LINUXでサウンド

ねらい

この授業では Linux で音(サウンド)を解析・合成するアプリケーションの使い方を学 びます。最初にサンプルの音の確認とハードウエアの確認を行います。次に音の性質やコ ンピュータで音を扱うための必要事項を学習します。これらの基本的な事項を参考にしな がら、Audacity というアプリケーションを使って、音の解析や合成の仕方について学習し ます。

(0) 準備

 (a) lftp コマンドを使って ftp.cs.tohoku-gakuin.ac.jp から必要なデータをダウンロードします。ここでは、/TG-Local/staff/matsuo/DTM07 というディレクトリをそっくりそのまま (ディレクトリ構造をそのまま)、自分のホームディレクトリに持ってきます。

ヒント:lftp を起動した後、適切なディレクトリまで移動して、 mirror コマンドで DTM07 をダウンロードしなさい。

(b) ホームディレクトリの下に、DTM07 というディレクトリができたか確認をしなさい。

(c) DTM07 というディレクトリの中には1つのファイルが有るはずです。確認をしなさい。 ファイル名:____

(d) DTM07 というディレクトリに移動した後に、tar コマンドを使ってこれらの2つのフ アイルを展開(解凍)しなさい。ファイルは、xxx.tgz とか xxx.tar.gz という形式の名前 になっていますので、xvzf というオプションを tar で指定しないといけません。

...\$ tar xvzf ファイルの名前

 (e) どのようなディレクトリができましたか? dtm-samples という名前のディレクトリ ができていますか? できていれば、そのディレクトリへ移動しなさい。
 注意:できていない場合は、TA や教員を呼び確認してもらいなさい。

(f) dtm-samples というディレクトリの中に、さらに幾つかのサブディレクトリができているはずです。xmms というアプリケーションを使って、この中の音声ファイルを再生して見ます。以下のコマンドを実行して見なさい。

...\$ xmms ./CV/be.wav

(g)「べ」という音は聞こえましたか?

右上の×印をクリックすると、xmms を終了させることができます。

(h) 注意:この項目は音量を調整したいときに行います。調整を行いたいときは、以下の説 明をよく読んでから実行しなさい。

説明:alsamixer というアプリケーションを起動すると、音量の調整をする事ができます。 起動をすると、画面の左上に Card、Chip、View、Item という4つの情報が表示されます。 カーソルキーを左右に動かすと、音量調整をする項目を移動することができ、対象となっ ている項目が Item という所に表示されます。今回は、Master という項目を調整します。 音量の調整はカーソルの上下で行うことができます。alsamixer は、ESC キーで終了する ことができます。(g)と(h)を繰り返して音量の調節をしなさい。また、mを押すとミュート (消音)になります。

…\$ alsamixer (ここでは最後に & をつけてはいけません)

練習1 alsamixer を用いて音の大きさを調節しなさい

 (i) 今あなたは、dtm-samples と云うディレクトリにいるはずですが、ここには幾つかのサ ブディレクトリができているはずです。どのようなディレクトリができていますか?それ ぞれのディレクトリの中には以下のような音ファイルがあります。

 PureTone:
 純音(波形が正弦曲線をなす音で、時報の音や音叉の音のような音)

 Vowel200ms:
 母音(あ、い、う、え、お). 普通に発話した音声とささやき声

 CV:
 子音(ば、だ、が等)

sound-samples : 動物の音声

others: 他の音(単語や文など)

(j) xmms は wav ファイルや mp3 ファイル (xxxxx.wav とか xxxxx.mp3 という名前の ファイル) が再生できます。

練習2 上記のディレクトリの中の wav ファイルや mp3 ファイルを xmms を使って再生 させなさい。まれに、xmms で再生できないものもありますが、そのときはファイルの選 択画面が現れるのでキャンセルしてください。

