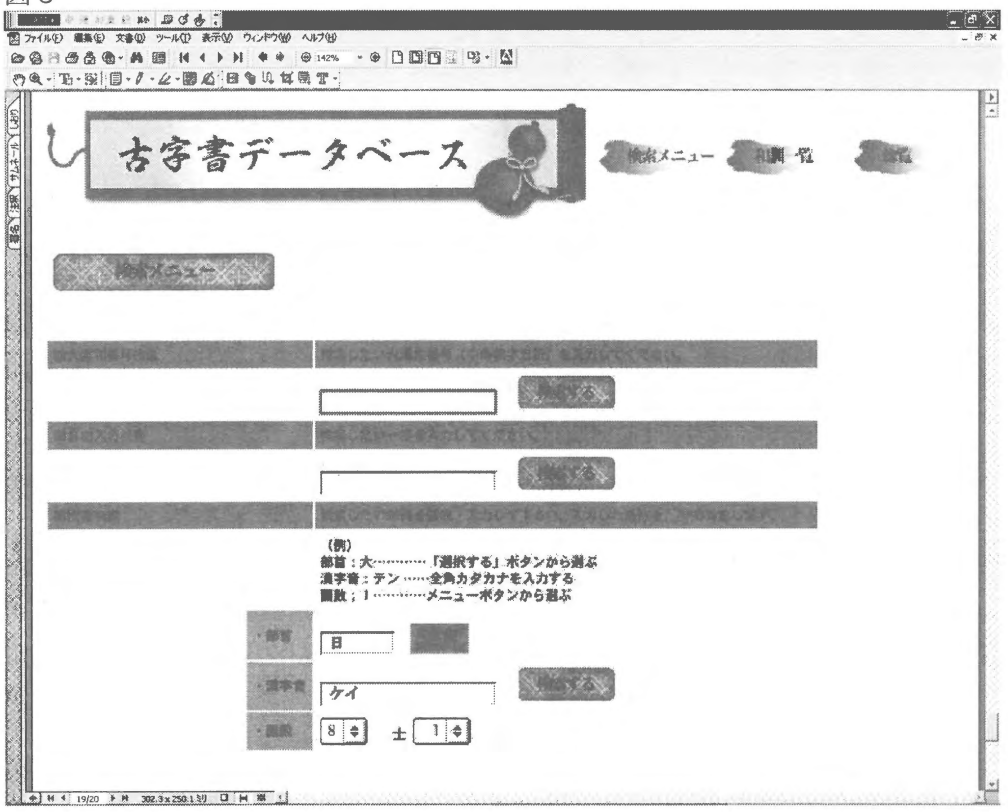
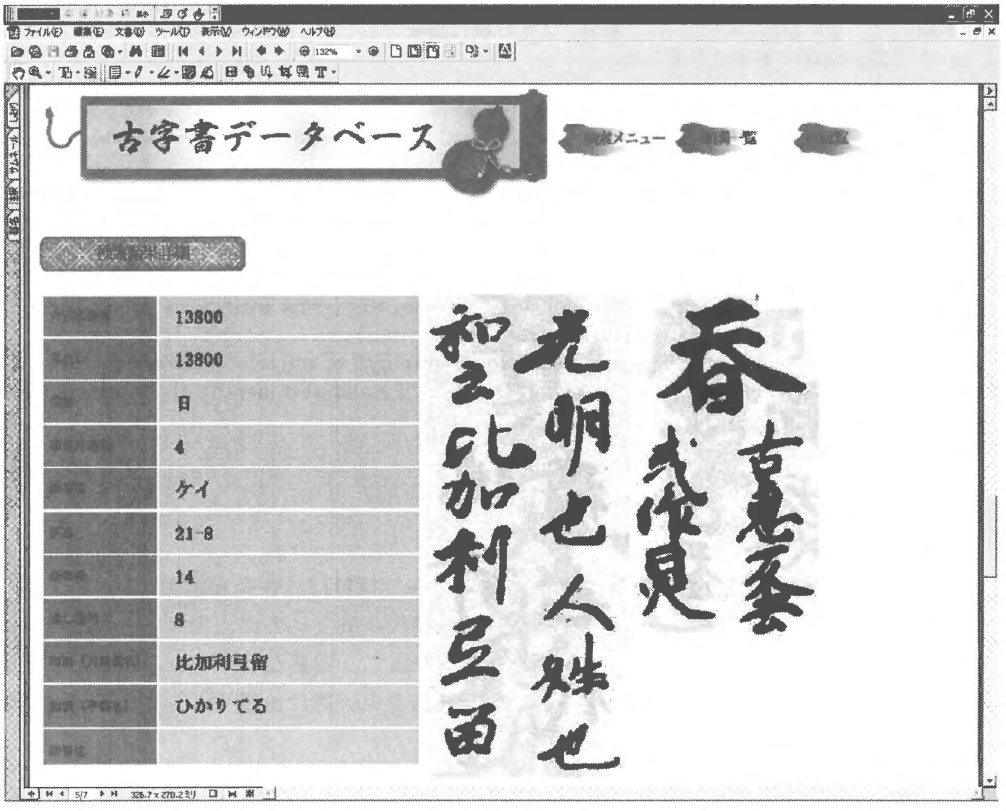

古字書データベース
検索メニュー
（例）
部首：大……「選択する」ボタンから選ぶ
漢字音：テン……全角カタカナを入力する
画数：1……メニューボタンから選ぶ
日
ケイ
8
1

図 4


古字書データベース
検索メニュー
13800
13800
日
4
ケイ
21-8
14
8
比加利豆留
ひかりてる

古今集データベースによる歌語の視覚化
Visualisation of the construction of poetic vocabulary using the database of the Kokinshū

山元啓史
Hilofumi Yamamoto
オーストラリア国立大学
The Australian National University
Japan Centre, Faculty of Asian Studies, BPB, ANU, Canberra ACT, 0200, AUSTRALIA

キーワード：歌枕 地名 和歌 モデル 視覚化 共出現パターン
Keyword: *Utamakura*, Place name, classical Japanese poetry, modelling, visualisation, co-occurrence pattern

概 要

本研究は、ひとつの和歌あるいはその現代語訳に共出現する任意の２語のパターンの頻度と重みづけを利用し、和歌語彙の構造を視覚化した。材料は古今集の和歌、1,000 首およびその現代語訳 10 種を用いた。視覚化は任意の２語それぞれの重みづけと、それら２語が各和歌に共出現した時の頻度を用いて、２語の結びつきの重要度を計算し、重要な共出現パターンからグラフで表現した。その結果、分析対象とした歌語について、和歌データだけを用いたネットワークは共出現頻度も少なく、２語の共出現関係も図からは推論しにくいが、現代語訳データもあわせて表示すると、注釈で述べられている関係を観察することができた。本研究は、主に語誌の研究として発展してきた現代語訳をもとに、共出現パターンを分析することにより、古今集の語彙の体系化を試みた。

Using mathematical analyses, this study examines the connotations of classical Japanese poetic vocabulary, especially those of *utamakura*, place names, with reference to the *Kokinwakashū* (in short, *Kokinshū*), the anthology compiled by imperial order in the tenth century (ca. 905). Using the frequency and weight of co-occurrence patterns between any two terms appearing in the translations of poems, I developed the structures of co-occurrence patterns of poetic vocabulary. I used 1,000 poems in the *Kokinshū* (OP, original poems; except for *chōka*, *sedōka*, and other different styles of poems) and the ten contemporary Japanese translations (CT). The models delineated by both OP and CT data could indicate the similar relationships between the terms with those which are described in the annotations. Since the co-occurrence patterns are indicated based on the significance of the two terms appearing in the ten CT, I could observe the strength of the relationships between the terms and the associations of terms in sub-network clusters. I could also identify the poems of the *Kokinshū* by retracing through the network models.

1 はじめに

本研究は、歌枕 (和歌に見られる地名) を視覚化によって分析する。材料は、古今和歌集（905 年、以下古今集あるいは KKS）とその現代語訳を用い、和歌あるいはその現代語訳文に共に出現する任意の２語の頻度と重みづけによって、すべての共出現パターンとその重みを計算し、重要なパターンを抽出し、視覚化した。「立田」と「吉野」に焦点をあて、視覚化を通して、同じ奈良県近隣に位置するふたつの地名の特徴を述べる。

和歌を鑑賞するためには、和歌に表記されている語や文法の理解だけでなく、それらに明記されていない含意の理解が不可欠である。含意は作品の内容、文脈、背景、受け手によって決定される意味であり、表記されている語から推論するためには、その世界の常識を用いる必要がある。和歌の世界の常識は一般的に辞典、注釈書などに詳しく記述されているが、歌語の体系的研究を行う上では、いくつか問題がある。たとえば、注釈の著者によって、その記述の多少や観点の差が見られること、どんな語と共に用いられるかという記述（共起情報）があまり見られないこと、などである[1]。本研究では、注釈を直接用いるのではなく、注釈に添えられた現代語訳を用いて、注釈と同様の説明が可能かどうかを試みる[2]。

歌枕とは広くは和歌に用いられた語のことであり、狭くは和歌に用いられた地名のことである。しかしながら、それは「単なる地名ではなく、特定の人事的観念が結合した地名」(片桐, 1970, 257) といわれている。つまり、歌枕は単にその土地の名前を示すだけでなく、景色、風物、季節などその土地に関する観念を持っているのである。本稿では「立田」「吉野」を対象語として取りあげ、それぞれが地名であるだけでなく、どのような「特定の人事的観念」が歌の中で示されているのか、について現代語訳文に共出現する２語の関係から分析する。「立田」「吉野」は、歌によく用いられる地名で、ともに奈良県に位置する土地である。古今集では、「立田」は、13 の和歌で、「吉野」は、22 の和歌で、使われており、頻度で第１、２位を占めている。

本研究は従来の注釈・語誌の研究を基盤とし、計算手続きにより、従来手作業では難しかった歌語の体系的な構造、いわゆる語彙の構造を明らかにするものである。

2　材料

古今集 (略 OP; original poems)、1,111 首のうち、長歌・旋頭歌などの雑体歌をのぞく 1,000 首、およびその現代語訳 10 種 (略 CT; contemporary translations) を利用した。現代語訳 10 種は、表 1 にある注釈書に収録されているものを利用した。底本はすべて定家本である。各著者による現代語訳の方法は、注釈書の序文や凡例において明らかに述べられていないものもあるが、だいたいにおいて逐語訳を基本に字句を補ったり、語順を任意に変更しているようである。表 2 に現代語訳の例を示す。

表 1: 古今集の注釈書 10 種の一覧

著者 (年)	ページ数	底本	現代語訳の方法
金子 (1933)	1105	定家本 (貞応二年)	逐語訳
窪田 (1960)	1449	定家本	逐語訳
松田 (1968)	1998	定家本 (梅沢本)	不詳
小沢 (1971)	544	定家本 (貞応二年)	語順・語法を変えている
竹岡 (1976)	2278	定家本 (伊達本)	逐語訳
奥村 (1978)	434	定家本 (貞応二年)	作者の作意を尊重
久曾神 (1979)	1260	定家本 (伊達本)	字句を補っている
小町谷 (1982)	407	定家本 (貞応二年)	不詳
小島・荒井 (1989)	483	定家本 (詁訓本)	不詳
片桐 (1998)	3022	定家本 (嘉禄二年自筆書写本)	逐語訳

[1] 語の体系的研究がなぜ重要であるかについては、田島 (2004, 19) に譲る。共起あるいは共出現パターン (co-occurrence patterns) といわれる。水谷 (1979) はおそらく公になっている最初の共出現パターンを利用した和歌の計量研究であろう。

[2] 注釈の利点は和歌に記述されていない語や表現が記述されていることである。欠点はその歌に直接関係のない関連情報があったり、他の歌で既に出た語の説明が省略されたり、簡略化されたりしていることである。

表 2: 現代語訳の例 KKS *298*

兼覧王	立田姫手向くる神のあればこそ秋の木の葉の幣と散るらめ
金子 (1933)	立田姫は秋の神だが、それすら暮れて行かれる折には、お手向けなさる道の神様があるからこそ、このように、御自身お染めなされた秋の木の葉が、手向けの幣のように散るのであろう。
窪田 (1960)	竜田姫は、手向けをするべき神があるので、そのつかさどる秋の木の葉が、幣のように散るのであろうよ。
松田 (1968)	竜田姫には、旅中、供え物をささげる道祖神があるからこそ、秋の紅葉した葉が、幣のように散っているのだろう。
小沢 (1971)	もはや秋の終りで龍田姫が帰り道にお着きになった。山の紅葉が色とりどりの幣となって散っているのは、姫が道の神様にそれを供えていらっしゃるのだな。
竹岡 (1976)	竜田姫が、旅にあって手向ける神があるからこそ、それであのように秋の木の葉が、幣として散っているのであろう。
奥村 (1978)	秋も終りに近づき、秋の女神の龍田姫がお帰りになる。姫が道中の安全を祈って手向けをなさる神があるからこそだろう、紅葉が幣のように散っている。
久曾神 (1979)	秋も終りに近づき竜田姫がお帰りになる際に、旅の安全を祈って手向けられる神があるからこそ、秋の紅葉が幣となって散っているのであろう。
小町谷 (1982)	秋の末近くなって帰り道についた龍田姫が、道中の無事を願って手向けをする神があるからこそ、秋の木の葉が幣となって散っているのだろう。
小島・新井 (1989)	竜田姫が手向けをする神さまがおられるからこそ、秋の木の葉のもみじが手向けものとして散るのでしょうね。
片桐 (1998)	秋をつかさどる龍田姫が旅立ちにあたって手向けをする神があるから、あのように、秋の木の葉が幣を撒くように散っているのだろう。

