

ビデオによる古文書の効率的画像入力法

柴山 守

(大阪国際大学経営情報学部)

星野 聡

(京都大学大型計算機センター)

ビデオ画像による古文書・古絵図の入力とワークステーション (WS) 上での入力画像の接合処理、JPEG標準による画像圧縮/伸張について述べる。分割入力されたビデオ画像は、WS上において、古文書画像の濃度のヒストグラム、古絵図に対するHough変換を用いた特徴抽出によって、隣接画像のマッチング時間が短縮でき、原画像が復元できる。ビデオ入力の有効性と共に特徴と問題点について指摘している。また、JPEG圧縮/伸張による画像データの効率的保存のために、パソコン用JPEGコードをWS上に移植し、ビデオ画像(640\*480画素)に対する圧縮率や所要時間について測定した。

An Effective Input of Ancient Documents by the Video Image

Mamoru SHIBAYAMA

Satoshi HOSHINO

Faculty of Management and  
Information Science,  
Osaka International University

Data Processing Center,  
Kyoto University

3-50-1 Sugi, Hirakata,  
Osaka 573-01, Japan

Yoshida Honmachi, Sakyo,  
Kyoto 606, Japan

An image input method for ancient documents and maps using the video camera is described. The mosaic operation and JPEG compression/decompression are utilized. The feature extractions based on the histogram for the document and the Hough transformation for the map in the mosaic operation are effective to reduce the necessary time for matching two neighboring images. The characteristics of the image input by the video camera are indicated. For the effective storage of images, a JPEG code for personal computers was implemented on the NEWS work-station. The efficiency of compression and its cpu time are measured.

## 【1】はじめに

マルチメディア情報処理に関する研究の進展と共に、近年動画像（ビデオ画像）が扱えるハードウェアやソフトウェアが数多く出現してきた。こうしたデジタル画像関係機器の技術革新と多様な装置の出現は、歴史・文学などの研究分野においても、研究支援ツールとして大いに役立つものと期待される。従来、古文書などの画像データは計算機処理において、主にイメージスキャナー装置により入力されている。しかし、大きな面積のものや厚みのある書物について入力しにくく、また必要な古文書が保管・展示されている場所まで入力装置を運ぶことは難しい。ビデオ機器の利用により、古文書を直接デジタル画像として計算機入力できれば、ハンディタイプのビデオカメラなどが入力装置として利用可能になり、入力作業は大いに軽減される。しかし、これらの入力法実現のために、入力された部分画像に対する一般的な画像の変換や画質改善の処理に加えて、特に

- (1) 画像データの接続・接合処理（原画像復元）、即ちモザイク処理、
- (2) 画像データの圧縮・復元・伸張、即ち効率的保存（記憶容量の節減）

が必要となる。

画像データの接続・接合処理は、隣接する画像の重複部分の画像を利用して、位置合わせと色合わせを行った後、隣接する画像を接合するが、(1) 位置合わせ及び座標系の統一、(2) 濃度補正、(3) 接合点探索、(4) 接合点周辺の平滑化などの処理が必要とされる [1]。空中写真のモザイク例においては、幾何変換と濃度変換の手法が報告されている [2]。画像圧縮・伸張については、静止画を対象にした J P E G (Joint Photographic Experts Group) 標準がカノープス電子 (株) 製ボード [3] や Mac 用の QuickTime、S U N における XVideo [4] で実現され、1/20～1/100 の圧縮が可能とされている。

本報告では、ディスプレイ上において古文書・古絵図の画像データが充分判読できる解像度と画質で扱えるよう分割して画像入力を行い、且つ、巻紙のように横に長い、掛け軸のように縦に長い、あるいは縦横共に長い古文書・古絵図をビデオ画像として入力するために、分割入力された古文書・古絵図の自動的な接続処理を考える。接続処理においては、入力画像から得られた 2 値化画像データを扱い、特に重複部分の特徴抽出やマッチング・位置合わせにおける実験と結果、及び、入力機器としてスチル・カメラ、ビデオ・カメラを使用した例につき、特徴や問題点について述べる。また、画像データベース作成上の必要から、J P E G 標準を実現したプログラム [6] のワークステーション NEWS-3720 (ビデオマップ・ボード NWB-254 装備) への移植を行ったが、それに関して実験結果を述べる。

