

(Ⅳ-19) EWSを用いた建設情報データベースの構築 におけるイメージデータの圧縮

東京理科大学 正会員 大林成行

東京理科大学 正会員 平野曉彦

○東京理科大学 学生員 橋本和記

1. はじめに

現在、建設分野においても工事实績情報など様々な情報をコンピュータライズし、それらを積極的に利用しようとする動きが活発である。特に、図面、工事写真や協議文書などイメージデータの有効的な活用に対するニーズが高いことから、イメージデータのデータベース化、すなわち画像データベースが注目されるようになってきている。

ところで、このようなデータベースは、実際には工事事務所クラスでの利用が中心であると考えられるため、安価で空調などの設備を必要としない小型の計算機を用いて構築されることが望まれる。従来、膨大なデータ量となるイメージデータを大量に取り扱う画像データベースは大型汎用機の範ちゅうであったが、最近のEWSやパーソナルコンピュータにおける処理性能の向上、磁気ディスク装置や光ディスク装置の記憶容量の増加、低価格化傾向により、小型の計算機システムを用いて画像データベースを構築することが可能となってきている。

しかし、実際に画像データベースを構築する上では幾つかの技術的課題があり、特に、イメージデータ特有の膨大なデータ量については、ハードウェアが飛躍的に進歩しているとはいえ、イメージデータの蓄積に際しては十分な検討が必要となる。そこで、本研究では、実際に道路工事で使用されている図面や工事写真を対象に画像データベースを構築することにより、その効果的な蓄積・利用方法、および実際のデータ量と圧縮による効果について詳細な検討を行なった。

2. イメージデータの利用例

一般に画像データベースといってもデータの表現方法、蓄積形態、検索方法などによりその形態は大きく異なるが、本研究では、イメージデータは全てラスタデータに変換した上で蓄積している。

また、イメージデータは外部ファイルとして磁気ディスク装置に常駐させ、そのインデックスデータをデータベースに登録するディレクトリ方式の蓄積形態をとっており、実際のイメージデータの検索はこのインデックスデータに対して行なわれることになる。したがって、図面上の各キロポストことの座標値データをインデックスデータとして登録しておくなどイメージデータ相互の位置関係を定義することにより、イメージデータ相互間でダイナミックな検索を行なうことができる。例えば、図-1に示すようにディスプレイ装置上に表示されている図面の任意の位置をポインティングデバイスで指定することにより、その位置の他の図面や工事写真、協議文書等を検索し、表示することができる。すなわち、従来、検索条件をキーボードにより入力していたのに対し、このように視覚的に検索し、必要な情報を参照することが可能となる。

3. イメージデータの圧縮

本研究では、イメージデータをラスタデータに変換する際の読み込み精度を、表示精度、表示領域およびデータ量の3つを考慮した上で、図面データについては200 μm 、写真データについては100 μm としている。この場合、データ圧縮を全く施さなければ、A4版の図面データで約1.5MB、写真データではR、G、Bの合成カラーとなるため18MBという膨大なデータ量になってしまう。そこで、本研究では図面データ、写真データそれぞれについて、多くの試行の結果、以下のようなデータ圧縮を行なった。

