

イメージ情報のデータベース化に伴う 数値化処理手法に関する研究

東京理科大学 正 員 大林 成行
東京理科大学 正 員 平野 曉彦
東京理科大学 学生員 ○竹内 浩昭
東京理科大学 学生員 市川 博一

1. はじめに

建設分野等における各種の設計図面また工事竣工後に記録として残される各種の完成図面や写真は、関連する数値・文字情報以上に有益な情報としての価値を持つ場合が多い。これらの図面や写真を効率的に蓄積し、効果的にコントロールできるシステムの実現は、維持・管理業務のみならず、調査・計画、設計等の多くの建設業務において大きな効果をもたらすものとして期待がかけられている。しかしながら、一般的には、図面や写真として取り扱われているいわゆるイメージ情報の効率的な蓄積手法や蓄積形態、あるいはイメージ情報の効果的な提供までを取り扱った統合的な処理手法は確立されておらず、解決しなければならない課題も多い。その中でも、特に、図面データや写真データの数値化については、ある程度の指針は得られているものの、データベース構築に即した検討課題については解決しなければならないものが多い。イメージ情報を対象としたデータベースの構築は、数値・文字情報を対象としたデータベースの構築と比較して、データの収集や数値化（デジタル化）、コンピュータへの蓄積手法、あるいは、データベースの骨格であるデータ構造等において、大きく異なる点が幾つか存在する。特に、イメージ情報に係わる課題は、ハードウェアに依存する技術的な問題も多く、ソフトウェアとの両面からの検討を十分に行いながら技術開発を行う必要がある。

以上の背景の下に、本研究では、イメージ情報のデータベース化を主眼に、具体的な研究題材として道路工事に係わる図面データと写真データを用いながら、入力・蓄積・出力に係わる一連の処理の流れに配慮しながらイメージ情報のデジタル化に伴う問題点とその対応策を中心に検討し、その成果を取りまとめたものである。

2. イメージ情報の範囲

(1) 取り扱うイメージ情報の範囲

一般に、イメージ情報の具体例としては、地図、設計図、写真、ランドサットデータによる画像、あるいは身の回りの絵画などもその範囲として挙げることができる。本研究では、これらの幅広いイメージ情報の中で、特に、建設分野におけるイメージ情報のデータベース化を念頭に、道路工事に係わる基本的な図面データや写真データを用いながら、様々な面からその処理手法について検討を行った。従って、ここでいうイメージ情報とは、上記に述べた図面データや写真データが中心となり、これらを整理すると以下ようになる。

① 図面データ

図面等に代表される線画として表現される情報である。具体例としては、設計図面や工法図等が挙げられる。本研究では、全体位置図(縮尺1:25000レベル)、一般平面図(縮尺1:1000レベル)、推定土質縦断面図(縮尺1:1000レベル)、工法平面図、工法断面図、工法詳細図、動態観測に関する位置図・断面図等を取り扱った。また、この他にも線画により表現されるデータとして動態観測に係わる各種のグラフ(沈下量グラフ、変位グラフ等)や特殊機械説明図、工程表

等も図面データの範囲に含め検討を加えた。

② 写真データ

線画では表現ができず、採色や濃淡を用いて表現される情報である。具体例としては、航空写真（カラー、白黒）や現場の記録写真等が挙げられる。本研究では、特定の路線に沿った13枚の航空写真（縮尺1:10000、カラー）を検討対象とした。

(2) イメージ情報の特徴

数値・文字情報においては、その個々の情報が意味を有するのに対して、イメージ情報は、利用者の視覚をととして認識できる情報である。また、イメージ情報は二次元的な広がりを持つ情報ということができ、そのため、数値・文字情報に比べて比較にならない程の大きなデータ量を持つ。従って、イメージ情報は、データの性質や構造、あるいはデータベースへの格納や蓄積の面から見ても、数値・文字情報とは全く異なった検討が必要となってくる。検討するに当たっては、イメージ情報の特徴を十分に把握することが前提であり、最も基本的な特徴を整理すると次のとおりとなる。

