第7回公開シンポジウム

人文科学とデータベース 「データ」を読む・観る・解く

2001年11月17日

第7回公開シンポジウム実行委員会 主催 関西学院大学情報メディア教育センター 協賛 人文系データベース協議会 後援

目 次

(特別講演)

	「 感性情報研究の動向とデータベース」・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	3	
	(一般講演)		
1.	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―・・・・ 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)	11	
2.	古文書文字列に対するキャラクタスポッティング・・・・・・・・・・ 橋本 智広 梅田 三千雄(大阪電気通信大学大学院 工学研究科)	29	
3.	日 商簿記検定問題の電子化・・・・・・・・・・・・・・・・・ 福田宏 小津稚加子(静岡県立大学経営情報学部)	39	
4.	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 魚住真司(関西外国語大学) ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	47	
5.	画像・音声の超高再現性 PC の開発・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	51	
6.	全国遺跡データベースの構築 2001 年度の動向・・・・・・・・・・・ 森本 晋(文化財研究所 奈良文化財研究所)	65	
7.	日本語キエルケゴール文献データベース・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	69	
8.	前方 後円墳データベース検索システムと地理情報処理・・・・・・・・・・ 西上昌治 小澤一雅(大阪電気通信大学)	73	
9.	貝類のマルチメディアデータベースの構築・・・・・・・・・ 高田茂樹 雄山真弓(関西学院大学 情報メディア教育センター)	81	
•	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・・・・・・ 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・・・・・・・ 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・・・・・・・ 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・・・・・・・ 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・・・・・・・ 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・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	90 91 92 93	

感性情報研究の動向とデータベース

Recent Trends of Kansei Information Processing and Kansei Database

井口征士 INOKUCHI Seiji

大阪大学大学院基礎工学研究科システム人間系

Department of Systems and Human Science Graduate School of Engineering Science, Osaka University Email: inokuchi@sys.es.osaka-u.ac.jp

あらまし:近年関心が高まっている感性情報 処理に関するプロジェクトを、ビデオを用いて紹介する。感性情報処理が唱えられて 10年以上になり、その間いくつかの大型プロジェクトが実施された。ヒューマンインタフェース、デザイン、アート創作、データ検索などに用いられる感性的アプローチを紹介し、後半で感性データベースについて触れる。

Summary: Recently in computer technology, there has been a great interest in how to deal with Kansei information. "Kansei" comes from a Japanese word, meaning sensibility, sentiments, the sense, emotion and/or feeling. Kansei information processing is effective for affective communication systems and intuitive retrieving system. This paper describes a survey of recent research projects dealing with Kansei information processing and also mentions Kansei database technology.

キーワード:感性情報、ヒューマンメディア、 人間中心システム、マルチメディアアート、 感性データベース

Keywords :Kansei information, Human media, Human centric system, Multimedia art, Kansei database

1. はじめに

近年、科学技術分野において大きな価値 観の転換が進んでいる。物質・エネルギー の世界でのこれからの主要価値は「環境」 であり、情報・社会における中心は「人間」 である。ここで社会における人間の役割を 考えると、従来の人間の役目は産業システ ムや工業システムを効率的に安全に操作す ることであり、その中心は常に機械であっ た。たとえ人間にサービスすることを目的 とするシステム、例えば自動販売機のよう な機械であっても、機械にとって都合のい い機能で作り上げられていた。1980年代後 半に生まれた「人間中心システム」の考え は、このような機械や産業や社会が、人間 にとって好ましい構造を持つべきであると の主張に基づいている。

このような人間中心の主張は必ずしも最近に始まったものではない。我々の身の周りで、ヒトと直接接する製品、例えば車、住居、服飾などは、単に「性能」といったモノの側からの評価だけでなく、それを扱うヒトの好みや嗜好といった感性的要素が

重要であり、このようなアプローチは、1980年代、日本が豊潤な時代を迎えたころから、多方面で芽生えている。例えば、クルマのデザインや快適性といった、本来の機能以上の付加価値に力を注ぎ始めた過程は、日本のある自動車メーカーが米国で行った講演で "Kansei" なる言葉を用いていることからも伺える。以来このような工学的アプローチは、広く産業分野で活用され、有用性を発揮してきた(1)。

近年改めて感性を情報科学の立場から見直す機運が生じたのは、HCI (Human Computer Interaction) や HI (Human Interface) において、人に優しいシステムを実現するニーズが発端になっているといえる。この背景には、機械には常識がなくフレキシブルな対応ができないといった知識処理レベルの問題があると同時に、機械は無機質で感情のような「ヒトの感じ方」にはまったく無頓着であることが挙げられる。人間の顔の表情を機械が感じ取ったり、あるいは機械が人間に豊かな顔の表情で語りかけるなら、より好ましいコミュニケーションが実現できるであろう。

2.感性情報処理の初期プロジェクト 感性情報処理という言葉が広く認知されるようになったのは、1992年にスタートした科学研究費重点領域研究「感性情報処理の情報学・心理学的研究」であった⁽²⁾。このプロジェクトは、当時大阪大学教授の辻三郎を代表者として、1992年から1995年にかけて実施されたプロジェクトで、全国の28大学から56研究グループが参加している。この研究プロジェクトの特徴は、情 報処理分野の研究者に加えて全研究者の3 分の1余が心理学者である点である。研究 班の構成は以下の通りであった。

総括班「感性情報処理の情報学・心理学的

研究」 班長:(阪大) 辻 三郎

A班「感性情報の基礎とモデリング」

班長:(京大)長尾 真

B班「メディアにおける感性情報の抽出・

表現」 班長:(阪大)井口征士

C班「行動空間の感性情報処理」

班長:(東大)中谷和夫

D班「コミュニントーションにおける感性情報処理の研

究」 班長:(東大)原島 博 E班「感性デザインおよび感性データベースの研

究」 班長:(慶大)鷲見成正

本研究でカバーしていた研究テーマは図1に示すように多岐に渡り、心理学の研究から、情報科学をベースとしたモデリング、さらにメディア技術、ヒューマンコミュニケーションやデザインなどにおける応用まで拡がっていた(3)。

本来、自然科学は新しい方法論や現象の 発見がシーズとなって、それをいかにして 実世界に応用するかという流れで進歩する が、感性情報処理はいささか趣を異にする。 むしろニーズが先行して実用技術が発展し、 それが学問の創成を促がすというプロセス で生まれてきた。その意味では、この基礎 的研究プロジェクトの前に、数多くの個別 応用技術の開発研究が、製品開発の分野で 有用性を認められ実用に供していた。

一方、さらに進んだ応用技術が通産省主

導の大型プロジェクトとして 1995 年に「ヒューマンメディア」というテーマで立ち上がっており、現在その評価が行われている。このプロジェクトは、当初幅広く9つのサブプロジェクトが企画されていたが、社会情勢の落ち込みから、3つの重要テーマに限って実施された⁽⁴⁾。

1] 次世代プラント用ヒューマンインタフェースの研究開発

代表者:溝口理一郎(大阪大学)

2] 感性エージェントとヒューマンメディアデータベースの研究開発 代表者:加藤俊一(電総研、現中央

大学)

3] 都市環境ヒューマンメディアの研究 開発 代表者: 竹村治雄(当時奈良先端 科学技術大学院大学、現大阪大学)

このプロジェクトの特徴は、知識に立脚して論理的理解を行う「知識メディア技術」、感情や感性をベースに認識や表現を行う「感性メディア技術」、物理的世界像の中で実体験する「仮想メディア技術」の3つを統合することを目的としたものである。個別には、この3つのメディア技術がそれぞれ上記1]、2]、3]の開発研究の核となっているが、いずれのプロジェクトも3つのメディア技術を整合的に含むことが要求されていた。

現実には、感性情報処理の実証研究は2]の感性エージェントのプロジェクトで実施されていたので、それに関して少し詳しく述べておく(5)。

「感性エージェントとヒューマンメデ

ィアデータベースの研究開発」では、共同作業による工業デザインを支援することを実問題として取り上げ、消費者や個々のデザイナーの感性的な特徴をモデル化する技術を核に、デザイナーと消費者が相互の感性的な表現を理解しながら、共同作業的に工業デザインを直感的かつ効率よく行えるようにする工業デザイン支援システム「感性工房」のプロトタイプシステムの構築を行うことを目的にしている。その主な要素技術としては、以下のものが取り組まれている。

- ①種々のマルチメディアコンテンツに対する感性のモデル化技術と、マルチメディアデータベースの内容検索技術(カラー写真、テクスチャ、絵画、都市景観、室内景観、3次元物体モデルをそれぞれ対象とした、例示による類似検索、イメージ語による感性検索など)。
- ②エージェントメカニズムを利用した自律 的な感性モデル化インタフェースと、イ ンターネットからの画像の感性検索への 利用技術。
- ③感性データベースシステムの構成法と動 的なデータ管理技術。
- ④仮想空間を利用した感性モデルの可視化 技術、および、直感的な共同作業空間の 構成法と工業デザイン支援システムのヒ ューマンインタフェースへの応用。
- ⑤カラー画像の高精度入力・提示技術と、 イメージ語による微妙な色調の制御技術。
- ⑥工業デザインにおけるカラーコーディネーション支援技術。

3. 最近のプロジェクト

日本学術振興会未来開拓事業の中で「感性的ヒューマンインタフェース」が 1999年から5カ年計画で現在進行中である。この企画は前出の科学研究費重点領域の延長線上に位置するものであるが、重点領域のように広い領域をカバーするものではなく、特定な分野に限定してトップダウン的にプロジェクトメーキングされたものである。ここでは、以下の4つのプロジェクトが走っている (6,7)。

1] 人間の内的知識と外的情報の統合的 な利用に関する認知科学的研究

リーダ:乾 敏郎(京大院情報)

2] マルチモーダル擬人化インタフェースとその感性基盤機能

リーダ:石塚 満(東大院工)

3] 情報知財の組織化と直感的アクセス の感性的インタフェース

リーダ:田中 譲(北大院工)

4] インタラクションによる相乗効果を 用いた感性創発世界の構築

リーダ:谷内田正彦(阪大院基礎工)

1]の「身体化による感性的認知」の研究プロジェクトでは、人間の対話行動を支えるメカニズムの解明やノンバーバルな情報伝達と感情や意図などの高次情報の抽出、さらにコミュニケーションインタフェースとしての眼、手、口の機能の解明を目的に、下記のような内容に取り組んでいる。

- ・ジェスチャ、パントタイムの機能とそれを支えるメカニズムの解明
- ・道具のイメージ化など創造性を支える基本機能のメカニズムの解明

・人の外的情報と内的情報の統合過程の解明

2]の「マルチモーダル擬人化インタフェース」では、マルチメディア時代のヒューマンインタフェースの新形態として、顔、姿を有し、音声会話能力を有する擬人化インタフェースを目指している。

- ・ユーザの感情を含む状態認知、感情モデル、適応的反応、感情に基づく自立的動作生成機能を含む擬人化インタフェース用の感性基盤機能の実現
- WWW情報空間の新しい接面となる擬人 化エージェントによる新しい情報として のマルチモーダルプレゼンテーションシ ステムの構築

3]の「感性的アクセスアーキテクチャ」では、知財流通メディアに存在する多様かつ膨大な知財を目的に応じて選び出し活用することを目的に、

- ・人間の直感的感性を最大限に活用する電子メディア情報知財の組織化とアクセス の技術の開発
- ・情報知財を流通・交換する場としてのメ ディア空間のアーキテクチャと、その開 発のためのアプリケーション・フレーム ワークの構築

4]の「感性創発世界の構築」では、仮想感性を持った複数のエージェントや多数の人間が参加する人工の共通の場「シナジェティック・フィールド」で、相互のインタラクションが引き起こす感性創発現象や感性増幅現象について検討している。

・遠隔地間デザイナーの協調によるシナジ

ェティック・インタフェース、社会教育 エデュティメント・電子コミュニティメ ディア、高齢者・障害者のコミュニティ 参加を支援するための身体・体感インタ フェースの実現

・身体センシング技術と生理モニタリング の要素技術の開発とインタラクションの ための「場」の構築。この「場」を介し たインタラクティブ創造活動を試行実験 しと「場」による感性情報の創発性の評 価

余談ながら、このプロジェクトを立ち上げる際に、英文でどのように表現するかと言うことが議論になった。研究推進委員は、すでにいくつかの国際会議やジャーナルで認知され始めた "KANSEI" を使いたいと主張したが、事務局の意向から英単語で表現することになり、結果 "Intuitive" が採用された。感性には直感 "Intuitive" と情緒 "Affective" の両面の機能が重要とされており、当初 "Affective" の方がヒューマンインタフェースの機能を表現できていいという意見もあったが、Intuitive に落ち着いた。上記3]4]のプロジェクトの対象としては、Intuitive でよかったのかもしれない。

1996 年以来継続進行中の「筑波大学感性評価構造モデル構築特別プロジェクト(プロジェクト長原田昭)」は、美術・デザイン作品などの実作品を遠隔地から鑑賞する実験を通して、鑑賞行動における人間の情緒的行動を含む心的状態の遷移に関するデータを解析することにより人間の作品鑑賞における心的・感性的態度モデルを明ら

かにすることを目的にユニークな取り組みが行われている⁽⁸⁾。

1998年に創設された日本感性工学会では、産官学を横断的に含む研究組織として40の研究部会を持ち、感性教育、感性デザイン、風土工学など幅広い活動を進めている(9)。

4. 最近のトピックス

ここで感性情報処理から生まれたアート& テクノロジーの試みを紹介する (10)。これは音楽の分野で、VR技術、感覚計測技術、生理計測技術を総合して、演奏者の意図をシステムが感知し、マルチメディアがバーチャルな音響・映像環境を生成するインタラクティブアートと呼ばれる新しい創作アートを目指したものである。楽器としては尺八をフィーチャーしたもので、多くのセンサを取り付けた尺八を特に「サイバー尺八」と呼び、次のような2つの目的を目指した前衛的プロジェクトであった。

- (1) 技術的な目的として、マルチメディア、 ヒューマンインタフェース、感性情報処 理の実地応用することで単に研究段階 では得られなかったさまざまな知見を 得る。
- (2) 芸術的な目的で、パフォーマーのジェスチャー・声・舞台条件に基づき、映像・音楽・照明などの各種メディアを協調的に動作させて、今までにない緊張感の高いアートを提示する。

このシステムは、センサ系、認知実行系、 プレゼンテーション系の3つの要素から成 り立っている。センサ系は尺八奏者の意図 を感知するものであり、音響ピッチ抽出センサ、タッチセンサ、曲げセンサ、角速度センサ、姿勢センサ、画像センサ、衝撃センサから、複数のセンサを組み合わせて利用するというマルチモーダルな構成を取ってくるサブシンボリック情報(例えば「深くおじぎをした」「大きく右肩をあげた」などを計測し、演奏家の意図を汲み取る。プレゼンテーション系において、認知実行系からの情報を受けて、シンセサイザー、エフェクタ、ビデオ、CG、照明などの実際操作が行われる。

ここでは、ライブ演奏によるコンポージ ング環境の有効性を実証するために、尺八 演奏家/作曲家志村哲らとともに、マルチ メディア作品「竹管の宇宙」の一連の作品 を制作している。

ダイナミックな動きを計測して、インタ ラクティブにアートパフォーマンスを支援 する事例としては、がある。これはイメー ジ情報科学研究所が新しい舞台表現を実験 するために、高安マリ子ダンスシアターと 共同で試みているもので、ダンサーに取り 付けたセンサによりジェスチャーを計測し、 映像・音楽・照明をインタラクティブにコ ントロールしている。マルチメディアおよ びコンピュータ技術を用いた舞台は他にも 多くあるが、その多くは時間制御を機械が 行い、人間がそれに合わせていた。音楽や ダンスのようなダイナミックな演出では、 真剣で息の合ったタイミングが重要であり、 機械によるプリセット的な制御では目的を 達することができない。ここではパフォー マー自身がメディア効果および時間を制御

しており、人間が主導権を持った極めて緊 張感の高い作品の創作システムとなってい る。

5. 感性データベース

感性を扱ったデータベースは、3つのカ テゴリーに分けられる。

- 1)感性関連の文献/機材などのデータベース
- 2) 芸術作品のデータベース
- 3)物理量/感覚量/心理量と感性語のデータベース

1) の例としては、1994年に通産省感性 産業研究会が編纂した感性データベースが 幅広い分野をカバーしている (11)。このデ ータベースは、感性の基礎となる感覚に焦 点を当て、感覚・感性に関する 16 の研究分 野における研究成果をコンパクトにまとめ たものである。またここでは、現在の開発 商品の中から感性商品と思われるものを収 集し、それらが感覚的にどのような価値を もっているのかについて分析し、可能な範 囲で感覚研究との結び付きをはかった。特 にこのデータベースにおいては「感覚価値」 という具体的な指標を基準として「感性商 品」にアプローチすることを試みている。 具体的には、視覚として光、形、空間、色、 聴覚として、音、音楽、音声、その他、味 覚、臭覚、体性感覚、皮膚、平衡、時間感 覚などで分類している。

収録されているデータは約300の感覚研究項目と関連文献情報、さらに約600の感性商品を集約し、感覚研究、感性商品開発における基礎的項目を収集している。

2) の例としては、音楽作品、美術作品な

どのデータベースのほか、楽器や衣装など 多くの関連資料が博物館的に整理されつつ ある。

3) の例としては、カラーコーディネートのための配色イメージスケール (カラーデザイン研究所) やHQLにおける人間感覚データベースなどが挙げられる。

感性情報を取り扱ったデータベースプロダクトとして、色彩と感性量の対応について扱った日本カラーデザイン研究所の一連の製品「カラーイメージスケール」「配色イメージワーク」が広く認められている (12, 13)。特にカラーイメージスケールに関しては、英語版も出版され国外でも評価されている。

色彩について、国民性の違いを過度に意識してきた読者の中には、このデータベースの感性情報が日本人と外国人の双方に通用することに意外性を感じる方が多いと思われるが、この書物が扱っている"Warm-Cool"と"Soft-Hard"空間上での感性特徴には国際的共通性が存在する。

「人間感覚計測応用技術」では、人間の感 覚を反映した快適な居住・職場環境や生活 製品等の設計を行うために、人間の諸感覚 を計測する技術や、快適性等を評価する指 標(物差し)を確立することを目的にデー タベース作りを行っている。前期では、人 間の感覚に影響を与える外的刺激を発生・ 制御するための模擬環境提示技術、刺激に より生起する生理変化を非侵襲で簡易かつ 定量的に捉える生理計測技術、人間の感覚 量と外的刺激や生理反応との関連を解明す る相関調査技術、等の開発を行い、

平成7年度からの第2期においては、開発

された各種生理・感覚計測技術を用いて、 人間感覚に関するデータを系統的に収集す るとともに、前述した人間感覚の指標化技 術、及びその指標を具体的事例に適用する 指標応用技術の確立を行っている。指標化 技術として、健康・安全を目指したストレ ス・疲労・覚醒の指標化、利便・快適を目 指した環境適合性と製品適合性の指標化を 具体的目標とした。

前述の文部省科学研究費重点領域プロジ ェクトのE班「感性デザインおよび感性データ ベースの研究」では、感性データベースは画 像データ等を含むマルチメディア・データ ベースの方向に向けて発展しているた(15)。 しかしなお依然として感性情報の扱いにつ いては未だモデリングが不十分のまま、ad hoc な試行の段階にとどまっていた。デー タベースにおける感性情報処理とは、具体 的には感性情報をまずモデル化することで ある。そのモデルが適切なものであれば、 モデルの枠内で効率よくキーを設定できる 感性データベースシステムが構築できる。 このシステムには、マルチメディア・デー タの分類・検索、また感性に応じたデータ の加工と提示、ユーザの意図、感性理解を 伴う協調的インタフェース等の内容が含ま れる。これら超大量かつ超多様な高次情報 である感性情報処理機能をもつ感性情報モ デリングと感性データベースを実現するた めの基本要素メカニズムに関する事例研究 として、仮面などの民族資料や、TVニュ ース映像などの動画コンテンツなどを対象 に支援ツールの開発を行っていた。

6. おわりに

製品や社会システムに贅沢な付加価値が 期待されていた 2000 年以前には、感性的手 法に大きな期待が寄せられていた。しかし ながら第1期を終えた段階で、感性情報に 対する新しい方法論が現れていない、知識 情報処理との違いが見えてこないといった 壁が立ちはだかり、実用的有用性が認めら れつつも、学問領域として認知される段階 に至っていない。

ただ、過去のこの種のプロジェクト評価においてデータベースの重要性が指摘されながら、データベース作成の困難さと、知識情報に比べて1桁も2桁も冗長な感性データベースシステムの実現性を疑問視する声もあった。この点に関しては、近年のハードウエアリソースの進歩には目を見張るものがあり、本格的な感性的データベースも近い将来、実現が期待されている。

参考文献

- 1) 原島 博監修、井口征士他:「感性情報処理」電子情報通信学会編ヒューマンコミュニケーション工学シリーズ、オーム社(1994)
- 2) 辻 三郎編:「感性の科学-感性情報処理へのアプローチー」、サイエンス社(1997)
- 3) 井口征士:「感性情報処理が目指すもの」 情報処理学会誌 Vol.35 No.9 pp.792-798 (1994)
- 4) 井口征士、橋本周司: ヒューマンメディアプロジェクトの概要、システム/制御/情報、Vol.42, No.5, pp.235-243 (1998)
- 5) http://www.tokyo.image-lab.or.jp/hm/
- 6) 感性的ヒューマンインタフェース公開

- シンポジウム報告書、2000/11/22 (2000)
- 7) システム制御情報学会「インタラクションによる感性的コミュニケーション特集号」、システム/制御/情報、Vol.45, No.6 (2001)
 - 8) 筑波大学感性評価構造モデル構築特 別プロジェクト研究報告集、(2000)
- 9) http://wwwsoc.nii.ac.jp/jske/main.htm10) 井口征士:「メディア技術を用いたインタラクションと芸術・福祉への応用」第28 回日本バイオフィードバック学会特別 講演(2000)
- 11) 感性産業研究会特別研究報告「感性データベース・概論データベース・データ 集」(1994年3月)
- 12) 小林重順:「カラーイメージスケール」日本カラーデザイン研究所、講談社 (1990)
- 13) 小林重順:「配色イメージワーク」、日本カラーデザイン研究所、講談社 (1995)
- 14) 「人間感覚計測応用技術プロジェクト (平成2年度10年度)研究のまとめ」社団 法人 人間生活工学研究センター(1999)
- 15) 「感性デザインおよび感性データベースの研究」平成6年度文部省重点領域研究「感性情報」E 班研究成果論文集(1995)

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

About Two Picture Series of The Tokaido 53 Stations, One Wood Prints by Hiroshige and The Other Oil Paintings by Presumed Kokan

> 荒木啓介(科学技術振興事業団、工博) ARAKI, KEISUKE 東京都千代田区四番町 5 - 3

Japan Science and Technology Corporation 5-3, Yonbancho Chiyoda-ku, Tokyo 102-8666 JAPAN 大畠洋一(資料提供、東海道研究家) OHATA, YOICHI(Tokaido Researcher) 横浜市保土ヶ谷区桜ヶ丘 1-10-13

1-10-13, Sakuragaoka, Hodogaya-ku, Yokohama-City JAPAN

[抄録] 歌川広重の世界的に有名な版画「東海道五十三次」は美術愛好家のみでなく、理科系の研究材料にも採り上げられているが、多くの謎も指摘されていた。最近構図や人物描写までよく類似している、司馬江漢の落款がある55枚の油絵が見つかったが、美術界は広重を真似た後世の贋作としている。その指摘点を再検討し構図を詳細に分析すると、多くの証拠から江漢の絵が広重版画の元絵であることは否定できず、この逆は不可能と分る。この前提に基づけば、広重画で従来から謎とされていた部分が解明できることを示す。

[Abstract] The Tokaido 53 Station Series, one of the world famous wood prints of Utagawa, Hiroshige is not only loved by art favorers but also studied or used by researchers or practitioners of science and technology. However some riddles have been long pointed out. Recently 55 oil paintings of Edo period with signs of Shiba, Kokan were found, 52 of which have many similarities in the composition and representation of the landscapes or the people drawn in them. Most Japanese artistic researchers or critics consider them as counterfeits by later painter(s) tracing Hiroshige's Ukiyoe printings. We minutely studied the reasons asserted by those critics and compared the compositions of corresponding pictures of two artists, and came to the following conclusions upon many evidences. Originality of Kokan's paintings can not be denied. Forgery of Kokan's paintings from Hiroshige's prints is impossible. Reversely, it can be said that Hiroshige referred to and modified Kokan's paintings as his Ukiyoe. According to these premises, such mysteries found in Hiroshige's prints can be clearly unriddled.

