

第19回公開シンポジウム

人文科学とデータベース

活用・継承する「データベース」と
「データベース」による新知見

2013年11月30日（土）

会場：立命館大学アート・リサーチセンター

主催：第19回公開シンポジウム実行委員会

共催：立命館大学アート・リサーチセンター

立命館大学文学研究科文化情報学専修

後援：人文系データベース協議会

目 次

●自由論題報告 I

「表情強度を用いたライフログ映像からの表情表出シーン検索」

森國淳司, 野宮浩揮, 宝珍輝尚 (京都工芸繊維大学 情報工学専攻) … 1

「空間的印象評価法のための分析手法 ―感性ワードに着目して―」

赤井俊介、宝珍輝尚、野宮浩揮 (京都工芸繊維大学) … 9

「The role of metadata in the Balanced Corpus of Contemporary Written Japanese in the analysis of register」

Bor HODOŠČEK (国立国語研究所), 山元啓史 (東京工業大学／カリフォルニア大学サンディエゴ校) …17

「RGB-D カメラを用いたテクスチャの特徴点と3次元点群との対応付けによるリアルタイム3次元形状復元」

鈴木遼平, 徐剛 (立命館大学大学院 情報理工学研究科 情報理工学専攻) …19

「浮世絵を対象とした多言語・異種データベースの横断検索」

Biligsaikhan Batjargal, (立命館大学衣笠総合研究機構補助研究員)

木村文則 (立命館大学衣笠総合研究機構研究員), 前田亮 (立命館大学情報理工学部) …27

【特設セッション】

立命館大学文学研究科・文化情報学専修連携企画

「活用・継承する「データベース」と「データベース」による新知見」

●招待講演

「データベースを発想力に変える方法」

高野明彦 (国立情報学研究所) …… 33

●事例発表

「19 世紀歌舞伎のビッグデータ分析の可能性 ―ヒット現象の数理モデルを用いて」

川畑泰子 (九州大学芸術工学府博士後期課程) ……35

「Web 公開中の伝統芸能興行データベースを使用した演目別上演回数の比較手法について」

坂部裕美子 (立命館大学文学研究科博士後期課程) ……41

「板木データベースの可能性―データベース連携のことなど」

金子貴昭 (立命館大学衣笠総合研究機構研究員) ……45

「画像データベースから生まれる新たな創作活動について」

前崎信也 (立命館大学グローバル・イノベーション研究機構研究員) ……49

●自由論題発表 II

「藤原京遺構データベースの構築」 宮崎良美（奈良女子大学古代学学術研究センター）……55

「奈良盆地前方後円墳データベースの構築と課題 ―WebGIS データベースのひとつの試み―」

出田和久, 石崎研二（奈良女子大学文学部）, 宮崎良美（奈良女子大学古代学学術研究センター）

……63

「広島県立美術館館蔵品データベース構築について」

福田浩子（広島県立美術館 学芸課）……71

「データベース活用におけるバザール型コミュニティ形成について」

藤岡洋（東京大学東洋文化研究所）……77

表情強度を用いたライフログ映像からの表情表出シーン検索

Emotional Scene Retrieval from a Lifelog Video Using Facial Expression Strength

森國 淳司, 野宮 浩揮, 宝珍 輝尚

Atsushi Morikuni, Hiroki Nomiya, Teruhisa Hochin

京都工芸繊維大学 情報工学専攻, 京都市左京区御所海道町

Kyoto Institute of Technology, Gosyokaido-cho, Sakyou-ku, Kyoto

あらまし: ライフログ映像データベースから重要なシーンを検索・提示することを目的として、表情認識に基づくシーン検索法についての検討を行う。従来では人物の表情の種類を識別することで表情表出シーンの検索が可能であったが、表情がどれだけ強く表出されているかは考慮されておらず、シーンの重要度の判断が困難であった。そこで、表情強度という表情表出の程度を表す尺度を導入し、よりの確に重要なシーンを検索できる手法を提案する。

Summary: For the purpose of retrieving the emotional scenes from a lifelog video database, and showing them, we perform the examination about the scene retrieval method based on the recognition of facial expressions. Conventionally, the retrieval of the emotional scenes has been performed by distinguishing the kind of the emotion of the person. However, the judgment of the importance of the scene is difficult without the consideration of the strength of the emotion. Therefore, we propose expression strength and the technique that can search the important scenes more precisely.

キーワード: ライフログ, 映像検索, 表情認識

Keywords: Lifelog, video retrieval, facial expression recognition

1. はじめに

日常生活上の様々なデータを取得・蓄積して、後に有効利用しようというライフログが提案されてきている[1]。ライフログとは文字通りに解釈すると生活の記録である。しかし、ライフログと一口に言っても、記録媒体、情報の種類などは多岐にわたる。例として、ブログ、GPS 記録、個人の健康状態などが挙げられる。新保はライフログの定義を「特定自然人の、特定活動に関する、特定の記録媒体に、自動的に、デジタルデータとして、包括的または連続的に記録(蓄積)し、それによって取得された、特定の個人に関する個人情報(個人識別情報)および個人に関連する個人情報に該当しない情報(非個人識別情報)の総称」としている[2]。

本研究では、ホームビデオのような映像データは手軽に作成することができ、またライフログとして有用な様々な情報を内包することができる点に着目し、ライフログ映像を対象としている。

従来はライフログ映像の取得、保存には高価な記録機器や大容量のストレージが必要であったためライフログの実用化が難しかったが、近年のマルチメディアデータ記録機器の高性能化やストレージの大容量化、価格の低下に伴い、個人でも簡単に大量のマルチメディアデータ(例えば子供の成長記録など)を作成できるようになった。しかしながら、多量の映像が蓄積されるので、映像の量が増える一方で、記録した映像を利用したくても、必要な映像を即座に引き出すことが出来ず、有効活用できていないというのが現状である。このような背景から、ライフログ映像の中から有用な映像を簡単に検索したいという要求が高まってきている。

映像に関する検索システムとして、YouTube[3]等の動画共有サイトが挙げられる。多くの動画共有サイトでは、各動画に対してタグを登録し、検索に活用している。これに対し、ライフログ映像は一般に個人が所有しているため所有者が映像一つ一つにタグを登録しな

ければならず、タグの登録に非常に手間がかかるので、タグを用いた検索は、ライフログ映像検索には適していないと考えられる。また、ホームビデオ等のライフログ映像では、ハンディビデオカメラなどで単純に撮影を行うだけでは、映像にタイトルをつけることや、フォルダごとに整理する事は可能でも、一つの映像の中でより重要なシーンを抽出することは難しい。1日中ずっと撮影を行なって大量のデータを得たとしても、実際にもう一度動画として視聴したいと感じるシーンはその中のごく一部である[4]。このことから長時間の映像の中からユーザが視聴したいと感じるシーンの検索方法が必要である。

そこで本研究では、ライフログ映像から効率的に重要なシーンを検索、提示することを目的とする。

本研究では、視聴者が視聴したいと感じるシーンの検索方法として、映像中の人物の表情を利用する。ユーザが検索したくなるようなシーンでは基本的に何らかの事象が発生し、シーン中の人物はその事象に対して何らかの感情を抱き表情を表出している可能性が高い。このことから、映像中の人物の表情に基づいてシーン検索を行う。

著者らはこれまでに、映像中の人物の表情を用いた、ライフログ映像からの印象的なシーンの検索に関する研究を行ってきた[5]。この先行研究の中では、動画内の人物の表情の検索はできるのだが、表情の程度を推定することができない。例えば、「笑顔」という表情は検索可能であるが、「笑顔」の中の「微笑み」から「高笑い」までというような感情の程度が異なるものが考慮されていない。また、実際に動画のシーン検索を行うにあたって表情の程度を考慮することで、より感情が大きく表出しているような事象を優先的に検索することができると考えられる。

そこで本論文では、このような表情の程度を数値化し、検索に活用することを目的として、顔の特徴点を用いて表情強度という値を導入して、被験者の表情の映像を取得し、表情強度の測定を行い、表情強度を用いることで表情の程度を推定する。

本論文の構成は次の通りである。まず2章で関連研究について述べ、3章では本研究で扱う顔の特徴点について説明する。その後4章で表情強度について述べ、5章で表情強度を用いたライフログ映像検索実験の内容と結果を示し、6章でその考察を行う。最後に7章でまとめを行う。

2. 関連研究

人間の感情を表すいくつかの表情について、その認識、分類のみならず、無表情から最大まで表情の程

度をおおまかに推定することを目的として、木村ら[6]は、顔の部位それぞれの特徴点を抽出し特徴量を作り出すのではなく、顔全体を一つのパターンとして表情認識を行う手法を提案している。処理に用いる顔画像を切り出しサイズの正規化を行った後に、ポテンシャルネット[7]と呼ばれる処理モデルを用いて解析を行っている。しかし、この手法では顔の画像を1枚ずつ人の手で適切に切り抜く必要があり、動画の解析で各フレームに同様の処理を行うのは非現実的である。また、俳優に作為的に演技してもらった表情と比べて、素人の曖昧な表情は認識結果にノイズとしてより大きい影響を与えることを言及している。そのため本研究で扱うようなライフログ映像で撮影が想定される曖昧な表情については不適であると考えられる。

3. 顔特徴点

本研究では、顔特徴点を用いて表情強度を推定する。顔特徴点は、顔特徴点抽出ソフトウェア Luxand FaceSDK 4.0[8]を用いて抽出しており、以下の顔の特徴点 59 点を用いている。図1にこれらの顔特徴点の抽出例を示す。

- ・両目付近 22 点
- ・鼻 9 点
- ・両眉 10 点
- ・口 14 点
- ・ほうれい線 4 点



図1 顔特徴点抽出例（四角い点が顔特徴点を表している）

4. 表情強度

表情の程度を数値化することを目的として本研究では表情強度を定める。本節ではその概要を示す。

4.1 表情強度の推定に用いる特徴量

前述の顔特徴点を組み合わせて、以下に示す10種類の特徴量を定め表情強度の推定に用いる。

- ・左眉と右眉の傾き f_1

左眉(右眉)の傾きは、左眉(右眉)上の5つの顔特徴点から最小2乗法(線形)により算出し、左右の眉がほ

ば対象であると考え、 $f_1 = (a_l - a_r)/2$ とする。ただし a_l は左眉、 a_r は右眉の傾きである。

・眉と目の間の距離 f_2

両眉の 10 点と上瞼の 10 点を両方共左から順に 1 対 1 対応で結んだ 10 本の線分の長さの平均値である。ただし、左目と右目の距離で割って正規化を行う。

・眉間の面積 f_3

両眉と両目からそれぞれ最も顔の中央に近い点を結んでできる四角形の面積である。ただし、左目と右目の距離の 2 乗で割って正規化を行う。

・目の面積 f_4

左目(右目)の 8 つの特徴点をつないでできる八角形の面積 b_l, b_r の平均値 $(b_l + b_r)/2$ である。ただし左目と右目の距離の 2 乗で割って正規化を行う。 b_l は左目、 b_r は右目の特徴量である。

・目の縦横比 f_5

左目(右目)の一番上にある点と一番下にある点を結んだ線分を目の縦の長さ l_v 、最も左と右にある 2 点を結んだ線分を目の横の長さ l_h とし、 $\tan^{-1}(l_v/l_h)$ を左右の目で平均値をとって縦横比とする。

・口の内側の面積 f_6

口の内側の 8 つの特徴点を結んでできる八角形の面積である。ただし左目と右目の距離の 2 乗で割って正規化を行う。

・口の外側の面積 f_7

口の外側の 8 つの特徴点を結んでできる八角形の面積である。ただし左目と右目の距離の 2 乗で割って正規化を行う。

・口の内側の縦横比 f_8

口の内側の点の中で一番上にある点と一番下にある点を結んだ線分を口の縦の長さ m_v 、最も左と右にある 2 点を結んだ線分を口の横の長さ m_h とし、 $\tan^{-1}(m_v/m_h)$ を縦横比とする。

・口の外側の縦横比 f_9

口の外側の点の中で一番上にある点と一番下にある点を結んだ線分を口の縦の長さ m_v 、最も左と右にある 2 点を結んだ線分を口の横の長さ m_h とし、 $\tan^{-1}(m_v/m_h)$ を縦横比とする。

・口角の上がつている度合 f_{10}

口の両端の点 x_1, x_2 と口の最も上の点と下の点 x_3, x_4 から次式によって求められる。

$$\frac{(y(x_1) + y(x_2)) - (y(x_3) + y(x_4))}{|y(x_3) + y(x_4)|}$$

ここで、 $y(x)$ は特徴点 x の y 座標を表す。 x_1 と x_2 の y 座標の平均値が x_3 と x_4 の y 座標の平均値より大きければ正の値をとるため、この値が大きいほど口角が上がっていると言える。分母は正規化のための項である。

以上、10 個の特徴量をライフログ映像中の各フレーム画像(10 フレームおきに抽出したもの)に対して求め、フレームごとに 10 次元特徴ベクトル $(f_{i1}, \dots, f_{i10})$ を構成している。ここで、 i はフレーム番号を表す。

4.2 表情強度の算出

4.1 節では人間が表情を表出する際に大きく変化すると考えられる 10 個の顔の特徴量を示した。しかし、どの特徴量を重視すればよいかは、表情の種類や個人差などに依存すると考えられる。そこで、各特徴量に対して主成分分析を行い、その固有値と主成分得点にもとづいて表情強度を定める。 i 番目のフレームの表情強度 I_i は、4.1 節で示した 10 個の特徴量 $(f_{i1}, \dots, f_{i10})$ に対して主成分分析を行い、得られた第 1～第 10 主成分得点 $(p_{i1}, \dots, p_{i10})$ を用いて次の式(1)のように定める。なお、主成分得点は平均 0、分散 1 に正規化する。

$$I_i = \sum_{j=1}^{10} \lambda_j \text{sgn}(s_j) p_{ij} \quad (1)$$

ここで、 λ_j は第 j 主成分の固有値である。また、

$$s_i = \sum_{j=1}^{[N/2]} \sum_{k=1}^{10} f_{r_{ij}k} - \sum_{j=[N/2]+1}^N \sum_{k=1}^{10} f_{r_{ij}k} \quad (2)$$

である。 N は映像に含まれるフレームの総数、 r_{ij} は $p_{i1} \sim p_{Ni}$ を降順に並べた時の p_{ij} の順位を示す。例えば、 p_{ij} が $p_{1j} \sim p_{Nj}$ の中で 30 番目に大きい値を持つ場合、 $r_{ij} = 30$ となる。また、

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{otherwise} \end{cases} \quad (3)$$

である。

一般に、顔面上の変化が大きい場合は、それぞれの特徴量 $(f_{i1}, \dots, f_{i10})$ の値は大きくなる傾向にある。したがって、式(2)の第 1 項 (p_{ij}) の値が大きいものから順に $[N/2]$ フレームの特徴量の値の総和が第 2 項(残りの $N-[N/2]$ フレームの特徴量の値の総和)よりも大きい場合は、 p_{ij} の値が大きいと表情強度が強いと考えられる。したがって、そのような場合($s_j > 0$)には、 p_{ij} に正の係数($\text{sgn}(s_j) = 1$)を乗じる。逆に、式(2)の第 1 項が第 2 項よりも小さい場合($s_j < 0$)には、 p_{ij} に負の係数($\text{sgn}(s_j) = -1$)を乗じる。また、式(2)の第 1 項と第 2 項が等しい場合($s_j = 0$)には、 p_{ij} から表情強度を求めることが難しいと考えられるため、 p_{ij} に($\text{sgn}(s_j) = 0$)を乗じて、表情強度が p_{ij} の値に影響されないようにしている。

5. 実験

本節では 4 節で示した表情強度を用いて、表情の程度を考慮した表情表出シーンの検出実験を行った結果を示す。

5.1 実験方法

被験者は著者の一人を含む男子学生 8 名 (21 才～24 才) である。8 名 (以降、被験者 A, B, C, D, E, F, G, H と表記する。) を 4 名×2 グループに分け、すべての被験者の前に顔が撮影できるように Web カメラを設置した。一回の撮影につき 10 分を目安にゲームの区切りの良いところまでトランプのハーツ、大富豪をそれぞれ 2 度ずつ行なってもらい、計 4 度の撮影を行い 8×4=32 個の表情映像データを得た。その後、別途、そのトランプを行なっている様子をホームビデオ風に撮影していた映像を被験者に視聴してもらい、検索したいシーンをその映像中からアンケートで回答を行ってもらった。撮影機材は以下のとおりである。

・表情撮影カメラ ELECOM UCAM-DLU130HWH
このカメラの主な仕様を以下に示す。

- サイズ: 640×480
- フレームレート: 30fps
- ファイル形式: WMV

なお用いた録画ソフトは Windows Movie Maker である。

・全体像撮影カメラ Sony HDR-CX560V

このカメラの主な仕様を以下に示す。

- サイズ: 1440×1080
- フレームレート: 30fps
- ファイル形式: MPEG4 AVC/H.264

次に、得られた表情映像に対して、著者の一人が目視による表情の種類と表情の程度の判定を行った。目視による表情判断基準を表 1 に示す。表 1 中では、「表情一覧」には表情の種類、「分類される表情」にはその表情に分類される代表的な表情、「特徴」にはその表情の判断基準を記述した。その後、4 節で示した表情強度の推定を行い、目視による判定と強度の照らし合わせを行った。

5.2 実験の結果

表 1 目視時の表情判断基準表

表情一覧	分類される表情	特徴
笑顔 1	微笑み、苦笑	口はあまり開けず唇の両端が上がる
笑顔 2	笑う	目尻が下がり、口を少し開けて笑い声をあげる。
無表情		基本 6 感情(喜び、驚き、悲しみ、怒り、恐れ、嫌悪)以外の表情

表 2 表情強度平均例 (ハーツ 3E)

表出表情	表出シーン	表情強度平均
		0.643
笑顔 1	0:54 ~ 1:04	-0.481
		-0.209
笑顔 2	2:00 ~ 2:07	0.410
笑顔 1	2:08 ~ 2:15	1.780
		-0.283
笑顔 1	2:30 ~ 2:37	0.062
		-0.027
驚き 2	2:52 ~ 2:55	0.823
		-0.115
笑顔 1	3:06 ~ 3:08	0.012
		-1.330
笑顔 1	3:44 ~ 3:46	-1.381
		-0.435
笑顔 1	4:06 ~ 4:12	-0.333
		0.210
笑顔 2	4:26 ~ 4:29	2.060
		-0.570
笑顔 1	4:38 ~ 4:45	0.795
		1.316
笑顔 1	4:50 ~ 4:52	-1.723
笑顔 2	4:58 ~ 5:06	1.786
		-0.070
笑顔 2	5:27 ~ 5:30	0.206
		0.037
笑顔 1	8:48 ~ 8:49	-0.170
		0.198
笑顔 1	9:30 ~ 9:35	0.479
		0.156
笑顔 1	10:01 ~ 10:07	0.323
		0.555
笑顔 1	10:10 ~ 10:13	0.096
		0.117
笑顔 1	10:40 ~ 10:42	0.391
		-0.001
笑顔 2	10:52 ~ 11:02	-0.033
		-0.173
笑顔 2	11:52 ~ 11:58	-0.207
		-0.414
笑顔 1	12:12 ~ 12:20	-0.228
笑顔 2	12:20 ~ 12:25	-0.033
		-0.172
笑顔 1	12:42 ~ 12:45	0.094
笑顔 2	12:52 ~ 12:58	-0.355

表 2 に実験で得られた結果の一部を示す。

表 2 は3回目にハーツを行った回における被験者 E の表出表情と表情強度平均を示したものである。ただし、表出表情は筆者の一人が表 1 に示す基準を用いて、目視で判断したものである。また、「表出シーン」は時間(分:秒)を示し、表情強度平均はそのシーン内の表情強度の平均を示す。

今回の実験ではトランプを行なってもらったため、表出していた表情はほとんどが喜び(笑顔)の表情であった。そこで、目視で表情を確認した際の判断基準として表 1 に記述する表情を用いた。

表 3 に 8 人の被験者の映像の中で強度の傾向が変わった 2 フレーム、計 16 フレームを示した。32 個の映像中から被験者 8 人それぞれから特に顕著な傾向が見られた動画を 1 つずつ選択している。

左端の列「ハ 1A」はハーツ 1 回目被験者 A のフレームを指す。「大 4E」は大富豪 4 回目の被験者 E のフレームを指す。表中の薄く色を塗っている f6,f7,f9 は、主成分分析時に第一主成分の主成分負荷量の絶対値が高かった上位 3 つの特徴量であり、濃い色を塗っている f1(被験者 G のみ f8),f2,f3 は第 2 主成分の主成分負荷量の絶対値が高かった上位 3 つの特徴量である。

表 3 各動画において特徴量の傾向が大きく変化した 2 フレームの各特徴量

	F 番号	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10
ハ 1A	6060	-0.090	0.345	0.196	0.061	0.310	0.184	0.021	0.336	0.031	0.045
	6070	-0.031	0.378	0.190	0.047	0.319	0.192	0.034	0.346	0.074	-0.041
ハ 2B	5470	-0.174	0.328	0.185	0.049	0.303	0.151	0.020	0.370	0.081	0.105
	5480	-0.270	0.304	0.197	0.045	0.284	0.191	0.045	0.508	0.170	0.000
ハ 1C	4610	-0.106	0.327	0.188	0.048	0.336	0.245	0.104	0.533	0.308	0.214
	4620	-0.221	0.346	0.225	0.044	0.306	0.171	0.045	0.389	0.090	-0.056
ハ 1D	14570	-0.177	0.317	0.196	0.038	0.299	0.203	0.087	0.491	0.269	0.303
	14580	-0.055	0.320	0.180	0.038	0.292	0.250	0.110	0.472	0.219	0.313
大 4E	11570	-0.273	0.306	0.188	0.030	0.204	0.203	0.078	0.326	0.135	0.450
	11580	-0.221	0.307	0.164	0.026	0.173	0.254	0.093	0.384	0.138	0.615
ハ 4F	10470	-0.134	0.325	0.203	0.036	0.296	0.182	0.028	0.361	0.078	-0.200
	10480	-0.081	0.328	0.198	0.035	0.269	0.166	0.016	0.377	0.020	-0.150
ハ 3G	14280	-0.140	0.290	0.183	0.044	0.312	0.152	0.019	0.433	0.077	0.211
	14290	-0.136	0.310	0.182	0.039	0.273	0.156	0.012	0.432	0.026	0.111
ハ 3H	2360	-0.122	0.352	0.208	0.050	0.303	0.254	0.064	0.464	0.159	0.040
	2370	-0.143	0.355	0.212	0.047	0.333	0.262	0.094	0.576	0.254	0.133

表 4 第 1 主成分 主成分負荷量

第1主成分	ハ 1A	ハ 2B	ハ 1C	ハ 1D	大 4E	ハ 4F	ハ 3G	ハ 3H
f1	0.080	-0.008	-0.355	0.047	0.051	0.357	0.036	-0.136
f2	-0.012	-0.322	0.241	0.015	0.433	-0.395	0.115	-0.053
f3	0.051	-0.293	0.439	-0.041	0.247	-0.368	0.015	-0.080
f4	-0.231	-0.425	0.325	-0.166	-0.324	0.340	-0.377	0.447
f5	0.103	-0.613	0.191	-0.342	-0.444	0.380	-0.488	0.006
f6	-0.914	0.893	0.895	0.923	0.910	-0.819	0.867	0.951
f7	-0.945	0.936	0.929	0.960	0.943	-0.828	0.934	0.969
f8	-0.556	0.187	0.375	0.859	0.516	-0.295	0.124	0.509
f9	-0.915	0.849	0.808	0.942	0.920	-0.757	0.874	0.927
f10	-0.518	0.651	0.504	-0.032	0.586	-0.049	0.412	0.664
固有値	3.217	3.598	3.229	3.548	3.727	2.697	2.970	3.632
寄与率	32.168	35.976	32.290	35.475	37.271	26.970	29.696	36.317
累積寄与率	32.168	35.976	32.290	35.475	37.271	26.970	29.696	36.317

表 5 第 2 主成分 主成分負荷量

第2主成分	ハ 1A	ハ 2B	ハ 1C	ハ 1D	大 4E	ハ 4F	ハ 3G	ハ 3H
f1	0.340	0.496	-0.099	0.585	0.585	0.567	0.672	0.155
f2	-0.841	-0.763	-0.884	-0.813	-0.690	-0.709	-0.680	-0.832
f3	-0.880	-0.891	-0.759	-0.929	-0.896	-0.762	-0.928	-0.872
f4	-0.603	0.126	0.027	-0.155	-0.454	0.464	-0.086	-0.598
f5	-0.367	0.144	0.270	-0.212	-0.438	0.209	-0.084	-0.505
f6	-0.118	-0.177	0.052	-0.046	-0.056	0.258	-0.217	-0.020
f7	0.074	-0.143	0.061	-0.015	0.046	0.472	0.008	0.044
f8	0.012	-0.155	0.597	-0.061	-0.030	0.063	-0.076	0.056
f9	0.111	-0.199	0.314	-0.031	0.085	0.468	0.053	0.120
f10	0.045	0.198	-0.258	0.465	0.162	0.513	0.335	0.020
固有値	2.129	1.813	1.969	2.158	2.059	2.438	1.958	2.109
寄与率	21.295	18.130	19.693	21.585	20.587	24.381	19.576	21.094
累積寄与率	53.463	54.106	51.983	57.060	57.858	51.352	49.272	57.411

表 6 第 3 主成分 主成分負荷量

第 3 主成分	ハ 1A	ハ 2B	ハ 1C	ハ 1D	大 4E	ハ 4F	ハ 3G	ハ 3H
f1	0.712	-0.354	-0.030	-0.130	-0.292	0.068	0.190	-0.441
f2	0.033	-0.117	0.046	0.079	0.285	0.302	0.060	0.303
f3	-0.206	-0.068	0.183	0.139	0.286	0.352	-0.032	0.358
f4	0.505	-0.676	0.817	-0.856	-0.634	0.642	0.712	-0.425
f5	0.308	-0.574	0.858	-0.768	-0.648	0.758	0.599	-0.747
f6	-0.060	-0.216	-0.139	-0.056	-0.278	0.153	0.204	0.061
f7	0.002	-0.179	-0.159	-0.078	-0.152	0.002	0.061	0.048
f8	-0.416	-0.513	-0.100	-0.168	-0.546	0.497	0.619	-0.110
f9	-0.048	-0.273	-0.157	-0.134	-0.154	0.023	0.163	0.019
f10	0.558	-0.019	-0.187	-0.095	0.252	-0.186	0.229	0.171
固有値	1.390	1.345	1.554	1.429	1.555	1.511	1.414	1.201
寄与率	13.904	13.453	15.540	14.291	15.550	15.108	14.136	12.008
累積寄与率	67.367	67.559	67.523	71.351	73.408	66.459	63.407	69.419

また、主成分分析を行い得られた第 1～第 3 主成分の主成分負荷量を表 4～表 6 に記述する。一番上の行におけるハ 1A は、1 回目のハーツにおける被験者 A の表情映像を、大 4E は 4 回目の大富豪における被験者 E の表情映像を意味する。

6. 考察

(a)表 1 の全シーンについて

笑顔 2, 笑顔 1, 無表情のそれぞれをすべてのシーンで表情強度の平均を取った結果を、表 7 に示す。

表 7 被験者 E～H の各表情の全シーンにおける強度平均

映像	笑顔 2	笑顔 1	無表情
ハーツ 3E	0.583	0.045	-0.040
ハーツ 3F	0.008	-0.104	-0.021
ハーツ 3G	1.230	0.317	0.544
ハーツ 3H	0.944	-0.298	0.227

平均値を見ると確かに、大きく笑っている笑顔 2 は笑顔 1 や無表情に比べて高く、強く表情が表出しているシーンに関しては表情強度が高い値をとることが分かる。しかし、笑顔 1 と無表情に関しては、無表情のほうが笑顔 1 より高いことも多く、表情表出の度合いが小さい時の精度は改善の余地がある。

(b)表 4～6 について

第一主成分に関しては、どの被験者に関しても、f6, f7, f9 の主成分負荷量が高く、全被験者について第一主成分は口の概形の成分といえる。

第二主成分に関しては、被験者 C 以外は f1, f2, f3 が強く、C に関しても f2, f3 の負荷量が高いため、全被験者について眉間の概形の成分と考えられる。

第三主成分に関しては、f4, f5 が全被験者について強く、目の概形の成分であると考えられる。

以上から、どの被験者についても主成分分析を行った後の傾向が似通っていると言え、表情の変化には口、目、眉間の概形が大きな影響を及ぼしている。直感的にも、人間が表情の変化を視覚的に感知する時は口や目を特に参考にすると考えられるので、概ね妥当な結果と考えられる。

(c)被験者がアンケートで実際に検索したいと回答したシーン

表 8 に、大富豪 4 回目において、被験者が実験後にアンケート回答した「検索したいシーン」の表情強度の平均を示す。その他は被験者が選択しなかったシーンすべての強度平均を意味する。

表 8 被験者がアンケートで検索したいと回答したシーン

(a)大富豪 4E

欲しいシーン	強度平均
1:13-1:22	4.905
2:02-2:07	5.302
3:08-3:12	0.185
6:22-6:27	0.095
7:02-7:12	-5.116
8:46-8:56	0.471
9:22-9:30	-2.330
その他	-0.038

(b)大富豪 4F

欲しいシーン	強度平均
0:45-0:55	3.546
1:25-2:10	5.269
6:08-6:15	-3.951
7:50-8:00	-4.079
8:10-8:15	-1.073
その他	-0.363

(c)大富豪 4G

欲しいシーン	強度平均
0:13-0:30	1.412
0:58-1:10	-0.992
1:10-1:30	-1.067
2:01-2:13	1.481
3:00-3:17	-1.839
4:36-4:46	-1.361
6:49-6:57	0.861
その他	-0.039

(d)大富豪 4H

欲しいシーン	強度平均
2:55-3:10	1.530
9:08-9:28	0.405
9:30-9:35	0.237
その他	-0.062

この結果から、被験者が検索したいと回答した全 22 シーン 中 13 シーンはその他のシーンよりも平均が高かった。このことから検索したいシーンの 6 割程度は、強

度平均によって検索できる可能性があると考えられる。また検索できなかったシーンでは、考察(b)に示したように表情表出度合いが小さいシーンが多く、うまく強度では判断ができなかった。この点については今後改善していく予定である。

7. まとめ・今後の課題

本論文では、ライフログ映像から効率的に重要なシーンを検索、提示することを目的に視聴者が視聴したいと感じるシーンの検索に用いることが出来る表情強度の提案を行った。その結果、表情表出シーンならびに被験者が検索したいと感じるシーンのある程度検出することが出来た。

今後の課題としては、表情強度に改善の余地が見られたので、表情強度の推定に使用する特徴量の改善を考えている。また、今回はあくまで 1 人のユーザが自分一人の表情を使うことしか想定していないが、ホームビデオのような映像では他の人物が映っている場合も十分考えられるため、そういった複数人の表情に対して表情強度を算出し、それら全員の強度を考慮した上で検索精度を向上させたいと考えている。

8. 参考文献

- [1] 相澤清晴, "体験の情報処理 - ライフログの取得と処理 -", 電子情報通信学会技術研究報告, PRMU, パターン認識・メディア理解 103(738), pp.284-290, 2004.
- [2] 新保史生, "ライフログの定義と法的責任 個人の行動履歴を営利目的で利用することの妥当性", 情報管理, vol.53, No.6, pp.295-310, 2010.
- [3] YouTube, <http://www.youtube.com/>.
- [4] Shogo Shimura, Yasushi Hirano, Shoji Kajita and Kenji Mase, "Experiment of Recalling Emotions in Wearable Experience Recordings", Advances in Pervasive Computing: Adjunct Proceedings of the Third International Conference on Pervasive Computing, pp.19-22, 2005
- [5] 野宮浩揮, 森國淳司, 宝珍輝尚, "ライフログ映像検索に向けた顔特徴点の位置関係に基づく効率的な表情表出シーン検出", 電子情報通信学会論文誌, Vol.J96-D, No.5, pp.1279-1289, 2013.
- [6] 木村聡, "表情認識とその程度推定", 情報処理学会研究報告. CVIM, [コンピュータビジョンとイメージメディア] 97(10), pp.99-106, 1997.
- [7] 松野 勝弘, 李 七雨, 辻 三郎, "ポテンシャルネットと KL 展開を用いた顔表情の認識", 電気情報通信学会論文誌 D II, No.8, pp.1591-1600, 1994
- [8] Luxand, Inc., "Luxand FaceSDK4.0," <http://www.luxan.com/facesdk>.

空間的印象評価法のための分析手法 －感性ワードに着目して－ Analysis Method for the Impression Evaluation Method by Space － Focused on the Impression Words －

赤井 俊介、宝珍 輝尚、野宮 浩揮

Shunsuke Akai, Teruhisa Hochin, Hiroki Nomiya

京都工芸繊維大学, 京都市左京区松ヶ崎御所海道町

Kyoto Institute of Technology, Goshokaido-cho, Matsugasaki, Sakyo-ku, Kyoto

あらまし: 現在、感性の評価手法として SD 法がよく使われている。しかしながら、感性というあいまいな事柄を考えると、SD 法では評価が困難なことがある。そこで、感性のあいまい性を考慮した評価手法として、空間的印象評価法の提案を行ってきた。具体的には、対象の印象を表す言葉である感性ワードが散りばめられた空間を用意し、そこから対象の印象に当てはまる部分を領域として囲んでもらうことで評価する手法である。この評価法を用いると、SD 法では評価できない印象も評価可能であるが、評価結果を単純に比較することができず、分析が困難である。そこで、本論文では、評価に用いられた感性ワードに着目した分析手法を提案し、それを実現するためのシステムを実装する。そして、以前に行なった空間的印象評価法を用いた画像の印象評価実験で得られた評価結果に対して分析を行う。また、その分析結果から、分析手法の有用性を示す。

Summary: The Semantic Differential method is often used as an evaluation method of Kansei. However, it is hard to evaluate because Kansei is vague. Therefore, Impression Evaluation Method by Space (IEMS) is proposed as the impression evaluation method considering the vagueness of Kansei. IEMS uses the space containing impression words. The impression of an object is specified by circling impression words that match the impression. The degree of matching of the impression is expressed by the darkness of the color. IEMS can also evaluate impressions that the SD method cannot. However, analysis of evaluation results is difficult because they cannot simply be compared. In this paper, an analysis method focused on the impression words is proposed. The system for the proposed analysis method is implemented. The proposed method is applied to the analysis of the evaluation results of impression evaluation experiment conducted before to show the effectiveness of the proposed method.

