作成日：2019年6月17日

画像差分法を用いたピアノの演奏動画を

midiファイルに変換するプログラムについて

法政大学　情報科学部　ディジタルメディア学科　3年

雨宮　聡

1．プログラム概要

　　このプログラムでは、ピアノの演奏動画を読み込み、その画像データから演奏した音を検索し、midiデータに変換するものである。動画の各フレーム画像から、画像の差分を計算し、一定量の差分が取れた場合にその座標、フレーム位置から弾いた鍵盤を読み取り、出力する。現在の開発段階では、拡張性を意識して開発したものの、添付したTest.mp4のみで動作するようになっている。

2．実行方法

　　・実行ファイルの場合

　　　　Launcher.exeファイルと同じディレクトリ内にTest.mp4を配置し、Launcher.exeを起動する。出現したコンソールにファイル名(Test.mp4)と入力する。入力が完了するとプログラムが実行され、そのディレクトリに各データを出力したdataフォルダと、Output.midiが出力される。

　　・スクリプトファイルの場合

　　　　Scriptフォルダ内にTest.mp4を配置し、Launcher.pyを実行する。その後ファイル名を入力するとプログラムが実行される。出力結果は実行ファイルと同様である。

実行には、Python3.6の動作環境と、opencv、numpy、pretty-midiなどのモジュールのインストールが必要である。

3．プログラム解説

　3－1　プログラム全体の流れ

各プログラムについて、スクリプトファイルをもとに解説する。

このプログラムの処理手順は以下の通りである。

(1) 動画を読み込み、各フレームの画像データに分解、保存する。

(2) フレームの画像から、鍵盤の座標を探索し、切り取る。

(3) 切り取った鍵盤の座標を、射影変換を用いて処理しやすくし、保存する。

(4) 変換した鍵盤画像をもとに、各鍵盤のx座標位置を検索、保存する。

(5．変換した各画像を順番に読み取り、画像差分を検索、大きく差分の出たx座標から弾かれた鍵盤を取得、保存する。

(6) 保存された鍵盤名、フレーム位置をもとにmidiファイルを作成する。

3－２　s00\_make\_img.py

このスクリプトでは、(1)の動作を実行する。Launcherから読み込んだパスにある動画を読み込み、各フレームに分解する。分解したフレームは、data/img/01image内に、フレーム番号.jpgという名前で保存される。



1　切り取られるフレーム例

data/img/01image/000.jpg

3－3　s01\_img\_exchanged

このスクリプトでは、(2)、(3)の動作を実行する。

　　(2)の動作について

　　　　鍵盤の位置は、get\_yy関数で検出する。この関数では、演奏が始まっている　状態の2枚のフレームの差分を取り、そのy座標を取得する。カメラを固定しているため、鍵盤以外のだひょうは変化しないことから、差分で撮れた箇所により鍵盤の上端を判別することができる。これで取得したy座標から、鍵盤の写っている範囲を切り取る。



　切り取りされた画像例

　　　　上記で切り取った範囲の中から、全体が写っている両端の鍵盤の座標を探索する。(get\_point関数)これをもとに射影変換を行い、すべての画像を真っすぐになるように調整する。この動作はmake\_img関数で実行している。これにより以下のように変換できる。