[キーワード] 東海道五十三次; 歌川広重、版画; 司馬江漢、油絵; 比較; 謎解き [Key-Words] Tokaido 53 Station Series; Utagawa Hiroshige Wood Printings; Shiba Kokan Oil Paintings; Comparison; Unriddling

1. 初めに

科学技術分野における専門分野間のディ スコミュニケーションの解消にこそ、デー タベース活用の意義があるのでは、の問題 意識から、その方法について検討してきた。 これには広く理系・文系の枠すらも越える 必要があろう。その中で偶々次のようなテ ーマに遭遇した。

有名な歌川広重の版画「東海道五十三次」 は、その美術的に優れた情景描写や構図で 広く世界中に知られ愛されてきたが、園芸 (小野) や土木建築 (坂井1-4) の専門家が江 戸期の樹木分布や景観を調べる研究に利用 したり、計算機画像処理の題材にしたり(小 沢1.2)、浮世絵を患者に見せる心理療法にも 利用する (白川) などが、科学技術振興事業 団の科学技術文献データベースから明らか になった。しかし理系の研究者は、広重は 実際には東海道を歩いてはおらず (鈴木)、 何らかの元絵を参考にして浮世絵風にアレ ンジしたのでは(浅野、稲垣、近藤、徳力)、の 美術界の議論や、広重の他の浮世絵には 次々と元絵と見られるものが見つかってい る (大久保) 事実には無頓着で、四季の変化 を取り入れた広重の優れた演出、と賛美し ている (小野)。

ところが 1992 年に岐阜の旧家から、司馬 江漢の落款がある 55 枚の油絵が発見され た。そのうちの52枚は広重の五十三次と偶 然の一致とは言えない類似性があり、江漢 絵から広重画を描くことは可能だがその逆 は先ず不可能であることが論じられている (對中1)が、美術界は大方これを無視する か、広重を模倣した後世の贋作、と決めつ けている。広重の東海道が自然科学・工学 の方面でも採り上げられているからには、 この問題を放置しておく訳には行かないだ ろう。しかしこの件は調べるほど微妙に錯 綜しており、そこに謎解きの面白さがある のも事実である。そこで浅学を省みず以下 のような調査・論考を行った。

2. 新発見の油絵東海道

2. 1 新発見油絵の特徴

司馬江漢は、1747年生まれ、1818年没、 生涯、少なくとも3度、春夏秋冬にわたり →東海道を旅行したことは記録により確か められる。一方、広重は1797年生まれで江 漢より 50 才若く、1858 年に没しているが、 「東海道」刊行の1833(天保4)年の前年 に東海道を旅したと伝えられるその春まで、 江戸の定火消しで江戸を離れられず、未だ あまり有名ではない絵師であった。

江漢と広重の年譜

1747 1788 1797 1799 1812 1818 1832 1833 1858 延享4 天明8 寛政9 寛政11 文化9 文政元 天保3 天保4 安政5 江漢: 出生 長崎行 近畿行 京都行 没 広重: 出生 旅行 五十三次 没

(注1) 江漢、長崎行:春-夏、近畿行:4-5月、京都行:2月立ち12月帰り

(注2) 広重、京都行(言い伝え): 1832(天保3)年7-8月のみ

新発見の油絵は、江戸と京都および53宿 ない類似性を持っているが、詳細に見ると 場の全 55 枚のうち 52 枚までが構図や描写 それぞれが微妙に異なっている。江漢油絵 された人物まで広重の版画と偶然とはいえ は遠近法や西洋の陰影法を駆使し、写実性 が高く曲線を多用し、立体感がある(横地)。 広重の方は平面的ではあるが人物がユーモ ラスに描かれ親しみやすいのが特徴である。

2. 2 美術界の見解

美術界・美術史家の反対意見は要約すると以下の通りである。

- (1) 画風が江漢のものとは異なる。江漢 が描くはずなし(成瀬不二雄氏、 2001/01/27の筆者との第1回電話にて)。 後世盛んな点描が見える(細野)。
- (2) 印譜(サイン)が違う。司馬峻の「馬」、 江漢の「漢」が違う。

絵の具に江漢の時代に無い新しいも のが使われている(クロムイエロー) (大久保純一氏、2001/01/27の筆者との第1 回目の電話にて)。

江漢の死後に普及したという白粉 「仙女香」の広告が「関」の本陣に見 えており、広重の模倣である決定的 な証拠である(稲垣。大久保純一氏、 2001/01/28の筆者との第2回目の電話に て)。

(3) 広重の五十三次は世界的に浮世絵の 価値基準となっており、元絵の存在 が正式に認められると大きな影響が ある。

2. 3 双方の方法論の違いについて

美術界の関係者とも接触し議論した際に 感じたことは、我々理科系との方法論の違いである。美術界の方法論の特徴を挙げる ならば、感性や経験に大きな比重を置き、 疑わしきは否定する傾向が強く、他からの 参入を嫌い、また美術商界と美術史学が未 分離といえる。

この第3の他からの参入を嫌う特徴は昔 からではなく、例えば筆者が絵具について 調べた「東洋絵具考」の著者、塩田力蔵氏 はその序文で「とかくの大革新は、却って 局外の好事家中より起こる例なれば、姑く概論として、自分一個の大局観を略述して 置いた」(塩田)、のような立場の専門家もかつては居たことを付記したい。

一方我々自然科学における方法論の特徴 は、以下のとおりであろう。

論理実証主義。絵画に適用すれば、素材・現場や時刻の実証。真実はいずれ誰かが明らかにする、の再現可能性、楽観主義。 疑わしきは保留し、誤謬により永久に失われる危険を回避する。新しい事態には新しい方法論で。

このような方法論の違いを念頭に置いて、協力可能性を探ることも、研究の動機の一つである。

3. 美術界への反論

3. 1 反論の項目

上記(1)については、あまり触れない。 ただ江漢は、日本画(狩野派)、浮世絵(鈴木春信)、漢画(南蘋派)、洋画(エッチングと油絵)を幅広く修得し、これらの様々な絵画を残しており、また**常に新しい画風を試みている**ことは「西洋画談」や佐賀の友人への手紙にも記されている。後年の作である「花鳥図」には点描の手法が見られ、さらに20番の「府中」には、江漢晩年の真筆「駿州柏原富士図」の麓にあると同様の、江漢に極めて特徴的な森林の形が描かれているのである。

(3)は学問とは無縁であり、美術界の内幕に容啄する意図はない。ただ未だ無名の広重が幕府から朝廷への「お馬進献の使節」に抜擢・同行を許されて夏に旅行し、出世作で最も有名な東海道を描き「蒲原」を雪景色としたのは、広重の天才のなせる技、と賛美するのは誉め過ぎであり、素人眼にも納得できない。が先ず(2)の諸点につき我々が調査した結果について述べ、さらに解説書等の論点も取り入れて江漢の絵と

広重の版画の大きな相違点を分析し、江漢 の絵によれば従来謎とされてきた再刻版の 理由や描いた場所が相当程度明らかになる ことを紹介する。

3. 2 美術界の見解への反論

(1) 印譜の違い

- ・真筆中にも、また急いで書いた自筆の 日記には、似ている「馬」や「漢」が ある(成瀬書より)。
- ・なお、印鑑は江漢の真筆「春宮図」のものと一致(對中2)。

(2) 絵の具、クロムイエローの問題

- クロムイエローは、1797年にフランス のヴォークラン(Vauquelin)が新しい 金属であるクロムを発見し、顔料史上 にエポックを作った(桑原利秀、安藤徳夫)。
- ・クロム: 1797 年、L.N. Vauquelin が Siberia 産のベニエン鉱 (紅鉛鉱) より 未知金属の酸化物として見出いだし、 この新金属が塩類として各種の色彩を 示すことから、ギリシャ語の $\chi \rho \omega \mu$ α (色) にちなんで命名された。(化学大辞典)
- ・ベニエン鉱 (紅鉛鉱): 鉛のクロム酸塩 鉱物。色にちなんでギリシャ語の κ ρ ο κ ο σ (saffron) に由来する。産状、 産地: 二次成鉱物としてリョクエン鉱、 ハクエン鉱、モリブデンエン鉱、デク ルワゾー石などと共生する。ソビエト Urals 地方 Beresov、ブラジル Minas Geraes の Goyabeira、アメリカ California 州 Inyo 郡 Darwin 鉱山。 組成、Pb(CrO4)、(中略) 暗トウ赤色、 オレンジ色、黄色、条コン: トウ黄色。 ダイヤモンド光沢ーガラス光沢。半透 明(以下略)。(化学大辞典)
- ・日本では古くから使われていた(成瀬不 二雄氏、2001/02/14の筆者との第2回電話に て)。

・古くからの天然黄色顔料である紅鉛鉱から新元素クロムが発見されたのが1797年、大量生産が20年後の江漢の没年である1818年(中右)。シカゴのマックローン研究所の鑑定でこれが分析されたとしても、江漢は生前に試作品を長崎経由で入手できた可能性があるし、輸入のベルリン青の使用は確認されている(成瀬 p372)。輸入の天然品を自分で工夫調合した可能性もあり、使用してあっても否定材料にはならない。

(3) 白粉「仙女香」の小さな広告札

寛政期(1789-1801)の歌舞伎名女形 3代目瀬川菊之丞の俳名、仙女にちなん で(高橋、花咲)、江戸の坂本氏が1807(文 化4)年頃から売り出した白粉(江戸博、 歌舞伎人名事典)。江漢が東海道を旅行した 1812(文化9)年頃に本陣に、上流階級 向けの広告札があっても矛盾しない。 1819(文政2)年に出版許可申請され、 1825(文政8)年に出版された「江戸買物独案内」で仙女香が確認できる(花咲)からといって、それ以前には存在しない ということにはならないし、浮世絵で「仙 女香」それ自体が広告として大きく描かれた後期のものとはあくまで扱いが異なるのである。

4. 構図からの江漢絵が先である決定的な証拠

01 **番**「日本橋」初版では人物が少ないが、 再版で江漢絵からインドネシア人等を追加 してにぎやかにした。

逆の、再版の人物から魚屋を除いて贋作者が江漢絵を描いたの説(稲垣)には無理がある。インドネシア人の出所や**魚屋を除く理由**が不明。

03番「川崎」初版の不自然な船頭の頭の向きや富士の見え方を再版で江漢に合わせ

て小修正。広重オリジナルならこんな修正 は敢えてしない。

04番「神奈川」再版では屋根の形を江漢に合わせ、また広重の時点では進捗していた埋め立ての様子を実景に合わせて追加した。もし逆に江漢絵が後世の贋作というなら、贋作者はなぜ日本橋や川崎では再版を真似し、神奈川や後述の小田原では実景に近い再版ではなく初版を真似たのか、説明が必要。

07番「藤沢」の遊行寺は、1831 (天保2)年12月27日に焼失し、再建は5年後 (大畠、藤沢市史)なので、広重は翌年の夏にも、五十三次刊行の前後にも見ることはできなかった。広重が江漢絵を真似た決定的な証拠の一つ。また江漢は構図の関係で寺を高い位置に描いた(江漢には他にもこの種の変更の事例がある)が、広重はこれを理解せず、実在しない階段を描き足してしまった。逆に江漢は実在の寺門を小さく描いているが、広重は見落としている。これも広重画から江漢絵を真似たのではない、決定的な証拠の一つである。

10番「小田原」初版では江漢が描いた相 模川と背景の大山を真似たが、箱根の山と 酒匂川が小田原により近くふさわしいとの 指摘を受けて実景に従い再刻しなおした。 しかし江漢は大山の麓にはあるはずのない 城は描いていないのに、広重はいずれも城 を描き加えている。ただし位置が異なり、 小田原城はもっと左の海寄りのはず。

11 **番**「箱根」 芦ノ湖の湖岸、駒ヶ岳の位置から、江漢が誇張して描いた「塔ケ島」が分かる。

12番「三島」三島神社の鳥居と灯籠の位置関係は江漢絵が正しく、また西方にある満月から早朝と分かる。

13番「沼津」東海道沿いの狩野川ではなく、支流の黄瀬川と分かり、満月の方向から早朝と判明。ここの描写手法の点描は、

江漢真筆の「花鳥図」にもある。

16番「蒲原」は雪景色であるが、夏に旅をした広重がこれを見事な冬の風景に描けるであろうか。江漢なら冬季にも東海道を歩いているので可能である。

17番「由井」富士の右側は広重のような 水平線はなく、愛鷹山や伊豆半島が見える はず。広重は実見していないので、江漢絵 の右先の部分を海にしてしまった。

32番「荒井」の双胴船を、広重には意味 不明であったのか、ムシロで隠した。広重 画から江漢絵を正しく描くのは先ず不可能 (對中1)。

5. 江漢絵の構図や描き方の巧みさ、立体感、 特徴

12番「三島」、13番「沼津」の陰影法。

15番「吉原」、44番「四日市」の巧みに曲線を用いた構図。

08番「平塚」、09番「大磯」、22番「岡部」、26番「日坂」などの立体感。

20番「府中」の山麓に江漢独特な森林の描写。

6. 状况証拠

- (1)1977(寛政 11)年刊行の「西洋画談」 末尾に「春波楼画譜」を刊行し、西洋 画法を解説する予定、との記述。
 - 一、春波楼画譜 西洋画伝、和蘭奇工、 天文地理の部分をして、観る者をして 煩わしからざらしむ 近刻 とあり、 「規矩術(定規とコンパス)を以て日 本の山水遠近の法を描く」画法を解説 する予定であった(未完、成瀬書より)。
- (2)佐賀藩士、山領主馬宛の1813(文化10)年6月12日付け書簡に
 - 一、此度和蘭奇巧の書を京都三条通富 の小路西の入、吉田新兵衛板元にて出 来申候。其中へ日本勝景色富士皆蘭法 の写真の法にて描申候(中野書より)。

7. 江漢絵の不思議な点と広重画との関係

全体の想定される関係

______ 江漢絵 =⇒ 広重初版 =⇒ 広重再刻版

- (1) 東海道名所図会(1797年刊行)にも 同様の構図の絵がある。江漢も広重 も江戸から遠い場所については、名 所図会を参照した。
- (2) 広重も後には小田原あたりまでは実 景を確認に出かけ、修正もしている。
- (3)縦に長く誇張された構図(藤沢、戸 塚、関)、実景の誇張や移動は江漢に よくある。

8. 広重画の問題点

- (1)季節の多様さは、想像のみで描ける
- (2) 実景とのズレや矛盾、さまざまな再 版の存在を説明しきれるか。
- (3)他の浮世絵には次々と見つかる元絵 ナルと主張し続けられるか。

9. 文化人・科学者江漢の謎と魅力

- ・天文学を研究、地動説を紹介(成瀬)。
- ・広い人脈 老中、松平定信、 仙台・紀州・岡山・佐賀藩主、 平賀源内 (芳賀)。
- ・通常会えない人物とも会見 ---間宮林蔵、大黒屋光太夫 (芳賀)。

10. 結論

(1) 江漢油絵は広重版画の元絵と考えて 矛盾しない。逆に江漢油絵を前提に

- すればこれまでの広重版画のさまざ まな疑問や二大傑作「蒲原」「庄野」 の謎も無理なく説明できる。
- (2) 広重の版画から全ての江漢油絵を贋 作することは不可能。
- (3)常識的に考えても、版画から実景に 近い高度な油絵多数を贋作するのは 不自然。

11.コメント

- (1) 広重の名アレンジャーとしての力量 は依然として高く評価される。
- (2) 江漢の広重への影響は美術界で既に 認識されつつある(成瀬 p378)。
- (3)問題の先送りは理系にも影響を及ぼ し、美術界の姿勢が問われよう。
- があるが、東海道は例外で、オリジ (4) 江漢作と見られる油絵は日本の重要 な文化遺産として大切にすべきでは ないか。

12.データベース構築について

- (1)自然科学-人文科学の多方面の協力 が望ましい。
- (2) データベースとは本来、周辺領域・ 関連情報を探すためのもの。 そのための総合的なあるいは特に関 連性を考慮したデータベースの構築 も必要。分類やキーワードの調整、 相互乗り入れ。共同研究、相互協力。

13.謝辞

快諾された對中如雲氏に感謝します。江漢 の黎明・25」から借用した。 真筆の絵は、集英社、日本美術絵画全集

本発表に際し、江漢絵の計算機コピーを 25「司馬江漢」、学習研究社、「日本絵画 広重版画は對中氏の著書等から転載した。

14. 参照図書・文献

浅野秀剛 浮世絵を読む・5 広重、朝日新聞社、1998

穴吹史士 名画日本史、朝日新聞社、2000.2001

稲垣進一 広重・東海道五十三次 - 成立とその謎、河出書房新社、2000

江戸博 江戸東京博物館パンフレット、今月の見どころ、1998.2月号、No.8

大久保純一 浮世絵を読む・5 広重、朝日新聞社、1998

大畠洋一 江漢図の風景と人物、藝術倶楽部、Vol.20, p8-11, 1997

岡村千曳 紅毛文化史、創元社、1953

小沢一雅1 浮世絵合成システムPICSの構成、情報処理学会研究報告、93,105, p45-52,1993

小沢一雅2 絵画生成のモデル化に関する一考察、情報処理学会論文誌、37,3,p.355-362,1966

広重・五十三次に見るみどころ演出法、造園雑誌、52,5, p241-246, 1989 小野良平

化学大辞典 共立出版、1993 (第34刷)

歌舞伎人名事典 野島寿三郎 編、日外アソシエーツ、1988

桑原利秀、安藤徳夫、顔料及び絵具、共立全書、1972

近藤市太郎 世界名画全集、広重と東海道、本文と別冊、平凡社、1960

坂井猛1 広重の浮世絵風景画にみる景観分類に関する研究、日本建築学会計画系論文 集、461号、p165-174, 1994

広重の浮世絵風景画に描かれた河川景観の構図に関する一考察、日本建築学 坂井猛2 会計画系論文集、482号、p155-163, 1996

広重の浮世絵風景画にみる樹木の構図的機能に関する考察、日本建築学会計 坂井猛3 画系論文集、507号、p165-171, 1998

浮世絵風景画に描かれた宿場の景観構成に関する考察、日本建築学会計画系 坂井猛 4 論文集、509号、p149-156, 1998

東洋絵具考 アトリエ社、1942 塩田力蔵

白川佳代子 母子浮世絵にみる「甘え」 心身症の新しい視点、日本小児医学会報、No.17, p99-103, 1999

鈴木重三 広重、日本経済新聞社、1970

對中如雲1 広重「東海道五十三次」の秘密、祥伝社、1995

對中如雲 2 司馬江漢「東海道五十参次」は真筆、藝術俱楽部、Vol.11, p20-24, 1995

高橋雅夫 化粧ものがたり、雄山閣、1997

徳力富吉郎 東海道-昔と今、第2版、保育社、1972

中野好夫 司馬江漢考、新潮社、1986

中右瑛 「江漢・東海道五十参次」の怪、藝術倶楽部、Vol.20、p15-18、1997

成瀬不二雄 司馬江漢-生涯と画業、本文編、作品編、八坂書房、1995

芳賀徹 日本の名著 22、中央公論社、1971

花咲一男 坂本氏・仙女香図説、化粧文化、ポーラ化粧品、1981,6月

藤沢市史 ふじさわわがまちのあゆみ、藤沢市

細野正信1 司馬江漢、江戸洋風画の悲劇的先駆者、読売選書、1974

細野正信 2 対談、司馬江漢の「東海道五十参次」、藝術倶楽部、Vol.6. p3-5, 1995

堀晃明 広重の東海道五十三次旅景色、人文社、1997

横地清 遠近法で見る浮世絵、三省堂、1995

O1 日本橋

江漢



広重初版

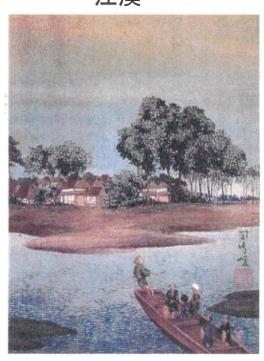


広重再版



03 川崎

江漢



広重初版



広重再版



04 神奈川

江漢



広重初版



広重再版

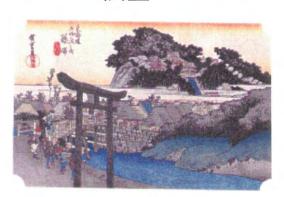


07 藤沢

江漢



広重

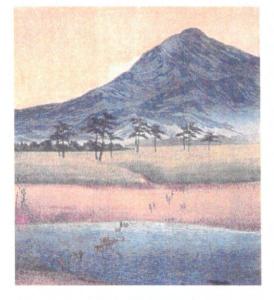


広重には実在しない石段、 遊行寺は天保2年に焼失

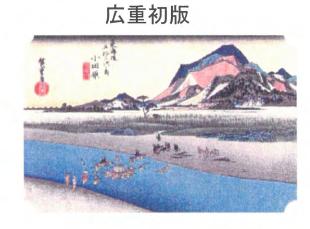
江漢には実在の黒門が鳥居左 柱中央に見える

10 小田原

江漢、大山と相模川



酒匂川と箱根山



広重再版



江漢



12 三島



鳥居と灯籠の位置関係は江漢が 正しい



13 沼津

点描手法の使用例

江漢

江漢•花鳥図





16 蒲原

江漢



広重



広重の最高傑作の一つ 冬にも旅した江漢、夏に旅の広重

17 由井

江漢





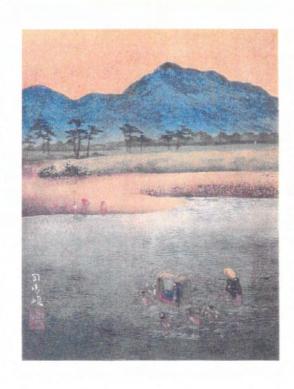


実際は右に水平線はない

江漢

20 府中







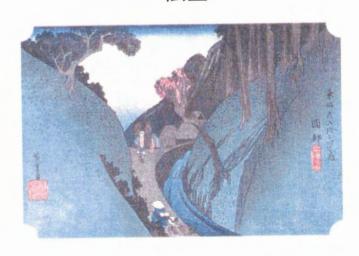
江漢独自の山麓の森林の形が酷似

22 岡部

江漢、立体感、写実性



広重



江漢、緩い奥行きの坂

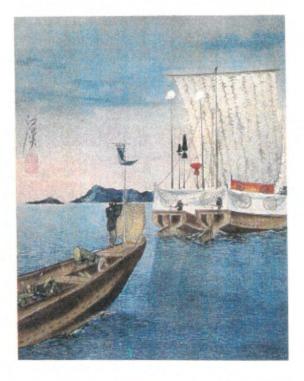




32 荒井

江漢、双胴船







ムシロで隠す

江漢

36 御油



広重



よりユーモラスだ

44 四日市

江漢、曲線の多用



広重、直線使用



46 庄野

江漢



広重



広重の最高傑作の一つ

48 関

江漢



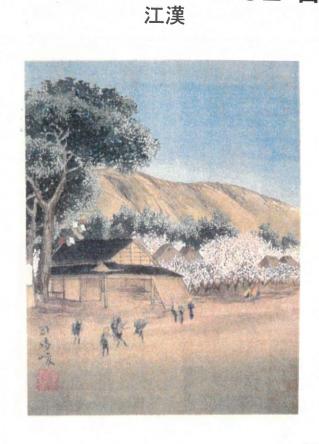




屋内に小さな仙女香の札 江漢の時代に既に存在

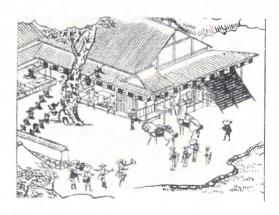
52 石部

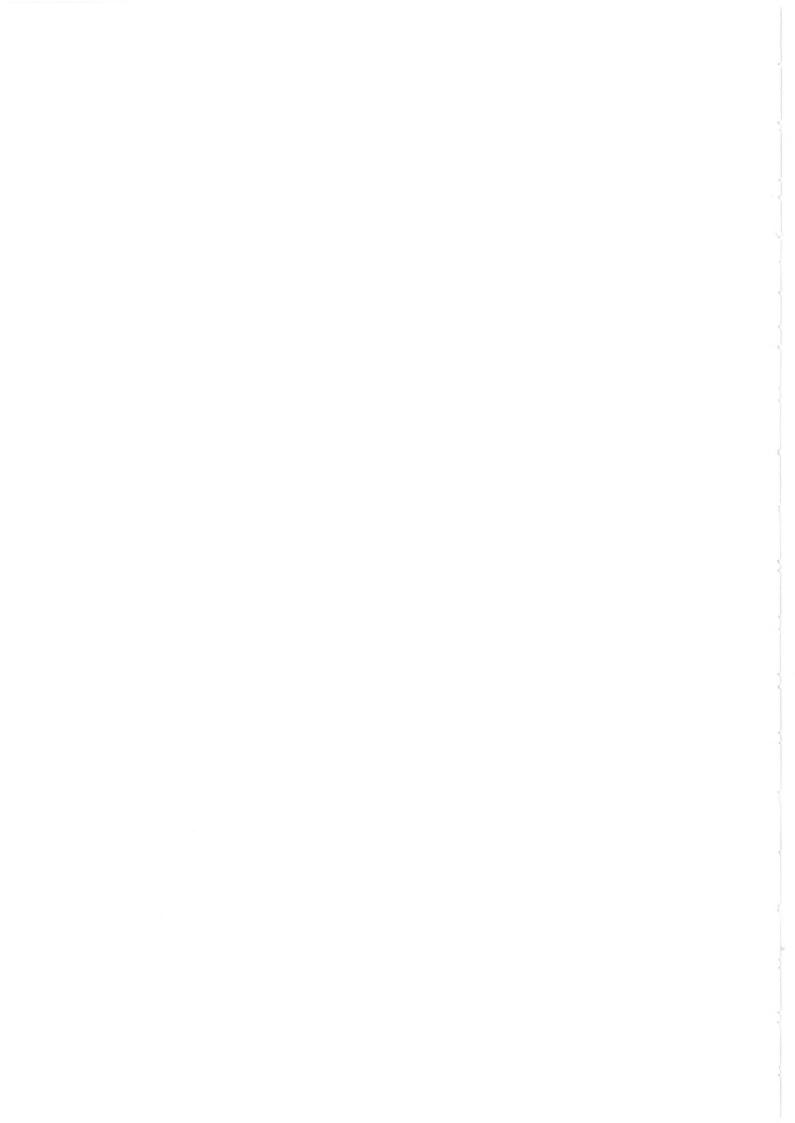






東海道名所図会





古文書文字列に対するキャラクタスポッティング Character Spotting for Histrical Document Character String

橋本 智広 梅田 三千雄
Tomohiro HASHIMOTO, Michio UMEDA

大阪電気通信大学大学院 〒 572-8530 大阪府寝屋川市初町 18-8 Graduate school,Osaka Electro-Communication University, 18-8 Hatsu-cho,Neyagawa-shi,Osaka 572-8530,Japan

あらまし:

本論文では、古文書文字列を対象として、古文書特有のつづけ字や食い込みに対処するために、認識過程を導入した文字切り出し手法を提案する。また、文字列から任意に指定する文字のみを抽出するキャラクタスポイッティングに注目し、指定文字のみを認識して抽出する手法を提案する。まず、初期の文字切り出しとして、文字列に対するラベリング処理により、連結成分ごとに領域分割する。この領域を矩形で囲み、高さや横幅といった矩形情報を基に個々の文字パターンを切り出す。次に、切り出された文字パターンに対し、個別文字認識する。この認識結果から切り出し失敗矩形を検出し、それに対し認識処理結果に基づく再文字切り出しをする。これによって、認識結果を考慮した切り出しが可能となり、最適な位置での切り出しが期待できる。その後、抽出対象とする文字のニューラルネットワークを用いて各文字パターンの誤差を算出し、抽出条件に適合する文字パターンをスポッティング結果として選出する。本手法において、「天保郷帳」を例にとった 615 個の古文書文字列を対象とした抽出実験では、抽出対象を 5 文字として、再文字切り出しを付加することにより、導入前には 87.58%であった平均抽出率が、導入後は 94.22%に向上した。

Summary:

This paper proposes a character segmentation and spotting method of histrical documents. In the segmentation method, the result of character recognition process is utilized to cope with the cursive scripts and the mutual encroachment of characters which are peculiar to the historical documents. In the spotting method, the previously assigned characters are only extracted from the characters string. As an early segmentation, the characters string pattern is divided into the same connected component by using the labelling processing. The area composed of the same component is surrounded with a rectangle and each character pattern is segmented each other by using the shape of rectangle such as height and width. Next, the individual character recognition is applied to the segmented pattern. From the recognition result the rectangle failed in the segmentation is picked up and the resegmentation is applied to the string contains this rectangle. Therefore, it is expected that the string is divided at the best position. On the other hand the neural network which corresponds to the previously assigned character is prepared. The error between input and output of the network applied to the segmented pattern is calculated and the pattern which satisfies the condition is extracted as a spotting result. From the extraction experiment applied to 615 characters strings, the correct extraction rate of 94.22% was obtained to 5 assigned characters by using the resegmentation process, but the rate was 87.58% without the resegmentation process.

キーワード:

文字認識,キャラクタスポッティング,文字切り出し,自己想起型ニューラルネットワーク,古文書

Keywords:

character recognition, character spotting, character segmentation, autoassociative neural network, histrical document

1 はじめに

手書き文字認識に関する研究は、様々な研究機関 で試みられており、数多くの認識手法が提案され、そ の技術は実用の段階にある. 一方, 人文学研究分野 では、古文書を対象とした OCR の実現を目指し、古 文書に対する認識手法の提案が期待されている[1].

現在、古文書画像データベースの構築においては、 史料の解読、文字データ入力に長時間の作業を必要 とする. この作業を自動化できれば、飛躍的に作業 時間を短縮でき、多量の古文書史料を短時間で、効 率よくデータベース化できる. そこで, 古文書を対 象とした OCR の研究が進められている [2]-[4].

しかし、古文書を認識対象とすると、文字のパター ン数に制限があるため、認識に使用する辞書作成に おいて、十分なデータ採取が困難となる. そのため、 認識対象となる文字が限定されてしまう. また, 手 書き文字認識のように認識に使用する辞書作成にお いて、不足するデータを新たに作成するといったこ とが不可能である. 従って、限られたデータの範囲 内で、古文書独自の認識手法を新たに考案する必要 がある. また古文書では、認識手法とともに、文字列 から個々の文字パターンに切り出す文字切り出しが での切り出しが期待できる. そして、抽出対象とな れに伴い高い認識結果を得にくいなどの問題もある.

本論文では、文字列を認識するだけではなく、文 字パターンをスポットする. 字列から任意に指定する文字を抽出する技法である キャラクタスポッティングに注目し、古文書文字列 から指定する文字がどの程度抽出できるか検討する から任意に指定する文字のみを抽出する手法を提案 する. 同時に、古文書特有のつづけ字や文字の食い 込み等を考慮した,認識過程を導入した文字切り出 し手法を提案する.