(1) 音声ファイルについて

音声のファイルはさまざまな形式で作られています。代表的なものとしては、.wav や.mp3 といった拡張子をもったものがあります。前者は MS-Windows の標準形式です。 後者は人が聴き取れる音(可聴音)だけを残すような方法で圧縮された音声ファイル(フ ァイルサイズが小さい)です。Unixの標準形式としては au があります。

(2) 音の性質

音は空気の圧力の振動です。ものを叩くとそのものに接している空気も振動します。こ れが空気中を伝播して我々の鼓膜のところで、空気の振動により鼓膜が振動することで音 を感じることができます。

●音の大きさ

この音による空気の振動がマイクに達するとマイクはこの変動を電圧の振動に変換しま す。音の振動が大きいと電圧も大きくなります。この音の大きさ(振動の大きさ)は音圧 レベル (Sound Pressure Level) で表され、単位はデシベル (dB) です。

◆実環境の音の大きさ

人が聴き取れる最小の音圧:0 デシベル、 ささやき声:20 デシベル 通常の会話:60 デシベル、 混雑した街:80 デシベル、地下鉄内:90 デシベル dBの値が20 大きくなると音の振動の10倍になることに対応します。したがって、60dB の音とは0dBの音の振動に比べて1000倍大きくなることに対応します。

● 音の振動数(周波数)

音は空気の圧力の振動である。単位時間(一秒)あたりのこの振動の繰り返し回数を音の振動数(周波数)とよぶ。周波数を表す単位にはヘルツ(Hz)という単位が用いられる。 この音の周波数によって我々は音の高さを感じることができ、低い音は周波数が低く、高い音は周波数が高い。加えて、人間が耳で聞くことができる音の周波数は、一般に 20Hz から 20000Hz くらいである。参考までに書くと、テレビの時報の「ピッ、ピッ、ピッ、ポ ーン」の「ピッ」が 440Hz、「ポーン」が 880Hz です。このような一つの周波数のみを含 む音を純音と呼びます。ここで、880Hz の音は 440Hz の音より1 オクターブ高いとも言い ます。

ある音にどのような周波数成分が含まれているかを解析することを「周波数解析」と呼び、どのような周波数がふくまれているかを表したものを「スペクトル」とよびます。た とえば100Hzの純音の場合、音の時間波形は下図の左のようになります。100Hzの音です ので、1秒間に100回振動する波となります。ですので、一つの波の時間は0.01秒となり ます。この場合、スペクトルは下図の右のようになります。100Hzの振動数(周波数)し か含まれないので、100Hzのところだけ大きさをもつことになります。



図1.100Hzの音の波形とスペクトル

練習3 音の高さと音の周波数の関係を確かめてみよう。dtm-samples/PureTones 内にい ろいろな周波数の音を用意しています。xmms を使って、いろいろな周波数の音を再生し てみよう(xmmsの使い方は先週配布された資料を見てください)。

たとえば、ファイル名「t100Hz.wav」というのは、100Hzの周波数の音です。「t1kHz.wav」 というのは 1kHz(1000Hz)の周波数の音です。

練習4 自分が聞き取ることが可能な最も高い周波数を調べてみよう。聞き取れた周波数 には括弧の中に〇をつけてみよう。

tlkHz (), t2kHz (), t4kHz (), t8kHz (), t12kHz (), t15kHz (), t16kHz (), t18kHz (), t20kHz ()

(3) 音をコンピュータで再生・録音するには

サウンドボードの役割はこの電圧の変動をコンピュータで扱えるデータに変換すること です。電圧の変動は時間的に連続変化だし、電圧の値も連続量です。このような量をアナ ログ量といいます。一方、コンピュータが扱えるのは離散量です。こちらのほうはデジタ ル量といいます。したがって、サウンドボードの大きな役割はアナログ-デジタル変換に なります。

