

WORLDチュートリアル

山梨大学大学院総合研究部・准教授

森勢将雅

mmorise@yamanashi.ac.jp

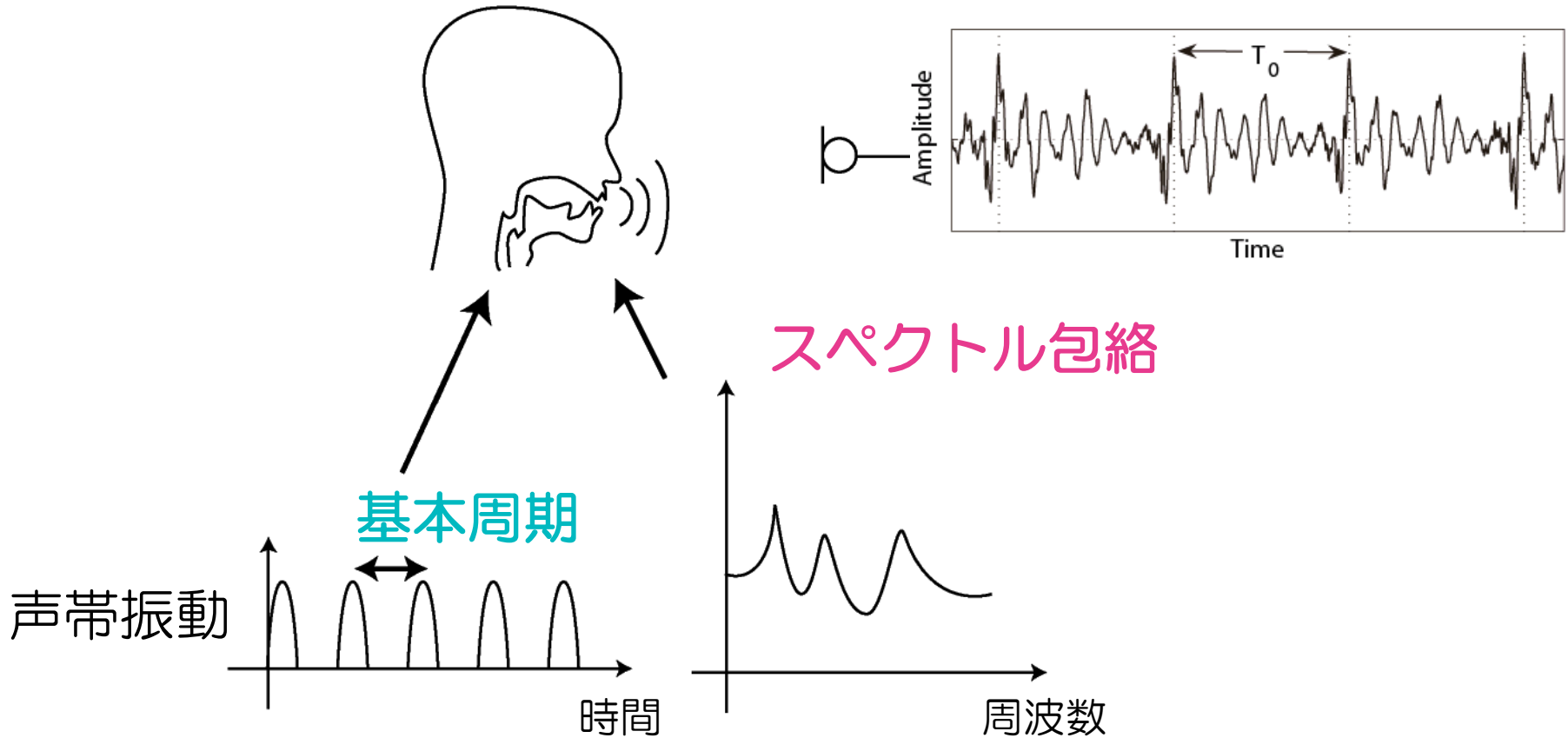
@m_morise (Twitter)

本日の概要

- 音声分析合成システムWORLDの紹介
 - 基礎的な説明と利用例について
- WORLDの導入と簡単な使い方（デモ）
 - Windows環境での導入例
- いくつかの音声分析，音声加工の実例