キーワード: 分析手法, 印象評価, 感性, あいまい性

Keywords: analysis method, impression evaluation, Kansei, vagueness

1. はじめに

近年、消費者の要求は機能や利便性といった物理的なものだけでなく、見た目の美しさなど消費者がどのように感じるかという感性による部分が重要になってきている。しかし、感性というものは個人によって異なり、あいまいで、定量化が困難であるという特徴があるため、的確に表現することは難しい。

そのような感性の評価手法として、現在 SD 法[1]がよく使用されている。SD 法は印象を数値化することで統計処理を可能にし、様々な分析を行うことを可能にしている。しかし、統計処理を行うために決められた範囲での評価が強制されており、実験者にとって都合の良い項目しかないといった、恣意的要素が介在してしまう恐れがある。そのため、被験者の感じる印象が表現

できない可能性がある。つまり、感性というあいまいな事柄の評価が困難な場合があると考えられる。従って、評価者にとって、あいまいな評価ができるような、柔軟な印象評価手法が必要である。感性の評価手法に関する研究はいくつか行われているが、確立するには至っていない[2-4]。

そこで、感性のあいまい性を考慮した印象評価手法として、空間的印象評価法 IEMS (Impression Evaluation Method by Space) という手法を提案してきた[5,6]。IEMS は、対象の印象を表す言葉(以下、感性ワードという)が散りばめられた感性空間において、対象の印象として当てはまる部分を領域として囲み、領域内に色を塗ることで、当てはまる度合いを評価する手法である。この評価法を用いることで、SD 法では評価できない印象も評価可能であることを示してきた[5,6]。また、感性ワードが移動も追加もされていない感性空間(ベースライン感性空間)に対しては、評価結果の各地点の色の濃さの平均や変動係数を求め、ベースライン感性空間上で示すことにより、評価対象の平均的な印象や評価のばらつきを視覚的に提示できることを示してきた[7,8]。しかし、感性ワードの移動や追加が可能であるため、評価結果を単純に比較することができず、分析が困難であるという問題がある。

そこで、本論文では、評価に用いられた感性ワードに着目した分析手法を提案し、それを実現するためのシステムを実装する。そして、以前行われた、画像の印象評価実験から得られた評価結果に対して分析を行い、考察する。

以降、2. で空間的印象評価法(IEMS)について説明する。そして、3. で感性ワードに着目した分析手法を提案し、4. で画像の印象評価実験から得られた評価結果に対しての分析について述べる。そして、5. で考察を行い、最後に6. でまとめる。

2. 空間的印象評価法(IEMS)

2.1 感性空間について

IEMS では感性空間という概念を用いる。ここでの感

性空間とは、人が対象の印象を評価する際に思い浮かべる空間のことであるとする。例えば、ある風景写真の印象を問うと、「爽やか」や「美しい」といった言葉で表現できる。また、SD 法のように、提示された感性ワードに合致している程度を示すこともできる。従って、心の中(感性空間)では、印象を表す項目(以下、印象表現項目という)が存在し、対象物がどの印象表現項目にどの程度合致しているかを判断していると考えられる。

印象表現項目として最もよく使用され、最も重要であるのは語であると考えられる。そこで、IEMS では印象を表現する語(感性ワード)を用いる。また、感性空間では、イメージの似ている感性ワードほど近くに配置されているとする。さらに、感性空間は常に変化する可能性があり、個人個人で異なるという性質があると考えられる。

2.2 IEMS の概要

IEMS では、感性ワードが散りばめられた平面を感性空間として用いる。2.1 で述べたように、感性空間内の感性ワードは、似ているものほど近くに配置されているとする。そして、感性空間から、対象の印象として当てはまる部分を領域として囲み、その領域内に色を塗ることで当てはまる度合いを表現する。つまり、濃く塗られている部分ほど、より当てはまると感じることになる。用いる色はグレー(白～黒)とし、濃さの表現には、塗れば塗るほどその部分が濃くなるブラシを用いる。

なお、共通の感性空間としてベースライン感性空間(図1)を最初に提供し、それをもとに必要なに応じて修正を加えていくこととする。ベースライン感性空間は、以下の手順で実験的に作成した[5,6]。

- ① 感性評価によく使用されている感性ワードの選定
- ② 感性ワードの非類似度のアンケート調査
- ③ MDS(多次元尺度構成法)[9]を用いた平面へのプロット

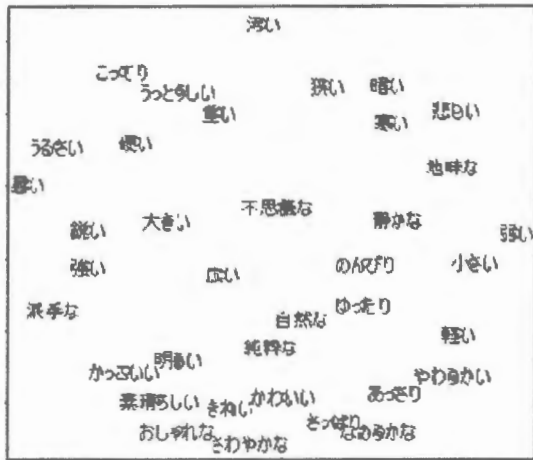


図1 ベースライン感性空間

2.3 IEMS での評価におけるルール

IEMS での評価におけるルールを以下に示す。

- ・ 対象の印象として当てはまる部分を囲み、当てはまる度合いが強いほど色を濃く塗る。
- ・ 当てはまると感じる感性ワードは必ず囲むようにする。
- ・ 1つの領域に任意の数の感性ワードを含めることができる。
- ・ 囲む領域はいくつでも良い。
- ・ 評価する観点が違う場合は、別の領域で囲む。
- ・ 感性空間内に評価したい感性ワードが無い場合は自由に追加することができる。
- ・ 似ているものほど近くに配置されているという条件のもとで、感性ワードの配置を自由に変更できる。
- ・ 同じ感性ワードを違う観点で評価したい場合や、領域が重なってしまう場合は、複数の感性空間で表現できる。

2.4 IEMS での評価例

IEMS での評価例を示す。ここでは簡単のため、図2左に示す単純な感性空間の場合を考える。この感性空間を用いて、ある対象の印象を評価した例を図2右に示す。この評価例の場合、対象の印象として、「広い」、「明るい」、「きれい」を感じており、「広い」と「明る

い」の間付近をより強く感じていることを表している。このように IEMS では、印象の度合いを色の濃さで表現するので、あいまいな評価が可能である。また、評価したい感性ワードが無い場合には自由に追加することができ、言葉として思いつかない場合でも、似ている感性ワードを用いて評価することが可能である。さらに、反対の意味の感性ワードを同時に評価することも可能である。

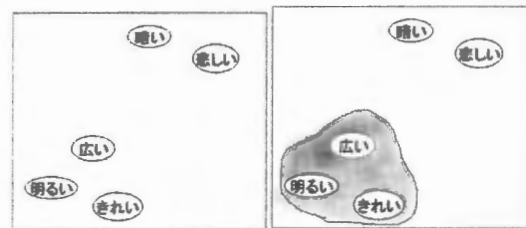


図2 感性空間の例(左)と評価例(右)

2.5 印象評価システム

IEMS での評価を実現するために、印象評価システムを実装した[5,6]。システムの開始画面を図3に示す。画面中央が感性空間を表しており、この感性空間上で評価を行ってもらう。また、このシステムでは、画面右側に画像を表示させることができ、画像の印象を評価することができる。

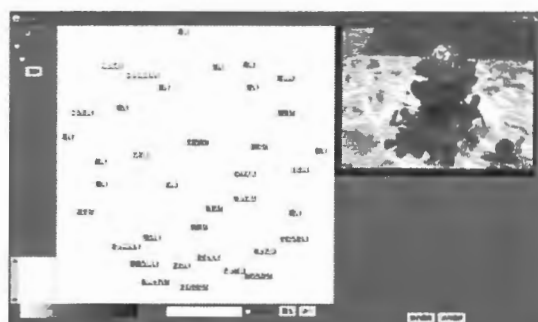


図3 印象評価システム

3. 感性ワードに着目した分析手法

IEMS での評価結果は、単純に比較することができず、分析が困難である。そこで、感性ワードに着目した分析手法を提案する。

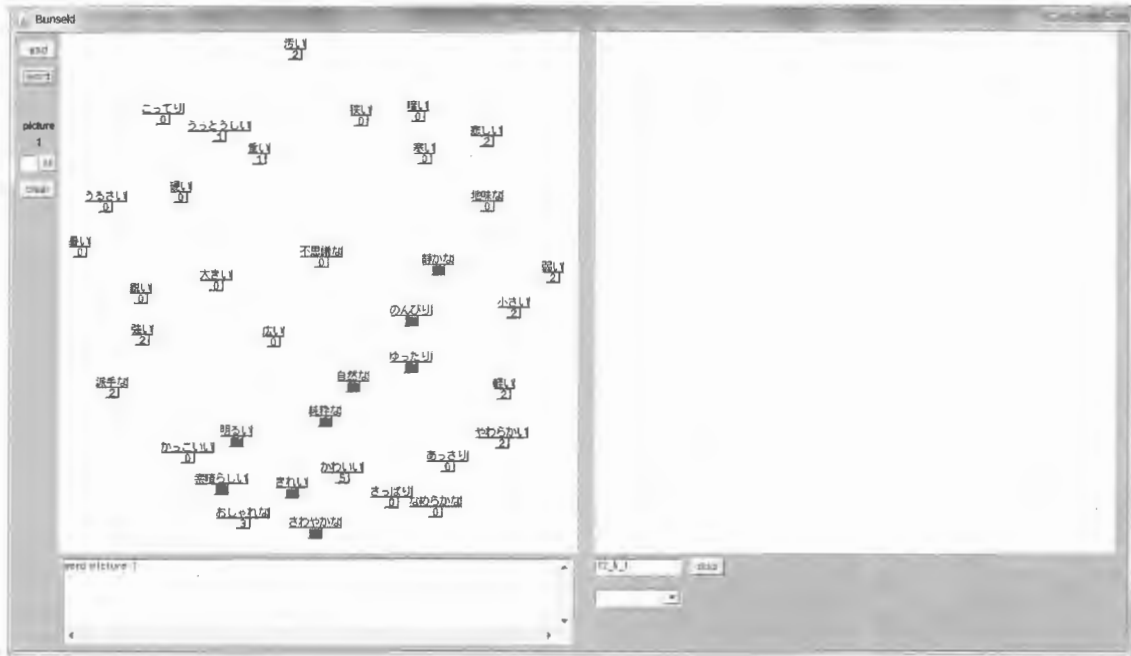


図5 分析システムの実行例(各感性ワードを囲んだ人数の表示)

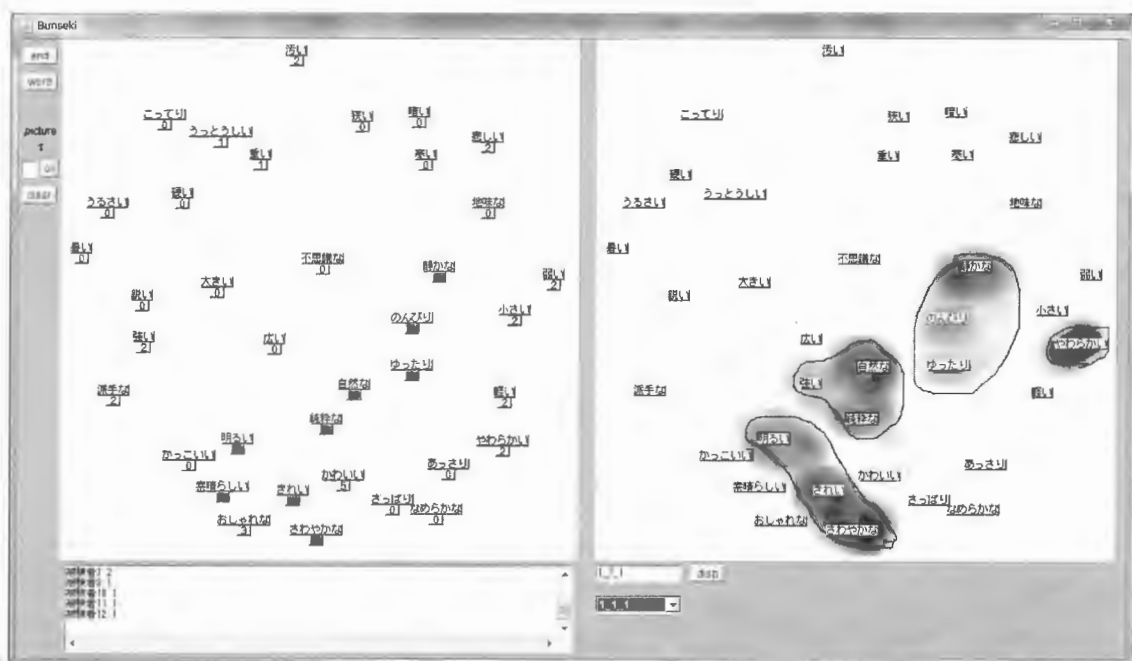


図6 分析システムの実行例(評価結果の表示)

4. 実際の評価結果に対する分析

以前行われた、IEMS を用いた画像の印象評価実験 [5,6] の評価結果に対して、実際に分析システムを用いた分析を行う。まず、画像の印象評価実験について簡単に説明し、次に分析結果を示す。

4.1 IEMS を用いた画像の印象評価実験

IEMS の有効性を示すために、画像に対する印象を評価してもらう実験を行なった[5-6]。実験の概要は以下の通りである。

・被験者: 20 歳～24 歳の工科系の大学生 12 名
(男性 10 名、女性 2 名)

・評価対象: 図7の計 5 枚の画像(画像の提示順序はランダム)

4.2 分析結果

4.1 の実験から得られた、各画像の評価結果に対して、感性ワードごとに囲んだ被験者数をベースライン感性空間上に表示させた。その結果を、図8～図12に示す。

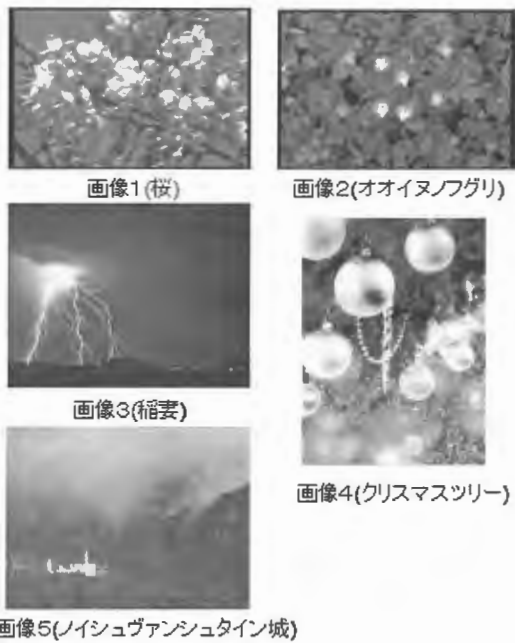


図7 実験に用いた画像[10-13]

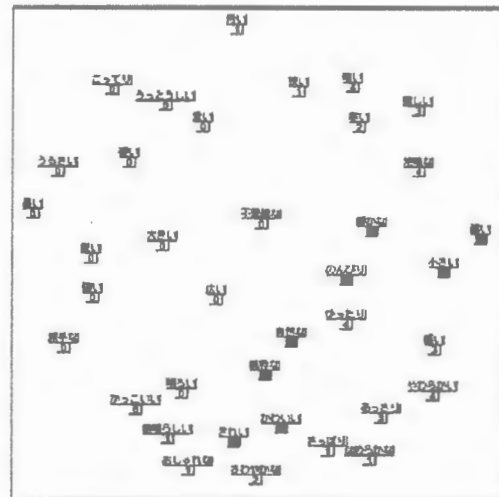


図9 画像2の分析結果

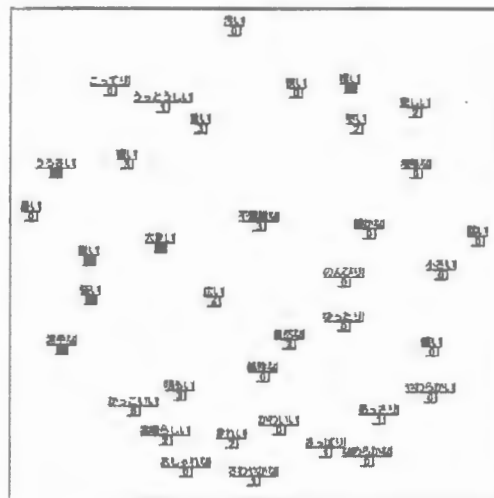


図10 画像3の分析結果

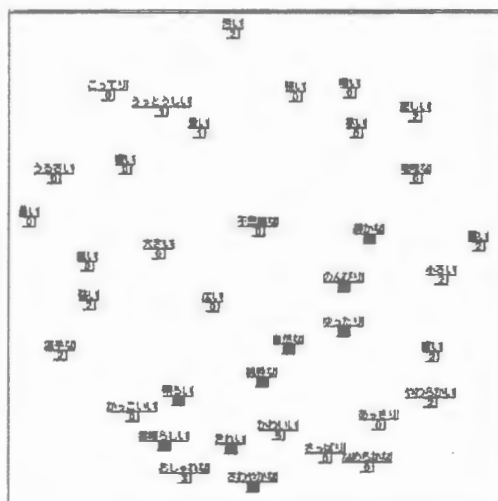


図8 画像1の分析結果

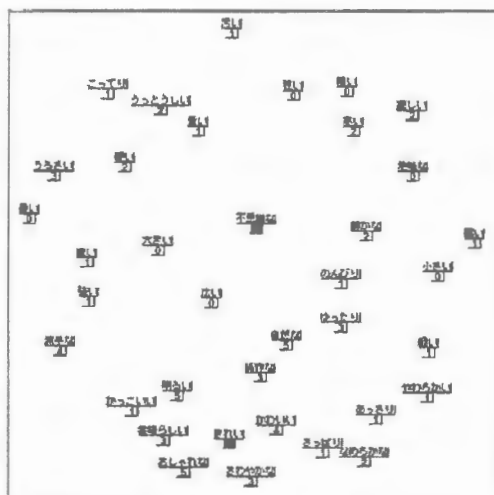


図11 画像4の分析結果

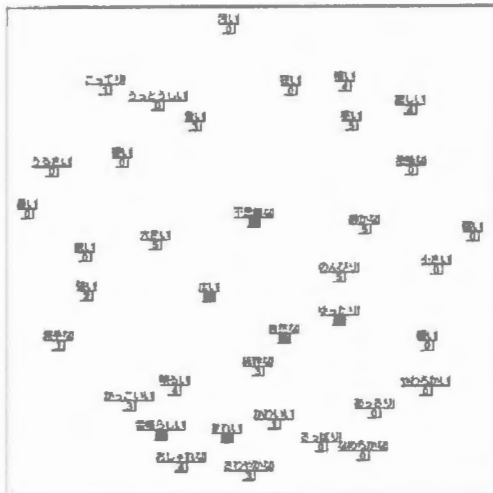


図12 画像5の分析結果

図8を見ると、画像1では、「静かな」から「自然な」、「きれいな」にかけての感性ワードが多く評価されている。特に「自然な」、「きれいな」は、ほとんどの被験者が感じていることが分かる。

図9を見ると、画像2では、画像1と同様に「静かな」から「自然な」、「きれいな」にかけての感性ワードが多く評価されている。しかし、画像1とは違い、「きれいな」よりも「かわいい」を評価している被験者の方が多く、「明るい」は評価されていない。また、ベースライン右上の「暗い」、「悲しい」を評価している人が画像1より多い。さらに、「小さい」、「弱い」などのベースライン右側の感性ワードも画像1より多い。同じ「自然な」が感じられる画像でも、違いがあることが分かる。

図10を見ると、画像3では、「強い」、「大きい」、「鋭い」などベースライン左側の感性ワードを多くの被験者が感じており、他の画像と大きく違うことが分かる。

図11を見ると、画像4では、「きれいな」、「不思議な」を評価している被験者が多いものの、少数の被験者しか評価していない感性ワードが多く、ばらつきがあることが分かる。つまり、個人差の大きい画像であることが分かる。

図12を見ると、画像5では、「自然な」、「きれいな」などを評価している被験者が多い。しかし、画像4ほどではないものの、少し評価がばらついている。

5. 考察

4.2で述べたように、囲まれている感性ワードを考慮することで、各画像の特徴を把握することができた。さらに、ベースライン感性空間上に視覚的に表示させることで、感性ワード間の関係も考慮することができ、特徴が分かりやすくなると考えられる。ベースライン感性空間上での感性ワードの配置は、一般的な感性ワード同士の関係を表している。距離が近い感性ワードほど、似ている感性ワードなので、囲まれている感性ワードはかたまることが多いと考えられる。画像1、画像2、画像3では、評価されている感性ワードがかたまっており、その画像の特徴が現れている。逆に画像4では、1～2人の被験者のみが評価している感性ワードが散らばっており、個人差の大きい画像であるという特徴が現れている。

現時点では、各々の感性ワードに対して、考察を行っているだけである。さらに、複数の感性ワードの関連に対しても考察を行うことができるようになれば、より細かな特徴も掴めると考えられる。そこで、ある感性ワードが他のどんな感性ワードと同じ領域に囲まれていたか、あるいは、複数の感性ワードを同じ領域で評価していた被験者数なども調べることができるよう、システムを改良させていく予定である。

6. まとめ

本論文では、空間的印象評価法 IEMS での評価結果に対して、感性ワードに着目した分析手法を提案した。また、その分析を自動で行い、分析結果を視覚的に表示させるシステムを実装した。さらに、そのシステムを利用し、IEMS で画像の印象を評価した結果に対して分析を行うことで、各画像の特徴を掴むことができた。

今後は、複数の感性ワードでも考察を行えるようにシステムを改良していきたいと考えている。また、感性ワードに着目した分析では、感じる度合いを表している濃さについて考慮することができない。そのため、領域内の濃さも考慮した分析手法も考えていく必要がある。

参考文献

[1] C. E. Osgood, G. J. Suci and P. H. Tannenbaum, The Measurement of Meaning, University of Illinois Press, (1957)

[2] 崔烘碩, 岡崎章, 製品の感性評価ツールの開発ー概念モデルの可視化を中心にー, 第 13 回日本感性工学会大会予稿集, I11(CD-ROM) (2011)

[3] 田崎慎也, 岡崎章, 感性を直接的に定量化する方法の提案, 第 10 回日本感性工学会大会予稿集, 12H-04(CD-ROM) (2008)

[4] 柏崎尚也, 感性の変化を捉える感性パラメータの提案, 感性工学研究論文集, Vol.4, No.1, 31-34 (2004)

[5] 赤井俊介, 宝珍輝尚, 野宮浩揮, 感性のあいまい性を考慮した印象評価法の提案, 第14回日本感性工学会大会予稿集, E5-1(CD-ROM) (2012)

[6] S. Akai, T. Hochin and H. Nomiya, Impression Evaluation Method Considering the Vagueness of Kansei, Proc. of 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD2012), 385-390 (2012)

[7] S. Akai, T. Hochin and H. Nomiya, A method for analyzing the results obtained through the impression evaluation method considering the vagueness of Kansei, Proc. of the 1st International Symposium on Affective Engineering (ISAE2013), 27-32 (2013)

[8] S. Akai, T. Hochin and H. Nomiya, Effectiveness of the Analysis Method for the Impression Evaluation Method Considering the Vagueness of Kansei, Studies

in Computational Intelligence, Volume 492, Springer, 187-201 (2013)

[9] 金明哲, R と多次元尺度法, ESTRELA10, 56-61 (2005)

[10] メイビス, PHOTO FREE WAY, メイビス CD-ROM 出版局, (1994)

[11] Moonpocket, 素材辞典 Vol. 102, 株式会社データクラフト, (2001)

[12] Moonpocket, 素材辞典 Vol. 186, 株式会社データクラフト, (2007)

[13] Moonpocket, 素材辞典 Vol. 177, 株式会社データクラフト, (2006)

The role of metadata in the Balanced Corpus of Contemporary Written Japanese in the analysis of register

ホドシチェク・ボル*, 山元啓史**

Bor HODOŠČEK*, Keiji Yamamoto**

*国立国語研究所,

**東京工業大学／カリフォルニア大学サンディエゴ校

Summary: This study proposes to evaluate the discriminatory power of the metadata contained within the Balanced Corpus of Contemporary Written Japanese (BCCWJ) for the modeling of linguistic variation (register). The available metadata is analyzed into several categories thought to influence register (NDC category hierarchy, gender, topic, media, etc.), which are then used to partition the documents within the corpus along different category groupings. The resulting similarity scores between the linguistic features of the category groupings reveal the relationships between--as well as the constraints and gaps within--the metadata, which is essential information for the reliable measurement of differences in register.



RGB-D カメラを用いた
テクスチャの特徴点と 3 次元点群との対応付けによる
リアルタイム 3 次元形状復元
KinectFusion Aided by Matching feature points
In Color Images

鈴木 遼平, 徐 剛
Ryohei Suzuki and Gang Xu
立命館大学大学院 情報理工学研究科, 草津市野路東 1-1-1
Ritsumeikan University, 1-1-1 Nojihigashi, Kusatu, Shiga

あらまし: 本稿では, 3 次元点群とテクスチャ情報を用いた 3 次元形状モデルの生成と自己位置姿勢推定手法を提案する. 距離センサより得られる 3 次元点群を用いた手法では, 壁や床のような平面の場所ではセンサの位置姿勢を求めることが困難である. また, 同形状の物体や回転体といった形状が変わらない物体などに対しても位置姿勢の推定が困難となる. そこで本提案手法では, テクスチャ情報の特徴点のフレーム間での対応付け, 同フレームにおける特徴点と 3 次元点群の対応付け, フレーム間の特徴点マッチングの結果を利用し, 3 次元点群の位置合わせを行う. これにより, 2 次元特徴または 3 次元特徴のどちらか一方がある場所でのセンサの位置姿勢を求めることができる.

Summary: As robots develop, importance of autonomous mobile robots in various fields increases. This paper proposes a new approach to real-time pose estimation and dense 3D reconstruction using both color images and 3D point clouds obtained by Kinect. At places with no 3D features such as walls or floors, it is difficult to estimate sensor poses by using only 3D point clouds. However, even in places with no 3D features, there are still 2D features in many cases which can be detected from color images. To aid point cloud fusion, we add feature point matching between two consecutive frames. 3D coordinates can be obtained from the depth image for these matched feature points. The 3D coordinates can then be used to align 3D point clouds. Thus, we can estimate sensor poses at places with either 3D features or 2D features. This new function is added to KinectFusion.

キーワード: 3 次元形状復元, 自己位置姿勢推定, 特徴点, SLAM

Keywords: 3D reconstruction, self-localization, feature points, SLAM

1. はじめに

1.1. 研究背景

近年の社会問題となっている少子高齢化により近い将来, 老人の介護や施設内での作業を行う人手の減少が予想される. また, 福島原発の事故現場のような人間の活動が制限された場所での作業の必要性が高まってきている. これにより, 現在の周辺環境を把握し, 自律移動を行うロボットの期待が高まっている. このような自律移動ロボットの実用化に向けた課題の 1 つに, 身の回りの物体や置物などの環境と自分の位置を認識するといったものがある. これは周辺環境の

3 次元モデルを作成し, さらに自分の位置姿勢を推定する SLAM (Simultaneous Localization and Mapping) [1] 問題と置き換えることが出来る. SLAM の手法には入力情報にテクスチャ情報を用いるものと, 距離情報を用いるものが挙げられる. ここで 3 次元モデルの作成のためには, 3 次元計測が必要である. 3 次元計測を行う方法として, カメラを用いる方法やレーザーレンジセンサなどの距離センサを用いる方法がある. カメラによる測定はステレオ法 [2] や SfM (Structure from Motion) [3] が挙げられる. これらは複数枚の RGB 画像からそのシーンの 3 次元形状復元

を行う。この測定は比較的手軽に行えるが、3次元形状推定のため計測精度があまり高くない。さらに、スケールが決まらないというスケールの不定性という問題も挙げられる。一方、距離センサによる測定は高速かつ正確であるが、カメラと比較すると高価である。しかし、近年 Kinect[4]の登場により比較的安価に距離センサの入手が可能となった。Kinect は Microsoft 社製の RGB カメラや距離センサ、マルチアレイマイクロフォンなどの様々なセンサを持ち、それぞれのセンサを同時に使用することが可能なデバイスである。テクスチャ情報(図 1)と距離情報(図 2)の両方の取得が可能である。カメラによる計測手法の自然特徴点のみの 3次元形状復元のため、結果が疎になりやすく、密な 3次元形状モデルの生成は困難である。今回の 3次元計測には距離センサを用いることとする。

距離センサを用いて物体の 3次元形状復元を行う。このとき距離センサで取得できる距離情報は 1 方向からの情報、センサが向いている部分のみのため、図 3 のように側面や反対側の形状は取得できない。このため、物体の周囲を抜けがないように複数回計測し、複数計測データの統合を行う必要がある(図 5)。計測データの統合を行う場合、図 4 のように計測データ間の対応を求め、位置合わせを行う。位置合わせとは共通部分の対応により、計測データ間の回転と並進の変換行列を求めることである。つまり位置合わせは距離センサ間の相対的な幾何関係、自己位置姿勢が求めることである。これらの位置合わせや統合手法はいくつか存在する[2]。しかし、計測データは 3次元点群のため、3次元における幾何関係を求める必要があり、重複領域の計測誤差や形状をどのようにするかという問題がある。また多くの計測データの位置合わせを行う場合には、累積誤差が発生する。累積誤差とは、2つの計測データ間では無視出来るような誤差でも、全体となると累積し大きな誤差となる。そのため 2つのデータ間計測データ全体で累積誤差が最小になるように最適化する必要がある。このように多くの処理を 3次元形状復元では行うため、メモリ及び計算に膨大なコストがかかる。そのためリアルタイムに位置合わせや自己位置姿勢推定を行う場合には GPU (Graphics Processing units)を用いた並列処理が必要となる。また、平面や球体のように 3次元形状特徴が少ない環境では、位置姿勢推定が困難という問題が存在する。



図 1 テクスチャ情報(RGB 画像)

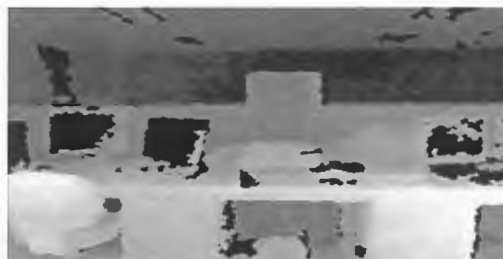


図 2 距離情報(深度画像)



図 3 Kinect が取得する
1 フレームの 3 次元点群



図 4 3次元点群同士の対応付け



図 5 3次元点群の位置合わせ

1.2. KinectFusion

距離センサを用いた手法に KinectFusion[5] がある。Kinect から取得出来る距離情報 (3 次元情報) を用いてリアルタイムに自己位置姿勢推定と図 4 のような実環境の密な 3 次元形状復元を行う。出力結果は 3 次元点群だけでなく、メッシュ化したモデルも出力可能となっている。通常、3 次元形状復元は計算コストが高く、リアルタイムでの処理は困難となっている。しかしながら、Coarse-to-Fine や GPGPU を用いることでリアルタイム処理を実現している。Frame-to-Model (フレームとモデル間の) 位置合わせを用いることで、累積誤差が少なく、安定して位置姿勢の推定が可能となっている。また Truncated Signed Distance Function (TSDF) を用いることで、モデルの表面推定や外れ値の除去を行っている (図 5)。しかし、距離情報を用いているため、壁や床のような平面や球体のように回転しても形状の変わらないものなどを対象とした場合、位置姿勢推定が困難になる問題がある。



図 6 KinectFusion の実行結果

左：取得した 3 次元点群
右：生成された 3 次元モデル

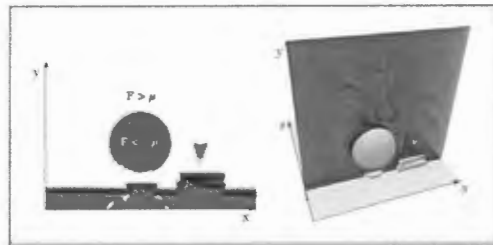


図 7 表面の推定

1.3. 問題定義

テクスチャ情報を用いた手法では、同形状の物体でも、色や模様といったテクスチャによる判別が可能である。しかし、3 次元モデルのスケールが定まらないスケールの不定性や照明変化に弱いという問題が挙げられる。一方、距離情報を用いた手法では、完全な 3 次元形状情報の取得が可能であり、照明変化にも強いという利点がある。しかし、平面や球体などの 3 次元特徴が少ない場合、自己位置姿勢推定が困難であることや、同形状や回転体の場合に誤対応が起きるという問題が存在する。本研究では、テクスチャ情報か 3 次元形状のどちらかに特徴が存在する場所での自己位置姿勢推定を行うことを目指す。

1.4. 研究目的

本研究では、テクスチャ情報と 3 次元情報の両方を用いて、自己位置姿勢推定及び 3 次元形状復元をリアルタイムで行うことを目的とする。テクスチャ情報と距離情報の両方を用いることで、一方の問題を他方の情報で補う。今回、RGB-D カメラからテクスチャ情報と 3 次元情報の両方を同時に取得する。1 つのピクセルが RGB-D、4 つの情報を保持している (図 8)。今回はテクスチャ情報 (RGB 画像) より特徴点の抽出を行う。フレーム間で特徴点のマッチングを行う。その後、特徴点マッチングによる対応を用いた 3 次元点群同士の位置合わせを行うことで自己位置姿勢の推定を行う。また、対応点が少ない場合は 3 次元点群のみを利用し、ICP アルゴリズムを用いて、位置合わせを自己位置姿勢推定を行う。得られた結果を用いて、世界座標系へと点群の統合やモデルの表面の推定を行う。今回、3 次元点群同士の位置合わせや表面推定の処理は KinectFusion を利用する。そのため、テクスチャ情報と 3 次元点群を用いた位置合わせ処理を KinectFusion に追加することで実現する。

本稿では、2 節で特徴点の処理について、3 節で 3 次元点群の位置合わせ、4 節で実験結果、5 節で結論及び考察を述べる。

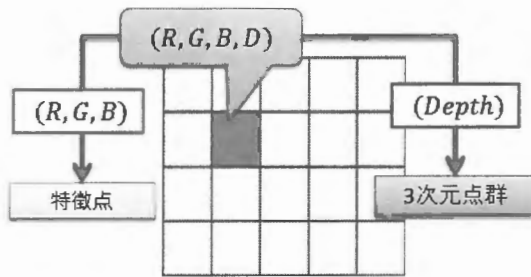


図 8 RGB-D 画像

2. 特徴点

テクスチャ情報である画像を用いて、自己位置姿勢を求める場合には画像同士の対応する特徴点の対応を求める必要がある。1999 年に Lowe らによって SIFT が提案された。SIFT は回転不変性、スケール不変性を持ちなおかつ高精度にマッチング可能である。しかし、SIFT は処理コストが高くリアルタイムなマッチング困難であった。そこで積分画像を利用し BOX フィルタを用いて高速化した SURF が提案された。近年、より省メモリかつ高速なキーポイント抽出や特徴量の記述が求められ、高速なキーポイントの抽出手法として FAST が提案された。また、高速なマッチングを行うためにバイナリコードを特徴量記述に利用した BRIEF[6] も近年提案された。しかし、BRIEF には回転不変性がないという問題点がある。

本研究では、リアルタイムでの処理を目的としているため、高速に抽出が可能であることと特徴量によるマッチングが可能であり、スケールや回転に対しロバストな手法が望ましい。そこで、高速に特徴点抽出や特徴量マッチングを行うことが可能な ORB(Oriented FAST and Rotated BRIEF)[7]を用いて特徴点の処理を行う。ORB は以下のような特徴を持つ。

- ・特徴点抽出が高速
- ・スケールに対しロバスト
- ・回転不変性
- ・特徴量をバイナリコードにより記述
- ・ハミング距離による高速なマッチング

以下の項から ORB について説明していく。

2.1. キーポイント検出手法

ORB のキーポイント検出手法には、高速に検出可能な FAST を用いる。FAST について簡潔に説明する。FAST は、パッチ中心の輝度値と周囲の画素の輝度値を比較し、明るいまは暗い画素が一定以上続く

場合はコーナーやエッジとして検出する。周囲 16 ピクセルの連続性を確認するために、上下左右 4 点の輝度値を調査することで、コーナーかどうかの判定を行う。更に高速な検出を行うために、学習した 3 分木によってコーナーの検出をする。木構造の検出器のため、高速に検出が可能となる。しかし、FAST はオリエンテーションの算出が出来ないため、回転に弱いという問題がある。回転に対してロバスト性を得るためには、オリエンテーションを算出する必要がある。オリエンテーションを基準にすることで、入力画像が回転をしていても、同一な特徴量として検出することができる。そこで ORB では、回転不変性を得るために Oriented FAST を提案している。オリエンテーションの算出にはパッチの輝度の重心とキーポイントの中心のベクトルを用いる方法である。モーメント算出を以下に示す。

$$m_{pq} = \sum_{x,y} x^p y^q I(x, y) \quad (1)$$

ここで、パッチ内の座標を (x, y) 、 p, q は各軸方向のモーメントを求めるための値であり、0 または 1 を取る。式(1)を用いて式(2)によりパッチの重心位置を算出する。

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (2)$$

算出した重心位置とパッチ中心との方向ベクトルをオリエンテーション方向とする。パッチの中心を O 、パッチの重心を C とするとオリエンテーションは $\overrightarrow{OC}$ となる。そして、オリエンテーションが作る角度は以下の式(3)で求められる。

$$\theta = \text{atan2}(m_{01}, m_{10}), \quad (3)$$



図 9 オリエンテーション方向

2.2. 特徴量記述

バイナリコードによってキーポイントの特徴量記述を

行う手法として、Binary Robust Independent Elementary Features (BRIEF)が提案されている。従来の SIFT や SURF では高次元の実数を用いていた。しかし、高次元の実数を使用した場合、メモリ容量の増加やマッチングの計算コストの増大などの問題がある。そこで、バイナリコードによって特徴量記述を行うことで、省メモリ化を図り、類似度計算にハミング距離を用いることで高速化を実現した。BRIEF について説明する。ノイズに対するロバスト性を得るために、パッチにスムージングをかけておく。パッチ内において選択された 2 点の輝度値の符号によりバイナリコードを作成する。以下に式を示す。

$$\tau(\mathbf{p}; \mathbf{x}, \mathbf{y}) = \begin{cases} 1 : \mathbf{p}(\mathbf{x}) < \mathbf{p}(\mathbf{y}) \\ 0 : \mathbf{p}(\mathbf{x}) \geq \mathbf{p}(\mathbf{y}) \end{cases} \quad (4)$$

バイナリセットを τ ，スムージングをかけたパッチを $\mathbf{p}$ とする。選択された 2 点の座標をそれぞれ $\mathbf{x}, \mathbf{y}$ とし、 $\mathbf{p}(\mathbf{x})$ はある $\mathbf{x}$ 座標における輝度値を示す。特徴量は以下の式を用いて、 n 列のバイナリコードとする。

$$f_n(\mathbf{p}) := \sum_{i=1}^n 2^{i-1} \tau(\mathbf{p}; \mathbf{x}_i, \mathbf{y}_i), \quad (5)$$

