

PHIGS による音声解析システムの開発

中西 宏文・鈴木麻紀子

(情報科学選修)

Speech Analysis System using PHIGS

Hirobumi NAKANISHI and Makiko SUZUKI

(Department of Information and Computer Sciences)

1. はじめに

今日, 半導体技術の進歩によりコンピュータのハードウェアは目まぐるしい進歩を遂げ, それに伴いそれまで計算処理量等の制約から実現不可能だった非常に高度な機能をもつソフトウェアシステムも次々と開発されている。このような状況の下で, ワークステーション上の基本ソフトウェアも年毎に充実し, ハードウェアの性能を100%引き出すべく改良が加えられており, 今回, 取り上げる PHIGS もコンピュータグラフィックス用ソフトウェアとして開発が行われたものである。今回は, このグラフィックスソフトウェア PHIGS の概略・使用法に関し解説するとともに, これを用いてワークステーション上で開発した音声解析システムについて述べる。

2. PHIGSについて

2.1 標準グラフィックスシステム

1984年に国際標準化団体 ISO (International Standard Organization) によって GKS (Graphical Kernel System) が標準のグラフィックスシステムとして採用され, 1985年になると, ANSI (American National Standards Institute) も標準として採用し, 急速に広まった。GKS は汎用は2次元グラフィックスシステムであり, 3次元への拡張も提案されたが標準化されるに至っていない。

これに対して, PHIGS (Programmer's Hierarchical Interactive Graphics System) は, 1982年に ANSI が, それまでの GKS で実現困難なグラフィックス分野に対応するものとして検討を開始したもので, グラフィックデータを階層化し動的な図形編集を可能とし, 高度な対話性を持っている。これによって図形データの要素単位での追加・削除・置換ができるようになっている。

2.2 PHIGS と GKS の比較

GKS は表示機能を中心にシステムが設計されているのに対し、PHIGS はモデリング機能を中心に設計されているのが大きな違いで、PHIGS では GKS の座標系にモデリング座標系が追加されている。

また GKS のデータ構造はセグメントと呼ばれるものの単一レベルであるのに対して、PHIGS では、グラフィックモデルはブロックを組み立てて全体を構成するメカニズムを採用している。このメカニズムをストラクチャと呼んでいる。PHIGS は基本要素やその属性によってではなく、基本要素を組織し操作する方法によってその描画機能を発揮するようになっている。

さらに、GKS のセグメントは属性の一部が一括変換できるだけで、作られたセグメントは原則的に変更できないが、PHIGS ではストラクチャの一部でも随時変更できる。また、PHIGS が GKS と異なる点の1つとしてデータ機能と表示の分離がある。これは、CSS (Central hierarchical Structure Stone) に図形データを格納することと、CSS 内の図形データをワークステーションに表示することが別個に制御されるということである。GKS の基本要素と属性の結合は一旦定義されると変更には複雑な手続が必要だが、PHIGS では出力基本ストラクチャは表示されるときだけ出力要素となるので、グラフィックスのデータの修正が容易にできる。

ただし、PHIGS は GKS で培った多くの概念を利用しており、基本的なプリミティブやその属性タイプ、論理入力装置モデルとワークステーションの概念は GKS のものに含まれている。

2.3 階層構造

PHIGS の大きな特徴である階層構造では、図形データは、まず初めに CSS と呼ばれるネットワーク形式のデータ構造に格納される。CSS の構成単位は前述したストラクチャである。これは、幾何学モデルの一部または全部を表現するものである。ストラクチャ内の各図形要素がストラクチャ要素である。

1つのストラクチャは、EXCUTE STRUCTURE という関数を用いることによって、他のストラクチャを作動させることが可能である。この関数の要素には、参照されたストラクチャの識別子が含まれる。この参照された関数は、次々に他のストラクチャを参照でき、このようにして作られたネットワークは CSS に蓄積される。

ストラクチャネットワークを完全に定義してから、ワークステーションに表示することをポスティングという。PHIGSでは、データ機能と表示機能が分離しているため、基本要素と属性の結合は表示されるときだけ出力要素となるので、データの修正が容易に行える。

2.4 出力基本要素と基本要素属性

PHIGS によって作成されるグラフィックスの出力は、出力基本要素と基本要素属性で構成されている。

出力基本要素は、ストラクチャのトラバーサル時に、2次元ないし3次元の線やマーカー・ポリゴンなどを生成する。PHIGS の出力基本要素としては、表1に示す7種類がある。