- ① 二次元空間的なデータ構造を持つ。
- ② 数値・文字情報に比べ、データ量が大きい。
- ③ 人の視覚により認識される情報である。
- ④ 出力に際しては、鮮明さや精度が要求される。

これらの点を踏まえて、図面データと写真データの特徴を整理すると以下ようになる。

① 図面データ

- ・サイズ、記載内容、記載精度、表現形態等がバリエーションに富んでいる。
- ・一本一本の線分が持つ意味よりは、線分どうしの相互の関連が意味を持つ。従って、線分ごとにコントロールする必要性が小さい。
- ・一般図面とグラフのようにデータにより記載精度が異なる場合がある。
- ・一般に2色（白地に黒い線等）で表現されている。
- ・余白部分が比較的多い。

② 写真データ

- ・面的な情報であり、形状だけでなく彩色や濃淡も重要な意味を持つ。

3. イメージ情報のデータベース化に伴う問題点とその対策

(1) データの収集

一般に、図面は、A0版やB0版といった大サイズがほとんどであり、加えてA3版程度の図面まで、サイズ、縮尺、記載精度、記載内容とも極めて多種多様である。数値・文字情報であれば、収集フォーマットや記載内容をある程度指定し、収集することが可能であるが、図面データの場合には、図面一枚あるいはその一部がそのままコンピュータに入力されるデータであり、これらの各種の図面データや写真データは使用できるハードウェアシステムのサイズや形態に合わせて収集しなければならない。そのためにはコピー機による複写図面を用いるとか、マイクロフィッシュ等により保管されている場合などは前処理操作で適当なサイズに変換する等の処理が必要となる。以上から、データの収集における主な留意点として次の2つが挙げられる。

まず第一に、イメージスキャナを用いて読み込むためには縮小しなければならない図面が多いことが挙げられる。本研究では、ハードウェアによる制約からA3版以下のサイズが必要とされ、このサイズを越える図面については、コピー機を用いて複写図面を作成した。この際には精度の低下に伴って記載内容が読みずらくなならないように配慮しなければならない。具体的な対応策としては

複写に用いる図面は原図面を用いることや一枚の図面の中でも必要な領域のみを複写することにより縮小率を小さくする等の工夫を施した。表示精度は、記載内容の細かさや要求される表示精度にもよるが、A3版までの縮小であればある程度の記載内容についての精度は確保される。この点については、読み込み精度と表示する際の精度あるいは領域等の観点を中心に検討を加えている。写真データについては、一般の航空写真のサイズは、A3版以下の大きさであるため、サイズの面からは特に問題はない。

第二は、図面データと写真データそのものとそれに付随する属性情報をもれなく効率よく収集することである。属性情報は、イメージ情報を管理するインデックス情報としての重要なデータの集合である。属性情報の収集に際しては、図-1に示すようなイメージデータ登録シートを作成し、属性情報は全てこのデータ登録シートに記載することとした。

(2) イメージ情報のデジタル化

収集された図面・写真データは、イメージリーダーやイメージスキャ

ナを用いた場合には、ラスターデータ(またはメッシュデータ)としてデジタル化(A/D変換)され、ディジタイザを用いた場合にはポリゴンデータとして変換される。本研究で対象とするような土木技術情報で取り扱われる図面やグラフは、線画の密度が非常に高く、また、図面を構成する線分一本一本は意味を余り持たず、線分をベクトルデータとしてデジタル化する必要性は小さい。また、ポリゴンデータやベクトルデータを作成する実際の作業面からも、人手を介してのディジタイザによる作業は多大な時間を要する。従って、本研究では基本的に、写真データはもちろんのこと、図面データはラスターデータとして取り扱うことにする。図面データをラスターデータとして取り扱うことにより様々な図面あるいはグラフや表が記載内容に係わらずそのまま自由に取り込むことができるようになる。また、データベースとしてのデータ構造もわかりやすいものとなりデータベースの構築を容易にする。しかし一方で、イメージ情報をラスターデータとして取り扱うことにより、データ量が非常に大きくなる等の問題点が生じる。次に、データ量を可能な限り少なくするための対策方法について述べる。