バイナリコードの長さは $n = 256$ としている。なお、選択するピクセルは、キーポイント位置を中心としたガウス分布により選択している。また、ノイズに対するロバスト性を得るために、パッチにスムージングをかけておく。

ORB では、高精度にマッチングをするために学習を用いてピクセルを選択している。選択するピクセル位置は、ペアのビット分散が大きく、なおかつ N 組みのペアの相関が低いときに特徴量記述能力が高いバイナリとして特徴記述に使用する。また、バイナリコードのパターン例を以下の図 10 で示す。相関の少ないペアの例である。ORB では、Greedy アルゴリズムを用いて、学習し選択を行う。

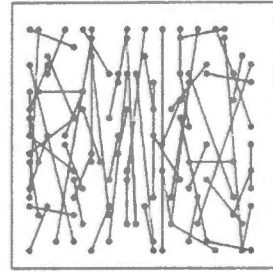


図 10 学習したパターン例

2.3. 特徴点マッチング

前述の特徴量を持った特徴点同士のマッチングを行う。各特徴点はバイナリコードで記述された特徴量を保持している。ある時刻 M における特徴点のセットを F_M とする。 F_M と F_{M-1} のフレーム間でのマッチングを行う。このとき、バイナリコード同士を排他的論理和 (XOR) で演算することができ、ハミング距離での高速な評価が可能となる。図 11 に例を示す。本手法では、特徴点マッチングの結果を位置合わせの処理で利用する。

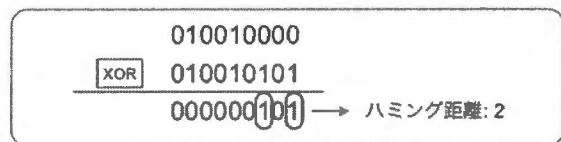


図 11 ハミング距離

3. 位置合わせと3次元統合

3次元モデルを生成するには、計測データの統合が不可欠となる。各カメラ座標系の計測データを同一座標系、つまり世界座標系へ変換することが必要となる。各カメラ座標系から世界座標系への変換には、自己位置を示す並進ベクトル $\mathbf{t}$ ，姿勢を表す回転行列 $\mathbf{R}$ を用いる。

本手法では、2 種類の手法を用いて、自己位置姿勢である回転行列と並進ベクトルの算出を行う。

3.1. 特徴点对応を利用した位置合わせ

ここでは前述で求めた特徴点对応を利用し、3次元点群同士の位置合わせについて述べる。壁や床のような 3 次元特徴がない場所でも、その場所に絵や模様のような 2 次元特徴が存在している場合に対応できるようにするためである。また、特徴点マッチングを行

った結果, 特徴点対応が 3 点以上ある場合にこの処理を行う。図 12 で示すように 3 つの対応を用いて, 位置合わせする。1 つ目の対応は RGB-D 対応である。今回 Kinect を用いているため, RGB カメラと Depth カメラの幾何対応が既知となっている。そのため, 同時刻のフレーム M おいて, RGB-D のデータが得られる。これにより, 特徴点とそれに対応する 3 次元点の登録することができる。2 つ目の対応は, ある時刻 M と M-1 のフレーム間における特徴点の対応である。2 節での処理により, 対応がわかる。\$F_M\$ の \$i\$ 番目の特徴点 \$F_M(i)\$ が持つ 3 次元点を \$X_{iM}\$ とする。ここで \$F_M(i)\$ に対応する \$F_{M-1}\$ 上の特徴点を \$F_{M-1}(i)\$ とすると, \$F_{M-1}(i)\$ が保持している 3 次元点 \$X_{iM-1}\$ と \$X_{iM}\$ も対応していると考えられる。この 3 次元点同士の対応が 3 つ目の対応である。これらを用いて, 以下の評価式で変換行列を求める。

$$E = \sum_{i=1}^n \|X_{iM-1} - (RX_{iM} + t)\|^2 \quad (6)$$

評価式 E を最小化することで, 回転行列と並進ベクトルを得る。また, 位置合わせの際に, RANSAC を用いることで, ロバストに特徴点の誤対応を除去している。

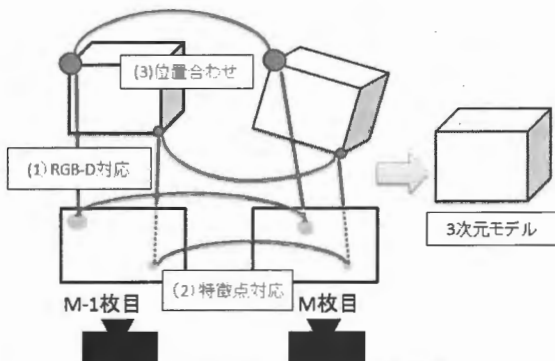


図 12 特徴点対応が存在する場合

3.2. 3 次元点群のみの位置合わせ

ここでは, 特徴点対応が 3 点未満の場合の位置合わせについて述べる。図 13 のように特徴点の対応がない場合, 3 次元点群同士のみの位置合わせにより, 回転行列と並進ベクトルを求める。特徴点の対応が 3 点未満しか存在しない場合, センサの位置姿勢を決定するのに, 特徴点の対応だけでは十分な拘束を得ることが出来ない。そのため, 位置姿勢の推定には, オリジナルの KinectFusion を使用する。しかし, 2

点の対応が存在している場合, 特徴点の対応を利用して, 並進ベクトルが定まる。そのため, 拘束条件を用いて回転行列の算出ができる。

KinectFusion では, フレーム-モデル間で ICP アルゴリズムを用いて, 位置合わせを行っている。フレーム-モデル間での位置合わせのため, 累積誤差を考えずに安定して 3 次元モデルの生成が可能となっている。このとき, 画像ピラミッドをモデル, フレームでそれぞれ作成する。そして, 画像ピラミッドを用いて, Coarse-to-Fine の考え方でセンサの位置姿勢の推定を行っている。また, 画像ピラミッドの各レベルにて, 距離関数を使用している。距離関数にモデルの点や表面からどれくらいの距離があるのかを登録しておくものである。表面を 0 とする。これにより計算コストの削減や, 外れ値の除去を行っている。

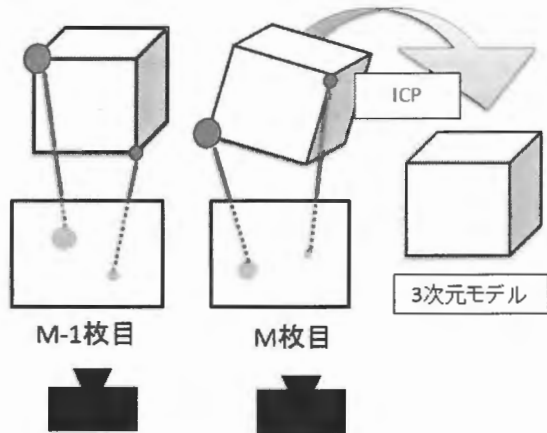


図 13 対応が 3 点未満の場合

3.3. 3 次元統合

この節では, 前節で求めた変換行列を用いて, 逐次的にカメラ座標系の 3 次元点群を世界座標系へと追加し, 3 次元形状復元を行う。ここでの処理はオリジナルの KinectFusion の処理を利用する。これから KinectFusion の処理の簡潔に説明する。KinectFusion では, 3 次元形状復元を行う際に, Truncated Signed Distance Function(TSDF), 切り捨て符号付き距離関数 \$F\$, を用いて, 復元されるモデルの表面推定を行う。各時刻の距離関数からも最も 0 に近い面, ゼロ交差を表面として推定する。このゼロ交差を求める際に, \$F > \mu\$ を空間, \$F < -\mu\$ を物体の内部として表現している。\$-\mu \leq F \leq \mu\$ の時に表面として推定している。これにより, 最もそれらしい表面の推定が可能となっている。

4. 実験

手法の有効性を示すために RGB-D カメラを用いて実環境の 3 次元形状復元実験について述べる。

4.1. 実験概要

本手法の有効性を検証するために、RGB-D カメラを用いて、実環境での 3 次元形状復元と自己位置姿勢推定を行う。実験には、平面での自己位置姿勢推定の困難の解決を示すため、ホワイトボードを対象とした実験を行う。

ここではホワイトボードに任意の文字や模様を描き、2 次元特徴であるテクスチャ情報を取得できるような環境で行い、3 次元形状復元ができていないか定性的に評価する。

4.2. 実験環境

本研究での開発・実験環境を表 1 に示す。実験で使用する RGB-D カメラは Microsoft 社の Kinect を用いる(図 14)。先にも述べたが、Kinect には IRONING プロジェクタ、IR カメラ、RGB カメラ搭載されており、640×480 ピクセルでテクスチャ情報と距離情報を 30fps で取得することが出来る。また、IR カメラと RGB カメラ間でのキャリブレーションができていないため、同じ位置姿勢からテクスチャ情報、距離情報を取得することが可能である。

表 1 開発・実験環境

CPU	Intel Core i7 3.9GHz
メモリ	16GB
OS	Windows7 (64bit)
開発環境	Visual Studio 2010
使用ライブラリ	OpenCV/PCL



図 14 Kinect

4.3. ホワイトボードを用いた実験

実験に用いる対象のホワイトボードを図 15 に示す。対象の周りを中心に撮影を行い、3 次元形状復元実験を行う。RANSAC を行い、誤対応を除去したフレーム間の特徴点マッチングの結果を図 16 に示す。提

案手法によるホワイトボードの復元結果を図 17 に示す。図 18 には、KinectFusion による復元結果を示す。特徴点実験の結果を表 2, 3 にまとめる。実験結果より、平面において従来手法よりも 3 次元形状復元されていることがわかる。

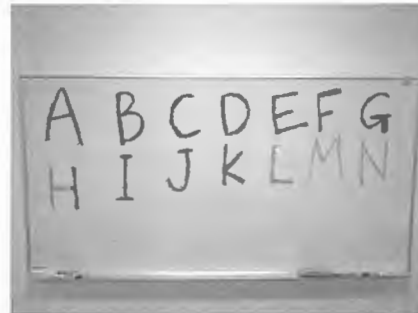


図 15 テクスチャを持つホワイトボード



図 16 特徴点マッチングの結果画像

表 2 特徴点の実験結果

	点数
特徴点 (m=100)	467
全対応点	458
対応点 (RANSAC 後)	349
誤対応 (RANSAC 後)	2
平均抽出点数	567 / f

表 3 特徴点関連の処理速度

抽出・マッチング (平均)	22.3(ms)
抽出・マッチング(最大)	26.2(ms)
位置合わせ(平均)	35.1(ms)
位置合わせ(最大)	38.4(ms)

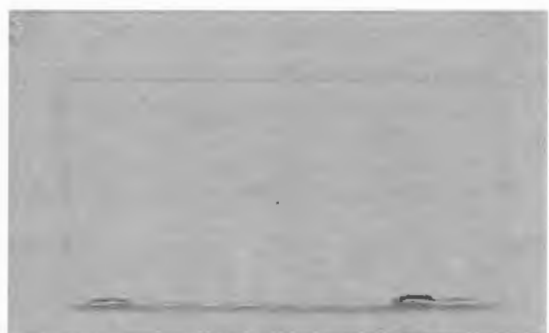


図 17 平面の復元結果 1 (提案手法)

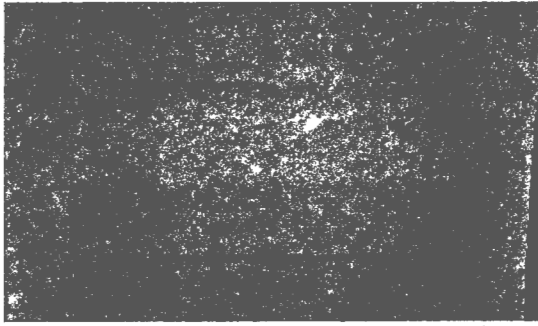


図 18 平面の復元結果 2 (KinectFusion)

5. まとめ

本稿では、テクスチャ情報、距離情報の両方を用いて 3 次元特徴の少ない場所において、センサの位置姿勢を推定し、3 次元形状復元を行う手法を提案した。テクスチャ情報において、局所特徴量の ORB を抽出された特徴点のマッチングに使用した。また、特徴点と対応した Kinect から同時に得られた 3 次元点を位置合わせに用いた。また、位置合わせの際に RANSAC を使用することで、特徴点同士の誤対応も除去することに成功した。これにより、オリジナルの KinectFusion において 3 次元特徴の少ない、平面や球体での位置姿勢推定で有効であると考えられる。2 次元特徴または 3 次元特徴のどちらか一方が存在する場所での SLAM を実現した。ただ、今後の展望としては、評価対象を増やし、定量的な評価や高速化を行っていく。

謝辞

本研究の一部は、私立大学戦略的研究基盤形成支援事業と芸術・文化分野の資料デジタル化と活用を軸とした研究資源共有化研究の支援を受けた。

参考文献

- [1] B. Williams, G. Klein, and I. Reid, "Real-Time SLAM relocalisation," *In Proc. 11th IEEE International Conference on Computer Vision (ICCV'07)*, pp.1-8, Rio de Janeiro, October 2007.
- [2] Gang Xu, Zhengyou Zhang, "Epipolar Geometry in Stereo, Motion and Object Recognition A Unified Approach," Kluwer Academic Publishers 1996.
- [3] Akihito Torii, Yuuki Hanzawa, Masatoshi Okutomi, "Easy-to-Use Online SfM System," *SSII2011*, pp.1-6, June 2011.
- [4] Microsoft, "Kinect, " <http://www.xbox.com/ja-JP/kinect>
- [5] Richard A. Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J. Davison, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Andrew Fitzgibbon, "KinectFusion: Real-time dense surface mapping and tracking, " *ISMAR '11*, pp.127-136, Basel, Switzerland, 26-29th October 2011.
- [6] M. Calonder, V. Lepetit and C. Strecha and P. Fua, "BRIEF: Binary Robust Independent Elementary Features, " *In Proc. European Conference on Computer Vision 2010*
- [7] E. Rublee, V. Rabud, K. Konolige and G. Bradski, "ORB: an efficient alternative to SIFT or SURF", *In Proc. International Conference on Computer Vision*, 2011.
- [8] Point Cloud Library: <http://pointclouds.org/>
- [9] OpenCV: <http://opencv.org/>

浮世絵を対象とした多言語・異種データベースの横断検索

Providing simultaneous access to diverse and multilingual Ukiyo-e databases

バトジャルガル ビルゲサイハン¹ 木村 文則¹ 前田 亮²

Biligsaikhan Batjargal¹ Fuminori Kimura¹ Akira Maeda²

¹ 立命館大学衣笠総合研究機構, 京都府京都市北区等持院北町 56-1

² 立命館大学情報理工学部, 滋賀県草津市野路東 1 丁目 1-1

¹ Kinugasa Research Organization, Ritsumeikan University,

56-1 Toji-in Kitamachi, Kita-ku, Kyoto

² College of Information Science and Engineering, Ritsumeikan University,

1-1-1 Noji-higashi, Kusatsu, Shiga

あらまし: 本論文では、浮世絵を対象として、多言語かつ異種のデータベースに対して同時に検索し、それらの検索結果を統合して利用者に提示する手法を提案する。これにより、入力インタフェースの異なるデータベースに対しても、一つの問合せで検索可能である。本システムではさらに、英語を用いる利用者が英語の検索語を用いて日本語のみで提供されているデータベースを検索可能にするため、英語と日本語間の言語横断検索の機能を提供している。

Summary: In this paper we introduce a prototype federated search system for accessing Ukiyo-e prints in different languages simultaneously, which retrieves diverse multiple databases in parallel and integrates multilingual retrieved results on-the-fly. The proposed system allows searching and browsing multiple Ukiyo-e databases in parallel as well as displaying diverse multilingual results in a user-friendly way. This system also has a facility for cross-language search between English and Japanese, which enables English-speaking users to search databases available only in Japanese using English query.

キーワード: 画像データベース, 浮世絵, メタサーチ, 横断検索

Keywords: image database, Ukiyo-e, metasearch, federated search

1. はじめに

近年、図書館、文書館、博物館など多くの文化施設がインターネット上でデジタル化された所蔵品の公開を始めており、その結果として、同種のデータを対象にしているにもかかわらず、実際には異種のデータベースが混在する状況となっている。本論文では、世界中に散在する浮世絵画像のデータベースに対して統合的なアクセスを提供する手法について述べる。浮世絵は、江戸時代における芸術の一つとして世界的に高い評価を受けている。最近の情報技術の進歩およびインターネットの普及に伴い、多くの浮世絵がデジタル化され、検索可能なデータベースや、美術館 Web

サイトにおける展示、あるいはスキャンされた画像のコレクションなどの形で公開されている。しかしながら、浮世絵画像の保存方法はコレクションによってまちまちである。その結果、インターネット上の浮世絵画像へのアクセスは、一貫性のない様々なインタフェースによって提供されているのが現状である。利用者にとっては、様々なインタフェースを持つ多様なデータベースから簡単に必要な情報を得ることは、非常に時間と手間がかかる状況になっている。

本論文では、我々が開発している、複数の異なる浮世絵画像データベースを同時に検索し、検索結果をその場で統合する浮世絵の統合検索システムについて

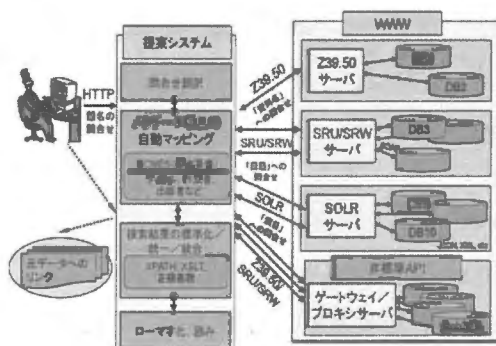


図 1: 提案システムの構成

て述べる。提案するシステムは、最新の情報技術を活用することで、複数の浮世絵データベースに対する検索や閲覧を可能とし、さらに利用者にとって使いやすい形での検索結果の表示を実現する。これにより、利用者の時間と手間を節約することができると考えられる。本システムではさらに、英語を用いる利用者が英語の検索語を用いて日本語のみで提供されているデータベースを検索可能にするため、英語と日本語間の言語横断検索の機能を提供している。

2. 浮世絵データベースに対する統合検索のアプローチ

我々が提案する横断検索システムの構成を図 1 に示す。図 1 に示すように、たとえば利用者がタイトルにある検索語が含まれる浮世絵を検索したい場合、本システムは複数の浮世絵データベースから、タイトルもしくはタイトルに相当するメタデータ項目にその検索語が含まれるものを検索する。同様に、作者・著者・主題などの他の項目による検索が可能である。提案システムの特徴は次の通りである。(1) 単一のインタフェースと一つの検索語で複数の浮世絵データベースへの同時アクセスを実現する。(2) 異なる言語で提供されている多様な浮世絵データベースのメタデータ項目のマッピングを行う。(3) 日本語で提供されている浮世絵データベースに対して英語での検索および閲覧を実現する。(4) 複数のデータベースから得られた検索結果をその場で集約・統合し、単一のインタフェースで利用者に提示する。

横断検索技術は、データベースへの問合せの結果生成された動的なページにアクセスすることが可能である [1]。さらに、Bergman [2] や Raghavan ら [3] の報告によると、主要な Web クローラは Web の表層にあるリソースに対してのみ収集・索引付けを行うが、Web の深層にあるリソース（例えば Web 上に公開されているデータベースに格納されている情報）については収集・索引付けを行わない。言い換えると、横断検索技術は、Web サイトや動的ページなどで提示される Web

の深層にあるリソース [2] [3] など、複数のデータベースを一つの間合せで同時に検索するために利用することが可能である [1] [4]。

我々の手法は、横断検索、すなわち分散情報検索の範疇に属する。複数のデータベースにアクセスするための他の手法として、事前に様々なリソースからデータを集める技術であるハーベスティングやクロールがある。これらの手法はたいいてい、頻繁な更新が要求され、かつ大規模なストレージに格納されている索引のマスターを使用する。頻繁に更新しなければ各データベースに格納された情報は最新ではなくなってしまうため、新しく更新された情報は検索に利用できないということも起こりうる。

2.1. 統合検索の実現のための課題

本節では、浮世絵画像の横断検索システムを実現するに当たって直面したいいくつかの課題と、それらをいかに解決したかについて述べる。このようなシステムを実現するには、「複数データベースへの同時アクセス」および「検索結果の集約と統合」という二つの大きな課題に対処する必要がある。以下でそれぞれの課題と解決策について述べる。

2.1.1. 複数データベースへの同時アクセス

提案システムのプロトタイプは、次に挙げる二種類のデータベースに対するアクセスおよび検索に対応している。(1) カタログ化・メタデータ作成・データ交換の方法が標準規格に準拠しているもの。(2) 独自のメタデータスキーマを持ち、それぞれ異なるインタフェースやプロトコルでアクセス可能な Web データベース。本システムでは、検索プロトコルの規格である Z39.50、Search/Retrieve Web service (SRW)、Search/Retrieve via URL (SRU)、OpenSearch などに準拠したデータベースに対する同時検索が可能である。本システムは、無料で利用可能なオープンソースソフトウェアである Pazpar2 [5]、YAZ [6]、SimpleServer [7] などをカスタマイズすることで開発している。これらを用いることで、本システムにおいて Web の深層にあるリソースも含んだ様々な種類のデータベースを同時に検索することが可能となる。

2.1.2. 検索結果の集約と統合

ここでは、複数の浮世絵データベースを同時に検索し、その検索結果を統合する際に問題となるいくつかの点について述べる。多様なデータベースから得られた検索結果を集約して統合することは、一貫性のない様々なインタフェースやプロトコルによってアクセスさ

れる多言語のデータベースに対しては特に困難である。各データベースからの検索結果は、データベースによって単純なものから詳細なものまで様々なものが存在し、さらに様々なメタデータ形式や HTML など様々な種類のフォーマットによって返される。

そのため、本システムでは検索結果をその場で統合するために、ダブリンコアメタデータ要素セット (Dublin Core Metadata Element Set: DCMES) に基づく単純なメタデータスキーマを採用した。具体的には、検索結果を単一のインタフェースによるリストの形で表示するため、<title>、<creator>、<subject>、<description>などの基本的な要素を採用した。

もし、ダブリンコアの基本要素に準拠したリモートデータベースが複数の出力フォーマットを提供するなら、我々は、情報の損失のない可能な限り詳細な元の情報源からの情報から返されるレコードを得ることを目指す。そのあと、返された結果と<title>、<creator>、<subject>などのダブリンコアの基本要素との対応付けのために、XSLT と正規表現が使用される。

標準規格に準拠しない Web データベースから返される検索結果は、ほとんどの場合 HTML である。これらの Web データベースに対しては、検索結果の HTML から必要な情報を抽出する「スクリーンスクレイピング」と呼ばれる技術を用いる。

提案手法のデータ抽出モジュールは、検索結果の HTML を読み込み、必要な HTML タグを抽出する。XPath と正規表現は、HTML 中の抽出対象個所の識別と出現位置の特定のために利用される。たいてい、HTML 文書中には重要な属性がいくつか含まれている。実際、<a>タグの多くは、提案システムが提供できるリソースへのリンクとしての URL へと対応付ける「href」属性を持っている。例えば、メトロポリタン美術館から得られる検索結果のレコードが以下のようなものであったとする。

```
<p><a href= "...">Azuma Bridge from Ko  
magatadô Temple, from the series One  
Hundred Famous Views of Edo</a></p>  
<p>Utagawa Hiroshige (Japanese, 1797  
-1858) </p><p>Date: 1857<br/>Accessi  
on Number: JP1023</p>
```

この例では、芸術家名とそれ以外のデータは、<p>タグおよびいくつかの任意の
タグにより分割されており、これを抽出パターンとして使うことができる。実際、<title>、<author>、<date>などのタグを含んでいる全てのフィールドは、データベースごとに固有の記述パタ

ーンを適用することより、XPath を用いて抽出することができる。HTML タグによるデータ分割が行われていない文字列から<author>の情報に相当するデータを抽出するためには、“(Made by);”, “(Print artist);”, “(Claimed to be by);”, “(Designed by)”, “(Print made by)”, “(Drawn by)”, “(Author)”, “(Published by)”, “Designed by”, “Formerly Attributed to”, “In the Style of”, “Original print designed by”などのラベルとしての表現を利用する。

さらに、日本語での標準規格非準拠の Web データベースの検索結果からタイトルや作者等の項目を同定するため、メタデータの自動マッピング手法 [8][9]を用いている。たとえば、項目名「画題等」「画題2」「役名」「外題」「所作題」「細目題」「主外題」「系統分類題」「演目(統合)」「画題統合」は、異なる意味を持つが、すべてタイトルに関する項目であるため、いずれも<title>にマッピングされる。

実際、この後の処理のために、提案手法は検索結果の HTML を記述されたメタデータのフィールドへと分割する。複数の異種の浮世絵データベースからデータの集約を行い、これらのデータのメタデータへのマッピングを行ったあと、提案システムはバイリンガルデータの生成を行い、検索結果の統合をし、利用者の見やすいように提示する。しかし、メタデータ属性<author>や<title>中にある全ての author 名や title 名を収集したとしても、メタデータの値はこの処理の段階では別々のデータベース間で異なる表現形式であるかもしれない。そこで、異なる表現を統一するため、重複排除、生誕日と死亡日の統合、姓と名の間へのコンマの付与などのような変換規則を適用することにより、さらなるデータの標準化が行われる [10]。

2.1.3. 異言語のデータベースに対する検索

英語と日本語を横断した検索の手法は、以前に我々が行った研究 [11]において実装を行った。このような特徴は非常に有用である。なぜなら、研究者が一つの言語内だけで検索・閲覧しているだけでは情報が不十分で、他の言語で記述された情報からさらなる追加の情報にアクセスする必要があるかもしれないからである。日本の美術館や公的機関において公開されているコレクションは、ほとんどが日本語でしか利用できない。それゆえ、日本語に不慣れな人は必要とする情報を見つけられないかもしれない。それに加え、浮世絵データベース中のテキストには、絵師の名前や用いられる用語には江戸時代独特のものが含まれているなど、かなり特殊である。

Home » Federated Search

utagawa hiroshige

Sort by and show per page. 61-90 of 1064

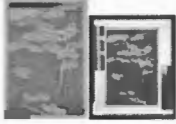
« 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 »

61. **Whale hunting at the island of Goto in Hizen.** Utagawa Hiroshige., 1859.

 Title transcription: Hizen gotō kujiraryō no zu
Authors: Utagawa Hiroshige;
Location: Library of Congress
Original records: [\[Link\]](#) [\[Link\]](#)

Information Whalers armed with harpoons hunting whales
• Title and other descriptive information compiled by Nichibunken-sponsored Edo print specialists in 2005-06
• Format: vertical Ōban Nishide
• From the series: Shokokumraisho hyakkei: One hundred famous views of the country
Edition: 1859; [\[Link\]](#); [\[Link\]](#); RISB&Tex

62. **Yoshino Mountain in Yamato Province.** Utagawa Hiroshige., 1859. »

 Title transcription: Yamato yoshinoyama
Authors: Utagawa Hiroshige;
Location: Library of Congress
Original records: [\[Link\]](#) [\[Link\]](#)

63. **Choshi Bay in Shimosa Province; Famous Places in the Provinces (Woodblock print).** Utagawa Hiroshige., 1854-1858. »

 Title transcription: Choshi Bay in Shimosa Province; Famous Places in the Provinces (Woodblock print)
Authors: Utagawa Hiroshige;
Location: Victoria and Albert Museum
Original records: [\[Link\]](#)

DATABASES

- British Museum (355)
- The Metropolitan Museum of Art (246)
- Museum of Art and Archaeology, University of Oxford (175)
- Victoria and Albert Museum (141)
- Boston Museum of Fine Arts (93)
- Library of Congress (54)

AUTHORS

- Utagawa Hiroshige (602)
- Utagawahiroshige (271)
- Suzuki Hiroshige (84)
- Utagawa Hiroshige II (84)
- Koshimura Hetsuke (6)
- Koshimurahetsuke (6)
- Utagawa Kunisada (6)
- Utagawakunisada Toyokuni III (4)
- Okazawa Taheiji (3)
- Ando Hiroshige [Utagawa Hiroshige; Ichiryusai; Ichiyusai; Ryusai] (2)
- Show 10 More Items

TIME SPAN

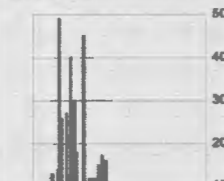


図 2: 横断検索システムのプロトタイプのスクリンショット。浮世絵の作者名を「Utagawa Hiroshige」として検索語を入力して得られた検索結果(英語表示)。

対訳辞書による問合せの翻訳を行う手法は、日本の芸術・文化分野の専門用語辞書を用いることにより採用している。提案手法における今後の課題として、例えば姓名のフルネームではなく名だけをキーワードとしたり、「絵師名 + 専門用語」などのような、様々な形式のキーワードに対応することである。例えば、もし利用者により「Kuniyoshi」という問合せが入力された場合、「Kuniyoshi」という問合せは「国芳」と翻訳され、複数の日本語の浮世絵データベースに対して検索が行われる。同様に、「国芳」という利用者からの問合せは、「Kuniyoshi」と翻訳される。提案手法における「国芳」と「Kuniyoshi」という二言語の問合せは同じデータベースへと送られ、各検索結果は一つの検索結果に統合される。

2.1.4. 二言語データの作成

本節では、どのようにして遠隔の浮世絵データベースに対する検索結果から二言語データを作成するかについて議論を行う。浮世絵のテキスト情報は日本語

または英語、時にはその両方が入手可能である。なぜなら、浮世絵は世界中に広く人気があるからである。

提案システムは日本語テキストに対して以下の追加のいくつかの課題を実行する。(1) 日本語テキストからの「読み」または「音訳」の抽出、(2) 可能であれば訳語の発見、(3) 利用者が見やすい方法での「読み(仮名またはローマ字)」、「訳語」、「漢字表記」の提示。

ある利用者は漢字表記による元のテキストを好むかもしれないが、別の利用者は翻訳のほうを好むかもしれない。またある利用者は漢字と仮名を用いた日本語を理解するかもしれないが、別の利用者にとってはローマ字表記のほうが読むことができるかもしれない。「読み」は、浮世絵に関係する専門用語の正確な発音を知るためには必要であるかもしれない。

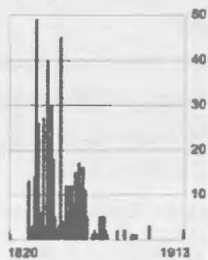
もし「読み」や「音訳」が元の情報源となるテキストから見つからない場合、日本語の形態素解析器である「MeCab」を利用して、元の日本語テキストから「読み」を生成する。日本の芸術・文化分野の専門用語辞書に掲載されている専門用語や芸術家の名前は、MeCab の辞書へ追加登録を行っている。デフォルトの

- DATABASES**
- British Museum (355)
 - The Metropolitan Museum of Art (246)
 - Museum of Art and Archaeology, University of Oxford (175)
 - Victoria and Albert Museum (141)
 - Boston Museum of Fine Arts (93)
 - Library of Congress (54)

AUTHORS IN JAPANESE

- うたがわ ひろしげ (692)
- 歌川広重 (271)
- うたがわ ひろしげ い (84)
- すずき ひろしげ (84)
- うたがわ くにさだ (6)
- こしむら へいすけ (6)
- 越村平助 (6)
- 歌川国貞 Toyokuni III (4)
- おかざわ たへいじ (3)
- Print Artis (2)
- あんど ひろしげ [うたがわ ひろしげ; いちりゆさい; いちゆさい; りゆさい] (2)
- Show 10 More Items

TIME SPAN



Home » Federated Search

utagawa hiroshige

SEARCH

Sort by **Relevance** and show **30** per page.

61-90 of 1064

« 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 »

61. ひぜん ごとうくじらりょうのず. うたがわ ひろしげ., 1859.



Title transcription: ひぜん ごとうくじらりょうのず

Authors: うたがわ ひろしげ;

[Other view](#)

Location: Library of Congress

Original records: [\[Link\]](#) [\[Link\]](#)

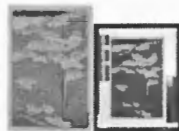
Informations Whalers armed with harpoons hunting whales

- Title and other descriptive information compiled by Nichibunian-sponsored Edo print specialists in 2005-06
- Format: vertical Oban Nishide
- From the series: Shokokumreissho hyakkei : One hundred famous views of the country

Edition: 1859; [\[Link\]](#), [\[Link\]](#);

RIS&Tex

62. やまと よしのやま. うたがわ ひろしげ., 1859. »



Title transcription: やまと よしのやま

Authors: うたがわ ひろしげ;

[Other view](#)

Location: Library of Congress

Original records: [\[Link\]](#) [\[Link\]](#)

63. Choshi Bay in Shimosa Province; Famous Places in the Provinces (Woodblock print). うたがわ ひろしげ., 1854-1858. »



Title transcription: Choshi Bay in Shimosa Province; Famous Places in the Provinces (Woodblock print)

Authors: うたがわ ひろしげ;

Location: Victoria and Albert Museum

Original records: [\[Link\]](#)

[Other view](#)

64. Fireworks at Ryôgoku Bridge, from the series One Hundred Famous Views of Edo. うたがわ ひろしげ., 1858. »

図 3: 横断検索システムのプロトタイプのスクリンショット. 浮世絵の作者名を「Utagawa Hiroshige」として検索語を入力して得られた検索結果(日本語表示).

MeCab の辞書では、日本の芸術・文化分野の専門用語に対しての「読み」を十分に提供できないためである。

ローマ字表記による「音訳」、英語による「title」、日本語「title」の英訳は、英語ページにおける日本のコンテンツを提示する際に使われる。一方、仮名表記の「音訳」、日本語による「title」、英語「title」の日本語訳は、日本語ページで提示する際に使われる。こうして利用者は、基本的な情報やオリジナルの浮世絵を見つけるためのリンクを保持した二言語での浮世絵のリストを提供される。

2.2. 横断検索システムのプロトタイプ

図 2 と図 3 は、我々が構築した横断検索システムのプロトタイプのスクリンショットである。現在のシステムでは、1915 年以前の日本の浮世絵を検索することができる。本システムでは、アメリカ議会図書館 (Library of Congress: 約 2740 件)、日本の国立国会図書館

(約 10,000 件)、大英博物館 (約 9,000 件)、ボストン美術館 (約 30,000 件)、ヴィクトリア&アルバート博物館 (浮世絵および浮世絵に関連する絵画や書籍 25,000 件以上)、立命館大学アート・リサーチセンター所蔵浮世絵検索閲覧システム (約 12,500 件)、メトロポリタン美術館 (約 4,200 件)、アシュモレアン博物館 (約 640 件)、シカゴ美術館 (約 2,100 件)、江戸東京博物館 (約 6,100 件) の浮世絵を検索対象としている。

浮世絵のテキスト情報やメタデータに関して、立命館大学アート・リサーチセンターは日本語のみで提供している。一方、江戸東京博物館や国立国会図書館は主として日本語で提供しているが、少ないながらも一部は英語でも提供している。欧米のデータベースでは、大半が英語のテキストで提供されており、日本語のテキストはごく一部でのみ提供されている。

提案した横断検索システムは、検索語に英語を用いても検索することができる。この特徴により、英語による単一のインタフェースと単一の検索語を使って日

本語のデータベースを検索・閲覧できるため、日本語の理解できない利用者にとって有用である[12]。

システムのプロトタイプのスクリンショット図2と図3は、浮世絵の作者名を「Utagawa Hiroshige」として検索語を入力し、その結果として得られた検索結果を示している。その結果、利用者は基本情報を含んだスニペットのリストにアクセスすることができる。また検索結果には、日本語または英語への切り替えリンクや、見つかった場合は浮世絵画像のオリジナルへのリンクも提供されている。基本情報のローマ字表記、英語での作品名、日本語作品名の英訳が、英語版のページに表示される。かな文字表記、日本語での作品名、日本語作品名の英訳は日本語版のページに表示される。

「Other view」ボタンは、利用者が表示される言語を手早く切り替えるためのものであり、このボタンを押すことで英語表示と日本語表示の間で切り替えられる。

3. おわりに

本論文では複数のデータベースを対象に、しかも複数の言語を対象に浮世絵データベースに対して同時並行で検索し、その検索結果の統合を即座にフロントエンドで行うという、プロトタイプシステムについて紹介した。このようなシステムは、言語横断検索という側面が日本語を理解できない利用者にとって有用であることはもちろんのこと、それだけではなく、単一のインタフェースで浮世絵データベースを検索・閲覧できることは利用者にとって時間節約の助けとなる。提案した手法は、浮世絵データベースのように様々な言語で記述されていたり、様々な書式・インタフェースから構成されているデータベースに対して、特に重要である。

我々は本システムをさらに改良し、単一の検索語・単一のインタフェースを介してより多くのデータベースに対して多言語横断検索が行えるようにすることを検討している。

謝辞

本研究の一部は、文部科学省私立大学戦略的研究基盤形成支援事業「芸術・文化分野の資料デジタル化と活用を軸とした研究資源共有化研究」、文部科学省科学研究費補助金基盤研究(C)「多言語デジタルアーカイブの統合検索に関する研究」(研究代表者:前田亮, 課題番号:24500300)の支援を受けている。

参考文献

- [1] Shokouhi, Milad and Luo Si. Federated Search, *Foundations and Trends in Information Retrieval*, vol. 5, no. 1, 2011, pp. 1-102.
- [2] Bergman Michael K. The Deep Web: Surfacing Hidden Value, *Journal of Electronic Publishing*, vol. 7, no. 1, 2001, pp. 1-17.
- [3] Raghavan, Sriram and Garcia-Molina, Hector, Crawling the hidden web, *Proceedings of the 27th International Conference on Very Large Data Bases*, 2001, pp. 129-138.
- [4] Callan, Jamie. Distributed information retrieval, *Advances in information retrieval*, vol. 7, no. 13, 2000, pp. 127-150.
- [5] IndexData, “Pazpar2,” 2007. [Online]. Available: <http://www.indexdata.com/pazpar2>. [Accessed: 10-Oct-2013].
- [6] IndexData, “YAZ,” 1995. [Online]. Available: <http://www.indexdata.com/yaz>. [Accessed: 10-Oct-2013].
- [7] Index Data ApS, “Index Data.” [Online]. Available: <http://www.indexdata.com/>. [Accessed: 10-Oct-2013].
- [8] Kimura, Fuminori, Takushi Toba, Taro Tezuka, and Akira Maeda. Federated Searching System for Humanities Databases Using Automatic Metadata Mapping. *Proceedings of the International Conference on Dublin Core and Metadata Applications*, 2009, pp.139-140.
- [9] Batjargal, Biligsaikhan, Fuminori Kimura, and Akira Maeda. Providing Universal Access to Japanese Humanities Databases: An Approach to Federated Searching System Using Automatic Metadata Mapping. *Journal of Zhejiang University-SCIENCE C*, Vol. 11, No. 11, 2010, pp. 837-843.
- [10] Batjargal, Biligsaikhan, Fuminori Kimura, and Akira Maeda. Metadata-related Challenges for Realizing Federated Searching System for Japanese Humanities Databases. *Proceedings of the 11th International Conference on Dublin Core and Metadata Applications (DC-2011)*, 2011, pp. 80-85.
- [11] Batjargal, Biligsaikhan, Fuminori Kimura, and Akira Maeda. Approach to Cross-Language Retrieval for Japanese Traditional Fine Art: Ukiyo-e Database. *Proceedings of the 14th European Conference on Research and Advanced Technology for Digital Libraries (ECDL2010)*, 2010, pp. 518-521.

