

歴史資料自在閲覧システムによる大規模資料群の画像閲覧方法の検討

鈴木卓治 安達文夫
国立歴史民俗博物館

国立歴史民俗博物館では、博物館資料の超精細画像をデジタル化し、歴史資料自在閲覧システムと組み合わせた「超精細デジタル資料」を研究や展示に利用している。古文書や錦絵のような大量の資料画像群を配列して超精細デジタル資料を作る場合、作成のための計算時間や作業記憶場所などが問題となる。本稿では、個々の画像について単体のデジタル資料を作成しておく、それを配列した資料群画像の超精細デジタル資料が高速に再構成できることを示す。

超精細デジタル資料データベースの可能性を検証するため、本館の近世小袖資料(約230点)および錦絵資料(約2600点)を例題として、超精細デジタル資料群を自動的に配列し表示するシステムを試作したので合わせて報告する。

Investigation of image browsing method of a mass of materials by the super high definition free viewing system

SUZUKI, Takuzi and ADACHI, Fumio
National Museum of Japanese History

1. はじめに

国立歴史民俗博物館では、博物館資料の超精細画像をデジタル化し、資料のどこでも望みの倍率で表示できる歴史資料自在閲覧システム[1]と組み合わせた「超精細デジタル資料」を作成し、研究や展示に利用している[2]。

超精細デジタル資料は、屏風や絵巻のような資料1点単位の作成が有効であることはもちろん、古文書や錦絵のような大量の資料群についても、適切に配列した画像を作成してやれば、画像情報の提供方法としてたいへん有効であろう[3,4]。先行研究として日文研の森洋久氏らによる「画像一覧表示検索システム」の例[5]をあげることができる。歴史資料自在閲覧システムの適用により、サムネイル画像によって画像群全体を把握する視点と、拡大画像によって各資料を熟覧する視点の両方を、視点移動と倍率変更という単純な操作によって切れ目なく提示することができるので、「画面上に画像をたくさん並べておき、目視により画像を探す」という、素朴ではあるが人間にとって自然な画像検索方法を実用的な性能で提供することができる。

大量の資料画像群から超精細デジタル資料を作る場合、1点単位の作成とは異なり、作成すべき資料の総数が確定できない場合(たとえば毎年計画的に購入する資料など)や、資料の画面上の配列が確定しない場合(たとえば資料の整理状況を逐一画像の配列に反映させたい場合など)が多々発生する。状況が変化するたびに、もとの画像群から超精細デジタル資料を作っていたのでは、作成に要する時間や作業用記憶領域などのコストが莫大なものになる。

本稿では、個々の画像について単体のデジタル資料を作成しておく、それを配列した資料画像の超精細デジタル資料群が高速に再構成できることを示す。

これはすなわち、デジタル資料のデータベースを構築し、利用者の望むように配列した資料群画像をただちに超精細デジタル資料として表示することができることを示している。

超精細デジタル資料データベースの可能性を検証するため、本館の近世小袖資料(約230点)および錦絵資料(約2600点)を例題として、超精細デジタル資料群を作成し、これらの資料番号の列を与えて、画像を自動的に配列し歴史資料自在閲覧システムを用いて表示するシステムを試作した。その結果についても合わせて報告する。

2. 超精細デジタル資料

たとえば、博物館資料の撮影写真(フィルムサイズは4×5ないし8×10を想定)から、フィルムの性能限界に近い解像度で画像をスキャニングしてデジタル画像を作成した場合、そのサイズはしばしば10000×10000画素を超える超精細画像となる。このような画像は、現在の一般的な画像表示ツールでは表示すら困難であり、プログラムの扱える上限を超えてしまうか、あるいは表示に著しく時間がかかってしまう。

歴史資料自在閲覧システムは、画像の表示場所と解像度が指定されると、その解像度より高解像度でかつもっとも近い解像度の画像の中から、表示場所を含む画像のファイルだけを読み込んで、画面に表示するシステムである。画像の呼び出しに必要な記憶容量は、画面の大きさに比例する量ですむ。また、呼び出しにかかる時間は、表示画像を同定するための計算時間と、ファイルの読み込みの時間との和となり、表示場所と解像度によらず一定となる。これらの工夫によって、たとえ100000×100000画素クラスの超精細画像であっても、一般的なパーソナルコンピュータ上での実時間(1秒あたり数コマ~10コマ以上)の表示更新を可能