3 方法

ある和歌あるいはその現代語訳に共出現する任意の 2 語のパターンの頻度と重みづけを利用し、和歌語彙の共出現パターンを視覚化した。和歌とその現代語訳はすべて β 単位で分割した[3]。ただし、地名に限っては、「立田山」を「立田」と「山」のように短い単位に分けた。さらに、分類語彙表 (中野, 1994) に準じたコードをそれぞれの語につけ、そのコードによって、和歌中の語あるいは現代語訳中の語が表記が異なっていても対応できるようにした[4]。視覚化には共出現パターンの重要度を計算し、重要な共出現パターンから順にネットワークで表現した。

共出現パターンの重要度 (*cw*; co-occurrence weight) は、(1) のように定義できる。

$$cw(t_1, t_2, d) = (1 + \log\ ctf(t_1, t_2, d)) \cdot \sqrt{idf(t_1) \cdot idf(t_2)} \tag{1}$$

ただし、前半の $1 + \log\ ctf(t_1, t_2, d)$ は、ある条件で抽出された資料の数 d で、t_1 と t_2 の 2 つの用語が共に出現した時の資料の数 (co-occurrent term frequency) を対数変換し、1 を加えたものである。また、後半の $\sqrt{idf(t_1) \cdot idf(t_2)}$ は、t_1 と t_2 の 2 つの用語の idf 値の幾何平均である[5]。idf (inverse document frequency) は式 (2) で定義できる (Manning and Schütze, 1999, 543)。

$$idf(t) = \log \frac{N}{df(t)} \tag{2}$$

ただし、N はすべての資料の数、ここではすべての和歌あるいは現代語訳文の数である。$df(t)$ (document frequency) は、語 t の出現する資料の数、ここでは当該の和歌あるいは現代語訳文の数である。$idf(t)$ は、

[3] 語の調査単位については、田中 (1971)、宮島 (1994) を参照されたい。宮島 (1994, 113) は、日本語の語彙調査で一番困ることとして、『「単語」という単位が確立していないこと』と述べている。

[4] 分類語彙表に準じた語にコードを付け、そのコードで作品中の意味成分を調査する提案や試みは、田島 (1995) や西端 (1996) などに見られる。

[5] 式 (1) は、$w = tfidf$ (Rocchio, 1971; Spärck Jones, 1972) を任意 2 語の共出現の重みを計算するために拡張したものである。頻度、tf の代りに共出現頻度、ctf を、idf の代りに 2 つの idf の幾何平均を用いた。

N と $df(t)$ の比を対数変換したものである (Spärck Jones, 1972; Robertson, 2004)。

視覚化の手順としては、対象語「立田」「吉野」を含む和歌をまず検索し、その和歌あるいはその和歌に対応する現代語訳文に共出現する任意の2語の頻度と2語それぞれの *idf* 値を使って *cw* を計算した。このように計算された2語の *cw* 値の高いパターンから順にグラフで描いた。描画には、Graphviz (Ellson et al., 2005) を用いた。

4 結果

本稿では、現代語訳を用いた分析について詳しく述べる。地名、人名のような固有名詞は、さまざまな文章に出てくるのではなく、特定のある限られたの文章に出てくる性質をもつため、結果的に *idf* 値の高い用語となる。「立田 (*idf*=8.52)」「吉野 (*idf*=6.08)」のように *idf* の高い語 (5–10) を調査の対象としてネットワークを描いた場合、対象語を中心に放射線状の図形 (spoke effect) が出ることがわかっている (Yamamoto, 2005)。しかし、対象語との共出現パターンを除くと放射線状の図形は取り除かれ、放射線状の図形に蔽われていた構造が見える[6]。本稿では、対象語との共出現パターンを削除した図、つまり、「立田」「吉野」との共出現パターンを除いた図を結果として示す。それぞれのネットワークモデルは語を示すノードとそのつながりを示す枝からなり、ノードの大きさは相対的な頻度を、枝に添えられた数字は2つのノードの共出現頻度を示す。

4.1 立田

図1は、「立田」のネットワークモデルである。「立田」のモデルは、大きく2つに分かれ、秋をイメージする「紅葉」を中心とするモデルと、「沖つ白波」の歌 KKS *994* を中心とするモデルである。

神様の山である立田山に向かって、色あざやかな紅葉が川一面に流れて行く情景を幣 (神にささげる布織物の贈り物) に見立てた歌が、ネットワークから読みとれる。このように紅葉の流れゆくさまを幣に見立てた歌は、KKS *298* (表2) や KKS *300* に見られる。

KKS *300* 神奈備の山を過ぎ行く秋なれば立田川にぞ幣は手向くる　—清原深養父

ネットワークの中にさらにネットワークが見られる。「神代」を中心とするサブネットワークには、KKS *294* の歌が見える。

KKS *294* 千早振る神代も聞かず立田川唐紅に水括るとは　—在原業平

上記から「立田」は秋を紅葉で彩る山であることがわかるが、そこには「神奈備」「神代」という語が含まれていることからもわかるように、「神」が存在していることが認められる。一方、「沖つ白波」のモデルは KKS *994* が基となっている。KKS *994* は伊勢物語 (23段) や大和物語 (149段) にも出てくる有名な歌である。

KKS *994* 風吹けば沖つ白波立田山夜半にや君が一人越ゆらむ　—よみ人知らず

「立田山」は想う人が越える山である。「白波」は「波が立つ」という意味関係に由来する「立田」の縁語である。この場合の「立田」は紅葉に代表される秋の「立田」ではなく、河内国から大和国への重要な交通路としての「立田」である (片桐, 1983, 251)。このことはモデルでは、図1のように2つのネットワークに分かれて描画されている。

[6] ゆえにモデル中には「立田」「吉野」の地名が見られないことに注意。対象語を含む歌が抽出された後、その歌に対応する現代語訳を用いて共出現パターンを作成しているので、すべての語は対象語と高い *cw* 値で結びついている。

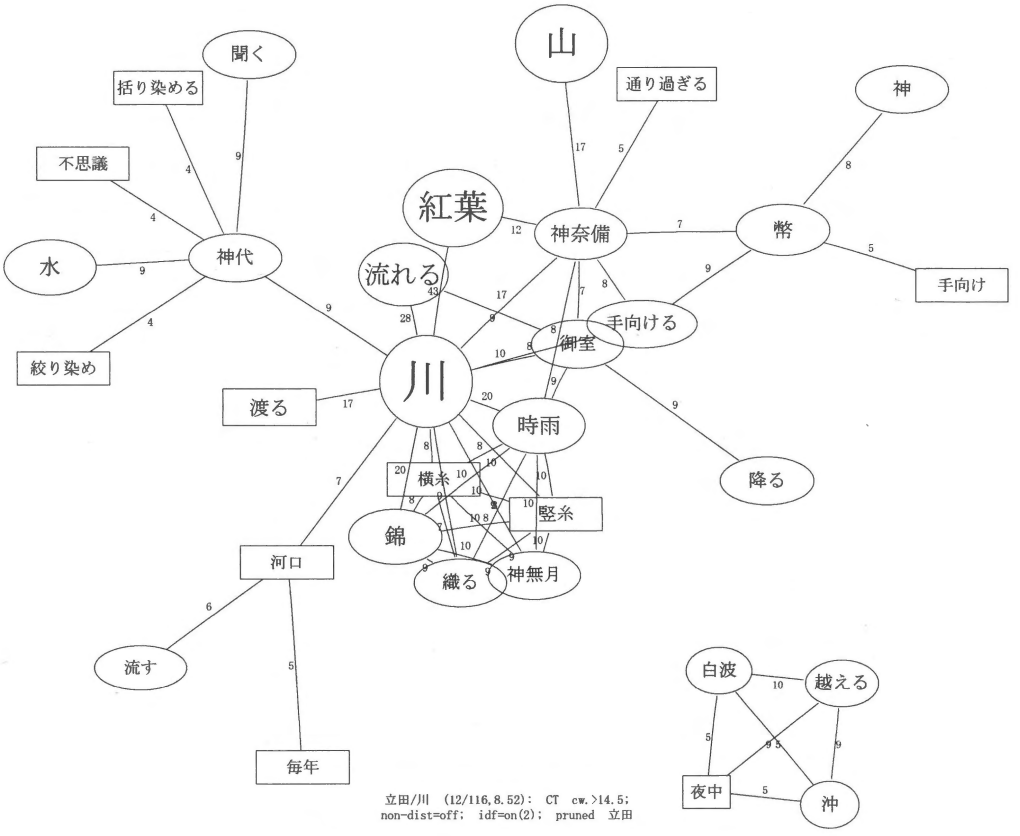


図 1: ネットワークモデル「立田」(川/白波): 立田のネットワークは、大きく 2 つの部分に分かれ、その「川」の部分と「白波 (右下)」の部分を描画。「立田」は OP では 12 回、CT では 116 回出現する。「立田」の *idf* 値は 8.52 。14.5 *cw* 以上を対象に描画。「川」は、28 ノード (OP 18 ノード、CT 10 ノード)、49 の枝からなっている。「白波」は、4 ノード (OP 3 ノード、CT 1 ノード)、6 の枝からなっている。各ノードの大きさは相対的頻度を示している。楕円中の語は OP にも出現する語で、矩形中の語は CT のみに出現する語。

4.2 吉野

図 2 は、吉野川を中心とするモデルで、図 3 は、吉野山を中心とするモデルである。

「吉野川」のモデルは、「烈しい」「流れ」「妹山」「男女」という語からもわかるように、男女間の心理の激しい変化が、「吉野川」の流れとして表現されている。

KKS *471* 吉野川岩波高く行く水の早くぞ人を思ひ初めて —紀貫之

「妹山」は「妹背山」ともいい、「妹背」とは「愛し合う女と男。夫婦」(広辞苑) の意味である。「吉野山」のモデルには「妹山」を中心とするサブネットワークが見える。「妹山」に流れ落ちる川とその谷の狭間を「男女」「夫婦」の「仲」「間」に見立てている歌 KKS *828* がモデルから読みとれる。

KKS *828* 流れては妹背の山の中に落つる吉野川の縦しや世の中[7] —よみ人知らず

一方「吉野山」のモデルは、「雪」「白雪」「山奥」「有明の月／里」「隠れ所」などの用語で構成されている。

[7] 図 3 には「たとえ」が見られる。これは「縦しや」の現代語訳であろう。「縦しや」は「吉野」の掛詞である。

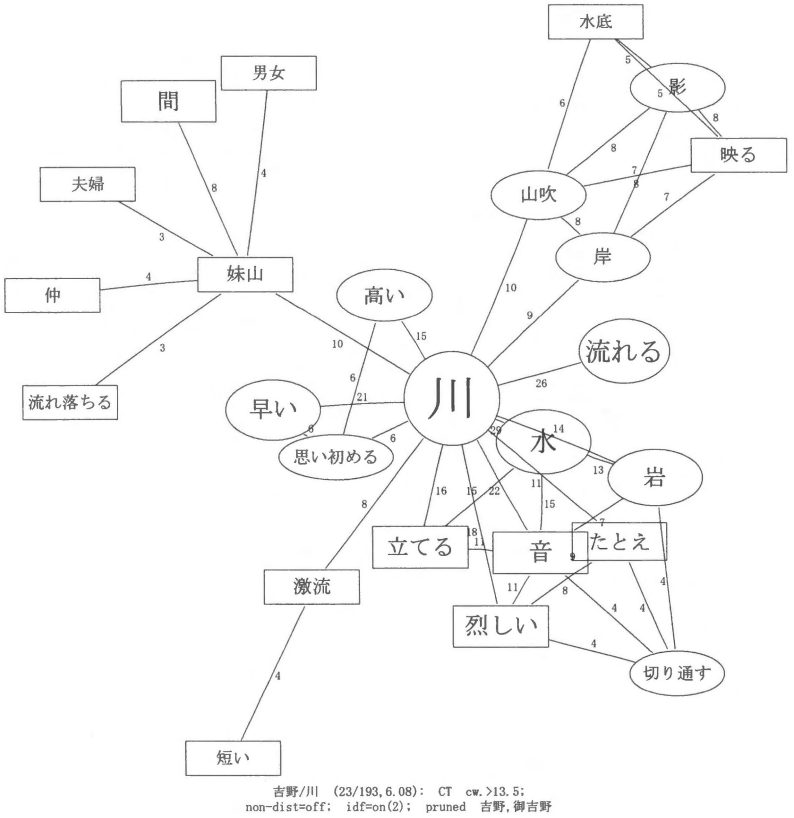


図 2: ネットワークモデル「吉野川」:「吉野」は OP では 23 回、CT では 193 回出現する。「吉野」の *idf* 値は 6.08 。13.5 *cw* 以上を対象に描画。このモデルは 25 ノード（OP 11 ノード、CT 14 ノード）43 の枝からなっている。各ノードの大きさは相対的頻度を示している。楕円中の語は OP にも出現する語で、矩形中の語は CT のみに出現する語。