文字列から個々の文字パターンを切り出すために, まず初期文字切り出しとして、文字パターンにおけ る連結成分を囲む矩形情報を基に統合, 分割を繰り返 特徴抽出に加重方向指数ヒストグラム特徴 [5] を用い た. また認識処理には, 柔軟な情報処理と高い汎化能 各文字パターンに対し個別文字認識する. ここでは, 力を持ち,人間の学習過程をモデル化した,ニューラ ルネットワーク(以下 NN と略す)を利用する. ここ NN を形成する. 初期文字切り出しでは、誤って切 NN[4] を用いた. さらに、認識処理結果から初期文 字切り出しにおける切り出し失敗矩形を検出し、認 入した再文字切り出しをする. 最終的に、対象文字 識処理を繰り返し実行することにより、最適な位置 値処理によって対象文字を選出する.

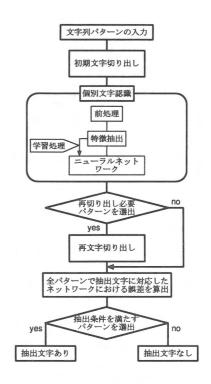


図 1: 処理の流れ

重要となる. しかし、古文書特有のつづけ字や文字 る文字に対応した NN を用意し、切り出された各文 の食い込み等から正確な文字切り出しが困難で、そ 字パターンを入力して誤差を算出する.この誤差を 閾値処理して抽出対象文字を選出し、対象となる文

> 抽出実験として、「天保郷帳」を例にとり、文字列 とともに、文字列からの各文字パターンに対する切 り出し率と抽出率の関係について検討する.

システムの概要

本システムの処理手順を図1に示す。まず、対象文 し、切り出し候補を得る. 個別文字認識においては、 字列において文字を個別パターン化するために初期 文字切り出しをする。これによって、切り出された 前処理、特徴抽出を施し、さらに学習処理によって では、古文書文字に対して有効とされる自己想起型 り出された文字パターンが存在することが多い、そ こで、該当する文字パターンに対し、認識過程を導 識処理結果に基づく再文字切り出しをする. この認 の NN で各文字パターンにおける誤差を算出し, 閾

初期文字切り出し

文字列から指定する文字のみを抽出するためには. 個別に文字を切り出す必要がある。初期文字切り出 しは、文字パターンにおける連結成分の高さ (heigh) や横幅 (width), 面積 (area) 等の情報に着目した切 り出し手法である.

処理手順は, まず, ラベリングにより文字パター ンの連結成分に対する外接矩形を求める. 外接矩形 とは、連結成分を囲む長方形のことである. この時 点では、「八」や「三」のような複数の領域から構成 される文字は、それぞれ連結性がないために各領域 は独立している. そこで, 各領域をグルーピングす ることで一つの領域とするために統合処理をする.

縦書きであるため、先に述べた「八」などの文字は 横方向の矩形同士をグルーピングすることで一文字 となる可能性が高い. 図2に統合処理における各段 階での条件を示す. 第一段階として, 図 2(a) のよう に対象矩形の高さ内に別の矩形があるとき統合する. 第二段階は,対象矩形と別の矩形が図 2(b) の位置関 係にあり、別の矩形の高さ比率が40%以上のとき統 矩形に対し、分割処理をする. 合する. この二つの条件は、矩形の高さに着目した り合う面積に着目して統合する. 第三段階では, 図 2(c) に示すように、対象矩形が別の矩形の幅よりも %以上で統合する. さらに, 第四段階として, 図 2(d) に示すような別の矩形が対象矩形の横幅よりも大き く、重なりあっている別の矩形の面積比率が50%以の平均をwidth.aveとする. 上のときに統合する. この第四段階に関しては、次

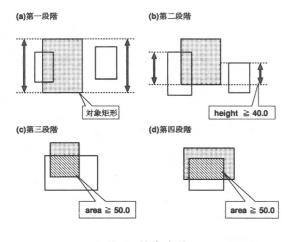


図 2: 統合方法

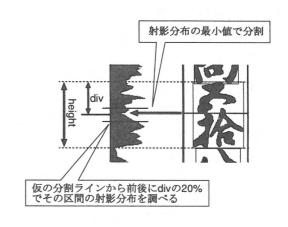


図 3: 分割方法

統合方法は、4段階に分けて行う.対象文字列が に説明する分割処理において、一回目の分割処理が 終了した時点で適用する. これらの統合条件の他に, 「三」を強制的に統合する条件も与えた.

> 次に, 統合処理によってグルーピングされた矩形 の中には誤って統合したものや、初期の段階でつづ け字や食い込みにより、二文字以上が同一領域であ るとみなされた矩形が存在する. そこで, これらの

分割方法は、二度の処理で完了する。まず、それ 統合方法であるが、第三段階以降は矩形同士の重な ぞれの分割対象矩形を選出するための条件を次のよ うに与える. 一回目の条件は, 矩形内に3個以上の 複数文字が存在すると考えられる矩形を選出するも 小さく, 重なり合っている対象矩形の面積比率が50 のであり, 二回目は, 矩形内に2文字程度が存在す ると考えられる矩形を選出する条件である.ここで, 全矩形の面積の平均を area.ave とし、全矩形の横幅

- 分割対象矩形条件:一回目
 - $height \ge width \times 1.2$
 - area > area.ave
- 分割対象矩形条件:二回目
 - width > height
 - width \geq width.ave
 - $area \ge area.ave$

これらの条件にあてはまる矩形を図3に示すように分 割する. まず初めに, height $\div width$ を求め, この値 を仮の矩形内文字数 (moji) とする. そして, height÷ moji で仮の分割ライン (div) を決定する. さらに, 仮 の分割ラインから前後に div の 20 %の大きさをとり、 その区間の射影分布を求める. そして, 最小値とな

るところを最終的に分割ラインとする. 二回目も同 様に分割する.

個別文字認識

文字切り出しによって切り出された各々の文字パ ターンに対し、個別字認識する.まず、前処理とし て, 孤立点除去, 大きさの正規化, スムージングに より文字パターンの均一化を図る. 次に,特徴抽出 を行う. ここでは、数多く提案されている特徴抽出 法の中から, 比較的高い認識率が期待できる加重方 向指数ヒストグラム特徴により特徴量を得る. さら に、得られた特徴量をもとに、NN を形成するため の学習処理を行う.

前処理 4.1

切り出した個々の文字パターンは大きさにばらつ きがある. そのため、各文字パターンの均一化を図 るために前処理を施す.まず,画像に含まれている 雑音を除去する孤立点除去、ばらつきのある大きさ を均一にする大きさの正規化, さらに, 大きさの正 規化によって凹凸の激しくなった文字の輪郭部を平 滑化するスムージングを施す. 図4に前処理前後の 文字パターンを示す.

4.2 加重方向指数ヒストグラム特徴

文字の輪郭部に着目した特徴抽出法である加重方 向指数ヒストグラム特徴は、次のようにして抽出す る. まず、文字パターンに対し輪郭線追跡を行いな がら、輪郭部に属する各画素に対し16の方向指数を 算出する. 方向指数の算出では、図5に示すように、 注目画素と連結している前の画素から注目画素をみ た方向指数と、注目画素から後の画素をみた方向指 タリングする. そして、ここでは8×8領域×4方向 数から注目画素の方向指数を算出する.

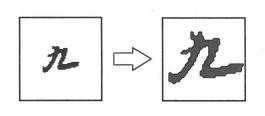


図 4: 前処理

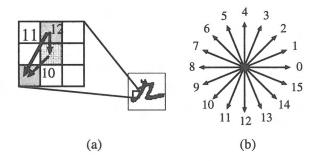


図 5: 方向指数の算出

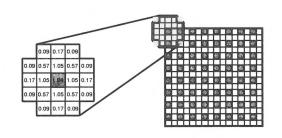


図 6: ガウスフィルタ

この例では、前の画素から注目画素を見た方向指 数は12となり、注目画素から後の画素をみた方向指 数は10となる. そこで, 両者の方向指数の平均をと ることで注目画素の方向指数11と算出する.

そして, 方向指数を算出後, 各方向指数に対して方 向圧縮する、まず、奇数方向に対し重み付けし、それ に対して前後の偶数方向を足しこむ処理により、16 方向から8方向へと圧縮する. さらに, 反対方向を 同一視することにより、4方向へと圧縮する、

次に, 領域圧縮として, 一定領域でヒストグラム を求める. さらに、図6に示すような、局所的な位 置をぼかす働きをもつガウスフィルタを用いて領域 圧縮する。ガウスフィルタは画素一つおきにフィル から成る 256 次元の特徴量を得る.

自己想起型ニューラルネットワーク 4.3

認識には、その柔軟で、かつ高い汎化能力から、文 字認識において利用されることが多い NN を使用す る. ここでは、特に NN の中でも古文書に対して有 効とされる自己想起型 NN を用いることとした.

自己想起型 NN は入力層と出力層のユニット数が 等しく, 入力パターンそのものを理想出力とするネッ トワークである. 従って, 教師信号には入力パターンそのものを与える. ここで, NN 形成における学習とは, 実際の出力と望ましい出力との差, つまり誤差を小さくするように重みを変更していくことを意味する. 学習には, バックプロパゲーション法 (BP 法) を用いた.

これは,出力層のニューロン値 O_k と理想的な出力である教師信号 T_k との二乗誤差を式 (1) により求め,その誤差が最小となるように出力層と中間層,中間層と入力層間のシナプスの重みを変更するものである.

$$e = \sum_{k} (T_k - O_k)^2 \tag{1}$$

図 7 にニューロンモデルを示す. ここで, 各ユニットに対する入力の総和を

$$net_i = \sum_j W_{ij} O_j \tag{2}$$

とする.次に、そのユニットの出力 O_i をとすると、

$$O_i = f(net_i) \tag{3}$$

となる. この出力関数には式 (4) のシグモイド関数 を用いる.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

重みの変化量は学習係数 η とすると,

$$\Delta W = \eta \left(-\frac{\partial E_p}{\partial W} \right) \tag{5}$$

で求めることができる. ここで, 出力層と中間層の 重みの変化量は,

$$-\frac{\partial E_p}{\partial W} = (T_k - O_k)f'(net_k)O_j \tag{6}$$

$$\delta_k = (T_k - O_k)f'(net_k) \tag{7}$$

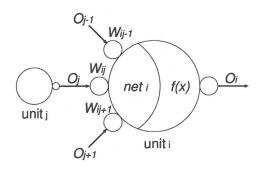


図 7: ニューロンモデル

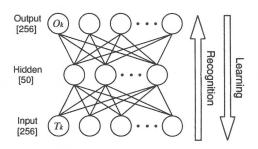


図 8: 自己想起型ニューラルネットワーク

 $\Delta W_{kj} = \eta \delta_k O_k \tag{8}$

で与えられる. また, 中間層と入力層との間では,

$$-\frac{\partial E_p}{\partial W} = f'(net_j) \sum_k (\delta_k W_{kj}) O_i$$
 (9)

$$\delta_j = f'(net_j) \sum_k (\delta_k W_{kj}) \tag{10}$$

とすることにより変化量は、

$$\Delta W_{ij} = \eta \delta_j O_i \tag{11}$$

となる.

このような重みの変化量を用いて、出力層と教師信号の間の誤差が小さくなるように、各ユニット間の重みを変更していくことで NN を形成する.

 (5) 図8に、ここで使用したネットワーク構成を示す. 各層のユニット数は、入力層と出力層が 256個、中間層は50個とした.また、このネットワークはカテゴリごとに形成するため、カテゴリの変化に容易に対応できる特徴がある.そのため、同時に他の文字
 (6) の影響を受けない学習が可能となる.

5 認識過程を導入した再文字切り 出し

初期文字切り出しでは、文字パターンの連結成分における矩形情報のみを利用するため、完全に文字が個々のパターンに切り出されていないことが多くある.ここで提案する切り出し手法は、古文書特有のつづけ字や文字の食い込みなどを考慮し、認識処理結果に基づき再文字切り出しすることによって、高精度な切り出しを実現する.

本手法では、次に示す条件を基に再切り出し候補 矩形を求める.

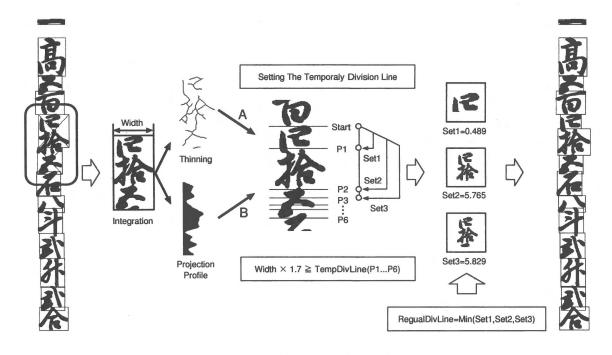


図 10: 分割ラインの決定方法

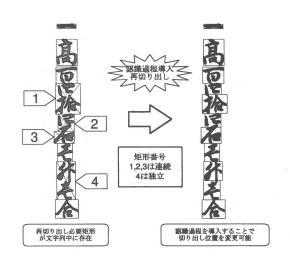


図 9: 再切り出し必要矩形

条件 α: NN における二乗誤差 > 2.5

• 条件 β: NN における二乗誤差 > 1.5

挙げられる.

- 再切り出しが必要とされた矩形の上下矩形のど ちらも再切り出し不要とされる場合(再切り出 し必要矩形が独立)
- 再切り出しが必要とされた矩形の上下矩形のい ずれかが再切り出し必要とされる場合(再切り 出し必要矩形が連続)

図9に再切り出し必要とされた矩形例を示す. こ こでは、初期文字切り出しの段階で得られた切り出 し候補に対し、条件 α を満たしていない矩形4個に 目印となる"×"が印されている. それぞれを矩形 番号 1,2,3,4 とすると、1,2,3 においては連続してい ることがわかる.一方,4は上下の矩形には目印が付 いていないため独立している. そこで, 再切り出し 必要矩形に対し再文字切り出しを適用する.

ここで, 連続している矩形の位置関係から互いに 少しでも離れている場合は独立しているとする. そ れ以外のときは、強制的に連続している矩形を統合 する. 図9では、1,2,3の矩形の位置関係は離れてい まず、切り出した各矩形に対して個別文字認識す ないために全て統合する. そして、統合後の矩形に る. このとき、NN における誤差をもとに、条件 α を 対して個別文字認識する. このとき、条件 α を満た 満たす矩形は誤って切り出されたとして再切り出し さなければ統合することによって適切に切り出され 必要な矩形とする.ここで、全矩形において再切り た矩形とする.しかし、条件 α を満たす矩形に対し 出しが必要とされる矩形の位置関係は、次のことが ては再分割する. これ以降の処理は再切り出し必要 矩形が独立している場合も同様にする.

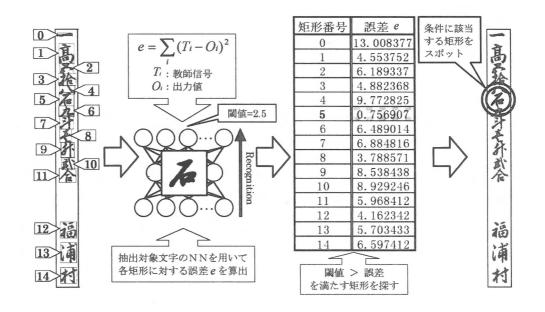


図 11: スポッティング処理

図 10 に再分割処理手順を示す。まず、対象となる ら矩形幅の 1.7 倍以内にある仮の分割ラインについて 影分布を求め、分布の平滑化を図るとともに、射影 値を一定の割合で小さくする.

ここで、あらかじめ対象矩形において、収縮法[7] により文字の線幅を求ておく. 求めた線幅を閾値と して、射影値が閾値より小さい部分を仮の分割ライ 定する.しかし、これでは、仮の分割ラインが対象 様に該当しているものがあれば仮統合し、認識する. とする矩形によっては複数になることが考えられる. そこで、仮の分割ラインが連続している場合はその 中間点を最終的な仮の分割ラインとする. また独立 している場合はそのままとする. これにより仮の分 割ラインを最小限に抑えることができる.

仮の分割ラインから分割ラインを決定するとき,矩 形の最上部から各々の仮の分割ラインまでの領域に 対し個別文字認識する. その際, 認識における誤差 が最小となるところを分割ラインとする.

まず、文字列中の一文字あたりのサイズには限界 があるため、対象矩形の高さに着目して、最上部か

矩形に対し, Hilditch の細線化処理 [6] によって矩形 誤差をそれぞれで求める. 図 10 では, 仮の分割ライ 内の文字パターンを線幅1で表し、横方向の射影分 ンが7本ある. そこで矩形幅 width の1.7 倍以内に 布を求める. このとき、射影値1となるところは文字 ある分割ラインを調べると、Start から P3 までが該 同士が食い込んでいないところである. そこで、射 当する. そのため、Start を基準として P1,P2,P3 の 影値1となる部分をピックアップし、これを仮の分 分割ラインまでの領域をそれぞれ Set1,Set2,Set3 と 割ライン A とする. しかし、これだけでは文字の食 して認識する. そして、認識における誤差が最小とな い込みに対応できない。そこで、原画像における射 るところを分割ラインとするため、Start から P1 ま での領域が一つの文字パターンであるといえる. そ れ以降は基準を分割ラインとした地点, すなわち P1 を Start として以後同様の処理で分割ラインを決定

最後に全矩形に対して個別文字認識し、条件βに ンBとする. 以上の処理により仮の分割ラインを決 該当する矩形を見つける. そして, 上下の矩形で同 その際、条件 α を満たさないならば統合する。そし て、最終切り出し候補を得る.

文字列からの指定文字抽出

文字列から切り出しによって得られた各文字パター ンにおいて, 指定する文字を選出するためには, 対 象文字に対応したネットワークを用意する必要があ る. これを利用して各文字パターンにおける二乗誤 差を算出し、この誤差から対象文字を選出する.抽 出文字の選出には閾値処理を適用する.これにより,

文字列中に対象文字が複数ある場合でも全て抽出す に示す、なお、各表の項目内容(内訳)を表1に示 ることが可能となる.しかし、別の文字を抽出する ことも考えられるため、各文字に対する適切な閾値 を決定する必要がある.

図11に抽出までの一連の処理の流れを示す。まず 初めに, 入力文字列に対し文字切り出しで得られた 各文字パターンに矩形番号を付ける. そして, 特徴 抽出で得られた特徴量を基に抽出対象文字の NN を 当矩形の文字パターンを対象文字として選出する.

しているため、図11中の表から誤差が2.5より小さの文字の最適な閾値となる. な矩形を探すと、矩形番号5が条件を満たす. そし て,この矩形を対象文字として選出する.

7 抽出実験

保郷帳」に収められている相模国に該当する、当時 下する.これは、対象文字が抽出されるのと同時に、 の各村における石高を示した文字列615個を用いた. 別の文字も抽出するためである. これらはイメージスキャナで解像度 500dpi で採取し、 一文字列あたりのサイズは 1140×100 pixel である. に、各対象文字での切り出し率と抽出率を表 7 に示 図 12 に文字列例を示す. 各文字列は前半に石高, 続 す. また, それぞれ最適な閾値もあわせて示す. これ いて該当する村が記されている. また, 対象文字を 文字列から任意に 100 パターン選出し、ニューラル ネットワーク形成のための学習パターンとして使用を得ている「村」に関しては、抽出率も同様に高く する. ここでは、文字切り出しの段階で比較的文字 なることがわかる. また, 石高部において後に続く パターンが正確に外接矩形で囲まれている(切り出 されている) ものを選出した. なお, 学習回数は200 回とした.

実験は, 文字列に含まれる文字のうち, 石高表記 に用いられる「石、斗、升、合」と「村」の5文字を対 象文字とした. そこで, 対象文字のみを正確に抽出 するためには, 各々の対象文字において最適な誤差 の閾値を決定する必要がある. ここでは、閾値を 6 段階に変化させて抽出率の変化を調べた. また, 閾 値の設定に伴い, 正確に対象文字のみを抽出する場 合や、別の文字を誤って抽出する場合などが考えら れる. そのため、閾値の変化により文字列から指定 する文字がどの程度抽出可能かについて実験した.

各対象文字における抽出実験結果を表2から表6

しておく、ここで、抽出率、切り出し率は

対象文字のみ抽出した数 対象文字を含む文字列総数 × 100 抽出率 =

対象文字が切り出されている数 × 100 切り出し率 = 対象文字を含む文字列総数

と定義した. 抽出率とは, 対象文字を含む全文字列 用意し、これを用いて各文字パターンに対する誤差 e において、対象文字のみを抽出したときの文字列数 を算出する.ここで、対象文字ごとに閾値を設定し、 の割合であり、切り出し率とは、対象文字を含む全 条件 閾値>誤差 を満たす矩形を探す. 最終的に、該 文字列において, 文字切り出しの段階で, 指定する 文字パターンが正確に外接矩形で囲まれている文字 仮に、文字列より「石」を抽出しようとする。ま 列数の割合とした。さらに、各対象文字の閾値に対 ず,「石」に対する NN を用いて各文字パターンにお する抽出率の変化を図 13 に示す. このときの各対象 ける誤差を算出する.このとき, 閾値を 2.5 と設定 文字において, 最高の抽出率を得るときの閾値がそ

各文字の抽出率の変化をみると、閾値が小さい段 階では抽出率も低いが、閾値を大きくするにつれて 抽出率は高くなる. これは、対象文字が適切に切り 出せていない矩形に対し、抽出条件が満たされてい ないためである. また, 正確に切り出せている矩形 に対しては低い段階でも正確に抽出できる. 一方、 閾 実験用データとして、江戸時代に書かれた書物「天 値を大きくすることで、あるところから抽出率は低

> さらに、切り出し率と抽出率の関係を調べるため より、切り出し率と抽出率は大きく関係するといえ る. その中でも, 切り出しの段階で高い切り出し率

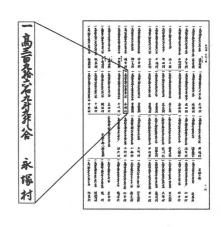


図 12: 使用データ

表 1: 結果表の内訳

	20 20 11 11 11 11 11			
	A	対象のみスポットした		
	В	対象と別の文字をスポットした		
内訳	C	別の文字のみスポットした		
	D	対象と別の文字をスポットした 別の文字のみスポットした どの文字もスポットしなかった		

表 6: 「村」に対する抽出結果

		1 4 3		, - ,	
閾値	A	В	C	D	抽出率 (%)
1.0	560	0	0	31	94.75
1.5	580	0	0	11	98.14
2.0	584	0	0	7	98.82
2.25	581	4	0	6	98.31
2.38	573	12	0	6	96.95
2.5	564	21	0	6	95.43

表 2: 「石」に対する抽出結果

_			1	4 7 - 3	/ 0 11	
	閾値	A	В	C	D	抽出率 (%)
	1.00	464	0	0	151	75.45
	1.50	540	0	0	75	87.80
	2.00	561	0	0	54	91.22
	2.50	572	2	0	41	93.01
	2.75	572	6	0	37	93.01
_	3.00	563	16	2	34	91.54
_						

表 7: 総合評価結果

対象文字	切り出し率 (%)	抽出率 (%)	閾値
石	89.92	93.01	2.50
斗	88.69	94.71	2.13
升	82.99	91.12	2.50
合	93.64	93.46	2.50
村	97.80	98.82	2.00

表 3: 「斗」に対する抽出結果

		20.	, -]	(~ V.)	1 01	
	閾値	A	В	С	D	抽出率 (%)
•	1.00	448	0	0	100	81.75
	1.50	497	0	0	51	90.69
	2.00	516	5	0	27	94.16
	2.13	519	7	1	21	94.71
	2.25	515	16	2	15	93.98
	2.50	491	47	0	10	89.60

表 4: 「升」に対する抽出結果

	24 10	. / 1 7	1-71	1 0 11	
閾値	A	В	С	D	抽出率 (%)
1.00	341	0	0	188	64.46
1.50	436	0	0	93	82.42
2.00	468	0	0	61	88.47
2.50	482	9	0	38	91.12
2.63	475	18	3	33	89.79
2.75	463	29	3	34	87.52

表 5: 「合」に対する抽出結果

	20.		1-11	1 0 11	41441
閾値	A	В	C	D	抽出率 (%)
1.00	346	0	0	189	64.67
1.50	453	0	0	82	84.67
2.00	489	0	0	46	91.40
2.50	500	8	0	27	93.46
2.75	495	13	1	26	92.52
3.00	486	24	1	24	90.84

文字が少ない「合」は、他の文字と比べると、切り出し率は高くなっている。一方、「升」は切り出し率、抽出率ともに低くなった。これは、「升」は前後の文字と重なっている場合が多いことや、文字パターンの射影分布が平均していることで、切り出し位置の推定に失敗したことが切り出し率の低下につながったといえる。

切り出しに失敗した文字列例を図14に示す.該当する文字列は、対象文字の外接矩形内に別の文字が存在しているため、正しく囲まれていないことがわかる.このような文字パターンは切り出し失敗とした.

次に, 再文字切り出しの効果を確かめるために, 導

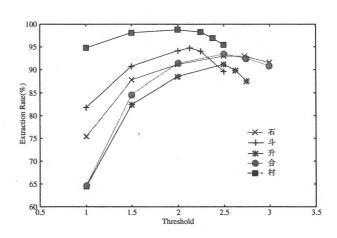


図 13: 各対象文字の閾値に対する抽出率の変化

表 8: 切り出し過程の違いによる抽出率の変化

対象文字	認識導入前(%)	認識導入後 (%)
石	84.23	93.01
斗	87.59	94.71
升	80.72	91.12
合	86.54	93.46
村	98.82	98.82

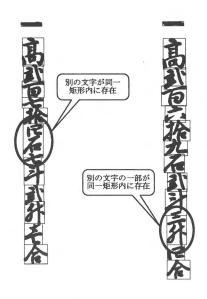


図 14: 切り出し失敗文字列

入前と導入後の各対象文字の切り出し率と抽出率を 比較する.表8にその結果を示す.これより,文字 列において,再文字切り出しを適用する部分は石高 部のみであるため,「村」に関しては変化がみられな い.しかし,その他の文字では導入前と比べると,導 入後の抽出率共は高い値を示すことが見受けられる. これは,切り出しの段階で認識過程を導入したこと の効果であるといえよう.

8 おわりに

本論文では、つづけ字や食い込み等が原因で、前後の文字が互いに影響しあう古文書文字列に対し、これらを考慮した文字切り出し手法を提案し、任意に指定した文字のみをいかに文字列から抽出できるかについて検討した。

その結果,提案手法では抽出対象文字の平均抽出率は94.22%,平均切り出し率は90.66%が得られた.これより,文字切り出しの段階で認識過程を導入し

たことが、切り出し率の向上つながったことがわかる。また、これに伴って高い抽出率が得られたといえよう。

キャラクタスポッティングでは、個々の文字を正確に認識するだけではなく、いかに適切に文字列から個々の文字を切り出すかという前段階が重要であるとわかった。今後は、毛筆書体の文字列認識を進めるにあたって、キー文字の抽出技術を向上させ、認識の高精度化につなげる検討を加える必要がある。また、他の書体の古文書にも応用できるように検討しなければならない。

参考文献

- [1] 山田奨治 "古文書 OCR 研究の現在"挑戦古文書 OCR, 人文学と情報処理, No.18, pp.2-5, 1998.
- [2] 日置慎治,上原邦彦,川口洋 "「宗門改帳」に記録された年齢表記の認識 "挑戦古文書 OCR,人文学と情報処理,No.18,pp.35-42,1998.
- [3] 和泉勇治,加藤寧,根元義章,山田奨治,柴山守,川口洋 "ニューラルネットワークを用いた古文書個別文字認識に関する一検討"情報処理学会研究報告, Vol.2000, No.8, pp.9-15, 2000.
- [4] 橋本智広,横田宏,梅田三千雄"自己想起型ニューラルネットワークによる古文書文字認識"電気関係学会関西連合大会論文誌, G13-14, 2000.
- [5] 鶴岡信治, 栗田正徳, 栗田昌徳, 原田智夫, 木村文隆, 三宅康二 "加重方向指数ヒストグラム法による手書き漢字・ひらがな認識"電子情報通信学会論文誌, Vol.J70-D-II, No.7, pp.1390-1397, 1987.
- [6] 手塚慶一,北橋忠宏,小川秀夫"ディジタル画像 処理工学",pp.139-142,日刊工業新聞社,1985.
- [7] 加藤寧, 根元義章 "ストローク情報に基づく手書郵便宛先の切り出しと認識"画像ラボ, Vol.8, No.8, pp42-45, 1997.