## 【2】古文書・古絵図画像の接続処理

ビデオ・カメラあるいはスチル・カメラなどを入力装置として古文書・古絵図の入力を自動的に行うには、入力後に画像データの接合・接続処理が必要となる。今回、入力した古文書は、茨城県土浦における江戸期（天保年間）の古文書（筑波大学岩崎宏之教授提供）である。

図 1 は、原文書（一部分）である。一度に入力すると解像度不足でディスプレイ上で判読できないため、1/4 ずつ入力した後、接続処理を施すこととした。各々の 1/4 画像データの内、右上、右下、左上、左下の順に 4 つの画像データについて図 2 (1)、(2)、(3)、(4) に示す。画像データはすべて 2 値化されたもので、古文書入力にはイメージスキャナー（E P S O N 製 G T 6 0 0 0 解像度最大 600dpi）を使用した。その後、スチル・カメラによる入力も行なった。また、古絵図入力には、ビデオ・カメラを用いた。

これらの古文書画像データの接合処理のためのマッチング手法として、隣接画像の重複部分内におけ



このピラミッド構造の画像に対して、(1)のSSDA法を用いたマッチングが効果的だと考えられる。しかし、この方法で解像度を落としすぎると誤探索の可能性がある。誤探索を防ぐために、また対象画像とテンプレート画像のサイズの大きい場合に演算時間を短縮する特徴量空間とピラミッド構造とを用いた画像の重ね合わせ法が知られている[7]。

### [3] 画像特徴レベルでのマッチング

線や領域、角といった画像特徴に基づいた照合は、(a)モデルを画像空間上に投影し、対象画像から抽出された特徴とのマッチングや対象物の認識、(b)モデルと対象画像に含まれる線や領域を抽出し、これら両者に含まれる特徴間の類似性を調べ、特徴レベルでのマッチング、(c)画像特徴の空間的関係を利用したマッチングで、線分の交差・分岐などの接続関係、あるいは近接しているという距離関係を用いた照合等が考えられる。

以上のような諸手法から、古文書・古絵図の画像マッチングを考える場合、(1)及び(2)の重ね合わせ法と、古文書における特徴抽出において(b)を用いること、古絵図では(b)、(c)が有効であると予測される。

### [3] 接合処理の実現と実験結果

マッチング法を実現するに際して、入力されたマッチング対象となる古文書画像は、対象画像及びテンプレート画像共にほぼ同一縮尺率であること、方向・角度が同一であると仮定する。

マッチングはテンプレート・マッチングの内、SSDA法を用いることとし、特徴抽出を行わない場合のテンプレート画像と対象画像における探索範囲は、つぎのとおりである。文書の上下部分接続において、テンプレート画像Tは入力画像の最上部200(W)\*50(H)画素、探索範囲は基準画像の下端部 400(W)\*160(H)画素とした。文書の左右部分接続において、テンプレート画像は入力画像右端部50(W)\*100(H)画素、探索範囲は基準画像左端部 300(W)\*200(H)画素とした。

#### [1] 古文書の特徴抽出

文書の左右部分画像の接続には、縦書きの特徴から形状の類似性によるマッチングを行うため、特徴抽出の機能を付加した。これを図3に示す。この結果から、図4に示すように、テンプレート画像は、入力画像の最右端から列として独立した部分  $u \times v$ 画素である。また、基準画像内の探索範囲は、画面の左端から探索して最初に見つかった列である。探索範囲は動的に決定され、ヒストグラムによって判定される列幅と高さH (200画素)の面積となる。

