

デジタル処理における前処理技術

吉田 真澄 棚橋 純一

(株式会社 富士通研究所)

1. まえがき

情報を処理する手段にはアナログ方式とデジタル方式がある。アナログ方式は光学技術や写真技術を土台とするもので、さほど精度を必要としない計算にはその並列実行性を利用して、簡便、かつ、実時間でこなせる利点をもつが、再現性、融通性、さらにはアナログ素子の調整などに欠点がある。一方、デジタル方式はコンピュータ技術を基礎とするもので、メモリや処理速度に欠点をもつが、処理の精度、再現性、融通性に優れている。

このように、アナログ方式、デジタル方式ともに長所短所をもつが、コンピュータ技術の向上に伴い、メモリや処理速度の問題は解決される方向にある。¹⁾²⁾ そのため、今日では情報処理の手段として、デジタル方式が主流になっており、この傾向は今後益々強まると思われる。

デジタル方式における信号処理の技術開発分野には、パターン認識、画像処理、帯域圧縮、人工知能、さらには生体系メカニズムの探求などがある。しかし、いずれの分野も未だ確固たる処理体系が構築されておらず、処理対象や処理目的に応じて適宜、対処している状態である。

本稿ではこうした現状のもとで、文字、図形、および音声を対象としたパターン認識、ならびに濃淡画像を対象とした画像処理をとりあげ、それらの分野で“前処理”として捉えられる概念と技術について紹介する。

尚、現在のデジタル信号処理において、“前処理”を一つの固定概念として捉えるのは困難であるため、本稿では筆者らのこれまでの研究活動の中で位置付けている“前処理”について述べる。

2. デジタル信号処理

2.1 理想システムのモデル

デジタル信号処理の目的は、人間相互のコミュニケーション機能を機械に代行させることにあるといえる。したがって、この実現にあたっては単なる符号処理にとどまらず、人間のもつ意思伝達機能を考慮した幅広い視野にたったシステム設計が必要である。こうしたシステムはまだ存在してないが、理想システム案としてShannonの通信模型がある。³⁾

図-1はその通信模型におけるコミュニケーション機能を、人間系に対比させたものである。これをみると、コミュニケーションは、機械系では“シンボル→シグナル→シンボル”、人間系では“概念・意味→変換→再現→特徴→認識→概念・意味”の推移によって行な

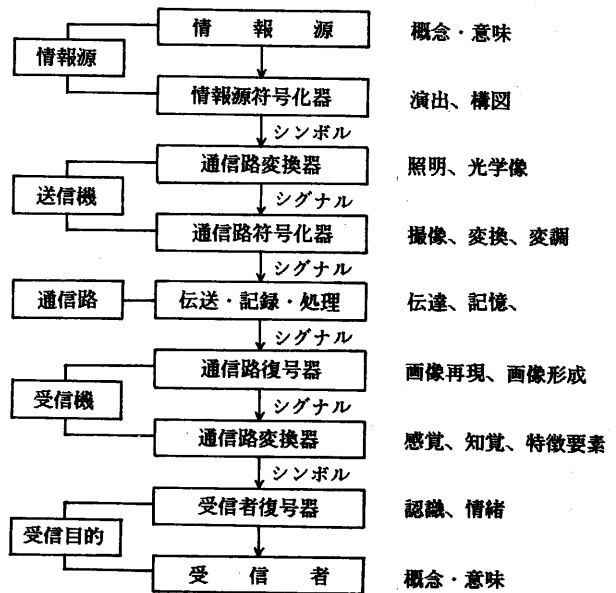


図-1 機械系と人間系

われることが分かる。

2. 2 現行システムのモデル

図-1に示したように、デジタル信号処理の基本は、人間系で行なわれるコミュニケーションの推移を、機械系の“シンボル→シグナル”、“シグナル→シグナル”および“シグナル→シンボル”からなる三種類の信号変換機能に置換することにある。これを今日の工学上の研究テーマ面から捉えてみると、“シンボル→シグナル”は入力、“シグナル→シグナル”は処理や記憶、“シグナル→シンボル”は処理、表示、出力(記録)、また、“→”は伝送などの技術開発にあたる。

