

4.2.1 スキャナーの読み取り

1) モードの選択

保有情報を方式 , によりデジタル化するが、その際のスキャナーのモードの選択基準を表 4.2.1.1 に示す。尚、原則としてカラーは使用しないので対象外とする。

方式 : ページ単位で全体を一度にデジタル化

方式 : 図表・写真をデジタル化した後、デジタルの文書のページに挿入

読み取る対象	線画モード	白黒ハーフトーンモード
a) 文書（小さな添え字等もある）		
b) 線画（線で構成される図表）		
c) 濃淡画（イラストや説明図など濃淡をもった図表）	×	
d) 白黒写真	×	

表 4.2.1.1 モードの選択基準

画像を白黒の2値で表現する a),b)と、グレーで表現する c),d)とに分けられ、a),b) = “線画モード” c),d) = “白黒ハーフトーンモード” を選ぶ。後者は白黒の2値で中間調処理をして読み取ってグレーを表現するものである。

読み取る対象によりモードを選べば良いが、問題は方式 に関して対象のページに a),b)と c),d)とが混ぜん一体となったものが多いことである。この場合は“白黒ハーフトーンモード” を選ぶが、a),b)を で示した様に PDF ファイル化した時にデータ量が若干増加する。この点については後出の表 4.2.2.2 と 4.2.2.3 (23,24 ページ) での比較検討を参照されたい。ページを a),b)と c),d)との領域に分けて読み取って合成するのは労多くして効果は少ない。

2) キャリブレーションの設定

スキャナーで読み取る時、印刷機器の種別によってキャリブレーションを設定する必要がある。しかし、閲覧する側がどんなプリンターを持っているのか分からない。そこで、現在主流になっているレーザープリンタとインクジェットプリンタの各々の場合について試行してみた。その結果、打ち出した画質上では明確な相違は見られなかった。結局、キャリブレーションには余り拘らなくてもよいことが分かった。今回は業務用としてよく使われるレーザープリンタをキャリブレーションとして設定する。

3) イメージ制御の設定

スキャナーの読み取りに際し注意すべきポイントはイメージ制御である。“線画モード”では[しきい値]、“白黒ハーフトーンモード”では[明度・コントラスト・ハ

イライト・シャドウ・輪郭強調]等の設定が必要である。

設定値は紙媒体の印刷状況によっても違ふし、スキャナーの機種間のバラツキでも変わる。今回の読み取りデータは他のケースの参考にはならないので公開しない。各々手持ちのスキャナーで読み取りながら、対象の印刷物と見比べてより良い設定状態を探るという手間のかかる作業を繰り返す必要がある。“線画モード”では[しきい値]を設定するだけであるから、数回の試行で決まるであろう。問題は“白黒ハーフトーンモード”である。複数の設定箇所があるので何回も試行を重ねて勘所を掴むことが要求される。中でも方式で表4.2.1.1のa),b)とc),d)とが混在したページを読み取る場合、a),b)を鮮明にするとc),d)の深度が浅くなり暗くなる。又、c),d)を鮮明にするとa),b)が薄くなる。この調整はかなり時間を取られるが、手間を省くと最終的に良いものにならない。逆に一度定まれば同様のページが同じ設定で処理出来るので大変効率が良い。

4) 解像度の設定

解像度に関しては、読み取る対象物の精細度や濃淡の再現性を吟味して決める必要がある。ある大学図書館における電子化の事例では、保存用のマスター画像を400dpi(256階調のグレースケール)で作成し、これから公開用の画像を200dpiで作成している。³⁾

保有情報の解像度を決めるに際して、図表や濃淡画、大小の文字等種々のケースが含まれるサンプルとして助成研究成果報告概要から四つの事例を抽出して試行してみた。

図4.2.1.1に於いて、事例1~3は“線画モード”事例4は“白黒ハーフトーンモード”で各々300dpi, 240dpi, 200dpiで読み取った場合の比較である。

