

液晶ペンタブレットを用いた音の生成システムの試作

A Prototype of Sound Generation System using LCD Tablet

出口 幸子
Sachiko Deguchi

高林 明広
Akihiro Takabayashi

近畿大学工学部
School of Engineering, Kinki University

This paper describes a prototype of sound generation system controlled by pen input on LCD tablet. This paper shows the result of the experiment on pen pressure. The standard deviation of pen pressure varies depending on user's experience. The appropriate time to get pen pressure is 200 to 500[ms], however, this system uses the pressure at 100[ms] to generate sounds. Next, this paper shows how to generate digital sounds using pen input on Windows. The pen pressure controls the volume, the y-coordinate of pen position controls the pitch, and the x-coordinate controls the timbre. When the pen input is detected, WAVE data is generated by synthesizing sine waves. Users can specify the ratio of harmonics and the decay function.

1. はじめに

液晶ペンタブレットは、画面に直接入力することで自然な操作が実現できるため、現在種々の分野で実用的に使用されている。一方、研究においては、文字や数式の認識方式の研究、ユーザインタフェースの研究[加藤 98]、等が行われている。また、音楽インタフェースとしても使われるようになって来ており、例えば MAX[IRCAM]を使った作品製作に用いられている。本稿は、液晶タブレットを音楽インタフェースとして使用するための基礎的な課題について報告する。

2. 液晶ペンタブレット

本研究では、ワコムの液晶ペンタブレット CintiQ C-1500X を使用し、ペンの位置情報と筆圧情報を利用する。CintiQ の筆圧は 512 レベルだが、タブレットドライバから情報が渡されたときの筆圧の最大値は 1023 である。今回は USB 接続としたので、93 点/秒の筆圧情報を得る。筆圧は、入力荷重そのままの値ではなく設定が可能であるが、本研究ではデフォルトの設定(入力荷重と出力筆圧がほぼ比例)を用いている。

2.1 タブレットプログラミング

開発環境は Microsoft Windows XP 上で Microsoft Visual C++ を使用している。タブレットインタフェースとして、米 LCS/Telegraphics 社が開発したタブレット用の API 群である WinTab を使う[LCS]。本研究では、入力を知るために、タブレットから WT_PACKET メッセージを送信させる方法を使用する。ウィンドウプロシージャが WT_PACKET メッセージを受信したら WTPacket 関数を使い、バッファを取り出す。本研究では座標、筆圧、およびパケット発生時間を取得している。

2.2 筆圧の計測

パケットの発生時間を取得する実験をした結果、正確に 1 秒間に 93 回サンプリングしていることが分かった。目的の筆圧に到達するまでに時間が掛かるので、どの時点の筆圧を用いればよいかを検討するため、筆圧を計測する実験を行った。

2.2.1 目的筆圧の指定

連絡先: 出口幸子, 近畿大学工学部電子情報工学科, 広島県東広島市高屋うめの辺1, deguchi@hiro.kindai.ac.jp

(1) 筆圧の平均値

入力の目的筆圧を 512(最大値の半分)と指定して、入力開始から 3 秒間、サンプリング点毎に筆圧を計測する実験を行った。図1は 400 回試行して時間と平均筆圧の関係を示したグラフである。表1に 100, 200, 300, 400, および 500[ms]時点(各々10, 20, 29, 38, 48 点目のサンプリング)の筆圧の平均値を示す。400 回試行と 100 回試行の値を比較すると大きな違いはないので、平均値を得るには 100 回試行で十分と考えられる。

同実験を異なる 3 名が 100 回試行し、筆圧の時間変化をグラフにしたところ、人によって入力の癖が若干あることが分かった。しかし、何れも約 500[ms]で目的の筆圧に達しており、また、約 200[ms]位からグラフの傾斜が緩やかになっている。この結果から、アプリケーションに応じて 200~500[ms]の筆圧値を取ればよいことが分かる。

(2) 筆圧の分布と分散・標準偏差

筆圧の分布を知るためにヒストグラムを作成した。400 回試行の 100, 200, 300, 400, および 500[ms]時点におけるヒストグラムを作成したところ、単峰で左右対称に近い形になった。