図-2は現行の各種技術をデジタル信号処理システムとしてまとめたものであるが、技術はお互いに独立している分けでなく、密接な関係を保ちながら一つの処理系を形成する。

2. 3 システムにおける前処理

デジタル信号処理システムにおいて、「前処理とは何か」を論ずることは大変に難しく、これまでにそれを明確

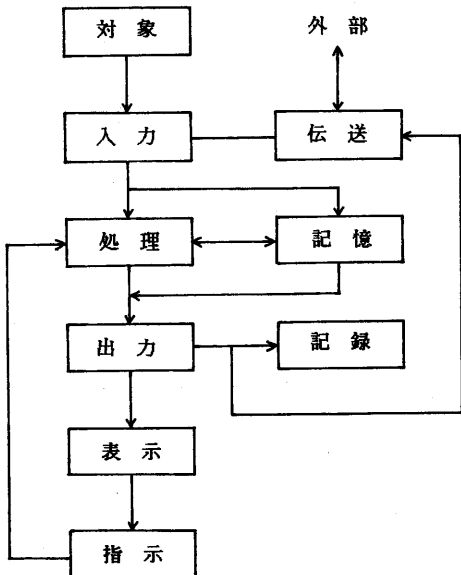


図-2 現行のシステム

に定義した例はない。この理由はシステムの基本が人間のもつ複雑な心理的思考過程を表現することにあるため、その実現に際しては図-2に示したように、異なる目的をもった数多くの技術を統合する必要があり、それによって生み出される相乗効果を明確に把握できないからであると思われる。

こうした背景を踏まえ、筆者らはシステムにおける“前処理”を技術面ではなく、人間の心理的思考過程に基づく機能面で定義することにした。

2. 3. 1 心理的モデル

図-3は図-1および図-2において、信号処理が行なわれる部分を、人間相互(送信者-受信者)の情報変換機能におきかえ、それを心理的なモデルで表現したものである。このモデルでは、送信者によって再現、あるいは形成された情報が、送信者と受信者に共通な感覚や知覚に依存した形に変換され、それが受信者の感情、認識、および記憶によって捉えられるようになっている。

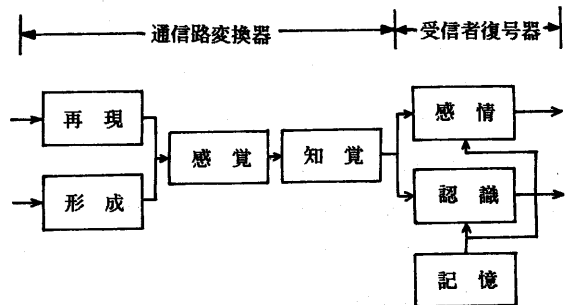


図-3 心理モデル

2. 3. 2 前処理の定義

図-3のモデルをもとに、筆者らはデジタル信号処理における“前処理”をつぎのように定義する。

◆定義……概念

送信者が自らの意思を受信者に伝達する手段の中で、受信者のもつ感情、認識、記憶に左右されない程度に信号変換を行なう行為。

◆定義……技術

ディジタル信号処理システムにおいて、情報の再現、および形成を行なう技術、ならびに、認識に直結した情報を得易くする技術。

2. 3. 3 前処理の機能

上記定義による前処理の意義、機能をまとめると下記のようになる。

(1) 情報の再現、形成

情報の再現、形成は人間のもつ視覚、聴覚、触覚などの感覚器官から得られる情報を、システム内で扱い易く、あるいは他人の理解を喚起するように変換するための技術に相当する。実際には以下の事柄に対処する必要がある。

- ・知覚面……送信者が対象物を直接感覚器で捉えたときに生ずる知覚をそのまま受信者に呼び起こさせる。
- ・感情面……受信者に美的感覚、快感、あるいは立体感などの感情を引き起こさせる。

(2) 認識のための処理

これは再現、あるいは形成された情報を、人間が認識し易くするためのもので、“シグナル→シンボル”への移行を促進する技術に相当する。実際には空間上、あるいは時間上に再現、形成された信号に対し、以下の事柄に対処する必要がある。