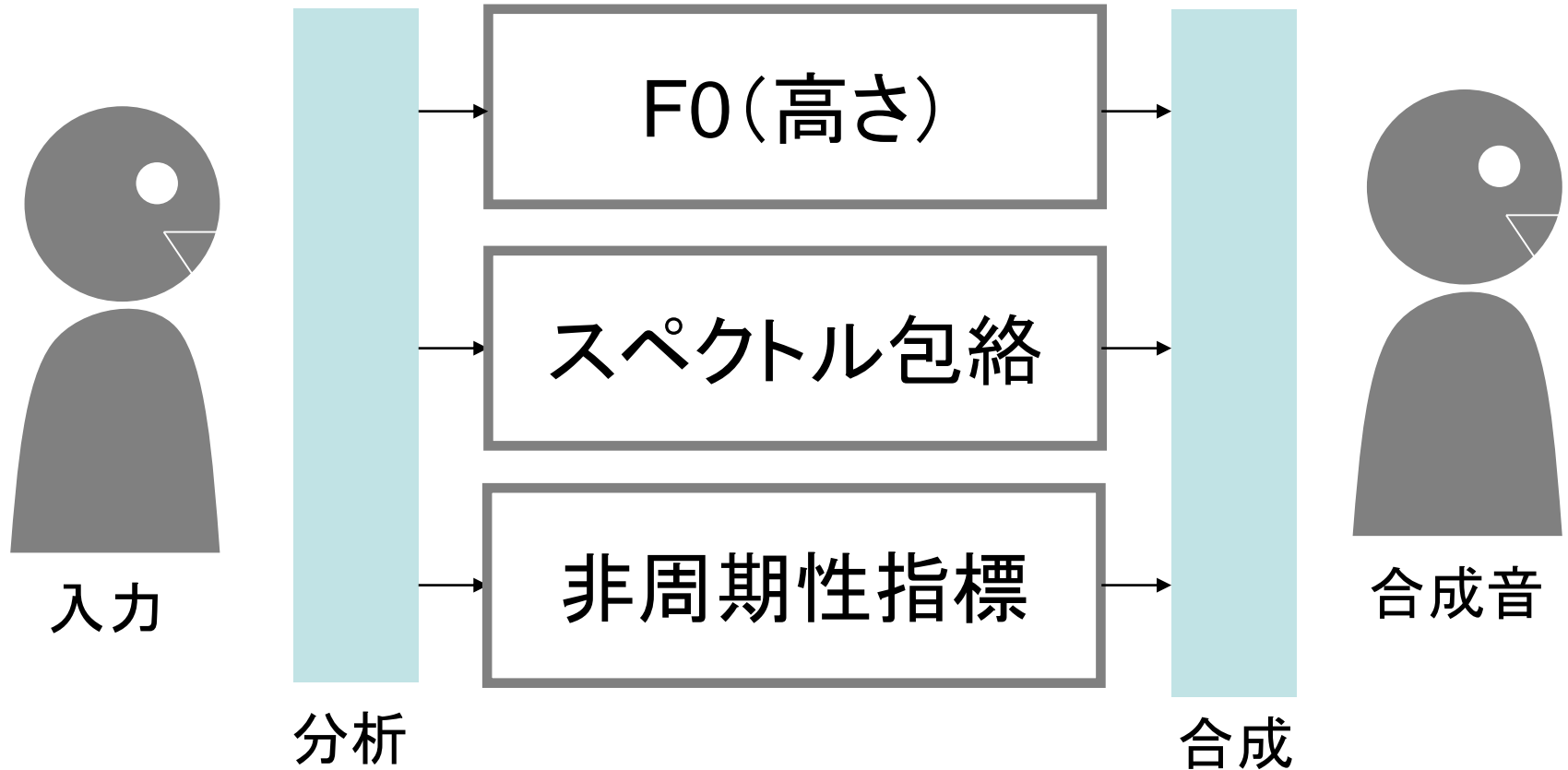
本資料では公開用に一部の情報を制限しています

音声の構成要素



- 声帯振動の揺らぎや雑音混入による影響
- 非周期性指標

WORLDの構成



- 音声を構成する3つの要素を抽出し再合成
 - 各要素は独立して制御可能

STRAIGHTとWORLD

- STRAIGHT

- 世界中で使われている最先端の技術
- 特許の扱いが面倒で商用利用は困難

- WORLD

- STRAIGHT以上の品質を達成
- 完全フリー（プログラムも修正BSD）
- 商用利用も可能

WORLDの特色

This repository Search Pull requests Issues Marketplace Explore

mmorise / World Unwatch 37 Star 155 Fork 47

Code Issues 2 Pull requests 1 Projects 0 Wiki Settings Insights

A high-quality speech analysis, manipulation and synthesis system <http://ml.cs.yamanashi.ac.jp/world/en...>

speech-analysis speech-synthesis vocoder Manage topics

232 commits 1 branch 0 releases 11 contributors BSD-3-Clause

Branch: master New pull request Create new file Upload files Find file Clone or download

47の派生
プログラム

論文のデータを再現可能なプログラムを
GitHub上に修正BSDライセンスで公開



WORLDを提案した論文は出版後1年で
自己引用を除き25件の引用

WORLDの利用実績（学術利用）

- Neural parametric singing synthesizer
（ポンペウ・ファブラ大学）
 - http://www.dtic.upf.edu/~mblaauw/IS2017_NPSS/
- The Merlin toolkit（エジンバラ大学）
 - <http://www.cstr.ed.ac.uk/projects/merlin/>
- Voice Synthesis for in-the-Wild Speakers
（Facebook AI research）
 - <https://arxiv.org/abs/1707.06588>