一方、ばらつきの尺度として分散と標準偏差を求めた。表2は 400 回試行、およびそれを 4 分割した 100 回試行の各々について、100, 200, 300, 400, および 500[ms]時点の筆圧の標準偏差を求めた結果である。どの時点においても、標準偏差は、最初の 100 回試行と次の 100 回はあまり変化がなく、その次の 100 回と最後の 100 回では順に減少する。これは、被験者がペン入力操作に慣れて行く過程を示している。しかし、同一被験者がこの実験の前に行った 400 回試行では、標準偏差の減少は見られなかった。この結果は、ユーザインタフェース実験における再現性の難しさを示している。

また表2から分かるように、100~300[ms]の範囲では標準偏差は同程度で、400[ms]以上では標準偏差が減少して行く。

2.2.2 筆圧の感覚と平均値

先に述べた実験では、時間と筆圧との関係を分析するために、目的の筆圧を指定する方法を取ったが、実際にペンを使う場合、指定された筆圧を入力することは稀であろう。そこで、自分が弱い、中位、強いと思う筆圧で入力したときの時間と筆圧の関係を計測する実験を行った。3 名がそれぞれ 100 回試行した筆圧の平均値をグラフにした。1 名のグラフを図2に示す。筆圧はそれぞれ異なるが、約 200[ms]からグラフの傾きが緩やかになってい

る。他の 2 名についても同様の結果になった。この結果は、目的の筆圧を指定しない場合も、指定する場合と同様、200[ms]以降の値を取るとよいことを示している。

3. 音の制御

3.1 正弦波の合成

音を特徴付ける要素として大きさ(振幅)、音高(基本周波数)、音色(波形)がある。本研究では、正弦波を次式のように合成して音を生成する： $W(t) = P k g(t) \{ A_1 \sin(2\pi f t) + \dots + A_i \sin(2\pi f_i t) + \dots \}$

(1) 振幅

振幅は筆圧に比例させる。音を鳴らすためには、入力から200[ms]後の筆圧を用いると遅れが目立つので、100[ms]後の筆圧を取ることにした。但し100[ms]時点では前述したように標準偏差が大きく、筆圧が十分に取れない場合がある。100[ms]時の筆圧値の標準偏差が約100であることから、その2倍の200を計測値に足して補正し、その値をPとして音の生成に用いる。一方、倍音や他の成分の正弦波を足すと振幅が大きくなるので、音量を調整するための係数kとして、 $C/\sum A_i$ を用いる。ここで A_i はi倍音の振幅、Cは定数である。

(2) 基本周波数

本研究では、画面の Y 座標に 3 オクターブの幅で、音階を構成する音(ド、レ、ミ、ファ、ソ、ラ、シ)を並べている。ユーザが画面上の音をペンで指定することで基本周波数が決定する。音階上の各音の周波数は音律により規定される。本研究では、平均律とピタゴラス音律の何れかをユーザが選択できる。

(3) 波形

倍音は10倍音まで合成でき、それぞれの割合を設定できる。i倍音の割合は整数のみでなく、小数第4位まで指定できる。例えば、基本周波数に1.26(2^{4/12})倍の音、1.5(2^{7/12})倍の音を合成すれば、三和音が生成できる。また、基本周波数に加え1.01倍の音などを指定すれば、うなりを表現することができる。

合成した波形に減衰関数を掛けることにより減衰を実現する。本研究では次の関数をユーザインタフェースで指定できる。

- ① $g(t) = \exp(-at)$
- ② $g(t) = 1 [0 \leq t \leq b], g(t) = 1 - a(t-b) [b \leq t \leq b+1/a]$
- ③ $g(t) = 1 [0 \leq t \leq b], g(t) = \exp\{-a(t-b)\} [b \leq t]$
- ④ $g(t) = 1 [0 \leq t \leq b], g(t) = \sqrt{1 - (t-b)/a} [b \leq t \leq b+a]$

3.2 ユーザインタフェース

音を制御するために作成したユーザインタフェースと設定用ダイアログを図3に示す。音の大きさ(振幅)、音高(基本周波数)、音色(波形)は、タブレット入力ではそれぞれ筆圧、Y座標、X座標に対応している。音色は図3に示すダイアログを表示して設定する。ダイアログの領域Aで、音の長さ、音律、および減衰関数を設定する。領域Cで倍音と他の成分を指定し、領域Bでそれらの振幅の割合を設定する。