- ・空間上……二次元、あるいは三次元的な空間上で再現された情報の中から、認識対象物のみを見出し、さらに、その構成要素の変動を補正する。
- ・時間上……音声などのように、時間的な経緯として再現された情報の中から、認識対象範囲を見出し、さらに、その構成要素の変動を補正する。

3. 前処理技術

3. 1 パターン認識

パターン認識とは本来、区別や判断といった人間の知的情報処理能力の根源をなす知覚行為である。したがって、

パターン認識における前処理は、受信者で行なわれる認知活動（識別）を促進させる行為といえる。

ところで、一般に我々人間がパターンを識別する際に行なう行為としては、「対象物、あるいはその一部分があるか否か」、「対象物、あるいはその一部分がそれを活用する環境下において、約束された形で表現されているか否か」などが指摘できる。

こうした行為が第2. 3. 3項の(2)にあたることは容易に推察できるが、これを今日の技術用語で表現すれば、前者はセグメンテーション、後者は正規化になろう。

以下、パターンとして、文字、図形、音声をとりあげ、それらの認識におけるセグメンテーション、および正規化技術を紹介する。

3. 1. 1 文字

(1) セグメンテーション

文字認識におけるセグメンテーションは、帳票から対象文字行(列)を見つけ、文字を一字ずつ切り出す技術にあたる。これまでのところ、手書き文字認識では枠内文字を対象としているため特別な技術はない。それに対し、印刷文字認識では文字間のピッチが小さいので、印刷機械のプリント不良などによる文字の位置ずれが生じ易く特別な工夫が必要になる。

(文字行の検出)

図-4は印刷文字認識における文字行の検出例である。図-4(a)は隣接する文字領域間の縦軸射影が重ならない場合の手法である。ここでは最初に、前もって設定しておく文字列検索ウィンドウを走査させ、文字を構成する黒点を縦軸に投影する。その後、投影された黒点の形状をもとに、文字の行間を区分する境界、ならびに、文字領域の中心を検出する。図-4(b)は縦軸射影が重なる場合の手法である。ここでは帳票の入力時に帳票の傾きを算出し、それをパラメータにして文字領域を検出する。この手法は印刷文字に限らず、手書き文字の場合でも、認識精度を上げる手段として良く用いられている。

(文字の切り出し)

文字の切り出しは対象文字の外接枠の位置を検出する処理に相当する。図-5は左右端の検出例であるが、図-5(a)は横軸射影が連続、図-5(b)は不連続の場合を表わす。図-5(a)は図-4(a)と同じように、あらかじめ設定した検索ウィンドウによって文字を切り出す。図-5(b)では文字ピッチ p 、文字内ギャップ g 、文字中心座標 X_c 、さらに、変動可能範囲 a を用いて文字を切り出す。実際には $g \ll p$ ならば g を無視するだけで良いが、そうでない時はつぎのようにして左右端を検出する。

$X_c \pm A$ の位置が黒点でない：この範囲内にある黒点列の左右端を求める。

$X_c + A$ 、 $X_c - A$ のいずれか、あるいは両方の位置が黒点である：黒点になっている方ではその最外側端を求め、黒点になっていない方ではその内側を調べて黒点列の端を求める。

(2) 正規化

文字認識における正規化は主に、文字の生成過程で生ずる変動分を除去する技術である。変動の種類としては位置、大きさ、太さ、回転、濃淡、および字体や字形などがあげられるが、今日の文字認識技術では字体や字形を除き、ほとんどの変動に対処できるといわれている。こうした変動に対し、現在良く使用されている正規化手法をまとめると
位置：重心や外接四辺形の中心を求め、それを基準位置に合わせる。