基本要素属性として、(a)幾何学性、(b)非幾何学性、(c)ビューイング、(d)識別子の4種類がある。これらについて簡単に述べる。

表 1 出力基本要素

基本要素名	定 義
POLYLINE	点列で定義される一組の結合線分
POLYMARKER	与えられた位置を中心とするある型のシンボル（群）
TEXT	与えられた位置に表示される文字列
FILL AREA[1]	単一の多角形の内部領域や各種の模様で塗りつぶされた領域で境界線を持たないもの
FILL AREA[2]	一組の多角形領域で各種模様で塗りつぶされている。境界線の有無は問わない。
CELL ARRAY[2]	個々にカラーを持つセルの2次元配列で、4辺からなるポリゴンの格子。
汎用描画プリミティブ	スプライン曲線・円弧・楕円弧等の特殊な幾何学出力を目的とする非標準プリミティブ

(a)幾何学性

幾何学性の属性は、図形そのものの形状や大きさに関係する。文字列中の文字や領域を埋めるパターンの大きさがその例である。また、この属性は独立的に設定可能である。

(b)非幾何学性

この属性は、図形の表現法や属性の構成している部分の形状・大きさに関する属性であり、線画・線の色・マーカー列の大きさ等である。

(c)ビューイング

この属性は、各ワークステーションに対応したビューイングや隠れ線・隠れ面消去のパラメータを選択する。具体例を表2に示す。

(d)識別子

識別子属性は、目に見えるものやピックできるものの形状を決定する。これには、ピックの識別と名前セットの2種類がある。この属性のうち、非幾何学性属性のみ制御には束表を用いる。例えば、SET TEXT REPRESENTATION（文字列の位置合わせの方法）という関数のパラメータを見ると、文字の高さについての引数がない。これは、文字の高さの幾何学性属性が、SET CHARACTER HEIGHT（文字の高さ）という幾何学性の層性によって設定されているからである。つまり、文字に関しては束表が使用されワークステーションにその内容が蓄積され、一番最後に設定された値によって非幾何学性の層性がセットされる。この他にも、領域群のエッジ・内部・折れ線・マーカー列にも束表が用いられる。

3. 座 標 変 換

グラフィックデータを構築し、実際にディスプレイ上に表示するために、PHIGSでは、5種類の座標系を持ち、これらの座標系の間で変換を行うことで、座標系間の移動が行われる。これらの座標系の関係は、図1のようになる。以下、各座標系について述べる。

(a)モデリング座標系

この座標系は、3次元空間を記述するためのもので、使用するワークステーションに

表2 ビューイングの属性一覧

View index	ビューを指定するための指標
HLHSR identifier	隠れ線・隠れ面消去の識別子
Face culling mode	表面を選択するモード
Light source state	輝度の状態
Depth cue index	奥行き量に応じて輝度を変化せる手法の指標
Colour mapping index	色のマッピング
Eyepoint index	視点の指標
Colour environment	色付けの環境
Dithering value	ディザ法の値
Set convexity flag	凸面体のフラグ設定
Highlight index	輝度強調の指標
Antialiasing mode	Antialiasing のモード

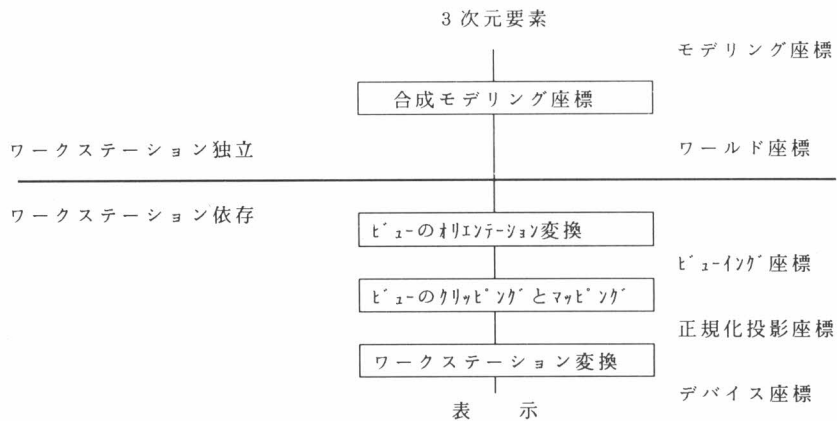


図1 PHIGS の変換パイプライン

依存しない座標系である。複数の図形（オブジェクト）を描きたければ、それぞれ別個にモデリング座標系を用意し、各座標系の中でそれぞれ図形を定義すればよい。

(b)ワールド座標系 [WC]