①図面データ

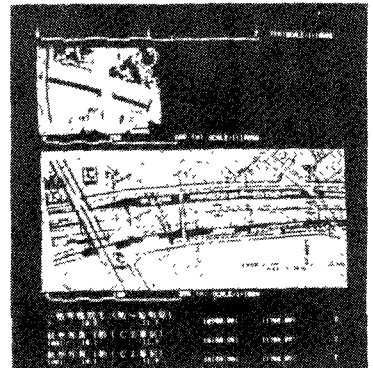


図-1 イメージデータの検索・表示例

ここでいう図面データとは白地に黒の線画で表現された図面、グラフ等を指しており、このようなデータに対しては①2値化処理、②ランレングス符合化、の2段階のデータ圧縮を行なっている。スキャナにより図面等をデジタル化したデータは、明確に白と黒の2値に分かれておらず、3値以上の濃淡画像となっている場合が多い。そこで、ある値を境にそれ以下を「0」、それ以上を「1」といった2値化処理を行なった。その結果、1画素を1bitで表現でき、データ量は1/8になる。さらに、図面データには余白部分が多く、黒の線画部分に連続性があることに注目し、ランレングス符合化により比較的簡単で効果的な圧縮を行なっている。2値化処理を行なったデータは白画素と黒画素の連続が交互に表われる。この画素の連続をランといい、データをこのランの長さ（ランレングス）で表現したものがランレングス符合化である。この圧縮方法は元のデータの精度を落すことなくデータ量の圧縮を行なえる点に特徴がある。データ量だけに着目した場合他の効果的な圧縮方法もあるが、本研究では、読み込み精度と表示精度も十分に考慮した上で、ランレングス符合化による圧縮を行なっている。

②写真データ

一般に工事写真等カラーのイメージデータにおいては、R、G、Bそれぞれの濃淡階調数を256(1B)とすることが多い。この場合、イメージデータのディスプレイ装置上での表示可能色は約1677万色にのぼり、極めて高精度な情報となるが、人間の視覚ではこれら全てを識別することは不可能である。本研究で対象とする画像データベースでは、画素1つ1つの濃淡値よりも相対的な濃淡階調によるテクスチャが重要となるため、人間の識別能力を越えた濃淡階調数は必要としない。そこで、まず、スキャナにより読み込まれたデータに対して濃度抽出を行なった上で、R、G、Bそれぞれの濃淡階調数を6とし、 $6^3=216$ 色でカラー表現することによってデータ圧縮を行なっている。このように、データの精度を落すことにより1画素あたり3Bであったデータ量を1Bに圧縮しているが、データの利用には支障がないことを確かめている。

以上のようなデータ圧縮による効果を表-1に示す。このなかで特に、図面データについては平均で1/10以下のデータ量にまで圧縮できており、建設情報を対象とした画像データベースの構築には効果的な圧縮であるといえる。

4. まとめ

本研究では、実際に小型の計算機を用いて画像データベースを構築し、その効果的な利用法を探るとともに、特にデータ量に着目し、実際のデータ量および圧縮による効果を確認することにより以下に示す成果を得た。

①イメージデータをデジタルデータとして蓄積することは、ポインティングデバイスを用いた検索を可能とし、イメージデータの有効的な利用の1つの手段となる。

②図面データにおける圧縮実績によれば1cm²あたりのデータ量は0.231KBとなっており、今回使用した72MBのハードディスク装置にはA4版の図面を524枚常駐させることができる。したがって、補助記憶装置の容量さえ確保できれば、小型の計算機を用いてもある程度のイメージデータの大量蓄積が可能となる。

また、画像データベースの利用環境としては、イメージデータを書類をめくる感覚で参照できることが望ましいが、表示に相当の時間を要するのが現状である。出力所要時間の大部分がデータ転送で占められており、この問題は小型の計算機固有のものではなく、大型汎用機などでも大差はないといえる。したがって、今後の課の課題として、データ転送時間を短縮するためにより効率の良い圧縮方法の検討が必要となる。

表-1 本研究で扱ったイメージデータのデータ量と圧縮による効果

図面データ	圧縮法	1画素1bit/1cm ² 濃淡画像	32レングス-1表現
	枚数(枚)	96	96
	総画積(cm ²)	29801.3(平均:310.4)	29801.3(平均:310.4)
	総データ量(MB)	71.040(平均:0.740)	6.720(平均:0.070)
写真データ	圧縮法	1画素3B合成A3-画像	1画素1B合成A3-画像
	枚数(枚)	12	12
	総画積(cm ²)	2306.8(平均:192.2)	2306.8(平均:192.2)
	総データ量(MB)	66.000(平均:5.500)	22.000(平均:1.833)

【参考文献】

- 1) 木戸手正継；画像データベース、オーム社
- 2) 大林、平野、竹内、市川；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、第12回電算機利用に関するシンポジウム講演集、1987
- 3) 大林、平野、竹内、市川；図形／画像データベースの概念設計とモデル構築、リモートセンシング研究所報告 No. 6、東京理科大学リモートセンシング研究所