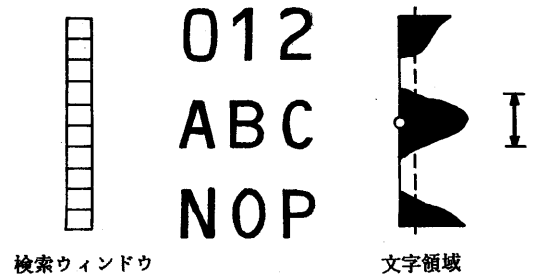
大きさ：重心まわりの2次モーメントや外接四辺形領域を算出し、それを基準サイズに合わせる。

太さ：細めや太め操作によって線幅を一定値に合わせたり、あるいは細線化する。

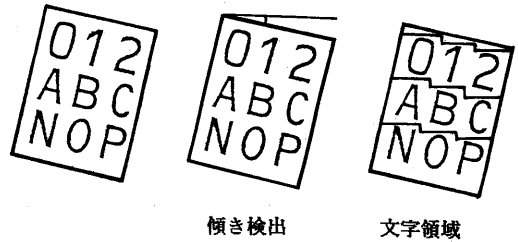
回転：主軸を座標軸に合わせたり、重心線を基準値に合わせる。

濃淡：最大濃度値や平均濃度値を算出し、それを基準値に合わせる。

などがあげられる。

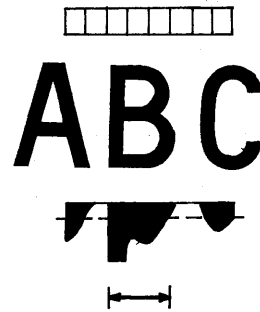


(a) 重畳しない場合

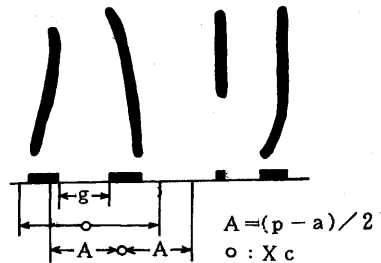


(b) 重畳する場合

図-4 文字行の検出例



(a) 連続パターン



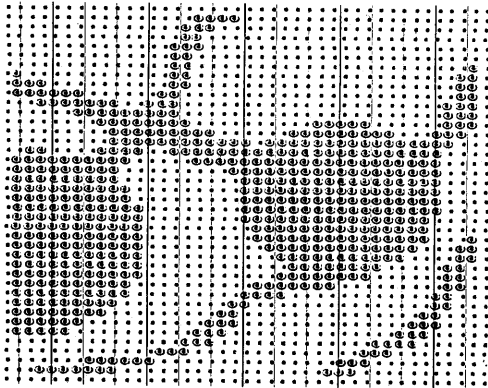
(b) 不連続パターン

図-5 文字の切り出し例

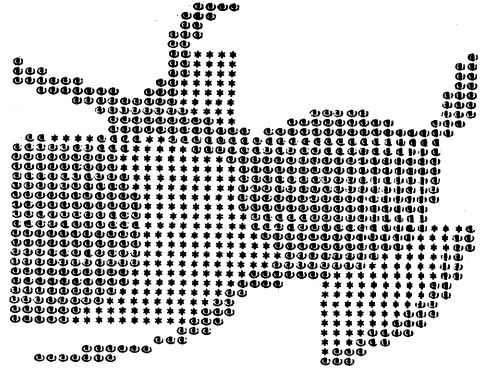
3. 1. 2. 図形

図形認識の前処理において、正規化はほぼ文字認識と同じであるが、セグメンテーションではかなり高度な技術が必要になる。この理由は図形認識の場合、対象物が常に孤立しているわけではなく、対象物以外の要素と接続したり、あるいは他の対象物の一部分が隣接したりするからである。これまでのところ、こうした悪条件でのセグメンテーション技術はまだ確立していないが、本稿では筆者らが試みた一例を紹介する。⁴⁾