路線名称		工事名称		工事KEY	
図面名称					
全体記載内容 (スケール)					
スケッチ メモ	スケッチ対象内容				
	スケッチ領域	L	XC	(cm)	
	スケッチピッチ	(mm/本)			
	サイズ	L	XC		
	70°-75°		1/4回転速度	(R/S)	
データ ベース 登録 メモ	工種コード				
	図面分類コード			図面分類下位コード	
	K.P.	地点	~		
	STA.		~		
	FILE NAME	Fn	Ft	Disk	
	図面KEY				
登録図面名称					
表示 チェック メモ					
備考	写真データ/SL/R(~)G(~)B(~)				

工種コード/100基本情報写真 10基本情報図面 20軟弱地盤処理工
30基礎工 40土留工 50土工
図面分類コード/100全体写真 10全体図面 21地質調査平面図
22土質柱状図 23地質推定断面図 31工法平面図 32工法断面図
41動態観測観測位置図 42動態観測断面図 43動態観測データ
50工程表 60その他
図面分類下位コード/【工法図面】1採石トレン 2引込マット 3載荷盛土 4引込トレン
57ロード 6バールトレン 7引込コンクリートパイプ 8引込トレン 9表層処理ネット
10引込ブラケット 11その他 【動態観測図】1沈下 2沈下(地表面) 3沈下(層別)
4変位 5変位(地表面) 6変位(地中) 7間隙水圧

図-1 イメージデータ登録シートの一例

(3) デジタル化の問題点の整理

イメージ情報のデジタル化はハードウェアが有する機能が最も直接的に影響するところであり、ハードウェアの性能を加味した技術的な検討が必要である。イメージ情報を取り扱う場合には、これらのハードウェア機能が大きく左右する。本研究ではイメージデータを出力する装置としてカラーディスプレイ装置を用いた。表-1にイメージスキャナとカラーディスプレイ装置の主な機能を示す。

イメージ情報のデジタル化に際しては読み込み精度の決定が非常に重要である。読み込み精度の選択は、データ量や表示される画像の精度、領域を直接に左右する。データベースに蓄えられたイメージ情報を利用するといった観点からは表示される画像の精度や一画面に表示される領域等が重要であり、データベースへの蓄積の面からは、データ量の増減は蓄積容量や蓄積する装置を検討する上で極めて課題となってくる。そこで、デジタル化の検討に際しての留意点を整理すると、①読み込み精度と表示精度、②読み込み精度と表示領域、③読み込み精度の3点となる。

① 読み込み精度と表示精度

読み込み精度と表示精度の関係は、細かく読み込む程その表示精度は高くなるが、収集した個々のイメージデータの記載精度や表示の際に求められる精度に充分配慮する必要がある。読み込み精度と表示精度の関係を図面データと写真データの代表的な縮尺である25000分の1、10000分の1、2500分の1、1000分の1、500分の1の5段階について検討した結果を表-2に示す。

表-1 イメージスキャナと
カラーディスプレイ装置の主な仕様

イメージスキャナ	
読み込み精度	25(40)、50(20)、100(10)、200(5)、1000(1) ($\mu\text{m}/\text{本}$) ($(\text{本}/\text{mm})$)
ドラム回転速度	1, 2 (回転/sec)
その他	適正密度調節機能、開口(71°-73°)調整機能
カラーディスプレイ装置	
表示画像	RGB合成カラー画像、シュードカラー画像
表示精度	1024×1024 ドット

表-2 読み込み精度と表示装置上で1画素が示す地上での大きさの関係
(1画素が示す地上での大きさは(m)単位)