《特設セッション》

立命館大学文学研究科・文化情報学専修連携企画

活用・継承する「データベース」と「データベース」による新知見

立命館大学アート・リサーチセンターは、2011年度まで、文部科学省グローバルCOE「日本文化デジタル・ヒューマニティーズ拠点」として、日本で初めてのデジタル・ヒューマニティーズを標榜した教育・研究活動を続け、大きな成果を上げてきました。2012年度以降も大学の支援を受けながら立命館大学日本文化デジタル・ヒューマニティーズ拠点として活動を続ける一方、この教育プログラムの資産を文学研究科の教育に引継ぎ、2014年度から文化情報学専修を立ち上げることになりました。ここでは、情報科学、とりわけデータベースの構築や運用を重視した、国際的な共同研究を活性化させるプロジェクト型教育が実施される予定です。

それにあたり、本セッションでは、人文科学と情報学の融合研究において、大きな成果を次々と発表しておられる高野明彦氏を講師に迎え、データベースが人文科学者らの新しい発想を生む活動を支援する方法についてお話していただきます。

また、続いて、立命館大学デジタル・ヒューマニティーズ拠点で蓄積された文化資源情報を活用し、研究活動を行っている若手研究者4名からの研究事例を発表してもらいます。データベースが人文学の研究にとって如何に有効であり、新しい知見や新しい活動を生んでいくのか、実践的な事例により説得力のある説明となるとことでしょう。

以上を受け、このセッションでは、実行委員会から高橋晴子(大阪樟蔭女子大学)、赤間亮(立命館大学)が加わり、新たな段階を迎えた人文科学におけるデータベースの位置づけについて議論します。今回は、有用なデータベースを如何に運営・継承していくのかという重要な論点についても、話題に取り上げます。

【招待講演】

データベースを発想力に変える方法

高野明彦氏

Takano Akihiko

国立情報学研究所

National Institute of Informatics



19 世紀歌舞伎のビッグデータ分析の可能性 — ヒット現象の数理モデルを用いて —

Possibility of analysis of "Big Data" of Kabuki play in 19th century using the mathematical model of hit phenomena

川畑泰子¹, 源田悦夫², 石井晃^{2*}

Yasuko Kawahata¹, Etuo Genda², and Akira Ishii^{2*}

¹九州大学 芸術工学部, 福岡市南区塩原4丁目9-1

²鳥取大学大学院工学研究科機械宇宙工学専攻, 鳥取大学 鳥取県鳥取市湖山町南4-101

¹Kyushu University, 9-1 Shiobaru, Minami-ku, Fukuoka

²Tottori University, 4-101, Koyamacyou-Minami, Tottori

あらまし: 歌舞伎は17世紀から19世紀の江戸時代における大衆娯楽だった。江戸時代の歌舞伎は、現代における伝統芸能のような立ち位置とは異なっていた。昨今における数学的モデルの方法で利用されているインターネットを通じたSNSのトピックは江戸時代には存在しない。江戸時代の人気役者の測定のために、私たちは江戸時代での唯一のメディアである紙媒体(浮世絵, 詩, 俳句, 本)を使用した。このアプローチでは、傑作の評判に関して、過去の人気俳優の研究を現在の視点からも、することが今後期待できる。

Summary: Kabuki was a popular entertainment in the Edo period in the 19th century from the 17th century. Kabuki in the Edo period was different from standing position as a traditional performing arts in modern. Methods of mathematical models in Nowadays has been selected topics in SNS through the Internet, they do not exist in the Edo period. In order to perform the calculation and measurement of popular artists of the Edo period, we were subject to publication of the paper medium is the only media at the time (Ukiyoe, poem, haiku, book) to there. With this approach, the study of past popular actor, with respect to the reputation of masterpiece can be also from the perspective of current.

キーワード: ヒット現象, 確率過程, 歌舞伎, 川柳, 浮世絵

Keywords: Hit phenomena, Stochastic process, Kabuki, Senryu, Ukiyo-e

1. 序

“ヒット現象の数理モデル”では今までSNSにおける口コミの数の時系列の変化を用いて、今まで曖昧となっていたエンターテインメントやイベントの評判が提示されてきた。今日では、このような推定はブログ、FacebookやTwitter、BBS(掲示板)、および世界中の他の同様のサービスの様なソーシャルネットワークサービス(SNS)に投稿されたコメントの数を使用することによって行うことができ、口コミの間の相互作用はデジタルデータからみることが出来る。私たちは、SNS社会でのネットワー

ク上の動きが現実の社会の動きに非常に類似していると仮定することができた。従って、我々は現実社会の観測データとして人間同士のコミュニケーションのデジタルデータを使用することができる。この観測可能な膨大なデータ(いわゆる“ビッグデータ”)を使用して、我々は社会科学への統計物理学の手法を適用することができる。本論では、この手法は現代だけでなく過去にも適応可能であると考え、江戸時代人気のあるエンターテインメントであった歌舞伎をテーマにした。江戸時代の歌舞伎は、現代のような伝統芸能としての立ち位置とは異なっていた。今日におけるヒット現象の数理

モデルのメソッドはインターネット上のSNS のトピックを選択しているが、それは、江戸時代には存在しない。計算および測定を分析するために、江戸時代の人気アーティストに関する当時の唯一のメディアである、紙媒体の出版物(浮世絵、詩、俳句、引き札、書籍)を対象とした。図1, 2のように大きな事件が起きた際のドキュメント数(話題数)のピークの類似性にも注目した。過去の事物の人気の考察をこのアプローチで、現在と同じ数理モデルで解析することを期待している。[1-3, 7]

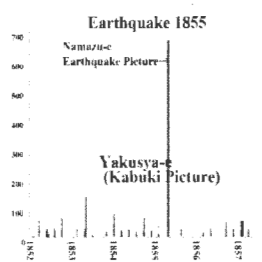


図1 1855年の安政の大震災時の地震に関するドキュメント数(話題数)

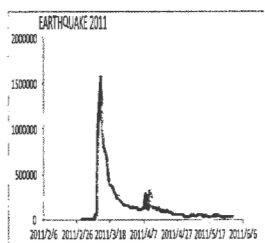


図2 2011年の東日本大地震時の地震に関するブログの総話題数

2. ヒット現象とは

ヒット現象とは、音楽や映画などのコンテンツが受け手の支持を得ることによって、発売あるいは公演と同時に販売数や観客動員数が爆発的に増え、ピークを過ぎると急速に減衰していく現象のことである。つまり、ヒットするということは、そのコンテンツが人と人の見る・聞く・話すといった触れ合いによって世間に広まっていくことである。様々なコンテンツをヒットさせるには人々の「心」をいかに動かすことができるかということがポイントとなる。ヒット現象がおこる過程での人々の「心」の動きをとらえ、人々の「心」にどのようにアプローチしていくかを考えることが、ヒット現象を解析する上で重要なことである。[3]本論では従来の口コミ情報におけるコミュニケーションの

量からの適応から歌舞伎役者の話題が発生している浮世絵・川柳・俳諧・小説の種数の推移から江戸の人々の「心」にどうアプローチさせ、そのコンテンツにおけるヒットにつなげていったかを考察するために理論を導入した。[1-3]

3. 話題の取得方法

①国内・国外にある出版物(俳諧・川柳・小説)の調査

②現存する浮世絵のデータベースからの調査

③美術館・博物館にある浮世絵の調査

を元に本研究を行った。

出版物の題名探しと全体数の確認には文献を参考に調査を行った。[8-11]

3-1. 浮世絵の数え方

浮世絵の登場する役者が誰であり、どの作品で出演したものか、また浮世絵には存在するが実際に出演はしていない役者・舞台においても話題の1つとしてカウントにいった。[10-15]

A. 浮世絵の選定の仕方

浮世絵を見ながら第一文献との比較、発行物の時期の選定、役割番付と絵本番付、もう一度文献との比較を行った。データベースの浮世絵の場合は出演作品や役者名・キャラクター名が明記されている場合があるが、誤りもあるため文献とのダブルチェックを行った。また、同じ時代に出たほぼ同じ絵であっても版木や塗り方が異なる場合がある。その場合はそれぞれを1としてカウントを行った。ただし、時期推定ができない浮世絵や役者の推定ができない浮世絵に関しては今回は話題数のカウントから省いた。役者の確認としては顔の推定に非常に有用であったのが三代目歌川豊国晩年の大作で元禄から幕末まで302人の役者の似顔を集成した『古今俳優似顔大全』と「浮世絵師列伝」と「浮世絵大事典」と「原色浮世絵大百科事典」と「浮世絵レ

ファレンス事典」である。さらに似顔判定のサンプルを増補。描かれた役者はだれか、役者絵における似顔判定に有用であった。役名の確認、登場作品の確認と出演作品、版元の確認までクロスチェックを行った。[13-15]

また以下の例を考慮した。

例1：模様が完全に異なる場合（色の違いは1枚1枚刷り師の彩色の際の押しの強さが異なるため考慮しないことと今回はした）

例2：同じ構図を使ったけれど背景が違う

（図のように朝の背景である場合と夕方の背景である場合とははっきりと明確な違いが有る場合がある）

例3：版元が違う、版元の認め印がない、幕府の認定印がない場合

（無許可で出版された浮世絵が一部存在しているため）

例4：文字情報が異なる

（デザイン自体は変わらないが記載されている文字が違う場合）

例5：肉筆の浮世絵に関しては1枚1枚が別物であるとカウントを行った。

番付の選択は以下の文献から行った。各公演ごとにいつ、誰が、どこの会場で歌舞伎の公演が行われるかが明記された番付や最良を対象に配布された番付、現代という映画のパンフレットのような役割を果たした演目の内容が明記されてある番付などがあつた。これらが無ければ浮世絵の内容や性質を考慮することができないため必要とした。[26-33]

看板の絵、お風呂や町中広告、月一のペースでの新作が生まれ、大当たりするとロングランでの公演を行っていた。江戸歌舞伎の人気役者を計算をするにあたって、広告宣伝費に相当する入力は舞台公演回数を元にした。当時出版された番付を元データとして計測を行った。[4]番付の選定は、日本国内にある大学、図書館、美術館、博物館のデータ

ベース全点より重複省き、選定調査を目視で行って数値を割り出した。そして今回の数理モデルの手法ではインターネットを通じたSNSにおける話題を選定したが、江戸時代にはもちろんそれらは存在しない。そこで、唯一のメディアである、紙媒体の発行物である浮世絵・川柳・瓦版・書籍を対象とした。調査対象としたデータベースは国内外にある美術館・博物館・個人所有・図書館・大学が所有する1849年から1855年の出版物とした。この時期を選んだ理由としては、人気役者の連続・死（中村歌右衛門、八代目市川団十郎、坂東うしか）によるアーティストの話題性が他のドキュメント数と比較し、大きかったこと、死絵の流行（役者の死に伴う浮世絵、話題の反映）があつたこと、鯉絵の流行（1850年～1855年にかけて地震が度重なった時期でもあり、地震にまつわる話題、浮世絵が多く出版された点）であつた。[6]また川柳・俳諧は幕末にかけて幕府批判（国の批判）に関する話題を弾圧されていたため、他の時期に比べ生活や文化にまつわるテーマの川柳が多くありケーススタディをするにベストであつた。また当時最も多く出版をしていた歌川豊国Ⅲ、歌川国芳の存在である。出版数が最も過密、公演ごとに出版する文献、俳諧などが非常にタイトなスケジュールで出版をしていた。そして、現代の日本で起きている度重なる大物歌舞伎役者の死と東日本大震災など比較しやすい時期であつたことである。[7-11]

4. 本手法の計算

本研究で数理的研究に用いた手法は、いわゆる回帰分析の類いとは全く違い、石井らによって提案された[1]ヒット現象の数理モデルと呼ばれる、統計物理学を基盤とした手法である。このヒット現象の数理モデルは今までに映画の観客動員数[1]を初め、地域イベント、音楽配信、AKB選抜総選挙、江戸歌舞伎の解析の初期段階の結果など多くのエンタテ

インメントの人気の数値的評価に应用されている [1,2,3,7].

このヒット現象の数理モデルでは,社会における人々1人1人が特定のエンタテインメントに抱く関心,興味,鑑賞意欲のようなものを数値化できるものとして定義し,その数値自体というよりその数値の時系列的な変動からそのエンタテインメントの人気・評判を定量的に評価するものである.図1を見てもらいたい.休日に公園かどこかでゆったりと新聞を読んでいた人が,近くで何かエンタテインメント的なイベントが始まったことに気付いて,ふとそちらを見たという状況のイラストである.この時,そのイベントに興味を持った心理とそれまでの新聞を読んでいた状況の心理の差を本研究で用いる,特定のエンタテインメントに対する興味・関心と定義する.このように定義すれば,人それぞれの性格や嗜好,背景やそれまでの人生の歩みのような要素は差を取る際に消去され,今着目しているエンタテインメント自体への興味・関心だけが取り出せると考える.

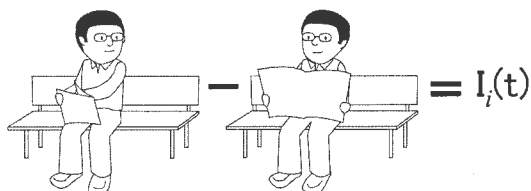


図3 特定のエンタテインメントに対する興味・関心 $I_i(t)$ の定義

次に,このような興味・意欲 $I_i(t)$ がどのように高まっていくかを,ヒット現象の数理モデルでは,3つのプロセスで考える.

- (1) 広告・宣伝による影響
- (2) 友人からの影響 (直接コミュニケーション)
- (3) 街の噂 (間接コミュニケーション)

詳細はこのヒット現象の数理モデルを提案した論文[1]を参照していただきたい.

まず,広告・宣伝によって直接的に人々の興味・

関心が呼び起こされる過程はイメージとして図2のようにテレビあるいはインターネット広告など外からの外力による影響として考え,次の項で数理モデル化する.

$I_i A(t)$

ここで, $A(t)$ は広告・宣伝の時系列的な変化であり,実際には広告出稿量[1]や,TV 露出の秒数[3, 42, 43],を用いるが,本研究のような江戸歌舞伎やコンサート等の場合は公演数自体[2,7,43-45]等を用いる.



図4 広告・宣伝による影響

次に直接コミュニケーションは,図3にイメージを示したように他の友人からの影響であり,その友人自身の興味・関心に比例して影響を受けると考える.さらに友人は1人ではないので友人全部の総和で影響を受けると考えると,次のように数理モデル化できる.

$$I_i \sum_j D_{ij} I_j(t)$$



図5 直接コミュニケーション

さらに,街の噂 (間接コミュニケーション)

は図4のような状況を想定して,特に知り合いでない2人が街角で話している話を第三者である i さんが小耳に挟んで興味を持つという考え方で数理モデル化する.従って,この部分は以下のような数式で表されるだろう.

$$\sum_j \sum_k P_{ijk} I_j(t) I_k(t)$$



図6 間接コミュニケーション

以上を取り纏め,さらに映画や公演などが公開後に日経つに連れて興味・関心が減衰していく項を付け加えると,以下のような数式が社会の中の1人の人 i さんについて立てられる.

$$\frac{dI_i(t)}{dt} = cA(t) + \sum_j D_{ij}I_j(t) + \sum_j \sum_k P_{ijk}I_j(t)I_k(t) - aI_i(t)$$

これが、本論文で用いるヒット現象の数理モデルの基本となる方程式であり、社会における1人の人の、特定のエンタテインメントに対する興味・関心の時間的な変動を記述する。これを社会全体の人々で平均すると（平均操作の具体的式展開は論文[1]を参照）、以下のような式になる。

$$\frac{dI(t)}{dt} = cA(t) + DI(t) + P I^2(t) - aI(t)$$

ここで $I(t)$ は社会の人々が特定のエンタテインメントについて抱く平均的な興味・関心を表し、 c は広告・宣伝の強さ、 D は直接コミュニケーションの強さ、 P は間接コミュニケーションの強さ、 a は興味・関心の度合いの減数の速さの係数となっている。

6. 計算結果

当時役者評判記といった、役者の善し悪しを専門家が評価した本が歌舞伎が行われていた上方（大坂・京都）、江戸において出版されていた。しかし、上演される作品の性質の違い、役者の評判の地域差があったため、上方歌舞伎と江戸歌舞伎のランキングは異なっていた。そこで、上方・江戸歌舞伎のそれぞれのランキングのどちらにも掲載された6人を選出を行い、その中のドキュメント数（話題数）の比較を行った。[7]

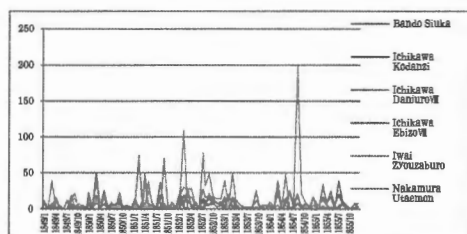


図7 選出した6人の役者のドキュメント数（話題数）

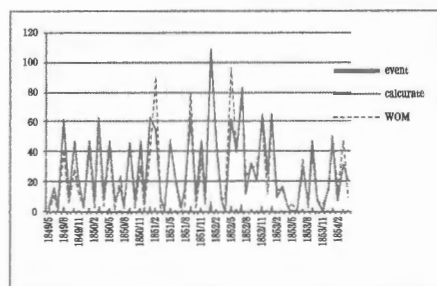


図8 8代目市川団十郎の計算結果

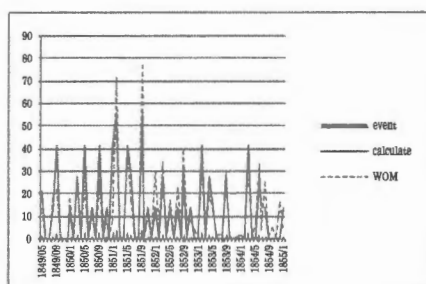


図9 坂東しゅうかの計算結果

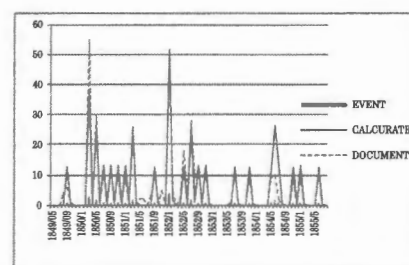


図10 7代目市川海老蔵の計算結果

7. 結論

私たちは現在のヒット現象も過去のヒット現象も全く同じ数理モデルでアーティストの評判を解析することができた。今回の手法では、まだまだサンプル数が完全とはいえないという問題があるが、本研究の試みを読んでくださった方や有志の方からの協力を願ひ、新たな歴史の見解の発見へ繋がり、世界中でも本手法の応用例が現れることで新たな発見や創造のヒントが現れることを願っている。

参考文献

- [1] Ishii, A. et al. The 'hit' phenomenon: a mathematical model of human dynamics

- interactions as a stochastic process. *New Journal of Physics* 14, 063018 (2012).
- [2] Y. Kawahata, E. Genda and A. Ishii, Revenue Prediction of music concerts Using the Mathematical Model of Hit Phenomena, the proceedings of the 2013 International Conference on Biometrics and Kansei Engineering (ICBAKE2013) 208-213
- [3] A. Ishii, S. Ota, H. Koguchi and K. Uchiyama, Quantitative analysis of social popularity of entertainments using mathematical model for hit phenomena for Japanese pop girl group AKB48, the proceedings of the 2013 International Conference on Biometrics and Kansei Engineering (ICBAKE2013) 143-147
- [4] Gerstle, C. A. Flowers of Edo: Eighteenth-Century Kabuki and Its Patrons. *Asian Theatre Journal* 4, 52-75 (1987).
- [5] Leiter, S. L. in *A Kabuki reader: history and performance* (ME Sharpe Inc, 3-359, 2002).
- [6] Nishiyama, M. & Groemer, G. in *Edo culture: daily life and diversions in urban Japan, 1600-1868* 23-76, 95-144, 181-228, (University of Hawaii Press, 1997).
- [7] "Edo Kabuki Ukiyo-e for Bigdata database", Available
(ア)at;
<http://data.linkedopendata.jp/storage/f/2013-08-30T04%3A47%3A01.086Z/edokabukirefernces.pdf>
- [8] "Library and Art Gallery and Museum and database of Ukiyo-e subject of investigation", Available
at;
<http://data.linkedopendata.jp/storage/f/2013-08-30T05%3A05%3A52.332Z/library-and-art-gallery-and-museum-and-database-of-ukiyo.pdf>
- [9] "Ukiyo-e Link List", Available
at;http://www.geocities.jp/web_ukiyoe/link2.html
- [10] Gerstle, C. A. Flowers of Edo: Eighteenth-Century Kabuki and Its Patrons. *Asian Theatre Journal* 4, 52-75 (1987).
- [11] Leiter, S. L. in *A Kabuki reader: history and performance* (ME Sharpe Inc, 3-359, (2002)
- [12] Nishiyama, M. & Groemer, G. in *Edo culture: daily life and diversions in urban Japan, 1600-1868* 23-76, 95-144, 181-228, (University of Hawaii Press, 1997).
- [13] Akama Ryo in *Vikutoria & arubato hakubutukan syozo ukiyoe zensakuhin kataaroginngu*, 2009).
- [14] Suzuki Tozo, Koike Syotaro. in *Fujiokaya nikki*. volumel-15 (Sanichisyobo, 1987-1995).
- [15] Takahashi Noriko. in *Study of literary and drama biological interpretation by the image database of Ukiyo-e : Elucidation of the historical development of the techniques of Mitate* (National Institutes for the Humanities National Institute of Japanese Literature, 2007).
- [16] Ryo, A., Fumiko, K., Ujlaki, P. & Noriko, H. The culture of play: kabuki and the production of texts1. *Bulletin of SOAS* 66, 358-379 (2003)

Web 公開中の伝統芸能興行データベースを使用した
演目別上演回数の比較手法について
A Comparison of the Number of Times of the Performance
Using the Traditional-Performing-Arts Database
Exhibited on the Web

坂部 裕美子

Yumiko Sakabe

立命館大学 文学研究科博士後期課程, 京都市北区等持院北町 56-1

Ritsumeikan University, 56-1 Tojiin-kitamachi, Kita-ku, Kyoto-city, Kyoto

あらまし: 現在 Web 公開されている最も充実した歌舞伎公演データベースを使用して「歌舞伎の演目別上演回数」を集計する際には、まず歌舞伎公演以外の公演の除外、演目の名寄せ等の処理が必要である。それらの処理を済ませて集計を行うと、「仮名手本忠臣蔵」が回数1位となる。だが、1幕1レコードという現行収録形式のままで度数集計を行った場合と、各幕に重みづけを行った上で集計した場合では、上位演目は若干異なってくる。Web 公開されている別のデータベースを使用して、文楽についても同様に演目別上演回数を集計し結果を比較すると、歌舞伎の場合以上に上位演目が変わる。歌舞伎も文楽も、重みづけをした上で行った集計結果の方が実感に近いという印象はあるが、どちらの方法も誤りとは言えず、「回数上位演目」の定義としてどちらを取るのが望ましいかという判断は今後の研究課題である。

Summary: When totaling "the number of times classified by program of Kabuki of performance" using the most substantial Kabuki performance database exhibited on the web, the exclusion of performances other than a Kabuki performance and the identification-of-multiple-accounts-under-the-same-name-as-a-single-entity processing of a program are required first. After the processing, it is found that "Kanadehon-Chushingura" ranks 1st in the total. But other higher rank programs are different by the ground total of the present form(one curtain is treated as one record), and weight attachment total(weight is attached to each curtain). When totaling the number of times classified by program of performance about Bunraku by the two same methods as Kabuki using the performance database exhibited on the web, and comparing the results, higher rank programs changes more than Kabuki. It seems that the result of weight attachment total is closer to feelings more than that of ground total in both case of Kabuki and Bunraku. Neither of the methods is errors, and the judgment which is desirable as a definition of the "higher rank program" is a future research task.

キーワード: 歌舞伎, 文楽, データベース, 度数集計, 重みづけ

Keywords: kabuki, bunraku, database, frequency total, weighting

1. 日本俳優協会版 DB を使用した歌舞伎興行集計

(1) データについて

筆者は主に、伝統芸能の興行データベースを集計した結果を使用して、興行の特徴や長期的推移を把握する、という研究を行っている。特に歌舞伎興行に

については、現在日本俳優協会が Web 上で無料公開している「歌舞伎公演データベース—戦後から現代まで¹」の制作に協力した経緯もあり、研究の中心的存在の一つとなっている。

¹ <http://www.kabuki.ne.jp/kouendb/>

このデータベースの網羅期間は、戦後すぐの1945年から「現時点で興行開催が決定している公演」まで（2013年11月時点で、2014年1月興行のデータが掲載されている）と幅広く、収録対象とする公演もかなり広範囲に及んでいる。データベースの「歌舞伎公演の範囲」には、以下のように説明が記載されている。

1. このデータベースでは、厳密な意味の“歌舞伎公演”のみならず、歌舞伎俳優が出演した商業演劇やミュージカル、現代演劇などの公演記録も掲載している。ただし「東宝歌舞伎」「東映歌舞伎」については、歌舞伎俳優が出演した公演のみを掲載した。なお、新派公演は今回は除外した。
2. 短期間の「研究公演」「試演会」「勉強会」「研修発表会」などは残念ながら調査が及ばなかった（「答会」など、一部入れたものもある）。

(2)「歌舞伎公演」以外の公演の除外

もともこのデータベースを作成した組織が「日本俳優協会」という主として歌舞伎俳優で構成された業界団体であり、例えばある歌舞伎俳優の生涯舞台活動を調べたい、というような場合には、この、内容の非常に充実したデータベースは相応に利用価値が高いと考えられる。

だが、このデータベースの収録内容を精査してみると、「歌舞伎興行の集計」を行う、という目的でここに収録された全データを使用することが適当といえるのか、疑問に感じるようになった。「歌舞伎公演」と見なすのには違和感が大きい公演データが相当数に上るのである。

「歌舞伎公演」の定義については大いに議論の余地があるのだが、ここでは取り急ぎ「違和感の大きい公演」を定義し、これらを除外することにした。

除外公演の判別ルールは以下のとおりとした。

1. まず、配役データの最初と2番目に記載されている俳優を別途リスト化し、いずれかが「歌舞伎俳優」（＝俳優協会所属で、新派でない人）でないものを抜き出す。配役不明の場合はここでの処理対象としない（積極的に除外しない）。

ここでは、前進座の俳優と、萬屋錦之介、大川橋蔵、市川雷蔵は「歌舞伎俳優でない人」とする（橋蔵・雷蔵は純然たる歌舞伎公演に歌舞伎俳優として出演した記録もあるが、配役の最初に挙げられるような俳優になるのは歌舞伎界を去った後のことであるため）。

2. 1.該当データのうち、以下に該当するものは「歌舞伎」扱いに戻す。

・総出演者、もしくは主な役の半分を超える人数が歌舞伎俳優である演目

・上記の判断が難しい「歌舞伎（舞踊）演目」で、複数の出演者のうち主たる踊り手である1人だけが歌舞伎俳優でないもの（橋蔵の「鏡獅子」など。ただし、非歌舞伎俳優が女優と組んで踊る道行ものなどは戻さない）

・歌舞伎の見方に関する解説で、出演者が歌舞伎俳優でないもの

3. 2.の処理後に残ったデータに「歌舞伎ではない」というフラグを付ける。

当初、公演データは12300件余りあった²が、このルールで処理を行った結果、(3)段階で930件余りのデータが除外対象になった。ちなみに、ここで除外対象となった公演を再チェックしたところ、スーパー歌舞伎は「リュウオー」だけが除外される（2番目の記載者が京劇俳優）、近年の公演では「赤い城 黒い砂」は歌舞伎扱いで残る（配役データの最初の2人だけが歌舞伎俳優）などが確認されるなど、除外ルールの機能不順もあった。だが、格上の役者の配役は「兵卒」「捕手」なども含めた配役一覧の最後に記載されることもあり、何番目までをチェックすれば条件に合う俳優を「合理的に」検出できるかは事前には推測できないため、当面はこの処理済みデータを使用して集計を行うことにした。

(3)上演演目の名寄せ

上演演目別の上演回数を集計するには、「演目名」変数を度数集計すればよいのだが、このデータはこのままでは使えない。それは以下の理由による。

1. 演目の表記違い

データ集計を行うと、新字・旧字の違いで表記の異なる演目は別データとして計上されるので、これらを合算しなければならない（俳優協会DBは「筋書の表記に準拠」が基本なので、元データ中には、例えば「假名手本忠臣蔵」というデータも存在する）。さらに、「加賀見山旧錦絵」（かがみやまこきょうのにしきえ）という演目は、「鏡山旧錦絵」と表して上演することもある。このような演目の名寄せ作業も必要である。

2. 別表記同演目の名寄せ

次に、歌舞伎独自の問題として、全く別の演目名が

² 先述のとおり現在Web公開されているデータベースには最新年次まで収録されているが、筆者が本報告で集計に使用したデータは2010年の興行までを対象としている。

付いていても、内容は同一の演目である、という場合がある(「弁天小僧(弁天娘女男白浪)」「十六夜清心」など)。さらに、いわゆる「助六」は、誰が主役の花川戸助六を演じるかで演目名が全く変わる、という約束がある。これらも合算しなければならない。

3. 「口上」の合算

演目名「口上」として記録されているもののデータ数は61回だが、実はこのほかに「襲名披露口上」「追善口上」「初舞台披露口上」という演目名でも数多の計上がある。今回のようなテーマの集計においては、これらをすべて別データとして計上するよりも、すべて合算するのが望ましい。

これらの問題には、各演目に集計用のユニークコードを別途設定し、「演目名」変数の内容をコード番号に変換してから集計する、という処理で対処した。

(4) 集計結果

ここまでの処理を経て集計された、歌舞伎興行における上演回数上位演目は表1のとおりである。なお、本論中では、「その舞台の実際の上演回数(=役者が役を演じた回数)」ではなく、「ある劇場のある月の公演演目として掲げられた頻度」を集計対象としていることに留意されたい。

表1 演目別度数集計・歌舞伎(単純集計)

演目名	データ数	構成比(%)
1 仮名手本忠臣蔵	674	5.10
2 義経千本桜	383	2.90
3 菅原伝授手習鑑	256	1.94
4 口上	200	1.51
5 勘進帳	187	1.42
6 京鹿子娘道成寺(「二人道成寺」含む)	162	1.23
7 春興鏡獅子	123	0.93
8 藤娘	114	0.86
9 天衣紛上野初花(「直侍」も含む)	113	0.86
10 青砥稿花紅彩圖	113	0.86
11 恋飛脚大和往来	112	0.85
12 一谷嫩軍記	111	0.84
13 妹背山婦女庭訓	108	0.82
14 元禄忠臣蔵	101	0.76
15 双蝶々曲輪日記	97	0.73
16 鬼一法眼三筋巻(「一條大盗腰」含む)	93	0.70
17 連獅子	90	0.68
18 伽羅先代萩	85	0.64
19 春夏秋冬	85	0.64
20 橋原平三嘗石切	83	0.63

2. 集計方法の改良

(1) データ構造に関する問題点

俳優協会版DBは、原則として1幕1レコードの形式になっており、例えばある月に忠臣蔵の通し上演が行われると、「忠臣蔵」の上演データが8~11レコード増える、という構造になっている。歌舞伎座には「一幕見」という制度もあるほど「幕」という構成単位は認知が高く、このデータ構造自体に問題はない。しかし、歌舞伎興行は「1ヶ月=1興行」と認識されるのが通例であるが、このデータ収録形式では、これまでに何興行開催されてきたのか、全く推測できない。

また、収録形式に関しては、別の問題も存在する。「ヤマトタケル」など近年の新作歌舞伎のデータは、「幕」単位の区分ではなく、公演全体が1レコードとして収録されている(現在Web公開中のデータで「昼夜同一公演」と記載のある公演にこの形式が多い)。この場合の1レコードと、「お祭り」などの舞踊の1レコードを同値に扱うのは、鑑賞者としての直感に合致しない集計、という印象がある。

(2) 簡易な「重み」づけ集計
この問題に正面から対応する方策としては、幕ごとの上演時間を入手し、その長さに合わせた重みをそれぞれの公演データに設定してから集計する、という手法が考えられる。だが、その作業ボリュームは膨大なものになる。

(2) 簡易な「重み」づけ集計

そこでより簡易な方法として、ある劇場の1ヶ月の上演分合計が「1」となるような、上演幕数に応じた重みを考える。(とりあえず、重みは全幕等分とする。)つまり、忠臣蔵が通し公演で8幕上演された場合、各幕を「0.125」分として扱うのである。また、先述の「公演全体が1レコード」となる公演は「1/1」で1扱いとなる。

「同一劇場・同一年月の公演データの数を数えて、1をその数で割った値を各幕データに付与する」というマクロプログラムを作成して「重み」を算出した後、「重み付き上演回数」を再集計した。

「同一劇場・同一年月の公演データの数を数えて、1をその数で割った値を各幕データに付与する」というマクロプログラムを作成して「重み」を算出した後、「重み付き上演回数」を再集計した。

(3) 集計結果

その結果を表2に示す。

表2 演目別度数集計・歌舞伎(重み付き集計)

演目名	データ数	構成比(%)
1 仮名手本忠臣蔵	94.82	3.94
2 義経千本桜	70.22	3.29
3 勘進帳	37.71	1.57
4 京鹿子娘道成寺(「二人道成寺」含む)	33.60	1.40
5 菅原伝授手習鑑	30.91	1.29
6 春興鏡獅子	29.02	1.23
7 口上	28.47	1.16
8 青砥稿花紅彩圖	27.80	1.16
9 恋飛脚大和往来	26.83	1.12
10 天衣紛上野初花(「直侍」も含む)	24.66	1.03
11 ヤマトタケル	23.11	0.96
12 藤娘	21.38	0.89
13 与話情浮名襷	20.15	0.84
14 連獅子	18.12	0.75
15 夏祭浪花鑑	17.09	0.71
16 身替座禪	16.72	0.70
17 元禄忠臣蔵	16.37	0.68
18 助六由縁江戸桜	16.28	0.68
19 傾城反魂香	15.88	0.66
20 お祭り	15.82	0.66

ただし、この時点での総興行回数(総興行月数)には、先述の「歌舞伎公演ではない公演」の開催分も含まれている。より正確に「歌舞伎の興行月数」を算出するために、「同劇場・同月の公演についてそれぞれ『総演目数』と『歌舞伎ではない演目数』の合計を算出し、『歌舞伎ではない演目』が全体の6割以上を占める公演は除外する」という処理を追加すると、本データ中には2360興行分のデータが収録されていたことが分かる。

板木データベースの可能性—データベース連携のことなど

Potentials of Printing Block Database through Linking with Relevant Databases

金子 貴昭

Takaaki Kaneko

立命館大学 衣笠総合研究機構, 京都市北区等持院北町 56-1

Ritsumeikan University, 56-1 Toji-in Kitamachi, Kita-ku, Kyoto

あらまし:本研究は、江戸時代の出版研究に関わり、従来顧みられることのなかった板木を研究資源として活用し、出版研究や板本書誌学を捉え直すことを目的としている。そのため、現在までに 11,000 枚におよぶ板木のデジタル撮影とデータベース構築を進めている。本稿では、板木のデジタル撮影手法、データベース構築、その利用について述べた上で、板木データベースと板木に関連するデータベースとの連携およびその可能性について探る。

Summary: This research aims to reconsider the study on commercial publishing in the Edo period (1603-1867) and Japanese bibliographical studies by utilizing printing blocks, as these studies have so far centred around research on printed publications. For that purpose, our project has constructed a digital archive of approximately 11,000 blocks. This paper presents a method of digitizing and archiving them in Printing Block Database. The paper also explores potentials of the database through linking with relevant databases.