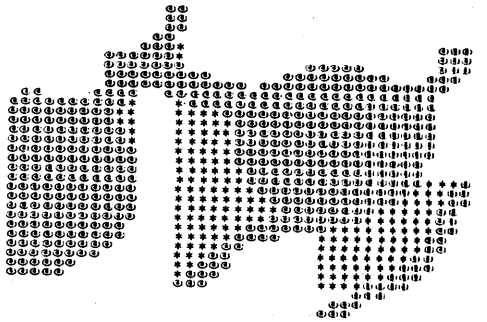
図-6 は一定の視野で捉えた図形の中から、黒ぬりされたパターンのみを抽出する処理を表わしたものである。この処理は原図形に対して、水平方向と垂直方向の各走査を交互に繰り返しながら逐次線分を除去し、最終的に残る部分をパターンとして抽出する構成になっている。特にここでは各段階の走査時に、黒点間を擬似的な点(図-6 中*印)で埋めた図形を生成するが、この生成のねらいは“線分は細い”、“対象物は塊まり”という概念を増幅させ、線分と対象物の差異をできるだけ大きくすることにある。尚、図-6 (d) の段階では複数個の塊まりが検出されるが、この例では単純に面積が最大なものをパターンとして抽出する。



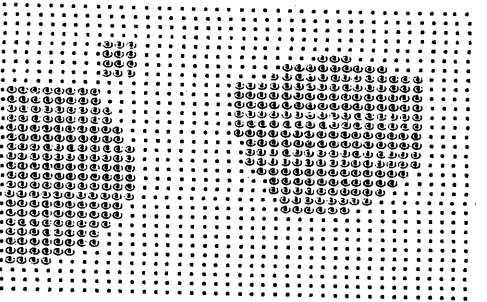
(a) 原 図 形



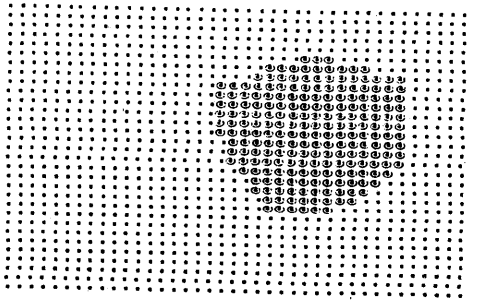
(b) 擬似点の生成



(c) 線分の除去



(d) 塊まりの抽出



(e) パターンの抽出

図-6 図形のセグメンテーション例

3. 1. 3 音声

(1) セグメンテーション

音声認識におけるセグメンテーションの方法は、認識の基準(単位)をどこに定めるかによって異なる。今日実用化が進んでいる単語単位の認識では、単語の始端と終端を見出す技術に相当し、まだ研究段階にある音韻単位の認識では、時間的に連続な音声波形を離散的な音韻に分割する技術に相当する。

〈単語単位〉

図-7は離散単語認識で良く用いられる始端、終端の検出法である。この方法ではあらかじめ音声パワーに対して二つの閾値(Th0, Th1)を設ける。一般にこれらの閾値は背景の雑音レベルを基準にして決められるが、Th0は母音の欠落、Th1は子音の欠落が生じない程度に設定される。実際の処理ではまず、Th0を用いて確実な音声区間(Ts1~Tel)を検出する。つぎに、Th1を用いて始端と終端を抽出するが、始端はTs1を始点として時間軸を逆上ったときに、音声パワーとTh1とが一致した位置Tsになる。終端はTelを始点として時間軸を進めたときに、音声パワーとTh1とが一致した位置Teになる。尚、この方法だけで正確な位置が決定できない場合もあり、今日では二組の閾値を用いて複数の音声区間を検出し、それによって認識性能の向上を図るような提案もある。⁵⁾

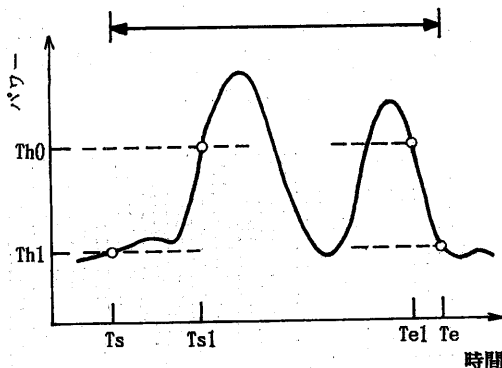


図-7 単語のセグメンテーション

〈音韻単位〉

これは連続音声認識研究の重要テーマの一つであるが、実現は極めて難しい。この理由として、人間のもつ発声器官の連続的な運動によって生ずる音声波形が、調音結合の影響により連続的に変化することなどが指摘されている。この分野の研究はまだ基礎検討の段階にあるが、これまでに、音声波形を無音、有声、無声区間に分類して母音や子音を抽出する、あるいは音声波形を一定間隔(フレーム)ごとにラベル化し、そのラベルによってフレーム間を合成するような検討が進められている。⁶⁾