WORLDの利用実績（製品利用）

- VoiceText（HOYA株式会社）
 - <http://voicetext.jp/voiceactor/>

- CeVIO（株式会社テクノスピーチ）
 - <http://cevio.jp/>

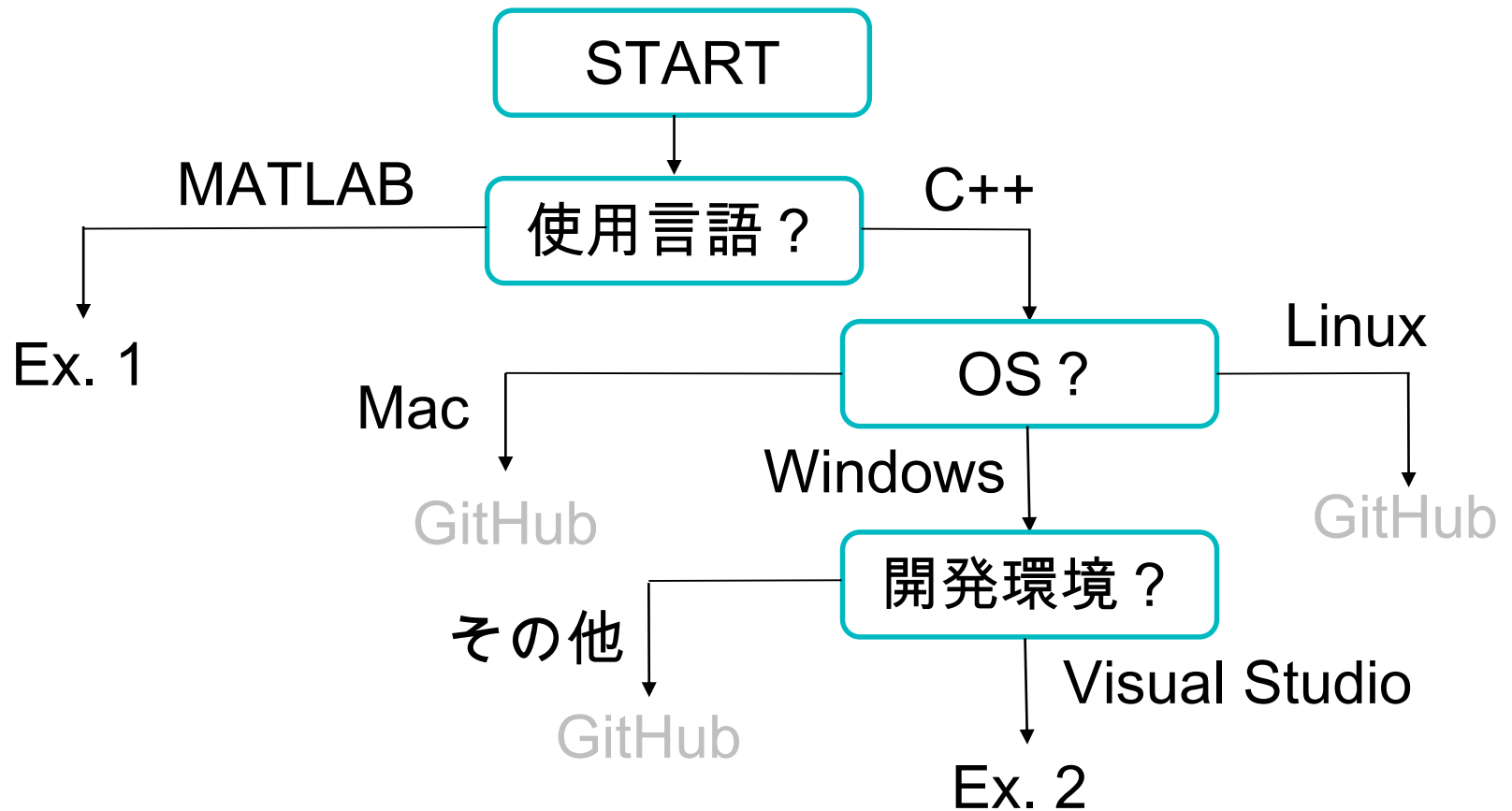
WORLDの利用実績（個人利用）

- VocalShifter
 - <http://ackiesound.ifdef.jp/download.html#vs>
- エレクトロ音楽ユニットmacaroom
 - <https://www.youtube.com/watch?v=5b47SueEWjY>
 - コンテンツ中に2名の歌手の歌声を混ぜるモーフィング技術を利用
- 徴音梅林（台湾のプロジェクト）
 - <http://projectmeilin.github.io/>