読み込み精度($\mu\text{m}/\text{本}$)とデータ圧縮率						図面・写真データの縮尺					
25	50	100	200	500	1000	1/25000	1/12500	1/10000	1/2500	1/1000	1/500
1						0.625	0.313	0.25	0.065	0.025	0.013
1/2	1					1.25	0.625	0.5	0.125	0.05	0.025
1/4	1/2	1				2.5	1.25	1	0.25	0.1	0.05
1/5	-	-				3.125	1.563	1.25	0.313	0.125	0.063
1/8	1/4	1/2	1			5	2.5	2	0.5	0.2	0.1
1/10	1/5	-	-			6.25	3.125	2.5	0.625	0.25	0.125
1/16	1/8	1/4	1/2			10	5	4	1	0.4	0.2
1/20	1/10	1/5	-	1		12.5	6.25	5	1.25	0.5	0.25
1/40	1/20	1/10	1/5	1/2	1	25	12.5	10	2.5	1	0.5

表から図面・写真データの縮尺を考慮して読み込み精度を決定する場合には、例えば、1000分の1の図面において要求する最小の長さを0.2m(実際の図面上では0.2mmに相当する)とするならば、表中において○で囲んだ数値から矢印をたどると、読み込み精度としては200($\mu\text{m}/\text{本}$)以上の精度が要求されることが判る。従って、おおよそ200($\mu\text{m}/\text{本}$)以上の読み込み精度であれば表示

精度は満足できることが表から読み取ることができる。

② 読み込み精度と表示領域

表示装置上で一画面に再現可能となる最大の表示領域（画像の大きさ）と読み込み精度および読み込み領域の関係は表-3のとおりである。

表-3 読み込み精度と表示装置上での表示領域
（画像の一边が示す長さは（Km）単位）

読み込むことのできる領域の 一边の長さ(mm)	読み込み 精度 ($\mu\text{m}/\text{本}$)	図面・写真データの縮尺					
		1/25000	1/12500	1/10000	1/2500	1/1000	1/500
25.6	25	0.64	0.321	0.256	0.067	0.026	0.013
51.2	50	1.28	0.64	0.512	0.128	0.051	0.025
102.4	100	2.56	1.28	1.024	0.256	0.102	0.051
204.8	200	5.12	2.56	2.048	0.512	0.205	0.102
512	500	12.8	6.4	5.12	1.28	0.512	0.256
1024	1000	25.6	12.8	10.24	2.56	1.024	0.512

A3版に縮小した1000分の1の図面データは、実縮尺では2500分の1である。この時、図面全体にわたって線画が記載されていた場合、図面の記載領域の大きさは約40cmであることから地上での最大の距離は約1Kmである。これを一画面に表示するには表中の矢印をたどり、500($\mu\text{m}/\text{本}$)の精度で読み込む必要のあることが判る。しかし、実際には表示精度も考慮しなければならない。また、表中の左端には、読み込み可能な領域を示した。表から、読み込み精度 25($\mu\text{m}/\text{本}$)で読み込むとすると、見たい画像を一画面に表示しようとするデータの間引かなければわずか25.6mm四方の領域しか読み込むことができないことがわかる。

③ 読み込み精度とデータ量

読み込み精度と一枚の図面・写真データの大きさとデータ量の関係を表-4に示す。

表-4 読み込み精度とデータ量（単位：MB）

読み込み精度($\mu\text{m}/\text{本}$)とデータ圧縮率						イメージデータの大きさ(mm×mm)				
25	50	100	200	500	1000	A3版 $\times 1/2500$ (297×420)	A4版 $\times 1/2500$ (210×297)	A5版 $\times 1/2500$ (148×210)	B4版 $\times 1/2500$ (252×358)	B5版 $\times 1/2500$ (179×252)
1						190.338	95.169	47.589	137.659	68.829
1/2	1					571.014	285.507	142.755	412.977	206.487
1/4	1/2	1				47.585	23.792	11.896	34.415	17.207
1/8	1/4	1/2	1			142.755	71.376	35.688	103.245	51.621
1/16	1/8	1/4	1/2	1		11.896	5.948	2.974	8.604	4.302
1/32	1/16	1/8	1/4	1/2	1	35.688	17.844	8.922	25.812	12.906
1/64	1/32	1/16	1/8	1/4	1/2	2.974	1.487	0.744	2.151	1.075
1/128	1/64	1/32	1/16	1/8	1/4	8.922	4.461	2.232	6.453	3.225
1/256	1/128	1/64	1/32	1/16	1/8	0.476	0.238	0.119	0.344	0.172
1/512	1/256	1/128	1/64	1/32	1/16	1.428	0.714	0.357	1.032	0.516
1/1024	1/512	1/256	1/128	1/64	1/32	0.119	0.059	0.030	0.086	0.043
						0.357	0.177	0.090	0.258	0.129