キーワード: 板木, 江戸時代の出版, デジタル化, データベース

Keywords: Printing Block, Publishing in Edo Period, Digitization, Database

1. はじめに

江戸時代には商業出版が成立し、大量の板本(出版物)が刊行された。出版産業は拡大の一途を辿り、その結果、1 点の板本を繰り返し印刷する必要が生じたことから、活字ではなく、板木を用いた整版印刷が主流となっていった。現代に伝わる板本のほとんどは、江戸時代以降の整版印刷によるもの、つまり板木で摺られたものといっても過言ではない。文学・宗教・学問など、江戸時代のさまざまな文化的事象は、板本を扱わずして述べることはできない。したがって出版研究はもとより、江戸時代の各分野研究においては、多かれ少なかれ板本を研究資源として扱わなければならない状況にある。

そのような状況であるから、板本を研究資料として客観的に扱うための方法論「板本書誌学」は広く深く蓄積され、板本は諸研究の基礎資料として利用されてきた。しかし、板本に比べて板木は、置き去りにされてきた感が否めない。板木は印刷の道具であるから、板木なくして板本は存在し得ない。江戸時代の出版権(板

株)は板木と連動するものであり、板木なくして版權を云々することもできない。つまり板木は、江戸時代の出版の根幹にあったにもかかわらず、ほとんど顧みられることはなかったのである。

筆者は、この板木を研究の主眼に置き、板本中心に行われてきた板本書誌学や出版研究を捉え直そうとしている。それに関わり、本報告では、板木のデジタル撮影、データベース構築やそれらがもたらす可能性について論じる。

2. 板木のデジタル撮影

板木をめぐる上述の状況は、板木が持つ特性に由来するものである。第一に、軽量の板本に比べて、板木は重たく嵩高い。表面には墨や埃が堆積しており、手や服が汚れてしまう。付いた汚れは、簡単には落ちない。第二に、色板を除き、板木はほぼ黒一色であるため、複製資料の作成が困難であり、資料流通の障害となる。第三に、板木には文字や絵が左右反転の状態で彫られており、板本との比較照合が難しい。

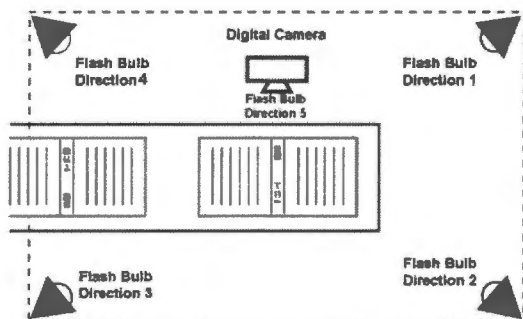


図1 板木デジタル撮影手法

筆者は、これらを克服するために、板木のデジタル撮影という手法をとった。またそれに際して、奈良大学博物館・奈良大学図書館が所蔵する板木約 5,800 枚をデジタル化・調査の対象とした（着手当初は約 5,000 枚）。このコレクションは、板木研究の第一人者・永井一彰氏（奈良大学文学部教授）が集積されたもので、江戸期の板元（出版元）が創業以来、運用・保管してきた一級資料である（注 1）。当該デジタル化プロジェクトは、立命館大学アート・リサーチセンター（以下、立命館 ARC）と奈良大学との共同研究としてスタートした。

撮影方法は、デジタル一眼レフカメラを用いた一般的な俯瞰撮影のセットであるが、ライティングには工夫を要した。フルカラー撮影とはいえ、被写体の板木は黒色である。板木の彫刻面を的確に記録でき、一様ではない板木の凹凸を捉えるライティングが必要であった。そのため、カメラ側からフラッシュ投光して得た標準画像に加え、複雑な彫刻面を記録するための斜光画像を別途撮影した。試行錯誤を重ねた結果、最終的には標準画像と 4 方向からの斜光撮影、合わせて、5 パターンのライティング手法を採用した（図 1）。

多くの場合、1 枚の板木には 4 丁または 2 丁分が彫られている。十分な画像解像度を保つために、1 カットに 1 丁ずつを収めることとして撮影を開始した。つまり、4 丁が含まれる板木を撮影する場合、4 カット×5 パターン＝20 コマとなり、分割撮影した場合には、板木表裏の全体像を記録したため、1 枚の板木を 22 コマで記録したことになる。板木の形式にはバリエーションがあり、単純計算はできないが、この方法で板木 5,800 枚のデジタル撮影を進め、約 90,000 コマのデジタル画像を集積した。

3. データベース構築と利用

デジタル化の次には、板木画像を自在に閲覧し、調査結果を蓄積していくためのプラットフォームが必要となった。これについては、板木 1 枚につき 1 レコード、1 丁につき 1 レコードとなる 2 つのデータベースを構築した上で両者を組み合わせ、web データベース「板木閲覧システム」（以下、板木データベース）（注 2）を構築し



図2 板木データベース 画像閲覧画面

た。板木データベースは 2010 年 2 月より、立命館 ARC のサーバで運用されている。

メタデータの詳細については割愛するが、板本書誌学に対応しつつ、板木独自の属性（板木の形式等）が記録できる構成としている。

板木データベース上では、あたかも板本を閲覧する場合のごとく板木を閲覧できるよう、ページめくり機能を備え、板木 1 枚ずつの閲覧、または、板本のタイトルごとに丁順にそくした複数板木の閲覧を選択可能な仕様とした（図 2）。また、板木と板本とを比較照合したり、彫られている情報を読み取ったりする上では、板木画像を鏡像で閲覧するほうが便利であり、彫りの状態を観察する場合には正像による閲覧が適切であり、それらの閲覧方法を切り替え可能な設計としている。

さらに、5 パターンライティングによる撮影を行ったため、それぞれのパターン画像を切り替えて板木を観察できる機能を持たせ、必要に応じて板木の全体像を呼び出す機能も付加した。メタデータの編集権限があれば、オンラインで随時更新可能であり、リアルタイムに最新の調査状況が反映されるようにした。

板木のデジタル化・データベース構築には労力を要したが、結果、先にあげた難点を克服し、板木を研究資料として細部まで観察するための強力なツールが出来た。筆者はこのツールを駆使し、関連資料とともに考察を進めている。それらの結果については、本稿では詳述を避けるが、板木の基本構造を明らかにし、板木という視点から見た板本の新観察手法を見出し、「印刷の道具」という理解を陵駕する板木の機能を明らかにしつつ、出版研究を実践している。

第 2 章に述べた板木の性質により、板木は原物による整理が難しいが、板木データベースは板木原物の管理・調査ツールとしても有効であり、2013 年 6 月に新聞各紙の紙面を賑わした上田秋成著作「冠辞統紹」の板木（注 3）、同じく秋成著作で「煎茶の世界」展（九州国立博物館）（注 4）に出展されている『清風瑣言』の板木などは、共同研究の枠組みにおいて、板木データベースを用いた調査が進展する過程で見出された板木である。

板木デジタル画像およびデータベースは、公開を前提としている。板木は板本とは異なり、原則1組しか存在しないため、どの板本のどの部分の板木が現存するのかを情報共有することは重要である。

なお、詳述を避けた箇所を含め、第2～3章の内容は拙著『近世出版の板木研究』(注5)に詳述している。合わせて参照されたい。

4. 関連データベースとの連携

前章に述べたように、板木データベースは日々調査結果が反映され、成長している。しかし、板木データベースと、すでにデジタル化されている他のコンテンツや構築済のデータベースが連携することにより、板木データベースや、データベースの利用環境は自ずと発展を遂げることになるだろう。

(1)板本

板木によって印刷されたものが板本であるから、当然ながら板木と板本は、コンテンツとして密接な関係を持つ。現存する板木は入木などの修訂技術による改訂版、または板木が再彫製された再版以降のものである場合が少なくない。板木と板本の諸本を比較照合し、版面の良し悪し、改訂の有無、初版と覆刻の関係など、版のバージョンを見極める場合、板木データベースとデジタル化された板本があれば、それらの照合は容易となる。

その実現のため、現在、板木データベースは、立命館 ARC の「ARC 書籍閲覧システム」(以下、古典籍データベース)(注6)とリンクしている。このデータベースでは、筆者が収集し立命館 ARC に登録した板本、立命館 ARC が画像の寄託公開を行っている奈良大学図書館および永井一彰氏所蔵板本など、板木に対応する板本を閲覧することができる。板木データベースで板木を閲覧し、その板木で摺られた板本を古典籍データベースでシームレスに検索・閲覧する機能は実装済みであり(図3、図4)、画像単位の比較対照は実現できていないが、現状で板木データベースと古典籍データベースを相互に参照する環境は整っている。

(2)絵画・浮世絵

板木と、板木で摺られた絵画や浮世絵(冊子体裁をとらない一枚物形態の資料を含む)は、(1)の板本と同様、強い関係性を持つ。

筆者が現在までにデジタル化を行った絵画・浮世絵の板木は、近代以降の作品もしくは複製が大部分を占めており、連携によって近世出版研究に関わるメリットは即座には得にくい。板本と同様、連携は比較的容易である。今後、立命館 ARC で稼働している「ARC 浮世絵閲覧システム」(注7)が連携の対象となる。



図3 関連データベースへのリンクボタン(非公開モード)



図4 板木データベースと古典籍データベース

(3)出版記録

京・大坂・江戸にはそれぞれ本屋(板元)仲間が組織されて出版を取り仕切っており、日常の業務日誌や先例の記録として必要な情報が書き留められた。

また、それとは別に、個々の板元が残した記録も現存している。板元の本分は、むしろ本を刊行することだったに違いないが、それを遂行するためにも、その活動のためには、板木の管理業務が必須であった。その結果として、出版記録上に板木の状態が記されたり、板木の管理に特化した蔵板記録や、板元間で行われた板木(著作権)売買の記録が生まれることになる。

これらの記録の中には、現存の板木に関わる記事が多く含まれている。それらを板木と突き合わせることで、板木がどのように運用されてきたのか、どのような経緯で伝存しているのか、といった板木の履歴を克明に追うことが可能となる。それはまた、板木が江戸期の出版において、どのような役割を果たしていたのかを明らかにすることにもつながるだろう。

上述の観点から、筆者は出版記録の画像・テキストデータ集積を鋭意進めており、板木データベースとの連携に向けて「出版記録データベース」の構築に取り組んでいるところである。

(4)板元

板木データベースに所収される板木の最終的な旧蔵板元は数軒でしかないが、現存する板木の伝存経緯に関わった板元、それらの相板元(共同出版者)を含めれば、現存板木の理解のために必要な板元は相

当な軒数にのぼる。また板木に限らず、板本・出版記録に目を移せば、出版を営みとした板元に関わる情報が必須となることは言うまでもない。

従来、板元に関わる研究ツールとして『近世書林板元総覧』(注8)があるが、板元に関わる総合的データベースの存在は聞かない。また同書は、研究の入り口で参照すべき資料であり、むしろ板木・板本などの諸資料を扱う上で、同書のブラッシュアップを行っていくのが利用者の責務であろう。

現状では、立命館 ARC 内で企画されている板元データベース(注9)との連携を模索している。

(5)研究文献

本稿冒頭に述べたように、板本研究は広く深い蓄積があり、先行研究は豊富に存在している。先行研究参照の必要性は本研究に限ったことではなく、データベース上で閲覧した板木や板本に関連する研究文献の情報を引き出すことができれば、有益であることは言うまでもない。

現状、板木データベースは、立命館 ARC が運用する「ARC 日本文化研究文献検索システム」(注10)と連携しており、閲覧した板木あるいは対応する板本に関わる文献情報を引き出せるようになっている。ARC の文献データベースは、研究メンバーが持ちよった文献情報を集積したものであり、網羅性は低いものの、関連性の高い文献情報ばかりが集積されている点において、ノイズが極めて少なく、連携しやすい。しかし、研究文献のオープンアクセス化が進む今日、CiNii Articles(注11)の API を利用した文献情報表示から文献そのものを閲覧させる方向性も検討している。

5. おわりに

板木データベース単体については、奈良大学所蔵板木を対象に構築が始まったが、2011 年度以降、調査対象を株式会社法蔵館、美術書出版株式会社芸艸堂、立命館 ARC、京都大学附属図書館、仏光寺などに拡大し、各機関の所蔵板木が板木データベースに収録され、収録板木枚数は 11,000 枚に増加している(注12)。いずれの機関も、遺産・資料としての管理運用に戸惑われている現状があり、板木データベースはこれらの所蔵機関をつなぎ情報交換を行うネットワークにもなり得る。事実、板木データベースの公開活動により、調査依頼・板木の寄贈・所蔵情報をいただくケースがでてきており、資料をめぐる人的ネットワークの形成にも寄与している。板木データベースは、板本研究ツールから、板本研究のハブへと成長しつつある。

データベース間の連携についてはまだ課題も多く、本格的な運用にはほど遠い。しかし、データベース構

築以前のデータを含めて、個々のデータベースが成長するにつれ、関連するデータベースは共鳴しあう。例えば板木データベースが成長すれば、古典籍データベースに資するところあり、その逆もまた然り、関連するデータベースが、リンクによってつながる連携の枠組みに加わることで、自ずと研究の基盤が広がっていく。筆者が構築に携わってきた板木データベースについていえば、前章(1)～(5)のデータベースと連携することにより、近世出版総合データベースへと昇華する可能性が見えてきているのである。

付記 本研究は JSPS 科研費 25770097 による成果を含む。また本稿は、第 2 回知識・芸術・文化情報学研究会(2013 年 2 月 9 日、於:立命館大学大阪キャンパス)における口頭発表の内容を一部に含む。

注

(注 1) コレクション構成には変遷があるが、奈良大学博物館編『板木さまざま 芭蕉・蕪村・秋成・一茶も勢ぞろい』(2013 年 6 月、奈良大学博物館)所収、永井一彰「はじめに」に示される内訳が最新である。

(注 2) 立命館 ARC「板木閲覧システム」
(<http://www.arc.ritsumei.ac.jp/db9/hangi/>)
※2013 年 11 月 8 日閲覧。

(注 3) 2013 年 6 月 19 日付各紙報道。

(注 4) 九州国立博物館トピック展示「茶の湯を楽しむ IV 特別編 煎茶の世界」(2013 年 10 月 1 日～12 月 1 日、於:九州国立博物館)

(注 5) 金子貴昭『近世出版の板木研究』(2013 年 2 月、法蔵館)

(注 6) 立命館 ARC「ARC 書籍閲覧システム」
(<http://www.arc.ritsumei.ac.jp/db1/books/search.php>)
※2013 年 11 月 8 日閲覧。

(注 7) 立命館 ARC「ARC 浮世絵閲覧システム」
(<http://www.arc.ritsumei.ac.jp/db/arcnishikie/>)
※2013 年 11 月 8 日閲覧。

(注 8) 井上隆明『改訂増補 近世書林板元総覧』(日本書誌学大系 76、1998 年 2 月、青裳堂書店)

(注 9) 松葉涼子氏(日本学術振興会特別研究員、南山大学)作成のデータなどが基となる見込みである。

(注 10) 立命館 ARC「ARC 日本文化研究文献目録」
(<http://www.arc.ritsumei.ac.jp/db10/bunken/search.htm>) ※2013 年 11 月 8 日閲覧。

(注 11) 国立情報学研究所「CiNii Articles」
(<http://ci.nii.ac.jp/>) ※2013 年 11 月 8 日閲覧。

(注 12) いずれも進捗中のプロジェクトであり、現在は画像・メタデータとも非公開である。

画像データベースから生まれる新たな創作活動について

Image Database as a Source of Creative Activities

前崎 信也

Shinya Maezaki

立命館大学 立命館グローバル・イノベーション研究機構, 京都市北区等持院北町 56-1

Ritsumeikan University, Ritsumeikan Global Innovation Research Organization

56-1 Tojiin kitamachi, Kita-ku, Kyoto

あらまし: 立命館大学アート・リサーチセンターでは、在外コレクション所蔵の日本の工芸品のデジタルアーカイブ・プロジェクトを2009年より行っている。本発表の前半ではプロジェクトの調査先機関、及び調査した作品点数やその内容の概要を紹介する。後半ではイメージデータベースの公開以降、新たに広がった人的ネットワークや研究・創作活動事例について述べる。それにより、デジタルアーカイブという行為が、いかに社会や芸術家に影響を与え、新たな文化の発信力につながっていくのかという可能性について検討する。

Summary: The Art Research Center, Ritsumeikan University, has been developing a digital archive of Japanese decorative arts since 2009. To begin with, this paper will introduce the latest report of the project including the list of affiliate institutions and the number of digitized works. The main purpose of the project is to create image databases for scholarly researches, but the project has also widened the international human network and resulted to form a number of cultural activities. The second half of the paper will introduce some of the examples and discuss the role of digital archive projects as a source of new creative activities.

キーワード: 画像データベース、陶磁器、竹工芸、漆器、コラボレーション

Keywords: Image database, Ceramic, Bamboo art, Lacquer, collaboration

1. はじめに

インターネット上で公開される画像データベースは年々増加し、学術研究に利用できるレベルの情報をそなえたものも多くみられるようになっていく。立命館大学アート・リサーチセンターでは、海外の美術館・博物館に所蔵される日本の陶磁器及び竹工芸のデジタルアーカイブ・プロジェクトを進めており、その成果として一部データを公開している。本プロジェクトの直接的な成果としては、これまで海外の所蔵機関の収蔵庫に

秘蔵されてきた多くの工芸品の実態が徐々に明らかになり、その学術利用が始まっていることである。しかし、データベース構築のために資料を調査し、その成果を公開するということは、学術の場以外にも影響を与える可能性を持つものである。

本発表では、まずこれまでの日本工芸データベース構築で得ることのできた成果を紹介し、その後、データベースの構築や公開によって学術の場以外に広がった文化的活動について紹介する。そうすることにより、



(画像1) 日本陶磁器データベース

URL: <http://www.arc.ritsumei.ac.jp/jcdb/>



(画像2) 日本竹工芸データベース

URL: <http://www.arc.ritsumei.ac.jp/bbdb/>

データベースが、人々や社会に与える文化的影響について検討する。

2. 陶磁器データベース・竹工芸データベース

立命館大学アート・リサーチセンターのホームページでは、日本陶磁器データベース及び日本竹工芸データベースの公開を行っている(画像1、2)。これらのデータベースは、2009 年より同センターで構築を進めている日本工芸データベースの内、一部の個人所蔵の作品を収録している。日本工芸データベース(非公開)は、欧米の博物館・美術館に所蔵されている日本の工芸品の画像データベース化プロジェクトの成果であり、現在までに工芸品 2,849 件、34,613 画像の調査撮影を終えた。成果の詳細は以下である。

イギリス

オックスフォード大学アッシュモリアン美術館

陶磁器: 280 件 1,940 画像

スコットランド国立博物館

陶磁器: 460 件 3,472 画像

セインズベリー日本藝術研究所

陶磁器・木工芸他: 76 件 533 画像

ドイツ

ベルリン国立アジア美術館

陶磁器: 134 件 1,209 画像

ハンブルグ工芸博物館

陶磁器・竹工芸: 901 件 8573 画像

リンデン民俗学博物館

陶磁器: 213 件、1,969 画像

グラッシー民俗学博物館

陶磁器: 253 件、1,546 画像

ドレスデン美術館磁器コレクション

陶磁器: 372 件、3,604 画像

ザクセン州立工芸博物館

陶磁器: 57 件、489 画像

アメリカ

ハワイ大学アウトリーチ・カレッジ

陶磁器: 19 件、142 画像

ミシガン大学美術館

陶磁器: 28 件、286 画像

個人コレクション

イギリス、ドイツ、アメリカ、日本等

陶磁器・竹工芸: 1,309 件、10,850 画像

3. 画像データベースの美術展覧会での活用

立命館大学アート・リサーチセンターでは構築した画像データベースを使い展覧会の補足的な情報として活用する試みを行っている。2009年12月から2010年4月にかけて英国バース市の Museum of East Asian Art (東亜美術館)で、企画展 Seifu Yohei and His Contemporaries: Meiji Ceramics in the Scholarly Taste が開催された。(画像3)

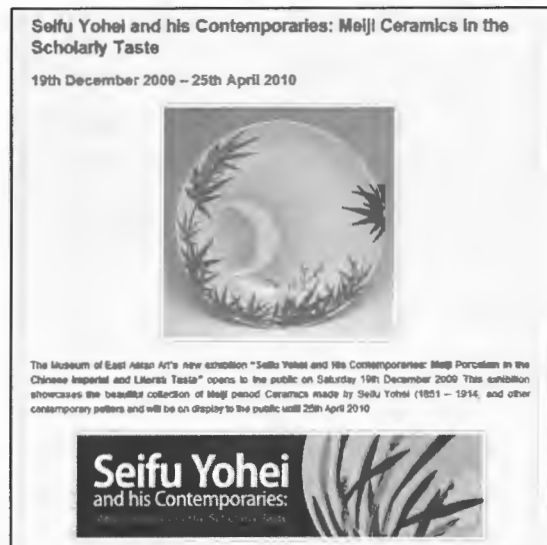


(画像3) 展示会場風景

本展覧会に出品された作品はすべて英国のディヴィッド・キング・コレクション所蔵のものであった。当時、同コレクションについてのデジタル化を進めていたことから、美術館に日本陶磁器データベースを利用した展覧会ウェブサイトの構築を提案し、承諾を得て提供することとなった。(画像4)。展覧会会場では見ることのできない作品の裏側や底部、作者の銘の画像が閲覧できる同サイトは、会の来場者からの高い評価を得た。同サイトは閉会後も過去の展覧会情報として、美術館のホームページからの閲覧が可能である。(画像5)



(画像4) 展覧会ウェブサイト 作品詳細画面



(画像5) Museum of East Asian Art 過去の展覧会紹介ページ

URL: <http://www.meaa.org.uk/past-events>

同様の試みとして、2014年夏には英国セント・アイヴスの Leach Pottery (バーナード・リーチ工房)において企画展 Matsubayashi Tsurunosuke in St Ives 1923-24 (仮題)を計画し、その場で展覧会用のウェブサイトの提供を予定している。本展覧会は大正時代にバーナード・リーチ工房で活動した京焼の陶工である松林蘊之助の遺した記録や作品を展示するという企



(画像6) 松林蘆之助関連資料データベース

URL:<http://www.dh-jac.net/db9/matsubayashi/index.html>

画である。発表者は2012年に松林蘆之助の遺した資料を「松林蘆之助関連資料データベース」として検索・閲覧可能なデータベース化を実施した。本展覧会では、このデータベースを利用して、松林蘆之助関連資料からバーナード・リーチ関連の全文献を表示することのできる展覧会用ウェブサイトを開発予定である(画像6)。

4. データベースから生まれる人的ネットワーク

発表者はこれまで日本工芸データベースの構築の為に欧米の博物館・美術館を中心に所蔵品の撮影を続けてきた。その活動の結果として海外に広範な人的ネットワークを築くことができたのは言うまでもない。しかし、データベースをインターネット上で公開したことや、展覧会の場で活用したことは、新たな人的ネットワークの創造にもつながっている。

日本陶磁器データベースで公開をしているのは、主に先述のディヴィッド・ハイアット・キング・コレクション所蔵作品が中心である。データベースの公開をきっかけにキング・コレクションの世界的な認知度が高まり、キング氏の元には世界中の日本陶磁器コレクターや研究者からコンタクトがあるという。データベースを構築・管理している立命館大学アート・リサーチセンターにも、



(画像7) カリフォルニア大学デーヴィス校での講義風景

北米の日本陶磁器コレクターからのコンタクトがあり、世界的な情報交換の場が広がり続けている。2013年10月には、サクラメント在住の日本陶磁器コレクターのオーガナイズにより、カリフォルニア大学デーヴィス校から招待され、近代日本陶磁器のアメリカ向け輸出とサンフランシスコ港の果たした役割についての講義を行った(画像7)。

5. データベースから生まれる創作活動:

事例1 姫路市大塩町 清風研究会

姫路市大塩町は明治期を代表する京焼の陶工として名高い三代清風與平(1851-1914)が幼少期を過ごした場所である。彼は京都の清風家に入った後、その才能を開花させ、陶工として初の帝室技芸員に推挙された。しかしながら、その活躍の場が主に京都であったため、大塩町がかつて日本の陶芸を代表した人物を輩出した町であるということについて、久しく忘れ去られていた。

発表者は陶磁器データベースの調査の一環で大塩町を訪れた際、町の公民館に三代清風與平作の『水彩磁鳳凰文大花瓶』が保管されていることを発見した(画像8)。町の公民館の押し入れにお宝が眠っていたという話は話題となり、発見の数か月後に同公民館で清風與平とその花瓶に関する講演を行った。この経緯は『神戸新聞』で紹介され、発見された大花瓶は、



(画像8) 三代清風與平作『水彩磁鳳凰文大花瓶』

URL: <http://www.arc.ritsumeai.ac.jp/jcdb/>

愛知県陶磁資料館(現:愛知県陶磁美術館)で2010年10月から11月に開催された企画展「帝室技芸員の技と美 明治の人間国宝—清風與平・宮川香山から板谷波山まで—」に出品されるに至った。

その後、大塩町では三代清風與平について学ぼうと町内の有志が集まり「清風研究会」を発足。以来、定期的に研究会を開催している。研究会では清風與平に関する文献の輪読会や、陶磁器関連の展覧会への訪問、三代清風與平が活動した京都市の五条坂へのバス旅行(画像9)など、地域の年長者を中心に大きな活動となった。発表者は清風與平に関する講演会を毎年1回、これまでに計3回行った。講演会では毎回地域に眠っていた新たな清風與平作品が持ち込まれ、大塩町には三代清風與平作品の一大コレクションがあると呼べるほどとなった。本研究会のこれからの目標は地域にある清風與平作品を集めた三代清風與平の展覧会を開催することである。



(画像9) 清風研究会 京都バス旅行、2011年4月

6. データベースから生まれる創作活動:

事例2 竹×漆 プライスコレクションの若冲

竹工芸データベースの構築ではハンプルグ工芸博物館所蔵の約170点、国内の個人コレクション約100点のデジタル化を行った。この他には、2011年、大阪の竹工芸を代表する一家、田邊竹雲斎家所蔵の文献資料等のデジタル化を行った。そして、一連の竹工芸関連のデジタルアーカイブ活動を通じて、田邊小竹氏と漆芸作家の若宮隆志氏と知り合い、竹と漆のコラボ



(画像10) 若宮隆志、田邊小竹作『提若』、2013年

レーション作品への協力を行うこととなった。

本企画は竹工芸と漆器作家のコラボレーション企画となり、そのモチーフに米国の日本美術コレクターであるジョー・プライス氏所蔵の伊藤若冲の絵画の使用許可を得て、「竹×漆 プライスコレクションの若冲」と題した展覧会の開催が 2013 年に実現した。作品は竹と漆を融合させた作品が中心。竹の上に高蒔絵をほどこす技法の再現や、網代網で若冲の絵画を表現した作品など、伝統的技法を用いながら新しい工芸の創出することを目標にした。

大阪高島屋・日本橋高島屋で展覧会を開催した。日本橋高島屋の会場では発表者をコーディネーターとしてギャラリートークを開催したが、同様のトークショーでは異例の数の来場者が訪れた(画像11、12)。この成功により、「竹×漆プロジェクト」は今後 2 年に1度の高島屋での企画となる事が決まり、毎回異なるテーマで 10 年間継続する予定である。



(図11、12) 日本橋高島屋でのギャラリートーク風景、
2013 年 6 月

7. おわりに

近年の人文科学研究においては、研究成果のデータベース化と公開が成果報告の前提のひとつとしてみなされるようになってきている。今後はそうして無制限に増え続ける学術データベース類を、いかにして有効に活用するかが問題となってくるだろう。しかし、データベースの活用をその構築者が前提として提供するのは正しい事であろうか。これらのデータベースが様々な目的を持って作られることとなったことと同じように、その利用方法も利用者によって千差万別であるはずである。むしろ本来の目的とは異なる方法で利用されるようになったデータベースがあるとするならば、それはその構築者の想定を越えたこととなり、文化的には意味のある広がりをもった重要な事例であると捉えることができるはずである。

本発表の後半で事例を挙げて説明したように、デジタルアーカイブを行う行為とは、人と人、人と物、物と物のつながりに、これまではなかった変化を与える可能性を有している。この関係性に対して与えられた刺激は、新しい創作活動へ発展する可能性を秘めたものである。デジタルをキーワードとしながらも、極めてオーガニックに予測不可能な方向へと向かう発展。もしもデジタルアーカイブやイメージデータベースに、学術的ではなく一般社会に対しての価値あるとするならば、新しい文化活動を生む源としての役割をその内のひとつとして捉えるべきなのではないだろうか。

藤原京遺構データベースの構築

On the establishment of a Historical GIS database for the *Fujiwara-Kyo* site, capital of ancient Japan

宮崎 良美

Yoshimi Miyazaki

奈良女子大学古代学学術研究センター, 奈良市北魚屋東町

Nara Womens' s University, Kita uoya nishi-mati, Nara, Nara

あらまし: 奈良女子大学古代学学術研究センターでは、藤原京城の遺構について Historical GIS データベースを構築している。このデータベースは、調査機関の保管する遺構実測原図ではなく、調査機関から刊行された発掘調査報告書を活用している。現在、約 1700 の調査区と 5600 件の遺構のデータが入力済である。本報告では、報告書とその掲載の遺構図から作成されたこれらのデータについて藤原京研究への GIS 活用の点から位置情報の精度等について検討した結果と、課題について報告する。

Summary: In Center for Research of Ancient Culture, Nara Women's University, I and colleagues are establishing a Historical GIS database for the *Fujiwara-Kyo* and other sites around there. This database has the contents of the excavation reports published by several research institutions, and has not original measured drawings. It has 1,700 excavated sites and 5,600 evidences' data at present. I present on the problems of establishing the database. In addition, the geographical information of site data in the database is accurate enough to make studies for the reconstruction of *Fujiwara-Kyo* using the GIS database.

キーワード: 藤原京, Historical GIS, データベース

Keywords: the Fujiwara Capital, Historical GIS, database

1. はじめに

「藤原京遺構データベース」は、奈良女子大学古代学学術研究センターで構築中の「奈良盆地歴史地理データベース」のひとつである¹⁾。

藤原京は、約 1300 年前に現在の奈良県橿原市東部・桜井市西部・明日香村北部にあった都である。持統天皇八年(694)十二月の遷都から和銅三年(710)三月の平城京遷都までの短い期間の都であったが、藤原京の時代は大宝律令の制定などを中心として古代日本の国家が整えられていく重要な時期にあたる。古代京都発達史においても、条坊、つまり中国の都に倣ったとされる基盤目状の道路と街区をもった計画的都城であり、さらにそれまで代替わりのたびに場所を変え作り替えられていた天皇の住まいであり執務などを行った宮を、初めて恒久的に一所に存続すること

を意図して造営されたのも藤原宮であった²⁾。そのため、藤原宮と藤原京についてその内部構造や、造営の時期、条坊施工の時期や地域的展開、人々の生活の実態などを明らかにすることは、古代日本形成過程を知る上で重要な鍵となる。

藤原宮・京について戦前より 80 年近く続いてきた発掘調査成果や近代以降の研究史の蓄積は非常に厚い。これを研究に存分に活用するために Historical GIS データベースの構築は不可欠である。そこで、奈良女子大学古代学学術研究センターでは藤原京遺構データベースの構築を継続して進めている。この経過と課題について中間的なものではあるが報告する。

2. 藤原京遺構データベースの概要

1) 藤原京の範囲

はじめに、遺構データベースへの入力対象範囲に関連して、藤原京城における発掘調査の歴史と復原案の変遷³⁾についてみておきたい。

藤原宮の本格的な発掘調査は、1934～1943 年日本古文化研究所の発掘調査により始まる。戦後、1966～1968 年の藤原宮跡調査実行委員会による発掘調査において大極殿や大垣など藤原宮の中心部の建物が確認された。この成果をもとに、岸俊男は、古代官道である横大路、中ツ道、下ツ道を京極とする南北 6 里・東西 4 里(1里はほぼ 530m)の地域に、南北 12 条東西 8 坊の復原案を示した⁴⁾。その後、この復原案にもとづく条坊道路の想定位置で道路側溝が次々発見されていった。

しかし、1979 年岸説藤原京の外側で京内条坊道路の延長線上に道路側溝が発見される例が相次ぎ、奈良県立橿原考古学研究所の秋山日出雄によって京城を拡大した大藤原京説⁵⁾が提唱され、各種の復原案も現われた。そして、1996 年、橿原市土橋遺跡と桜井市上之庄遺跡で相次いで京極道路が発見され、東・西京極が確定した。

これらの成果を受けて、中村太一氏⁶⁾と小澤毅氏⁷⁾により 1 条・1 坊を 1 里(530m)とする南北 10 条、東西 10 坊の正方形で、宮がその中央にあるとする藤原京復原案が提示された。2004 年に橿原市新口町で北京極道路が想定位置から発見され、中村氏・小澤氏の復原案が、最も有力な説となっている(図 1)。

2) 構築の目的と背景

藤原京および飛鳥地域については、奈良文化財研究所、奈良県立橿原考古学研究所をはじめ複数の調査・研究機関によって発掘調査が行われている。1969 年に奈良文化財研究所が継続的な藤原宮跡の発掘調査を開始してからでも半世紀近くが経ち、各調査・研究機関から刊行された報告書は膨大な数にのぼる。しかし、これらの調査成果を利用する場合、1 冊ずつ精査することになり、非常な手間がかかる。そこで、これらの検出遺構や遺物などの情報がデジタル化、データベース化されることで研究の効率化がはかれることが期待される。発掘調査のデジタルアーカイブや、遺跡・以降の地理情報の管理などの効率化も進むであろう。GIS のスケールフリーで紙地図の判型の制約が内シームレスであるという特長をいかすことで、遺跡や遺構を詳細に検討することも、藤原京城あるいは奈良盆地全体の中で俯瞰するように分析す

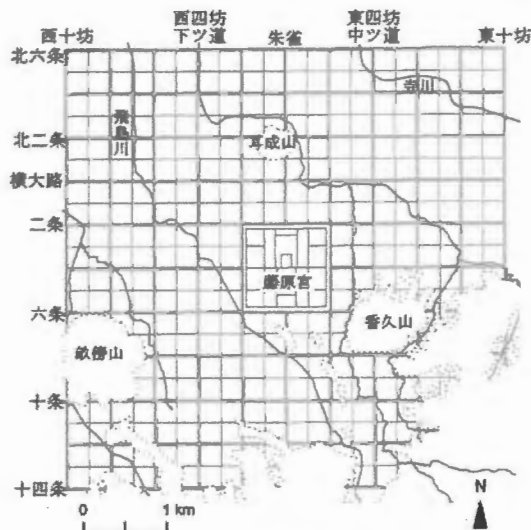


図 1 藤原京城

『奈良文化財研究所紀要 2013』p119 図をもとに作成

ることも自在となる。

宮都地域については、平城京・平城宮⁸⁾や、平安京⁹⁾についてのものがある。藤原京・飛鳥地域に関して一般利用が可能なデジタルの遺跡情報には、奈良文化財研究所の遺跡データベースや、奈良県立橿原考古学研究所の「奈良県遺跡情報地図」などがある¹⁰⁾。このなかで、本センターでは GIS を活用した藤原京遺構データベースの構築に着手し、一般利用も可能なかたちでの公開を目指している。

本センターは発掘調査を担う調査・研究機関ではないため、一般に利用可能な刊行物をデータベースの基礎資料とすることにし、複数の調査機関による刊行物を対象とする。また、地理情報としての遺構データの基礎資料となる遺構図は、精度等の点で調査機関が保管する遺構実測原図を直接スキャンしてデジタル化することが望ましいが、大判の図面を大量にスキャンすることは時間や経費、調査図面の保存の面で容易ではない。そのため、可能な限り刊行物所収の遺構図を利用することにした。このような形での遺構データベースは発掘調査機関に属さない研究者でも制作可能なものであることから、その工程においてどのような問題があるのか検討し報告することも、本データベース構築の目的とした。