日商簿記検定問題の電子化

Electric hyper-text for the bookkeeping licensing examination of the JCCI

福田 宏,小津稚加子 Hiroshi FUKUDA and Chikako OZU

静岡県立大学経営情報学部 〒422-8526 静岡市谷田 52-1

School of Administration and Informatics, University of Shizuoka, 52-1 Yada, Shizuoka 422-8526

あらまし:日本商工会議所の簿記検定試験問題を電子化することによって、社会的ニーズの高い簿記学習を支援する e-learningシステムを開発した.

Summary: We develop the e-learning system for the bookkeeping by converting the JCCI (Japan chamber of commerce and industry) examination to the electric hyper-text. Studying bookkeeping technique is the gateway to the business world. Especially JCCI examination is regarded as the most qualified license. キーワード: 簿記,日本商工会議所,検定

キーワード:簿記,日本商工会議所,検定試験問題,電子化,e-learning,e-book, HTML.

Keywords: bookkeeping, JCCI (The Japan Chamber of Commerce and Industry), licensing examination, e-learning, e-book, HTML.

1 はじめに

筆者の一人(小津)は、静岡県立大学経

営情報学部で簿記・財務会計領域の教育・ 研究に従事してきた経験から、高校および 大学学部必修科目において理科系科目を重 点的に履修した大学生,大学院生,社会人 大学院生のための簿記教育, および IT (Information Technology)時代における効 率的な簿記学習システム開発の必要性を痛 感してきた. 前者は経営学部, 商学部(以 下,単に経営学部という)における簿記教 育ではみられなかった傾向に関連がある. つまり,経営情報学部生は基礎教育で,経 営学部学生と比較して理科系科目を多く履 修しているために計算能力には優れている ものの、経済現象と結びついた経済取引を イメージできない. また簿記学習に不可欠 な単純な反復練習を嫌う傾向にある.

また、日本の会計制度が国際会計基準に 調和化する方向で抜本的に改革されるなか、 その基礎教育としての簿記教育は、会計プロフェッショナルを目指す学生、再教育を 希望する社会人のみならず、将来企業に就 職しようとするすべての学生にとっても必 要不可欠である.企業会計の知識がパブリック・セクターにも摘要される必要性が指摘されるなか,簿記は,自治体,病院といった非営利組織の経営に携わる人すべてが身につけるべき教養である.つまり,社会的にきわめてニーズが高い領域である.

本研究の目的は、ITを使った効率的な簿記学習のための e-learning システムの開発である。上に述べたように簿記学習システムは社会的にきわめてニーズが高いため、すでに多くのシステムが存在する。先行研究[1-2]を検討した結果、著者らは独自の簿

神記検	€ - Microsoft Internet Explorer				<u> </u>
)編集化)表示公 お気に入	***************************************	***************************************		#3
2 * m)	- 0 0 0 0 0 0	y Pruz@je	C¥desktop¥kinouchi_CDR¥kentr	ei¥3-91.frm.html	<u> </u>
第1問	記(制限時間:全間で2時間) (20点) 取引について仕訳しなさい。 対	ただし、勘定科目は次の中 当度預金 未収商品 買掛金 位性 定定完和益を 上げ、た。 割引料 ¥27,000 差別((1株の額面金額: ¥50,00 のた検列欄(購入日:平月	受取手形 前私会 建物、金 達物、通信如果計額 支払手数料 固定資産売却指 取った内金 ¥50,000 を差3 、た手取金を当座預金とした 200)を1様につき ¥75,000、 物等2月1日、取得原価: ¥3	売 掛 金 立 備	料 ¥18,000 を含めて月末 額法、耐用年数:5年、残
5. 建	物について固定資産税 ¥125,				~ Ch U/co
later > 800	f to \$ \ \ \.				· · · · · · · · · · · · · · · · · · ·
	The state of the s	が			
あなたは 3級 第1間	/100点) 総得点 過去が この問題は過去に解いていません 第91回解答用紙] /20点) 採点				
あなたは 3級 第1問	たの問題は過去に解いていません 第91回解答用紙] /20点)採点	在 訳	D-101	45	
5なたは 3級 51間	にの問題は過去に解いていません 第91回解答用紙】	andanana da ana ana ana ana ana ana ana	(資方料目	金額	
5 年 3 年 6 1 間	たの問題は過去に解いていません 第91回解答用紙】 /20点)採点 借方料目	在 訳	資方科目	3	
846 81 181	だの問題は過去に解いていません 第91回解答用紙】 /20点)採点 	在 訳	() () () () () () () () () () () () () (•	
5なたは 3級 馬1間 1	第91回解答用紙 /20点) 採点 (音方科目	在 訳	· · · · · · · · · · · · · · · · · · ·		
あなたは 3級 第1問 1 2 3	第91回解答用級] /20点)採点	在 訳	(京方科目) (京方子 (方方子 (

図1 電子化された日商簿記検定問題

記教育体系をシステムに組み込むのではな く、既に存在する日本商工会議所の簿記検 定試験問題(以下,日商簿記検定問題とい う) [3-4]を電子化することによってシステ ムを構築することとした. その理由は、定 評のある検定試験問題を使うことで、質の 高い学習効果が自動的に期待できること、 社会的に最も認知された資格検定の準備も 同時に行えること、そして、それゆえ普遍 的なシステムを構築できるからである. ま た、インターネットへの接続は可能ではあ るが、システムの必須条件とはしないこと とした. このような電子化の結果, パソコ ンまたはインターネット上での学習が可能 になるので、いつでも、どこでも好きなと きに学習でき、また、同じ問題を反復学習 できるので、伝統的な簿記学習と組み合わ せて、授業時間以外の副教材としての高い 利用価値が期待できる.

2 電子化された日商簿記検定問題

日商簿記検定問題は1級,2級,3級,4級の4段階に分かれている.1級が最も難易度が高い.本研究で対象とするのは,このうち高校および大学学部必修科目で扱う内容に相当する2級及び3級である.これらの検定試験は年に数回全国各地の商工会議所で開催され,2001年2月25日で97回目を迎える.

日商簿記検定問題の電子化は、パソコン 等のプラットホームに依存しないように、 また、ブラウザで閲覧することができイン ターネットとの親和性の高い HTML(Hyper Text Makeup Language)に よっておこなう、HTMLの中には、 JavaScript 等のスクリプト言語を埋め込む こともできるので、これによって、同時に インタラクティブな電子化も達成できる.

日商簿記検定問題は問題用紙と解答用紙から成る.問題用紙は紙媒体に印刷されたものを単にHTML化する.解答用紙は解答欄を含んだ「表」である.解答欄には2種類あり、複数選択肢の中から選択して記入する選択解答欄と、計算した数値を記入する数値記入解答欄である.解答用紙をHTML化する際には、ブラウザで解答を選択あるいは入力できるようにする.また、採点基準通りに採点を行い得点を表示し、さらに間違った個所を指摘するようにする.

図1に1998年2月の第91回3級の日商 簿記検定問題を HTML 化し,ブラウザ, Microsoft Internet Explorer 6.0で表示した例を示す。ブラウザ画面は上下に2分割され、上半に問題、下半に解答欄が表示される。画面は紙媒体の問題・解答用紙と同じようにデザインされている。問題・解答用紙のブラウザでの表示方法には、この他、2つのウインドウを用いる方法、縦に2分割する方法等が考えられるが、2つのウインドウを用いるとアクティブウインドウの切り換え操作が予想以上に面倒であること、解答用紙の「表」が横方向の一覧性を必要とすることから図のような画面の利用方法を採用した。

3 解答用紙の機能

電子化された解答用紙の機能を詳しく述べる.まず、選択解答欄は、HTMLのいわゆるセレクトボックス・オブジェクトで実装する.セレクトボックス・オブジェクト

は図2に示すように選択肢をプルダウン形式で表示する.

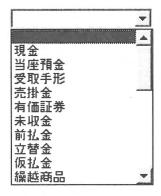


図2 セレクトボックス

次に、数値記入解答欄は、図3に示すような HTML のいわゆるテキストボックス・オブジェクトで実装する。テキストボックスにはキーボードから自由に文字列を入力することができる。

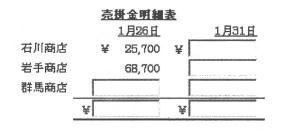


図3 テキストボックス

正誤の指摘及び採点は、HTML のボタン・オブジェクトを押すことによって、JavaScript 言語で記述された正誤判定関数と採点関数を呼び出しておこなう. 本研究で扱う日商簿記検定問題の2級、3級の場合、正解の多様性がないので、解答の正誤判定は、単に解答欄に選択・記入された解答が正答と一致しているか否かによっておこなう事ができる. すなわち、正誤判定関

数は、選択解答欄については選択が正答と一致しているかどうかの判定、数値記入解答欄については、記入された数値が正解と数値として等しいか否かを判定する。ただし、例外として、図1のように、表の一つのセルに複数の選択解答欄が含まれている場合には、セルの中での選択解答欄と数値記入解答欄の対にたいして正解の順序は問われない。

正誤の判定結果は、誤りについてのみ、 図4に示すように、選択解答欄については ×印で、数値記入解答欄については!印で 指摘する.

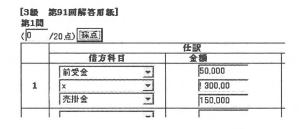


図4 誤答の表示

日商簿記検定問題の2級,3級の問題は,100点満点で5題の問題から成る.各問題の配点は解答用紙に図4のように記されている.この配点はさらに問題内の解答欄のグループ(部分集合)に対する得点に分解される.この分解と採点基準も問題ごとに異なるから正解と同様 HTML に埋め込んでおき採点関数で利用する.採点関数は,問題ごとの得点,及び総得点を計算表示する.

ところで、HTMLでは、スクリプト言語 によって上のような処理を行うことはでき るが、安全性の観点からローカル・ファイ ルへの書き込み操作は行えないようになっ ている. したがって、HTMLで電子化され た解答用紙では、原理的に解答欄の状態を ローカルに保存(保持)する事ができない.

ローカル・ファイルへの書き込みの唯一の例外として、4KB以下のクッキーと呼ばれるテキスト・ファイルを作る機能がある.そこで、電子化された解答用紙では、大きな記憶容量を必要としない、問題ごとの得点と総得点、解答日時のみをクッキーに保存し、図1に示すように、過去に同じ問題を解いた時の得点、および、その一覧を表示できる簡易データベース機能を実装した.

電子化が HTML でおこなわれたにもかかわらず、インターネットに接続されない、本格的なデータベース・サーバーを利用できない環境においても、本システムではこのような簡易データベース機能を利用することができる。なお、現在のシステムにはデータベース・サーバーへの接続機能は実装されていない。

4 電子化システム

日商簿記検定問題の電子化のうち,問題 用紙の電子化は通常の印刷文書のテキスト 化と同じように行うことができる.我々が 行った手順と使用した機器は以下の通りで ある.

- (1.1) 問題用紙をイメージスキャナーで スキャンする.
- (1.2) 市販の OCR (Optical Character Recognition) ソフトでスキャンし た画像をテキスト化する.
- (1.3) テキストを Microsoft WORD 2000 (以下単に WORD) に読み込み,

手作業で元の問題文に近いフォーマットに整形する. また, OCR の変換ミスを修正する.

(1.4) WORD で HTML 形式に変換する. ただし、再修正が可能なように WORD 形式(*.doc)の文書ファイル も保存しておく. この二重の保存は WORD の標準機能に無いので WORD に搭載されているマクロ言 語である VBA (Visual BASIC for Applications)によって行う.

一方,解答用紙の電子化では,前節に述べた,選択解答欄,数値記入解答欄,解答の正誤判定機能,採点機能,簡易データベース機能を埋め込まなければならない.これらを HTML と JavaScript でゼロから記述することが困難であることは言うまでもない.しかし,逆に,解答用紙をスキャンするだけで,これらの機能を自動的に組み込むシステムを開発することもまた,困難である.

我々は、解答用紙が「表」であること、WORD の作表機能が優れていること、および、WORD のユーザーが大変多いことから、WORD を利用して、電子化解答用紙を能率的に作成するシステムを開発することとした。 開発した WORD による電子化解答用紙作成システムは以下のようである.

- (2.1) 表は WORD で作成する. 簿記で使われる表の形式は決まっているので、VBA によって表の雛型を自動的に作る.
- (2.2) 正答以外の文字は表の中にそのま ま書く.

- (2.3) 選択肢は表の前に二重の中括弧で 囲んでカンマ区切りで、{{現金,当座 預金,...}}のように書く.
- (2.4) 選択解答欄は二重の大括弧で正答 を囲んで[[仕入]]のように書く.
- (2.5) 数値記入解答欄は二重の小括弧で 正答を囲んで((30,000))のように書 く.
- (2.6) もとの解答用紙と同じように問題番号の右に小括弧でくくって(20点)のように配点を書く.この点数が採点関数に読み込まれ採点に使われる.
- (2.7) 借方と貸方が明確な表に対しては, 貸方科目と貸方金額の欄を範囲指 定して,フォントの色を緑色に変え る.これは,先に述べた正解の多様 性の例外に対処するためである.
- (2.8) 表を仕訳問題,清算表とそれ以外に 分類して,表の上に二重のイコール で囲って分類を書く.すなわち,仕 訳問題に対しては==仕訳問題==, 清算表に対しては==清算表==,そ の他の表には分類は書かない.
- (2.9) 清算表に対しては、勘定科目の前に「※」を付ける.その他の表については、採点対象の二重括弧の前に「※」をつける.採点対象は公表されていないのでどこでもかまわない.配点を割り切れる数だけ「※」をつける.仕訳表については「※」は不要.

図5にこの規則で書いた仕訳表の例を示す.

{{現金,当座預金,受取手形,売掛金,有価証券, 手形貸付金,繰越商品,前払金,仮払金,未収金, 立替金,備品,支払手形,買掛金,手形借入金, 前受金,借受金,未払金,預り金,備品減価償却 累計額,資本金,売上金,売上,仕入,給料,租税 公課,減価償却費,受取手数料,支払手数料,受 取利息,支払手形,固定資産売却益,固定資産 売却損}}

第1問(20点)

==仕訳問題==

		仕 訳		
	借方科目	金額	貸方科目	
1			[[前払金]]	
	[[仕入]]	((300,000))	[[受取手形]]	
			[[買掛金]]	

図5 仕訳表の例

以上の規則にしたがって WORD ファイルを完成したら、問題用紙と同じように WORD 形式と HTML 形式の両方で保存する.

保存された HTML ファイルは、C++で書かれた変換プログラムによって、JavaScript の組み込み、選択解答欄、数値記入解答欄、採点ボタン、簡易データベース機能の追加が行われ、インタラクティブな HTML ファイルに変換される。ここで、このような変換プログラムは文字列処理、ファイル処理の得意な perl で書かれることが多いが、perl インタプリタは WORD のプラットホームである Microsoft Windowsには標準で用意されていないので、我々はコンパイル可能な C++を用いて、WORD と同じ条件で動作する実行可能プログラムを

作成した.

WORD を使うことのできる数名の学部 学生に対して、本システムの使用説明書を 渡し、簡単な説明を行った後、解答用紙の 電子化を依頼したところ、ほとんどの学生 が数時間で1回分の電子化された解答用紙 を作成できたとのことであった.

以上のようにして表1に示す3級33回, 2級20回の日商簿記検定問題を電子化した.

表 1 電子化済日商簿記検定問題

年度	月	3級	2級
1999	11	第93回	第 93 回
	6	第 92 回	第 92 回
1998	2	第91回	第91回
	11	第90回	第90回
	6	第89回	第89回
1997	2	第 88 回	第 88 回
	11	第87回	第87回
	6	第 86 回	第 86 回
1996	11	第 85 回	第 85 回
	6	第84回	第 84 回
1995	11	第83回	第 83 回
	6	第 82 回	第 82 回
1994	11	第81回	第81回
	6	第80回	第 80 回
1993	11	第 79 回	第 79 回
	6	第 78 回	第 78 回
1992	11	第77回	第77回
	6	第 76 回	第 76 回
1991	11	第75回	第 75 回
	6	第74回	第74回
1990	11	第73回	
	6	第72回	

1989	11	第71回	
	6	第70回	
1988	11	第 69 回	
	6	第 68 回	
1987	11	第 67 回	
	6	第 66 回	
1986	11	第65回	
	6	第 64 回	
1985	11	第63回	
	6	第 62 回	
1984	11	第61回	

5 今後の課題

本研究では、社会的ニーズの高い簿記学習を支援するe-learningシステムとして、日本商工会議所の簿記検定試験問題を電子化するシステムを開発した.このようなアプローチをした理由は、日本商工会議所簿記検定試験問題が日本の簿記教育において最も信頼されている検定試験だからであり、それゆえ普遍的なシステムを構築できると考えたからである.電子化されたシステムは、結果的には、日商簿記検定問題のデータベースとみなすこともでき、また近年注目を集めているe-bookとみなすこともできる.

本システムの教育システムとしての効果 測定はまだ行っていないが、静岡県立大学 経営情報学部で簿記を履修している何人か の学生に使ってもらい、自由記述の簡単な アンケートを実施した結果、紙媒体の問題 集と比較すると、一覧性は悪いが、正解で なかった問題の答えを見ずに採点ができる、 ゲーム感覚で気楽に取り組める等の記述が みられた. 本システムの問題点は、利用した検定試験問題が、公の検定試験ではなく、日本商工会議所が著作権を有する検定試験問題であったことである。したがって、日本商工会議所の許可なくしては利用できないシステムとなってしまった可能性が高い。著作権問題については現在調査中であり、静岡県立大学経営情報学部の講義の副教材として内部利用することも控えている。

本システムの今後の課題は、(1) インターネット上にデータベースサーバーを設置し、データベースにクッキーに蓄積された利用者の簿記学習状況を蓄積すること、(2) 逆にデータベースによってクッキー情報を更新すること、(3) 利用者の学習進捗状況をデータベースで集中的に把握できるようにすること、(4) 著作権者と連絡をとり、著作権の保護機能を内蔵させ、広く一般に公開することである.

参考文献

- [1] 山本達司「LAN を用いた簿記・原価計算教育支援システムの構築」追手門経営論集 1,337—351,(1995).
- [2] 金川一夫「九州産業大学におけるパソコンを用いた簿記システム教育―簿記システム学習ソフトの開発と応用・・」簿記学会年報第14号,70—75,(1999)
- [3] 税務経理協会編「[平成13年版第1版] 日商簿記検定試験3級出題傾向と対策」 税務経理協会.
- [4] 税務経理協会編「[平成13年版第1版] 日商簿記検定試験2級出題傾向と対策」 税務経理協会.

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 TELEVISION NEWS DATABASE IN THE U.S.: HISTORICAL OVERVIEW OF THE VANDERBILT TELEVISION NEWS ARCHIVE

> 魚住真司 Shinii UOZUMI

関西外国語大学 〒 573-1001 枚方市北方鉾町 16-1 Kansai Gaidai University 16-1 Kitakatahoko-cho, Hirakata city Osaka 573-1001

あらまし: 米国におけるテレビニュース・データベースの構築は、ある私立大学とその卒業生の大胆な試みに端を発する。 米国テネシー州ナッシュビル市に位置するヴァンダービルト大学 (Vanderbilt University) 付属の The Vanderbilt Television News Archive (略称 VTNA)では、1968 年以来、三大ネットワークのニュース番組を毎日ビデオ収録し、一般に公開してきた。1972 年には各ニュース番組の放送日と抄録を掲載した Television News Index and Abstracts を刊行し、VTNA の存在を世にひろく知らしめた。蓄積されたニュース番組の抄録を作成する過程で、VTNA のスタッフは業務の円滑化と利用者の利便向上をめざして、いつしかテレビニュース・データベースの構築をこころざすようになった。1994 年にはそれらをウェブ上に公開し、ここにインターネットにつながっている全世界の端末からオンライン検索可能なテレビニュース・データベースを完成させている。

今ではマスコミ研究者のみならずひろくその名の知られているVTNAであるが、設立当初は法的な困難を数多く乗り越えなければならなかった。特に、ほぼ3年間にわたって争われたCBS対Vanderbilt訴訟は、VTNAにとって設立以来最大の試練になるとともに、後に学術研究機関主導によるテレビニュース・データベースの構築が可能になるかどうかの分水嶺でもあった。

Summary: More than thirty years ago, the unique project was launched by the private institution in the United States: it was the effort to collect television news programs and open them to the general public so that

people can study and learn more about broadcast news media. The VTNA, Vanderbilt Television News Archive, one of the institutions at Vanderbilt University in Tennessee, has been video taping the daily news programs of ABC, CBS, and NBC since 1968. In 1972, it published something called Television News Index and Abstracts which is the world's first monthly tv-news database magazine. The magazine contained full of data such as on-air date and abstracts of each news program. It was so useful not only to the people who wanted to use VTNA but also to the VTNA staff itself. Now the magazine has been transformed into web pages that enable researchers around the world to do on-line search what was on the past news programs.

Despite the networks' claims that the Archive violates network copyrights, the Archive opens its collection to the public since its start. The VTNA and the Archive's founder had to fight for a few years in order to justify and secure the status of the archive.

キーワード: テレビ、ニュース、資料館、著作権、アメリカ 合衆国

Keywords: television, news, archive, copyright, United States of America

1.VTNAの設立過程

VTNAの設立者は、ヴァンダービルト大学法学部を1933年に卒業し、当時メトロポリタン生命保険会社(Metropolitan Life Insurance Company)の地区マネージャーを務めていたポール・シンプソン(Paul C. Simpson)である。シンプソンは、他のマスコミ研究者同様、テレビ報道の世論に及ぼす影響について強い関心を抱いており、テレビニュースを検証したいと強く望んでいた。¹⁾

ヴァンダービルト・テレビニュース・アーカイヴは1968年8 月5日、ヴァンダービルト大学図書館の一室において、三大ネットワークの定時ニュースの録画を開始した。三大ネットワークのうちNBCとABCは当初、ニュース番組を著作物として登録するような動きも見せず、シンプソンの計画についても、いわば成り行きを見守っているかのようであった。一方、シンプソンの試みを、当初は黙認するかのように見えたCBSは、アーカイヴの知名度が上がるにつれて難色を示すようになった。2

2.CBS, Inc.対Vanderbilt University事件

VTNAのかたくなな態度に対しCBSは、VTNAによるCBSニュースの収録ビデオ貸し出し禁止と、それらテープの「返還」を求めて訴訟を起こすことにした。CBS, Inc. v. Vanderbilt University (Civil Action No. 7336)³⁾である。本件は、アメリカ合衆国連邦地裁の一つ、テネシー中部地方裁判所(The United States District Court for the Middle District of Tennessee, Nashville Division)に1973年12月21日提訴され、1976年12月20日に双方からの提訴取下げで終焉を迎えるまで、ほぼ3年にわたって争われたことになる。

本件はCBS、Vanderbilt双方から提出されたファイル(趣意書や証拠物件等の書類をとじたもの)の日付をもとに、大きく三期にわけることができる。第一期は、CBSからの提訴があった1973年12月から、被告(ヴァンダービルト大学)が一連の抗弁を通して、一回目のサマリー・ジャッジメント(Summary Judgment、略式裁判的)を申し立てるまで。それから半年をおいて、5 第二期が原告CBSの告訴内容の追加・修正の申し立てがあった1975年6月から、その要求に対し被告が裁判所に却下の申し立てを行う1975年12月9日まで。第三期が、被告から裁判所に二回目のサマリー・ジャッジメントを求める1975年12月22日から、その申し立てが却下され原告の追加修正された告訴内容に対し、被告の回答がなされた1976年5月まで。なお、被告から最後のファイルが提出された1976年5月19日をもって本件には新たな展開はなく、同年12月に訴訟終結となる。

第一期 (1973/12/22~1974/12/20) の要点は、テレビニュース 番組が現行の著作権制度の中で、著作物として認められるかが 争点となったことである。原告は、夕方の定時ニュース番組*CBS* Evening News with Walter Cronkiteの1973年8月6日放送分を、 米国著作権局 (Register of Copyrights) に映画 (Motion Picture) として著作権登録し、ニュース番組であても著作物である旨主 張した。これに対し被告は、ニュース番組の著作権の保護は現 行著作権法では予定されていない旨の抗弁を一貫して行い、原 告の主張する争点は存在しないとしている。またこの時期、被 告はVTNAの存在価値を認める教育団体 (Council on Children, Media and Merchandising) をアミカス・キュリィ (Amicus Curiae、証人⁶⁾) として法廷に招くよう申し立て、成功し、VTNA の業務が公共の利益 (Public Interest⁷⁾) に貢献している旨主張 した。

第二期(1975/6/19~1975/12/9)の要点は、訴えの追加・修正の内容にある。原告は、CBS Evening News with Walter Cronkiteの1973年8月6日放送分に加え、1975年2月17日分を著作権局に登録申請し受理されたことから、ニュースの定義がどうであろうが、ニュース番組は現行著作権で保護されると主張した。つまり、ニュースの定義に関してはこれ以降、一切議論しないという原告の意思表示である。これに対し、被告は1975年11月19日、これまでの経緯をまとめたHistory of the Caseという一遍の文書を提出し、ニュース番組が現行著作権法上の「出版物(Publication)」たりえていないことを再度主張した。

第三期(1975/12/22~1976/5/19)の要点は、被告からの二度 目のサマリー・ジャッジメント申し立てについて、裁判所は「被 告は争点が存在しないことを十分に証明していない」との理由 で、その申し立てを1976年3月31日に却下したことである。つま り、裁判所としては、ニュース番組が現行著作権制度のなかで は何故保護されえないのか、被告にそれを証明する義務がある と判断したようである。これを受けて被告は1976年5月19日、追 加・修正された訴えに対して、回答書を提出した。なお、これ 以降は原告・被告ともにことの成り行きを見守ることとなる。 この頃から議会で著作権法改正案が成立する見通しとなったか らである。

そして本件は、結局1976年9月に著作権法改正案が議会を通過し、VTNAの業務に法的な後ろ盾が成立すると、CBSからの提訴取下げで終結する。この訴訟から解放されたとき、ヴァンダービルト大学の事務長(Secretary)で初代アーカイヴ運営委員の一人のロバート・マクガー(Robert A. McGaw)は、「アーカイヴの運営は、2年前にCBSが我々を訴えにでた当時と比べると、そう変化はない。しかしながら、これは必ずしも喜ぶべきことではない。訴えに対抗するために浪費した時間と金を、アーカイヴ運営の向上に役立てたかった」

9 と語ったという。このことから、いかにVTNAが訴訟問題で疲弊していたかうかがい知れる。

3.1976年著作権法改正とベイカー修正

米国の議会図書館(Library of Congress)は、1909年著作権法が、複製技術の発達など技術革新のため時代に相応しなくなったとし、1961年以来議会に抜本的改正を求めてきた。その後、議会は1974年にも改正を試みたが、失敗に終わっている。9 しかしついに1976年9月30日、上院法案第22号(S 22)が通過し、改正に至った。同法案の特徴は、出版物に限らず、ひろく著作権を保護している点と、その一方で、「フェア・ユース (公正使用) じり」という規定を設けて著作権に制限を加え、著作権者の利益の追及と、社会の利益とのバランスをはかっている点である。

ところで、1976年法案の原形となる1974年法案(S 1361)においては、テレビニュース番組の著作権も保護されることになっていた。しかし、この1974年法案は原型のままでは許諾なしの録画は全て著作権侵害となり、資料館によるテレビニュースの収集は、テレビ局の顔色次第になるという恐れがあった。これに対し、テネシー州選出の上院議員ハワード・ベイカー(Howard Baker=現駐日米国大使)は、1974年法案に対して修正案