としている。

自在閲覧システムが扱うのは、階層化デジタル画像と呼んでいる構造の画像ファイル群である。これは、元画像とそこから作成される解像度の異なる複数の縮小画像、および画像間の関係を表わす管理情報によって構成される。

本稿では「デジタル資料」という言葉を、階層化デジタル画像と歴史資料自在閲覧システムの組み合わせを表わす語として用いる。また、超精細画像から作成したデジタル画像をとくに「超精細デジタル資料」と呼ぶことにする。

階層化デジタル画像のもっとも素朴な構成法は、最高解像度の画像から、 $1/a$ 倍（たとえば $a=2$ ）の縮小

画像を順に作り、これをさらに適当な大きさ（たとえば 384×384 ）に分割して、別々のファイルに分けて JPEG 形式の画像として格納しておくことである（図1）。各ファイルは、倍率および画像内の位置とサイズを属性にもつ。なるべく表示に必要な画像データだけを読み込みたい、という観点からは、分割画像のサイズは小さいほうがよいが、一方分割画像の個数は急速に増加し、表示画像同定のコストは上昇する。また、ファイルの読み込みは、読み込む総容量が同じならば、ファイルの個数の多いほうがより読み込みの時間がかかるため、分割画像のサイズをある程度大きく（たとえば 512×512 など）して画像数を減らす方が、表示更新速度の低下を防ぐ場合がある。

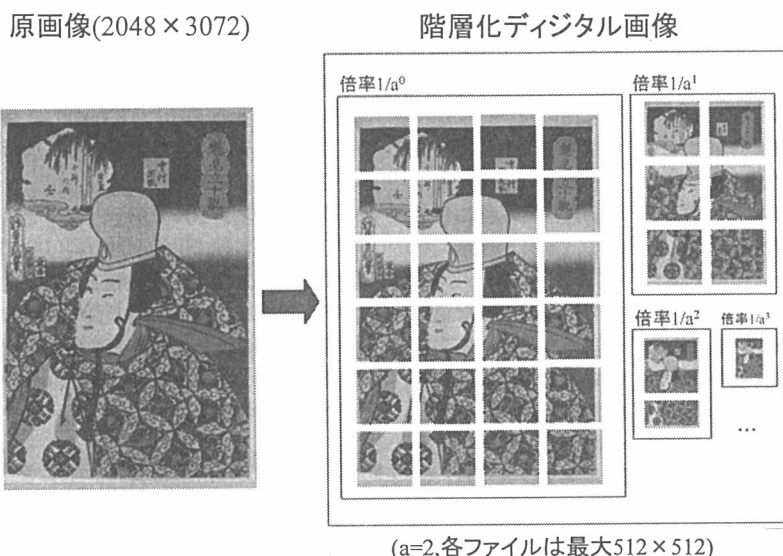


図1 階層化デジタル画像

3. 階層化デジタル画像群の動的な再構成

いま、ほぼ同一の条件で撮影した資料写真から同一解像度でスキャンして作成した n 個の階層化デジタル画像群があるとする。個々の階層化デジタル画像に含まれる画像の縮小率はすべて等しいものとする。このとき、この n 個の階層化デジタル画像を配列して再構成することは、個々の分割画像の表示位置を再構成することにほかならない。すなわち、図2において、階層化デジタル画像 i の中のある画像 k が、 i 内で (x_k, y_k) の位置を左上隅とするように配置されているとき、 i の左上隅が (x_i, y_i) となるように再構成された場合、 k の位置は再構成画像全体の中では $(x_i + x_k,$

$y_i + y_k)$ となる。(画像単独では、左上隅は常に $(0,0)$ として扱っている。) i の画像サイズを (w_i, h_i) とすれば、再構成された画像の表示において、表示範囲が $(x_i, y_i) - (x_i + w_i - 1, y_i + h_i - 1)$ の範囲にあるときだけ、 i にアクセスすればよい。したがって、表示範囲が指定されたときに、該当する画像ファイルを特定するための手間は、階層化デジタル画像のサイズが画面サイズに比べて充分大きいときは、たかだか4つの階層化デジタル画像を調べればよい。縮小率の高い画像については、4つ以上の階層化デジタル画像を調べることになるが、このとき各階層化デジタル画像において表示すべき分割画像はただ1枚となっているので、表示位置はただちに確定する。