結果から言うと200dpiでも大体良さそうだが、細かく見ると紙媒体と比較して次の様な問題がある。

事例1: 字体がにじんだ印象になる。

事例2: グラフの曲線部分のジャギー(凹凸)が目立ってくる。

事例3: 直線に乱れが生じて汚くなる。

事例4: 濃淡の画質にざらつきが目立ってくる。

300dpiでは、ほぼ紙媒体と同等の品質が維持でき、問題点は解消されている。240dpiでも少しは改善するが、紙媒体と比較するとやはり品質が落ちる印象がある。結局、画質を維持するためには300dpiを採用することとした。

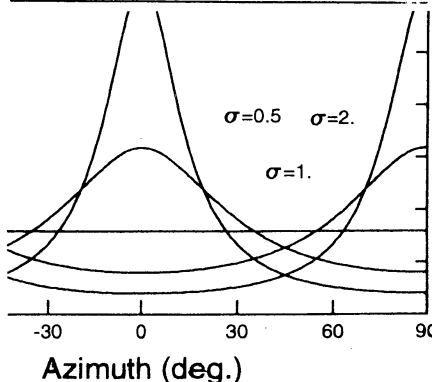
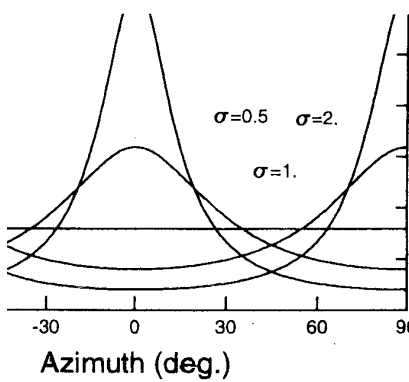
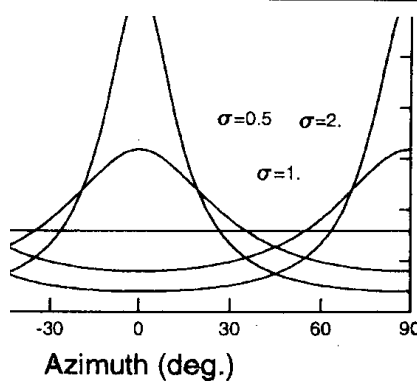
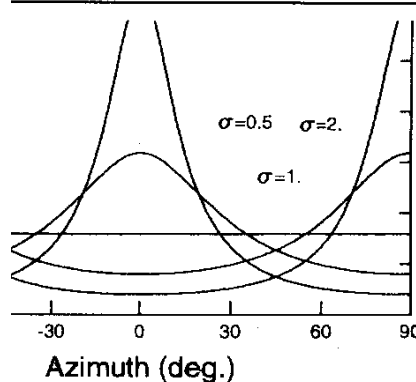
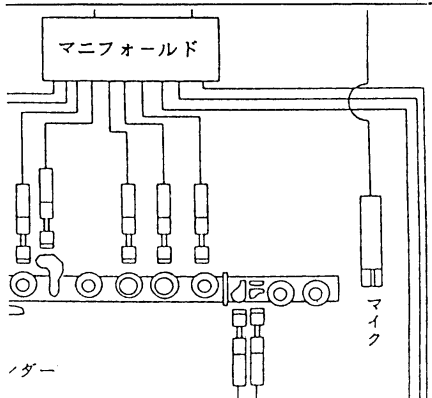
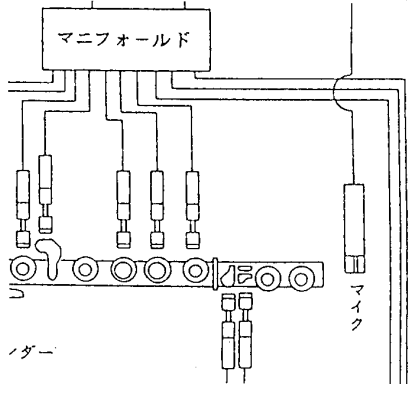
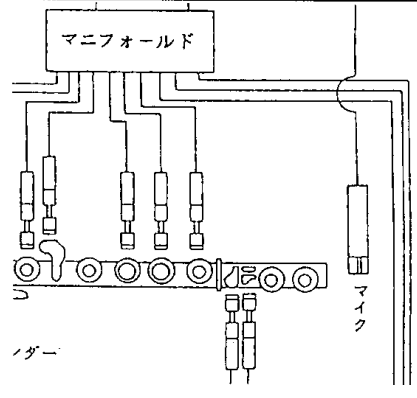
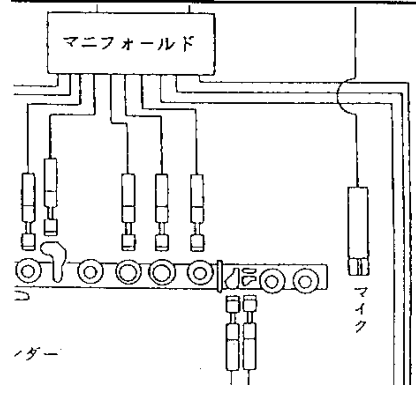
	紙媒体	300dpi	240dpi	200dpi
事例 1	製作しセンサの写真を示す。いずれも偏向する横波を同時に発生する。ども変化する磁場が生じ、この変化を打ち永久磁石が作る静磁場と作用しあい、の衝突などの相互作用によって運動しコイルと永久磁石の静磁場形状によゆいため、y方向に偏向する横波が発、磁石が作る静磁場は放射状に広がれローレンツ力 F_z および F_y が生じ縦	製作しセンサの写真を示す。いずれも偏向する横波を同時に発生する。