WORLDの導入

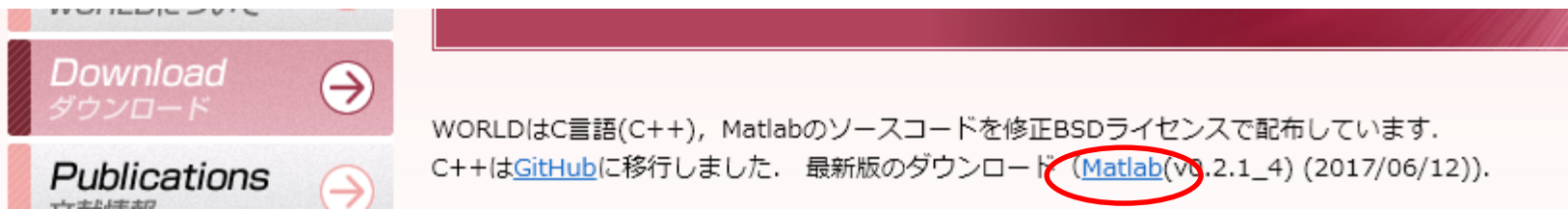
導入方法について



本日は初心者を対象にEx. 1, 2をサポート

MATLAB版の導入手順

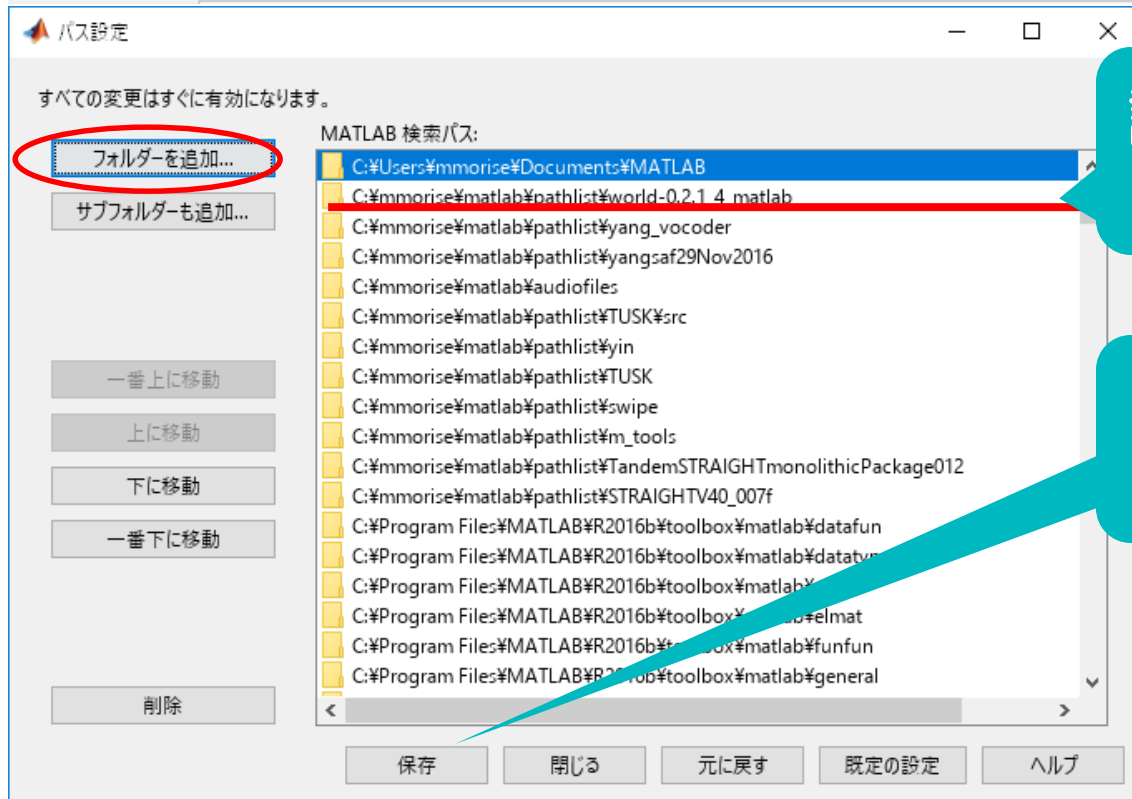
- MATLABはインストール済みとします。
 - <http://ml.cs.yamanashi.ac.jp/world/index.html>
 - Downloadから赤丸をクリックでダウンロード



フォルダを解凍し任意の場所に設置
(MATLABのインストールフォルダなど、
頻繁に動かさない場所を推奨)

MATLABの設定

・パスを設定する



設置したフォルダを参照して設定

設定後には必ず「保存」を押す

動作チェック

MATLABのコマンドラインで

```
>> TestWORLD;
```

```
>> soundsc(x, fs); % 入力音声の再生
```

```
>> soundsc(y, fs); % 合成音声の再生
```

合成音が再生されれば導入は完了

Signal Processing Toolbox必須のため注意

Windowsでの導入手順

- Visual studioを利用
 - 恐らくVisual Studio CommunityエディションでOK
 - 本資料は、Visual Studio 2015でチェック済み
- 「Visual studio communityエディション」で検索してインストール