ここで文書の各列幅を $W_i$ 、列間の空白幅を $S_i$ とすると、ヒストグラムによる特徴を用いない場合、文字列がn列存在するとき、X軸方向の幅Wは

$$W = \sum_{i=1}^n (W_i + S_i) \quad (2)$$

となり、テンプレート画像の探索範囲は

$$\sum_{i=1}^n (W_i + S_i) - (u - 1) \quad (3)$$

となる。ここで、 $u$ はテンプレート画像の列幅であり、 $u \leq W_i$ である。

ここで、ヒストグラムからの特徴をマ

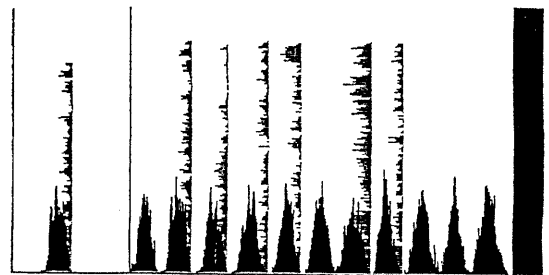


図3 ヒストグラムによる特徴抽出

ッチングに用いると、テンプレート画像の探索範囲は、つぎのようになる。

$$\sum_{i=1}^n [W_i - (u - 1)] \quad (4)$$

この結果、抽出した文書の特徴を利用することにより、X軸方向の探索範囲を

$$\sum_{i=1}^n S_i + \sum_{i=1}^{n-1} (u - 1) \quad (5)$$

減少することができ、マッチングの高速化に有効である。

### [2] 古絵図の特徴抽出

図1 (b) の古絵図の接合処理における画像の重複部分には、地図の線分が含まれる。この線分を抽出し、マッチングに利用すれば有効であると考えられる。この線分の抽出には、Hough変換の適用を試みた [5]。Hough変換は、x-y 空間におけるある直線の式は

$$\rho = x \cos\theta + y \sin\theta$$

と表現でき、この直線が画像上の点 (x0, y0) を通るとすると

$$\rho = x_0 \cos\theta + y_0 \sin\theta$$

の関係が成り立つ。ここで ρ-θ 空間を考え、上式で表される ρ-θ 空間の軌跡は、x-y 空間における点 (x0, y0) を通るすべての直線群を表す。この操作を画像全体に施し、ρ-θ 空間において軌跡が集中している点 (ρ0, θ0) を求めることにより、直線分が抽出される。図1 (b) に示す古絵図では、基準画像の最左端の重複部分において線分抽出を行った。

### [3] 実験結果

上記で実現したプログラムは、ワークステーション N EWS3720 (主記憶8MB) のXウィンドウ上で動作し、以下の実験を行った。作成したプログラムによる接続結果を、図5に示す。

第1に、文書の上下部分の接続処理におけるSSDA法にピラミッド探索法を併用してみる。テンプレート画像は、200×50画素である。実験結果を表1に示す。実験結果が示すように、解像度が粗くなるに従って、マッチング時間が1/4づつ短縮されることが判る。

なお、解像度を1/4以上に粗くすると正しくマッチングされない。ピラミッド探索法において、一般的には、最上層の一致により、最下層まで順次解像度が密になるようマッチングされるが、本実験ではその必要がないことが判った。

第2に、左右部分の接続処理において、文字列の特徴抽出を行うことによるマッチングの所要時間は105msであり、特徴抽出をしない場合の1119msと比較して、約

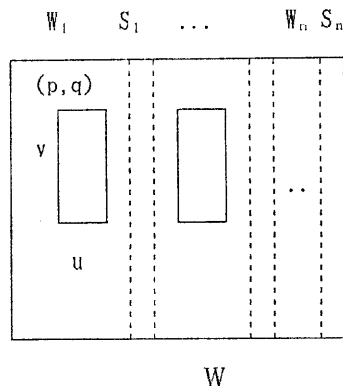


図4 古文書の縦書き特徴

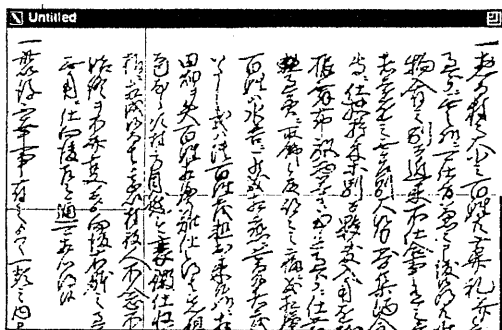


図5 自動接続処理の結果

表1 ピラミッド探索法におけるマッチング時間 (単位: ms)