(Amendment 1803)を議会に提出する。ベイカーは、かねてからテレビニュースの収集と研究が必要だと考えており、かつて議会図書館にそれを行わせるための法案を提出したこともある人物である。第93回米国議会における修正案提出理由説明で、ベイカーはヴァンダービルト・テレビニュース・アーカイヴの社会的貢献度を高く評価しており、これが後に、ベイカー修正を「ベイカー・ヴァンダービルト修正」と呼ばせるきっかけにもなった。ちなみにこの1974年法案は、日程上、下院で審議未了となり廃案になったが、ベイカー修正は、1976年著作権法案にそのまま引き継がれた。

1976年法案は、1976年9月30日に成立し(発効は1978年1月1日)、ここにヴァンダービルト・テレビニュース・アーカイヴの運営方針は、法的に認められることとなった。具体的には、合衆国法典第17巻(United States Code Title 17)の「資料館の職員は、その職務の範囲において、全種類の著作物についてそのコピーまたはレコードを複製し、これを公に頒布する」ことが許される、という部分に結実している。この1976年法案を受けて、CBSはVanderbiltに対する提訴を取り下げた。

4.データベースの完成とアーカイヴの現状

1972年以降、VTNAは、収録したニュース番組の索引と抄録を 掲載したTelevision News Index and Abstractsという冊子を、 全米の大学や研究機関に向けて毎月発刊していた。この冊子に よって、どのネットワークのニュース番組が、いつ、何を報道 したのか、VTNAのスタッフはもちろん、全米の研究者にも一目 瞭然となっていた。このことはVTNAの、テレビニュース資料館 としての完成度と知名度を一気に上げた。その一方で、知名度 の向上はVINAのスタッフにとって業務の拡大を意味していた。 この頃からVINAのスタッフは、日々増大する問い合わせの対応 に忙殺されるようになっていく。限られたスタッフ数しか確保 できないVINAにとっては、何らかの対策が必要であることは明 らかであったが、それにはもうしばらく時間を要した。

1989年、VTNAは大学所有のベイカー・ビルディング (Baker Building)という、大型の建物へ移転することが実現した。それまではヴァンダービルト大学図書館の一室で業務を行ってきたが、増加する一方のビデオテープの保管場所がついにつきたのであった。ベイカー・ビルディングでは、7階の新しいオフィスに加えて、地下に約180平米の倉庫を得ることになり、逼迫するテープの保管場所の問題は、ひとまず解決を見た。また、この移転を機に、アーカイヴの業務にコンピュータの導入がはかられた。VTNAのスタッフは手分けして、これまで蓄積されてきたニュース番組の抄録をコンピュータに入力していった。この地道な作業によって、次第にテレビニュース・データベースが構築されていった。

1994年10月、VTNAはデータベースの検索機能を備えたウェップページの公開に踏み切り、それとともにTelevision News Index and Abstractsを廃刊させることにした。ウェッブページにVTNAの業務に関するほぼ全ての情報を掲載したことで、初歩的な問い合わせは激減した。その一方で、ビデオテープの貸し出し請求の件数が飛躍的な伸びをみせた。ちなみにデータベースを公開した当初は、世界中から、週に100を超える貸し出し請求があったという。現在は、繁忙期を除き、週におよそ30程度に落ち着いているという。主に3人の非常勤スタッフと数名の学生がこれらの請求を処理している。また、1991年に勃発した湾岸戦争以来、CNNの「ワールド・ニュース(World News)」についても収録を始めたため、さらに、大統領選挙の際のテレビ討論会等、特別番組も守備範囲としているため、現在5人の常勤職員は、それら番組の索引と、抄録作りで多忙な日々を送っている。

ところで、VTNAのウェブページには次のように書かれている 部分がある。¹¹⁾

Texas Monthlyの1991年11月号によると、CBSニュースのアンカーであるダン・ラザー(Dan Rather)は、次の二つに対して責任を負っているとインタビューに答えたー「視聴率とヴァンダービルト・テレビニュース・アーカイヴに対してだ。」

つまり、過去に訴訟まで起こしてVINAを追い詰めたCBSであるが、今となってはその看板アンカーでさえVINAの存在価値を高く評価しているということである。VINAが乗り越えてきた試練は、社会にとって様々な教訓を投げかけていると言えよう。

注および参考文献

- 1) このあたりの経緯については、シンプソン自らが記した Simpson, Paul C. Network Television News: Conviction, Controversy, and a Point of View. Franklin, Tennessee, Legacy Communications, 1995に詳しい。
- 2) この頃、Efron, Edith のNews TwistersやHow CBS Tried to Killa Bookといった、CBSにとって歓迎できない内容分析調査報告書が出版された。このような背景にも留意したい。
- 3) 本件の公判記録は既に連邦地裁ナッシュビル局にはなく、ジョージア州の連邦記録文書センター(Federal Records Center)で保管されている。
- 4) 東京大学出版会の『英米法辞典』によると「略式裁判」という翻訳は厳密性に欠け、「重要な事実について真正な争点がなく、法律問題だけで判決できる場合に、申立によりなされる判決」であるという。
- 5) ここで6ヶ月間も停滞したのは、議会で著作権法改正の動き があり、双方ともこの動きに関与(聴聞参加等)していたから と思われる。
- 6) 前出『英米法辞典』によると「証人」という翻訳はは厳密性 に欠け、「裁判所に係属する事件について裁判所に情報または 意見を提出する第三者、法廷助言者」であるという。
- 7) 放送に関する訴訟では、Public Interestが最も重視される、 キーワードである。放送局は、Public Interestに貢献することを 前提に免許が与えられているからである。
- 8) Simpson, 前掲書。120p.
- 9) 土井輝生. アメリカ新著作権法の解説. 東京, 第一法規出版, 1981. 4p.
- 10) ただし、フェア・ユースの範囲がケースによってまちまちであることから、フェア・ユース抗弁には慎重論もある。
- 11) tvnews.vanderbilt.edu

画像・音声の超高再現性 PC の開発

Development of The Ultra Hi-Fi Graphics and Sound PC

片岡 裕
Yutaka Kataoka
kataoka@res.otani.ac.jp

大谷大学文学部人文情報学科,京都府京都市北区 Otani University,Kita-ku, Kyoto City, Kyoto-Fu

あらまし: 高精度デジタル静止画像データは、マイクロフィルムや印刷より分解能・明度・色調で、より高い再現性を保証可能であることが証明されている[1、2、3]。同様に、96KHz24Bit 以上のデジタル音声データは、44.1/48KHz 再生で誤差が含まれる波形推測を必要とせずに、4 KHz 以上の周波数において高精度再生が可能である。また、デジタル・ハイビジョンは、1920×1080ピクセルの基本分解能を持つ。このような高精細・高精度データは、PCを用いて再生する以外に、フル・フォーマットでの再生が不可能であることは言うタサでの再生が不可能であることは言うタ共有と評価には、極めて高い再現能力を持つPCが必須なのである。

しかし、PCでは、内部の電気的ノイズ及び 冷却 FAN やモータの振動・騒音により、高再 現性を実現し得ないと言われてきた。そのた め、PC内で種々のノイズの影響を検証[4]して 排除し、同時に高安定化する研究を行った。 その結果、画像(RGB 各 8Bit 範囲¹)は誤差が測 定限界以下、音声デジタルアウトは 0.5n 秒以 下のジッタ、機械的騒音は 10dB 以下の、Ultra Hi-Fi PC を開発した。さらに、この Ultra Hi-Fi PC を、確実かつ安価に作成する方法を確立し た。

キーワード: デジタル博物館、デジタル図書館、高再現性、古典学、HTPC

Keywords: digital museum, digital liblary, high fidelity, philology, HTPC

1 再現性からの Hi-Fi PC の必要性

データベースのコンテントは、マルチメディア化が進行しており、再生装置の対象がテキストのみのでは不十分である。コンテンツの主要フォーマットは、静止画、動画、音声と多岐にわたる。さらに、それぞれのフォーマットは細分化され、非常に高精度のデータを記録可能となっている。これら種々のフォーマットのデータは、ネットワークでのアクセスを前提としており、フォーマットに依存しない再生機器としての PC の使用が前提となっている。

PC で再生可能なデジタルデータと比較し、マイクロフィルムや写真印刷(静止画)・NTSC ビデオ信号(動画)・CD(音声)のなどの既存メディアのフォーマットでは、記録および再生精度に大きな制限がある。

静止画

高精細・高精度のデジタル画像データは、マイクロフィルムや印刷より高い再現性を持っている[1、2、3]。

マイクロフィルムの撮影時には、一般に可 読性の保証まではされず、モノクロームであ り、汚れや着色がある場合、資料の可読性は 極めて劣化してしまう(図 1)。さらに、フィル ムのデュープリケーションで、再現性が低下 する。マイクロフィルム・リーダの投影画像

¹ ATI 社の Graphics CPU では、RGB 各 10Bit の DAC を内蔵する。

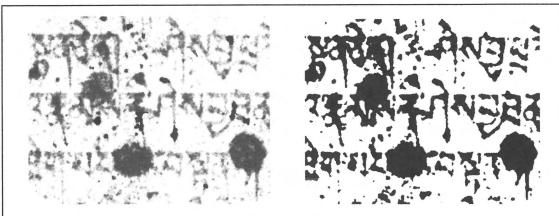


図1. モノクローム・マイクロフィルム(右)による情報の欠落

も色温度に制約があり、カラー・マイクロフィルムの再現性は、著しく損なわれる。

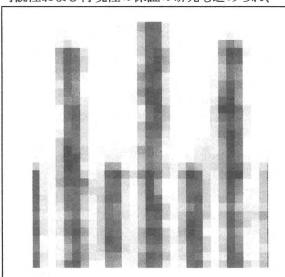
写真印刷は、印刷のカラー分解法が、1ピクセルを複数点の基本色点に分解するため、実質的な解像度と明度階調が低下してしまう。さらに、インクの粒子径を小さくすると、インクの厚みが薄くなり、吸光度が低下し、コントラストが低下する。また、インクの粘度が温度依存性であるため、低温の季節では硬化を防ぐために溶媒量が増加し、結果として色が薄くなる。当然、用紙の白度によって色相が変わってしまう。

静止画のデジタル化と公開は、資料の共有と参照の容易性から、貴重書を中心として実施されている。静止画の高精細化および高精度化は、資料の可読性を上げるに留まらず、可読性および再現性の保証の研究も進められ、

実際にデータベースの構築が進行中である[1,2、3]。さらに、画像データの補正に関しても、特定の印刷装置ではなく、標準光源での状態として記録可能である[3]。これらの研究の結果、資料の可読性(分解能)を保証し、同時に色再現性と明度階調を肉眼に近いところまで再現可能である。従って、再現性の差による画質の差は想像以上に大きく、その例を図2に示す。図2は、CRT像を撮影し、ピクセルサイズの領域に分割し平均したものである。ピクセルのジッタが大きい通常のPC(右)では、図2の様に、細部の情報が著しく欠落する。

動画

VTR、LD、DVD などの動画は、NTSC ビデオ信号を再生するためのテレビ受像機で画像が再生されるため、横方向の分解能は、720ピクセル程度に低下してしまう。縦方向の限



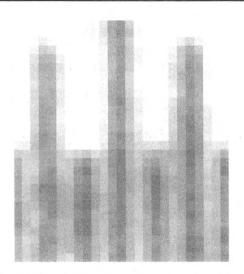


図2. 再現性の差による画質の低下

界は、実質 480 本の走査線である。同様に衛星放送によるデジタル・ハイビジョン放送でも、NTSC 互換受像機では、1920×1080 ピクセルでの再生は不可能である。NTSC 再生では、明度階調が十分な範囲ではなく直線性も低い。

PC は、デジタル・ハイビジョン画像の 1920 ×1080 ピクセルを持つ動画を、そのままのピクセル数で再生可能である。この動画は、MPEG2 圧縮で 30MBPS のデータ量となり、100BASE のネットワークでの PC への転送が可能であり、インターネットでのストリーミング放送も実験段階にある。デジタル・ハイビジョンの分解能は、医療用に留まらず、フィルム映画のデジタル化(テレシネ)、テキスタイルや遺跡の記録など、極めて用途が広い。また、今後デジタル・ハイビジョン・カメラは、24コマ/秒のプログレッシブ方式(ノンインターレース方式)に変更され、プログレッシブ・スキャンの PC での再生により適合する。

なお、DVDのNTSCビデオ信号での再生では、マイクロビジョン方式のコピー・プロテクション信号が含まれ、VTRテープと同様、画室が劣化することが多い²。

音声

CD は、44.1KHz16Bit、DAT は、48KHz16Bit のリニアーPCM 記録である。最も多用される CDでは、18KHzまでの記録が可能であるが、 高域では波形を記録するサンプリング数が相 対的に減少し、誤差が著しく大きくなる。従っ て、CDやDATでは、そのままDA変換し18KHz をカットするフィルタを通すと、4KHz以上で は数%を越える誤差となる。そのため、CDや LD のデータの再生時には直接 DA 変換を行わ ず、再サンプリングを行い、サンプリング・ レートを 8 倍から 64 倍に上げ(オーバー・サ ンプリング)、さらに、サインカーブ補間を行 うことが一般的である。これでは誤差が残る ので、DSP(Digital Signal Processor)を用い、波 形の推測を行い、オーバー・サンプリング時 の補間を行う方式も多い。また、弱音では、

16Bit の直線性では、山・谷を2Bit 程度で記録することとなり、やはり誤差が蓄積する。オーバー・サンプリングを行う過程で、16Bitから20Bit、もしくは24Bitに拡大して、弱音域の補間を行うことが一般的である。このようなAD変換時の技術は、サンプリング・レートが低く、ダイナミックレンジが小さいことに原因があり、失われた情報の推測に過ぎず、誤差が減少するわけではない。

DVD-Video もしくは DVD-Audio では、 192KHz24Bit 記録が可能であるが、コピー防 止のため、いまだ再生技法が明確に決定され ていない³。

PCでは、市販されている AD・DA コンバータは、最大 192KHz24Bit サンプリングまで可能である。96KHz 以上のサンプリング・レートでは、波形の真値との差が少なく、波形補正後の誤差も大きくない。また、同時録音・再生も8 チャンネル以上が可能である。従って、メディアの規格に左右されずにハイビット・ハイサンプリングレートの記録と再生が可能である。

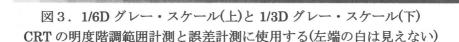
Hi-Fi PC と Hi-Fi 再生装置の必要性

このように、PCでの再生は、既存のメディアの制限を受けず、より簡便に高再現性データの再生が可能である。従って、PCを用いて高精度・高再現性データを再生しなければならず、高再現性 PC、すなわち、Hi-Fi であること保証した PC が必要となるのである。

PC の再生信号は、アナログの映像信号とデジタルの音声信号が中心となる。アナログ映像信号では、RGB 各色の階調は、MPEG2 の再生ではカラーデコーダと DAC が 10Bit であり、Bit 換算で各 10Bit が可能である。しかし、デジタル映像信号転送では、ビット・レートが高くなりすぎるため 8Bit 階調であり、再現性に劣る。そのため、映像の表示装置としては、CRT が選択される。音声に関しては、Dolby Digital 5.1 など種々のフォーマットがあり、再生には専用のデコーダを搭載したアンプを必

² 現在、コピー・プロテクション信号を除去 することは違法行為である。

³ PC による 96KHz24Bit 以上の DVD 音声の デジタル出力は、現在ではコピー・プロテク ションのために不可能である。



要とする。従って、これらの装置の再現性と 共に PC の再現性を考えなければならない。

特に映像の再生装置である CRT とプロジェクタの性能差は極めて大きく、適切な機器を選択しなければ、PC 出力信号の画質の評価すら不可能である。図 3 は、画像出力機器の性能を評価するためのグレー・スケールである。1/6D 階調のグレー・スケールを全て表示可能である表示装置は少ない。1/3D 階調グレー・スケールで、3.5D の範囲が表示可能であり、1/6D 階調グレー・スケールで図の全階調が表現可能でなければ、静止画の正しい表示はできない。本論文の印刷では、図 3 は正しく表示されないので、正しいグレー・スケールによる CRT 調整の詳細は、下記の URL を参照していただきたい。

http://www.hifi-pc.com/crt.htm

液晶は、シャッターから光が漏れて階調表現が正確ではなく、アナログ画像信号を AD変換して再生するため、さらに誤差が大きくなる。CRTは、21インチでシャドーマスクピッチの小さい円形シャドーマスク形式でないと、1920×1440ピクセルの分解能を得ることはできない。アパチャーグリル型 CRT では、縦横の分解能が異なってしまう。最近の平面に近いアパチャーグリル方式では、左右と中央のグリル間隔が異なり、高い再現性を持ち得ない。プロジェクタでは、9インチ管を使用した3管式でなければ、1920×1440ピクセルの表示で十分に電子線を絞りきれない。

音声の DA 変換では、外付けの AV アンプを使用するだけでなく、その精度に十分注意が必要である。 1 Bit 方式 DAC では、ジッタによって PIM 歪が発生してしまうため、倍音

の位相が狂って聴こえてしまう。同様に、スピーカでは、エンクロージャの振動により発音するタイプのスピーカでは、正確な再現が困難である。ヘッドフォンによる再生では、頭部に装着した時の周波数特性を考慮した規格を満たす製品でないと、再生周波数全域に渡っての正しい音圧を得られない。

2 無対策 PC の再生誤差

無対策 PC では、PC 内のノイズによって出 力誤差は極めて大きい。RGB 各 256 段階の 8Bit-DAC の画像出力は240 以下には誤差が含 まれ、正しい暗部色調は望めない。ピクセル の位置揺らぎは、近接2ピクセルの重なり合 いとして、横方向 1920 ピクセルでは 60%を 超える(図 2)。Sync 信号(輝度基準・掃引時間 基準)の誤差は10%以上である。デジタル出力 音声では、最大で 10n 秒を越えるジッタであ り、高再現性再生に必要な最大で 1n 秒以下に はならない。ジッタが 1n 秒では、マルチチャ ンネル音声再生での各音声チャンネルの相関 は、聴感上で明らかなほど劣化する。なお、 Dolby Digital 信号では、出力装置が時間管理 のマスターとなるため、AVアンプによってジッ タの排除はできない。そのため、PC のデジタ ル音声信号のジッタは、最後まで再生品位に 影響を与えてしまう。PC 内の DAC を使用し た場合、特殊な製品を除き 24 Bit 再生では下 位 4 Bit に相当する誤差を含み、さらに電気的 ノイズが混入して S/N 比が劣化し、評価に十 分な品位は持ち得ない。

また、PC の再現性だけでなく、PC からの 画像と音声信号の再生装置の再現性も、前項 に示したように極めて重要である。図書館や博物館では、実測した数値を示せる精度を持った CRT やスピーカを設置すべきである。

3 PC の電気的ノイズの原因

PCの出力信号の劣化は、電気的ノイズが原因である。しかし、PCのノイズの原因と対策をPC全体にわたって追求した研究はないが、ノイズの主因は、スイッチング電源のリップル電圧とデジタル信号路が輻射するノイズであると考えられてきた。

電気的ノイズは、コモンモードとノーマル モードに分類され、どちらも輻射と考えれば、 電流量の増減の絶対値に比例する[5]。電流の 実測値は、デジタル信号が 5mA 程度なのに対 し、CPU や I/O チップ、メモリへの電源電流 は約0.5Aから最大12A程度で約100KHz~数 GHz の断続的電流である。HDD や DVD ドラ イブへの電源電流も, 0.5A から最大 4A に達 する断続的電流である。これら断続的大電流 の経路からのノイズは、信号や電源由来のノ イズより 1000 倍以上大きい。 同様に FAN や モータには, 逆起電力吸収用と放電防止用の キャパシタが無く、12V 駆動であるため、強 いノイズを発生している。これらのノイズは、 デジタル信号にジッタを加え、多大な回数の エラーとリトライをチップ内とチップ間で発 生させる。さらにアナログ回路の混変調歪を 極端に増加させ、リニアリティーを減少させ ている。

マザーボード上の CPU は、800MHz 駆動の Pentium-3 で約 20W を消費する。実際の使用 時では、負荷の平均値は DVD 再生などでも 40%を超えることは稀であるが、上述の様に 消費電流の短時間の増減は極めて大きい。メモリおよび North Bridge(CPU、AGP バス、メモリ、PCI バスへのデータ転送機構)も同様の電流消費パターンである。これらの LSI への給電系は、1500 μ F 程度の比較的高周波特性の良いアルミ電解コンデンサで瞬間的な電流変動に対応するようになっている。しかし、消費パターンの分析から GHz にまで及ぶことが判り、150KHz 程度のリップルまでしか効果

がないアルミ電解コンデンサでは、周波数特性として十分ではない。このため、AGPバスおよび PCI バスにリップル電圧が発生し、クロックジッタなどの影響を与え、ピクセルが重なり合って画像の明度・彩度を下げ、デジタル音声出力のジッタを増加させる。

AGPバスおよびPCIバスには、DDコンバータで電圧変動が抑制されているが、バッファとなるアルミ電解コンデンサの容量が不足している場合が多い。最近のグラフィクス・ボードのメモリは 64MB の高速メモリであり、グラフィクス・ボード上の CPU も大電力を消費するので、アルミ電解コンデンサでは周波数特性的にも十分ではない。同様に、高度なサウンド・カードも比較的消費量の大きい LSIを搭載しており、挿入してある PCI バスの位置にキャパシタによる電流供給の安定化と電圧変動の抑制が必要である。

PC 内は給電電流路が長く、振動に由来する ノイズは、数 10mV に達する低周波成分であ る。これは、直接的に画像信号の Sync 信号の 時間軸上と電圧(明度基準信号)のジッタとなり、 同時にデジタル音声出力のジッタとなる。

さらにこれらのノイズ源は PC 内に広く分散しており、それぞれの輻射が空間で合成され極めて大きいノイズとなる。この合成ノイズは、繰り返しの無いパルス状となり、オシロスコープでの確認は極めて困難である。また、磁性体は全て、入力周波数より高い周波数で、強く2次輻射することが判明したため、PC筐体内の鉄やステンレスまでがノイズ源となってしまう。

4 機械的な再現性低下の要因

上述の輻射系ノイズに加え、再現性低下にメカニカルな要因がある。HDD(ハード・ディスク・ドライブ)は、動作中に共振点を持ち、リードエラーが多発する記録領域がある。この領域では、ヘッドのサーボが振動によって激しく動作し、消費電流も増加している。このような領域は、HDD内に数箇所あり、HDDの固定方法によっては、リード速度が10分1以下にまで低下することがある。従って、ハ

イビット・ハイサンプリングレートの音声再生や動画再生では、時間的な遅れが発生してしまうことなる。リード時間の遅延は、デバイスドライバのスピンロック時間を長くし、CPU 負荷を上昇させ、画像・音声用のデバイスドライバの処理時間を圧迫する。結果、画像信号と音声信号のジッタとして現れる。

DVDD(DVDドライブ)も、HDD 同様に機械的な共振点を持ち、DVDメディアの偏心によって、レーザー・ヘッドが読み出しに時間を要する領域ができてしまう。このような領域は、DVDメディア上で5、6から30以上あり、画像のデコード時間を奪うため、画像を極端に劣化させ、さらに音声が画像に同期しているため、音声までがジッタを含んでしまう。DVDDは、再生中加減速を繰り返し、そのための歳差運動も大きい。DVDDの機構のほとんどは、ゴムによるダンパーで動くように支持されるため、この運動によるリードエラーとサーボの増加は無視できない。

このような機械的な共振と機構そのものが 原因となるノイズは、時間軸上の誤差だけで なく、サーボ電流を増加させ、グラフィクス・ ボードとサウンド・カードへの給電電圧を変 動させてしまい、S/N 比を劣化させる要因に もなっている。

5 ノイズへの対策方針

上述のように、PCのノイズは給電路、輻射 および振動由来であることが判明した。従っ て、次の方針と手法でノイズをカットし精度 を向上した。

- 1. ノイズ対策が可能で、その結果が反映される高精度・高品位部品を選択する⁴。
- 2. 磁性体を排除し、100Hz 以上の電波・磁場を遮蔽する Al 板で、電源、ドライブ、マザーボード、ドータ・カードを分離する。
- 3. 硬度の異なる AI 板による制振合板を、振動モードにあわせて作成して筐体内部を補強し、ドライブから振動の発生と伝達を阻止す

る。

- 4. HDD と DVD ドライブの内部に制振処理をしてサーボ量を減らし、電流消費を下げる。
- 5. 大電流かつ断続電流消費部品・チップの 近傍に低 ESR 大容量コンデンサを接続し、給 電路の電流・電圧変動を減少させる。
- 6. 電源・給電路の高周波インピーダンスを 上げ、さらに給電ケーブルは全てツイスト⁵し, 高周波の輻射を減少させる。
- 7. エアフローを設計し、FAN(コンデンサ付き)を電源内のみとし、FAN はスポンジで浮かせ、サイレンサーを設け、振動・風切り音を取る。
- 8. 筐体内の Graphics、Sound card 部では、近 傍のケーブルを排除する。

さらに、2次輻射の無い非金属非磁性体製 ノイズ吸収剤(開発中)で、給電路と AI 板表面 のノイズ反射を吸収し、シールド効果を高め た。

6 部品の選択

Ultra Hi-Fi PC の開発は、部品の選択によって始まる。前項で述べたように、高精度の部品を使用しなければ、改造・改良後にも良い結果を得られない。また、直接的に信号を生成するグラフィクス・ボード、サウンド・カードだけで性能が決定されるわけではない。

PC 筐体

最も精度と S/N に影響を与える部品は、PC 用筺体である。「OWL-PC-61」ブラック(オウルテック社)を選択した。

本筐体は、インナー・ケースの剛性が比較的高く、改造に適した構造であることと、全体がAl合金でできており、2次輻射が少ないことによる。現在では、最も適したケースは、Silent Master(Japan Value 社)である。

電源

電源は、300W の静音電源(Seven team 社)である。これは、ファンが1個であり、輻射が平均的で、改造が容易であるため選択した。

⁴ 部品によって、100倍程度ノイズ量が異なり、 10倍程度精度も異なる。

⁵ PC では, グランドに平均約3A の電流が流れており,シールドケーブルを使用するとシー

電源は公差が大きく、全体の部品の組み合わせで最適のものを選択し選別すべきである。 なお、PCではグランド電流が流れることに注意が必要である。

マザーボード

マザーボードは、筐体についで、ノイズの下限を決定する重要な要因である。 CUV4X-E(Asus Tech 社)を選択した。本マザーボードは、CPU と North Bridge への給電に多数の 1500μ Fのアルミ電解コンデンサを使用しており、余裕がある。マザーボード上で使用する電圧は、12Vを除き、全て DD コンバータによって生成され、リップル電圧を下げられる構造になっている。また、フェライト製チョークコイルなどの輻射が激しい素子とコネクタ類が、出力信号を生成する部品が挿入される AGP バスおよび PCI バスから離れている。さらに、付加コンデンサを取り付けるスルーホールがある。

グラフィクス・ボード

グラフィクス・ボードは画像信号を生成するため、直接精度に影響を与える部品である。RADEON64MBVIBO(ATI 社)を選択した。このボード上のDACは、10Bitで、静止画で使用される上位8Bitの誤差が極めて少ない。また、MPEG2再生では、カラーマップが10Bit分あり、動画部のDA変換は10Bitを使用することができる。このため、全て8Bitである他のグラフィクス・ボードより高再現性を考慮した設計となっている。なお、比較的公差が少ないボードであるが、飛び込みノイズの影響を受けやすい。なお、本ボードは、グラフィクス専用バスであるAGPバスに挿入される。サウンド・カード

サウンド・カードもグラフィクス・ボードと同様に直接再生信号の精度に影響を与える。 RME DIGI 96/8 PST(RME 社)を選択した。本カードは、最大 96KHz24Bit 再生と記録が可能であり、アナログ入出力機能を持っている。本カードは、高再現性を目標に設計されており、高いポテンシャルを持っている。特にデバイスドライバの性能が良く、CPU の負荷が