ども変化する磁場が生じ、この変化を打ち永久磁石が作る静磁場と作用しあい、の衝突などの相互作用によって運動しコイルと永久磁石の静磁場形状によゆいため、y方向に偏向する横波が発、磁石が作る静磁場は放射状に広がれローレンツ力 F_z および F_y が生じ縦	製作しセンサの写真を示す。いずれも偏向する横波を同時に発生する。ども変化する磁場が生じ、この変化を打ち永久磁石が作る静磁場と作用しあい、の衝突などの相互作用によって運動しコイルと永久磁石の静磁場形状によゆいため、y方向に偏向する横波が発、磁石が作る静磁場は放射状に広がれローレンツ力 F_z および F_y が生じ縦	製作しセンサの写真を示す。いずれも偏向する横波を同時に発生する。ども変化する磁場が生じ、この変化を打ち永久磁石が作る静磁場と作用しあい、の衝突などの相互作用によって運動しコイルと永久磁石の静磁場形状によゆいため、y方向に偏向する横波が発、磁石が作る静磁場は放射状に広がれローレンツ力 F_z および F_y が生じ縦
事例 2	 <p>$\sigma=0.5$ $\sigma=2$ $\sigma=1$</p> <p>Azimuth (deg.)</p>	 <p>$\sigma=0.5$ $\sigma=2$ $\sigma=1$</p> <p>Azimuth (deg.)</p>	 <p>$\sigma=0.5$ $\sigma=2$ $\sigma=1$</p> <p>Azimuth (deg.)</p>	 <p>$\sigma=0.5$ $\sigma=2$ $\sigma=1$</p> <p>Azimuth (deg.)</p>
事例 3	 <p>マンフォールド</p> <p>マイク</p> <p>ノダー</p>	 <p>マンフォールド</p> <p>マイク</p> <p>ノダー</p>	 <p>マンフォールド</p> <p>マイク</p> <p>ノダー</p>	 <p>マンフォールド</p> <p>マイク</p> <p>ノダー</p>