WORLDそのものはインストール不要

音声分析変換合成の例

MATLAB版

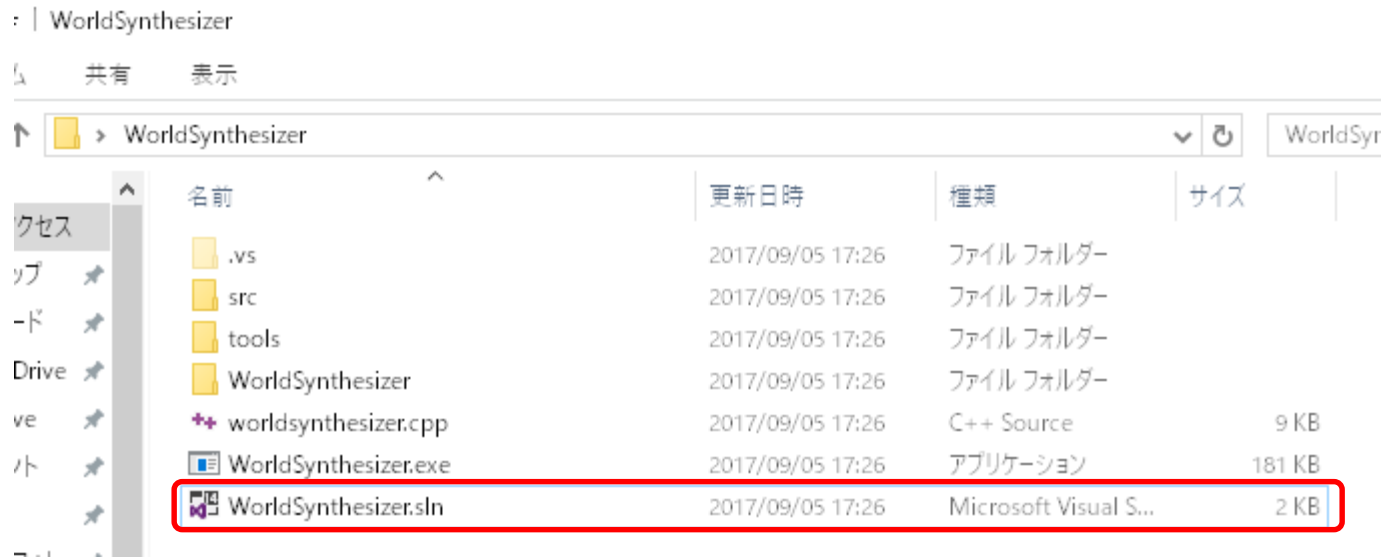
- <http://ml.cs.yamanashi.ac.jp/media/20170921/>
- で必要なファイルをダウンロード可能
- 作業フォルダにファイルを配置（あるいはWORLDのフォルダにおいてもOK）

>> Help WorldSynthesizer で書式を確認

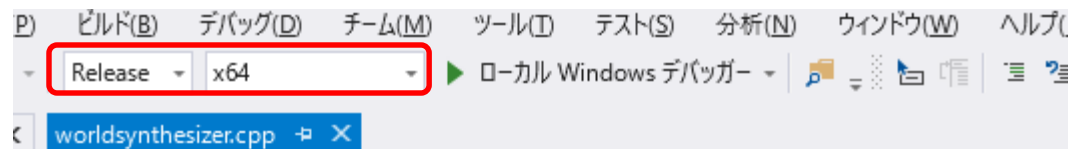
>> WorldSynthesizer('vaiueo2d.wav', 'output.wav', 1, 1, 1);
順番に「入力ファイル名」「出力ファイル名」「FOの変化率」「フォルマントの変化率」「発話時間の変化率」