KKS *332* 朝朗け有明の月と見るまでに吉野の里に降れる白雪　—坂上是則

「隠れ所」を中心とするサブネットワークが見られ、「吉野山」が世の中（人間社会）を嫌った人の隠遁の地であることがわかる。このサブネットワークを構成する語はほとんどが現代語訳文にのみ出現する語であることから、「隠れ所」「隠遁」という意味は和歌そのものからは、

KKS *950* 御吉野の山のあなたに宿もがな世の憂き時の隠れ家にせむ　—よみ人知らず

のように「隠れ家」という語から読みとれる。片桐 (1983, 434) によると、奈良時代までは実際に吉野山に赴くことがあまりなかったがゆえに、山岳信仰の地・隠遁の地としての吉野山であったが、そのような神秘的なイメージは雪をいただく山々の姿とマッチして、吉野山といえば雪がよまれるというようになった、ということである。モデル中の「雪」「山奥」「隠れ所」の語が、結果的にこのようにネットワークで表現されている理由は、上記の説明でわかる。

ところで「吉野の桜」は有名ではあるが、「桜」はこのモデルには見られない。片桐 (1983, 436) によると、吉野山は、平安時代中期までは、雪の名所であるが、桜の名所としての吉野山のイメージはそれほど強くはなく、吉野山と桜の関係が決定的なものになるのは、新古今集を代表とする数々の歌を詠んだ西行の時代、といわれる。古今集のデータにも「吉野」と「桜」を詠んだ歌があるにもかかわらず、このモデ

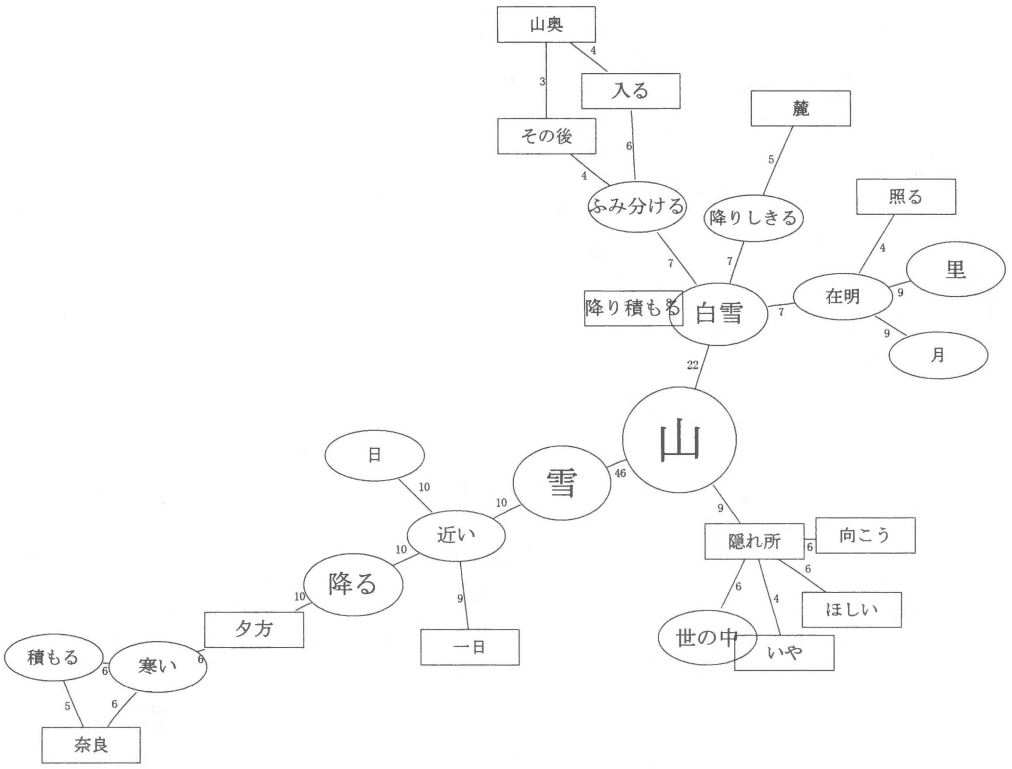


図 3: ネットワークモデル「吉野山」:「吉野」は OP では 23 回、CT では 193 回出現する。「吉野」の *idf* 値は 6.08 。13.5 *cw* 以上を対象に描画。このモデルは 27 ノード（OP 14 ノード、CT 13 ノード）28 の枝からなっている。各ノードの大きさは相対的頻度を示している。楕円中の語は OP にも出現する語で、矩形中の語は CT のみに出現する語。

ルに「桜」が含まれていないのは、共出現パターンの重みづけ (*cw*) がうまく機能しているものと思われる。しかし、それをより確実に結論づけるなら、新古今和歌集（1201 年）を分析し、そこに「桜」が「吉野山」のモデルに含まれるかどうかを検証するとよいだろう。

5 考察

「立田」と「吉野」を対象語として分析を行った。ここまで、それぞれ歌語としての意味・用法をモデルから推察し、注釈と比較しそのモデルの妥当性を検討してきた。モデルが正しいという前提で、改めて「立田」と「吉野」を比べてみると「立田」は神の山、紅葉／秋の山であるのに対して、「吉野」は男女の山、人間心理の激しさを象徴する川、世の中に疲れた人が隠遁する山、と見てとれる。つまり、「立田」の神とは対照的に「吉野」は世の中（人間世界）の歌枕として目に映る。同じ奈良県に近隣する 2 つの歌枕であっても、モデルから考察すると非常に異なる観念が含まれていることがわかる。

本稿では詳細に述べなかったが、分析対象とした歌語について、和歌データだけによっても視覚化してみた。和歌データだけを用いたネットワークは共出現頻度も少なく、2 語の共出現パターンも図からは推論しにくいものであった。しかし、現代語訳データもあわせて視覚化すると、注釈で述べられている関係

がモデルから読みとることができた。共出現パターンは重要度によって表示されるため、１０種の現代語訳に共通して示されているもの、示されていないもの、結びつきの強弱も観察できた。ネットワークの中にはさらにサブネットワークが見られ、そのサブネットワークを構成する語群が、どの和歌に起因するものなのかも特定できた。また、同じ対象語であっても意味・用法の異なる場合には、別の群として描画されていることがわかった。

6 おわりに

本研究では、主に語誌の研究として発展してきた現代語訳をもとに、古今集の語彙の体系化を試みた。共出現パターンを作成し、共出現パターンの重要度 (cw) にしたがって、パターンを描画し、視覚化を試みた。視覚化されたモデルを通して、語の意味・用法の違いを確認することができた。

参考文献

Ellson, J., Gansner, E., Koren, Y., Koutsofios, E., Mocenigo, J., North, S., Woodhull, G., Dobkin, D., Alexiev, V., Lilly, B., Scheerder, J., G., D. R., and Low, G. Graphviz — graph visualization software. `http://www.graphviz.org/` 2005; The page cited on 4th Feb. 2005.

金子 元臣 1933. **古今和歌集評釈: 昭和新版** 東京: 明治書院 .

片桐 洋一 1970. 歌枕の成立 —古今集表現研究の一部として— [**古今和歌集** 小町谷照彦, 日本文学研究資料刊行会 編] 第 1 版 pp. 256–266 東京: 有精堂 , [1976]「国語と国文学」(1970) 昭和 45 年 4 月号.

片桐 洋一 1983. **歌枕歌ことば辞典** 角川小辞典, 第 35 巻 東京: 角川書店 .

片桐 洋一 1998. **古今和歌集全評釈 上・中・下** 全 3 巻 東京: 講談社 .

小島 憲之, 荒井 栄蔵 1989. **古今和歌集** 新日本古典文学体系, 第 5 巻 東京: 岩波書店 .

小町谷 照彦 1982. **現代語訳対照 古今和歌集** 東京: 旺文社 .

窪田 空穂 1960. **古今和歌集評釈（上中下）** 東京: 東京堂出版 .

久曾神 昇 1979. **古今和歌集 全訳注** *(1)-(5)* 講談社学術文庫 .

Manning, C. D. and Schütze, H. 1999. *Foundation of statistical natural language processing* Cambridge, Massachusetts: The MIT press .

松田 武夫 1968. **新釈古今和歌集（上下）** 東京: 風間書房 .

宮島 達夫 1994. **語彙論研究** 東京: 麦書房 .

水谷 静夫 1979. 用語による梅・桜の歌の弁別. **計量国語学** *12*, 1–13.

中野 洋 1994. **分類語彙表／フロッピー版** 国立国語研究所言語処理データ集, 第 5 巻東京: 大日本図書 .

西端 幸雄 1996. 「歌物語」３作品の使用語彙の比較 [**「歌物語」語彙の数量的分析と研究**] 第 1 版 pp. 3–18, 文部省科学研究費：重点領域研究 「人文科学とコンピュータ」研究成果報告書.

奥村 恒哉 1978. **古今和歌集** 東京: 新潮社 .

小沢 正夫 1971. **古今和歌集** 第 13 版 日本古典文学全集, 第 7 巻 東京: 小学館 .

Robertson, S. 2004. Understanding inverse document frequency: on theoretical arguments for idf. *Journal of Documentation 60*, 503–520.

Rocchio, J. J. 1971. The SMART Retrieval System: Experiments in Automatic Document Processing in *Relevance feedback in information retrieval* Salton, T. G. (ed.) 1 edn. pp. 313–323 Englewood Cliff, NJ: Prentice-Hall .

Spärck Jones, K. 1972. A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation 28*, 11–21.

田島 毓堂 1995. 語彙と単語. **日本語学** *14*, 4–11.

田島 毓堂 2004. 比較語彙論 [**語彙研究の課題** 田島 毓堂 編] 第 1 版 pp. 1–20 大阪: 和泉書院 .

竹岡 正夫 1976. **古今和歌集全評釈（上下）** 東京: 右文書院 .

田中 章夫 1971. 新聞語彙調査の同音語と同形語 [**電子計算機による国語研究** *III*] 第 1 版 国立国語研究所報告, 第 39 巻 pp. 121–145 東京: 秀英出版 .

Yamamoto, H. 2005. A mathematical analysis of the connotations of Classical Japanese poetic vocabulary. Ph.D dissertation, The graduate school of the Australian National University.

PDFファイルを用いた歴史的資料のデジタル化とデータベースの試作

An Approach on Interactive Database using Annotation of PDF Documents

竹内 さおり　　白川 哲郎
TAKEUCHI Saori　　SHIRAKAWA Tetsuroh

大阪樟蔭女子大学　学芸学部，東大阪市菱屋西 4-2-26
Osaka Shoin Women's University, 4-2-26, Hishiyanishi, Higashiosaka, Osaka
takeuchi.saori@osaka-shoin.ac.jp

あらまし：本稿では，樟蔭学園デジタルアーカイブズ構築の一部として，昭和戦前期の学園広報誌『樟蔭學報』のデジタル化とデータベース試作の取り組みについて報告する．『樟蔭學報』は，1936年（昭和11年）8月に創刊され，創刊号から第三巻第三号までの合計19冊が現存する．本研究では，資料をデジタル化して保存・整理することとこれらを授業で活用することの二つを目的とする．授業で活用するための工夫として，データベース本来の検索や検索結果を閲覧する機能に加え，内容の充実に利用者が参加できるような仕組みを考えた．具体的には，PDFファイルの注釈を用いて，既存のキーワードに利用者の保持する情報やコメントを付加することを可能にし，インタラクティブな仕様を実現した．

Summary: In this paper, I will report on digitalization and the making of database for the magazine 'Shoin Gakuhou', for the first term of the war of the Showa era. This project is a part of constructing the educational institution digital archives of Shoin Gakuen. 'Shoin Gakuhou' started in August, 1936, and presently, there are nineteen books consisting of volume 1 to volume 3, the third issue. There are two aims for digitalizing these magazines to preserve and arrange them, and to use them in the classroom. In addition to the user-friendly interface of accessing information, the program of the database is designed for student users to add new information and comments easily to the contents. They are expected to type in their own information under existing keywords of the data. This interactive specification was achieved by using the annotation function of PDF documents.