は合成カラー画像に対応するデータ量である。

表から、A3版の図面を100($\mu\text{m}/\text{本}$)の精度で読み込む場合には11.896MBのデータ量が必要になることが判る。また、カラー写真データをR(赤色)、G(緑色)、B(青色)を用いて合成カラー画像として表示する場合には3倍の35.688MB(約2400フィートMT1巻に相当)ものデータ量となる。従って、いずれの場合にもデータ量を少なくするための対策が必要である。

(4) 検討例

以上述べてきた3つの課題は、各々独立に検討し得るものではなく、相互の関連を充分考慮しなければならない。また、これらの他にも実際の作業を行う際に生じる問題点もいくつか存在する。ここでは、これまで述べてきた検討結果を踏まえて図面データと写真データについて実際に適用した検討例を示す。

① 図面データのディジタル化

本研究での図面データの読み込み精度は、前述の結果を考慮した上で200($\mu\text{m}/\text{本}$)とした。従って、表-2から実縮尺1000分の1と2500分の1の図面上においては、0.2mと0.5mの表示精度(一画素が示す地上での大きさ)が確保され、これは地図上では0.2mmに相当し精度としては適当であると思われる。一方、表-3からは表示領域について、1000分の1の図面で205m、2500分の1の図面で512mであることが判る。しかし、これら2つの図面のA3版サイズの最大の領域は、約400mと1000mであり、一画面に入り切らない。この時、表示領域を優先させると、表-3から読み込み精度は500($\mu\text{m}/\text{本}$)となるが、実際には表示精度に重きを置くため、一画面で入りきれない画像については複数画面に分割することにする。あるいは、表示する際にイメージデータを2分の1に圧縮することにより画像の全体を表示することとし、逆に詳しく見たい時には見たい領域を拡大する方法で対処する。データ量については、A3版サイズ全域に記載されている場合には200($\mu\text{m}/\text{本}$)の読み込み精度に対しては約3MBのデータ量(表-4参照)となるが、実際には図面全体が読み込み対象となることは少なく、必要な領域に絞り比較的多く存在する余白部分を除いた結果、本研究で取り扱った各種の図面の平均データ量は1枚当たり0.8MBであった。

② 写真データのディジタル化

写真データの読み込み精度は、種々の検討結果より100($\mu\text{m}/\text{本}$)が適当である。この場合、表示精度は航空写真(縮尺1:10000)の場合、地上解像度1.0mが確保され、写真上では0.1mmに相当する。読み込み領域は写真により異なるが、本研究では路線から周囲300mを対象領域として設定した。従って、約100×200mmが読み込み領域となる。この時、データ量は、R(赤色)、G(緑色)、B(青色)を用いた合成カラー画像の場合は1画素についてR、G、Bの3Byteのデータ量となり1枚当たり6MBが必要となる。しかし、これではデータ量が膨大となるため、216色のシュドカラー画像を用いることにした。この結果は、1画素を1Byteで表現できるためデータ量は3分の1となり1枚当たり約2MBとなる。この場合、色調についても合成カラー画像と比較して見劣りはしないことが検討の結果判明している。

4. イメージ情報のデータベース化

4.1 イメージ情報のデータベース化

これまで、図面データと写真データの収集から実際にこれらのイメージ情報を用いて読み込むまでに生じる課題を詳細に検討してきた。これらの結果を踏まえて、データベース化を念頭に置き、イメージ情報のディジタル化からデータベースへの登録までの手順を整理した結果を図-2に示す。