解像度	1/1	1/2	1/4	1/8	1/16
CPU 時間	4120	1039	274	76	22

1/10に短縮されることが判った(解像度:1/8)。

第3に、図6に示すように、上下部分の接続処理において、テンプレート画像を基準画像(入力済み)におくか、あるいは入力(接続対象)画像におくかの選択が重要である。基準画像との重なり部分が小さい時には、テンプレート画像を入力画像上におき、それに応じた幅のテンプレート画像に動的に変化させることが考えられる。

なお、古絵図の線分抽出による接合処理については、現在実験中である。

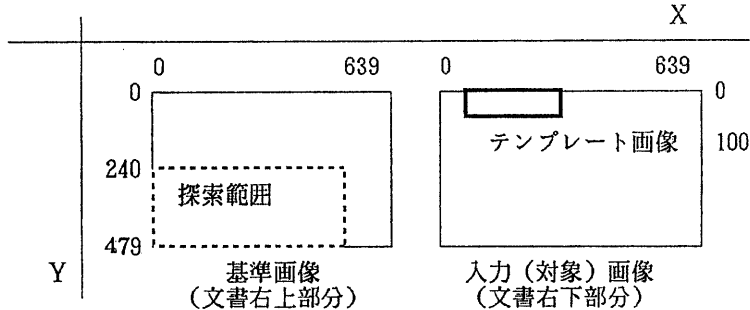


図6 上下部分のテンプレート画像と対象画像

#### 【4】ビデオ画像入力の特徴と問題点

スチル・カメラやビデオ・ムービ・カメラなどのビデオ信号による画像入力は、イメージスキャナー装置などと異なった特徴や多くの問題点がある。古文書や古絵図入力の実験結果や経験を通して得た特徴や問題点についてまとめてみる。画像入力時点における特徴や問題点は

##### 【1】背景色の調整と照明ムラの除去(図8参照)

実際の古文書入力においては撮影条件が異なり、また撮影環境が良くないことが考えられる。ビデオ機器のホワイト・バランス調整による背景色の適切な設定が必要である。入力画像に十分なコントラストを得るためには、照明が必要となるが、強力な照明よりむしろ様な照度が重要である。

##### 【2】ビデオ・カメラの被写体に対する移動速度

撮影時の被写体に対するカメラの移動速度が、入力画像に直接影響を与える。

図7に示すように、1秒間の被写体上での点PQ間の距離 $\delta$ は、 $\delta/p$ の「流れ」現象を起こす(ここで、Oはカメラ位置、pはフレーム数/秒である)。例えば、カメラから被写体までの距離2m、カメラを1秒間に $\theta = 1.8$ 度回転すると、被写体上では6cmの移動距離になり、60フレーム/秒では被写体で1mmの「流れ」現象が起こる。これはビデオカメラの飛越し(インターレース)走査で30コマ/秒(2フレーム/コマ)に起因するものと考えられ、ズームによる撮影で影響される(図1(b)参照)。カメラを低速度で移動・回転させる精神的負担は大きく、入力対象位置ではカメラを静止させることが必要になる。

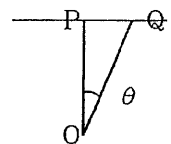


図7 カメラと被写体の位置関係

##### 【3】光学レンズ系の放射量歪みとカメラの入出力特性

航空写真に見られるような画像の中心点に対する点対称の放射量歪み等が考えられる。また、ビデオ・カメラで画像入力を行う場合、撮影管の特性により、入射光量とビデオ出力信号との関係が線形でな