ルドに流れる電流で,輻射ノイズが増大する。

極めて小さく、遅延も少ない。

DVD ドライブ

DVDD(DVD ドライブ)は、GD7000(日立)を 選択した。GD7000 は、最も動作音が小さい。 DVDD は、動作音を除いて、機種ごとの大き な精度上の差は少ないが、構造から公差が大 きい。従って、複数から選択することになる。

HDD

IC35L020(IBM)を選択した。この HDD は、7200RPM で、騒音がやや大きく、適してはいない。最適な HDD は、2.5 インチのモバイルPC 用である。3.5 インチ HDD では、流体軸受けを使用した5400RPMの静音型が適している。HDD ケース

HDD は、振動と騒音が激しくそのまま装着すると機械的ノイズを取れない。そのため、Smart Drive(グロウアップジャパン社)を選択した。騒音が問題でなければ、PH-35B PRO(Japan Value 社)が最も制振性とノイズ遮断性に優れている。

IDE ケーブル

フラットケーブルでは、輻射の反射がある ため、丸型ケーブルの GCWF-100S(Justy 社) の 29cm を選択した。本ケーブルは、ATA100 対応で、HDD と DVDD に 1 本づつ使用して いる。

CPU

動画再生の品位にも着目した場合、ピーク電流の大きな CPU を選択することはできないので、Celeron633(cC0 コア)を FSB=100MHz とし950MHz で駆動している。

CPU ヒートシンク

高さ 85mm 幅 60mm 長さ 60mm の特殊な ヒートシンク(サーマルコンポーネント社)を選 択した。ヒートシンクが高いと輻射が大きく なるが、輻射防止が可能となったので、振動 防止のため、電源のファンの排気量で冷却可 能な熱交換能であるのでこれを用いた。

メインメモリ

256MB の SDRAM(Winbond 社)で、低消費電力の選別品を選択した。消費電力が大きい物を選択すると、リフレッシュ時に大電流が流れ、AGP バスと PCI バスへのリップルが増加する。

ネットワーク・カード

基板が小さく、高周波の輻射が比較的小さいので、EtherExpress Pro100+(Intel 社)を選択した。

その他

FDD(フロッピー・ディスク・ドライブ)は、OS に Windows2000(SP2)を選択したため、接続していない。

電源ケーブルは、4線対向スタッガー結線 で、アース線でコイル状に全体を巻き、アー スに関して高周波フィルタ特性を持つケーブ ルを自作した。

電源環境

部品ではないが、電源環境が良くなくては、 高再現性再生は不可能である。壁の電源コン セントは、HUBELL 社製の US Hospital grade を使用し、第1種アースに近い接地抵抗のアー スを接続した。

7 ノイズ対策加工

部品を選択し、組み合わせてもUltra Hi-Fi PC にはならない。 すなわち、一般的な PC の組み立てではノイズ対策にはならず、部品に対する加工と配線の工夫が必要となる。

PC 管体

インナー・ケースに硬度と厚みの異なる AI 板を取り付け、全体を制振合板化した。特に振動の激しい HDD と DVDD の取り付けは、側板部に剛性の低い AI 板で制振合板化し、ダンプ機構付き板バネ構造とし、振動エネルギーを積極的に摩擦熱に変換すると共に捻れの反動を吸収した。HDD と DVDD は、ゴムワッシャを介してポリカーボネート製ネジを使用し、高周波振動の吸収を行った。

3.5インチベイに無共振煉瓦を硬質スポンジで浮かせて固定し、共振周波数をずらし、共振の鋭さであるQを分散した。Qと共振周波数の分散は、振動エネルギーの伝達を阻止する。

マザーボードの取り付け板に対し、厚さ3mmと厚さ2mmのAl板を取り付け、間にCu箔(t=0.1mm)を挟み、グランド電流をCu箔に通すと共に、表面輻射抑制と制振合板化を行っ

た。

電源部、ドライブ部、マザーボード部を Al 板(t=2mm)で分離し、輻射の影響を除いた。純 銅は、輻射の反射が大きく使用できない。

結果、ほとんどの振動が吸収され、マザーボードには伝達されなくなった。なお、振動防止のため、冷却ファンは取り付けない。

磁性体およびステンレスは全て、500MHz から1GHzの2次輻射が大きく、画質・音質 に大きく影響するため全て排除し、ネジなど は、黄銅とした。

冷却ファン

冷却ファンには、 $10 \mu F$ の無極性アルミ電解コンデンサを取り付け、給電ケーブルをツイストした。

電源

グランドに高周波ノイズが伝達されることを防ぐため、グランドへ3ターンのコイルで結線した。冷却ファンの振動を防止するため、硬質スポンジでケースから分離した。コンデンサおよびヒートシンクを非磁性体のノイズ吸収材で覆い、輻射と振動を押さえた。ケースの内側および外側は、同上のノイズ吸収材で覆い、輻射を約20dB下げた。

給電配線

給電ケーブルは、最も大きな輻射の原因である。ループ面積を最小とすると同時に、輻射を打ち消し合わせるため、全てツイストした。ツイストペア線は、ノイズキャンセルに対する効果が極めて大きい。

給電ケーブルは、マザーボードから離し、マザーボードへの給電ケーブルも最もマザーボードから離して配線した。なお、給電ケーブルは、全て固定し、振動による発電を避けた。

HDD

HDDの制振合板が薄いための共振であるので、制振材を貼り、共振を防止して、HDDケースに取り付けた。HDDケース内のスポンジの硬度を調整して移動量を制限し、発熱の伝達を向上した。結果、振動が抑制され消費電流が減少した。

DVDD

ヘッドの微振動がリードエラーの原因とな

るため、ヘッドに鉛箔(t=0.3mm)を貼り、モータ周辺に微振動抑制のため、制振材を貼った。 結果、リードエラーが極めて減少した。当然 消費電流が減少し、発生する電気的ノイズも 騒音も激減した。

ドライブ用給電ケーブルフィルタ

HDD と DVDD は、有機半導体固体電解コンデンサ(佐賀三洋工業株式会社)、積層セラミックコンデンサ(日本ケミコン)を取り付けた高周波フィルタを介して給電ケーブル(5V12V)に接続した。なお、有機半導体固体電解コンデンサの容量は、ドライブの消費電流によるが、500 μ F 前後で良い。積層セラミックコンデンサは、50V10 μ F である。

この高周波フィルタによって、約 25Hz~20MHz 程度までの断続的な電流消費によるノイズを抑制可能である。

マザーボード

CUV4X-E の CPU への電流供給用 $1500 \, \mu \, F$ のアルミ電解コンデンサ3本のうち2本を $1500 \, \mu \, F$ の有機半導体固体電解コンデンサ(佐賀三洋工業株式会社)に置換し、残り 1 本を $2200 \, \mu \, F$ のアルミ電解コンデンサ(HD 型: ニチコン株式会)に置換した。

同様に、North Bridge への給電用 1500μ F の アルミ電解コンデンサ 3 本を、1 本の 1500μ F の有機半導体固体電解コンデンサ(同上)に置換し、1 本 1500μ F のアルミ電解コンデンサ(同上)、 1 本を 2200μ F のアルミ電解コンデンサ(同上)に置換した。

CPU と North Bridge のコンデンサの高周波 リップル除去特性を向上することで、AGP および PCI バスへのリップル電圧が減少する。しかし、有機半導体固体電解コンデンサを多用すると、インピーダンスが低下しすぎ、AGP バスと PCI バスへの給電能力が低下する。そのために、HD 型の高性能アルミ電解コンデンサを並列に使用している。アルミ電解コンデンサには、特性向上のため、 $10 \mu F$ の積層セラミックコンデンサを付加した。

AGP バス給電部のアルミ電解コンデンサ (1000 μ F)を 560 μ F の有機半導体固体電解コンデンサ(同上)に置換した。この置換で、グラフィクス・ボードへの給電能力が増加すると

同時に電圧リップルが減少し、極めて画質が 向上する。

PCI バスのサウンド・カードへのスロット 部にあるスルーホールに 560μ F(4V)と 390μ F(6V)を付加し、サウンド・カードへの給電能 力を上げ、さらにリップル電圧を減少した。 この付加で、デジタル音声出力のジッタが激減した。

さらに、非磁性体ノイズ吸収材で、平行配 線部を全て覆い、輻射を減少した。

これらの加工で、North Bridge と South Bridge のエラーも減少した。

CPU ヒートシンク

振動と輻射の防止のため、非磁性体ノイズ 吸収材で覆った。電源の排気の背圧を利用し て冷えるようにダクトをつけた。

グラフィクス・ボードおよびサウンド・カード、ネットワーク・カード

非磁性体ノイズ吸収材で全体を包み、外部からの輻射ノイズの遮断と、ボード上の素子間でのノイズの影響を避けた。このノイズ吸収材は研究途上のものである。良いノイズ吸収材を使用すれば、その効果は劇的である。しかし、一般的な特定の周波数用の磁性体を主としたノイズ吸収材では、2次輻射のため、かえって劣化してしまう。

ネットワーク・カードは、WOLの機能を使用するケーブルを接続すると劣化する。

グラフィクス・カードは AGP バスに、サウンド・カードは PCI バスの3番に、ネットワーク・カードは、PCI バスの6番に挿入し、それぞれの干渉を最小とした。

IDE 配線

非磁性体のノイズ吸収材で覆い、振動を防止して接続した。

8 結果

騒音は 10dB 以下となり、ほとんど無音である。残った騒音の原因は、グラフィクス・ボードのファンの風切り音である。

振動は、DVDが不定速回転中でもほとんどない。振動の影響は、サウンド・カードのアナログ出力電圧の変動で計測できるが、測定





図4. 未対策 PC(上)と Ultra Hi-Fi PC(下)の比較(1920×1440): 同一部品構成

値には現れなくなった。使用はしないがアナログ音声出力の S/N は、約 20dB を良くなった。これは高級 DAC 並みである。

静止画は、直線性のある階調範囲が 3.7D を 越え、さらに RGB 0 と、3.7D が識別可能となっ た。無信号走査線の輝度が最もノイズを示す が、明度での比較で 20%以下まで減少した。

ピクセルジッタは、1920×1200 ピクセル表示(60Hz)で、ピクセル間の重なり合いとして、20%以下になり、1 ピクセル幅の黒と白を完全に分離可能となった。ジッタとしての直接計測は、ほとんど不可能なレベルであるため、CRT を直接撮影して、重なり合いを計測した値であるため CRT の変動を含む。従って、実際のジッタはさらに少ない。CRT は、現在最高の性能を持つF980使用し、3段のノイズフィルタを挿入した極めて安定した電源環境で使用し、D1(Nikon)を用いて撮影した。

静止画だけでなく、最もノイズに影響される DVD 動画でも、色相・分解能・明度共に極めて高精度となり、ほとんど誤差を含まない結果となった(図 4)。

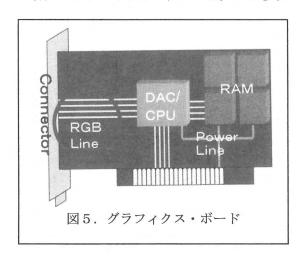
音声デジタルインとアウトでは、0.5n 秒以

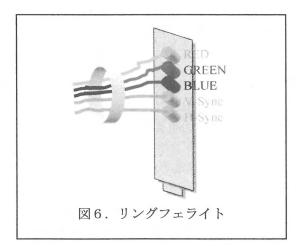
下のジッタとなり、デジタル記録・再生に専用機器として使用可能である。なお、ワードクロックジェネレータを使用して外部同期とすれば、さらにジッタが減少し、録音・再生マスターに匹敵する。

DVD 再生の音声は、動画が同期マスターであり、通常では動画再生の速度が、コマによって時間内に再生されず、大きなジッタを持つ。この改造によって、全体にわたって遅延が減少し、コマ遅延は全く起きず、またボード上のクロックジッタも極めて少ないため、DVD専用再生機より良い結果となった。

以上の結果から、PCの出力信号劣化は、デジタル信号自身の輻射やDACの直線性の低さではなく、給電ケーブル由来のノイズ、振動、マザーボード上のリップルであることが証明された。すなわち、始めから精度の高いグラフィクス・ボードとサウンド・カードを選択し、ノイズ対策を行えば、それらの回路の変更は不要であると言える。

なお、図 5 に示す円内のアナログ配線部が 最もノイズを吸収する部位である。また、マ ザーボードと同じ速度で振動しないと、ジッ





タを生じるのもこの部分である。さらに、本研究によって、画像ケーブルのリングフェライトが著しく画質を劣化させることが判明した(図 6)。しかし、空間ノイズの大きい環境下で水平周波数が 130KHz 以上の高性能の CRT をリングフェライトが装着されていない映像ケーブルで PC と接続すると、大変危険な状態になる。

2m 以下の PC と CRT の接続では、D-SUB コネクタの影響は、高画質用 D-SUB to 5 BNC 変換部を持つ専用ケーブルを使用すれば、大きな劣化はないことが判った。遠距離では、映像ケーブルの品位が極めて大きく画質に影響してしまう。特に、各 BNC 同軸 75 Ωケーブルのインピーダンス・ミスマッチングによる反射と、電気的信号長の違いによる同期誤差が映像を劣化させる。

3m 以上の接続では、AES/EBU ケーブル(110 Ω)や BNC ケーブルでは、高周波特性が劣化してしまい、音声デジタル信号の立ち上がり電圧加速度が低下し、ここにノイズが混入することで、劇的にジッタが増加してしまう。従って、遠距離では光ケーブルの使用が好ましい。

なお、本実験による計測で、鉄筐体の AV アンプその他 AV 機器は、48KHz 以上のデジタル音声信号を入出力する時、多大な輻射を出すことが判明した。そのため、鉄のラックに入れることは、他の機器への影響のため、極めて好ましくない。本実験で作成した PC は、このような輻射は極めて小さく、他の機器への影響はほとんど無く、そのままスタジオ等での使用が可能である。

9 結語

ノイズを特定し排除することによって、Ultra Hi-Fi PC を実現できた。ノイズの排除には、 給電部で電流消費変動の大きい能動素子の直 近で変動を吸収することが最大の効果を示す ことが証明された。結果として PC からの輻 射も測定限界に近づいた。また、方針を明確 に示せた事により、安価に Ultra Hi-Fi PC の作 成が可能となった。 本指針および改造は、他のデジタル機器の 高品位化にも応用可能であった。特に、AV ア ンプ等のように、消費電流の大きい LSI を使 用した機器での向上は非常に大きい。そのた め、再生装置側の高再現性の研究を開始した。

オーバーオールで高再現性を実現した研究および教育の効果は、極めて大きく研究と教育の範囲そのものを変えてしまう。例えば、NTSC 再生では、細部の消失のため、映画の撮影技法などを正確に示せないが、1920×1080ピクセルでの DVD の再生では可能である6。データの精度のみならず、再生環境の精度向上が切望される。

10 謝辞

本研究は、大谷大学が所蔵している北京版 大蔵経等の貴重図書のデジタル化による可読 性を保証した資料共有データベースの作成研 究から支援されている。プロジェクトを推進 しておられる草野教授、山本講師、松川講師、 箕浦研究員以下、学生諸君に感謝する。特に、 京都国立博物館の主任研究員赤尾氏に深く感 謝する。

本研究は、多数の企業によって精度向上の協力を得ている。本研究のような複合的な研究は、多数の専門家による協力がなければ、成果を得ることが極めて困難であることを明記しておきたい。高精度写真撮影および高精度デジタル静止画に関しては、Kodak、株式会社ニコン及び、(株)堀内カラー、大日本スクリーン製造(株)、コメット株式会社に感謝する。CRTの解析に関しては、株式会社ナナオの宮保氏および研究開発の方々に深く感謝する。伝送系ケーブルに関しては、カナレ電気株式会社に感謝する。ノイズ吸収材、パーマロイ、フェライトコア等のノイズ吸収材、パーマロイ、フェライトコア等のノイズ対策に関しては、株式会社トーキン、日立金属、TDKに種々の教えとご協力を頂き深く感謝する。

特に本研究に重要なノイズと有機半導体固

⁶ フィルム媒体では、オリジナルの Region の DVDでないと、フィルムのデュープリケーショ ンでの劣化と、圧縮の差で、劣化が激しいこ とに注意が必要である。

体電解コンデンサ特性に関して、佐賀三洋工業株式会社の喜多川氏と三洋電子部品(株)の三好氏に御教示と御協力を頂き、深く感謝する。高性能アルミ電解コンデンサに関しては、ニチコン株式会の藤馬氏(技術部)にご教示を頂き、同社営業統括部4課の湯地氏と平地に御協力を頂き、深く感謝している。

最後に、非破壊検査株式会社の宮部氏にノイズ測定に関して多大の御教示と御協力を頂き、深く感謝する。

11 参考文献

- [1] 柴田, et al. 北京版チベット大蔵経の高デジタル画像化:写真撮影過程. 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」第6号, pp.72-79, 1998年10月.
- [2] 柴田,et al. 北京版チベット大蔵経の高再現性デジタル画像化:高精度スキャニング過程,情報処理学会「人文科学とコンピュータ研究会」第43回,「人文科学とコンピュータ」,43-6,pp.43-50,1999年7月16日.
- [3] 片岡, Hi-Fi デジタル画像の作成: 明度と 色調の1段階での補正,情報処理学会,「じんも んこん2000」2000 年12月15-16日.
- [4] 片岡, 高精度デジタル画像の高再現性表示, 京都大学大型計算機センター第64回研究セミナー, 口頭発表及び報告書(ISSN 0910-3201), pp.74-85, 2000年3月(24日).
- [5]トランジスタ技術 SPECIAL No.64,1998 年,CQ 出版社.

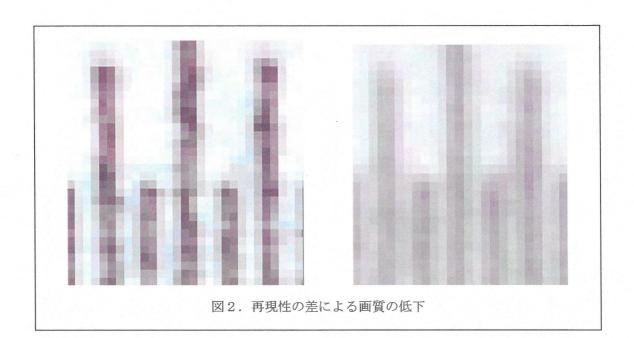
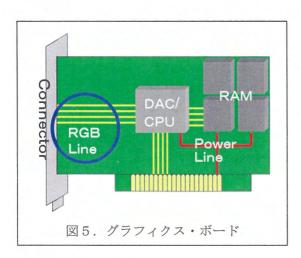


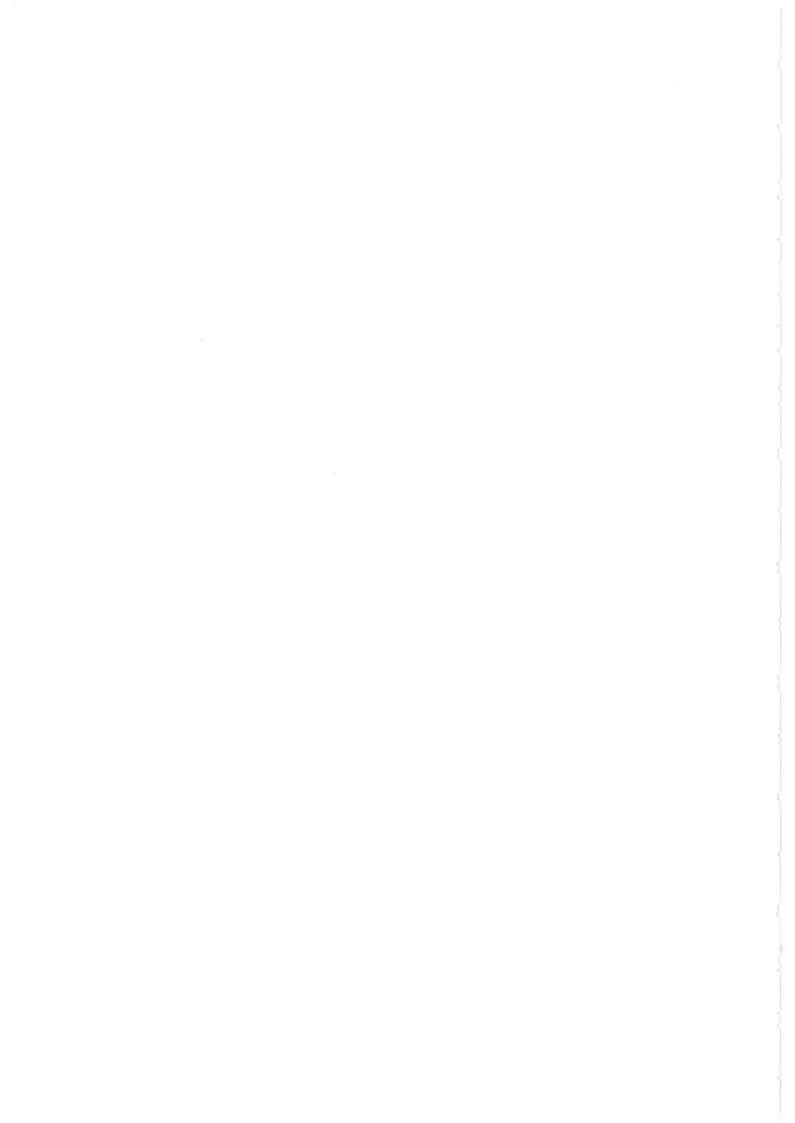




図4. 未対策 PC(上)と Ultra Hi-Fi PC(下)の比較(1920×1440): 同一部品構成







全国遺跡データベースの構築 2001年度の動向 Database of Archaeological Sites in Japan, 2001

森本 晋 Susumu MORIMOTO

奈良文化財研究所文化財情報研究室、奈良市 National Research Institute for Cultural Properties, Nara Nara 630-8577

あらまし: 奈良文化財研究所では1987年度より全国遺跡データベースの構築を行っており、1998年度からはインターネットを経由しその成果を公開している。2001年度は位置データの入力とともに遺跡の定義に関する研究も進めている。

Summary: National Research Institute for Cultural Properties, Nara (ab. Nabunken) began to study about the database of archaeological sites in Japan in 1987. More than 400 thousand sites are existed in Japan, and Nabunken has registered about 240 thousand of sites in the database. Now one can get the data via the Internet home page of the Institute (http://www.nabunken.jp).

キーワード: 遺跡 データベース、考古学、位置情報、定義

Keyword: site, database, archaeology, location, definition

1 「全国遺跡データベース」とは

考古学にとって重要な情報は、遺跡、遺物、文献である。その筆頭として、研究上も埋蔵文化財行政にとっても遺跡に関する情報は、必要不可欠なものである。

1987年度より調査検討が始まった文化財データベースの一環として、全国規模で網羅的な遺跡データベースについて、奈良文化財研究所(以下、奈文研)が設計・構築を行っている。

構築の経緯については、昨年発表している通りで、 ある。現行システムでの一般公開は1999年9月28日 より行っており、利用者は登録や課金なく自由に検索が可能となっている。

構築初期の遺跡地図データに加えて、各地の新規 の遺跡地図や文化財地名表、古墳の集成といった各 種遺跡地名表から情報収集を行い、レコードの追加 や内容の更新を行っている。また、遺跡の発掘調査 報告書に添えられている報告書抄録の内容によるデー タベース更新を合わせて行っている。

2 内容

全国遺跡データベースの設計については、『埋蔵 文化財ニュース』75号に詳しい。

基本的な考え方は次の通りである。奈文研はデータの核となる部分を提供する。都道府県や市町村などの遺跡を調査する機関にとって必要な独自項目があれば、ローカルシステムを構築して、全国システムとリンクさせて利用する。

研究者個人が必要とするような独自のフィールドについても個々の研究者においてローカルシステムとして構築してもらうのが最も効率的なやり方であると考えている。

昨年度と同じく今のところ、全国遺跡データベースは次の27フィールドで構成している。

ID

市町村ID

種別

名称 (漢字)

名称 (かな)

所在地コード

所在地

境界文化財ID

所有者種別

主な時代

指定区分

調査・発掘の有無

文献の有無

地形図番号

緯度

経度

時代·遺跡種別

立地

現況

保存状況

面積

遺跡地図番号

群集遺跡ID

遺構概要

遺物概要

発掘概要

その他概要

3 遺跡の定義

遺跡とは何かの定義には明快でない部分があり、 それがデータベース構築の妨げとなっている面がある。

新しい時代について、どこまで遺跡と認定するか機関によってばらつきがある。太平洋戦争関係などの構築物については、遺跡と認めるか認めないかいろいろな立場があるようだ。そこまで新しくなくても、明治時代や江戸時代の遺跡を、遺跡として認定しているか否かは一定していない。

古い時代に関しても、評価の仕方は様々である。 遺物が若干散布しているだけで遺跡と認定されていることもあれば、わずかな遺物分布では遺跡と認定されていないこともある。丘陵地において地表の隆起が古墳に近い形状をしている場合に、これを積極的に古墳として登録している所もあれば、そうでない所もあり、全国的にみれば統一はとれていない。

広域にわたる遺跡についても、遺跡として登録している場合とそうでない場合がある。たとえば、条 里地割や街道などがこれに相当する。

4 1件の遺跡とは

遺跡の数の数え方にもむつかしい問題がある。数え方とはデータベースにおいては何をもって1レコードとするかである。1件の定義がむつかしい遺跡には複合遺跡、広域遺跡、群集遺跡がある。逆に比較的容易なものに、古墳、窯跡、寺院跡がある。

4.1 古墳の場合

古墳は一般にわかりやす墳丘を持っており、その 範囲も確定しやすい。古墳と他の遺跡、たとえば住 居跡や中世の墓などが重複することもあるが、それ らとの区別は困難ではなく、古墳を古墳としてひと つの遺跡と認識することは容易である。

だが、上記のように断言できない問題が存在する。 古墳で遺骸を埋葬した部分を主体部と呼ぶ。多くの 古墳で、主体部は複数存在する。埋葬施設として、 ひとつのはっきりとした遺構をなしている横穴式石 室においても、複数回の追葬が行われることが普通 にみられる。一般には何体埋葬されていても、墳丘 がひとつであれば、ひとつの古墳として認識されて いるようである。1体しか埋葬していない古墳も、5 体6体埋葬している古墳もどちらも1レコードとする ということである。

しかし、埋葬は墳丘上に行われるとは限らない。 古墳を囲む周溝の中に埋葬が行われることがある。 古墳が孤立していれば、こういった周溝内埋葬がど の古墳に所属するのか決定しやすい。複数の古墳が 隣接していて、周溝を共有するような場合は、周溝 内埋葬をどちらに含めて考えるのか、という問題が 生じる。

また、墓壙が墳丘から離れて単独に築かれることがある。墳丘にごく近い位置であれば、その古墳に属するものと解釈できるが、距離が離れれば、古墳に帰属させてよいものか判断がわかれるところである。

同じような型式の主体部が墳丘上にあれば、その 古墳の構成要素とされ、当該古墳のデータとして処 理されるのに、墳丘からはずれているために同等に は扱うことができない。

では、この埋葬は単独で1レコードとして取り扱うのであろうか。そのような例は少ないように思う。 〇一古墳群、〇一古墳群1号墳、〇一古墳2号墳と登録していき、墳丘外の埋葬は1号墳や2号墳には所属させずに、古墳群に入れて処理する、あるいは、〇 ○遺跡を古墳群と重複させて定義してそれこに含めるといった扱いがなされているようである。

古墳とよく似ている遺構のひとつに、方形周溝墓がある。一般には方形周溝墓は、1基ごとに遺跡として登録することはない。方形周溝遺構としてとらえられているものには、墓と断定できないものが数多くある、というのもひとつの理由であろう。また、周溝を共有して、ひとつひとつの周溝墓の範囲を確定しづらいものがある、という理由も考えられる。