キーワード：デジタルアーカイブズ，PDFファイル，注釈，データベース，双方向性
Keywords: digital archives, PDF files, annotation, data base, interactive

1.はじめに

1917年（大正7）に樟蔭高等女學校として開学した樟蔭学園は，2007年に創立90周年を迎える．学園草創期の資料は，近現代の女性史及び学校教育史を考察するうえで貴重な資料であるとともに，大阪樟蔭女子大学大学史を理解するための重要な教材である．

平成15年度より継続的に進めてきた樟蔭学園草創期資料のデジタル化とデータベース化に関する研究は，作業手順をほぼ確立することができ，進捗状況は順調である．着手から3年目にあたる平成17年度は，保存と整理の作業そのものよりもデジタル化した資料やデータベースを教材として活用することにポイントをおき，使い易さや楽しさをとり入れる工夫をした．1936年（昭和11）創刊の学園広報誌である『樟蔭學報』を対象とし，デ

ジタルコンテンツの利点を生かし，知識の共有も容易に行える教材の作成に取組んだ．

本稿では，PDF(Portable Document Format)[1]ファイルの注釈機能を用いたインタラクティブなデータベースの試作について報告する．以下，2節で研究の背景と目的を述べ，3節で提案手法を詳細に説明する，4節で『樟蔭学報』の内容にふれ，5節で基本作業手順と具体例を示す．6節では評価実験について述べ，7節で関連研究を挙げる．最後に，考察と今後の予定を述べる．

図1：『樟蔭學報』創刊号の表紙

2.研究の背景と目的

本研究は，樟蔭学園の歴史的資料を授業で活用することで，学生の教育効果を期待するとともに母校への愛着を増加させたいという筆者等の考えからスタートした．平成16年度には，データベースソフト(Microsoft Access)を用いた『設立二関スル書類』のカード型データベースを試作し，紙資料のデジタル化とデータベース化を行った[2]．しかしながら，テスト運用の段階で以下のような指摘を受けた．

1. 1枚のカードで資料1ページ分の画像と各項目のテキストデータを見ることができるが，資料全体のイメージは把握しにくい．
2. 文字や画像が小さすぎて見えにくいので，教材としては相応しくない．
3. 画像をそのまま複製して転用される可能性があるので，データでの配布やWebで公開することは好ましくない．

これらの指摘を改善するために，データベースソフトに依存することや画像データとテキストデータを並べて表示する型に固執せず，コンピュータの画面上でパラパラとページを繰るように閲覧できる仕組みの実現と検索の容易さ，データでの配布やWebで公開する際に不都合が生じないように配慮することを目的に，新たな手法を検討した．

3.提案手法

本研究では，PDFファイルの特徴に着目し，データベース構築への利用を検討した．

PDFファイルは，プラットホームや使用環境に影響を受けず，利便性の高い文書フォーマットであり，文書の配布にPDFファイルを利用する頻度も高くなっているので，利用価値が高く有用なファイル形式とあると考えられる．また，PDFファイルは，文書ファイルからだけでなく，画像ファイルからの変換も可能であり，変換後のファイルに遜色がないばかりかファイルサイズが小さくなるためネットワーク上で扱い易い．検索に関しては，既存の全文検索機能が利用できるが，文書ファイルから変換した場合に限られるため，OCR装置やデジタルカメラで入力したデータは対象外とされていた．

今回のように，画像ファイルから生成したPDFファイルの文字列を検索する方法として，注釈機能を利用した手法を提案する．つまり，キーワードを注釈として追加し，注釈を手がかりにPDFファイルの中身を検索しようという発想である．

注釈を利用するもう一つの理由は，データベースの作成者だけでなく，利用者も注釈を追加することに参加できるからである．

具体的には，以下のような方法である．

- PDFファイルの作成者がつけたキーワード

を用いて検索した利用者が新たなキーワードを注釈として追加する．

- 閲覧した利用者が重要と思われる情報や関連した資料についての注釈を追加する．

利用者が注釈を追加することでキーワードを増加することができる．さらに，その人が何を重要と考えているかも明らかになる．

つまり，利用者が積極的に参加できる環境を整えることができれば，知識を共有するという観点からも効果が期待できる．例えば，図書館で借りた本を教材として利用していたときに，重要だと思ったことや他に調べた資料のことを直に書き込むことはできない．他の利用者が何を理解して何を重要と考えたかという見えない情報や参照した関連資料の記録は明らかにならない．

しかしながら，このような情報が役立つ機会は多いように思われる．紙の上では実現できないならば，PDF ファイル上で実現したいと考えた．

提案手法では，Amazon のお薦め情報の提示[3]に類似した状況を機械的なアルゴリズムからではなく，人と人のつながりから実現する．

教員と学生及び学生同士のコミュニケーションツールの役目を果たす可能性も高く，教育効果も期待できる．教材を介し知識を共有できる場を提供するとともに，注釈の蓄積は原資料に付加価値を付けることもでき，有用性も増加する．

一方，画像の複製や転用に関する件は，PDF ファイルを作成する際にセキュリティ設定を変更することにより防止が可能であることを確認した．

利用者の編集や印刷の操作に制限を加えることが可能になるので，画像データであっても保護することができる．データでの配布や Web で公開する場合に不安とされた点が解消でき，ネットワークを介した遠隔授業等での利用も予定している．

従来のデータベースのように，蓄積したデータを利用するだけでなく，動的な仕様[4]を実現することも可能かもしれない．

4.『樟蔭学報』について

『樟蔭學報』は 1936 年（昭和 11）に創刊された樟蔭学園広報誌であり，創刊号から第三巻第三号までの合計 19 冊が現存する．創刊号掲載の「樟蔭學報発刊の辭」には，在学生二千名，樟蔭高等女學校本科・専攻科・高等科の卒業生同窓会員（緑蔭会員）三千数百名，樟蔭女子専門學校卒業生同窓会員（緑翠会員）一千数百名，それぞれの保護者を含めた合計一萬数千名の関係者に向けて，親交を温めるために発行されたと記載されている．内容については，樟蔭高等女學校及び樟蔭女子専門學校の年中行事の紹介，経過報告，緑蔭会及び緑翠会の会記や会員の消息，生徒の詩歌や文章の発表，緑蔭会員及び緑翠会員の作品発表，保護者の声と学校側の反響，論説，文苑等となっている．行事や校内で撮影された写真も多く掲載されていて，服装や行事の様子などもよくわかる．当時の様子がうかがえる広告などもたくさん含まれ，非常に興味深い資料であると思われる．なお，図 1 に示した創刊号の表紙は，カラー印刷である．

5.基本作業の手順

図 2 に基本作業の流れを示す．基本作業は，二つの部分から成る．第 1 段階では，『樟蔭學報』の各ページをデジタルカメラで撮影して画像を保存し，学園資料調査カード（紙媒体）に必要な項目を記入する．画像データとテキストデータを 1 つずつ作成していく地道な作業である．第 2 段階では，画像データを修正し，PDF ファイルに変換してから，画像の撮影時に抽出したキーワードを注釈として付加する．今回の作業では，各号ごとに PDF ファイルを作成した．第 1 段階は白川ゼミ卒業生の松田憲子さんに担当を依頼し，第 2 段階は竹内が担当した．本稿執筆の段階で，創刊号からの第一巻五号までの 5 冊分が完成し，現在も作業を継続中である．以下に，詳細な説明と具体例を示す．

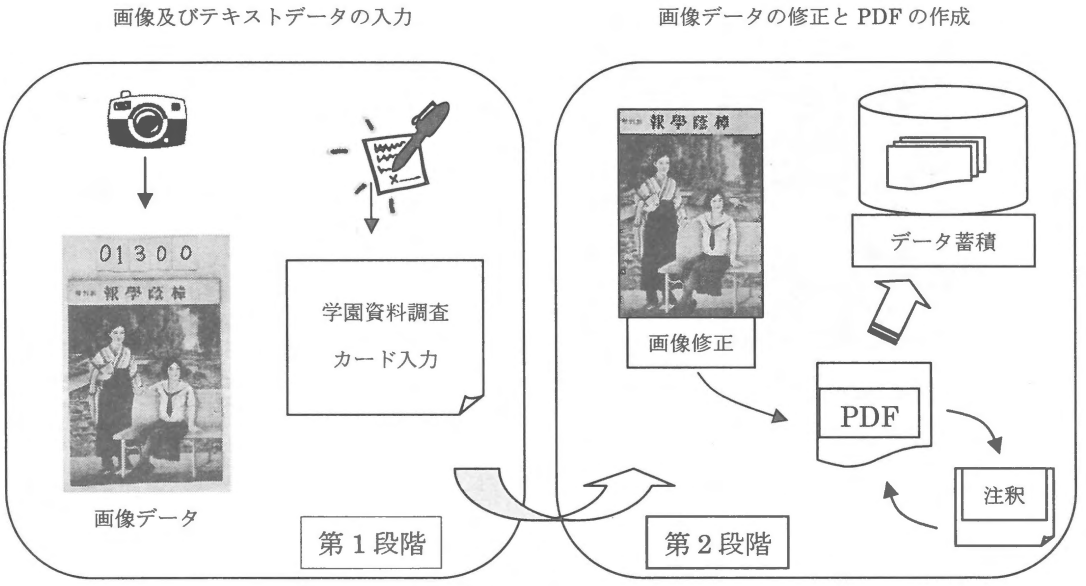


図 2：基本作業の流れ

5.1 画像の撮影と学園資料調査カードの記入

画像は，1 ページずつ順番にデジタルカメラで撮影した．デジタルカメラの誤操作によるデータの消失を避けるために CD 内蔵のデジタルカメラ[5]を使用した．学園資料調査カードの様式は，前回と全く同じ A4 サイズの用紙に「カード No.，調査日，調査者，名称，資料名，形状，材質，サイズ，年記，内容，キーワード，写真，特記事項，学園整理台帳における記載の有無，記載者」の 15 項目で作成したものを使用した．学園資料カードの記入（手書き）と Microsoft Excel のワークシート入力の両方を行うことに，冗長的な感じは否めないが，作業方法の変更は行わない．第 1 段階で特筆することは，撮影後にハードディスクへ保存する方法でデータの二重化につとめていたにも関わらず，ハードディスクのクラッシュによりデータの一部を消失してしまった点である．解決策としては，毎日決まった時間に別のハードディスクにバックアップを作成するタイプ[6]に更新した．データの消失は，保存データが増加すればするほど，大きな痛手となる．今後も，データの管理やバックアップの方法については，その都度，信頼できるハードウェアを採用していくこと等で対応したい．

5.2 画像の修正と PDF ファイルの作成

図 3 に PDF ファイルの例を示す．デジタルカメラで撮影した jpeg 形式のファイルは，コントラストと色調の微調整を行い，画質の良い状態で保存し直した後，1 冊分ずつまとめて PDF ファイルに変換する．まとめて変換することで，1 枚ずつバラバラに撮影した画像をひとまとまりとして扱うことができるので，1 冊の本のようにパラパラと閲覧することが可能となる．画像の修正には Adobe Photo Shop を PDF ファイルの作成には Adobe Acrobat を使用した．PDF ファイルに変換した後に，キーワードとして抽出した語を注釈（ノート）の追加を行う．

5.3 注釈を用いた検索と拡張性

図 4 に注釈を用いた検索の例を示す．「旅行」をキーワードに創刊号を検索したところ，3 件の結果が得られた．該当する「東京地方修学旅行」のページと注釈が表示され，注釈のマッチした部分が太字になる．ここでの検索は，注釈として付加された語だけを対象にしているので，検索文字列と全く同じもしくは一部が，注釈と同じ文字列の検索にすぎない．ここでは，関心のある箇所を閲覧

図3：PDF版『樟蔭學報』の例

し始める入口の意味合いで検索結果を捉えたい．前述したように検索で得られた結果が最終目標ではなく，利用者が注釈にキーワードを追加することやコメントを書き込むことによって，知識を共有し再利用するための手がかりである．

例えば，演習形式の少人数のクラスで使用する場合，ネットワーク上のPDF版『樟蔭學報』に同時に複数の学生が各自のコンピュータからアクセスして，閲覧しながら授業に参加する．授業時間内にキーワードや個々のコメントを注釈として追加することができなかったとしても，授業時間以外にも同じように利用できるので，時間的な制約も受けない．教授者はコメントに対するアドバイスを与え，議論の活発化をはかる．

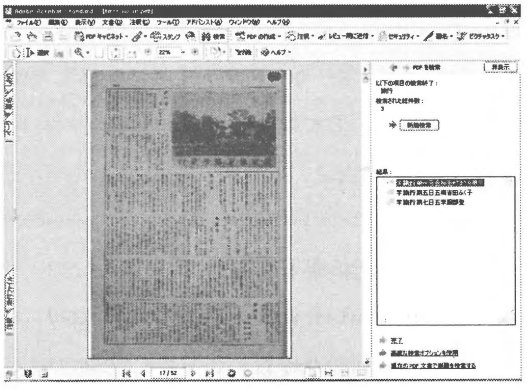


図4：注釈を用いた検索の例

5.4 作業の進捗状況

創刊号から第三巻三号までの合計19冊の撮影と学園資料カードの記入が完了し，テキストデータの入力と最近になって見つかった緑蔭会員[7]及び緑翠会員[8]向け冊子の特集記事を撮影中である．画像の修正が終了したものから順にPDFファイルへの変換作業を済ませているので，テキストデータが揃えば注釈付けは完了できる．