ワールド座標系もモデリング座標と同じく3次元でデバイス独立の座標系であり、オブジェクト間の相対的関係を表示する。

(c)ビューイング座標系 [VRC]

ワールド座標系をビューイング変換によってマッピングしたもので、デバイス独立でワークステーション依存の座標系である。この座標系は原点としての視参照点 (VRP) を定めて、UVN 空間を定義する。これは、幾何学で一般的に使用する XYZ 座標系と類似の物で、ビューアップベクトと呼ばれるベクトルでV軸方向を定義し、視平面法線ベクトルは視平面に対して垂直で、このベクトルがN軸方向となるものである。

(d)正規化投影座標系 [NPC]

この座標系は、VRC をビューマッピング行列によって変換したもので、全体の長さを

1 になるように正規化されており、デバイス独立でワークステーション依存の座標系である。

(e) デバイス座標系 [DC]

デバイス座標系は、NPC をワークステーション変換によって変換した座標系である。まず、図形の変換を可能にするためにワールド座標系とは別の新たな空間を用意するが、これをデバイス空間と呼ぶ。図形は、まずワールド座標系で定義されデバイス座標系に変換され、その結果が実際に表示される。

5. PHIGS の論理入力装置

PHIGS では、アプリケーションプログラムが論理入力装置を制御して図形の入力を行う。論理入力装置はワークステーション識別子、入力クラス、デバイス番号の3つの引数によって定義される。

入力クラスでは、入力する値の形式で決定され、位置入力・点列入力・実数値入力・選択値入力・ピック入力・文字列入力の6種類がある。

また論理入力装置は、動作モードと呼ばれる、要求・抽出・事象の3種類のモードによって操作される。

それぞれの論理入力装置には、(a)初期値 (b)トリガーとメジャー (c)プロンプトとエコータイプ (d)エコーエリアとエコー値 (e)プロンプトとエコータイプを含んだデータレコード 等が含まれる。以下、これらの概略について説明する。

(a) 初期値

それぞれの論理入力装置には、初期値の設定が必要である。これらの初期値は、ワークステーションをオープンする際にワークステーション記述テーブルから値を獲得する。

(b) トリガーとメジャー

論理入力装置の正確な動作を記述するために、図2に示すようなトリガーとメジャーという概念を使用した論理入力値のモデルが示されている。要求モードはトリガーの発火により応用プログラムに1つの値を返す。また、リクエスト機能の呼出しの間、対話状態が継続する。

抽出モードでは、トリガーは働かずサンプル機能呼び出す毎に値が返される。対話状態においてはサンプル機能を繰り返し呼び出せる。

事象モードでは、トリガーの発火により、値と入力装置の識別情報が待ち行列に送られ、AWAIT EVENT 機能により取り除かれる。このモードにおける待ち行列は、0個以上の情報をもつ FIFO 行列であり、AWAIT EVENT 機能とは待ち行列から以前の情報を取り除き新しい情報を得られる状態にしておくための機能である。

(c) プロンプトとエコータイプ

それぞれの論理入力装置は、最低1個のプロトタイプとエコーを持つ。プロンプトとは、論理入力装置を利用できるということをオペレータに示すもので、エコーとは、論理入力装置の“メジャー”をオペレータに直接知らせるためのものである。また、エコータイプとは、どのようにしてエコーを示すかという手法である。