研究史から見て、方形周溝墓が認定されたのが新 しく、そのために古墳とは扱いが違っているという 側面もあるのではないか。

4.2 複合遺跡の場合

種別や時代の異なる遺跡が、同一の地理的範囲内で重なっている場合、複合遺跡と呼ぶことが多い。 ただ、何をもって複合遺跡とするのか、厳密な定義 かある訳ではないようだ。

複合遺跡について、それぞれを別々に遺跡として登録している機関もあれば、そうでない機関もある。どういう場合には同一遺跡と判断し、どういう場合には別遺跡とするのか、について判断が必要であるが、その基準は統一されていない。

たとえば、京都府向日市では、鶏冠井(かいで) 遺跡と長岡京とはその範囲が重複している。重複し ている地域で弥生時代の遺構や遺物が発見されれば それは、鶏冠井遺跡に属すると判断される。長岡京 期の遺構や遺物が発見されれば、それは、長岡京に 属すると判断される。平面位置が同一でも、発見さ れたものの属性によって所属遺跡が異なる。

平城宮跡の下層からは、弥生時代や古墳時代の遺構・遺物が発見されることがある。神明野(しめの) 古墳のように独立して扱われている遺構もあるが、 多くは「平城宮跡下層」と呼ばれている。

複合している遺跡の一方が、宮跡のように地理的・時間的に明確な範囲を持っている場合には、発見された遺構・遺物の帰属について、判断が比較的容易である。また、一方の遺跡種別が古墳のようにわかりやすい場合は、同じ時期であっても、古墳と住居跡のように弁別はむつかしくない。

集落跡が重なり合っている場合、同一遺跡とみなすのかどうか、判断がより困難になる。認識しやすい時間差がある場合、たとえば、縄文時代の集落と中世の集落が範囲を重複している場合には、それらを別々の遺跡としてとらえることも比較的容易である。しかし、古墳時代前期の集落と古墳時代後期の

集落が一部重複していても、一般には別の遺跡とは とらえない。遺跡の連続性をどのように判断するか が問題となる。

遺跡の調査が進んでいる時には、遺構や遺物の分布範囲が詳細に知られて、時代・時期による遺跡の範囲の変遷がわかる。すなわち、調査が進んでいなかったり、遺跡の保存状態が悪い場合には時代・時期別の遺跡範囲については手かがりが乏しく、複合している遺跡を「解きほぐす」ことは不可能となる。

たとえ遺構の保存状態が良くても、発掘調査で見 つかるすべての遺構を、小さな穴のひとつひとつま で、時代・時期を決定することは実際には不可能な ことが多い。

遺物についても、破片が小さい、あるいは、特徴 的な部分ではないといった理由で、時期の決定が困 難なこともよくある。

4.3 広域遺跡の場合

広域にわたる遺跡として、条里地割や都城の京域を例にあげることができる。これらの遺跡について全体をひとつの遺跡、すなわち1レコードととらえることの是非が問題となる。

広域遺跡は、面積が広いために当然、種別の異なる遺跡が重複する率が高くなる。平城京の中には古墳もあれば弥生時代の遺跡もある。

平城京域内で、かつ、奈良時代に限っても遺跡の 複合が発生する。たとえば、西隆寺は奈良時代後期 に建立された寺院であり、その寺域は奈良時代前期 には一般の邸宅地であった。したがって、その範囲 内の奈良時代前期の遺構については、む西隆寺下層 とか西隆寺以前と呼ぶことはできても、西隆寺とと らえることは正しくない。

平城京をひとつのレコードととらえ、かつ京内の 寺院もそれぞれに、ひとつのレコードととらえるの が適切であろう。では、寺院以外の部分はどうであ ろか、広い面積を発掘して「長屋王邸」と推定して いる地域は、その部分としての高いまとまりを示し ているので、別レコードとすることも考えられる。 東市なども同様である。さらに、行政上の扱いから すれば、史跡指定されている宮跡庭園は、ひとつの 遺跡として独立して扱うべきであろう。

都城においては、条坊道路と都市の区画(坪)それぞれを、別々に登録するという考え方も可能である。「平城京左京三条二坊五坪」でひとつの遺跡とする、といった具合である。もちろん、その場合においても道路の交差点の扱いをどうするのか、決め

なくてはならない。全体のレコード数が多くなるという欠点もある。

別の種類の広域遺跡の例として、道路をあげることができる。たとえば、江戸から京都まで連なる「東海道」は、ひとつのレコードでよいのだろうか。各地での調査について、データベースの中に記述することを考えれば適切ではなかろう。しかし、東海道が通過する、すべての市町村単位で分割して、それぞれ1レコードとして登録するのが、はたして最良のやり方かは検討が必要である。

条里地割や範囲を確定しづらい、埋没している水田などの耕作地といった、非常に広範囲にわたる遺跡、水辺の祭祀跡や狩猟場といった範囲の決定が困難な遺跡についても検討を深める必要がある。

4.4 群集遺跡の場合

群集遺跡の扱いも機関によっていろいろである。

古墳群を例にすると、10基からなる古墳群があったとして、1)全体をひとつのレコードとしている場合、2)10基それぞれをひとつのレコードしている場合、3)10基それぞれに加えて古墳群も1レコードをなしている場合、4)古墳群中の主な古墳のみに別途レコードを割り当てている場合、というように取り扱いには統一がない。全体のレコード数は1以上、11以下となる。

大きな古墳群が小支群に分かれている時には、組 み合わせがいろいろと考えられるので、さらに複雑 になる。

遺跡の数え方のところで述べたように、古墳はひとまとまりを認識しやすい遺跡である。よって、個々の古墳それぞれを1レコードすることは重要であると考える。しかし、前述のように、古墳と古墳の間には、発掘調査しなければわからない、関連する遺構が発見されることも多い。上記3)のように、個々の古墳に加えて古墳群全体もひとつのレコードとして記述するのが、よりよい方法であると考える。

このやり方を採用した時は、古墳が何基あるかといった検索に対して正しい結果が得られるようにデータの記述の仕方に工夫が必要である。

須恵器や瓦などの窯跡についても、古墳と同じように範囲の認識が比較的簡単と考えてよいので、レコードの単位も同様に処理してよいのではないか。

古墳や窯跡が1基ずつ認識されるのは、それらが 遺物研究の単位となっていることからも望ましいこ とと考える。

5. 今後に向けて

遺跡の登録を遺跡データベース側から眺めると様々な改良が必要なことがわかってきた。その多くが「古くて新しい」問題であり、簡単に解決できるような性質のものではない。

考古学の研究にとっても、行政にとっても、遺跡・遺物・文献の3分野のデータベースを充実させることは緊急かつ重要な課題である。遺跡データベースを充実させていくとき、「遺構データベース」や「調査データベース」の構築が必要となってくる。

遺構データベースは、方形周溝墓や竪穴住居跡といった、遺構ごとに作成されるのが実際的である。

調査データベースは、平城宮跡第201次調査といった、調査の単位を1レコードとして作成されるべき ものである。

ただ、遺跡データベースと完全に独立させた形で 「遺構データベース」「調査データベース」を作る ことも困難であり、協調をとる何らかの方策を練ら なくてはならないであろう。

遺物・文献のデータベースを含めて、多くの人々の努力によって全国に蓄積されている、全体的な考古情報を活用するための努力は、当然、奈文研が単独で成しうるものではない。地域分担などの「交通整理」と全体を把握するための、考古情報交換標準の研究、情報の情報たるメタデータの整備、その公開のためのクリアリングハウスの設置が、いよいよ求められている。

参考文献

奈良国立文化財研究所埋蔵文化財センター 『埋蔵 文化財ニュース 75 全国文化財データベースにつ いて』 1992.9

奈良国立文化財研究所埋蔵文化財センター 『埋蔵文化財ニュース 97 遺跡GISと遺跡情報』 2000.5

森本 晋 「全国遺跡データベースの構築」 公開シンポジウム「人文科学とデータベース」2000.12 森本 晋 「Database of Archaeological Sites in Japan」 Asia GIS 2001 2001.6

日本語キェルケゴール文献データベース Database for Kierkegaard Literature in Japanese

橋本淳、平林孝裕 Jun HASHIMOTO, Takahiro HIRABAYASHI

関西学院大学神学部

School of Theology, Kwansei Gakuin University

662-8501 兵庫県西宮市上ヶ原一番町 Uegahara-ichibancho, Nishinomiya, Hyogo, 662-8501

あらまし:「日本語キェルケゴール文献検索システム」は、1906 年から 1985 年に至る日本語キェルケゴール文献(原著書の翻訳ならびに研究文献ほか)の全容をふくむ文献検索システムである。その特徴は、第一次文献と第二次文献の有機的な連関が考慮した立体的な検索を可能にした点にある。又、主要データには、研究史に担う意義が付されている点でも、きわめて有用なデータベースとして構築されている。

Summary: "Database for Kierkegaard Literature in Japanese" is a search engine for Japanese translations of Kierkegaard's works and the literature on Kierkegaard available in Japanese covering the period 1906 through 1985. To provide maximum usefulness as a database the system was designed with a search engine that is capable of performing multi-level relational searches between primary and secondary sources. Additionally, for major entries the database provides a useful commentary on its significance.

キーワード:日本語キェルケゴール文献、文献データベース、立体的検索

Keywords: Kierkegaard literature in Japanese, literary database, multi-level search

序.

19 世紀デンマークに生まれた天才、キリスト教思想家セーレン・キェルケゴール (Søren Kierkegaard, 1813~1855) は、20 世紀初頭に再発見されるととも

に、その思想の予言者的な性格ゆえに高い関心を呼び、 のちのヨーロッパ精神史の歩みに大きな足跡をのこした。その影響は、キリスト教神学にとどまらず、哲学、 心理学、そして文学へときわめて広範な領域で見出す ことができる。

キェルケゴールの影響は、ヨーロッパにとどまらず 我が国でも著しいものがある。日本におけるキェルケ ゴール受容は、明治 39 年にまでさかのぼり、以来、 わが国では、その質量ともに、欧米に比肩しうるキェ ルケゴール文献が認められる。

1. 「日本語キェルケゴール文献検索システム」

日本語キェルケゴール文献に関する目録作成は、かねてより萬里小路通宗氏の手ですすめられ、その成果が逐次、公表されてきた¹。1984年に関西学院大学では、キェルケゴール文献システム作成共同研究会を組織し、同氏の長年にわたる成果と共に、これに新規データを加え、日本語キェルケゴール文献に関する綜合的なコンピュータ検索システムの開発に着手した。

「日本語キェルケゴール文献に関するコンピュータ検索システム」² (以下、キェルケゴール文献検索システム)は、関西学院大学および文部省からの研究補助を得て、1987年に一先ず完成をみ、すでに実用に供されている。

この「キェルケゴール文献検索システム」は、1906年(明治39)から1985年(昭和60)に至る日本語キェルケゴール文献の全体を網羅する我が国最初の綜合的なキェルケゴール文献検索システムである。所収文

献の総数は、1427件にのぼる。入力されたデータは大別して二つのカテゴリーに分類される。

A. 原著作の翻訳 (177件)

キェルケゴールの公刊著作あるいは日誌・遺稿集及 び手紙などキェルケゴールの第一次資料に関する 日本語訳(単行書、単行書の一部としての所収のも の、雑誌に所載のものなど)

B. 研究文献 (1250件)

キェルケゴールに関係して日本語で書かれた第二次的文献・資料(単行書、単行書の一部として所収のもの、雑誌・月報所載のもの、新聞記事、書評、文献目録、年表、会議報告、また学術大会における研究発表要旨など)。さらに諸外国の研究書・雑誌論文の日本語訳もこれに含む。

また、検索システムの特徴は、「利用者のためのガイドブック」³に紹介されているように、以下の点に見られる。

- ① 明治 39 年から昭和 60 年に至る、日本語キェルケゴール文献の全容(翻訳・研究・書評・文献目録・年表・新聞記事ほか)が、綜合的に把握できる。
- ② 漢字表示によるメニュー方式でできており、コン ピュータに不慣れな研究者であっても、容易に利 用可能なシステムであること。
- ③ 著者検索・項目検索等の一般的な用途の他に、第 一次資料と第二次資料の相互の関連検索が可能 となる、立体的なシステムになっていること。
- ④ 主要データには注記が付されて研究史に担う意 義が解説されているほか、初出の雑誌論文がその 後に単行書へと展開する経緯等も指示される。

2. 検索の実際

以上のような特徴をキェルケゴール文献検索システムは有するが、実際の検索方法には次の5種類が用意されている。

- (1) 原著作の日本語翻訳に関する検索、
- (2) 研究文献の検索
- (3) 内容項目に関する検索
- (4) 著者(訳者)に関する検索
- (5) 関連検索

(第一次資料と第二次資料の相互関係に かかわる検索) 検索のための初期画面は、(1)原著作の翻訳、(2)研究 文献、それぞれに関する検索を選択するか、または、 現著作の翻訳と研究文献の全資料について著者(訳者) 名からの検索ができるよう設計されている(図1を参 照)。

キェルケゴールの原著作は、文学から神学ならびに哲学・心理学などの多様なジャンルに及び、さらに彼が遺した膨大な日誌記述・遺稿が、これに加わる。したがって原著作の翻訳に関するデータベースはこのような事情に対応したものでなければならない。そこでキェルケゴール文献検索システムでは、原著作を9つのカテゴリーに分類し(図1のAの右を参照)、カテゴリーを絞り込んだ上で、著者(訳者)名により検索を進めることができる。また、この原著作の検索には、利用者の便宜を考慮して、その他として「キェルケゴールの言葉」「キェルケゴールの祈り」「キェルケゴールの引用語句、モットーなど」が含まれている。

研究文献の検索は、主として内容項目による検索となっている(図1のBの右を参照)。キェルケゴールの伝記的な問題について詳細なメニューが準備されているほか、思想については美学・哲学・宗教思想・政治社会といったテーマから、利用者はその関心に従ってさらに個別的なテーマへと検索を行うことができるよう工夫されている(図2. コード表の例を参照)。

著者(訳者)名による人名検索では、姓をカタカナ 読みで入力して行うが、また姓の一部分からでも検索 が可能である。また姓・氏名(例えば「スギヤマ ヨシム」) を同時に入力して検索することも可能である。

しかしながら、この文献検索システムのもっともすぐれた点、実際にキェルケゴール文献を検索する上で 利用者に有益な点は、検索システムの立体的な構成に ある。

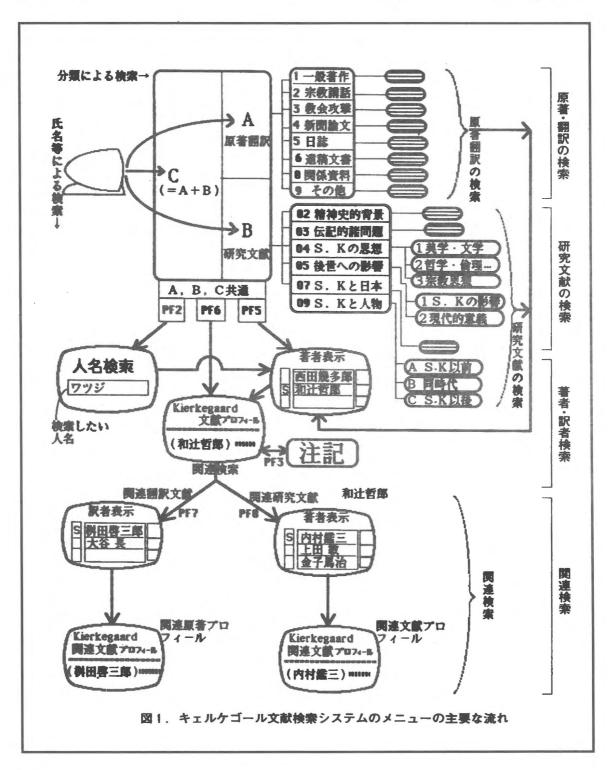
一般に、文献検索システムが構築される場合、第一次資料(著作などの翻訳)と第二次資料(研究文献など)とがそれぞれ独自のシステムとして設計され、この両者の有機的な連関を図ることは疎外されがちである。「キェルケゴー文献検索システム」では、当初からこの二つを一体化して統合的な検索を可能にするよう意図され、第一次資料と第二次資料とが相互的に関連検索が可能となるよう設計されている。

したがって原著作の翻訳に関する検索の結果から、 直接にこれと関連する研究文献をさらに検索すること が可能であるし、また逆に、研究文献の検索からその 内容に結びつくキェルケゴールの著作種別へと関連検 索も可能となっている。この目的のために、研究文献 ほかの第二次文献は、その内容を精査し、これに関連 性を有する著作などの第一次資料へと接続された。

以上のように、キェルケゴール文献検索システムは、 簡便な操作によって利用者に近づきやすいだけでなく、 その内容においても、キェルケゴール理解の深まりに おおいに寄与するものである。

3. 国際版データベースの試み

これまで関西学院大学共同研究会による「キェルケゴール文献検索システム」について報告してきたが、これは日本語での利用を前提において構想されている。その一方で近年、キェルケゴール研究がますます国際化し、海外との研究者との研究協力・交流がさかんになると共に、明治以来の豊かな日本におけるキェルケ



ゴール研究の成果を広く海外に紹介するようにとの要望もたかまりつつあった。オーゼンセ大学(現・南デンマーク大学)のモーテンセン(Finn Hauberg Mortensen)教授は、このような日本のキェルケゴール研究の紹介を試み4、そのなかで主要な日本のキェルケゴール文献の目録を所収したい旨を強く希望された。これに呼応する形で再び共同研究会が組織され、1994年から1995年にかけて日本語キェルケゴールデータベースの国際版が準備された。その際の用語は、キェルケゴール研究における国際共通語ともなっているデンマーク語が採用された。これらはテキストファイル(または印刷物)として頒布された。(したがって、これには現在、検索システムは付属していない。)データとして収録された内容は、

- ① 原著の日本語訳:1930-1979年
- ② 研究書(単行書と単行書の一部として所収 されたもの):1915-1994年
- ③ 雑誌論文:1906-1979年

ただし、これは既存のキェルケゴール文献検索システムに収録されたすべてではなく、書評・新聞記事などは収録されなかった。

4. 課題と展望

今後の課題としては、まず「日本語キェルケゴール 文献検索システム」について、既入力の日本語データ ベースを再点検し不備を是正すると同時に、昭和61 年以降の新規データを補完することが求められる。

さらに日本語文献データベースを整備した上で、国際版データベースの完成を期すべきものと思われる。

いっぽう今日の情報通信技術の進展は目を見張るものがあり、その成果が、また考慮されなければならない。すなわち、現在、「日本語キェルケゴール文献検索システム」は、大型計算機を利用して TSS 端末機から利用する形態で運用されているが、これをそれぞれの利用者のパーソナル・コンピュータでの利用での前提とした改良システムへ、またはインターネットを介して WEB をもちいた検索システムに移行することが必要となってこよう。

- 1 『キェルケゴール研究』 (キェルケゴール協会), 創刊号, 第7号, 第8号, ほかを参照。
- ² 1984 年度関西学院大学共同研究費および昭和 61 年 度文部省科学研究費補助金による研究成果である。 また、その作成にあたっては(株)アシストの協力を 得ている。
- 3 関西学院大学キェルケゴール文献システム作成共同 研究会編「日本語キェルケゴール文献に関するコン ピュータ検索システム利用者のためのガイドブッ ク」昭和63年.
- ⁴ Finn Hauberg Mortensen, Kierkegaard Made in Japan, Odense: Odense UP, 1996.
- ⁵ Søren Kierkegaar Litteratur i Japan, 1906-1979, redigeret ved Jun Hashimoto, Michimune Madenokoji og Takahiro Hirabayashi, 1995.

内容コードの形式: 〈内容コード(1) 2 析〉 + 〈内容コード(1) 1 析〉 + 〈副コード 3 析〉 「キェルケゴールの思想」 〈内容コート*(1) 04〉 から検索するケース (図 1 の B の右のコート*)

]- °	内容コード(2)	1
ı	1	美学、文学など	
	2	哲学、倫理思想、論理学など	\rightarrow
	3	宗教思想(キリスト教理解)	
	4	政治・社会に関するもの	
	9	上記以外	

図 2. 研究文献検索のための コード表の例

内容コー	内容コード(2): 2の副コード表				
副小,	内容	副叶	内容		
001	哲学者としてのキェルケゴ ール、哲学的著作	009	自己、精神		
002	ヘーゲルとの関係・批判	010	倫理的実存、倫理思想、倫		
003	実存概念、実存の三段階、 実存弁証法、実存的思惟、		理的実存、決断・選択、誠実		
	生成	011	例外者、単独者、個と普遍		
004	時間論、永遠、歴史、瞬間	012	デモーニッシュなもの、悪		
005	真理、主体性、現実性、 内面性、可能性	013	論理学、逆説、懐疑		
006	死、不安、苦悩、自由、 ニヒリズム、無、人間学、	014	伝達の方法、間接的伝達、 ソクラテス的なもの		
	絶望	015	教育、学問		
007	反復	_			
008	イロニー、フモアー	999	上記以外のもの		

前方後円墳データベース検索システムと地理情報処理

Information Retrieval from the database of Japanese Ancient Tombs and Geographical Processing

西上昌治 小澤一雅 Masaharu NISHIUE and Kazumasa OZAWA

大阪電気通信大学大学院 〒572-8530 大阪府寝屋川市初町18-8 Graduate school, Osaka Electro-Communication University, Neyagawa·shi, Osaka 572·8530, Japan.

あらまし

本論文では前方後円墳データベースのための検索システムとそれにかかわる地理情報処理について述べる。最初に前方後円墳データベースを構成する主な項目について説明する。次にインターネットに接続することなくローカルマシンで有効に活用するための検索システムについて述べる。最後に前方後円墳に関する研究に必要と思われる地理情報処理機能について述べる。

Summary

This paper presents an information retrieval system for the archaeological database of Japanese ancient tombs and its related geographical processing. First, important items (or attributes) of the database are illustrated. Second, the query expression specially designed for the retrieval system is presented. Finally, discussion is made on what types of geographical information processing should be needed for archaeological studies on the Japanese ancient tombs. A tentative work for the geographical processing is presented as an example.

キーワード 考古学、データベース、古墳、検索エンジン、地理情報処理

Keywords archaeology, database, ancient tomb, information retrieval, geographic information system (GIS)

1 はじめに

前方後円墳は約1500年前に各地の豪族たちの権力の象徴としてつくられた巨大な墳墓であって、北海道・青森・秋田・沖縄を除く43都府県に遺存している。

本研究の対象である前方後円墳データベースは約5800基の古墳を収録しており、約9割を前方後円墳が占めている. 古墳データは1基ごとに、時代を示す"築造推定時期", 古墳の形を示す"古墳形態", 古墳の大きさを示す"7部位", 古墳の遺存している場所を示す"都府県", "所在地", "緯度・軽度"などの項目を設け各データを原簿カードに記述している. データベース化にあたって、最低限必要とされる41項目について電子化作業を進めている.

1995年より、本データベースは Web 上で公開している[1,2]. しかし、現状ではすべてのマシンがインターネットに常時接続できる状況にあるわけではない。そこで、Web に関係なくローカルマシンで前方後円墳データベースを利用できる環境も依然必要と考えられる.

ただし、データベースから所望の古墳にかかわる情報を得るための検索をローカルマシンのみで行うには、Web上でよく見られるテキストフィールド型の入力欄のみの検索方法では効率がわるい。しかし、インターネットブラウザで扱える Web 上での検索エンジンと同様に、できる限り OS に依存しないものがよいと考えられる。本研究ではこうした点を考慮した新しい Java を用いた検索システムについて述べる。

2 前方後円墳データベース

2.1 古墳データ

古墳データは参考資料[3,4,5]をもとに原簿 カード(図 1)を記入することによってデー タの整理を行ってきた. 古墳データは1基の 古墳に対して 41 項目を設定している(付録 表 1). 主要項目を以下に記す.

古墳名

古墳名を漢字で記述し,またカタカナで読みを示す. さらに将来検索システムの英語化

を見越してローマ字表記も併用する.

緯度·経度

緯度・経度は,各基ごとに後円部の中心の 位置をデジタイザで測定した計測値である.

推 副 图			14:3	5 E :	29 2 2	71
271	13	25 27201	28 34	电	3394	2
2 4 8	1+	25 17 95 JA 26	# 27		almiter:-	dider
古場	名	仁愧天皇陵	大仏		#	
英雄	名					
所 在	所 在 岫 堺市大仏馬					
古 坳	蟾	新台馬古遺跡				
古場別	劍	南方後丹墳	立地	A	丸	
*	素	克·無	方		h	
隋	鄮	@:景	吳峰底		~	
賽	石	6 ·無	推定時期	*	訓	
過	进	の 銀・銀・	增丘長	48	£ .	m
埴	輸	②· 號	前方配模	30	0.	m
		門筒、人馬	前方部高	3	4.	ini
查	10	·無	後円都径	24	4.	m
		德、金纳联刀、甲胄、引汉国	後門都高	3	6.	m
		ጀር ት, ከ ሃት	《LL本語編	16	3.	m
			〈以相解唐	:	25.	m
			外 雅			m
石材	貫	有·無	周 液	1	₹ 3	1
石材	建地		745 FER.	錘	,	m
₹ 5	室	**	集成段数	前方	部 4	段
(1	置)	聖欠会(前部)	殷	鄉	1個 4	股
(方(2)		参考文献			
欄		有·無	巨大白塊			
(88	E)					

図1 原簿カード

古墳形態

古墳形態は円墳や方墳などシンプルなものも存在するが、図2の前方後円墳、前方後方墳、帆立貝式古墳、双方中円墳、上円下方墳の5種を扱っている。

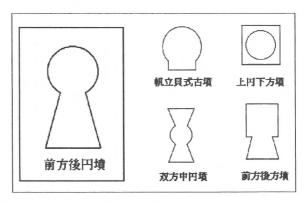


図 2 古墳形態

基本形状

前方後円墳の形を考えるために、その基本 形状を決定づける7部位(図3)が重要であ る。それらを測定することによって築造推定 時期(表1)の推測をすることもできる。

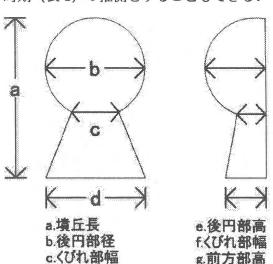


図3 7部位

その他の項目

d.前方部幅

他のデータ項目として、古墳の向きを示す "方向",巨大古墳に近接する小規模な古墳で、その巨大古墳に関連して営まれたとされる "倍塚",古墳の表面の装飾にかかわる"葺石" や "埴輪", 古墳内部の "遺物" や "石室" などの項目が存在する.

表 1 築造推定時期

	分類	時期
	4 世紀	前期
1.0	4世紀~5世紀	前期~中期への移行期
e	5 世紀	中期
	5世紀~6世紀	中期~後期への移行期
f	6 世紀	後期

2.2 前方後円墳データベース

原簿カードをもとに電子化支援ツールを用いて電子化したものである. 古墳データの 41項目をそれぞれ "#"で区切り記号としている. また, 1項目内に複数のデータを記述する場合は";"を区切り記号としている. ファイルは1行(1レコード)に1基の古墳データを記述する方式でテキスト形式であり, 北海道,青森県,秋田県,沖縄県以外の 43都府県をそれぞれ別のデータファイルとしている. 現在は古墳データの最終修正および点検を行っている段階である.

27001#モリガミ;キツネヅカ#森上;狐塚#Morigami_tomb;Kitsunezuka_tomb#0#大阪#豊能郡能勢町森上吉芝44#27322#LBA#345743#1352406#妙見山#帆立貝形古墳#0#有#0#0#有#無#0#丘陵頂#南#0#後期#26#10#0.8#17#3#9#0#0#0#0#0#0#0#0#0#遊佐和敏『所謂「帆立貝式古墳」地名表』(1982年)。

27002#タジリ#田尻#Tajiri,tomb#0#大阪#豊能郡能勢町上田尻西原 2 1 #27322#。』A#345757 #1352654#妙見山#前方後円墳#0#0#0#有#無#0#0#山頂#南東#0#0#29.6#10.7#2.29#15.8#4.