C++版

- 赤丸のアイコンをダブルクリックで起動



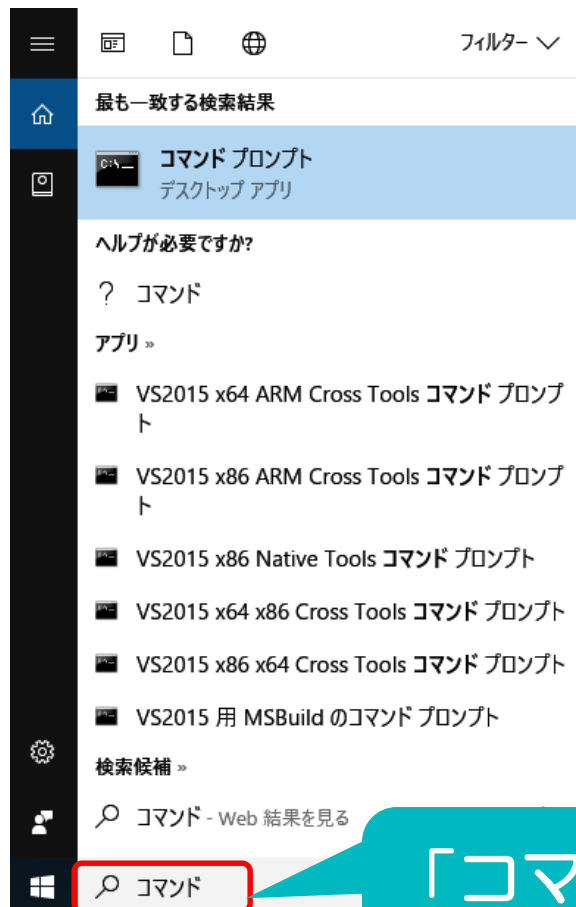
「Release」 「x64」にしてF7でコンパイル



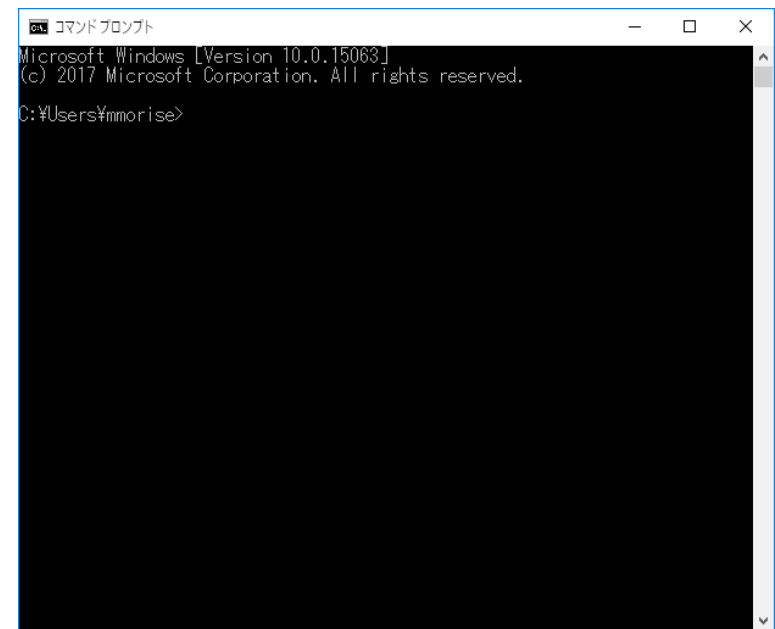
現フォルダにx64フォルダができるので
その下のReleaseに.exeが作成

実行方法（1 / 3）

- コマンドプロンプトを開く



「コマンド」まで入
れば出る



コマンドプロンプトの実行画面

実行方法 (2 / 3)

ホーム 共有 表示

C:\Users\mmorise\Desktop\WorldSynthesizer

アドレスをコピー

アドレスを選択して
右クリック

元に戻す(U)

切り取り(T)

コピー(C)

貼り付け(P)

削除(D)

すべて選択(A)

右から左に読む(R)

Unicode 制御文字の表示(S)

Unicode 制御文字の挿入(I)

IME を開く(O)

再変換(R)

drive x64
x64
worldsynthesizer.cpp
WorldSynthesizer.exe
WorldSynthesizer.sln
WorldSynthesizer.VC.db
WorldSynthesizer.VC.VC.opendb

9 KB
181 KB
2 KB
220 KB
0 KB

コマンドプロンプト

```
Microsoft Windows [Version 10.0.15063]  
(c) 2017 Microsoft Corporation. All rights reserved.  
C:\Users\mmorise>cd "C:\Users\mmorise\Desktop\WorldSynthesizer"
```

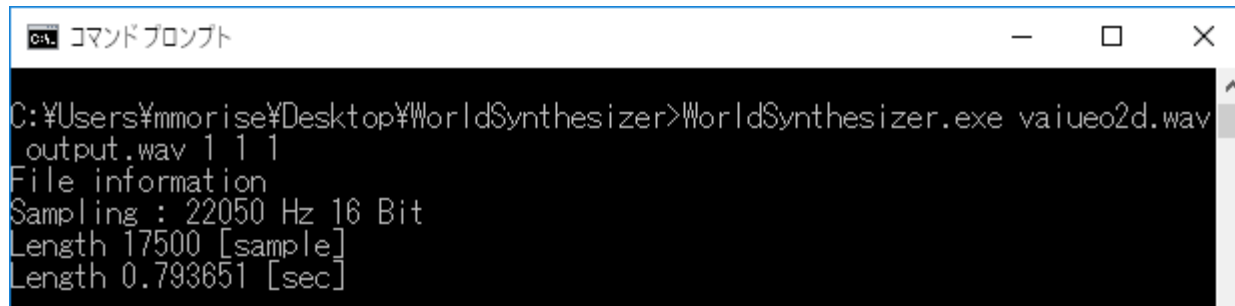
最後に「”」を打って
Enterを押す

「cd ”」まで打ち

右クリックでペースト

実行方法（3 / 3）


- 書式はMATLABと同じだがファイル名を「」で囲む必要がない



```
cmd コマンドプロンプト
C:\Users\mmorise\Desktop\WorldSynthesizer>WorldSynthesizer.exe valueo2d.wav
output.wav 1 1 1
File information
Sampling : 22050 Hz 16 Bit
Length 17500 [sample]
Length 0.793651 [sec]
```