また、前節でみたとおり藤原京城の復原については変遷があるが、最も広域をカバーしていることもあり中村氏・小澤氏説藤原京城と、飛鳥地域を含めた地域をデータベースの対象地域とする。大藤原京説以前

のものでも、この地域で行われた発掘調査成果はさかのぼって入力することを方針としている。

3) 遺構データ入力作業

遺構データの入力作業は次の通りである。

- ①発掘調査報告書中の遺構図等をスキャンしてデジタル画像とする。
- ②図中に記載された位置情報(平面直角座標系による国土座標)をもとに、GISのジオリファレンス機能により幾何補正を行い、GISソフトウェアに取り込む。
- ③②の画像データ上で、調査区や遺構の形状をトレースする。
- ④トレースした調査区や遺構のデータの属性情報を入力する。GISソフトウェア上で遺跡名称や遺構番号などの概要、典拠となる報告書名や遺構図データのファイル名等を入力するが、GISソフトウェアは長文のテキストデータを扱うのには適していないため、検出遺物や特記事項などの詳細な情報は、入力や集計等に便利な表計算ソフト(Microsoft Excel)を使い、入力・保存している。

なお、藤原京や飛鳥地域という広範囲をカバーするため、トレースの対象は建物遺構、塀や柵の遺構、溝遺構に絞り、建物、塀や柵は遺構図中に復原された柱筋を、溝は上端線のみとしている。

以上により、藤原京遺構データベースは、GISソフトウェアに取り込まれた遺構図の画像データ、調査区・トレンチの区画形状をトレースした調査区データ、遺構をトレースした遺構データと、これに関連づけられる表形式の詳細情報データより成っている。

現在のデータ数は、調査区データは約1700箇所、遺構データは建物、塀や柵、溝を合わせて約5600件である。入力済データについてみると、最も古いものは奈良国立文化財研究所による『飛鳥・藤原宮発掘調査概報4』(1974年)所収の調査区であり、最新は2010年前後刊行の報告書所収の調査区である。だが、橿原市、桜井市、明日香村の教育委員会等による発掘調査成果は、報告書のバックナンバーの入手の都合や国土座標の確認等の事情もあって1990年代以降が中心であり、1980年代以前は奈良国立文化財研究所と奈良県立橿原考古学研究所・奈良県教育委員会による報告書が中心となっている。

3. 遺構図の取り込みに関わる問題

ここでは、報告書掲載の遺構図をGISソフトウェアに

取り込む際に確認された問題点などについて整理しておきたい。

長期間にわたる発掘調査報告書をみていくと、同一の調査区や遺構が、発掘調査が行われた際の遺構図のほかにも、後に複数の調査区にまたがる遺構の配置を示すために集成図にまとめられたり、隣接する調査区の遺構図に参考として示されたりすることがしばしばある。これらの集成図、遺構図等をGISソフトウェアへ取り込むと、同一の調査区であるにもかかわらず、互いに位置が食い違ったり、わずかながら回転したりしている。遺構図全体がずれているので、速報や概報の刊行後に調査が継続されるなかで柱穴や溝遺構の形状が変わったためにずれて見えるものとは考えにくい。また、掲載図の注記にある縮尺に対して、実際の図の縮尺が2%程度小さい場合もある。

このずれや縮尺の縮み等は、報告書の誌面上で概ね1~2mm程度に収まっていて、縮尺は1/100~1/700であるから、実際の距離に換算すれば0.1~1m程度のずれがあることになる。このようなずれは同一の調査区・遺構が複数の報告書の掲載図に示されるほぼ全てに見られることから、これが報告書掲載図に一般的に含まれる「誤差」とみなしてよいだろう。

逆に言えば、これを超える大きさのずれやゆがみは、国土座標の誤りなど何らかの問題があるとみて、ずれの原因を調べ修正するのがよいだろう。このような例としては、複数の調査区にまたがる溝遺構が溝幅以上にずれて接続できないなどがあり、上層・下層の遺構図が重ならないなどによって気づく場合もある。

なお、藤原京遺構データベースの場合、複数年次の報告書に同一の調査区・遺構が掲載される場合、国土座標の記載がないものをのぞき、なるべく初出の遺構図を使用し、ずれが顕著な場合は調査機関に確認していただくようにしている。

しかし、ここ最近の間に発掘調査が行われた調査区の遺構図どうしについては、若干のずれ・ゆがみはみられなくなってきたように思われる。まだ遺構図の件数が少ないので言い切ることはできないものの、申し添えておきたい。

4. 藤原京遺構データベースによる国土座標の計測値の検討

1) 調査機関による測量値との対照

本データベースは、調査機関が保管する遺構実測原図に比べる縮尺が小さいこと、また、先にみたように、

遺構図データについて若干の「誤差」があること、さらに、国土座標などに地表面に換算して数 m 程度の何らかのエラーがあっても、隣接する調査区がない場合など気づく手がかりがない可能性もあることなどから、このまま GIS を活用した条坊の復原的研究などの空間分析に利用することについて不安がある。

そこで、発掘調査を行う調査・研究機関が公表した遺構の座標値と、同じ遺構について、藤原京遺構データベース上で国土座標を計測して、比較しについて検証しておきたい。

藤原京城については、奈良国立文化財研究所(当時)が、奈良県立橿原考古学研究所、橿原市教育委員会、桜井市教育委員会と協力し、『藤原京条坊資料(1998)』を刊行している¹¹⁾。その後、一部のデータ訂正を行った「【資料編】藤原京主要条坊遺構の概要」¹²⁾も出されている。これは、調査機関が検出された藤原京の条坊道路の道路心や道路側溝心の国土座標を集成したものである。藤原京遺構データベースの計測値をこれらの座標と比較し、大きな齟齬がなければ、今後データベースを利用した計測した値もそれなりに信頼してよいと考えられよう。

また、小澤毅氏が『藤原京条坊資料(1998)』の条坊の座標に一部追加修正を加えながら、桜井市の吉備池廃寺周辺の条坊復原について、回帰分析に基づく検討を行っている¹³⁾。入倉徳裕氏も、藤原京の中心部を通る条坊道路について方位のふれや規格等について、『藤原京条坊資料(1998)』等をもとにした詳細なデータ解析を行っている¹⁴⁾。両氏の論文中的資料にあった条坊の国土座標から「【資料編】藤原京主要条坊遺構の概要」の座標に追加・修正を行い、比較用のデータとした。

そして、藤原京遺構データベース上で、「【資料編】藤原京主要条坊遺構の概要」等に挙げられた条坊の検出地点に対応する箇所のうち、まず条坊大路¹⁵⁾について、報告書の入手や遺構図の取り込みができていないものを除き、条坊を検出した 62 箇所の調査区について、道路の中心線、側溝の中心線や中心線の交点等を描画し、両者の距離について比較した。この際に、先述のデータベースの「誤差」をふまえ、両者の距離が 1m に収まるかどうかを目安にしたところ、57 箇所(約 92%)の調査区で、1m 未満であった。

1m 以上のずれがあった 5 箇所について、ずれの原因はまだ明らかにできておらず、今後の課題である。ただ、うち 1 箇所は、後年の報告書において調査区

内の大垣の遺構について測量成果の誤りが指摘されていた。これが当該の遺構にのみあてはまるのか、調査区全体に及び、この道路遺構についても位置が変わってくるものなのかどうかについては確認できていない。この例では、測量成果の指摘は発掘調査報告から約 15 年に指摘があり、周辺の調査に関連する報告書を長期にわたって精査することも必要であることを実感させる。余談となるがアルバイトなど複数人で入力する際にこのような情報の抽出や、過去の調査成果にさかのぼって修正の要不要の検討などがあることも入力作業に応じて留意すべきであろう。

ひとまず、藤原京遺構データベースの調査区・遺構の国土座標について、何らかの問題が 1 割程度見込まれるものの、9 割程度の遺構については若干の「誤差」に注意すればよいらしいことが確認できた。

2) 条坊道路関数化への応用

さらに実際の条坊復原的で用いられる回帰式の算出を通して、将来的に藤原京の遺構群に関する GIS を用いた研究に「【資料編】藤原京主要条坊遺構の概要」の情報を更新するような資料集成の刊行をまっことなく、本データベースを利用することも見込まれることもあり、検討を加えておきたい。

先に挙げた小澤氏、入倉氏¹⁶⁾は、条坊方位のふれや条坊道路の規格や割り付け等に関連して、回帰分析を行って条坊道路を関数化し、つまり国土座標で表される一次関数に表し、これをもとに詳細な分析を加えているが、この論文に条坊の回帰式も示されている。両氏が関数として示した条坊道路について、前節と同様に、本データベース上で改めて検出された道路側溝心の座標を計測して回帰式を求め、結果を両氏によるものと比較しながら、利用の可能性や問題点について整理、検討することにしたい。

なお、対象となった条坊遺構検出地のなかには世界測地系による国土座標が示された遺構図も含まれるが、比較の都合上、ArcGIS10.1 を用いて日本測地系に変換し、計測を行った。前節で国土座標の数値に若干不安があると確認された箇所については、あらかじめ「【資料編】藤原京主要条坊遺構の概要」等による座標に置き換え、それら以外の事由に起因する問題について確認できるようにした。また、両氏は道路心についても回帰式を示しているが、ここでは、データベース上で描画・計測できる側溝心の計測に絞り、道路心の座標や溝の両肩が検出されていない場

表1 藤原京遺構データベースの計測値による条坊道路の傾角

	条坊道路	計測位置	計測 点数	関数式	差違	傾角 傾度
東西道路	一条大路	北側溝心	6	$Y=\tan(0^{\circ}14'7.31'')X-165727.43$	0.61	0.933
	一条大路	南側溝心	6	$Y=\tan(0^{\circ}12'53.61'')X-165721.63$	1.45	0.934
	二条大路	北側溝心	4	$Y=\tan(0^{\circ}17'22.33'')X-165944.03$	0.49	0.933
	二条大路	南側溝心	5	$Y=\tan(0^{\circ}11'45.45'')X-165953.5$	1.14	0.933
	三条大路	北側溝心	5	$Y=\tan(0^{\circ}32'57.67'')X-166130.55$	1.03	0.933
	三条大路	南側溝心	3	$Y=\tan(0^{\circ}33'1.55'')X-166133.4$	0.00	1.000
	四条大路	北側溝心	12	$Y=\tan(0^{\circ}55'47.55'')X-166372.11$	2.43	0.933
	四条大路	南側溝心	6	$Y=\tan(0^{\circ}37'41.45'')X-166353.22$	1.15	0.933
	五条大路	北側溝心	5	$Y=\tan(1^{\circ}5'33.53'')X-166452.36$	1.16	0.955
	六条大路	北側溝心	5	$Y=\tan(1^{\circ}22'12.55'')X-166664.75$	0.75	0.933
南北道路	六条大路	南側溝心	6	$Y=\tan(1^{\circ}43'43.12'')X-166556.55$	4.50	0.933
	西三坊大路	東側溝心	5	$Y=\tan(0^{\circ}25'33.52'')X-15444.24$	1.71	0.937
	西三坊大路	西側溝心	6	$Y=\tan(0^{\circ}14'21.53'')X-15312.37$	2.39	0.935
	西二坊大路	東側溝心	3	$Y=\tan(0^{\circ}14'43.4'')X-15358.77$	0.00	1.000
	西一坊大路	東側溝心	3	$Y=\tan(0^{\circ}20'53.55'')X-15357.06$	0.05	0.975
	西一坊大路	西側溝心	4	$Y=\tan(0^{\circ}15'50.2'')X-15451.51$	0.22	0.937
	朱雀大路	東側溝心	3	$Y=\tan(0^{\circ}40'55.5'')X-15333.55$	1.52	0.975
	朱雀大路	西側溝心	6	$Y=\tan(0^{\circ}17'13.45'')X-15273.33$	1.51	0.931
	東一坊大路	東側溝心	4	$Y=\tan(0^{\circ}55'45.2'')X-20005.11$	0.54	0.937
	東一坊大路	西側溝心	6	$Y=\tan(0^{\circ}55'57.03'')X-19303.14$	1.03	0.936
	東二坊大路	東側溝心	3	$Y=\tan(0^{\circ}42'23.16'')X-15330.93$	21.44	0.922
	東二坊大路	西側溝心	2	$Y=\tan(1^{\circ}17'12.55'')X-20634.67$	-	-

合の溝心の推定などは行っていない。

回帰分析には、Microsoft Excel の分析ツールである「回帰分析」を使用した。得られた回帰式は表1に示した。なお、平面直角座標系においては、Xは東西方向の、Yは南北方向の座標を示す。また、藤原京城周辺では、国土座標の方眼方位の南北軸を基準として、角度が1°ふれると、1km先では南北軸から約17.5m、5km先では約87.3mずれることになる。以上をふまえて、回帰式の結果についてみていくが、比較のために引用する数値は、小澤氏と入倉氏の注13)と14)に挙げた論文中的表および付表による。

①東西道路

一条大路：データベース上で国土座標を計測することができたのは北側溝について両端を西二坊から東四坊、南側溝は西三坊から東四坊にかけてとする区間である。計測点数は表1に示している。入倉氏の付表には東七坊の国土座標が挙げられているが、その出典となる遺構図は得られておらず、分析対象となった区間が氏のものより短い。しかし、入倉氏による回帰式は $X=\tan(0^{\circ}13'35'')Y-165714.44$ であり、表1の北側溝心の傾き $E14'7.31''N$ や南側溝心の $E12'53.61''N$ はこの傾き

から大きく外れることはなかった。

二条大路：データベース上では、「【資料編】藤原京主要条坊遺構の概要」以後の調査成果も加えたことで、南北側溝とも西十坊から西二坊までと、両氏よりも長い区間で計測点を得ることができた。入倉氏による道路心の回帰式は $X=\tan(0^{\circ}17'08'')Y-165953.46$ 、小澤氏による南側溝心の回帰式は $X=\tan(0^{\circ}17'08'')Y-165961.5$ であり、北側溝はほぼ同じ値となったが、南側溝心は約6'程方眼方位に対する傾き小さく表われた。

三条大路：西二坊、東二坊と、北側溝のみ東十坊で側溝心が計測できた。ほぼ小澤氏の検討対象区間と同じであるが、本データベースでは東四坊付近の遺構データを欠く。しかし、小澤氏による北側溝心の傾き $X=\tan(0^{\circ}32'39'')Y-166131.8$ 、南側溝心の $X=\tan(0^{\circ}35'56'')Y-166124.1$ と比べると、北側溝はほぼ同じ、南側溝は約3'傾きが小さいという結果になった。

四条大路：東一坊から西十坊にかけて計測点が取得できた。入倉氏論文付表にある東四坊の地点はデータを欠くが、2001年の四条シナノ遺跡の調査成果により西十坊の座標を加えている。入倉氏による

回帰式は、 $X=\tan(0^{\circ}34'04'')-166394.21$ で、データベースの計測結果によるものは、南北側溝とも2~3' 傾きが大きい。

五条大路：西二坊から西三坊の約1坊分の区間で、南側溝では近接する2点しか得られなかったので、北側溝のみ5箇所を計測した。入倉氏の検討したのとはほぼ同じ区間である。1度を超える大きな傾きとなったが、氏の回帰式 $X=\tan(1^{\circ}20'00'')-166418.39$ に比べると、約13' 傾きが小さく表われている。一般に両端の距離が短い場合に、回帰線の傾きが大きく振れることがあるが、短い区間で異なる計測点によるためとみておきたい。

六条大路：西四坊～東二坊の区間で計測できた¹⁷⁾。入倉氏論文付表の東三坊については遺構図データを欠き、検討区間は若干短くなった。入倉氏の回帰式は $X=\tan(1^{\circ}24'22'')-166661.63$ で、南側溝では約16' の傾きの差がみられた。六条大路については道路幅員に関して問題があることが指摘されており¹⁸⁾、このことが南側溝の傾きの差となっているものとみられるが改めて検討したい。北側溝は氏の回帰式に近い数値が得られた。

②南北道路

西三坊大路：東西側溝とも一条から九条の比較的に距離で計測ができた。東側溝は、入倉氏の $Y=-\tan(0^{\circ}29'06'')-19619.37$ とは約3.5' の傾きの差がある。

西二坊大路：東側溝のみ四から九条で計測された。結果は氏の $Y=-\tan(0^{\circ}19'48'')-18912.05$ と約5' の差となった。西二坊大路については入倉氏も座標が少ないことを指摘しており、この差については検討を待ちたい。

西一坊大路：西側溝は一条から七条まで、東側溝は一条でのみ計測できた。耳成山の北では遺構データが取得できず、全て耳成山の南である。西側溝は入倉氏の $Y=-\tan(0^{\circ}16'38'')-18496.06$ に近く、東側溝は計測区間が短いので傾きの差が大きく出たものと考えられる。

朱雀大路：西側溝は3~8条で、東側溝は3~5条の区間で計測できた。北面中門・南面中門などは除いて計測している。東西側溝のふれが異なるが、入倉氏の指摘するように朱雀門南側で道路心から西側溝、東側溝までの幅が異なることによるものであろう。氏による回帰式は $Y=-\tan(0^{\circ}37'39'')-19248.72$ である。

東一坊大路：二、四条で計測ができた。距離は短いですが、小澤氏の検討区間にほぼ重なり、回帰式も氏の $Y=-\tan(0^{\circ}54'20'')-19793.0$ と近い結果となった。

東二坊大路：東側溝は四、八条、西側溝は四、七条で計測したが、東側溝は残差が大きく、相関係数も低い。東二坊大路は、前節で検出されずれが1mをこえる調査区の1つを通る。座標については置き換えをしているが、そのほかの遺構の位置関係など報告書等も確認し検証が必要であろう。

以上のように、藤原京遺構データベース図上で計測された道路側溝の座標についての回帰式と小澤氏・入倉氏による回帰式と比較したところ、両氏とほぼ同じ区間で座標が得られた条坊道路では、計測点数にかかわらず、かなり近い結果が得られたものと考えられる。また、東二坊大路のように、座標に何らかの問題がある可能性がある箇所について、相関係数や残差等の手がかりが得られるらしいこともわかった。

このほかいくつか検討すべき点もあるが、六条大路や五条大路のように条坊の直線性や道路幅員に関わる研究上の課題との関連で検討が必要なものもあり、改めて検証を行うことにしたい。

5. 藤原京遺構データベースを用いた検討例 —主題図の作成

ここでは、簡略ながら藤原京の条坊道路検出地について、遺構データベースを使ってマクロスケールで遺構を概観することが容易になる例を紹介した。データベースの使用例を紹介したい。

現在、入力された調査区・トレンチデータは、藤原京内で約1000箇所である。このうち、条坊道路推定線にかかる調査区・トレンチは小規模なものも含めて約630箇所を超えるが、藤原宮造営以前の先行条坊とそれ以前に施工されていた先々行条坊を含めて、約240箇所で条坊遺構が検出されている。

これをGISソフトウェアによって分布図に示すと図2のようになる。これをみると、条坊遺構は藤原宮跡とその付近での検出例が圧倒的に多い。発掘調査が早くから継続して行われてきたことが、検出例の数に反映されていることがわかる。藤原京城では、下ツ道つまり西四坊大路から飛鳥川、十条大路付近にかけて比較的空間的にまとまって条坊が検出されている。この西四坊大路に沿っては北四条から十二条大路まで道路遺構が検出されている。特に東四坊大路である

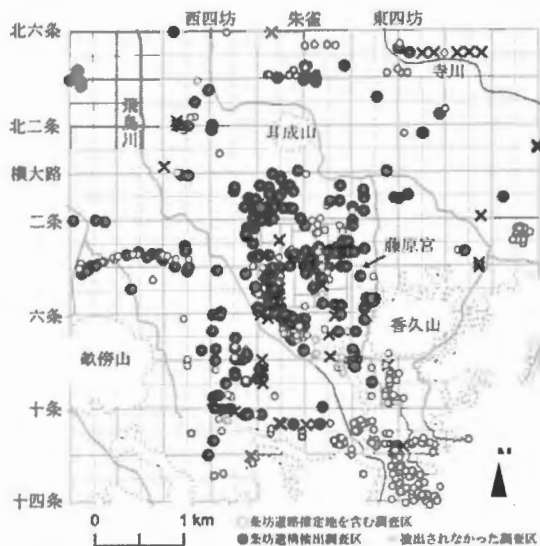


図2 藤原京条坊遺構の検出地

中ツ道の南延長線にあたる東南部で、条坊遺構がほとんど検出されていないことと対照的である。また、畝傍山の北側では道路工事に伴って線状に遺構が検出されている。一方、京の北部や東部では発掘調査数自体が少ないこともあり、検出例は多くない。

一方で、条坊遺構が検出されなかった、あるいは条坊遺構に関する言及がない調査区をみると、条坊遺構検出地のすぐそばなどにもあり、検出地と消失地に空間的な隔たりがあまりない様子がうかがわれる。一般的に発掘調査で検出される条坊道路は削平された状態で検出されるとのことであり、削平の度合いが異なるだけで、後世に一樣に削平を受けた可能性がうかがわれる。これについては、遺構の埋没深度や削平の状態などを検討することで明らかにできよう。

香久山の南方から南の明日香村にかけての地域は、藤原宮跡と同じくらの発掘調査が実施されているが、条坊遺構についての言及は多くない。これは香久山の南麓に大官大寺跡があり、その周辺も紀寺をはじめとする寺院跡があり、さらに南は飛鳥時代に時計台がおかれた水落遺跡、京城の南端付近は宮殿遺跡である石神遺跡にあたるためと考えられている。また、条坊遺構が失われた箇所をみると、条坊遺構検出地よりも河川に近い箇所にある場合も比較的多いようであるとわかる。

6. おわりに

最後にそのほかの課題について述べておきたい。

・国土座標の記載について

藤原京・飛鳥地域において、発掘調査報告書掲載の遺構図等に平面直角座標系による国土座標が記載されるようになるのは、1970年代からである。その後、段階的に、他の調査機関の報告書に国土座標が記載されるようになってきた。しかし、国土座標記載以前の発掘調査には、藤原宮中枢部や、飛鳥地域の中心部などの重要な遺構を対象とするものが多く、遺構データベースの有用性を高める上で必要なデータであり、遺構データベース構築上の重要な課題であり、有効な解決策を見つけれないでいる。

ただ、近年、古い調査区の隣接地で発掘調査が行われる際に、過去の調査区も一部含めて調査し、その位置情報の記録が進められている¹⁹⁾。これによる遺構図を含む調査報告を手にする機会が増えており、速効的ではないが確実な解決策として、これらの蓄積が待たれる。

・発掘調査報告書等で用いられる用語の分類・整理

長期かつ複数の調査機関による発掘調査報告書をみていくと、調査の経緯、遺構や遺物の検出状況などにも応じて、調査区ごとに報告で述べられる内容や重点がおかれる事柄、表現等に違いが多くある。入力される属性情報は報告書の記載に忠実であることが基本であるが、実際にこのようなデータ群についてテキスト検索を行う場合に、思い当たる用語で検索を繰り返す必要があり、非常に不便である。例えば、7世紀半ばの遺構について検索するならば、「7世紀中期」「7世紀中葉」「飛鳥時代」「藤原宮期以前」「藤原京期」などの用語で探すことになり、不便である。

本データベースの場合、藤原京地域では藤原宮や条坊に伴うかどうかを時期区分の柱であることから、「藤原宮・京期前」「藤原宮・京期の直前」「藤原宮・京期」「藤原宮・京期の後」「藤原宮・京期のかかなり後(中世・近世)」の5段階にゆるやかに区分し、これを各データに加えているが、同様に、遺構の時期以外についても遺構について使用される用語を集計し、類似の用語を包括する分類をつくり、これを検索の一助とすることが望まれる。

また、研究の進展に応じて条坊道路の認定や遺構の用途をめぐる解釈なども変わる。報告書だけでなく、最新の研究成果からも、随時、情報が追加されていくのが望ましい。

・遺構検出面の標高

GISデータベース活用の上で、遺構検出面の標高や、地表面からの埋没深度、土層、遺構の削平の状

況など、過去の地形・環境の復原のための資料もあるのが望ましい。この遺構検出面の削平の状況については、廃都後の藤原京の変容・耕地化の展開を明らかにする手がかりとなる可能性もある。しかし、一般的に、報告書には調査区壁面の土層断面図は示されるが、収載の遺構図に面的に検出面の標高や、溝・柱穴などの削平の状況を知るための深さ等が記入されることはあまりない。このような情報を詳細に知ろうとすれば、遺構実測原図にあたることになる。報告書を活用するデータベース構築としては、このような情報の収集も課題のひとつである。

以上、藤原京遺構データベースの現状と課題について報告した。本報告の際には具体的な数値を示しながら、掘り下げて報告する予定であるが、特に、入力作業で問題に行き当たるが多かった遺構図の精度について検討を行い、留意すべき点はあるものの、藤原京域のようなマクロスケールで予察的分析を行うには問題のないことを確認できたと考える。

このことは、藤原京域の調査・研究機関の発掘調査と作成される報告書の内容や遺構図の精度の高さを改めて確認したことにもなる。これらの成果を存分に活用することで、さらに藤原京に関する研究が進展していくであろう。

付記

本稿の作成にあたっては、奈良女子大学 21 世紀 COE プログラム「古代日本形成の特質解明の研究教育拠点」および平成 23～25 年度科学研究費補助金（基盤研究（B）「古代都城・都市をめぐる環境論」（研究代表者：館野和己、課題番号：23320137）の成果を使用した。

注および引用文献

- 1) 「奈良盆地歴史地理データベース」については次の文献で紹介されている。拙著(2010):「奈良盆地歴史地理 GIS データベースの構築と課題」古代学1、pp55-68。出田和久(2012):「奈良盆地歴史地理データベースの構築とその利用」(HGIS 研究協議会編『歴史 GIS の地平』、勉誠出版、pp197-207。
- 2) 木下正史(2003):『藤原京ーよみがえる日本最初の都城』中公新書、中央公論社。木下正史、佐藤信編(2010):『飛鳥から藤原京へ』、吉川弘文館など。

- 3) 深澤芳樹(2012):「藤原京の成立」遺構解釈の一例」、史林、95-1、pp38-73。
- 4) 岸俊男(1969):「京城の想定と藤原京条坊制」(奈良県教育委員会『奈良県史跡名勝天然記念物調査報告第二五冊 藤原宮ー国道一六五号線バイパスに伴う宮域調査ー』)pp119-126。
- 5) 秋山日出雄(1980):「藤原京の京城考ー内城と外京の想定ー」檀原考古学研究所紀要 考古学論攷第4冊、pp1-8。
- 6) 中村太一(1996):「藤原京と『周礼』王城プラン」、日本歴史 582、pp91-100。
- 7) 小澤毅(1997):「古代都市「藤原京」の成立」、考古学研究 44-3、pp52-71。
- 8) 森本晋(2007):「平城宮・平城京の調査と GIS、条里制・古代都市研究 22、pp19-32。
- 9) 河角龍典ほか(2007):「宮都研究と GIS」、条里制・古代都市研究 22、pp1-18。
- 10) このほか奈良産業大学の「藤原京 CG 再現プロジェクト」などもある。(URL:<http://www.nara-su.ac.jp/archives/fujiwaracg/index.html>。)
- 11) 奈良国立文化財研究所飛鳥藤原宮跡発掘調査部(1999)『藤原京条坊資料(1998)』。
- 12) 奈良国立文化財研究所(2000)『奈良国立文化財研究所学報 第60冊 研究論集XI』。
- 13) 小澤毅(2003)「藤原京条坊と寺地」(奈良文化財研究所『吉備池廃寺発掘調査報告ー百済大寺跡の調査ー』)pp.220-230。
- 14) 入倉徳裕(2008)「藤原京条坊の精度」(奈良県立檀原考古学研究所編『檀原考古学研究所論集第十五』、八木書店)pp.421-450。
- 15) 道路幅員等の検討から、条坊の大路は岸説藤原京以来の偶数大路が本来の大路で、奇数大路は小路であるとする説が有力であるが、ここでは従来説大路の計測を行っている。今後、その他の小路等に対象を広げて検討する予定である。
- 16) 前掲注 13)、14)
- 17) 六条大路の計測にあたっては入倉徳裕・小澤毅(2008)「藤原京六条大路の再検討」(奈良文化財研究所紀要 2008、pp.16-17)に基づき、高所寺池遺跡で検出された六条大路の北側溝を SD9359、南側溝を SD2915 とした。
- 18) 前掲注 17)。黒崎直(1997)「藤原京六条大路の幅員について」奈良文化財研究所年報 1996、pp26-27。
- 19) 奈良文化財研究所の松村恵司氏のご教示による。

奈良盆地前方後円墳データベースの構築と課題 —WebGIS データベースのひとつの試み—

On The Construction of The Database of keyhole-shaped mound, *Zenpoukouenahun*, and its Problems :An Attempt as A WebGIS-Database

出田 和久*・石崎 研二*・宮崎 良美**

Kazuhisa Ideta*, Kenji Ishizaki*, Yoshimi Miyazaki**

*奈良女子大学文学部, 奈良市北魚屋西町

Nara Women's University, Kitauoya-Nishi-Machi, Nara

** 奈良女子大学古代学術研究センター, 奈良市北魚屋東町

Nara Women's University, Kitauoya-Higashi-Machi, Nara

あらまし: WebGIS データベースの現況を簡単に紹介し、従来の遺跡・文化財データベースが WebGIS データベースとしての特色を必ずしも十分には発揮していないことを指摘した。ついで、奈良盆地の前方後円墳データベースの WebGIS 化について、GIS の特色である空間分析が容易である点を活かすために、検索結果を地図の表示することを重視し、また利用者が簡単に分布図を作成できるように操作性を高める工夫を行ったことを紹介した。利用者がその関心に従って様々な分布図が容易に作成できることにより、前方後円墳への理解を深めると同時に高次の関心へと誘う効果が期待できる。今後は、内容のさらなる充実と対象範囲を拡大して全国規模の検討が行えるようにすることが課題である。

Summary: The purpose of this paper is to present the devices which we made for raising usability so that a user can create a distribution map easily on the WebGIS-database of keyhole-shaped mound, *Zenpoukouenahun*, and problems which will be caused when a user uses it and a user going to use it more highly. In addition, I pointed out that many database of remains and cultural property did not necessarily show the characteristic as the WebGIS database fully. When users can create easily the distribution map about various *Zenpoukouenahun*s according to their concerns, we think that the effect invited to high order concern is expectable at the same time they deepen an understanding to the *Zenpoukouenahun*.

キーワード: WebGIS, データベース, 分布図, 歴史地理学, 前方後円墳

Keywords: WebGIS, database, distribution map, historical geography, the *Zenpoukouenahun*

1. はじめに

従来、歴史情報のデータベースの多くは単に文字ベースのテキストデータから構成されていた。しかし、歴史情報とひとくちに言っても文字史料だけではなく、絵巻物や古絵図、さらには考古資料など多様な資料が用いられるようになってきた¹⁾。このような動きを反映して歴史情報のデータベースにも図像をデータとして有するものが登場してきた。東京大学史料編纂所の

「史料編纂所蔵肖像画模本データベース」や「古写真データベース」、「史料編纂所蔵荘園絵図模本データベース」や「東洋文庫所蔵」図像史料マルチメディアデータベースなどはその代表例と言える。

このように歴史情報は多様化しているものの、歴史学ではその情報の起源地や所在地についての関心は希薄なようで、これらのデータベースにおいても必ずしもそれら図像やテキストに関連した地理情報が結合されているわけではなかった²⁾。

これに対して近年、特に 1995 年の阪神・淡路大震災以降 GIS の有効性が注目されたこともあり、GIS を活用した情報提供が多くみられるようになってきた。そのような中で最近では WebGIS を活用したデータベースが多くの自治体で遺跡の分布や情報の検索機能を実装して公開されるようになってきている³⁾。

以上のような動きを受けて、奈良女子大学古代学術研究センターでは 21 世紀 COE プログラム「古代日本形成の特質解明の研究教育拠点」(2004 年度～2008 年度)において構築した GIS を活用した「奈良盆地歴史地理データベース」の改良と充実に取り組んでいる。なかでも「前方後円墳データベース」は単に古墳の位置を示し、その属性に関する詳細情報を示すにとどまらず、ユーザーが 2 項目の属性とさらにそのサブ項目を選択し、それらの組み合わせを容易に変更して検索でき、その分布状況を知ることができるとともに詳細情報を必要に応じてクリックして表示できるユニークな WebGIS データベースとなっている。本報告では、このデータベース構築の考え方とプロセスを紹介するとともに、今後の課題について述べることにしたい。

1. 1. WebGIS データベースとしての奈良盆地歴史地理データベース

1) WebGIS データベースの現状—埋蔵文化財・遺跡データベースを中心に—

WebGIS とは、インターネットやイントラネット上で、Web ブラウザを通じて GIS を利用可能にする仕組みのことで、最近地方自治体において住民に各種ハザードマップ等に関する防災情報や福祉・医療情報など様々な情報を公開・提供するのに活用されている。WebGIS データベースとは、この WebGIS を利用したデータベースのことであり、データベースから必要なデータを検索し地図上に表示できる。この利点を活かして位置情報が欠かさない遺跡や文化財に関する情報の公開に WebGIS が利用されているといえよう。

遺跡データベースは、奈良文化財研究所が全国を網羅して構築し 1996 年から公開しており、緯度・経度により位置情報も示しているが地図による表示機能は備わっていない。当時は日本における GIS 普及の初期にあたり⁴⁾、GIS の利用は余り意識されていなかったものと思われる。その後公開された「地方官衙関係遺跡データベース」では位置を示した地図がポップアップで表示されるようになり、位置情報を分かり易く示しており WebGIS の代表例とも言えるものとなった。

また、国立情報学研究所による文化遺産データベースは全国を網羅しているものの、位置については所在地の簡単な住所を記載するにとどまっている。文化庁の国指定文化財データベースでは位置については地図表示のリンクがある。このように大規模な機関が関与して構築された遺跡等に関するデータベースであっても位置情報に対する認識や扱いには差がみられるのが現状である。

このほか、地方自治体やその関連団体による GIS を利用した遺跡や埋蔵文化財に関するデータベースも公開されているものも多くみられるようになったが、その多くが所在地、時期、名称、種別によって検索し、位置を地図上に示すということを主としたものである⁵⁾。たとえば、比較的よくできていると思われる岡山県の例では、『おかやま全県統合型 GIS (地理情報システム)』を公開し、これによりさまざまな情報を提供している。文化財情報についても、国・県指定文化財と埋蔵文化財(遺跡)の位置や内容を検索して調べることができる仕様となっている。WebGIS を活用したデータベースでは、地図により位置情報を表示するに際して、多くが GoogleMaps をはじめとする、プロバイダ提供の地図 API を用いている。しかし、上記の岡山県の例でも、地図で表示するのは位置情報が主となっており、static な分布図の表示にとどまっており、必ずしも GIS の特色を十分に発揮しているものとはなっていないといつてよいであろう。

2) 奈良盆地歴史地理データベースの概要と WebGIS データベース

奈良盆地歴史地理データベースは、古代を中心にしつつもそれ以前の古墳時代の開発状況をも視野に入れた遺跡情報、さらに中世に下がる文献史料もデータベースに加え、奈良盆地という比較的広域を対象として歴史的景観とその変遷を明らかにし、日本の古代都城の特質解明に資することを目的としたものである。本データベースは、歴史地理研究の基本資料となる小字地名の GIS データベース「小字データベース」、それと密接に関連する「条里・条坊関連史料データベース」や「藤原京遺構データベース」、「前方後円墳データベース」、「古墳時代集落遺構データベース」の考古資料をデータの核とするデータベース及びこれらとは若干性質を異にするが奈良盆地の歴史地理に関連して自然景観など有意義な情報を読み取ることができる『万葉集』の歌碑が建立されている場所と歌の解説等を付した「万葉歌碑」データベース

ースの 6 サブデータベースから構成され、部分的に web ページで一般に公開している。

奈良盆地の歴史的景観とその変遷を究明することに資するために構築しているデータベースではあるが、飛鳥時代以前すなわち古墳時代に関しては文献史料が存在しない。しかし、地域の開発状況は古墳や集落遺跡などの遺跡分布等からある程度知ることができる。たとえば古墳に関してみると、いつの時期に、どのような古墳が、どこに造営されているかを知ることが、古墳造営の場所が集団の存在する場所から大きく離れていないと考えれば、その周辺に古墳造営の主体となった集団の存在を推定することが可能であり、地域の開発状況を概観する手がかりとなるだろう。特に前方後円墳の築造は地域の有力な首長層に限られていたと考えられ⁶⁾、奈良盆地は古墳時代にヘゲモニーを握ったヤマト王権の所在地であり、その盟主や関係者の墳墓とされる前方後円墳の地域的展開とその実態を把握することにより、開発状況の検討に資することが期待できる。したがって、前方後円墳に関して多様な視点から分布のありようを探ることが必要となるので、GIS の利用による空間分析はそのための有力な方法となると考えられる。このようなことを視野に入れながらデータベース化を進めた。