(d) エコーエリアとエコー値

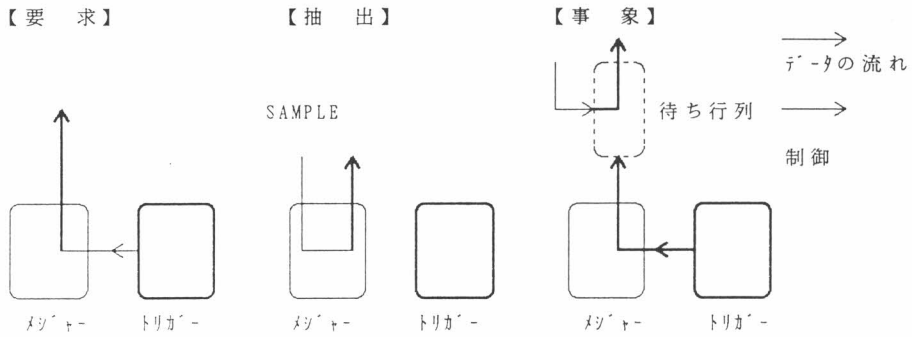


図2 論理入力値のモデル

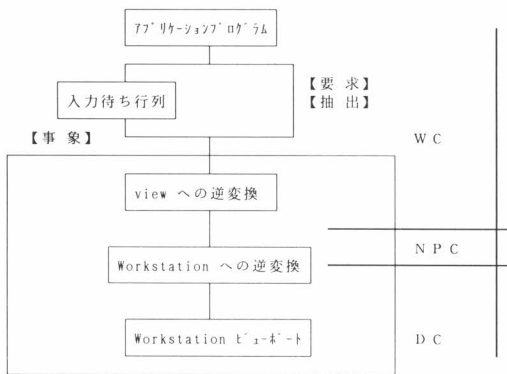


図3 座標入力の変換

ようにデータレコードに入っている。

ロケーターストックの入力クラスで得られる座標は、自動的に画面上の座標からワールド座標系に逆マッピングされる。これは、デバイス座標系から正規化投影座標系への逆変換も含む。図3にこの様子を示す。

6. ワークステーション上での音声解析システムの開発

近年、信号処理技術の進歩によりコンピュータを用いた音声処理システムの研究が急速に進み、音声合成・音声認識・話者識別・低ビット伝送等、応用分野が拡充している。このような研究を行う上で必要不可欠なのが、マイク等から音を入力してA/D変換を行ない、計算機上の主記憶に音を2値化データとして保存でき、それに対して様々な演算を行ない特徴抽出等の行えるシステムである。以前このためのシステムをパソコン上で開発したが、主記憶の関係から収集できるデータ量が300KB程度に制限され、またCPUとA/D、D/Aコンバータの能力から、一度に変換できるデータの上限が32767個までであったためそれ以上のデータを収集するには、変換操作を繰り返す必要がある、これに必要なCPUクロックサイクルの間のデータが欠落するといった欠点があった。また、各種の演算

それぞれの論理入力装置は、画面上のどこにプロンプトを表示させるかを明確にできる。エコーエリアは、座標値の最大・最小を用いて表す。
(e)プロンプトとエコータイプを含んだデータレコード

論理入力装置は、初期化する関数の中でプロンプトやエコータイプについて付加的な情報を得るためのアプリケーションを持つ。

例えば、付加的な情報の中でカーソルの色を決定できる。付加的な情報はこの

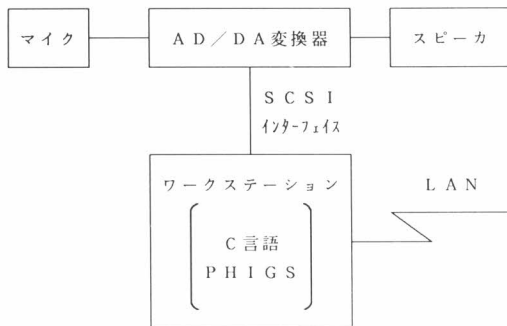


図4 ハードウェア構成図

処理を行う際にも演算時間の点で難点があり、こういった諸問題を解決するために、今回新たにワークステーション上で同種のシステム開発を行ったものである。

6.1 基本機能

今回開発したシステムのソフトウェア基本機能として、以下のものを備える。

(a)A/D, D/A 変換

この機能は、音声の入出力の際の基本となる機能で、マイクから入力した音のアナログ信号をコンピュータ内部で処理できるデジタル信号に2値化したり、またその逆の操作を行って音を再生するための機能である。ただし、実際の変換はワークステーションに外付けするA/D, D/A変換器によって行われるため、ここではその変換器を制御するためのドライバソフトウェアをコールできる機能を持つものをソフトウェアとしては用意する。