3.3 WAVE データの作成と再生

本研究で試作したプログラムでは、入力がある度に計算してWAVファイル形式のデータを作成し、メモリ上のサウンドイメージとして音を鳴らす[田辺 01]。前述した方式で正弦波を合成し、それをサンプリングしてWAVEデータを作成する。基本周波数の最大は1047Hzで、10倍音まで合成するので周波数の最大は10470Hzであるから、サンプリングレートを24KHzとしている。

WAVE再生の手順を次に記す。①再生デバイスをオープン

し、②デバイスとバッファを結びつけ、③バッファを再生デバイスに書き込む。④ヘッダを開放して、⑤再生が終わっていれば再生デバイスをクローズする。本研究では同時発音数(デバイス数)を10としている。なお、マルチバッファリングを使えば、音を途切れなく再生させることができるし、負荷を分散できるので、本研究ではトリプルバッファリングを用いている。

4. おわりに

液晶タブレットを音楽インタフェースに使用するための基礎的な課題を研究した。筆圧の時間変化を計測する実験を行い、200~500[ms]の筆圧値を取ればよいことが分かったが、音の生成には100[ms]の値を用いた。一方、正弦波の合成により音を生成する方式について検討した。振幅、基本周波数、波形は、それぞれペンの筆圧、Y座標、X座標で制御した。ユーザが倍音と他の成分および減衰を指定できるようにした。

今後の課題としては、入力補正の再検討、およびペンによる音の長さの制御が挙げられる。さらに、波形合成における物理モデル、録音データの利用、等も検討したい。また、本システムを旋律合成に応用することも検討する予定である。

参考文献

- [加藤 98] 加藤直樹, 中川 正樹: ペンユーザインタフェースの設計のためのペン操作性の検討, 情報処理学会論文誌, Vol.39, No.5, 1998.
- [IRCAM] IRCAM: <http://freeware.ircam.fr/>
- [LCS] LCS/Telegraphics: <http://www.pointing.com/Resources.html>
- [田辺 01] 田辺義和: Windows サウンドプログラミング, 翔泳社, 2001.

表1: 筆圧の平均値

	100[ms]	200[ms]	300[ms]	400[ms]	500[ms]
001-100 試行	284.7	416.0	450.4	476.6	492.4
001-400 試行	273.9	396.6	437.0	470.3	490.2

表2: 筆圧の標準偏差

	100[ms]	200[ms]	300[ms]	400[ms]	500[ms]
001-100 試行	105.5	111.5	100.9	85.7	74.5
101-200 試行	111.8	121.7	109.3	94.8	83.8
201-300 試行	91.9	94.3	86.9	74.5	63.6
301-400 試行	76.5	73.4	67.1	57.7	46.6
001-400 試行	98.0	102.4	92.9	80.2	69.1

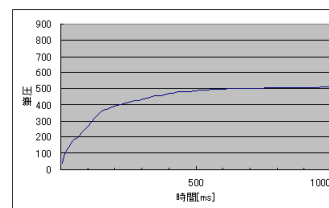


図1: 筆圧の時間変化(筆圧指定有)

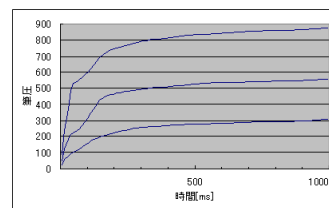


図2: 筆圧の時間変化(筆圧指定無)

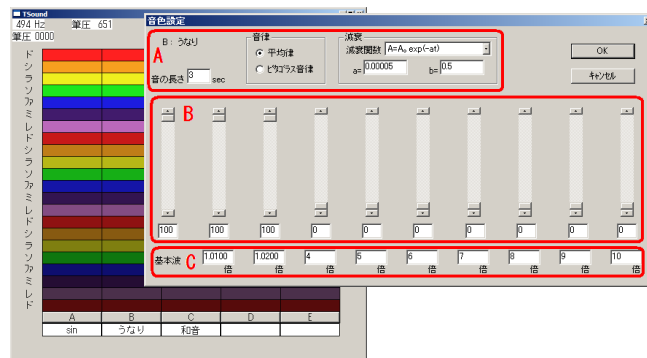


図3: 音の制御のユーザインタフェースと音色設定用ダイアログ