6.評価実験

本研究における成果を平成18年度に白川が担当する授業で活用する予定であり，現段階で評価実験を行った．以下では，評価実験の方法，被験者の反応とアンケートの集計結果について述べる．

6.1 評価実験の方法

被験者には，大阪樟蔭女子大学学芸学部日本文化史学科1回生8名と2回生9名に協力を得た．「基礎ゼミ」の授業時間に，研究の主旨や目的を説明し，実験に協力してもらった．まず，冊子体の『樟蔭學報』とPDF版『樟蔭學報』の両方を見せる．前者は，手にとって自由に触れられるようにし，後者はコンピュータの画面をテレビに映して，キーワードによる検索の様子と注釈を追加する方法について，操作を交えて説明した．実際に，情報やコメントを追加して，注釈を複数の利用者で共有して活用できる仕組みについて説明した．30分間程度，コンピュータの操作を試した後に，簡単なアンケートに回答してもらった．

6.2 被験者の反応

同じように実験を進めたにも関わらず，1回生のクラスと2回生のクラスでの反応は全く異なった．1回生のクラスでは，関心を示す者が多く，非常に活発な議論がなされた．一方，2回生のクラスでは冊子体の『樟蔭學報』にもコンピュータの画面や操作のいずれにも，関心を示す者が少なく，議論も殆どされなかった．以下に，1回生のクラスで議論された内容を列挙する．

- 校歌の楽譜のページで，クリックしたら校歌が聞けるようにしたらいいと思う．

- ♦ 校歌を歌う機会も少ないし，歌えないから．
- ♦ このページの写真の建物が現在はどうなっているかが見られると楽しい．
- ♦ これって，正門のとこの建物？
- ♦ 伊賀駒吉郎って誰れ？
- ♦ 校長先生って書いてある．
- ♦ この本が発行された頃はどんな時代だったんですか？
- ♦ 二二六事件のあった年です．（教授者）
- ♦ 時代背景などがわかるように，年表と一緒に表示できると面白い．
- ♦ 広告の文章が面白い．
- ♦ 蚊が取れるから，カトール？

6.3 アンケートの結果

以下にアンケートの集計結果を示す．アンケートの質問は，予め回答群を用意した 3 問と自由回答 1 問の合計 4 問である．

1. 今日のようにデジタル教材を使う授業について，どう思いますか？【複数回答可】

回答群	人数
面白い	15
面白くない	0
楽しい	7
楽しくない	0
積極的に参加したい	5
あまり参加したくない	0
難しそう	2
操作が面倒	2
現物(本)のほうが良い	3
現物(本)より扱いやすい	8

2. 樟蔭学園の歴史に関心がありますか？【選択】

回答群	人数
ある	1
少しある	13
どちらかというと無い	3
全く無い	0

3. デジタル教材に書き込みをして，知識を共有することについてどう思いますか？【選択】

回答群	人数
関心をもった	9
少し関心をもった	8
全く関心が無い	0

4. 自由な感想を何でも【自由記入】

- もっとデータの数が増えるのがたのしみ*です．
- パソコンを使って自分の知りたいことを一瞬で調べられるのは大変ベンリ*だと思いました．
- 昔のことが分かってよかったと思う．
- デジタル教材を積極的に使ってほしいと思いました．
- 私は本を読むことが好きなので現本*を読むほうが好きですが，最近はマンガもデジタルで読む時代になっているので，どんどん作るべきだと思います．
- 検さく*がしやすそうです．
- 少し画面が見にくかったので，見やすくなればいいと思いました．
- パソコンでデジタル教材として扱うのはすごいと思った．
- どんどん新しい教材ができるのは賛成です．
- 樟蔭の歴史を知ることができるしおもしろそうだと思います．
- 現物より，多くの人が見れるのは良いと思いますが，正直*見やすさは本の方が上*だと思います．
- どういう授業になるのか一度体験してみたい．
- デジタルなら本物のように壊れやすくないので，積極的に活用できそうだと思います．
- キーワードを入れたら調べたいもの*がでてくるというのは便利

* 自由回答の箇所については，被験者が記入した表記のまま転載した．

- 他の資料でも，デジタル教材にしたらおもしろそうだと思いました．
- 校歌の所*に曲がながれるようにしてほしいです．

7.関連研究

國學院大學日本文化研究所で進められている中村等[9]の研究は，デジタルアーカイブシステムパッケージ「でんとうなび®」[10]を利用したWeb版データベースの公開である．國學院大學が所蔵する考古学・民俗学・国文学・文化財学・神道学等の分野の重要な業績と膨大な研究資料の写真資料を中心に，冊子体目録・CD-ROM版データベース・Webデータベースの構築を実現した．ユーザー用メモ欄を設けるなどの工夫が参考になった．

また，大橋[11]の南山大学図書館における収蔵の貴重書『ローマ法大全』をデジタルアーカイブ化に関する報告では，映像を交えた資料紹介のWebページ作成について詳細に述べられている．なかでも，エンドユーザーの立場から見て利用しやすい，かつ独創的な発信内容を満載した利用型のアーカイブ構築したい旨の将来構想には，共感するところが大きい．

多様な資料構造に対応したデジタルアーカイブシステムについて，依田等[12]はメータデータツリーの生成という手法で，検索の柔軟性と拡張性を実現した．利用者の要求に即した情報を網羅的に収集するためには，章や節といった構成要素の単位まで検索できなければならない．検索者の意図とメタデータ間の関係を利用して，検索方法を使い分ける仕組みに注目した．

8.考察と今後の予定

本稿では，昭和戦前期の学園広報誌『樟蔭學報』をデジタル化し，PDFファイルの注釈機能を利用したインタラクティブなデータベースとして作成し，教材として活用することについて述べた．

先行研究[2]におけるカード型データベースのテスト運用で，指摘された問題点を改善し，従来型の画像データとテキストデータを並べて表示する型式に固執せず，コンピュータの画面上でパラパラとページを繰るように閲覧できる仕組みと利用者が内容の充実に参加できるインタラクティブな仕様を実現した．

さらには，データで配布する場合やWebで公開を想定して，データのセキュリティ面の強化にも努めた．

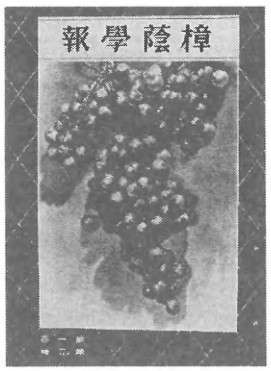
今後は，残り14冊分の注釈を早急に作成し，利用者側の環境を整備した後，実用に備えたいと考えている．

今後，検討しなければならない事項として，以下のような点を挙げておく．

1. ネットワーク上においた同一ファイルに複数の利用者が同時に注釈を付加したときに，更新が正常に行われるか検証する必要がある．
2. データの増加とともにバックアップの作成は必須であり，とくに慎重に取り組みたい．
3. 1冊ずつ個別にPDFファイルに変換したが，1冊単位としたことで検索範囲を狭めてしまい，検索結果を得にくくなっているので，19冊全てを1つのPDFファイルとして作成し直すことも検討する．
4. コンピュータの操作に不慣れな利用者を支援するためのマニュアルづくりにも対応したいと考えている．
5. PDFファイルの検索には，インデックスをつける方法やNamazu[13]による全文検索の方法などが紹介されている情報[14]も得られたので，今後の検討材料としたい．

付録

第一巻第二号から第五号の表紙と創刊号の広告



謝辞

本研究の遂行にあたり，多数のアドバイスをいただいた大阪樟蔭女子大学学芸学部英米文学科小森道彦助教授並びに藤澤良行助教授に心より御礼申しあげたい．また，画像の撮影等データ入力の作業を担当してくれた奈良女子大学大学院人間文化研究科博士前期課程1年の松田憲子さんと評価実験の被験者として協力を得た方々に深謝する．

本研究は，大阪樟蔭女子大学平成17年度特別研究助成費によるものである．ここに記して謝意を表す．

参考文献

[1] http://e-words.jp/w/PDF.html
[2] 竹内さおり，白川哲郎，『設立二関スル書類』のデジタル化とデータベース試作の取り組み－樟蔭学園草創期資料のデータベース化とその活用(2)，大阪樟蔭女子大学（学芸学部）論集第42号，p213-219，2005
[3]「この本を買った人は，こんな本も買っています」という案内．
[4] 清水宏一，治田嘉明，デジタルアーカイブの実践と今後への課題，情報処理学会研究報告，2003-CH-60，p1-8
[5] CD マビカ(MVC-CD500)，SONY 製
[6] Link Station(HD-160LAN)，Direct Station(HD-160U2)，BUFFALO 製
[7] 樟蔭高等女學校本科・専攻科・高等科の卒業生同窓会員
[8] 樟蔭女子専門學校卒業生同窓会員
[9] 中村耕作，黒崎浩行，小川直之，杉山林継，学術調査資料の整理・公開システムの構築－写真資料を中心に－，情報処理学会研究報告 2005-CH-66，p15-22
[10] http://www.toshiba.co.jp/efort/product/digital/index_j.htm
[11] 大橋直美，貴重書『ローマ法大全』と南山大学デジタルアーカイブ，南山大学図書館紀要第8号，2003，p109-116
[12] 依田平，渡邊隆弘，大月一弘，鳩野逸生，岩杉大輔，多様な資料構造に対応したデジタルアーカイブシステム－神戸大学電子図書館アーカイブ検索システム－，情報処理学会研究報告，2003-FI-73，p45-52
[13] http://www.namazu.org/
[14] PDF HACKS，Sid Steward 著，千住治郎訳，オライリー・ジャパン発行，2005
※上記URLの最終確認日は平成17年11月7日である．

階層に基づく遺物データベースシステム
Relic Database System Based on Top-Down Relationships

宝珍 輝尚
Teruhisa Hochin
大阪府立大学大学院 理学系研究科，堺市学園町 1-1
Osaka Prefecture University, 1-1, Gakuen-cho, Sakai-shi, Osaka

あらまし：本論文では，遺跡からの遺物・遺構の出土分布を表示する多様なシステムを，遺跡と遺物・遺構間に存在する上下関係を考慮することによって統一的に扱うデータベースシステムについて検討する．ここでは，分布表示における，「全体」，「検索対象」，ならびに，「集約単位」を表すそれぞれのデータ実体と表示要素のデータ実体を柔軟に対応付け可能とすることにより，遺物等の分布表示システムを汎用的にするとともに，実行時に動的に「全体」，「検索対象」，ならびに，「集約単位」を変更可能とする．

Summary: There is a variety of the systems presenting the distributions of relics obtained from remains. This paper studies on the variety of these kinds of system. By considering the top-down relationships among relics and remains, the variety of the systems is tried to be treated with a uniform framework. The elements called the whole, the aggregation unit, and the target are introduced as the elements for presenting the distribution of relics. Mapping the element on remains to one of these elements can bring us the followings merits. The variety of the systems presenting the distribution of relics can be uniformly treated. From the point of the view of information system, this seems to be preferable. The displaying elements can be changed in the retrieval time. The system could be very flexible.