(2) 正規化

音声認識における正規化は文字や図形と異なり、周波数や時間的な次元で生ずる変動を吸収する技術である。今日の音声認識では主に、声質と長さなどの変動が扱われている。

声質の変動は人間のもつ声帯や声道長の相違によって生ずるが、音声認識システムにおいて、声帯による変動はスペクトルの傾斜のばらつき、また、声道長による変動はホルマント周波数のばらつきとして現われる。現在、スペクトルの傾斜に関しては周波数軸上でのDP整合法の適用、あるいは調音パラメータ上での声道長補正などが試みられている。しかし、いずれもまだ完全な正規化が実現できる段階ではなく、今後の研究成果が期待されている。¹⁷⁾

これに対し、長さの変動はDP整合法によってほぼ解決したといわれているが、装置化に際しての演算量や辞書容量などに問題を残している。そのため、今日ではDP整合法の簡略化をめざした研究も進められている。

3. 2 画像処理

画像処理はOAやFAに代表される今日の情報化社会の担い手として精力的な研究が展開されており、その対象も宇宙観測から地下資源探索まで様々である。それに伴い、技術開発の内容は、人間の目視処理の自動化はもちろん、目視では不可能と思われるような処理の実現をめざすなど実に幅広くなっている。

こうした状況のもとで、画像処理における前処理を一定の論理に沿って分類することは大変難しい。しかし、これまでに、画像処理を応用分野ごとに区分し、それぞれの分野の中で前処理を捉える、あるいは画像処理を画像工学とパターン認識に分け、前処理を画像工学の中で捉えるといった報告がある。⁸⁾⁹⁾ いずれの報告も前処理を画像処理技術全体の中で捉えているわけではないが、基本的には誤っていないと思う。そこで、こうした報告を参考しつつ、

(1) 情報の再現、形成

- ・知覚面……雑音除去、幾何学的補正、画像の復元、
画像の変換、濃度補正

- ・感情面……画像の強調、画像の合成

(2) 認識のための処理

- ・空間上……雑音除去（狭義）、画像の変換（狭義）、
幾何学的変換

などになる。現在、これらの技術を支える技法に関しては既にSPIDERなどで解説されているが、¹⁰⁾ 以下に、情報の再現例として画像の復元、画像の強調、また、認識のための処理例として画像の変換モデルを付記する。

3. 2. 1 画像の復元

図-8は画像の復元モデルであり、このモデルは送信者の情報 $F(x, y)$ が伝送系 $H(u, v)$ で劣化したときに逆フィルタ $M(u, v)$ で復元することを表わす。このような処理はリモート・センシングや空中探査などの画像において、大気散乱、光学系の歪などの補正に良く使用される。

3. 2. 2 画像の強調

図-9は画像の強調技術の一つであるヒストグラム変換モデルである。ここでは出力画像のヒストグラムが人間の持つ視覚特性 $Z(g)$ の逆関数 $P_g(g)$ になるように入力

画像を変換するので、画像は人間の感性に強く捉えられるようなコントラストになる。この処理は露出不足の写真画像やX線画像のコントラスト強調などに良く利用されている。

3. 2. 3 画像の変換（狭義）

認識に直結した画像の変換としてはエッジ検出が有名である。エッジ検出手法は数多く存在しているが、すべての基本は図-10のモデルになる。尚、現在では検出されたエッジから線や面を抽出しようとする研究も進められているが、まだ、すべての画像に対処できる技法はない。