イメージスキャナによりA/D変換されたイメージデータは、磁気テープから磁気ディスク装置に転送される。続いて、図面データは2値化されたラスターデータに変換され、また写真データは、

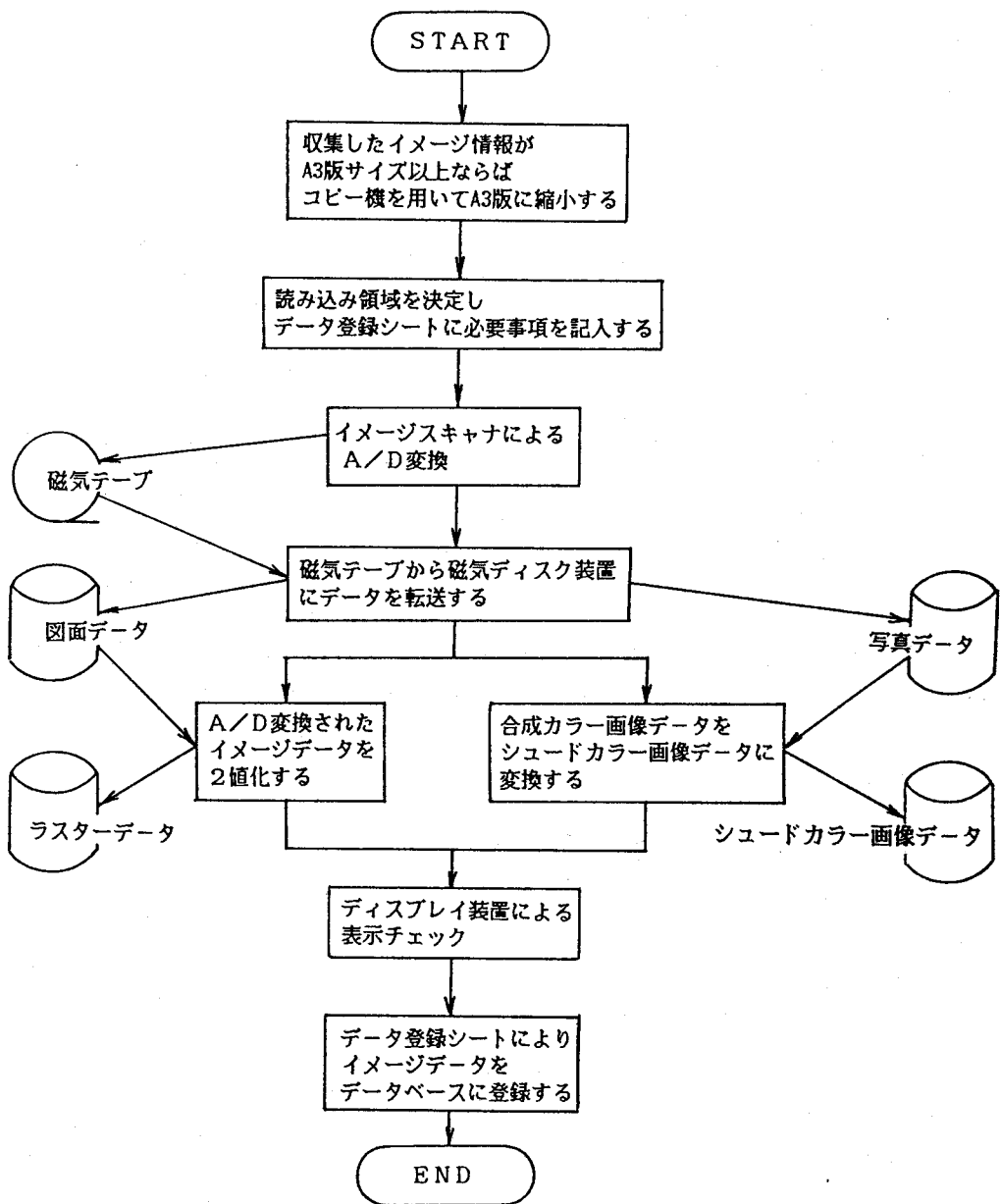


図-2 イメージ情報のデジタル化を中心としたデータベースへの登録までの流れ

R G Bによる合成カラー画像データから 216 色のシュードカラー画像データに変換されることによりデータ量が3分の1に減少する。このようにして作成されたイメージデータは最終的にカラーディスプレイ装置に出力され確認される。最終的に仕上がったイメージデータは磁気ディスク装置にファイルとして格納されデータ登録シートに記入されている属性情報とともにデータベースに登録されることになる。