図 4.2.1.1 300 ~ 200dpi で読み取った事例 1 ~ 3 注) 配信用では紙媒体を 480dpi で読み取った。

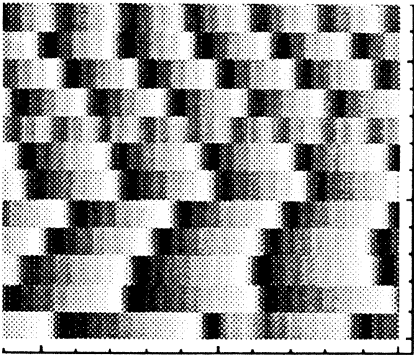
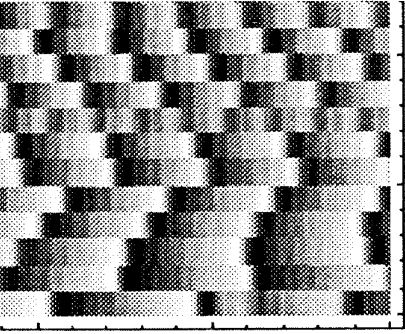
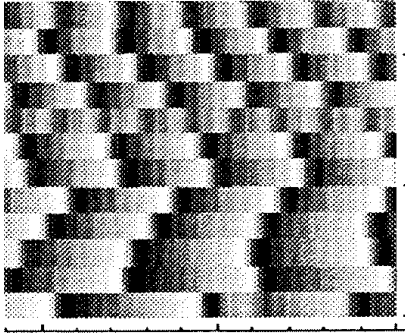
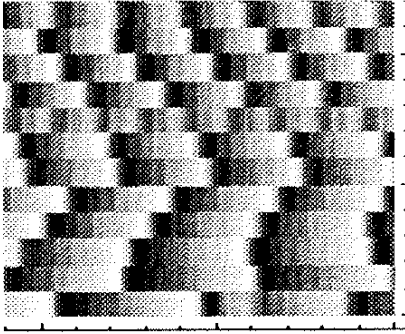
	紙媒体	300dpi	240dpi	200dpi
事例 4				

図 4.2.1.1 (続き) 300 ~ 200dpi で読み取った事例 4 注) 配信用では紙媒体を 480dpi で読み取った。

5) ファイル形式の選択

スキャナーで読み取ったデータのファイル形式は、代表的な GIF と TIFF を取り上げる。この選択の判断は PDF に変換したときのデータ量の相違が問題になってくる。次節の 4.2.2 に於いて PDF ファイル化の際の圧縮オプションの設定にからめて検証するが、結果的には双方に明確な差はなかった。

6) “白黒写真モード”について

モードの選択に於いて、スキャナーには“白黒ハーフトーンモード”の他に 256 階調でグレーを表現する“白黒写真モード”があり、これで写真を読み取ることも考えられる。しかし、以下の理由で採用しなかった。尚、ここでの実験のファイル形式は GIF である。

- ・写真を 300dpi で読み取ってデータ量を比較すると、“白黒ハーフトーン”：“白黒写真” = 1 : 10 程度となり、圧倒的にデータ量が増える。PDF 変換する際に JPEG 等の圧縮オプションを駆使して高水準に圧縮しても 1 : 2.7 程度にしか縮まらない。写真はそれだけでなくデータ量が多く、この差は配信の際の大きな負担になる。
- ・解像度を落とした場合を次ページの図 4.2.1.2 に示す。1)の“白黒ハーフトーン” 300dpi で読み取ったもの(データ量 = 57KB)と比較して、2)の“白黒写真” 180dpi で読み取ったものを JPEG で高水準に圧縮すれば 3)に示すようにデータ量は 47KB に減る。しかし、劣化を伴う高圧縮の影響もあってピントはかなりぼけてしまう。圧縮前の 2)データ量 = 276KB でもピントの点でやはり少しぼけている。データ量と画質の双方から判断して“白黒ハーフトーン”が優れていると判断できる。
- ・閲覧側にモノクロのプリンターしか無い場合でも“白黒ハーフトーン”ならベストな印刷が可能である。“白黒写真”はグレーが表現できるプリンタードライバを備えている必要があり、保有資料の様にモノクロで印刷して見ることが多いものに制約を設けることは得策ではない。

尚、“白黒写真”は文字には全く不向きであり、方式 の様にページ全体を対象とするとお手上げである。以上の検討結果から 300dpi のピントを維持して最もデータ量が少なく方式 にも対処できる“白黒ハーフトーン”を採用することにした。



1) “白黒ハーフトーン” 300dpi (57KB)



2) “白黒写真” 180dpi (276KB)



3) 高水準で圧縮後の“白黒写真” 180dpi (47KB)

図 4.2.1.2 “白黒ハーフトーン” と解像度を落とした“白黒写真” との比較

注) 配信用では写真の PDF 化は画像情報の損失がない ZIP で圧縮した。