●録音の場合は、

音源 \rightarrow マイク \rightarrow アンプ \rightarrow A/D 変換 \rightarrow PC

●再生の場合は、

PC → D/A 変換 → アンプ → スピーカ → 空気中

音声信号をデジタルデータとして表現する方法には、さまざまな方法があるが最も基本 的なものとして PCM (Pulse Code Modulation)がある。PCM では、アナログ信号から一定 時間ごとに値を取り出し、その値を量子化(デジタル量に変換)することによりデジタル 信号に変換する。PCM の概略図を図に示す。PCM の音声の品質は、時間方向にどの程度 細かく値を取り出すか(サンプリング周波数)と取り出した値をどの程度の解像度で量子 化するか(量子化ビット)により表現される。例えば、音楽 CD ではサンプリング周波数 が 441000(ヘルツ)で、16 ビットの PCM が用いられています。この場合、一秒間に 44100 のサンプル点で記録され、各サンプルにおいて-32768 から 32767(2 の 16 乗=65536 点) の整数値となります。例えば、10000Hz のサンプリング周波数で、3 ビットで量子化(-8 から 7 までの値に符号化) した場合は次のような図で表せます。



図2.サンプリング周波数と量子化

(4) Audacity を使って音の解析と合成

Audacity の起動(Language は Nihongo を選択)



図 3.起動画面

(この図は Windows 用のものなので細かい配置等には異なる部分があるかもしれません)

(a) 音の読み込みと再生

ファイルにマウス・カーソルを合わせ、左ボタンを押すと選択項目の一覧がでてくる (PULL-DOWN-MENU プルダウンメニュー)。その中で import →audio を選択する。

音声ファイル(***.wav)を選択すると音が読み込まれ、下図のような新しい窓ができる。 ここでは、「vowel_a.wav」を読み込む。音を再生するには、再生ボタンを押す。波形の途 中から音を再生するには、再生させたい時間のところでマウスを左クリックし、再生ボタ ンを押す。この音声は 0.2 秒(200ms)で、母音の「あ」です。



図4. 音の読み込みをした結果

(b) 波形を詳しくみる

横軸の拡大 → 🌶 (ズームイン) 横軸の縮小→ 🌶 (ズームアウト) 窓にあわせるように横軸を調整する → 🍳 (Fit project in window) 練習5 ズームインした後に、「Fit project in window」ボタンを押す。ズームインした後

にズームアウトを行う。

縦軸の拡大 → 縦軸メモリを左クリック(縦軸のズームイン)

縦軸の縮小 → 縦軸メモリ右クリック(縦軸のズームアウト)

実際にやってみよう:縦軸の「0」の辺りにマウスを移動し、左クリックを行い、その後 に右クリックする。

(c) 波形の一部を選択する

波形 Window の所にカーソルを持って行き選択したい部分のはじめの所で左ボタンを押 し、押したまま選択したい最後の部分まで移動し、左ボタンを離す(選択したい部分をド ラッグする)。再生ボタンでこの選択部分を再生できる。

選択した部分だけ表示したいときは、 🚑 ボタン(Fit Selection)を押す。

練習6 横軸の 0.05 から 0.15 まで選択し、選択した部分のみを画面上に表示し再生する。 (d) 選択した波形の一部を wav ファイルとして保存する

編集のプルダウンメニューから コピー を選択する。ファイル のプルダウンメニューから「Export Selection」を選択し、ファイル名を入れて保存する。必ずファイルの種類を「wav」 にしてください。「保存」ボタンを押すと、「Edit the metadata for the uncompressed file」 という画面がでてきて付加的な情報を入れることも可能です。

練習7 上記で選択した部分を「vowel_a1.wav」として保存する。上記の付加的な情報はいれないで保存してください。

●補足説明:窓を消すには、窓の左上にある×印にマウスをあわせ左ボタンを押す(今は 行わないでください)。

(e) 合成音を作成する

ファイルのプルダウンメニューから新規を選択し、新しい窓を作成する(または、[Tracks] ⇒[Add New]⇒[Audio Track]を選択する)。製作のプルダウンメニューから Tone を選択す る。新しい窓が出てきて、波形の種類(Waveform)と周波数(Frequency)と音の大きさ(振幅) と音の長さ(Length)を聞いてくるので、そのままの値で「OK」を押す。ここで行った[製 作]-[Tone]とは Frequency で指定された周波数の音を作成します。ノイズのような音を作る には、Tones ではなく、「Noise」を選択し、「Noise Type」を「white」とします。次にあ る時間幅のみに音を作成します(デフォルトは 30 秒)。