キーワード：遺物データベースシステム，上下関係，遺跡，柔軟性
Keywords: relic database system, top-down relationships, remains, flexibility

1. はじめに

近年のコンピュータの進歩はめざましく，考古学へのコンピュータの導入が盛んに行われている．考古学データベースも構築され公開されるようになってきている[1-5]．遺跡は地理上に存在するので，遺跡の情報は良く地理情報と関連付けて管理される．また，地図上に遺跡を表示するなどして考古学的な考察を行うことも行われてきている[6-8]．遺跡からの遺物の出土分布を表示するシステムもこの範疇のシステムである．

ここで，遺跡からの遺物等の出土分布表示にも様々なものがある．例えば，ある地域を表示しその中に存在する遺跡で指定された遺物が出土しているか否かを地図上で表示するシステムもあれば，ある遺跡の特定の区画を表示しその区画の中から指定された遺物の出土状況を表示するシステムもある．遺跡からの遺物の出土分布を行うという点では同じであるが，同じシステムとは言い難いほどシステム的には異なるものと考えられる．このような多様なシステムをそれぞれに実現するのは効率的ではない．情報システムの観点から見ると，これらの多様なシステムを統一的な枠組みの下で統一的に扱う必要があると考えられる．

著者は，遺跡からの遺物等の出土分布表示を行う多様なシステムを対象とし，遺跡と遺物間の上下関係に着目して多様性に対処する方法について考察してきている[11,12]．ここでは，遺跡と遺物に関係する事物の上下関係を明確化し，これらの上下関係と表示における上下関係を対応付けることによって汎用性を持たせようとしている．

本論文では，遺跡からの遺物・遺構の分布表示を行う様々なシステムの多様性を克服し様々な遺物・遺構分布表示システムを統一的に扱うことを目的として，上記の枠組みに基づいたデータベースシステムの実現について報告する．

以下，2．では，まず，遺跡と遺物・遺構間の上下関係について述べる．次に，3．で考古学データベースシステムの実現について述べる．そして，4．で実行例を示す．最後に，5．でまとめる．

2．遺跡と遺物・遺構間の上下関係

システム全体の実体の実体関係図を図1に示す．ここでは，実体クラスを四角形で表し，関係を線分で表している．また，1：多の多側を線分の端の黒ひし形で表している(この表記は通常の実体関係図の表記と異なるので注意されたい)．例えば，地域には遺跡が複数あるといったぐあいである．発掘調査の調査区を区画に分けているので，「発掘調査」クラスと「区画」クラスの関係は1：多である．また，“堺市”は“大阪府”に含まれるという関係もあるため，地域は地域自身と1：多の関係を持っている．また，クラス間の関係を表す線分の端の黒丸は，0個か1個の要素があることを表す．

さらに，図1において，実線の四角形は基本クラスを表す．基本クラスは，本システムに必須のクラスである．これに対し，点線の四角形は導出クラスを表す．導出クラスは，場合によって必要となるクラスで，ビュー機構等により基本クラスや導出クラスから導出されるクラスである．図1では，既知の導出クラスを図示している．また，三角形はスーパクラスとサブクラスの関係を表す．なお，導出クラスについては，そのもととなるクラスとの間の関係のみを記述している．例えば，「遺物群」と「遺物」の間の関係である．その他のクラスとの関係は，煩雑になるので図示していない．

「全体」，「集約単位」，ならびに，「対象」は，それぞれ，表示における全体，対象をまとめる単位，ならびに，検索の対象を表す[11]．ここでは，これらのクラスのデータ実体と「表示要素」クラスのデータ実体を柔軟に対応付けることで，汎用的な表示システムとしようとしている．また，「領域」の情報として地図情報が不可欠であるが，情報が多く重要であるので「領域」と分けて表現している．

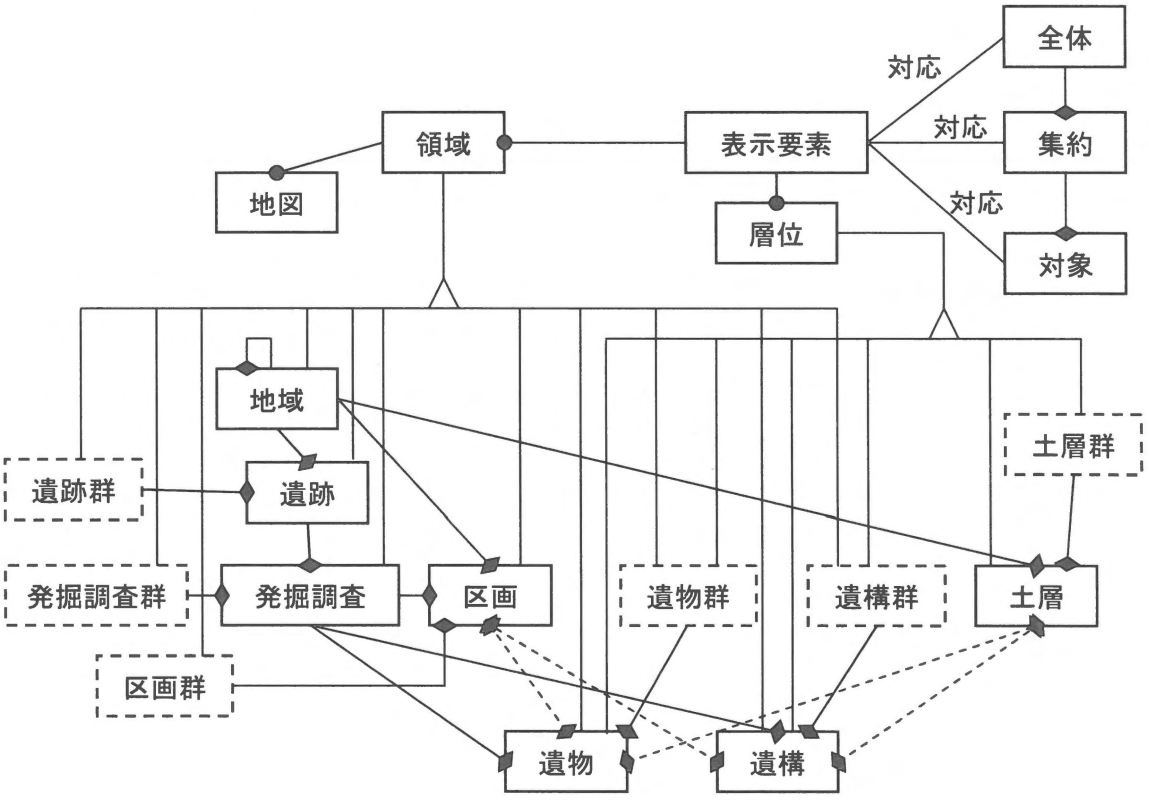


図1　全体の実体関係図

3. 実現について

3. 1 要求

本システムでは，実行時（検索時）に，「全体」，「集約単位」，ならびに，「対象」の指定を変更可能としたい．例えば，“遺跡”を「全体」とし，“発掘調査”を「集約単位」とし，“遺物”を「対象」として検索した後，「対象」を“遺構”に変更して検索するといったことを可能にするということである．

また，「全体」，「集約単位」，ならびに，「対象」には，実体クラスを対応可能とする．すなわち，基本クラスのみではなく導出クラスも対応させることを可能とする．導出クラスは管理者によって定義可能とすることを考えているので，固定的な基本クラスのみでなく，任意のクラスを対象とするよう考慮する必要がある．

さらに，可搬性を考慮し，オブジェクト指向データベースといった特別なデータベースではなく，ごく普通の関係データベースを用いることができるようにする．

また，特別なソフトウェアを用意しなくても実行できるよう，ウェブブラウザを用いて実行できるものとする．

3. 2 情報の管理

3. 1で述べた要求を満足するためには，検索時に動的にSQL文を生成する必要がある．SQL文は，関係データベースのための標準データベース言語であるSQLで書かれた命令文である．

3. 2. 1 定義情報の管理

検索のためには，検索対象としてどのようなテーブル（ここでは，クラス）が存在し，各テーブルにはどのような属性が存在するかという情報が必要である．

この情報は，システムで独自に管理しておくこともできれば，データベースシステムに問合せを行って得ることもできる．ただし，検索対象以外にも多くのテーブルが存在するので，検索対象としてどのようなテーブルを対象とするかをシステムで独自に管理している方が効率的である．

また，検索に使用される属性はどれかという情報も把握しておく必要がある．これは，テーブルには，システム上必要な属性等，検索に使用されない属性が含まれているからである．

3. 2. 2 表示情報の管理

検索条件の情報を得るために，ウェブブラウザ上でこのための表示を行う必要がある．例えば，文字列を入力してもらうためのテキストフィールドを用意したり，既存の値から選択してもらうためのメニューを用意したりするということである．このための情報は，3. 2. 1で述べた定義情報と重複するものもあるが，例えば，文字列でも，テキストフィールドで入力してもらうのか，メニューで入力してもらうのかという情報や，その際のテキストフィールドの長さ，ブラウザで表示する上での項目名といった情報が必要である．

3. 2. 3 上下関係の管理

図1に示したような実体クラス間の上下関係を管理しておき，実行時に情報を得る必要がある．これは，「全体」と指定された実体クラスから，「集約単位」と指定された実体クラスを通って，「対象」と指定された実体クラスへのパスを求める必要があるためである．

これは，構成関係を管理することになるが，構成関係を表現するには2つの方法がある．一つ目の方法は，子のテーブルのキーに親のキーを含ませる方法である．つまり，子のテーブルのキーを構成する属性集合に親のキーを構成する属性集合を包含させるのである．これにより必然的に親を指定したことになる．二つ目の方法は，親のキーを構成する属性をキー以外の属性としてテーブルに含める方法である．これは外部キーとなる．

また，単なるクラス間の上下関係ではなく，上下関係がどの属性によって対応付けられるかという情報も必要である．これは，外部キーの管理と同様である．

3. 3 結合に関する処理

3. 3. 1 クラスの結合

検索に当たっては，複数のテーブルを結合して検索することが不可欠である．ここで，図1にも示したように，テーブル間は単なる上下関係ではない．例えば，

「発掘調査」は「遺物」の直接の親になっているが，「区画」を通して間接的にも「遺物」の親になっている．このように経路が複数有り得るので単に経路をたどればよいというわけではない．

結合回数が多いと検索時間が長くなる可能性が高いので，結合回数はできるだけ減らさなければならない．したがって，テーブル間の上下関係を表すグラフから，あるテーブルを開始点としたときの最短経路を求める必要がある．ここでは，ダイクストラのアルゴリズム[13]を用いて最短経路を求める．

結合条件は，3．2．3で述べた上下関係を表す情報を利用して生成する．

3．3．2　自己結合処理

通常の関係データベースシステムは再帰問合せをサポートしていないので，「地域」のように自己参照を行っているテーブルには特別の対処が必要である．

ここでは，最大の繰り返し回数を与えて，最大その回数までの結合を繰り返し，開始の実体データから目的の実体データが得られるまでの（自己）結合の回数を求め，これを利用することにする．

3．4　SQL 文生成処理

SQL で検索を行うための SELECT 文を作成する．ここでは，検索対象となるテーブル，検索結果とする属性，テーブルの結合条件を含む検索条件，ならびに，グループ化の単位を指定する属性を求めなければならない．

3．4．1　検索対象テーブル

検索対象とするテーブルは，「全体」から「集約単位」までのテーブルの結合と，「集約単位」から「対象」までのテーブルの結合に要するテーブルとして求める．「全体」が自己参照を含む「地域」の場合は，3．3．2で述べた自己結合処理を行う．その他の場合は，3．3．1で示した結合処理として扱う．

「全体」，「集約単位」と「対象」の指定とテーブルの定義によっては，結合回数の削減を図れる可能性があるが，これは今後の課題である．

3．4．2　検索条件

検索対象に対する検索条件は，概ね，文字列か数値かによって処理し分ける．文字列の場合は，LIKE 述語を使用した文字列部分検索を行う．

結合条件は3．4．1の処理の過程で求められる．

3．4．3　検索結果

検索結果は，グループ化の単位を指定する属性と，個数を求めるための COUNT 述語となる．従って，グループ化の単位を指定する属性が求まれば求まることになる．

3．4．4　グループ化の単位

グループ化の単位は「集約対象」としたテーブルの行である．したがって，「集約対象」としたテーブルのキーを構成する属性を求めれば良い．

4．実行例

現在，PHP5 [14]を用いて実装を行っている．プラットフォームは MacOS X 10.3 サーバで，Apache2 を用いて開発を行っている．

試作版の検索開始画面を図2に示す．画面の上部に，「全体」，「集約単位」，ならびに，「検索対象」を表示している．図2では，「全体」は「発掘調査」で“朝倉氏遺跡第50次発掘調査”，「集約単位」は「区画」，「検索対象」は「遺物」である．

次に，「全体」として指定されたもの（ここでは，“朝倉氏遺跡第50次発掘調査”という名称の「発掘調査」のデータ実体）に登録されている地図を表示し，そこに「集約単位」として指定されたもの（ここでは，「区画」）を表示している．図2では，集約単位は四角形の集まりを表示しているが，遺構や（広がりを持った）面として登録された遺跡の場合は，多角形の集まりが表示される．

その下に「検索対象」に対する検索条件を指定するテキストフィールド等を配置している．図2では，テキストフィールドのほかに，メニューやラジオボタンを例として示している．「検索」ボタンを押下することにより検索