いことを考えねばならない。画質に影響する程度については、今後の課題である（図8参照）。

〔4〕2値化、雑音除去と濃度変換処理

隣接画像のモザイクや特徴抽出においては、2値化処理が必要となる。ビデオ画像は、スキャナ入力と違って一般的に画質が良くなく、複雑な2値化処理、平滑化や雑音除去が要求され、適切な濃度変換を行って画質改善を行わねばならない。図9に、スチル・カメラ入力（図8（a））の2値化画像を示す。

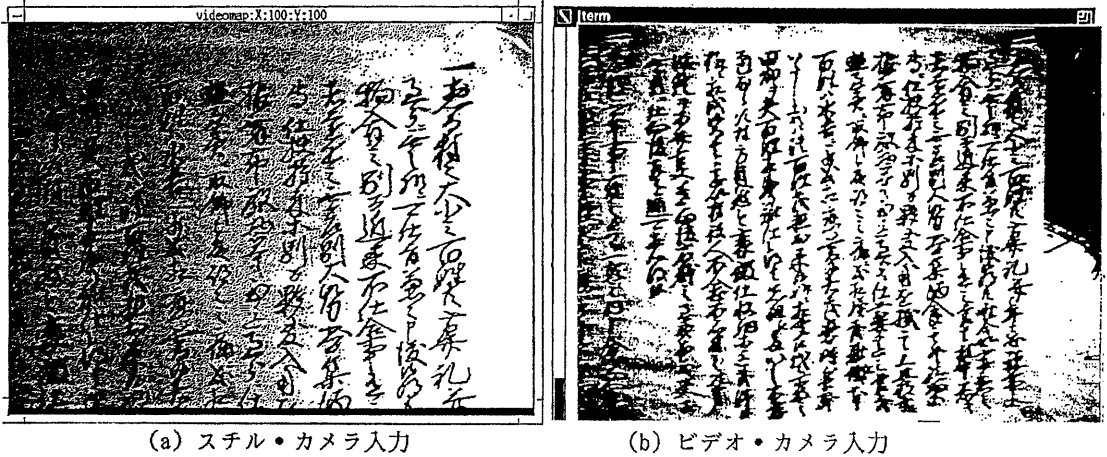


図8 照明効果による入力画像の違い

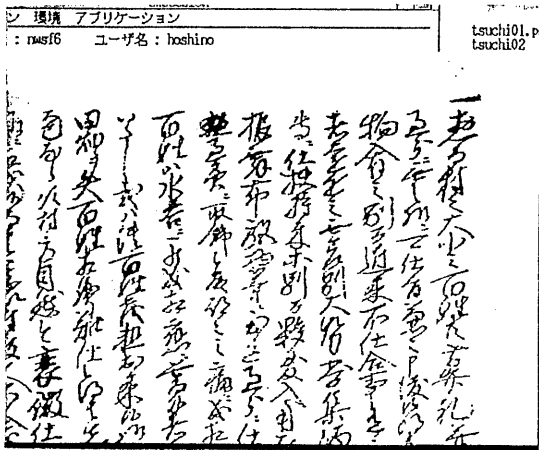


図9 2値化画像（スチル・カメラ入力）

〔5〕J P E G画像圧縮・伸張

画像データ圧縮は、画像データベースを構築する際に必要となり、大容量の画像データとなる古文書・古絵図は、画像圧縮が必要とされる一つの例である。最近、パソコンやWS上で実現されているJ P E G標準の圧縮符号化方式の概要について述べる。

J P E G (Joint Photographic Experts Group) 標準は、カラー静止画像符号化方式の標準化を進めるC C I T TとI S O両研究グループの符号化方式で、1991年8月国際標準化草案となった。

入力画像は、8\*8画素/ブロックに分割され、R G B表示系からY U V (YCrCb)表示系に変換を行いながらU V要素の間引きを行う。この結果から、

離散コサイン変換(D C T)を行い、D C T係数を求め、量子化テーブルを用いて量子化される。つぎに、ジグザグ・スキャンによりハフマン符号化を行って圧縮画像データを得ることになる。伸張処理は、この逆符号化である。