練習8 音の長さが2秒で、周波数が1000Hzの音を作ってみよう。

→製作のプルダウンメニューか Silence を選択し無音区間を作ります。次に横軸の1から3 までを選択する (左ボタンを押しながらマウスを移動)。次に製作の「TONE」を選択する。 Frequency を 1000 に変更して「OK」ボタンを押す。再生ボタンを押し確認する。 (f) 簡単な操作(カット、コピー、ペースト、トリム、デリート)

これらの操作は全て編集のプルダウンメニューの中から選択できます。

○カット(Cut):選択した部分を削除します。ただしクリップボード(一時的なデータ保管 所)に保存するのでほかのところに貼り付け(ペースト)可能です。

○コピー(コピー): 選択した部分をクリップボードにコピーします。

○ペースト(ペースト): クリップボードの内容をマウスのあるところに貼り付けます。

○トリム(Trim): 選択した部分以外を削除します。

○削除(Delete): 選択した部分を削除します。クリップボードには保存しません。

練習9前の練習で作った音の一部を選択し、カットする。無音区間にマウスをもってい き、ペースト(貼り付け、ペースト)を行う。次に一部をコピーし、無音区間にペースト する。一部を選択し、トリムを行う。

(g) mp3 ファイルへの出力(wav ファイル)への出力

ファイルプルダウンメニューから「Export」を選択する(波形全体をファイルに出力)。 選択した部分のみをファイルにするには「Export Selection」を選択する。「Export File」 という新しい窓がでてきたら、ファイル名とファイルの種類を選択して保存を行う。

練習10 音の一部を選択し、「vowel_a1.mp3」として保存しなさい。

音にいろいろな効果を加える方法を学ぶ前に、音声等の基本的なことを学びます。

◆ 複雑な波形

純音は一つの周波数しか含まれないので、スペクトルは一本の線として表されます。次 にもっと複雑な波形をみてみます。複雑な音の波形でも。ある一定のパターンを時間的に 繰り返すものがあります(下図をみてください)。このような波形を「周期波形」と呼び、 その繰り返しの時間間隔を「基本周期」(ここでは 0.005 秒)と呼びます。そして、その逆 数が「基本周波数(ピッチ)」(=「基音」、ここでは 200Hz)です。周期波形のスペクトル は下右図のように 200Hz,400Hz,600Hz,800Hz,1000Hz の音が含まれており、全ての周波 数が 200Hz の倍数となっています。ここで 200Hz 以外の周波数の音を「倍音」と呼びます。





図5.倍音をもつ音の時間波形とスペクトル

◆ 音声のスペクトル

音声も複雑な波形の一つで、特に母音(「あ」「い」「う」「え」「お」)の部分は「周期波 形」の一つです。

練習11 音声の周波数解析をしてみよう。波形の全てを選択して、解析プルダウンメニ ューから「Plot Spectrum」を選択します。新しい窓(Frequency Analysis)にスペクトル が表示されます。次に「512」という部分を「2048」にしてください。「リニア周波数軸」 という部分を「対数周波数軸」にしてください。下図のようなスペクトルが表示されます。



線形周波数の場合(Linear Frequency)



対数周波数の場合(Log Frequency)