(b)ファイル入出力

A/D変換されたデータは、計算機のメモリー上に2進数データとして展開されているが、このデータを永久的に保存しておくためにハードディスク上のファイルに書き出したり、ハードディスク上のデータを読み込んだりするのがこの機能であり、また特徴分析の行われた結果等も出力できる機能を有する。

(c)ラベリング

人間の音声等は、通常複数個の音を連続して発声することで1つの文を構成し意志の伝達を行う。しかしながら、音声分析を行う場合には逆に連続した発声の中から1発声単位であるところの音節や音韻、または単語・文節等を抽出して扱いたい場合が多い。このような場合に、機械的に自動で区切りを抽出する自動ラベリングの研究も行われているが、基本としては、音声波形やスペクトル特徴を表示して、マウスカーソルにより視察で目印をつける必要がある。これを実現するための機能がラベリング機能である。

(d)特徴抽出

音声処理を行う際には、A/D変換された原波形そのものを扱うのではなく、周波数軸上のスペクトル分布を基にしたパラメータに変換して、認識等を行う場合がほとんどである。このため、本システムでも原波形に対し、FFTスペクトル係数・LPCメルケプストラム係数等のパラメータを求める関数を付加した。

6.2 ハードウェア構成

本システムを構成するハードウェアは、独立型のAD/DAコンバータ、PHIGSを利用できるワークステーションと、マイクアンプ・スピーカー等からなる。AD/DAコンバータはSCSIインターフェイスに接続可能なもので、接続するワークステーションに応じたドライバソフトが提供されている。構成の概略を図4に示す。

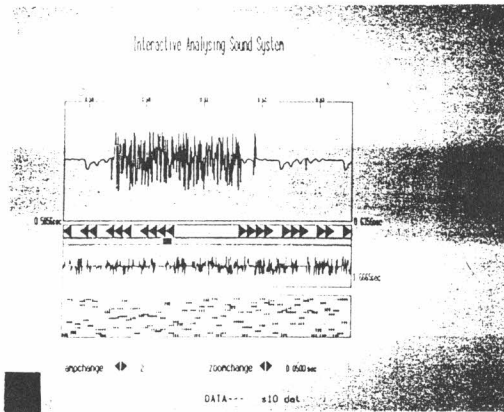


図5 音声解析システムの画面表示

6.3 操作環境

本システムではワークステーション上のディスプレイに各種情報を表示し、マウスを使用してプルダウンメニュー形式により機能を選択しながら操作を行っていくものとした。画面の一例を図5に示す。この図は、ある音声データがメモリー上に展開された後、FFT分析を行ったもので、図の上から部分波形・スクロールバー・全体波形・全体波形に対するFFT結果が表示されている。また、その下のマークをマウスボタンでクリックすると表示されている部分波形の振幅幅や表示時間の変更が行えるよう

になっている。

7. ま と め

今回、業界標準となるべく3次元グラフィックスシステムであるところのPHIGSに関しての概説とPHIGSを利用した音声解析システムについて述べた。PHIGSに関してはソフトウェアメーカの提供する英文マニュアルしか資料がなく、まだまだ今回記述した内容以外にも豊富な機能が利用できるが、それらに関しては必要に応じてマニュアルを参照していただきたい。

また、残念ながら業界標準となるべくPHIGSであるが、他のソフトウェアやOSの例にもれず、メーカー毎に機能拡張を行ったため、ワークステーションのメーカーが異なると互換性が保てない、といった状況になっている。さらに、1991年頃からは新たな試みとしてPEX (Phigs Extention to X) と呼ばれるソフトウェアが登場し、移植性の点でPHIGSよりも優れているために、今後の流れとしては、2次元システムはX libで、3次元システムはPEX libを用いて記述するといった状況になりつつある。今後は特殊な用途を除いては、この流れの中でグラフィックスアプリケーションを開発することが望ましいものと思われる。

(平成5年9月13日受理)

参 考 文 献

- [1] “A Beginner’s Guide to Using HP-PHIGS” Hewlett Packard 社
- [2] “HP PHIGS Graphics Techniques Volume 1, 2” Hewlett Packard
- [3] “HP PHIGS FORTRAN Binding Reference Vol 1, 2” Hewlett Packard
- [4] 伊福部達 “音声タイプライタの設計” CQ 出版社
- [5] 齊藤収三・中田和男 “音声情報処理の基礎” オーム社