- 分析音声のデータが表示される

簡単な音声変換はこのコマンドで実現可能！



音声の基本周波数分析の例 ～バッチ処理による一括分析～

利用方法

- f0analysis.exeとanalysis_all.batを分析したいファイルのあるフォルダに置く

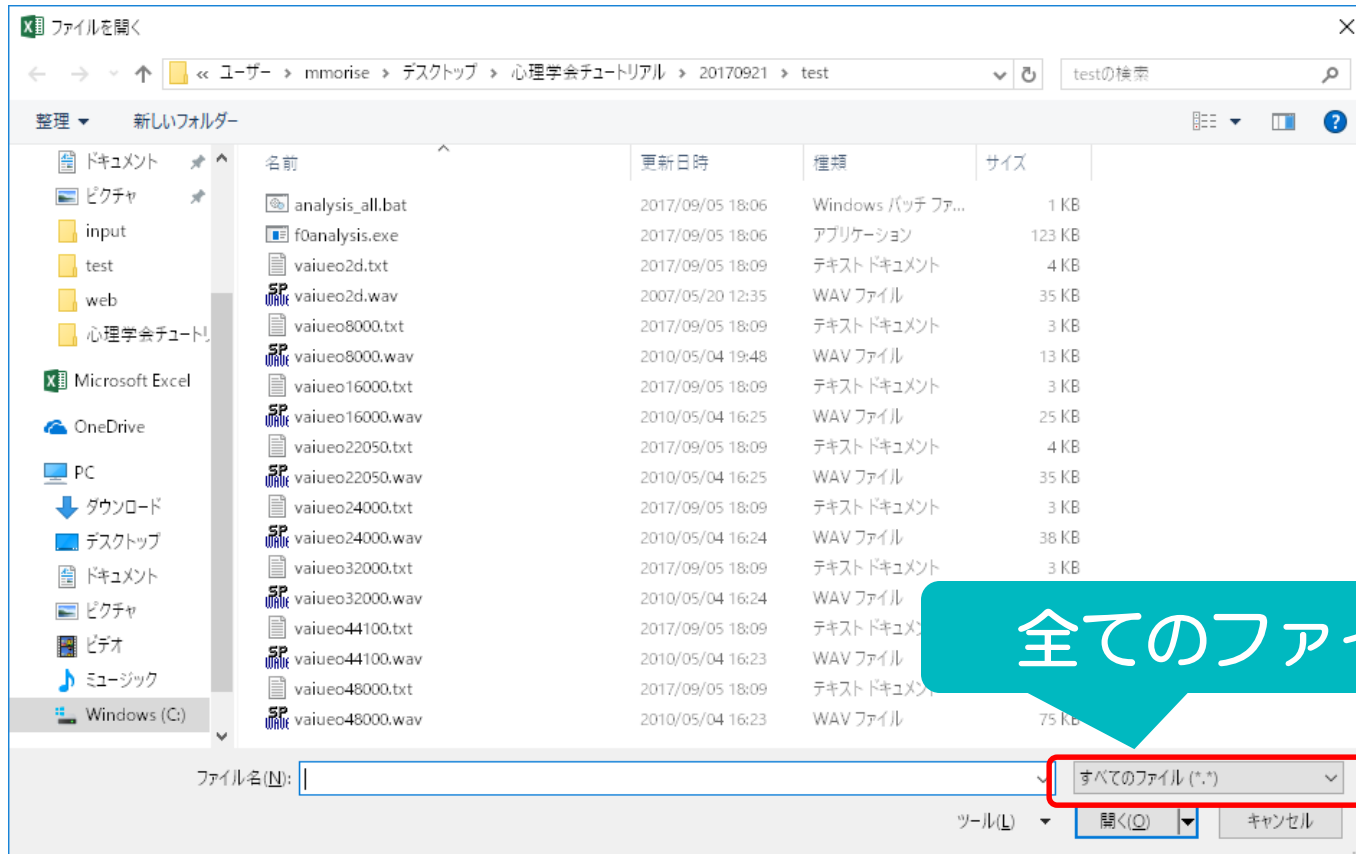


名前	更新日時	種類	サイズ
analysis_all.bat	2017/09/05 18:06	Windows バッチ ファ...	1 KB
f0analysis.exe	2017/09/05 18:06	アプリケーション	123 KB
vaiueo2d.wav	2007/05/20 12:35	WAV ファイル	35 KB
vaiueo8000.wav	2010/05/04 19:48	WAV ファイル	13 KB
vaiueo16000.wav	2010/05/04 16:25	WAV ファイル	25 KB
vaiueo22050.wav	2010/05/04 16:25	WAV ファイル	35 KB
vaiueo24000.wav	2010/05/04 16:24	WAV ファイル	38 KB
vaiueo32000.wav	2010/05/04 16:24	WAV ファイル	50 KB
vaiueo44100.wav	2010/05/04 16:23	WAV ファイル	69 KB
vaiueo48000.wav	2010/05/04 16:23	WAV ファイル	75 KB

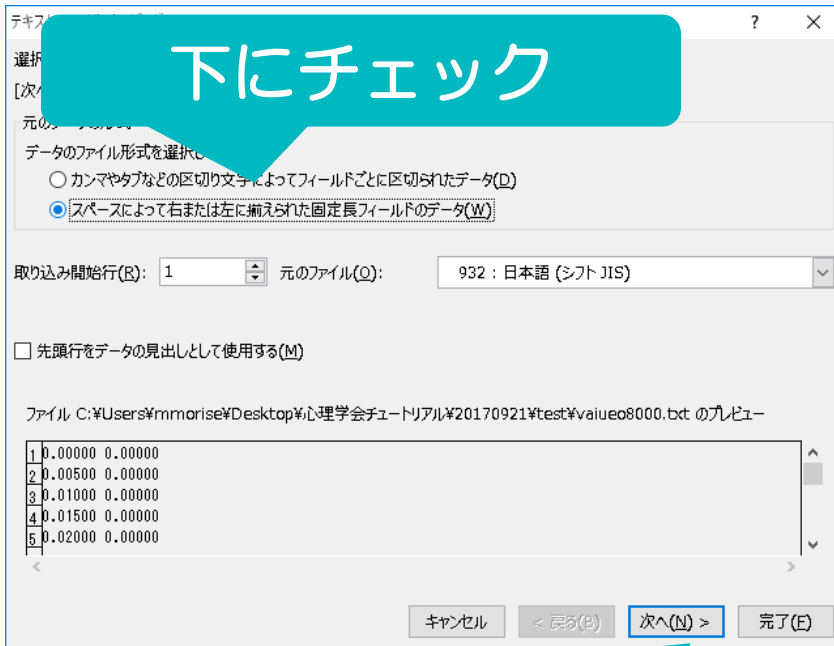
この状態でanalysis_all.batをダブルクリック
自動的にフォルダ内の全ファイルを解析