3) 前方後円墳に関するデータベースの現況

前方後円墳については、早くも 1995 年に小沢一雅氏により学術的な貢献を目指してホームページが開設され、その後前方後円墳研究会(代表:小沢一雅氏)の編集で改訂されてきた前方後円墳データベースが全国を網羅して著名である(図-1)。他にも一般の関心も強く古墳ファンによって各地で様々なデータベースが構築されているが、個人で構築していることから労力に限りがあり、範囲が一地方(県内)等に限られ、そ



図-1 前方後円墳研究会のホームページ

こで利用可能な情報は規模や時期など限られた情報であることが多い。これらのなかには、Google マップ等を利用して詳細な位置表示がなされ、現地踏査に大きな便宜を提供してくれるものもある⁷⁾。しかし、検索機能を有するものでも地名や時期等による検索結果を示して、そこから知りたい古墳名をクリックすると個別の古墳の位置が示されるというもので、ある地域における前方後円墳の分布を示すという性格は有してはいない。

前方後円墳の Web 公開データベースの嚆矢といえる前記の前方後円墳研究会によるデータベースは、地域(府県別)を選択してから古墳名、所在地のほか墳丘長、時期、さらには墳丘の部位の規模や内部主体、外表施設の有無など多様な組合せで、しかも全国規模で検索可能なデータベースであるが、検索の結果は該当する古墳名の一覧が示され、見たい古墳名をクリックすることで位置が表示される。情報工学の専門家が中心になって構築されたものであるだけに、機能面では他の Web データベースとは格段の差がある。GIS が普及する前に構築されたデータベースであるので、古墳の分布状況を見るには適していない。

しかしながら、歴史地理学からみると、貴重な検索結果の一覧を示すだけでなく、検索項目を満たす古墳がどのような分布状況を示すのかは、地域的な視点からは重要な情報であり、検索結果全体を地図上に表示するシステムが望まれるのである。そこで、GIS の特色をできるだけ活用した、検索結果を分布図で示す WebGIS データベースの構築を目指すことにした。

2. WebGIS データベースとしての前方後円墳データベースの構想

1) GIS の特色を活かしたデータベースの考え方・構想

GIS の最大の特色は空間分析が容易にできる点であるといっても過言ではなく、たとえばデータベースにより検索して得られたデータ(事象)を地図上に簡単に表示できる。そこで、前方後円墳のデータベースを構築するにあたり、この特色を活かした WebGIS データベースとすることを目標とし、空間分析の手始めに検索結果を一覧表の形で表示する代わりに地図上に分布図として表示するようにした。地図上に分布を示すことは、前方後円墳に限られたことではないが、地図に示された事象の周囲との関連や立地について判読し得るので、そこから得られる土地の情報の豊富さに注目することになる。判読できる情報・内容は利用可能な地図の縮尺や情報内容にもよるが、通常であれば平野・丘陵・台地、土地の傾斜といった地形条件(図-2 参照)に加えて道路や鉄道などに関する交通条件、あ



図-2 都市圏活断層図(国土地理院)の一部
断層によるケルンバットを利用して後円部が築造されていることが読み取れる。

るいは集落の位置や規模さらに地名などからは社会的・歴史的・文化的条件などが読み取れる。

したがって、次に問題となるのが前方後円墳のデータベースとしてどのような属性データを格納すれば、この特性を活かしたデータベースとなるかが重要な問題点となる。つまり、意味のある地理的分布を表示したりや空間分析が可能となるようにしたりできるような属性をデータとして格納しなければ、その特色を十分に活かすことができない。しかし、データベースに入力する属性の種類が多くなると、その組合せも格段に増加するので、予めどのような属性および属性の組合せが地理的に有効であるかを決定することは、一般には困難である。

次に、データベースのプランニングに当たって考慮した点は、研究レベルでも利用可能なものとすると同時に、地域貢献に資するという点であった。現在大学には地域貢献が求められていることから、成果を地域社会に還元することを念頭に置き、一般にも関心の高い前方後円墳のデータベースを構築し公開するからには、一般の人々が特段 GIS について知識や使用経験がなくても簡単に利用できるようにする必要がある。つまり、使用するに際していかに気軽にアクセスできるユーザーフレンドリーな WebGIS データベースの環境を用意するかということであった。

2) 地理的分布・空間分析を考慮した WebGIS データベースの要件

奈良女子大学古代学学術研究センターでは、研究上の必要性から既に構築していた前方後円墳データベースの経験を基礎として、上記のような構想の下に新たな WebGIS データベースを構築し公開することとした。

最近では GIS への理解が深まりつつあるとはいえ、一

般に GIS ソフトウェアは高価であり、ArcGIS をはじめとする GIS ソフトウェアは、操作にはある程度の習熟が必要であり、いわゆる文系人間にとってはファイル構成も複雑で、mxd と shp ファイルの区別も難しく思われるようである。したがって、歴史系 GIS 研究の副産物あるいは GIS 環境整備の一環として GIS データが作成・公開されても、一部の研究者の間でしか利用されず、歴史系 GIS の発展や研究者・理解者の増加には容易には結びつかないという現実と直面することになる。

そこで前記のように GIS の大きな特色である空間分析の容易さを活かして、直接的には地理的分布の面白さに利用者の関心を誘導するデータの見せ方をセットにして、WebGIS データベースを公開することが望まれる。これには、アプリケーション操作やブラウザの制約を利用者に感じさせないようにユーザビリティを高めるとともに、①前方後円墳データそのものを見る、②分布図(主題図)を作る、③主題図から前方後円墳を理解する、④空間分析的思考をする等が可能となるような工夫が必要となろう。ここで列挙した①～④の事項は基本的には、この順序でより高度な分析へと進んでいくことを念頭に置いている。

また、研究レベルにおいても利用価値のあるデータベースとするためにはある程度データベースに収めるデータの属性項目や各項目のデータ量が大きくなる。しかし、それが余り大きくなると利用するパーソナルコンピュータのスペックにもよるが、Web ページの画面がスムーズに動かなくなったり、表示に時間がかかりすぎたりする。そうすると、移り気な利用者はすぐに別のページへと移動することになる。

以上のような諸要件を満たすにはどのような WebGIS データベースとすればよいか、具体化画面設計やプログラムについては Web ページ制作の業者に委託するとともに適宜打ち合わせをしながら開発することとした。

3. 奈良盆地前方後円墳データベースの構築

1) Web ページ制作—専門業者への委託

上記のような要件を満たし、デザインやレイアウトを洗練されたものとして、ユーザーが利用するに当たって抵抗感が少なく、アプローチしやすいものとするために専門業者に委託した。制作の過程での打ち合わせを通じて種々助言を得られたことは有意義であった。

2) GoogleMapsAPI の使用

GoogleMapsAPI を Adobe の GoogleMaps 用の Flash

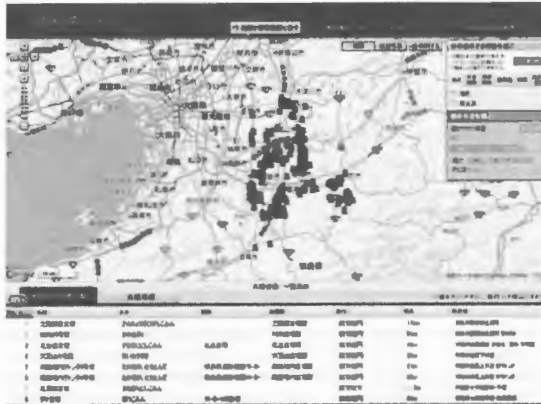


図-3 初期画面

「分布表示する項目を選ぶ」というウィンドウを右上におき、直感的に操作できるようにした

で使用することにより専用の地図表示ソフトウェアの購入の必要がなくなる。さらに、大きな効果として日常的に使用しているインターネットブラウザを使用するので、利用者にとっては複雑な操作は不要で、利用上の心理的バリアが小さいと考えられた。

3) 画面設計と操作手順

画面設計や操作手順は、利用者がマニュアルなどを一切読むことなく直感的に扱えるものを目指す。委託業者のニューロマジックの担当者の助言も得ながら検討を進め、以下のようなものとする事とした。

①初期画面には古墳アイコンを表示し、利用者が全体のデータ数やその範囲、密度などを把握できるようにする。

②メニューは分布図作成と検索、ベースマップの切り替えなどに絞り込み、シンプルにする。

③操作パネルなどは初期画面で画面上に表示し、直感的に「これをクリックしたらよいらしい」と利用者に思わせる。具体的には初期画面に「分布表示する項目を選ぶ」としたウィンドウを表示している(図-3)。

④操作パネルは常に画面上に表示し、タブ表示などで常に所在が分かるようにしておく。つまり、操作に関わるメニューなどが隠れてしまつて、どこにあるのか見つけられないことがないようにする。

⑤メニュー名は「分布表示する項目を選ぶ」など具体的に、利用者が何のための操作をしているかを分かりやすくする。

⑥分布図(主題図)作成時に選択解除、あるいは初期状態に戻すボタンを用意する。

⑦後述の副葬品や形象埴輪のパイチャートは、凡例ウィンドウ上のパイチャート(の該当箇所)を直接クリックして表示項目を選択する。これはチェックボタン



図-4 ポリゴン表示の画面

やラジオボタンを使わないことにより、マニュアルを読まずに直感的に操作できるようにするためである。

4) 利用者が空間分析にも興味を持つようにする表示の工夫

①点表示から面表示へ

ただ単に点記号(ドット)によって分布を示すだけではなく、利用者が興味を持つように表示するために、小縮尺では点記号による表示にならざるを得ないが、大縮尺では前方後円墳の平面形や規模が見ただけで分かるようにポリゴンの表示も取り入れた(図-4)。前方後円墳のポリゴン作成には手間を要したが、利用者に興味を持ってもらうためには視覚的な効果が必要である。また、ベースの地図をGoogleMapsの空中写真モードに簡単に切替えることができるので、墳丘の平面形・規模をみるとともに周辺の地形環境などが理解しやすくなり、古墳の立地等に関心が向けられることが期待できる。

②3D 画像の表示

また、やはり視覚的にも関心を惹起するために、一部の古墳について墳丘の3D画像(アニメーション)を加え、立体的にも前方後円墳を見ることができ

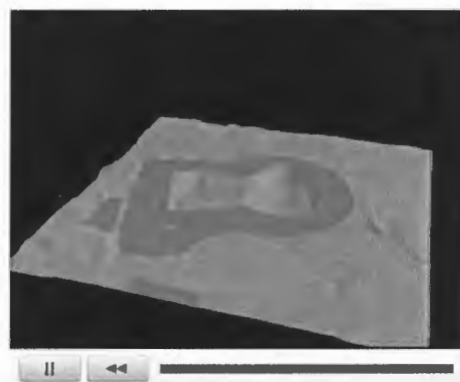


図-5 3D 画像の表示画面(宝来山古墳)

るようにした。これは国土地理院が提供する基盤地図情報(数値標高モデル)を用いて作成した⁸⁾。ポップアップで表示される詳細情報のウィンドウに組み込み、再生ボタンをクリックするとアニメーションを閲覧できるようにした(図-5)。

③利用者による分布図(主題図)の作成

利用者が表示したい項目と表示方法を選択して、主体的に分布図を作成することを可能とした。表示する項目は2項目選択できるようにしたが、その理由は3項目以上とすると組合せのパターンが多くなりすぎてかえって主題図の読み取りに困難を来すと考えられたためである。地図上に表示する際には記号の大きさと形状にも配慮し、墳丘長や石室の大きさのような数値データについては値を段階区分して表示することで分かり易くした。また、副葬品および形象埴輪については、有無のほかにそれぞれ種類による組合せをパイチャートで表示できるようにして興味を深める工夫をした(図-6)。

④古墳の詳細検索

検索の結果分布が示された古墳の内、概要を知りたい場合は当該古墳のアイコン・ポリゴンをクリックすると情報を見ることができ、さらに詳細情報を知りたい場合には「詳細情報」をクリックするとポップアップ画面に詳細情報が表示されるようにした。また、別の古墳の概要を見る場合は、その都度ポップアップ画面を閉じなくても別の古墳をクリックすると、自動的に閉じられて新しい古墳の概要がポップアップ表示される。これにより利用者が多数の古墳についての概要を見ようとする場合のストレスをかなりの程度緩和できると思われる。

5)高度な利用のための工夫

画面の下部において「古墳検索」のスペースを設け、古墳の詳細検索をテキスト入力により行えるようにした。これは古墳名、古墳群名、形状、所在地、出土品、特記事項という項目を選択して、検索したい内容をテキストで入力して検索し、結果を地図上に点滅表示するようにしている。2項目の組合せでも検索できる。たとえば、出土品で三角縁神獣鏡、所在地で天理市と入力して検索すると黒塚古墳が点滅する(図-7)。これは、「分布表示する項目を選ぶ」という簡便さを第一にした検索とは異なり、より高度な検索を可能としようとするもので、項目を増やして研究レベルでも使用可能となることを目指している。

4. 今後の課題—終わりにかけて

前方後円墳は全国的に分布し(ただし、北海道、青森県、秋田県、沖縄県には知られていない)、その総数は5200基に達するという⁹⁾。今回報告したのは、その分布の中心地のひとつである奈良盆地に所在するものを対象としており、僅か325基にしか過ぎない。全国的に展開した前方後円墳に関して、現在の奈良盆地を対象としたデータベースはきわめて局地的であるので、特に研究レベルでの活用を考慮した場合に、利用価値は非常に小さなものとならざるを得ない。したがって、対象地域を拡大し全国規模にすることが、まず第一の課題である。このことに関連しては、データの量にも考慮する必要がある。現在の奈良盆地に限定したデータの場合、エクセルデータのファイル容量はおよそ300kBに過ぎないが、これが全国規模になると単純計算で4.5GB前後になると予想される。この場合、



図-6 出土品構成のパイチャートによる表示と分布

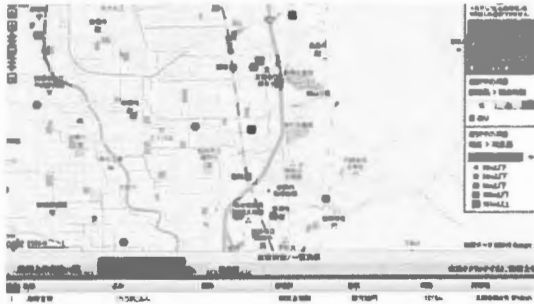


図-7 「古墳検索」結果の画面

web の利用者が使用するパーソナルコンピュータ等のスペックにもよるが、データが重くて動かなくなる可能性が危惧される。その場合、全国規模で検索して動作がスムーズにいくようにするための方法を検討することが必要となる。

前方後円墳は、時期により石室の構造や規模、立地などに変化が見られることが、これまでに ArcGIS を利用した空間分析により確かめられた¹⁰⁾ので、古墳の詳細検索に「時期」の項目を追加し、より有効な検索が行なえるように修正するなど、検索項目や属性データの整理についてさらに検討を進めることも必要である。

本データベースは上述のように Google Maps API for Flash を利用しているが、これに対する正式サポートが 2011 年 9 月に終了した。奈良盆地歴史地理データベースは、Flash を使用して 2008 年度から部分公開を開始しており、サポート終了後も使用してきたが、2014 年 9 月 2 日で機能停止することになっているので、それまでにこれを Google Maps JavaScript API (V3) に移行しなければならないことも課題となっている。

[付記]本研究には奈良女子大学 21 世紀 COE プログラム「古代日本形成の特質解明の研究教育拠点」、平成 19 年度～平成 21 年度科学研究費補助金(基盤研究(C)、研究代表者:出田、課題番号:19520674)、および平成 25 年度～平成 27 年度科学研究費補助金(挑戦的萌芽研究、研究代表者:出田、課題番号:25580174)の成果の一部を使用した。

注及び参考文献

- 1) たとえば絵巻物や古絵図を資料とした研究では、黒田日出男による『姿としぐさの中世史』(平凡社、1986)や『中世荘園絵図の解釈学』(東京大学出版会、2000)などの一連の研究、荘園絵図では小山靖憲『中世村落と荘園絵図』(東京大学出版会、1987)や金田章裕・石上英一他編『日本古代荘園図』(東京大学出版会、1996)などがある。

- 2) 上記のデータベースのうちでは「史料編纂所蔵荘園絵図模本データベース」がデータの項目として「地理」を掲げているが、所在地の住所を記すのみである。地図とリンクさせてその凡その位置が示されれば解釈の助けになると思われる。

- 3) 例として、「おかやま全県統合型 GIS」
URL:<http://www.gis.pref.okayama.jp/map/top/index.asp> をあげることが、できよう。

- 4) 日本において GIS が普及するのは阪神淡路大震災以降で、本格的には 21 世紀になってからと考えてよいであろう。このことは、たとえば北海道 GIS・GPS 研究会の設立が 2008 年であることや埼玉県 GIS 普及推進研究会が 2003 年に設立されたこと等から首肯できるであろう。

- 5) たとえば千葉県、富山県、奈良県、岡山県など太宰府市教育委員会、向日市埋蔵文化財センター。これらは基本的には遺跡地図の情報をそのまま WebGIS データベース化したものであるため、ベースマップが 2.5 万～1 万分の 1 であるものが多く、また内容が大まかであったり、複合遺跡などではポリゴンが重なりあって操作に不都合であったりするために、スムーズに操作できないことも多い。

- 6) 都出比呂志(1991):日本古代の国家形成論序説—前方後円墳体制の提唱—、日本史研究 343、pp.5-39。

- 7) たとえば、「埼群古墳館」(URL:<http://sgkohun.world.coocan.jp/>)など。

- 8) 来山古墳の標高データは 21 世紀 COE プログラム「古代日本形成の特質解明の研究教育拠点」の成果を利用した。

- 9) 広瀬和雄:『前方後円墳国家』、(角川選書)、角川書店、2003 年、p.7。

- 10) 拙稿(2009):「近畿地方における前方後円墳の分布論的検討—「奈良盆地歴史地理データベース」の利用による墳丘規模と副葬品の時期別分布を中心に—」、古代学(奈良女子大学古代学学術研究センター発行)、1、pp.35-44

拙稿(2010):「九州地方における前方後円墳の分布論的検討—墳丘の規模と内部構造・副葬品の時期別分布を中心に—」、査読無、奈良女子大学地理学・地域環境学研究報告、VII、65-76

主要参考文献(上掲以外)

- 小沢一雅(1988):『前方後円墳の数理』、雄山閣
近藤義郎(2005):『前方後円墳の起源を考える』、青木書店

- 近藤義郎編(1992)『前方後円墳集成』近畿編、山川出版社
- 近藤義郎(1983):『前方後円墳の時代』、岩波書店
- 都出比呂志(2005):『前方後円墳と社会』、塙書房
- 津村宏臣(2006):「文化財と時空間情報科学」(東京文化財研究所国際文化財保存修復協力センター編『文化財の調査研究および保護に対する地理情報システムの利用』所収)、pp.13-31
- 奈良県立橿原考古学研究所編(2001):『大和前方後円墳集成』、奈良県立橿原考古学研究所
- 広瀬ほか編(2004):『古墳時代の政治構造：前方後円墳からのアプローチ』、青木書店
- 森浩一編(1986):『前方後円墳の世紀』(『日本の古代第5巻』:貝塚茂樹・江上波夫・司馬遼太郎監修)、中央公論社

広島県立美術館館蔵品データベース構築について（事例報告）

Constructing the Database of Collection of Hiroshima Prefectural Art Museum

福田 浩子

Hiroko Fukuda

広島県立美術館学芸課，広島市中区上幟町 2-22

Hiroshima Prefectural Art Museum, 2-22 Kaminobori-cho Naka-ku, Hiroshima

あらまし：広島県立美術館は、2011-12年に館蔵品データベースを構築した。これは館内の行政LAN上で学芸員が使用する管理用データベース、web公開用データベース、そして館内来館者閲覧用データベースの3つを諸事情のために非同期連動させた画像を含むデータベースである。今回は、行政LANの制約のためにデータ更新内容を同期連動させない当館の新データベース構築事例を紹介したい。

Summary: Hiroshima Prefectural Art Museum constructed the database of Collection in 2011-12. It is constructed by three image databases as the management database, the database for website, the database for museum visitors, those are asynchronised and interlocked. Here I introduced our case of database construction.

キーワード：人文科学，データベース，画像データベース，美術館，web公開

Keywords: humanities, database, image database, art museum, website

1. データベース構築の契機

今回の作品データベース構築については、厚生労働省による雇用創出基金事業の緊急雇用対策一般公募事業のひとつに、有限会社アイ・ティー・エスによる「広島県立美術館収蔵品検索サイトの構築事業」が選定されたことから始まった。同事業の提案者は、前述の有限会社アイ・ティー・エス（代表取締役：田川育伸氏）で、同社は企業独自のシステム構築を主に行っている広島市内のソフトウェア開発会社である。それまで当館との事業実績はなく、当館側も予期していない事業案であった。公募選考途中以降に設けられた当館側の担当が筆者であった。

緊急雇用対策の事業であるから、データベース構築の各工程において雇用を生み出すことが目的のひとつであり、システム開発を中心的に行った田川氏をはじめ、美術館側との各種調整役の藤原氏はもちろん、多数の外

部の方の力をお借りして、今回のデータベースが作り上げられた。

2. 今回のデータベース構築以前の状況

昭和43(1968)年に開館した広島県立美術館は、平成23(2011)年度末現在で4,399点の美術作品および美術資料を所蔵し、保管維持すると同時に展示の用に供している。平成8年のリニューアル・オープンを控え、広島県ゆかりの美術作品、1920-30年代の美術作品、日本を含むアジアの工芸という3本の重点収集方針に則ってコレクションの充実をみた。平成2(1990)年以前の館蔵品に関する情報は冊子体の目録に掲載されていたが、それ以降の収集品についての情報は作品台帳と作品カードが主力であった。

(1) 作品台帳

台帳記載の項目は、受入年月日、作者名、生没年、作品名、制作年、納入者、購入価額、備考である。場合によって各作品の担当学芸員が追記している場合がある。

(2) 作品カード

作品カードの項目は、表面に大分野、中分野、備品記号、中分野の作品番号、員数、作品名、制作年、寸法、材質技法、制作者、制作者の生没年、所属、納入者（納入者、寄贈者、その他）、備考、写真貼付欄、写真管理欄、裏面には参考文献、使用歴欄がある。

作品カードとポジフィルム、その他の関連資料（主な素材は紙）はカットフォルダーに1作品分ずつ収め、カードキャビネットに収納して管理している。また、作家＝制作者についての情報は、別に作家ごとのカットフォルダーを使用して保管している。

(3) 冊子体の館蔵品目録

冊子体として最後に印刷・発行された館蔵品目録は、平成2(1990)年発刊の館蔵品目録である。これにはそれ以前に収集した計1,594点がモノクロ写真とともに収録されている。現在とは分野ごとの作品内容や作品数がまったく異なっており、現状の主要コレクションの多くが掲載されていないため、すでに役割を果たした目録といえる。本目録発行に前後して、リニューアル・オープンに向けて重点収集が開始されていたが、新たなコレクションの情報を一覧できる目録は現在までついに発行されていない。

(4) 館蔵品データベースシステム

当館は平成8(1996)年リニューアル・オープンにあたり、館蔵品データベースをはじめとして各種の機器を導入した（経緯と詳細については、角田新「広島県立美術館の情報機器について」『広島県立美術館研究紀要』第5号、2001年）。なお、当時の館蔵品データベースについてはp.3に記述がある。

その次段階として、平成18(2006)年度以来、今回のデータベース構築まで使用してきたデータベースが富士通のMusethequeV3 Light/SSであり、現在も館内公開用館蔵品検索システムとして稼働中である。今回新たに構築したデータベースとの関係については、後述する。Musetheque自体は博物館・美術館資料の管理に特化したデータベースであるから、本来ならば館蔵品すべてをこのシステムで管理すべきところであった。しかし、そう

ならなかった大きな理由のひとつは、広島県庁の職員が使用するネットワーク（以下、県庁LANと呼ぶ）の制限によって、Musethequeを接続することができなかったことである。そのため、Musethequeは日常的に更新が必要な作品管理には充分活用できず、館内の図書室に設置された来館者用作品検索システムとして稼働している。

3. データベースの枠組み

今回のデータベース構築に当たっては、当館で使用中の来館者用データベース Musetheque を残すこととし、管理用データベースは県庁LAN上の各学芸員のパソコンで使用できること、かつ、インターネットを使用した当館ホームページ上でもweb版館蔵品検索システム(web公開用作品データベース)を公開することを目標に、「仕事に使える」データベースの構築を目指した。

そして、3つのデータベース

管理用データベース

館内公開用 Musetheque

web 公開用データベース

を連携させる方針がアイ・ティー・エスより提案された。

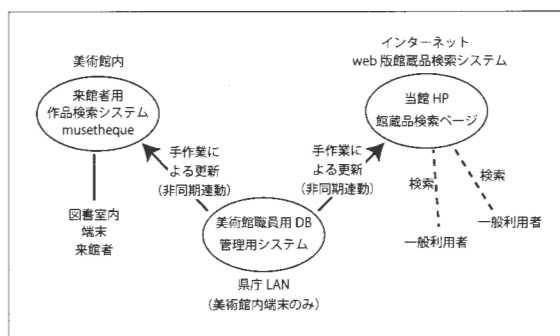


図1 データベース関連図

このような複数のデータベースは変更内容が随時同期されるのが常識的だろう。しかし、学芸員が使用しているパソコンが県庁LANに接続している事情により、管理用データベースの更新内容は他の2つのデータベースに同期するのではなく、ハード的に切り離した状態を維持し、更新作業はUSBメモリを介して手作業で行うという、あまり例のない方法をとることとなった。webデー

データベース公開後の更新作業は、年間数回、必要に応じて担当者が行っている。

4. web 公開用作品データベース

インターネットを使用できる環境ならばどこからでもアクセスして当館の館蔵品に関する基本的な情報を得られる web 公開用作品データベースは、当館にとっては初めての試みであった。

データベースへのリンクは、当館ホームページのトップページと、所蔵作品展タブのすべてのページに置いた。トップページでは、右カラムの所蔵作品紹介枠の下に館蔵品データベースへのバナーを貼っている。

館蔵品データベースに入る前には、当データベースの定義内容、著作権について了承したことを確認するボタンを押すこととなる。このため、各作品情報の html ページへの外部からのリンクはできない仕様となっている。また、館蔵品検索ホーム画面の背景には、当館を代表する作品である《伊万里柿右衛門様式色絵馬》2体を置き、web 版データベースへの来館者へ親しみを感じていただけるようにデザインを工夫した。

このホーム画面からは6種類の検索方法を選べ、各検索タブ選択時も他の検索タブが表示されており、いつでも他の検索方法へ移行できる。



図2 館蔵品検索ホーム

(1) おすすめ作品を見る

まず、「おすすめ作品を見る」では、当館を代表する作品群を一覧して、ビギナーや特に検索目的を持たない訪問者が気軽に楽しめるページとした。初期画面ではテ

キストで右上のボタンで「サムネイル表示」へ切り替えられる。また、同ページ「スライドショー」リンクからは「おすすめ作品」に設定されている作品が次々にスライドショー表示される。



図3 おすすめ作品一覧



図4 おすすめ作品サムネイル一覧

(2) 厳選30の作者を見る

このタブでは、30人の作者リンクから、その制作者の作品を一覧表示する。

(3) 作者名で検索

作者名での検索は、作者名の漢字と読み（カタカナ）による検索枠を上部に設けている。下部にあ行～わ行のタブを作り、作家名一覧からも選択できる。

当館は国内作家だけでなく、海外作家や海外の工芸作品を所蔵しているが、これらも「ウズベキスタン」や「エルサリ族、トルクメン」のようにカタカナ読みでタブ内リストに表示している。

(4) 作品名で検索

作品名での検索結果は、作者名同様に漢字または読み(ひらがな)の部分一致で30件ずつ表示される。

(5) 連想ワード検索

連想ワード検索は、アイ・ティー・エスによる提案により実現した、他館のデータベースにない特徴を持つ検索方法である。

作品名に含まれない、作品イメージによる単語が登録されていて、抽象的な言葉や文様、地名などでヒットする。ひらめきによる作品の印象が言葉として当てられているので、当然のことながら、時には検索者がイメージするものとは異なる結果となる可能性もある。

(6) 分野で検索

最後は分野別の検索である。大カテゴリから小カテゴリを選ぶと、検索結果として作品が30件ずつ表示される。

例えば、絵画内のカテゴリとして日本画、油彩画、版画、水彩画、素描、パステル画、平面造形、拓本があり、工芸内カテゴリとしては陶磁、染織、金工、漆工、木竹工、ガラス、七宝、人形、甲冑刀剣、その他の工芸といった主に素材による分類がある。

分野は、館蔵品管理のため使用している分類を採用している。なお、絵画カテゴリの平面造形には、もしかすると別分類にした方がよいかもしれない写真や工芸的平面作品の一部も過去の収集経緯の中で含まれている。

(7) 個別作品ページ

上述の6種類の検索方法の結果、表示される個別作品の詳細ページには、作品の基本情報として作品名、作者名、材質・技法、寸法、制作年、作品解説の項目を置いている。学芸員が分担執筆した作品解説はweb版データベース公開時点で約150点掲載し、順次追加している。



図5 作品詳細ページ

左側の作品画像をクリックすると、拡大画像が表示される。web公開用画像は、縦横の長辺を500ピクセル、72dpiとし、電子透かしを入れた。

5. 管理用作品データベース

前述のように、県庁LAN上にある管理用作品データベースは学芸員が主に作品管理に使用する。web公開用館蔵品データベースや館内来館者用データベースの基礎データを管理するものである。

(1) ユーザー管理

県庁LANにIDを持つ学芸員およびその他の美術館職員に、データベースへのログインは個別のIDとパスワードを使用し、それぞれの使用状況が管理者モードからは見えるようになっている。また、ユーザフループ管理として、カテゴリごとにアクセス者を制限することもできる。

(2) 作品情報管理

このデータベースの根幹である作品情報の管理を行う部分である。各作品情報は作者名、作者名(カタカナ)、作品名、分類記号、分類記号と組み合わせた作品番号から検索して表示できる。

作品情報検索結果として表示されるリストの頭のチェックボックスにチェックすることで、サムネイル表示や作品目録作成などに活用するためのcsv形式での出力ができる。

作品を選択するとその作品の詳細情報が表示される。

ページ1では、分類番号、作品名、制作年、寸法、材質・技法など。ページ2では、web 版作品データベースの「連想キーワード検索」のキーワード、web 版や館内検索の公開フラグ、「おすすめ作品」掲載フラグ、作品解説など。ページ3は、備考や貸出歴の記録となっている。貸出歴はさらに別ページでそれぞれのデータを記入する。

図6 作品詳細情報 1

図7 作品詳細情報 2

図8 作品詳細情報 3

このシステムでは、館内用と web 用にそれぞれデータをパッケージとしてエクスポートするため、画像データのファイル、そして作品情報（テキスト）の 1 件ずつのデータに対して、館内用と web 用に公開するか非公開とするか選択できるようになっている。この選択は各作品情報詳細画面でも行うが、公開フラグ一括登録画面ではリスト形式の状態で公開・非公開の区分を変更できる。一多数の作品を扱う場合に便利な機能である。

(3) 分類マスタ

分類マスタは作品の分類を整理して、作品情報に割り当てる基本情報管理である。館蔵品は美術館所蔵作品として管理されると同時に、県の物品管理規則に基づいた管理も行われるので、作品としての分類番号と県の物品という意味の備品番号とが併存する。この分類マスタは、作品情報を管理する際に、これらの分類や番号を扱う手間をできるだけ簡便にするものである。

(4) 作者情報管理

作者情報を管理するページである。検索後に表示される作者のリストから作者詳細情報を開く。作者という名称ではあるが、作品によっては「アフガニスタン北西部」のように地名や民族名が作者名として便宜的に扱われている。館によっては、地名や民族名は備考欄に入れられるが、当館の場合は制作者情報としているため、このような処理となっている。

(5) データの構造

このデータベースの構造は次のようである。

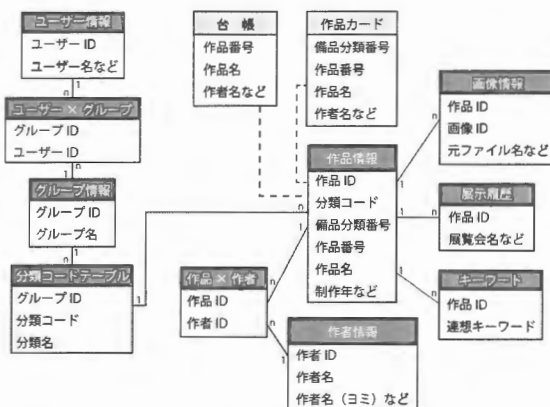


図9 データ構造の概念

実際のリレーショナルデータベースの各テーブルの項目はさらに複雑である。台帳と作品カードはデータベース内には存在しないが、登録データの作成にあたって基本情報となったものである。

6. EUC ビュー

管理システム内に登録されたデータは、データベース外から EUC によってデータ抽出ができる。EUC とは、エンドユーザーコンピューティング (End User Computing) の略で、一般ユーザーが生産性の向上をめざし、コンピュータを使い簡単にデータベースを検索・抽出することを意味する。管理システムを使わずに、Microsoft Access 等からデータを抽出し、Excel にデータコピーして利用することができる。今回のデータベース構築にあたり、開発者側からは基本的な作品情報や作者情報など、12 種類の EUC データが提供された。また、Access を通して登録データのすべての項目を抽出することができ、当初は予想していなかったきわめて重宝な機能である。

7. 管理システムのバックアップ

管理システムのデータのバックアップは毎月 1 回、管理サーバー内で定期的に行われている。管理サーバー内のバックアップなので、別媒体への退避は手動で行う。

8. おわりに

今回のデータベース構築によって、現在までの作品情報がデジタル化された。このデータベースを将来も活用していくためにはデータの追加更新を継続し、最新データを維持することが大切である。また、機器の故障や災害時にこれまで培ったかけがえのないデータを消失することをできるだけ避けるため、データのバックアップを心がけたい。データの保護という観点からは、例えば遠く離れた場所にサーバーが設置されている文化遺産オンライン等の外部データベースに参加することによって、一部のデータだけでも確保しておく可能性を考え、現在までのコレクション情報を文化遺産オンラインへ登録している。

本稿では、このデータベースの構築に関わる経緯や内容について記録としてまとめようとした。

著作権の問題や作品写真の有無により、作品画像情報は公開できないものもあるが、テキスト情報は可能な限り登録し、web 公開を継続している。当館の館蔵品に関心を持つ人が web を通じて情報に接していただくことを願う。作品情報の管理とその情報を web 版データベースによって公開することで、当館館蔵品の活用を進めることができると信じたい。

謝辞

美術品への理解と情熱をもって当データベースを開発された有限会社アイ・ティー・エスの田川育伸氏および藤原真澄氏、多大なる協力を得た当館総務課・学芸課全職員、連想キーワード割り当て等を受け持たれた加藤真理氏、データの更新および文化遺産オンラインへの情報登録に協力された向井知美氏をはじめ、お名前を記すことのできない館蔵品データベース構築と維持作業に関わったすべての関係者の皆様に感謝します。

参考文献

- 2001 角田新「広島県立美術館の情報機器について」『広島県立美術館研究紀要』第 5 号、広島県立美術館 (http://www1.hpam-unet.ocn.ne.jp/htdocs/imagess/about/BulletinNo05_Kakuda2001.pdf) にも掲載。2013. 1. 31 参照)
- 2004 村田良二「東京藝術大学大学美術館収蔵品データベースの構築」、『アート・ドキュメンテーション研究』第 11 号、pp. 61-73、アート・ドキュメンテーション学会
- 2010 佐藤克己「新潟県立近代美術館・万代島美術館所蔵作品データベースの作成について」『新潟県立近代美術館研究紀要』第 9 号、pp. 36-39. (<http://www.lalanet.gr.jp/kinbi/collection/bulletin/09/sato.pdf>, 2013. 1. 31 参照)

データベース活用におけるバザール型コミュニティ形成について

About the Bazaar Type Community on Database Utilization

藤岡 洋

Hirosi Huzioka

東京大学 東洋文化研究所, 東京都文京区本郷 7-3-1

The University of Tokyo, 3-7-1 Hongo Bunkyo, Tokyo

あらまし: 人文系研究機関でデータベースの乱立が始まってから早 15 年近くたつ。今回はグローバル COE、21 世紀 COE の前身、COE 形成プログラムで構築したデータベースを振り返る。その上で、なぜデータベース継承が難しかったのか、問題を整理したい。さらに、データベース活用は幅広い利用という観点とそれに関連する技術が主流を占めているように見受けられるが、今回はバザール型コミュニティを仮定し、その視座から上記2点の可逆的に考えてみたい。

Summary: This paper attempts to explain why database succession has been difficult. At the moment, almost 15 years after the initial increase and flooding of databases have passed. In many ways there is a strong case for stopping their activity, i.e. because of the lack of concept for a programmer or a manager in manual documents or technical reports. In this paper, I elaborate the meaning of a "player" in database utilization. In order to do so, I present an overview of "concept map for database" and "bazaar type community on database utilizations".