19#9#0#0#0#0#0#0#0#0#0#0#0

27006#ミナミヅカ#南塚#Minamizuka_tomb#0#大阪#茨木市宿久庄#27211#...A#345022#13532 56#高槻#前方後円墳#0#0#0#有#有#無#有#有#丘陵腹#南#0#後期#50#0#5#30#5#0#0#0#0#0# #0#円筒#横穴式石室#組合式石棺2#0*

図4 前方後円墳データベース(大阪)の一部

3 検索システム

本論文で報告する検索システムでは都府県, 古墳形態,築造推定時期の3つを選択形式, その他をキーワード入力形式の検索方式を採 用した.項目別の論理演算にはANDを使用 している.

3.1 選択検索

全国に対して検索項目を"古墳形態",検索 文字を"前方後円墳"のみの単純な検索を行った場合,膨大な該当データ数が返ってくる ために扱いが繁雑になる。そこで各都府県別 に選択を行えるように今回はチェックボック ス型の都府県選択形式を採用した。しかし, 複数の都府県にまたがる古墳を検索するため には,チェックボックスのみでは扱いづらい と思われる。そこで,該当地方すべてを検索 対象に選択する場合の手間を省くために,地 方別のボタンを付加した。"古墳形態"と"築 造推定時期"はデータの種類の数が決まって いるのでリスト型の選択形式を採用した。

3.2 文字検索

入力する検索文字(キーワード)と対応し

たデータ項目に含まれているかどうかについてパターンマッチを行う。1 項目内で複数のキーワードを入力できるようにし、それぞれの項目ごとに"必ず含む (AND)","何れか含む (OR)","含まない (NOT)"の3種類の論理演算を行えるようにした。

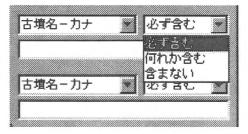


図6 論理演算の選択

3.3 数值検索

7 部位等の数値型の検索では、例えば "257m の墳丘長の古墳を調べたい"というような検索要求は稀である。大方は"100m 以上 400m 以下"のような特定の範囲や、"200m 以上"または"400m 以下"などの検索を行うものと想定した。そこで、それぞれ"100・400"、"200 <"、"400 >"と入力することで数値検索を行えるようにした。

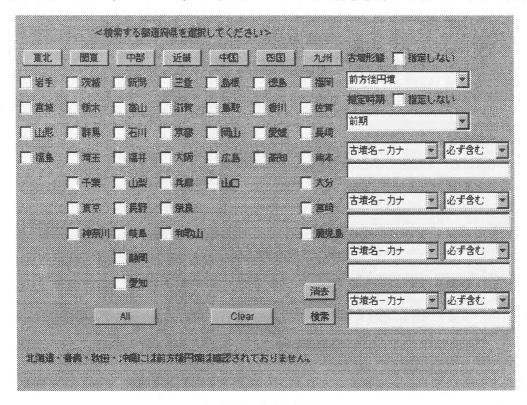


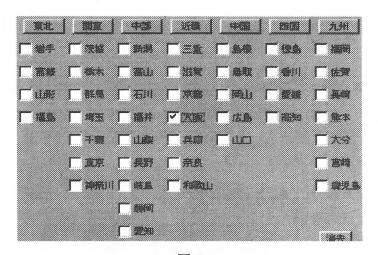
図 5 起動画面

4 検索実行例

実行例として,仁徳天皇陵を検索する.下記のキーワードを用いて検索システムの検索過程 を示す.

キーワード

都府県"大阪府"古墳形態"前方後円墳"古墳名(漢字)"天皇"墳丘長"300m以上"



Step1

都府県のチェックボック ス"大阪"にチェックを入 れる.

図 7

TABLE	福定しない
前方後円墳	.
推動報 🔽	物面しない
前期	
古墳名-漢字	* 必ず含む *
天皇	
墳丘長	● 必ず含む 💌
300 <	

図 8

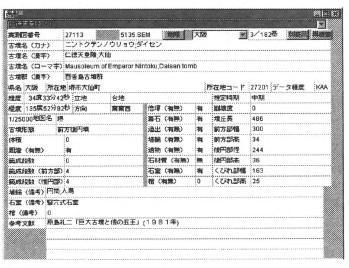


図 9

Step2

古墳形態は"前方後円墳" 推定時期には"指定しない" をチェックし、検索項目を "墳丘長"、検索文字を "300 <"と入力し、"検索 "ボタンを押す.

Step3

該当する古墳がある場合 は左図のように検索結果フレームが表示され、該当する古墳の総合計と都府県別の総合計,1基の詳細データが表示される。内容を表示したい古墳を上部のリストボックスで選択する。

5 地理情報処理機能

該当する古墳データの緯度・経度が存在している場合に限り、その古墳がどの場所に遺存しているかがわかる.しかし、文字による所在地表示のみでは視覚的に把握し難い.この課題については国土地理院発行の数値地図を用いることによって可能とした.以下、数値地図 250m メッシュ (標高)を用いた地理情報処理機能について述べる.

5.1 数值地図

数値地図 250m メッシュ (標高) は 2 万 5 千分 1 地形図を基図にして,等高線から計算によってメッシュの中心点の標高値が収められた数値標高モデルデータ (DEM) である.日本全土を緯度 (南北) 方向で 7.5 秒,経度 (東西) 方向で 11.25 秒ずつに区切り,その土地の高さを数値で表示したものである.実距離は約 250m である.DEM の標高値は,海抜を基準に記録されていて 0.1m 単位で表現されている.1 ファイルにつき緯度 320 個,経度 320 個の計 102,400 個のデータが存在する.

5.2 地理表示

DEM を用いることにより、古墳が一体どの ような地理状況の上に遺存しているかが視覚 的に理解しやすい表現が可能になる. まずは 2次元での地理表示に取り組む.1データを1 ピクセル, つまり縦 320 ピクセル横 320 ピク セルを1ファイルとして表示する. なお,標 高値の色は一般的な地図帳と同様の配色を行 う (表2). また、2次元上でより立体的な表 示とするために左に隣接するデータよりも値 が低い場所を影とする. 古墳を正確に表示し ようとしても、最大で2ピクセルが限度であ るが、逆に1ピクセルでは認識し難いので簡 易的に 2x2 ピクセルの点で表示する. また, 立体的表示機能としてクオータービューの地 理表示を試作する. 1 方向のみの表示ではな く, "北東", "北西", "南東", "南西"からの 表示を可能とし、できる限り高さによる死角 を減らすようにする.

表 2 地理表示の配色

区別	標高地(m)	色
海	-999	青
湖	0~	水色
陸	1~20	緑
	21~50	黄緑
	51~100	薄い黄緑
	101~200	薄い肌色
	201~400	肌色
	401~600	山吹色
	601~1000	薄柿色
	1001~1500	薄茶色
	1501~2000	赤茶色
	2001~3000	茶色
×	3000~3776	こげ茶色

例として大阪府の南部に遺存する仁徳天皇 陵の位置表示結果を示す。



図 10 地理表示 (2D)

図 10 では 2 次元での地理表示の 1 例を示したが、東京湾の平均海面をもとにした水準原点を基準として、各地の海抜高度を測量されているので、国土内にある湖沼は指定しない限り湖沼として表示されない。この例として滋賀県にある琵琶湖などがあげられる。逆に水準原点よりも低い場所は海として表示している。これは図 11 でも同様である。

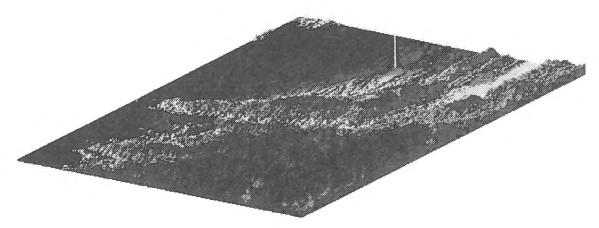


図 11 地理表示 (クオータービュー:南東)

5.3 地理情報処理

まず,該当古墳から半径を指定してその円の範囲に分布する古墳を表示する円内検索がある。例えば,仁徳天皇陵から半径 10km 以内に分布する古墳の表示である(図 12)。一方、2 地点間の断面図を表示する断面表示機能も必要である。この機能の応用として,可視領域表示が可能である。可視領域表示とは,入力された地点から,人が立っている状態で見ることができる地点だけを表示する機能である。

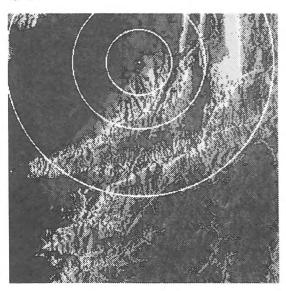


図 12 古墳から半径 10,20,40km の円

6 今後の課題

今回構築した検索システムは、ほぼ当初の目的を達成したが、問題点も残している.

(1) 再検索を行うごとにフレームが増加する.

(2) 検索結果の表示方法は 1 基の古墳データ のみの表示を行うので該当古墳ごとにデ ータを比較することができない.

検索機能の強化を目指して上記の2点を今後 の課題とする.

参考 URL

小澤研究室

http://www.ozlab.osakac.ac.jp/ 古墳検索エンジン

http://www.ozlab.osakac.ac.jp/cgi-bin/kSear ch/search_kofun.cgi

参考文献

- [1]H.Fujita, T.kato and K.ozawa, Towards a Multimedia Database of Japanese Ancient Tombs, *Proc. of VSMM 98*, 281-285, 1998.
- [2]三浦・小澤,考古学データベースにおける 検索エンジンの研究,第 6 回公開シンポ ジウム「人文科学とデータベース」論文集, 73-80,2000.
- [3]小澤 一雅,『前方後円墳の数理』, 雄山閣, 1988.
- [4]近藤 義郎,『前方後円墳集成·近畿編·』, 山川出版社, 1992.
- [5]大塚 初重,『日本古墳大辞典』, 東京堂出版, 1989.

付録表1 古墳データ項目

データ項目	データのタイプ	日頃アーク頃日 例:仁徳天皇陵の場合
実測図番号	数値	27113
古墳名(カタカナ)	文字	ニントクテンノウリョウ;ダイセン
古墳名(漢字)	文字	一つ・ファン・ファン・ファン・ 仁徳天皇陵;大仙
古墳名(ローマ字)	文字	Mausoleum of Emperor Nintoku;Daisen tomb
古墳群	文字	百舌鳥古墳群
県名	文字	大阪
所在地	文字	堺市大仙町
所在地コード	数値	27201
データ精度	文字	LBA
緯度	数値	1332927
経度	数值	343342
1:25000 地形図	文字	堺
古墳形態	文字	前方後円墳
陪塚(有無)	文字	有
葺石(有無)	文字	有
造出(有無)	文字	有
埴輪(有無)	文字	有
遺物(有無)	文字	有
石材質(有無)	文字	有
石室(有無)	文字	有
棺(有無)	文字	不明
立地	文字	台地
方向	文字	南南西
崩壊度	文字	完存
推定時期	文字	中期
墳丘長(m)	数值	486
前方部長(m)	数值	300
前方部幅(m)	数值	34
後円部径(m)	数值	244
後円部高(m)	数值	36
くびれ部幅(m)	数值	163
くびれ部高(m)	数值	25
体積	数值	不明
周濠(有無)	文字	有
築成段数	数值	不明
築成段数(前方部)	数值	4
築成段数(後円部)	数值	4
埴輪(備考)	文字	円筒;人;馬
石室(備考)	文字	竪穴式石室
棺	文字	不明
参考文献	文字	原島礼二「巨大古墳と倭の五王」(1981年)

貝類のマルチメディアデータベースの構築

-西宮市貝類館との共同研究-

Construction of Multimedia Database for Shell

- Collaborative Research with Nishinomiya City Shell Museum -

高田茂樹 · 雄山真弓

Shigeki Takada, Mayumi Oyama

関西学院大学 情報メディア教育センター

Center for Information & Media Studies Kwansei Gakuin University

あらまし:

西宮市貝類館に所蔵されている貝類の標本とその学術的情報のマルチメディアデータベース化、およびそのCD-ROM化、ホームページ化を試みた。本研究の意義は、時とともに劣化する紙や標本の形でのみ保存されている情報をデジタル化して、半永久的に保存することを可能にした点と、それをより多くの人々に公開することを可能にした点である。以下にデータベース化作業の特徴とその流れについて述べる。

Summary:

We have constructed the Multimedia Database for Shell and opened the Database by CD-ROM and Web Pages in this collaborative research with Nishinomiya City Shell Museum. This attempt enabled us to keep the important samples and scientific data into digital data and to retrieve the data. This is the report how we have worked for the Database.

キーワード:マルチメディアデータベース、貝類、画像処理、情報公開

Keywords: Multimedia Database, Shell, Photo retouch, Release of Information

1. 研究の背景:

西宮市内には現在菊池貝類研究館と西宮市 貝類館があり、両施設に所蔵されている貝類標 本は故黒田徳米博士が新種として確定した模式 標本の約230点を含む合計約4万点の膨大なも のである。菊池貝類研究館が1995年の阪神大 震災で物理的に多大なダメージを被り、貴重な コレクションを早急に保存する必要性が生じたこ とを契機にコレクションの整備が進み、1999年に 西宮市貝類館が新設された。我々は同じ西宮市 にある大学として、貴重な標本・資料を保存する ことに協力を申し出、さらにそれを地域の住民に 公開することでいくらかでも地域に貢献したいと の思いから、菊池貝類研究館と西宮市貝類館お よび国内の貝類研究者や専門家との連携を取り、 貝類のマルチメディアデータベースを構築すると いう研究を計画した次第である。

2. 本研究以前の状況:

我々が菊池貝類研究館および西宮市貝類館を最初に訪ねたとき、貝類に関する学術的情報は紙の上に記録されたデータ状態のままで保管されており、阪神大震災下にあったため標本自体の保存状態もあまり良いものではなかった。また、それらの貝類に関する学術的情報は一般に公開されておらず、来館者が自由に貝類についての情報を検索することができない状態であった。

関西学院大学情報メディア教育センターでは、 国内の貝類研究家である山田まち子氏の協力 により、山田氏が作成された微小貝約 800 種の データベースを変換し、1998 年 2 月よりサーバ 上で公開するということを行ってきた。その記録 によると、毎日世界中からの利用があり、そのア クセス数は1年間で延べ約 14,000 件に上った。 このことから、貝類の情報公開を地域の人々の みでなく、世界の人々に広く情報を公開すること も視野に入れて、西宮市貝類館との共同作業に よりデータベースの作成を始めることにした。

3. 研究の目的:

本研究の目的は、西宮市貝類館に所蔵されている標本をデジタル映像化した画像と共に、貝の形状や貝の学名、和名、属番号、綱、亜綱、目、属、亜属、種名、オーサー名、科名、科番号、仲間、国、産地、分布、生息場所、個数、図鑑類、貝の形、提供者などの項目を、日本語と英語によりデータベース化することである。また、構築したデータベースを、インターネット経由で検索できるシステムとして公開することにより、世界中の科学教育や貝類に感心のある人々に、貝類に

ついての正確な情報を提供することも目的であ る。

4. 研究の意義:

貝類に関する学術的情報をマルチメディアデ ータベース化して保存することは、大切な標本 や学術的情報を半永久的に守るという意味で学 術的意義のあることであり、さらにデータベース 化することで保存したデータを来館者に公開し、 自由に検索してもらうことが可能になるという意 味で地域的貢献ができるものであると考えている。 また、日本語と英語でデータベース化を行い、イ ンターネット上で検索できるシステムの開発を行 うことは、地域のみならず国内外の貝類研究者 や環境問題研究者に学術的寄与を行えることに なり、科学教育や貝類に感心のある人々に、貝 類についての正確な情報を提供することができ ることにもなる。完成すれば、世界でも類をみな い貝類データベースとなりうることからも、この研 究は大きな意義があるものである。

5. 研究の特色:

当研究の特色は、マルチメディアを取り入れたデータベースの構築を目指している点である。つまり、貝類の情報検索を行った場合に、その貝の学術的詳細データと共に、複数の方向から撮影された標本写真や立体回転映像などを同時に見ることができるなど最新技術を取り入れたデータベースを作成する点である。また、日本で採取された貝類を中心にデータベース化を行うため、同じ貝であっても採取された地域による違いを見ることができ、その採取された地域の環境問題などにも参考となりうるデータを提供することができるため、地域社会問題にも貢献できると思われる。国内の貝類をデータベース化した後は、国外で採取された貝類のデータベース化へと拡張する予定である。

6. 研究方法:

以下に実際の研究作業の流れを詳しく説明 する(表1参照)。

まず初めに、基本項目(通し番号、和名、学名、 綱、目、科、産地、模式産地、採集者、採集年月 日、分布、生息場所、標本番号、備考およびそ の英文内容)が記入できる A4 サイズの作業シー トを用意し、種類ごとに分けられた標本からそれ ぞれの貝の学術的な情報を貝類辞典などから 調べて各シートに記入する。次に各シートに記 入された標本番号の貝を標本箱より取り出し、撮 影台に固定して向きを変えながら3方向(種類に よっては2方向)より撮影する。撮影は、ある程度 以上の大きさの標本はデジタルカメラを用い、微 小貝の撮影には顕微鏡セットと 35 ミリフィルムカ メラを用いて行う。撮影用の照明は、デジタルカ メラの場合は自然光と建物用蛍光灯を、顕微鏡 撮影の場合は照明装置を用いる。また撮影ミス などを考慮し、各標本の方向ごとに複数枚ずつ 撮影を行う。撮影後の作業としては、デジタルカ メラの場合はコンパクトフラッシュメモリからパソコ ンヘデータ転送を行い、35 ミリフィルムカメラの 場合は現像後のネガをフィルムスキャナによりデ

ジタル化を行う。次にデジタル化した貝の写真 (付録資料図 3 参照)をフォトレタッチソフトの Adobe Photoshop を使用して、RGB (赤・緑・青の 光の要素)のレベル補正、彩度補正、コントラスト 補正をひとつひとつ目で確認しながら行い(付録 資料図5参照)、実物の貝の色になるように調整 する。最後に背景の色を完全な黒色に変更して、 統一された大きさの画像サイズに切り出し JPEG 形式ファイルとして保存する(付録資料図 4 参 照)。次に各シートに記入されたデータを、 Microsoft Access 形式(MDB 形式)のデータベ ースにフォームを利用して入力する。写真画像 データは OLE オブジェクト形式による貼り付けを 試みたが、ファイルサイズが大きくなりすぎるため、 VBA コントロールにより直接 JPEG ファイルを読 み込み、イメージオブジェクトとして表示させる方 法を採用した。

このようにして基本項目のデータの入力をした 後、その他の詳細項目のデータを貝類関連の辞 典などから調べてデータベースに追加入力をす る。

これらの一連の操作を繰り返し行いながらデータを追加するのが研究作業の流れである。

貝の収集:西宮市貝類館、菊池貝類研究館より貝を運搬作業

貝の分類:資料選定、データ記入、編集作業

写真撮影 :カメラによる貝の撮影作業

データ入力:データベースへ貝の情報データ入力、修正、変換作業

画像入力:スキャナによる貝の画像入力、ファイル転送作業

画像補正 :ソフトウエアによるデジタル画像データの補正、加工作業

データベース化:貝の情報データ、画像の統合最適化作業

データ確認 :入力済みデータの確認作業

WWW へ変換 :データベースから WWW ホームページへの変換

システム開発:貝類データベース管理・検索システムの開発

CD-ROM 化 : 貝類データベースシステムの CD-ROM 化

表1 研究作業の流れ

7. 研究の現状:

まず、西宮市貝類館所蔵の黒田徳米博士が 新種として確定された約230点の標本目録のデ ータベース化(付録資料、図1参照)を終え、現 在はそれ以外の標本約 4 万点の中から陸産の 貝のデータベース化(付録資料、図2参照)に取 りかかっている。次に西宮市貝類館研究員の 方々にも協力していただき、検索キーや画面の 構成などについてご意見をいただいた上で、 Visual BASIC を用いて Windows OS 上で利用で きる貝類データベース検索システムの開発を行 った。現在検索システムのデータ確認作業の最 終段階にきており、今年度の貝類学会で CD-ROM 版データベースとして発表する予定で ある。また、我々はこのデータベースを多くの人 に利用してもらうために Linux OS 上の PostgreSQL データベースサーバにこれらのデ ータを登録し、PHP(正式名称は、"PHP: Hypertext Preprocessor")言語を用いて Web 対 応のデータベース検索システムの研究も行って いる。PostgreSQL は、普通の SQL テーブルで は保存できないデータをデータベースに保存す るためのインターフェースを持っているため、ラ ージオブジェクトにアクセスできるデータベース エンジンである。ラージオブジェクトはテーブルと インデックスの組合せとして保存され、テーブル からは OID(Object ID) 値で参照されるため、 これを利用するとサイズの大きな画像データを、 データベースのテーブルへ容易に保存すること ができる。この PostgreSQL + PHP による検索シ ステムの研究は現在も続行中であるが、情報提 供サービスをできるだけ早く実現させるため、デ ータの入力が完了したものについて順次インタ ーネット上で公開する予定である。現在、微小貝 のデータを公開している URL は、 [http://shell.kwansei.ac.jp/~shell/]である。

8. 作業上の問題点

まず、人的労力の必要性がある。データベー ス化の対象となる貝類標本の総数が多いため、 総データ量が多く、資料選定、データ作成、入 力、確認などの作業に多くの時間を費やしてい る。特に多くの作業量が必要と思われるものとし て、学術的データの資料検索とデータ整理があ げられる。つまり、それぞれの貝の学術的な情報 を貝類辞典などから収集し、1つの貝につき詳 細な 50 以上の項目データを日本語と英語で項 目別に整理し、各データをパソコンで入力すると いう作業である。次に作業効率の問題がある。も うひとつ多くの作業量が必要と思われるものとし て貝の写真撮影とその写真のデジタル化および 画像編集作業があげられる。つまり、一つの貝を 3 方向から撮影し、写真をデジタル化し、デジタ ル化された画像をフォトレタッチソフトにより色調、 コントラスト、画像の大きさなどを調整し、背景色 を統一した後データとして登録する作業である。 撮影対象の貝は小さいものが多く、ピンセットを 用いながら貝を固定したり方向を変えたりする必 要があるため、これらの作業は当初の予想よりも はるかに時間がかかり、大変な作業であることが 判明した。貝の撮影とデジタル化した画像の色 調、コントラスト、明るさ、大きさなどを調整し、より 美しい画像を作ることに多くの時間が必要であっ た。1999 年度は貝を、一般のフィルム式カメラで 接写撮影し、現像後フィルムスキャナを用いてデ ジタル画像を作成し、画像調整をするという作業 を繰り返し行っていたため、作業効率が悪かっ たのではないかと思い、2000 年度は主にデジタ ルカメラを使用して撮影を行った。これによりフィ ルムからデジタル画像化をする作業時間が短縮 されたが、デジタルカメラで撮影した画像はホワ イトバランスが不安定なため、画像調整に多くの 時間が必要になり、作業効率を上げることはほと んどできなかった。

9. 今後の研究計画

データベース作成は時間のかかる作業であり、 特に本研究で扱っているデータは莫大な量である。完成まで先が長いが、継続して少しずつでも データを蓄積することが重要である。今後もこれ ら一連のデータ入力作業を根気よく続け、最終 的には他に類を見ない高度で有用なデータベースになることを目指して、取り組んでいきたいと 考えている。またデータ表示方法のひとつとして、 地図に貝の採集地を表示し、広域地図→貝の 採集地周辺詳細地図→貝の採集地周辺写真→ 貝の採集地写真→貝の標本写真のようにズームイン的に表示できるシステムの開発にも取り組ん でいきたいと考えている。

撮影データ

カメラ: NIKON CoolPix 950 (1/2 inch 211 万画素 CCD)、ズームニッコールレンズ、f=7~21mm (35mm 判換算 38~115mm 相当) F2.6~4 付

撮影時使用画素数:1024×768 (XGA)

画質モード: Fine(約 1/4 圧縮 JPEG)

ホワイトバランス: 蛍光灯用 露出制御: プログラムオート

フォーカスモード:マクロオートフォーカス

ストロボ:発行禁止

参考文献

- 1. 原色日本陸産貝類図鑑,東 正雄著,保育 社
- 2. 原色日本貝類図鑑, 吉良哲明著, 保育社
- 3. 続原色日本貝類図鑑,波部忠重著,保育社
- 4. 学研生物図鑑『貝I』,波部忠重·奥谷喬司 監修,学研
- 5. 学研生物図鑑『貝Ⅱ』,波部忠重·奥谷喬司 監修,学研

付録資料:

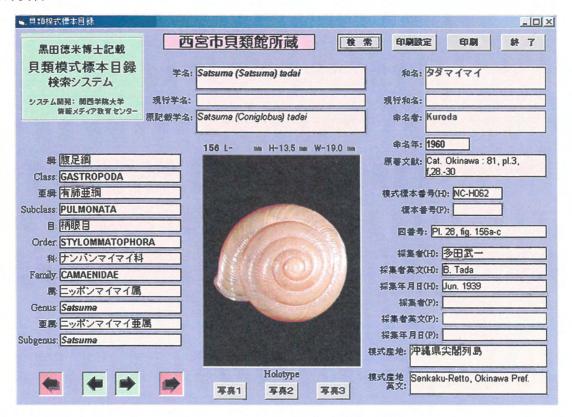


図1 黒田徳米博士記載・貝類模式標本検索システムによる検索結果画面



図2 Microsoft Access による検索結果画面



図3 デジタルカメラによる撮影写真(オリジナル)

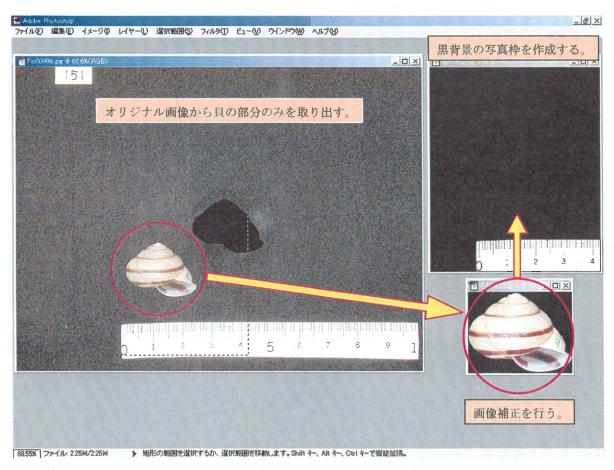


図 5 Adobe Photoshop による画像補正作業画面

■ 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

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 芝勝徳(神戸市外国語大学) (pp.55-56)

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)

- 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.611-70) 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)

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) 主催:人文系データベース協議会

第7回公開シンポジウム実行委員会

議長: 小沢一雅 (大阪電気通信大学)

出田和久 (奈良女子大学)

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)

黒川隆夫 (京都工芸繊維大学)

柴山 守 (大阪市立大学) (大阪樟蔭女子大学)

高橋晴子 都司達夫 (福井大学)

(静岡大学) 中谷広正 中村敏枝 (大阪大学)

八村広三郎 (立命館大学) 濱 裕光 (大阪市立大学)

深海 悟 (大阪工業大学) (福井大学)

宝珍輝尚 正木久仁 (大阪教育大学) 委員長:雄山真弓(関西学院大学)

委員:小沢一雅(大阪電気通信大学)

江澤義典 (関西大学)

中谷広正 (静岡大学) 中村敏枝 (大阪大学)

加藤哲弘 (関西学院大学)

于 康 (関西学院大学)

平林孝裕 (関西学院大学)

人文系データベース協議会 第7回シンポジウム「人文科学とデータベース」

発行日 2001年 11月17日

発行所 第7回公開シンポジウム実行委員会

> 〒662-8501 西宮市上ヶ原1-1-155

関西学院大学情報メディア教育センター ・文学部

雄山真弓

ovama@kwansei.ac.jp

印刷所 有限会社 仁川印刷所

西宮市室川町1番25号

電話 (0798)26-0308