分析結果の読み込み

- Excelで読み込み可能

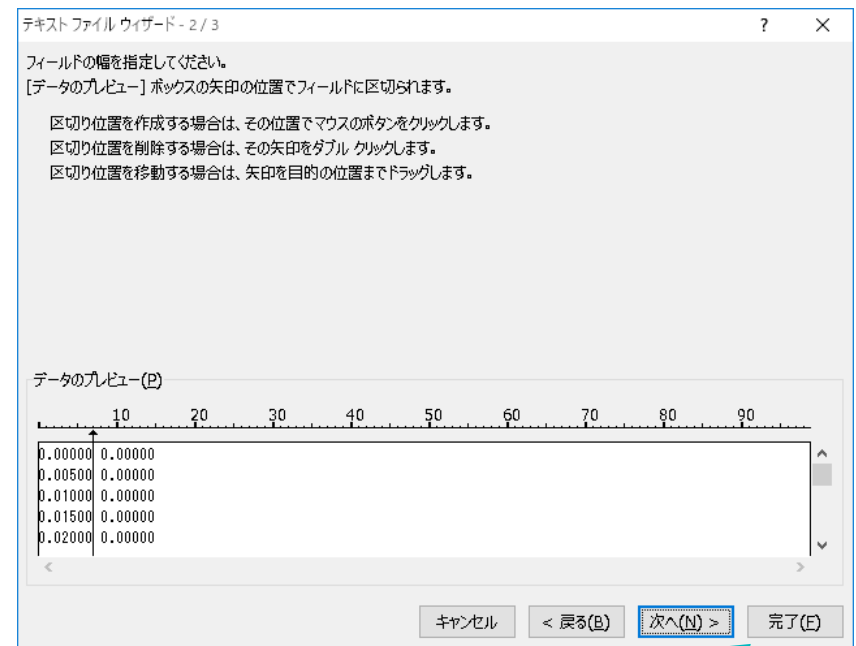


読み込み時の注意点



下にチェック

次へをクリック



完了をクリック

1列目は分析時刻，2列目がF0
無声音の場合は0が入っている

WORLDの特徴

- 分析に失敗する場合パラメータ調整が可能
- **音声モーフィング**などさらに高度な加工も可能
- 合成分析の精度は**世界最高**でありながら、特許がないため誰でも自由に利用可能
 - 開発者が現役かつ日本人なので、日本語での問い合わせが可能（いつでもご連絡ください）

おわりに

意見募集

- 実験で使うにあたりどのようなソフトウェアが好ましいかの意見を募集しています
- 望ましい分析パラメータ（フォーマットなど）についても同様に募集しています
- 様々な分野で利用可能なツールの開発を目指して今後も研究開発を進めていきます

参考文献 1

- WORLD本体について（ [2]は非周期性指標推定法です ） .
「WORLD [1] (D4C edition [2])」 と引用してください .
- [1] M. Morise, F. Yokomori, and K. Ozawa: WORLD: a vocoder-based high-quality speech synthesis system for real-time applications, IEICE transactions on information and systems, vol. E99-D, no. 7, pp. 1877-1884, 2016.
- [2] M. Morise: D4C, a band-aperiodicity estimator for high-quality speech synthesis, Speech Communication, vol. 84, pp. 57-65, Nov. 2016. <http://www.sciencedirect.com/science/article/pii/S0167639316300413>

参考文献 2

- スペクトル包絡推定法

[3] M. Morise: CheapTrick, a spectral envelope estimator for high-quality speech synthesis, *Speech Communication*, vol. 67, pp. 1-7, March

2015. <http://www.sciencedirect.com/science/article/pii/S0167639314000697>

[4] M. Morise: Error evaluation of an F0-adaptive spectral envelope estimator in robustness against the additive noise and F0 error, *IEICE transactions on information and systems*, vol. E98-D, no. 7, pp. 1405-1408, July 2015.

参考文献 3

- 基本周波数推定法Harvest

[5] M. Morise: Harvest: A high-performance fundamental frequency estimator from speech signals, in Proc. INTERSPEECH 2017, pp. 2321–2325, 2017. http://www.isca-speech.org/archive/Interspeech_2017/abstracts/0068.html