4. 2 イメージ情報の蓄積と出力

(1) イメージ情報の蓄積

ディジタル化された個々のイメージデータは補助記憶装置に格納・蓄積される。これまで述べてきたように、イメージデータはデータ量が数値・文字情報に比べ非常に大きく大容量の記憶装置を必要とする。本研究では、図面データ約 130枚と写真データ13枚について検討した結果、各々のデータ量は約 104MBと35MBとなった。イメージ情報を対象としたデータベースを構築する際には、このような大容量に対応する補助記憶装置が不可欠であり光ディスク装置やCD-ROMディスク装置の積極的な利用を進める必要がある。また一方、データ構造も数値・文字情報と大きく異なり、本研究では、イメージデータそのものとこれらに付随する数値・文字で表現される属性情報に分けた上でデータベース化を進めている。図-3に蓄積形態の概念図を示す。各々を分けて蓄えることにより、光ディスク装置やCD-ROMディスク装置を利用する上で柔軟に対応できるとともに、イメージデータを蓄えるための容量が確保しやすくなることやイメージデータのインデックス情報としての属性情報をデータベース化することにより、データベース・システムとしての構築が容易となる。

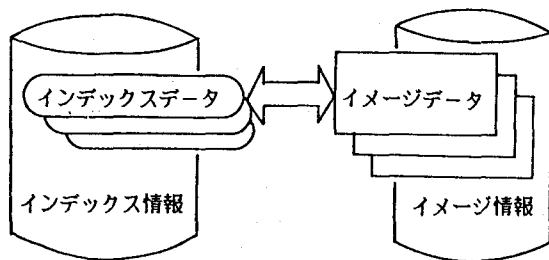


図-3 イメージ情報の蓄積形態の概念

(2) イメージ情報の出力

イメージ情報の出力に際しては表示領域や蓄積されている個々のイメージデータの再現性が重要な課題となる。この点については、イメージ情報のディジタル化の段階において検討を行っているが、最終的にはデータベースの利用形態を考慮した上で更に検討を加える必要がある。本研究では、イメージ情報を対象としたデータベースを念頭に置き、蓄積されたイメージデータを利用精度に応じて間引いて表示したり、イメージデータを分割して複数画面による制御等の基本的な機能について整備した。

5. まとめ

本研究では、土木技術情報として用いられている、図面・写真類のいわゆるイメージ情報の収集から出力までの一連の処理手法について特にイメージ情報のディジタル化を中心に述べた。イメージ情報を対象としたデータベースを構築する際に、情報の質に直接影響を及ぼすイメージデータのディジタル化について数値を詳細に実証的に検討したことにより、今後の実作業面での指針が得られた。今後は、検討対象を他の種類のイメージ情報へと広げ、1つ1つの処理手法を要素技術として確立していくことが重要である。特に、ハードウェアの技術的発展をにらみつつ、イメージ情報の取り扱いのポイントであるコンピュータへの入力、蓄積、更には大容量のデータ転送等も合わせたシステムの検討が必要となる。

【参考文献】

- 1)木戸出正継；画像データベース，オーム社 2)尾上 守夫；標準画像データベース，第8回画像工学コンファレンス論文集，昭和52年 3)石坂 公一；画像等情報システムの整備について，JACIC情報，(財)日本建設情報センター，198 6，Vol.1，No.3 4)大林 成行；建設業者のためのデータベース講座，JACIC情報，(財)日本建設情報センター，198 7，Vol.2，No.2