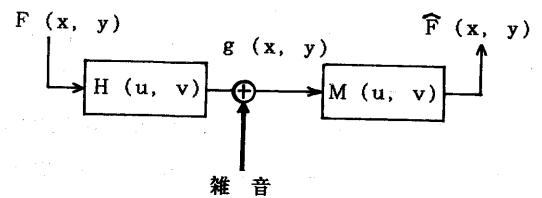


図-8 画像の復元モデル

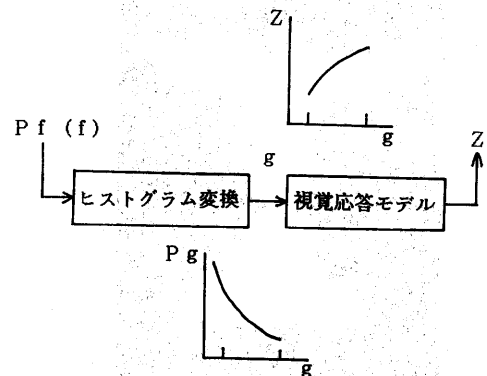


図-9 画像の強調モデル

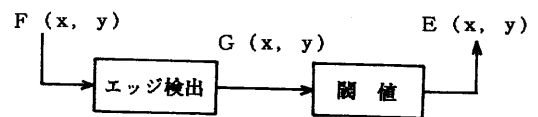


図-10 画像の変換モデル

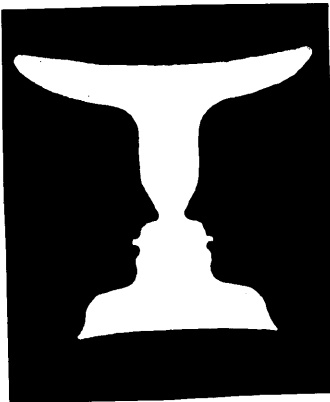
4. むすび

以上本稿ではデジタル信号処理における前処理に関し、システム面からの定義付けを行なうとともにその実例をパターン認識と画像処理で紹介した。

特に本稿において筆者らは、デジタル信号処理を送信者と受信者の意思伝達手段とみなし、送信者側からみた前処理をとりあげた。しかし、実際には受信者側からみた前処理も存在する。例えば、認識の過程においては受信者の意思は無視できないものであり、そこから生ずる前処理としては、識別のためのモデル設定、あるいはデータベースの構築などがあげられる。

このように前処理は信号処理の立場をどこに定めるかによって様々な形態が考えられ、ちょうどRubinの花びんみたいなものになろう。

おわりに、日頃ご指導いただく山田部門長、佐藤部門長代理、ならびに今回の資料作成にご協力いただいた岩田、山本、奈良、後藤、木島、井上の諸氏に深謝する。



参 考 文 献

- 1) R.L.ノイス：マイクロエレクトロニクス，別冊サイエンス“マイクロエレクトロニクス”，pp. 6-15 (1979) .
- 2) 木戸出，坂上：パイプライン方式と完全並列処理方式が増えた最近の画像処理装置，日経エレクトロニクス，1982年7月19日号，pp. 179 -212.
- 3) 有本著：現代情報理論，電子通信学会，pp. 1-5.
- 4) 松浦，吉田，岩田：地図上の対象物の認識，昭和55年度電子通信学会全国大会論文集，3-341 (1980) .
- 5) Y.Nara,K.Iwata,J.Tanahashi : Large Vocabulary Spoken Word Recogniton Using Simplified Time-Warping Patterns, ICASSP82, pp. 1266- 1269 (1982) .
- 6) 中川：連続音声の認識および会話音声の理解，音学誌 Vol.31, No.5, pp. 223 -231 (1981) .
- 7) 中嶋：デコンボリューションによる声道形の推定と適応型音声分析システム，音学誌，Vol.34, No.3, pp. 157 - 166 (1978) .
- 8) 尾上編：医用画像処理“画像処理システム”，朝倉書店，pp. 10- 15.
- 9) 尾上編：医用画像処理“画像処理ソフトウェア”，朝倉書店，pp. 23 -39.
- 10) 電子技術総合研究所：SPIDER USER'S MANUAL, 1980.