⇒もう少し詳しく「あ」のスペクトルをみると下図のようになります。周波数軸上で等間 隔に音が含まれていることがわかります。



基本周波数 (ピッチ) は約 130Hz です。

参考:成人男性の基本周波数の平均 は125Hz で、成人女性の基本周波 数の平均は2倍の250Hzです。

図6.母音「あ」のスペクトル

練習12 母音の基本周波数を調べてみよう。

母音は Vowel200ms の中にあります。vowel_*.wav というのが母音で、*が a,i,u,e,o となります。()のなかに調べた基本周波数をかいてください。

b() Hz, v() Hz, j() Hz, z() Hz, z() Hz, z() Hz

ヒント:図6のようなスペクトル(対数周波数表示にしてください)を表示し、基本周波 数に対応する山の頂点付近にマウスを移動させてみてください。

練習13 同じディレクトリの中に whisper_*.wav というファイルがあります。これらの 音にたいしても音を読み込んで、スペクトルを表示させてください。また、実際に音を再 生し聞いてみてください。練習12で用いた音(vowel_*.wav)と音の聞こえ方、時間波形、 スペクトルの違いをわかる範囲でかいてみてください。

(h) 表示の変更や消音の方法



複雑な音声の場合、時間的にスペクトルが変化します。その場合全時間に 対する周波数解析をしてスペクトルを見るよりは、時間ごとにどのような 周波数成分が含まれているか (スペクトルの形)をみたほうが良い場合が あります。Audacity ではそのような表示を行うことができます。左図の ▼のところを左クリックすると、プルダウンメニューがでてきます。その 中の「スペクトル表示」を選択します。横軸は時間のままですが、縦軸が

周波数となり、色によって音の大きさを表しています。大きさは「白」→「赤」→「青」 の順で小さくなります。

音を鳴らさない(消音)する場合には、「MUTE」を押します。複数の Audio Track を開 いている場合に、一つの Audio Track のみを音を出すときには「Solo」を選択します。

(i) エフェクト(音を加工します)

効果のプルダウンメニューにあります。加工したい部分を指定し、選択してください。

(あ)大きさを変える(用いる音:母音)

● 増幅(Amplify): 波形の大きさを変えます。

練習14 母音の大きさを変えてみよう。

注意:音の波形(縦軸の値)は-1から1の範囲で設定可能です。これを超える場合、「OK」ボ タンを押せなくなります。 ● バス・ブースト (Bass Boost...): 低い周波数を強調します。

周波数(Frequency)とブースト(Boost)を調節します。たとえば周波数が 200 ヘルツ(Hz)で、 ブーストが 12dB の場合、200 ヘルツより低い音の周波数(振動数)成分を 12dB だけ大き くするということになります。

練習15 母音に対してバス・ブーストをしてみよう。

(い) 音のスピード・テンポ・ピッチを変えます。

音のスピードとは音を再生する速さに、ピッチとは基本周波数(周期波形の周期の逆数) に、テンポの変更とは音に含まれる周波数情報は保持しながらスピードだけ変更すること に対応します。Effectの中の「スピードの変更」、「テンポの変更」、「ピッチの変更」を選択 します。

● ピッチの変更(Change Pitch):音のテンポを変えずにピッチだけ変更します。

● スピードの変更(Change Speed): 音のスピードを変更します。

● テンポの変更(Change Tempo): ピッチは変えずに、テンポだけ変更します。

練習16 単語("とうほくがくいん")に対してピッチ・スピード・テンポの変更をそれ ぞれ行い、波形の変化と聞こえの変化があるかを確認し、()の中に書いてください。音 のファイルは Others の中の tango.wav です。テンポ・スピード・ピッチの変更は入力した 時間波形にたいして行ってください(テンポの変更率を 50 にした後に、次の変更を行うま えに、「Ctrl」+「z」を押して変更前の状態にもどしてください)。

・テンポの変更 ⇒	Percent Change を 50 にする	\Rightarrow	()
・テンポの変更 ⇒	Percent Change を-50 にする	\Rightarrow	()
・スピードの変更⇒	Percent Change を 50 にする	\Rightarrow	()
・スピードの変更⇒	Percent Change を・50 にする	\Rightarrow	()
・ピッチの変更 ⇒	Percent Change を 50 にする	\Rightarrow	()
・ピッチの変更 ⇒	Percent Change を-50 にする	\Rightarrow	()