キーワード: 人文科学, データベース

Keywords: humanities, database

1. 大規模科研でのデジタルアーカイブ構築を振り返る

平成 11 年度から 15 年度までの 5 年間、東京大学人文社会系大学院には象形文化研究拠点(以下、拠点)が設置され、その主要プロジェクトの一つとして国内外の古代ローマ史考古学者、美術史家、情報学者などが協力してデジタルアーカイブを作る試みがなされた。このプロジェクトは文部科学省 COE 拠点形成プログラムとして実行された。COE 拠点形成プログラムは、グローバル COE、その前身である 21 世紀 COE のさらに前身となるプロジェクトであった。この拠点形成プログラムの目的は次のようなものであった。

「高度情報化社会においては、情報科学の技術的進展に応じて情報通信基盤の整備だけでなく、情報内容(コンテンツ)が重視されるようになるのは必然であるが、近年喧伝されている「コンテンツ重視」の主張は、とすれば、情報の量的増大のみを招来するに留まっている。他方、情報の出典の検討と情報の質の検証という観点からみれば、人文学は対象とする原資料ごとにその基本的操作法を方法論として確立している。本研究は以上のような現状認識に立って、高度情報

化社会において最も重要なものとなる情報のコンテンツに対して、基礎研究である人文学の立場から寄与しようとするものである。研究の機軸とするのは、人文学がその基盤として確立してきた原資料批判の方法論(史料学・文献学を含む資料学)を再認識して統合し、その総合の上に立ってコンテンツの学としての情報学と融合させることにある。そのために、これまでの人文学で確立されてきた原資料批判の方法論とその成果の蓄積である学術情報を、人文学の基本的学的方法論の観点から更なる体系化と統合を行い、象形文化アーカイブへと集約する」[1]

1.1 アーカイブの概要

ポンペイデジタルアーカイブはこの拠点プロジェクトの中心的なコンテンツであった。20000 枚ほどのポジフィルムから大型スキャナ(Heidelberg, Imacon など)を使ってのデジタル化と PPM(Pompei Pitture e Mosaici)とそれを元とした最新の文献研究成果の統合を目指した。データベースの果たす役割は、いうまでもなく、この「統合」部分であった。筆者は創設当初には拠点に所属していなかったため詳しい経緯は知らないが、日

立制作所の DIS 開発と同時並行的になされており、すでにプロトタイプが出来上がっていた。[2]



図 1. 日立 TSS

図 1. のとおり、当初のアーカイブはスタンドアロンのタイプであった。おそらく博物館などでの展示公開を念頭に置いていたと思われる。その仕様は、2000 年初頭としては珍しくタッチパネルを採用しており、利用者はローマ帝国全図から地図階層をたどっていき、イタリア、ナポリ、ポンペイ、地区、建物、部屋、壁、最終的にはこの領域では標準となっている 9 分割区分の領域画像一覧が画面下 2 分の 1 に表示される。そのサムネイルをタッチすることで地図から拡大画像に切り替わり、サムネイル表示部分に PPM 情報やそれに基づく最新の学術情報などの各種学術情報がタブ形式で表示される、というものであった(参照: 図2)



図 2. 拠点アーカイブ(例: 最下層)

ただこのスタンドアロンタイプは仕様上というよりも、構築プロセスに大きな問題があった。画像の著作権について日立に全面的に譲渡することに躊躇した拠点側が、それをコントロールするために、1) 候補を挙げ、2) 合議にかけ、3) CD-R に焼いて日立に送るという方

法が採られていた。日立側でデータベースに載せる作業に数週間から数ヶ月を有した。

ここからアーカイブ構築を完全に拠点側で構築する決定がなされたのは拠点形成後 3 年以上たってからのことであった。

1.2 システム

拠点の選択肢は日立のデザインを再現してウェブブラウザ上に載せるものであった。

日立との連携で生じた問題を解消する目的もあって、すでにサーバの管理運用の補助を委託していた外部業者に文字検索インターフェイスを委託し、データ(画像とテキスト)入力インターフェイスを別の関連業者に委託した。委託業者との関係は、日立と違い、サーバ管理者の知己であり、問題だった構築に関する障害はなくなった。アーカイブ本体は古代美術史の専門知識を持つ拠点スタッフによって一人で構築された。それぞれ PSS (Pompei text Searching System), DIS (pompei Data Input System), TopV (pompei TOPographical Viewer) として新たな船出となった。アーカイブに載せられるデータ内容もこれに伴い急速に充実していった。年数回、イタリアから研究者を招聘し、集中的にテキストデータ、画像データそのものについて深夜まで議論がなされ、その結果は以前とは比較にならないスピードで更新されていった。

だが、スピードがあがると同時に、次の問題が生じた。すでにこの頃は最終年度に入ろうとしていて時限との戦いがすでに始まっており、一言で言えば「もっとスピードを」ということであった。

またウェブベースに載せたことで公開に際しての準備も潜在的な問題として残されていた。以前日立との間で生じ、その後保留したままになっていた著作権の問題だった。ポンペイには江戸時代と同じほどの長さの発掘の歴史があり、かつ未だ発掘作業が行われている。ゆえに、拠点のアーカイブには非公開状態の画像も多く載せられており、これをそのまま公開してしまっているのかについては議論が深められる余裕がなかった。

これらの諸事情を抱えたまま拠点形成プログラムは最終年度を迎えた。ポンペイデジタルアーカイブはいわゆる「落ち穂拾い」を必要とする状態で最後の国際シンポジウムを主催した。幸いなことに、その後研究代表者であった青柳正規先生(現:文化庁長官)が国立西洋美術館館長として着任されるという契機と、東京大学学環情報学府の坂村健先生のご好意もあり、東京大学病院分院跡地の研究室でアーカイブ構築は継続されることとなった。

この時期には時限との戦いに加えコストの問題が生じている。前任サーバ管理者の知己によって作られた上記2つのシステム (PSS, DIS)の改良、メンテナンス・機能追加要望に割くコストはこの時期の拠点にはすでに難しい問題になってしまっていた。PSS については、php+postgresql という当時のスタッフで十分メンテナンス・改良に耐えうる環境で書かれていたため、問題は最小限に抑えられたが、DIS については、Microsoft Visual Basic (以下、VB)で書かれていた。当時、VBに明るいスタッフがいなかったため、データ入力・編集に関する要望に応えられなくなってきた。このため、DISを再度書き換えることにした。この部分は筆者が担当し、PSS, TopVと同じ php+postgresql で新たに書きなおした。DISは DECS (pompei Data Edit ComponentS)として生まれ変わった。この間にすでに2年が経過している。

DECSは単に研究者のためのデータ入力システムを目ざしてはいなかった。TopVではコンテンツの内容・配置について自由に変更ができなかったし、これまでの経緯を踏まえると人材に頼るシステムはいずれ破綻する恐れからだった。そこで、DECSはTopVすなわちコンテンツそのものにアクセスをして、すぐさま研究者の研究成果をコンテンツに反映できるようにしていったようにした。

だが、拠点はその後1年をもって終了する。COE拠点形成プロジェクト開始から数えれば計8年にも及ぶプロジェクトだったが、内実は試行錯誤の連続であった。しかもすでに同時期、イタリアでは同様に画像を全面に押し出したデータベースが幾つか公開されており、その数年後には Google map / Google earth によってポンペイを利用してそれをパノラマ化して散策できるようにさえなっていた。拠点のコンテンツはこれらの事例と理念からしてもまったく異なるもののはずだったが、それを差別化する方策を拠点スタッフでは当時見いだせなかったし、また見いだしても実行に移すことはもはや不可能であった。

拠点消滅の直接の原因は時限とコストという外的なものだったが、拠点アーカイブ独特の事情として画像の使用許諾の問題があった。この手の問題は現在のデータベースにはあまり見られないので今後のデータベース開発には参考にならないであろう。その一方で、データベース継承については、いくつか反省すべき点が残されたままだった。一つは継承そのものに対する考え方である。拠点アーカイブは学術的価値、データの規模からして当然学術世界から注目されるであろうという甘い見込みがあったことは否めない。しかし、これはデータベースを構築するすべてのプロジェクトに

当てはまるのではないだろうか。その先には継続利用がイコール、コンテンツの存続につながるという甘い見込みにつながっている。確かに、継続して利用されるウェブサービスはその技術はもとより、利用され続けることがその存続の大事な要素になってはいるだろう。しかし、学術データベースにとってそのような形での存続は必ずしも求められていないし、開発者も求めている場合が多い。もう一つはデータベースの完成度とといったときの意味である。これは本稿の意図の中核部分であるが、データベースの完成度は基本利用者本位で図られる。だが、それがプログラム言語で書かれているかぎり、データベースがその後置かれる状況如何にかかわらず、再生可能なデータベースこそが継続性を備えたものなのであり、そうでないものはデータベース開発者が考える以上に深刻な問題をすでに抱えてしまっていると言わざるを得ない。そして我々拠点スタッフにはそのような発想がそもそもなかったのである。

2. 現在のデータベース活用に関する所見

とはいえ、いま述べてきた状況は2013年現在でも散見できる。相変わらず科研費データベースでも、データベースの構築、という文字が踊っている。これらのデータベースが数年後、十数年後停止に追い込まれるかもしれないという予測に対する対策に関するアイデアを目にすることは少なく、現況に対して楽観的すぎはしないかと心配になる。

ただデータベースの保存ないし継続的運用を視野に入れたデータベースの活用についてのアイデアは、少しずつ裾野を広げつつあるようにも見受けられる。

- 公開方法のバリエーション。モバイル環境の変化に応じたデータベースの利用範囲の拡張。
- 出版。データベースとの紐付けに関して。
- ワークショップ。公開に留めない積極的な活用方法の模索。
- データベース同士の連携など。

前者2つに関しては拠点アーカイブでも試みていたことであったので、紹介しよう。

2.1 公開方法のバリエーション

拠点アーカイブでは、最終年度にある意図をもって、二つの実験を行っていた。その意図とは、アーカイブを外に持ち出すことで、研究のフィールドにより近いところでのデータベース活用の可能性を探ることであった。

そのために、まず 2009 年に国立西洋美術館で開催された「古代ローマ帝国の遺産」展 [3] でのギャラリートークを利用して予備的な実験を行った。



図 3. 国立西洋美術館でのコンテンツ配信実験

図 3.には二つの PC が見られるが、実は演壇の講演者が説明に使っている TopV は、手前で聴講している女性の PC から発信されている。

この実験が成功して後、続いてポンペイに出向き同様の実験を行った。この時は、学術データベースが外に持ち出されることで得られる効果を考察するためだった。ポンペイでは研究者が指導のため学生を連れて公開停止中の遺構を訪れるが、その際の補助ツールとして、TopV を野外に持ちだしてもらったわけである。実際にこの時点でデータにして検討すべき課題が幾つか見つか、その場で大学院生と教員とで議論がなされるといった光景が見られた。これをメモする学生がおり、編集に関して柔軟性を増していた DECS を利用すれば、データ監修そのものに役立つであろうという新たな展望も寄せられた。この実験は成功だったといえてよいが、残念なことにその後すぐ拠点アーカイブは停止してしまった。

一つ脱線するならば、拠点が閉鎖されてまもなく、2010 年に初代 iPad が発売されたが、その宣伝で、ポンペイの遺構(とキャプションにあった)に入って土まみれの作業員が iPad を見て何かを確認するという画像(参照:図 4)が使われていた。

これを見た瞬間、まさに著者が想像していた図が目の前に提示されたことに衝撃を受けるとともに、明らかにこれは問題のある嘘の広告だと分かった。少なくとも同時ポンペイに公衆無線 LAN は届いていなかったからである。しかし、いずれにせよ、同様の使い方は当時の iPad 開発者たちも考えていたという、何か背中を押されたような感傷とも、拠点閉鎖の寂寥感ともいえない複雑な気持ちになったことを覚えている。



図 4. 初代 iPad 発売時の広告

2.2 データベースとの紐付けされた出版

学術データベースの活用法として、次に挙げられるケースが多いのはデータベースの資産を使っの論文執筆、著作刊行であろう。



図 5. POMPEI INSULA OCCIDENTALIS P.167

拠点でも最終的に両国の研究者が合同で取り組み一つ大きな成果をあげている。Masanori Aoyagi -Umberto Pappalardo "POMPEI (REGIONES VI-VII) INSULA OCCIDENTALIS -Volume Primo-", 560 pages, Valtrend Editore, 2006 である。

この著作には一つ、特徴があった。使用された画像(モノクロ・カラー)には、拠点アーカイブで監修・管理していた 16 桁の固有 ID が振られている。この刊行物から他の研究者が、オリジナルの画像情報に接しようとすれば、拠点アーカイブにアクセスし、その他関連

情報は誰でも調べることができた(参照:図5)。このことはアーカイブと出版物はセットとして扱うことができる可能性を示そうとしていたわけである。データベースが継続運用されていれば、さらに出版物そのものが学術成果の記録(ログ)として位置づけられたであろう。

しかし、プロトタイプとしての著作物刊行は単発で終わってしまった(そのタイトルどおり、予定では第2巻、第3巻が少なくとも予定されていた)こと、そして繰り返すが上記著作物もデータベースが運用を停止してしまったことで狙った効果が半減してしまった。現在ではこのIDの意味は消失しているわけである。データベースが停止すると書物の意義までが落ちてしまう現代ならではの現象の一つであろう。

2.3 アウトリーチとしてのワークショップ

それでもデータベース活性化の手段としてそれをつかった論文・著作活動を後押しすることが有効であるとするならば、もっとドラスティックに考えることもできる。すなわち、データベース使用を登録制にして、必ず論文ないしレポートの形でフィードバックを要求する、というものである。だが、まず利用者はタイトルだけでどのようなデータが収められているかも分からないようなデータベースに登録などしないだろうし、先に中身も見せないうちからフィードバックを求めるようなハードルに真っ向から付き合ってくれるとは到底思えない。

ところが、学術データベースを質で測るという観点に立てば、次のような試みは魅力的である。著者が現在所属している東京大学東洋文化研究所には10年近くの運用実績のある「アジアパロメータ」というアジア比較社会学のためのデータベースがある(参照:図6)。このデータベースはアジア各国に住む20歳以上69歳以下の全ての成人を母集団として対象とし、基本的に調査員による面接聴取法によって有効回収数が必要サンプル数程度になるように、各国ごとに期待される回収率を勘案して標本規模を800から2,000サンプル採り、層化多段無作為抽出法と割当抽出法を用いて、1.社会基盤整備、2.経済状況、3.生活における価値観・満足度、4.社会的行動の価値観・満足度、5.アイデンティティ、6.政治意識、7.健康状態のデータを蓄積したものである。

このアジアパロメータを使ったデータベース活用法がユニークなのは、当該データの価値を十分理解している教員たち(代表は東京大学情報学環・東洋文化研究所園田茂人教授)が日本はもちろん、中国や台湾や韓国、それにアメリカ各地で講習会を開き、学生にデータの使い方を指南した上で、ワークショップ開催、コンソーシアム設立、コンテスト主催などを通じて、

研究者や学生を労力を惜しまず鼓舞して、優れた論文を世に送り続ける試みを行っていることである。

実際その研究成果[4],[5]は、日本の社会学会の中でも一定のポジションを得るにいたっていると聞き及んでいる。

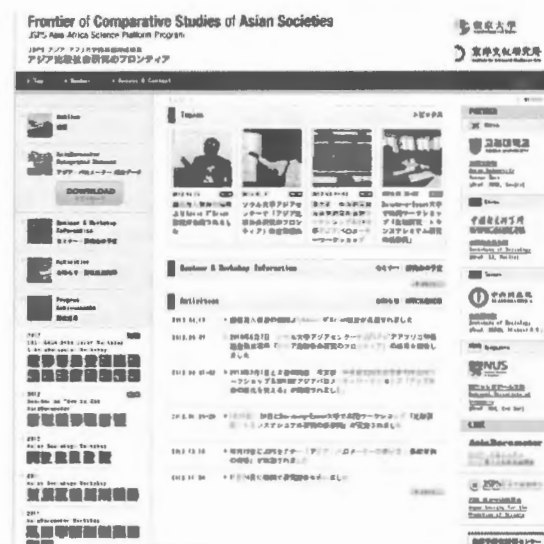


図6. 東京大学東洋文化情報センター内
アジアパロメータのフロントページ

しかしながら、すべてのデータベースにこのアジアパロメータのような活用法が適用できるかというと、まず現実的に難しいことは容易に想像がつく。

第一に、このようなデータベースの活用には構築とは別の、ある意味では構築以上の「労力」が必要になることである。アジアパロメータの例をみるに、主催する教員がこの所業を義務や負担といった後ろ向きの姿勢で行った時点で、試みは失敗する可能性を孕み始めることは明らかである。その問題がないとしても、年に何度も海外に出向いて各種プランを練り直し、実施し、修正をする必要があろう。

第二に、その時その時に労力を厭わない「人材」の有無に大きく左右されざるを得ない、ということである。もしデータベース維持のために労力を厭わない人間がいても、もし当該データベースの周囲にその活用に意欲的な人物がいなければ、休止するか忘れ去られるか、いずれかの事態に陥ることになるからである。そのような事態に陥るデータベースはどのような方法で休眠させて置くべきなのか、そしてまた時機が訪れたときにいかにして迅速それらを活用できるようにする体制にするべきなのか、問題はさらに根の深い問題に行き着くことになる。これについては本稿の最後にもう一度触れようと思う。

3. 継続活用を目指すデータベース周辺の技術変化

以上は、著者が実際に関わった、ないし関わっている機関におけるデータベースの具体的な活用事例である。

数少ない事例ではあるものの、その底流には次のような事実があるように思われる。

データベースというものは、その形態はあまり変わらないにもかかわらず、いつでも進化していく技術との比較によって古く感じられたり、新しく感じられたりするものである。例えば画像を使ったデータベースは、本来であれば画質の向上への対応につねにさらされる。数年前には高精細の最新カメラを使っていたものだったり、撮影方法を工夫して単体のカメラでは実現不可能だった画像を合成し格納していることを特長としたりしていたデータベースがあつという間に陳腐化してしまう例を個人的にも何度か目にしてきた。

また管理運用に関してはデータベースを格納するサーバ技術の変化も激しい。近年の象徴的な技術変化として は VM が挙げられよう。VM の急速な進化に伴い、クラウドという言葉もすっかり定着した。コストと安定性と情報保護とをトレードオフする選択肢が増えてきたといえるだろう。

しかし、安定した管理運用が活用のための基盤作業だということは現在も変わらない。サーバデーモン関連のトレンド、セキュリティ保持、インシデント情報の収集と迅速な対応体制を保持しておくことは、現在もこれからも必要なことであろう。

ただ、サーバ関連の技術が進化ないし変容しても、やはりそこでも相変わらず時限とコストの問題は関わってくる。この問題を解消するために現在は、活用に活路を見出すか技術に活路を見出すかの傾向は今後も続くはずである。

その手法としては、上記3つの個人的に内実を知りうる事例とは別に、最後に挙げた複数のデータベースを効率的に連携させる試み(データマインニングによるデータベース研究など)などが現在では代表的なのかもしれない。この手法は、連携した複数のデータベースを利用することによって、かえって、データベースの必要性を浮かび上がらせ、運用停止や管理保持へのカンフル剤の役割を果たす可能性をもつ。ただ、この手法では一つ二つのデータベースが停止してしまっても仕方がないという前提も含まれるかもしれない。また、例えば「デジタルアーカイブのコンテンツが地域の製品のブランド化と結びつく」といったより実践的な試みによってデータベースの維持を促すといった事例が、総務省が昨年にも出した「知のデジタルアーカイ

ブ社会の知識インフラの拡充に向けて一提及及びデジタルアーカイブの構築・連携のためのガイドライン」[6]に挙げられていたりする。が、しかしこれは学術データベースというデータベース領域には適切な例とはいえないかもしれない。

とまれ、これらもまた、データベースの内容のみを担保として維持管理していくだけではなく、活用方法の工夫こそがデータベースの継続に寄与するというスタンスかと思われる。

4. データベース活用の別の視点

本稿では、異なるスタンスでデータベースを活用する方法を提案してみようと思う。

それは次のような前提から始まる。データベースというものはいずれ開発・構築が終わり、運用も終わるものである。

本稿ではあえて活用しているから止められないというスタンスをとらない。またはデータの内容によっては止めてしまっても構わないというスタンスもとらない。データベースは少なくとも構築当初は、構想者ないし開発者が何かしらの価値を付与して半自動的に情報を関連付けるデータの集合体である。したがって、本来であれば止めてよいデータベースは一つもないはずである。それにも関わらず管理運用の維持はやはり難しい。なぜだろうか。おそらくデジタルデータベースというものはそもそもコストや時限の問題で運用をいずれ止めるを得ないものなのである。

これを前提とした上で、「停止に追い込まれないようにするためには」ではなく、「停止に追い込まれてしまっても」という視点からデジタルデータベースの管理運用ならびに活用を考えることはできないだろうか。

データベースがもし

- 迅速かつ的確に、再生できる機構を備えている

ものであれば、データベースの構築、管理運用、継続活用に関して新たな知見が生まれるのではなかろうか、というのが本稿の趣旨である。

4.1 再生の観点から見たプログラムソースと仕様

仮に停止に追い込まれたデータベースを考えよう。何が残されている可能性が高いだろうか。おそらく、報告書や仕様書の類、データ本体、ソースコードであろう。あるいは当時の開発者や利用者の記憶も「残されたもの」といってよいかもしれない。

データスキーマについては報告書や仕様書が残されていれば、参照できる可能性が高い。停止していても保存媒体が壊れていなければ、形式に問わず、取り出すこともできるだろう。ソースコードはどうであろうか。これもデータスキーマと同様、保存媒体が読み込み可能ならば取り出すことは容易だ。だが、問題はソースコードが読解できるかどうかだ。マイナーな言語あるいは再生を夢見る者が読めない言語である可能性は大いに考えられる。こうしてみるとソースコードに関しては、取り出すことはそれほど難しいことではないにせよ、再生の観点からみると、優れた報告書や仕様書には価値の点で及ばないと本稿は考える。

ソースコードより報告書や仕様書の方が価値の点上である、という主張はプログラマからすると少し違和感を感じるかも知れない。しかし、例えば E.レイモンド『伽藍とバザール』[7]中、ソフトウェア開発の教訓その9.として全く同じことが言われている。

9: 賢いデータ構造と間抜けなコードのほうが、その逆よりずっとまし。

またもやフレッド・ブルックス本の第 11 章から。「コードだけ見せてくれてデータ構造は見せてもらえなかったら、わたしはわけがわからぬままだろう。データ構造さえ見せてもらえれば、コードのほうはたぶんいらない。見るまでもなく明らかなだから」

実はその理由は簡単なものである。例えばボタンアクション一つを考えればよい。利用者はボタンを押すことに一つのアクションを割り当てているわけだが、プログラムではそうはいかない。単純なケースを考えても、変数に蓄えられたデータを引数として再度配列化したり、エラーチェックをしたり、出力プログラムに引数としてそれを渡したり、といった内容を書いていることが多い。ひょっとすると、サブルーチン化した関数を使いつつ、考えうる限り最適な結果を、自分なりにエレガントな方法で導き出そうとしているかもしれない。これらをソースコードから逐一読み解くより、開発者が何をしていたのか、それを適切に表現できる手段があれば、その方が効率的で方法としてもスマートなのは明らかである。

データベースの再生の場合でも、どうやら報告書や仕様書に頼るのが良さそうである。

そうするとデータベースを再生には迅速さという観点に加え

- どんなプログラム言語にも対応できる、再生のための機構を備えている

ことが同時に重要なポイントとして挙げられるよう。

そして、可能な限りの確・迅速に、プログラム言語に縛られずデータベースを再生できる仕組みの鍵は報告書あるいは仕様書にあると考えられる。

5. 仕様にかえれ コンセプトマップの提案

単なる成果報告ではなく、再生を予め考慮した報告書ないし仕様書作りとはどのようなものを目指せばよいだろうか。

本稿ではそのために 4 つの構成要素からなるデータベースのコンセプトマップなるものを提案したい。

考えられる 4 つの構成要素とは以下のものである。

1. 状態推移図
2. 画面推移図
3. プログラムフロー図履歴
4. データスキーマ
5. (データ本体)

重要度から並べていったが、これまでの記述経緯も踏まえ、順序を逆に、4,5.から触れていこう。

5.1 データスキーマ

当然ながらこれはデータベースの命ともいべきものである。ここにはデータベース構想者の思想が詰め込まれているといってもよい。繰り返しになるが、最近ではリレーショナル構造をもつソフトを使うことが多いので、このスキーマの出力(例えば sql97 仕様であれば、desc)はいとも簡単に取り出すことができよう。データベース開発者の中にはスキーマを報告書や仕様書に載せることに違和感を感じるという者もいるかもしれないが、データベーススキーマはそれだけでは意味をもたない単なる形式、枠組みである。もちろん、全データが報告書や仕様書、できれば別途電子媒体の中に残されていれば、当初の目的である確・迅速なデータベース再生に役立つだろうが、データスキーマは報告書や仕様書には必須である。プログラムとの関係からこれは外せない。

5.2 プログラムフロー図とその履歴

次に 3.プログラムフロー図の履歴を取り上げよう。

プログラムフロー図というのは、非難されることも多いが、それにもいくつかの記述方法があると思う(例えば

古典的なダイアグラムによるフロー図などが、次のようなものである。

プログラムソースはたいいていブロックごとに記述されている。ここでは、ソースに直接あるいは別レイヤで視覚的にブロックを書き入れて、ブロック同士の相関を描画したものをプログラムフロー図として想定している(実際に市販されている解説書でこの手の方法をとるケースも多い)。

さきほどプログラムソースコードよりも報告書・仕様書を、と述べたが、次に述べる状態偏移図、画面偏移図の補助手段として、歴代開発者、メンテナの思想を追尾するために、プログラムフロー図をその歴史とともに残していくのはデータベース再生にとって強力な補助手段となりうる。

プログラムフロー図とデータベーススキーマは、開発者、代々の管理運用者ごとにログとして残しておいた方がよい。プログラムであればこれが changelog として記録されるところだが、あえて冗長性を増やしてより視覚的に保存される方がよいだろう。これらが重層的に記録されていることは、停止から再生までの期間が長ければ長いほど、役立つ可能性も生まれてくると考えられる。とりわけプログラムフロー図の履歴は種類の異なる言語で書かれた場合にはデータベース自体の資産になりうる。

5.3 画面偏移図と状態偏移図

最後にこのコンセプトマップ、報告書や仕様書にとって欠いてはならない 1. 状態偏移図と 2. 画面偏移図を取り上げる。先にボタンの例を述べたが、状態偏移図も画面偏移図もユーザアクションとプログラムアクションの相関図である。状態偏移図は開発者目線の、画面偏移図は利用者目線のそれといってもよいかもしれない。とまれ、両者はいずれもいわばデータベースの機能を凝縮し、簡略化したものといえる。

まずは画面偏移図である。

これは高い確率でデータベースのマニュアルに付されることが多い。それゆえ、おそらく構築後に残される可能性も高い。ただ、画面偏移図を作成するとき、利用者目線と同時にメンテナ目線も考慮することが必要かもしれない。簡易化することは効果があるが、メンテナにとって必要な情報が落とされていては何の役に立たないものになりかねないからである。報告書などを書く場合に、この部分に積極的に開発プログラマは関わるべきである。

次に状態偏移図である。

これも簡単なものだけに大規模な報告書には開発者が付す場合も多い。例えば図7.は拠点アーカイブ

PSS のための状態偏移図であるが、この図だけでは例えば Pompei/Rome が何を指しているのかがわからないが、もしこれが効果的に画面偏移図と連動していればその意味は自ずとわかるであろう。

さらに、これらに加えてデータスキーマがあれば、もっと具体的に多くのプログラマにコンテンツ再生を可能にするものと思われる。

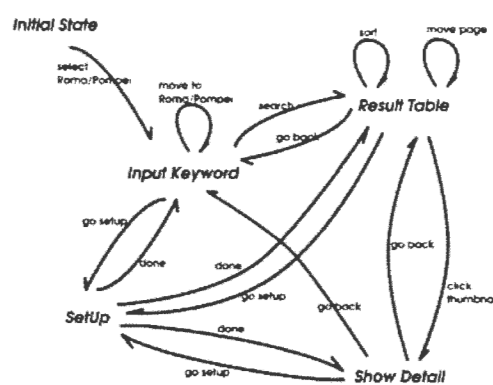


図7. 状態偏移図の例

重要なのは、ここに特定の言語の記法に関する知識は可能な限り排除されている点である。

図7.にも “select” といった一部 sql 文を想起させる部分があるが、それでもここで使われている機能は “move to”、“sort”、“search”、“go to”、“go back”、“done”、“click” などごく単純な用語で示されている。これらの機能の実装はそのつどのプログラマに委ねられている。

どのように実装してもよい、というスタンスでプログラムの記録が残されていることは、データベースを再生させようとする者にとって、時として自身のプログラム技術以上の効果をもたらす。すなわち、迅速なプロトタイプの実現である。そして、この者は実装が固まった時点でもういちど自分が手本とした状態偏移図を眺め、追加・変更した機能について修正を施す。

こうすることでデータベースは単に機能するもの以上に、「発掘可能な所産」となるのではないだろうか。

5.4 コンセプトマップ

以上 1. から 4. の要素を効果的に関連付けたものを本稿では仮にコンセプトマップ(参照: 図8)と呼ぶことにする。

仮に、といったのはこれがまだ構想段階であるからで、実際に運用中ならびに停止してしまったデータベースに果たして本稿のアイデアを実際に検分する必要があるからである。

ただ構造がシンプル極まりないだけに検分手順も大
方見通しがついている。

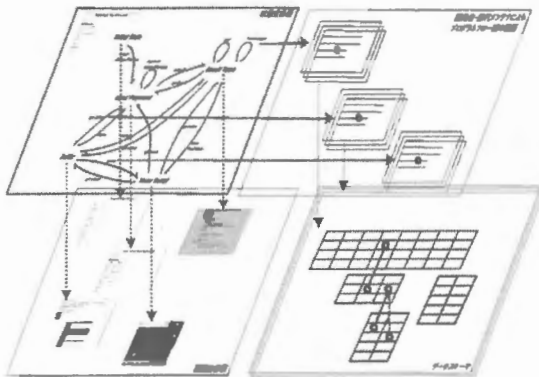


図8. コンセプトマップ(案)

1. 「運用中の」小規模なデータベースを素材にコンセ
プトマップを開発者に依頼。マップ作成にはどれ
ほどの手間がかかるかを実測。改善点を検討。
2. 1.を別言語で再開発。構築時間の実測。改善点の
検討。
3. 1.を今度は「停止している」が当時の開発者と連絡
がとれる小規模なデータベースで実施。
4. 3.を 2.を踏襲して実施。元開発者に検証を依頼。
5. 1.から 4.の成果を整理し、「開発者とのコンタクトも
なく、運用も停止している」データベースにて再生
を実施、検証。

データベースの構造は今後も進化していくべきであ
る。コンセプトマップは、データベースを再生しようと目
論む者のアイデアを、できるだけ迅速に、そして実行
に移す際の心的負担を少なくして、反映させるという
目論見を元に構想されたものである。

6. コンセプトマップから、データベースのパザ ール型形成の可能性を探る

実はこのコンセプトマップ構想にはもうひとつの可能
性を込めている。それはデータベースの活用形態の
新たなオプションとしてのパザール型コミュニティ形成
である。

コンセプトマップ構想の実現には、運用を停止した
データベースを対象にしているにも関わらず、多くの
手間と時間がかかるだろう。

だが、このこと自体がデータベースを「活用」している
ことにはならないだろうか。

コンセプトマップはデータの内容を含むものではない。
よって、著作権等には抵触する確率が低い。これを特
定のデータベース固有の資料として保存することには

ほとんど意味がない。それでは、もしこれをデータベ
ースの構造に関心をもつ研究者が共有したらどうだろ
うか。共有化の結果、構造の類型化がなされ、それを参
考に管理運用中のデータベースの改善点や機能追加
のヒントにはならないだろうか。あるいは停止してしま
ったデータベースの「発掘」の重要な緒にはならないだ
ろうか。

従来のデータベースの活用のプレーヤといえば、無
自覚にその利用者を指してきた。換言すれば、データ
ベースの活用方法というのは、データベースの利用法
と同義であった。だが、このコンセプトマップ構想は、
開発者、管理運用者、将来のデータベース開発者を
巻き込む。しかも、将来のデータベース開発者に、複
数のコンセプトマップを提示できれば、より効率的もし
くはより斬新なアイデアの種を提供できる可能性すら
あると考えられる。

繰り返しになるが、コンセプトマップ構想ないしコンセ
プトマップ仕様の策定に見込んでおいてよいことは

- データベース活用のフィールドに新たなプレーヤ
が自然と加わる

ことである。

というのも、この手法にある程度の見通しが立てられ
た暁には、先に挙げた E.レイモンドが掲げた「パザ
ール型」開発と同じ環境をデータベース開発者・再生者
が得るからである。

この開発形態の優位性はすでに広く認められている。
その特長として頻繁な更新回数と対象(ソフトウェア)
の譲渡のスムーズさが挙げられる。

運用停止の危機に陥っている、ないしはすでに停止
してしまったデータベースにとって、さらには今後の停
止の危機に備えるためにも、データベースのコンセプ
トマップ公開、収集はマイルストーンになりうる。

なるほどデータベースを最初に構築する際は、確か
にデータに関する監修者が必要で、そのニーズに応
えることが必須条件であるが、その後はプログラマがプ
ログラムを共同で開発する形態にそれを持ち込むこと
ができるのである。

データベースを活用することと、必要に応じて再生さ
せることは、今後のデータベース開発、管理運用に多
少なりとも貢献できるアイデアではないかと本稿は期
待するものである。

参考文献

- [1] 東京大学大学院人文社会学系研究科象形文化研究拠点, 「文部科学省中核的研究拠点 COE 形成プログラム報告書象形文化の継承と創成に関する研究」, 2004.
- [2] 神内俊郎・浜田長晴・池庄司伸夫, 「デジタルイメージシステム (DIS) の開発とその応用」, 日立評論 デジタルアーカイブ, 1997.
- [3] 国立西洋美術館, 「古代ローマ帝国の遺産 栄光の都ローマと悲劇の街ポンペイ」, 2009
- [4] 園田茂人編, 『勃興する東アジアの中産階級: アジア比較社会研究のフロンティア』勁草書房, 勁草書房, 2012
- [5] 園田茂人編, 『リスクの中の東アジア: アジア比較社会研究のフロンティア』勁草書房, 2012
- [6] 総務省, 「知のデジタルアーカイブー社会の知識インフラの拡充に向けてー提言及びデジタルアーカイブの構築・連携のためのガイドライン」, <http://current.ndl.go.jp/node/20515>, 2012
- [7] Eric Steven Raymond 著, 山形浩生訳, 『伽藍とバザール』, 光芒社, 1999

主催：第19回公開シンポジウム実行委員会

人文系データベース協議会

共催：立命館大学アート・リサーチセンター
立命館大学文学研究科文化情報学専修

議長：出田和久（奈良女子大学）
副議長：加藤常員（大阪電気通信大学）
副議長：川口 洋（帝塚山大学）

後援：人文系データベース協議会

第19回公開シンポジウム実行委員会

委員長：赤間 亮（立命館大学）

委員：高橋晴子（大阪樟蔭女子大学）

丸川雄三（国立民族学博物館）

後藤 真（花園大学）

研谷紀夫（関西大学）

河角龍典（立命館大学）

赤間 亮（立命館大学）
江澤義典（関西大学）
及川昭文（総合研究大学院大学）
桶谷猪久夫（大阪国際大学）
小沢一雅（大阪電気通信大学）
金田明大（奈良文化財研究所）
小島篤博（大阪府立大学）
五島敏芳（京都大学）
後藤 真（花園大学）
小森政嗣（大阪電気通信大学）
柴山 守（京都大学）
関野 樹（総合地球環境学研究所）
高橋晴子（大阪樟蔭女子大学）
田窪直規（近畿大学）
中谷広正（静岡大学）
八村広三郎（立命館大学）
原正一郎（京都大学）
深海 悟（大阪工業大学）
寶珍輝尚（京都工芸繊維大学）
三宅真紀（大阪大学）
村上征勝（同志社大学）
森下淳也（神戸大学）
師 茂樹（花園大学）
山元啓史（東京工業大学）
米澤 剛（大阪市立大学）

人文系データベース協議会 第19回公開シンポジウム「人文科学とデータベース」

発行日 2013年11月30日

発行所 第19回公開シンポジウム実行委員会

〒603-8577 京都市北区等持院北町 56-1

立命館大学アート・リサーチセンター

赤間 亮（シンポジウム事務局） Email:rat03102@lt.ritsumeai.ac.jp