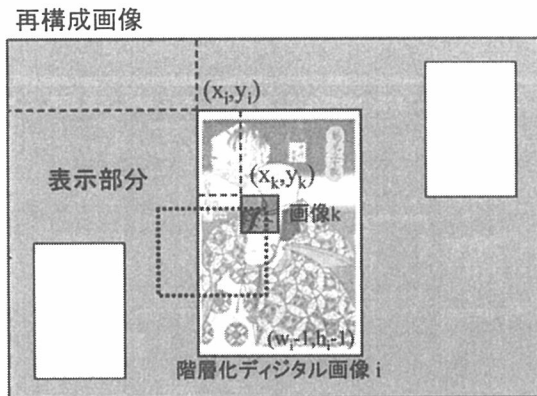


図2 階層化デジタル画像の配置と座標の関係

4. 超精細デジタル資料群閲覧システムの試作

超精細デジタル資料群閲覧システムの試作を行なった。このシステムは、Microsoft Excel のマクロとして実現されており、資料ごとに作られたデジタル資料の一意名のリストを与えると、資料画像を自動的に配列した「超精細デジタル資料群」の再構成を実施し、歴史資料自在閲覧システムを起動して、画像を表示するものである。図3に、超精細デジタル資料群閲覧システムの画面例を示す。

歴史資料自在閲覧システムの現在の実装は、ただひとつの階層化デジタル画像を表示するシステムとなっており、表示画像を同定するための情報として、階層化デジタル画像を構成する分割画像ファイルの属性値リストを入力として受け取る。(属性値は、画像ファイルのパス名、縮小率、位置、および大きさの情報を含む。)そこで、超精細デジタル資料群閲覧システムは、個々の資料における属性値リストを読み込み、画像配置情報の再計算を行なって新たな属性値リストを作成し、歴史資料自在閲覧システムに渡す。すなわち自在閲覧システムは、ひとつに合成された階層化デジタル画像を表示していることになる。

例題データとして、本館の近世小袖資料(約230点)および錦絵資料(約2600点)の画像をとりあげた。これらはいずれも、銀塩写真(リバーサルフィルム)により撮影された画像から作成された ProPhotoCD の BASE16 画像(2048×3072)を元画像としている。図4に、各資料群の全体を表示した例を示す。

いまのところ歴史資料自在閲覧システムには、ただひとつの属性値リストしか渡せないで、資料数が増えるにつれ、属性値リストの再作成の手間が無視できなくなる。また、属性値リストが非常に大きくなるので、表示に必要な画像を特定するためのコストが動作速度に対して無視できなくなってくる。この問題は、

歴史資料自在閲覧システムが複数の階層化デジタル画像を認識できるようになれば、前節で述べたように、ただか4つの属性値リストを検索すればよいため、容易に解決することができる。

また、図4のような資料群全体の表示において、現在の実装では、個々の資料のごく小さい画像を、表示されている資料の数だけ読み込んで表示しているが、このときファイルの読み込み起因する表示更新の遅延が無視できない。この問題は、表示画面に比してサイズの小さい分割画像について先読みしてキャッシングしておくとか、作成した資料配列を何度も繰り返し利用するならば、合成画像を作って読み込み枚数を減らす、などの方策によって解決できる。

5. 配列された錦絵群の目視による検索

錦絵研究の専門家(本館教員)に、錦絵画像を配列した画面を見せ、目視による検索が可能かどうかを尋ねたところ、17インチCRT、解像度1024×768

(うち画像表示部分は1022×556)の場合、およそ1000枚(縦23枚×横45枚)程度であれば、画像のおおよその種類を判別できるとの回答を得た。また、錦絵資料の超精細デジタル資料群を、本館において10月21日から10月30日まで開催された特別企画「歴史を語るサイエンス」において、他の超精細デジタル資料と合わせて展示に供し、来場者に、指定した錦絵を探してもらうことを試みたところ、(本稿執筆時点で企画展示開催中のため確実とはいえないが、)縦10枚(全体では約200枚)もしくは縦5枚(全体では約50枚)の表示倍率が多く選択されているようであり、利用者がこのぐらいの枚数を適当と感じて画像を探している様子が伺われた。専門家および非専門家がどのように目当ての画像を探し当てるかについて、引き続き検討をすすめたい。