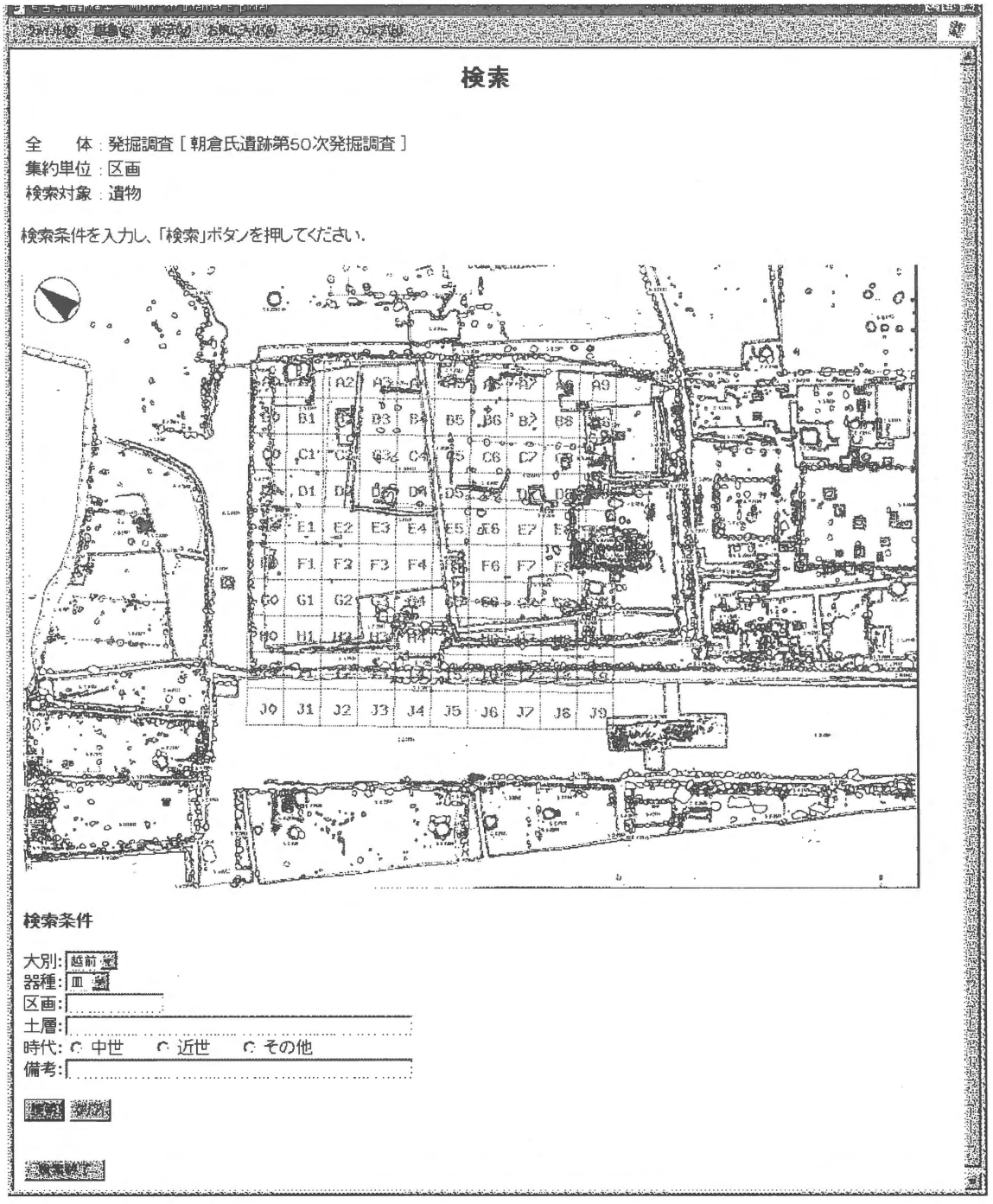

検索
全　　体：発掘調査［朝倉氏遺跡第50次発掘調査］
集約単位：区画
検索対象：遺物
検索条件を入力し、「検索」ボタンを押してください。
検索条件
大別:
器種:
区画:
土層:
時代: 中世 近世 その他
備考:

図2　検索開始画面の例

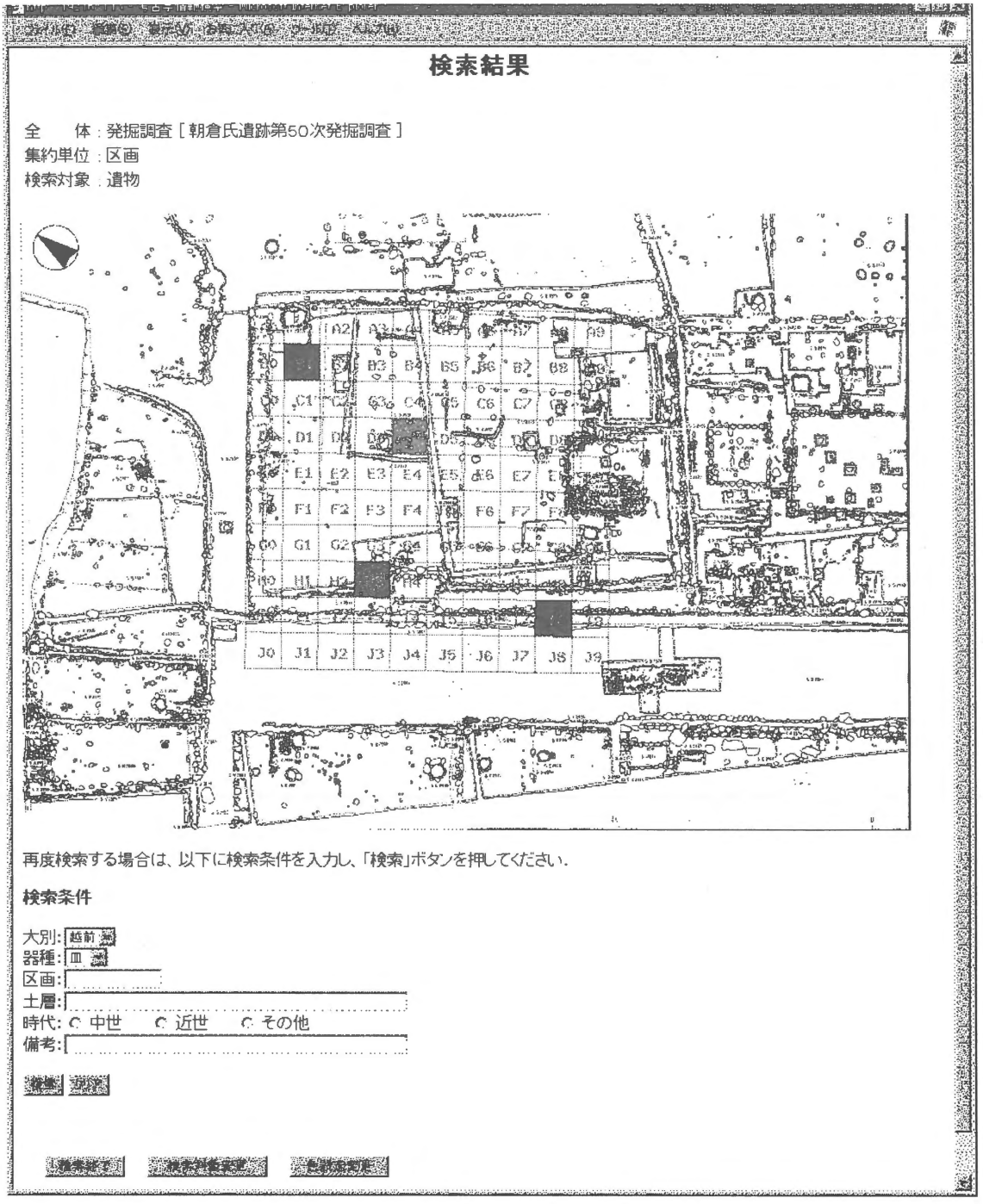

検索結果
全　　体：発掘調査［朝倉氏遺跡第50次発掘調査］
集約単位：区画
検索対象：遺物
再度検索する場合は、以下に検索条件を入力し、「検索」ボタンを押してください。
検索条件
大別:
器種:
区画:
土層:
時代: 中世 近世 その他
備考:

図3　検索結果の画面例

の実行を行う.

検索結果の画面例を図3に示す. 図では分かりにくいが,「集約対象」ごとに検索条件に合致した遺物の個数を色分けで表示している. 現在は表示していないが, 色と個数の関係を表す凡例を表示する予定である. また, 個数に対する色の設定を変更可能とする予定である.

検索結果が満足のいくものでない場合は, 検索条件を修正して再度検索を行うことが可能である.

検索対象(「全体」,「集約単位」,「対象」)を変更する場合は検索を最初からやり直すことになる. これは, 表示する地図,「集約単位」として表示する多角形, あるいは, 検索条件が変更になるためである.

5. おわりに

本論文では, 遺跡からの遺物の出土分布を表示する多様なシステムを対象とし, その多様性を, 遺跡と遺物・遺構間に存在する階層性を考慮することによって克服し, 統一的に扱う枠組みに基づいたデータベースシステムの実現について報告した. ここでは, まずシステムに対する要求を示し, 検索に関する実現について述べた. ここでは, 管理すべき情報, 結合に関する処理, ならびに, SQL 文の生成を取り上げ, 重要な事項について説明した.

本論文では, 試作中の考古学データベースシステムの中心的な機能である検索について述べた. データベースを構築・維持管理してゆくには, データの登録・修正といった機能は不可欠であるが, これらについては述べてこなかった. 今回は述べていないが, 検索にあたって必要最低限必要となる, 地図, 遺跡や区画等の登録機能は試作済みである. 今回の試作システムはウェブブラウザで動作することを前提としているが, ウェブページの記述に使用されるHTMLを使用すると, データの記述と表示のための記述が混在する. これは, データの追加により表示を変更したり, 本格始動にあたり表示を整えたりすることを考えると効率の良い方法とは言い難い. ここでは, データの記述と表示のための記述を分離して実装するために, XML と XSL[15]を用いて試作している.ただし, 現在, XSL をサポートしているウェブブラウザは限られており, 可搬性という点で問題ではある. しかし, データ登録では, 同様の表示を行うことが多く, XSL を用いることでかなり効率的に実装できるという感触を得ている.

また, 試作中の考古学データベースシステムは, 考古学研究者に使って頂けることが重要である. この点で, 使い勝手の評価は重要である. 早期に使い勝手を評価し, 改善点を探る必要があると考えている.

さらに, データの国際標準化に対する対応[16]も必要である. ダブリン・コアで規定されている項目はもちろん, 博物館資料の記述に必要な項目の登録を可能とするようにする予定である.

謝辞

データの収集, 分類作業, ならびに, 考古学データの管理に関する議論などでお世話になっている福井県立一乗谷朝倉氏遺跡資料館の水村伸行氏に感謝いたします. また, 本研究は, 一部, 文部科学省科学研究費補助金(課題番号 16019201)による. ここに記して謝意を表す.

参考文献

[1] 及川 昭文:考古学データベース－過去を復元するマルチメディア技術－,情報処理, Vol. 38, No. 5, pp. 388-391 (1997).

[2] 八重樫 純樹:思考の道具としてのパーソナルコンピュータ,第2回考古学におけるパーソナルコンピュータ利用の現状, pp. 37-41 (1989).

[3] 山田 康晴:遺跡における遺物出土地点のデータベース化,第3回考古学におけるパーソナルコンピュータ利用の現状, pp. 22-30 (1990).

[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, pp. 377-382 (1997).

[5] 宝珍 輝尚, 都司 達夫:考古学データベースシステムの現状と課題, 電子情報通信学会誌, Vol. 85, No. 3, 171-175 (2002).

[6] 横山 隆三, 千葉 史:地理情報システムを用いた遺跡データベース構築,情報考古学, Vol. 3, No. 2, pp. 29-40 (1997).

[7] 及川 昭文:貝塚データベース-インターネットによる公開とコラボレーション-,日本情報考古学会第10回大会発表要旨, pp. 13-18 (2000).

[8] 加藤 常員:高地性集落遺跡データベースにもとづく古代ノロシ通信路の推定,人文学と情報処理, No. 19, pp. 46-51 (1999).

[9] 三宮 健, 岡安 光彦, 吉川 正俊, 植村 俊亮:考古学情報のデータモデルとその実装,情報考古学, Vol. 6, No. 2, pp. 11-18 (2000).

[10] 宝珍 輝尚, 都司 達夫:利用者適応型考古学データベースシステムに関する一考察,情処研報 CH54-8, pp. 51--58 (2002).

[11] 宝珍 輝尚:遺跡と遺物・遺構表示の階層性に関する一考察, 第 10 回公開シンポジウム「人文科学とデータベース」, pp. 27-32 (2004).

[12] 宝珍 輝尚:遺跡と遺物・遺構表示の階層性に関する一考察, 日本情報考古学会第19回大会発表要旨, pp. 34-39 (2005).