〔1〕J P E Gソフトの移植と利用実験

J P E GソフトのWSへの移植は、1991.12 発表の稲村 [6]によるJ P E G準拠画像データ圧縮/伸張システムのソフトを、ソニーNEWS用ビデオマップボードNWB-254 (R G B各5ビット) 装備機種で動作するようにした。移植に際し、つぎに示す修正を行った。

- (1) 圧縮対象画像の最大 640\*400画素をビデオ画像に対処可能なように 640\*480画素まで拡張、
- (2) ビデオマップボードRGB各32(5ビット)階調を256階調に変換、
- (3) YUV表示系の間引き・変換における繰り返し(for文)操作減少による所要時間の短縮、
- (4) フレーム・バッファ入出力をX-Window上のイメージ処理に書き換える。

移植後、実現したシステムは、XWDフォーマットによる保存ファイルを入力画像データとし、WSに入力と同時にCRTに表示・圧縮を行ってXWDヘッダーを付加して、新たなファイルに保存する。圧縮時に指定可能なパラメータは、間引き率と量子化率である。

本システムによる圧縮実験は、スチル・カメラで撮影し、NEWSのX-Window上で動作するxvmapで保存(640\*480画素、各32階調)したファイルを使用した。実験結果は、間引き率3、量子化率3において640\*480画素の圧縮に要するCPU時間は、18.35秒、圧縮率37.3となった。実験結果は、圧縮に要するCPU時間は間引き率に影響され、圧縮率は量子化率によって決まる。

JPEG標準による圧縮/伸張はハードウェアによる実現が望ましい。ハードウェアによる例として、最近SUN Sparc IPX上にXVideoボードを搭載してのJPEG圧縮において、400\*300画素の書き込みが10~15フレーム/秒、読み出しが60フレーム/秒と報告されている[4]。

## 【6】おわりに

分割入力した古文書の部分画像のマッチング処理において、残差逐次検定(SSDA)法に加えて、ピラミッド探索法を併用したプログラムの実現と実験結果、及び、ビデオ画像のJPEG標準による画像圧縮/伸張システムの移植について示した。実験結果から、古文書の接続処理においてピラミッド探索法が照合時間の短縮に有効であること、文字列の特徴抽出の結果からテンプレート画像の探索範囲を限定はマッチング時間の短縮に有効であること、古絵図におけるHough変換についても述べた。

SSDA法によるマッチング時間の短縮には、特徴抽出と特徴空間の利用が重要である。特に、探索範囲の縮小やテンプレート画像の決定のために用紙の端点や文字の書き出し位置の自動認識が必要である。また、テンプレート画像の大きさの動的変化による接続効率の向上が計れる。これらは今後の課題である。また、ビデオ画像による古文書・古絵図の入力では、輝度の差による歪み補正や接合処理における濃度補正、接合点周辺の平滑化、画質改善が不可欠であり、ビデオ画像入力における特有の問題を解決しなければならない。

本研究は、科学研究費補助金試験研究(B)「東洋学研究における研究者用マルチメディア情報CD-ROMの実用化」の下で進めた。

## 【参考文献】

- [1] 東大出版会：画像解析ハンドブック、東大出版会、pp.461-467
- [2] 河田悦生、森克己：空中写真の接続処理に関する検討、信学会技術研究報告、IE82-11、1982
- [3] カノープス電子(株)：JPEG準拠画像圧縮ライブラリUSER'S GUIDE、1991
- [4] 下條真司：UNIXマルチメディア事始め7、UNIXマガジン、1992.5
- [5] 田村秀行：コンピュータ画像処理入門、総研出版、pp.149-153、1985
- [6] 稲村 浩：Cで実現するJPEG準拠画像圧縮/伸張システム、インターフェース、CQ出版、pp.183-203、1991.12
- [7] 曾根光男、寺田聡、板内正夫：特徴量空間とピラミッド構造とを用いた画像の高速重ね合せ法、電子情報通信学会誌D Vol.J71-D, No.1, pp.102-109, 1988.1
- [8] 遠藤俊明：カラー静止画像の国際標準符号化方式-JPEGアルゴリズム-、[6]に同じ