練習17 母音や文に対してもピッチやスピードやテンポを変えて聞いてみてください。

(う) 音の大きさを次第に変えます。「フェードイン」、「フェードアウト」を選択します。

● フェードイン(Fade In):音を次第に大きくします。

● フェードアウト(Fade Out):音を次第に小さくします。

練習18 文(例: Others/bun.wav または Others/mlk3.wav)を読み込んでフェードイン、 または、フェードアウトを行ってみよう。

(え)フィルタ(周波数ごとに音の大きさをかえる)

● イコライゼーション:周波数ごとに音の大きさを変えることができます。

練習19 母音(または子音、単語)などに対して FFT フィルタをかけて聞いてみよう。
 →全ての周波数で大きさが一定となっていますので、300 ヘルツから 4000 ヘルツ以外を小
 さく(縦軸の値がマイナス)に変更して「OK」を押してください(数字の目盛は大まかに
 しか表示されていませんので、大まかでかまいません)。

(お) 他いろいろ

● エコー (Echo): 減衰するエコーを連続して音に加えます。(用いる音:単語、文)



図7.エコーをつけた音の波形の例

練習20 文に対してエコーをかけて聞いてみよう。音の全範囲を選択し、遅延時間を0.1 に、減衰ファクターを0.5 にして聞いてみてください。いろいろ値を変えて行ってください。

● 上下を反転(Invert):波形の符号を反転します。

- リピート(Repeat):選択した部分を指定した回数だけ繰り返します。
- 前後を反転(Reverse):選択した部分を時間軸上で反転します(用いる音:母音、子音)

練習21 母音や子音に対してインバートや繰り返しやリバースを行ってみよう。子音に 対してリバースを行った場合、どのように聞こえたかを「」の中にかいてください。

(6) Audacity を使った簡単な音声合成

我々が話している言葉には母音と子音があります。日本語の場合、母音は「あ、い、う、 え、お」の5つで、それ以外は子音となります(ただし、「ラ行」「ヤ行」は半母音)。たとえ ば、「ば」という言葉のばあい、「b」という子音と「a」という母音の組み合わせになりま す。子音は、「b,d,g,p,t,k」の破裂音、「m,n」の鼻音、「s,h」のノイズ音にわけれます。し たがった、子音と母音の組み合わせによって単音節を合成することができます。ここでは サブディレクトリ「CV」にある音ファイルを用いて、色々な音声を合成します。このよう な音声合成をすることにより、音声がどのような音から構成されているかを理解してみま しょう。

練習22 「べ」という音が、母音「え」にしか聞こえないように後半部分だけをきりだ しなさい(範囲を決定しなさい)。次に「か」という音の前半部分だけを切り出して、前に 切り出した「え」とつなげて、「け」と聞こえるようにしなさい。

練習23 練習22と同じようにして、「せんだい」という音を合成しなさい。合成した音 声を wav 形式で保存し、ファイル名は s0657***.wav とします。

wav ファイルへの保存方法: audacity で合成音声を作成し、ファイルメニューの Export (または時間範囲を指定して Export Selection) を選択しファイル出力する。必ずファイ ルの種類として wav を選択してください。合成した音声を wav 形式で保存し、ファイル名 は sound-s0757***.wav とする必ず拡張子が wav であることを確認してください。

「せんだい」という音を作るヒント:たとえば、「せ」の音を合成するには、「sa」の「s」の部分と「be」の「e」の部分(または「e」の一部切り出し)をつなげるように合成します。「ん」は「な」の「n」の部分のみを切り出します。「い」は「bi」の「i」の部分を抽出します。

レポート提出方法

提出先と提出方法: yakurai:/home/staff/matsuo/reports-2007 にファイルを保存してくだ さい。ファイル名を自分の学生番号としてください。

提出方法の説明

⇒提出したいファイルを一つのディレクトリに保存し、このファイルがあるディレクトリ でいかのような操作を行う。

lftp –u s0757*** 157.118.89.3

cd /home/staff/matsuo/reports-2007

put sound-s0657***.wav ←提出するファイルを put でおく

ls (提出できていることを確認してください。)

exit