[13] A. V. エイホ, J. E. ホップクロフト, J. D. ウルマン(大野義夫訳):データ構造とアルゴリズム, 培風館 (1987)

[14] 日本 PHP ユーザ会:http://www.php.gr.jp/

[15] 高橋麻奈:やさしい XML, ソフトバンク (2001)

[16] 和久田聖衣, 八重樫純樹:デジタルアーカイブの調査研究－博物館情報の標準化動向を中心に－, 静岡大学情報学研究, Vol. 10, pp. 127-146 (2004)

■公開シンポジウム「人文科学とデータベース」1995年プログラム
1995年12月25日(月曜日)、26日(火曜日)　大阪電気通信大学(寝屋川キャンパス)

特別講演

古地震データと活断層(pp.1-4)
寒川旭(通産省地質調査所)

一般講演

Intelligent Padシステムを用いた歴史学研究支援データベースの構築(pp.5-12)
赤石美奈・中谷広正・伊東幸宏・阿部圭一・田村貞雄(静岡大学)

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(pp.13-18)
小林努・加藤常員・小沢一雅(大阪電気通信大学)

視点に依存する属性付け機構をもつ木簡研究支援システム
－構造化型データベースの概念－(pp.19-28)
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)

古典籍とJIS漢字
－テキストの本文校正とのかんけいについて－(pp.29-36)
當山日出夫(花園大学)

手書き文字時系列筆跡パターンの一解析と今後の計画(pp.37-42)
東山孝生・山中由紀子・澤田紳一・中川正樹(東京農工大学)

絵画DBとイメージ検索
－浮世絵の線画表現とデータ圧縮効果－(pp.43-48)
濱裕光・志賀直人(大阪市立大学)

画像データベースの自然言語インターフェースについて(pp.49-54)
伊東幸宏・中谷広正(静岡大学)

多視点距離データを用いた3次元形状モデリング(pp.55-60)
横矢直和(奈良先端科学技術大学院大学)・増田健(電子技術総合研究所)

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(pp.61-66)
上村隆一(福岡工業大学)

「歌物語」語彙の数量的分析と研究(pp.67-74)
西端幸雄(大阪樟蔭女子大学)

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(pp.75-82)
亀田弘之(東京工科大学)・藤崎博也(東京理科大学)

社会調査結果の視覚化データベース(pp.83-88)
吉田光雄(大阪大学)

「間」に関するデータベースの構築(pp.89-98)
中村敏枝(大阪大学)

方言音声データベースの作成と利用に関する研究(pp.99-104)
田原広史・江川清・杉藤美代子・板橋秀一(大阪樟蔭女子大学)

■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
1996年12月21日(土曜日)　大阪電気通信大学(寝屋川キャンパス)

招待講演

三浦梅園の主著『玄語』のデータベース化と解読の試み
－江戸時代のハイパーテキストを読み解く－(pp.1-10)
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(pp.11-22)
黒川隆夫(京都工芸繊維大学)

一般講演

Shape from motionを応用した什器類の立体データ作成(pp.23-34)
中島重義・岡本次郎・濱裕光・細川省一(大阪市立大学)

古地図に描かれた内容のデータベース化の試み(pp.35-44)
出田和久・正木久仁・小方登・山近博義(奈良女子大学)

考古学のためのデータベースシステム(pp.45-54)
宝珍輝尚・中田充・白井治彦・都司達夫(福井大学)

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(pp.55-56)
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(pp.57-66)
原田章・吉田光雄(大阪大学)

『方言認知地図』プログラムと統計処理地図(pp.67-78)
ダニエル・ロング(大阪樟蔭女子大学)

■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
1997年12月20日(土曜日)　大阪電気通信大学(寝屋川キャンパス)

招待講演

「邪馬台国大和説を科学する」

倭人社会と卑弥呼の王権(pp.1-10)
吉田晶(岡山大学名誉教授)

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－(pp.11-17)
広瀬和雄(奈良女子大学)

一般講演

歴史学研究支援システムの構築(pp.19-30)
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－(pp.31-38)
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とシステム
－知見と仮説に基づく再構造化－(pp.39-46)
森下淳也・大月一弘(神戸大学)・上島紳一(関西大学)・大庭脩(皇學館大學)
杉山武司(姫路獨協大学)

短編推理小説の論理構造の分析(pp.47-54)
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－(pp.55-64)
モハメド・アミラン・ブイヤン・阿古弥寿章・濱裕光(大阪市立大学)・松平進(甲南女子大学)

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－(pp.65-72)
濱裕光(大阪市立大学)

■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
1998年12月9日(土曜日)　大阪電気通信大学(寝屋川キャンパス)

招待講演

邪馬台国は北九州甘木市付近にあった(pp.1-22)
安本美典(産業能率大学)

特別講演

選挙研究とデータベース(pp.23-24)
三宅一郎(関西大学)

一般講演

中期インド・アリアン聖典のデータベース(pp.25-34)
逢坂雄美・山崎守一(仙台電波工業高等専門学校)・宮尾正大(室蘭工業大学)

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(pp.35-42)
加藤常員(大阪電気通信大学)

データベース倫理について(pp.43-52)
江澤義典(関西大学)

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(pp.53-59)
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－(pp.61-70)
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－(pp.71-78)
橋本隆之・青木功介・釣裕美(インテックシステム研究所)

■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
1999年12月18日(土曜日)　関西大学総合情報学部

招待講演

顔の認知研究における顔データベースの利用(pp.1-10)
加藤隆(関西大学)

Recent Research on Paleolithic Arts in Europe and the Multimedia Database (pp.11-22)
Cesar Gonzalez, Roberto Cacho Toca (University of Cantabria, Spain)

一般講演

Photo VR 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－(pp.23-34)
深沢武雄(テクネ)

バーチャルリアリティによる遺物探訪　3D Archaeo-Copter (pp.35-38)
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(pp.39-46)
結城宏和・宝珍輝尚・都司達夫(福井大学)

文字データベースのための文書の構造化と意味管理(pp.47-58)
横田一正・三宅忠明・国島丈生(岡山県立大学)
劉渤江(岡山理科大学)・田槙明子(リョ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(pp.59-66)
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(pp.67-78)
芳野学・都司達夫・宝珍輝尚(福井大学)

■第6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
2000年12月20日(土曜日)　静岡大学情報学部

特別講演

「ええじゃないか」研究を振り返って(pp.1-2)
田村貞雄(静岡大学)

一般講演

文化資源情報における Topic Map の適用と評価
－『源氏物語』画像 DB を使ったモデル作成の試み－(pp.3-10)
長瀬真理(静岡大学)

文字冗長度による日本文学分析(pp.11-16)
福田宏(静岡県立大学)・山下泰弘(科学技術振興事業団)・勝矢光昭(静岡県立大学)

広領域分野資料の横断的アーカイブ論に関する基礎研究(pp.17-26)
八重樫純樹(静岡大学)

ジャイナ教聖典のデータベース
－特殊フォントで表現されたデータの PDF－(pp.27-34)
逢坂雄美(仙台電波工業高等専門学校)

パスカルデータベースシステム（1）(pp.35-42)
白石修二(福岡大学)

吾妻鏡データベースの構築(pp.43-52)
安道百合子(国文学研究資料館)

全国遺跡データベースの構築(pp.53-62)
森本晋(奈良国立文化財研究所)

地理情報システムを用いた城下町の復元的研究
－彦根城下善利組足軽屋敷地図を中心として－(pp.53-72)
生方美菜子・濱崎一志(滋賀県立大学)

考古学データベースにおける検索エンジンの研究(pp.73-80)
三浦宙明・小沢一雅(大阪電気通信大学)

■第7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
2001年11月17日(土曜日)　関西学院大学情報メディア教育センター

特別講演

感性情報研究の動向とデータベース(pp.3-10)
井口征士(大阪大学基礎工学部)

一般講演

広重の版画と江漢作らしい油絵
－二つの「東海道五十三次」について－(pp.11-27)
荒木啓介(科学技術振興事業団)　資料提供、大畠洋一(東海道研究家)

古文書文字列に対するキャラクタスポッティング(pp.29-38)
橋本智広・梅田三千雄(大阪電気通信大学)

日商簿記検定問題の電子化(pp.39-46)
福田宏・小津稚加子(静岡県立大学)

米国におけるテレビニュースデータベース構築の歴史的・法的経緯と現状について
(pp.47-50)
魚住真司(関西外国語大学)

画像・音声の超高再現性PCの開発(pp.51-63)
片岡裕(大谷大学)

全国遺跡データベースの構築　2001年度の動向(pp.65-68)
森本晋(奈良文化財研究所)

日本語キエルケゴール文献データベース(pp.69-72)
平林孝裕・橋本淳(関西学院大学)

前方後円墳データベース検索システムと地理情報処理(pp.73-80)
西上昌治・小澤一雅(大阪電気通信大学)

貝類のマルチメディアデータベースの構築(pp.81-88)
高田茂樹・雄山真弓(関西学院大学)

■第８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
2002年12月２１日(土曜日)　帝塚山大学

特別講演

・空間コンテンツの検索とプレゼンテーション
田中克巳（京都大学）

自由論題報告

・尾張藩士・朝日文左衛門の生活行動空間
－GISを用いた『鸚鵡籠中記』の分析－
村田祐介（名古屋大学研究生）

・３次元空間を共有するマルチユーザCSCW環境と
仮想考古遺跡ウォークスルーへの応用
坂田義則（立命館大学院生）・八村広三郎（立命館大学）

SYMPOSIUM“人文科学における空間情報の利用”

Ⅰ．古地図・衛星画像の分析と利用

・高精細絵図画像データの利活用　－阿波国絵図・徳島城下絵図を例に－
平井松午（徳島大学）

・分散型GIS「GLOBALBASE」の実装
森洋久（国際日本文化研究センター）

Ⅱ．考古学におけるGIS利用

・古環境復原のための考古学情報クリアリングハウスの構築
—Java Scriptと Dynamic HTMLを使用して—
河野一隆（九州国立博物館(仮称)設立準備室）・塚本敏夫（元興寺文化財研究所）・
魚津知克（大手前大学）

・平安京における空間情報システムの整備と条坊復原
宮原健吾（京都市埋蔵文化財研究所）・内田賢二（ライカジオシステムズ）

・中世都市の景観と構造
鋤柄俊夫（同志社大学）

■第9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
2003年12月20日(土曜日)　大阪工業大学　情報科学部

特別講演　テーマ：著作権

・他人の著作権を侵さないために　－ 法律的視点からディジタル著作権
岡村久道(英知法律事務所　弁護士)

・自分の著作権を守るために　－ 技術的視点から
佐野睦夫(大阪工業大学)

一般講演

・コンピュータグラフィックスを用いた花型学習システム
三原比呂美，西尾孝治，小堀研一(大阪工業大学)

・デジタルアーカイブとデータベース構築
清水宏一，山口豊博(京都デジタルアーカイブ研究センター)

・絵巻物を利用した心理検査支援システム
池田瑞穂，雄山真弓(関西学院大学)

・浮世絵を通してみた江戸時代女性の人体表現について
森下あおい(成安造形大学)，黒川隆夫(京都工芸繊維大学大学院)

・縄文語による地名語源の解釈－山名の例を中心に－
永田良茂

・奈文研航空写真検索システム(NARS)について
森本晋(独立行政法人文化財研究所奈良文化財研究所)

・兼永本古事記・出雲国風土記データベースの構築
松本智子(国文学研究資料館研究情報部)

・心理学入門のWeb自習教材として項目反応理論を適用した項目プール（データベース）の開発
田崎美弥子(東京理科大学)

・感性に基づく動画検索について
井田俊博(福井大学大学院)，宝珍輝尚(大阪府立大学)，都司達夫，樋口健(福井大学)

■第 10 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

2004 年 12 月 18 日(土曜日)　大阪府立大学　総合科学部　数理・情報科学科

特別講演

・考古学データベースとＸＭＬ…… 1
八重樫 純樹（静岡大学）

招待講演

・中世の国際都市 堺……11
續 伸一郎（堺市立埋蔵文化財センター）

一般講演

地名分析および考古学データ表現

・百名山からの縄文地名解釈……17
永田 良茂

・遺跡と遺物・遺構表示の階層性に関する一考察…… 27
宝珍 輝尚（大阪府立大学）

人文科学と柔構造データベース

・考古学情報のＸＭＬ記述とデータベースシステム～データ統合と横断検索へのＸＳＬＴの利用
……
小笠原 和慶、八重樫 純樹（静岡大学）

・柔軟な構造を持つデータベース管理システムを用いた万葉集検索システムの構築法…… 33
中田 充（山口大学）

前方後円墳データ分析

・前方後円墳の築造と方位観念・社会的背景の復元にかんするデータベースの活用…… 43
北條芳隆（東海大学）

・前方後円墳の墳形計測と築造企画…… 53
小沢一雅（大阪電通大学）

主催：第１１回公開シンポジウム
実行委員会

後援：人文系データベース協議会

委員長：高橋晴子（大阪樟蔭女子大学）
委　員：西端幸雄（大阪樟蔭女子大学）
田原広史（大阪樟蔭女子大学）
竹内さおり（大阪樟蔭女子大学）
田窪直規（近畿大学）

議長：小沢一雅　（大阪電気通信大学）
出田和久　（奈良女子大学）
上原邦昭　（神戸大学）
江澤義典　（関西大学）
及川昭文　（総合研究大学院大学）
雄山真弓　（関西学院大学）
加藤常員　（大阪電気通信大学）
川口　洋　（帝塚山大学）
柴山　守　（大阪市立大学）
高橋晴子　（大阪樟蔭女子大学）
都司達夫　（福井大学）
辻田忠弘　（甲南大学）
中谷広正　（静岡大学）
中村敏枝　（大阪大学）
八村広三郎（立命館大学）
深海　悟　（大阪工業大学）
宝珍輝尚　（大阪府立大学）

人文系データベース協議会 第１１回公開シンポジウム「人文科学とデータベース」

発行日　2005年12月3日
発行所　第１１回公開シンポジウム実行委員会
〒577-8550　大阪府東大阪市菱屋西4-2-26
大阪樟蔭女子大学　学芸学部
高橋晴子（シンポジウム実行委員長）
Email:takahashi.haruko@Osaka-shoin.ac.jp