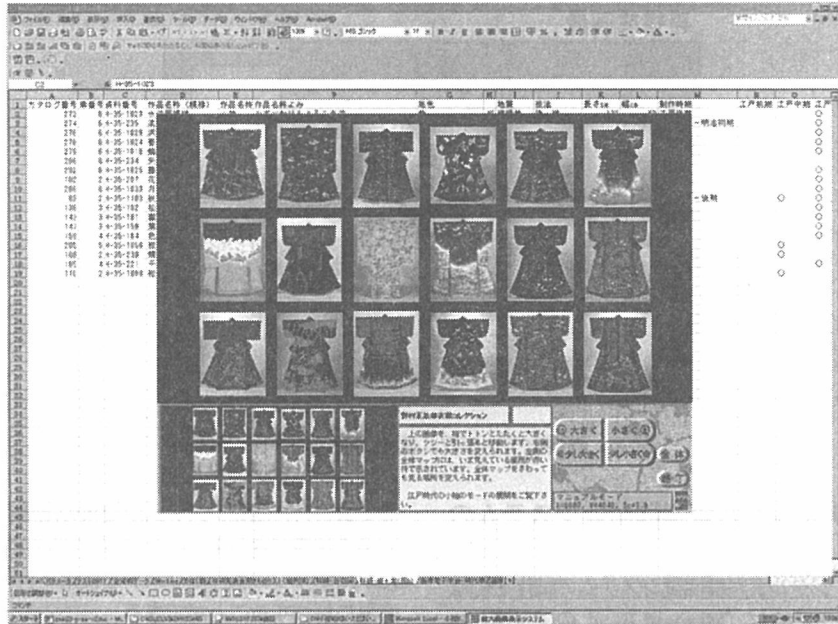
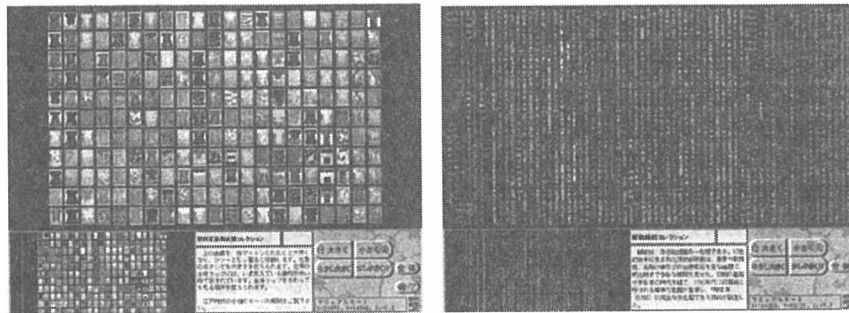


図3 超精細デジタル資料群閲覧システムの画面



(a) 近世小袖資料

(b) 錦絵資料

図4 超精細デジタル資料群の表示

6. おわりに

画像データベースにおいて、多量の画像群から望みの画像を探すためには、画像に付与された属性情報(テキストや数値)を用いるのが一般的である。これに対し、歴史資料自在閲覧システムを用いて資料群の画像から目視により探すという方法は、画像の情報が未整理もしくは未確定である場合でも画像情報の提供が可能になるという点で有効であろうと考えられる。引き続き検討を続けたい。

なお、この研究は一部、文部科学省科学研究費補助金特定領域研究(課題番号 14023229)の補助を受けている。

参考文献

- [1] 鈴木卓治, 安達文夫: 歴史研究・展示用画像表示システムの機能に関する検討, じんもんこん 2001 論文集, pp.229-234(2001).
- [2] SUZUKI, T: Super-high-definition Free Viewing Digital Materials for Historical Research, Proceedings of the Sixth REKIHAKU International Symposium, pp.73-80(2003).
- [3] 安達文夫, 鈴木卓治, 宮田公佳: 歴史資料自在閲覧システムの検索画像表示への適用の検討, 画像電子学会第31回年次大会予稿集, pp.29-30(2003).
- [4] 安達文夫, 鈴木卓治, 宮田公佳: コレクション資料閲覧への歴史資料自在閲覧システムの適用, ADADA2003 概要集, pp.88-89(2003).
- [5] 山田奨治: 国際日本文化研究センターにおけるデジタル画像の利用, Science of Humanity Bensei Vol.39, pp.105-109(2002).