

Python ライブラリを用いた実験データの高効率な処理システムの構築

大嶋元啓^{1†} 中山勝之² 小原弘之¹ 坂村芳孝¹

¹富山県立大 ²オープン CAE 勉強会@富山

Development of Experimental Data Processing System with Python Libraries

Motohiro OSHIMA^{*†} Katsuyuki NAKAYAMA^{**} Hiroyuki OHARA^{*} Yoshitaka SAKAMURA^{*}

^{*}Toyama Prefectural University ^{**}Open CAE@Toyama

Abstract

In this study, a data processing system for spectroscopic measurement of the dielectric-barrier-discharge plasma jets was developed by using python libraries. The chemical species for the spectrum peaks were identified and the vibrational temperature were calculated from the spectrum data. It was then found that the chemical species could be identified automatically, and the vibrational temperature is agreed with the calculation result by hand.

Keywords: Python, Data Processing, Matplotlib, NumPy, OpenCV, Pandas, Python-pptx, Scikit-learn, SciPy

1. はじめに

近年、高速度ビデオカメラ、デジタルオシロスコープの高性能化により、これまで把握できなかった短時間における物理現象を高空間分解能あるいは高時間分解能で把握することが可能になった。しかし、測定データは膨大であり、これらを効率的に処理することは難しい。著者らはこれまで、画像処理には `imagej`、数値データ処理には `Matlab` や `Fortran` を用いて処理してきた。これらソフトウェアは素晴らしい処理機能を併せ持っているが、複数の画像データを数値データに変換し、数値データから輝度の分散を実験条件ごとにまとめて出力するといった複雑な処理を行う場合は、複数のソフトウェアを使用あるいは手動でソフトウェアを操作して処理する必要がある。

著者らは大気圧誘電体バリア放電ヘリウムプラズマ噴流（以下、DBD プラズマ噴流）の撮影画像、スペクトルからプラズマ噴流と周囲気体との関係を明らかにしてきた⁽¹⁾⁽²⁾。DBD プラズマ噴流の制御因子は印加電圧、繰り返し周波数、ヘリウム流量など多岐にわたる。そのためこれらの制御因子を変化させて実験を行った場合、測定データが大量に生成される。これまで測定データの解析は前述の通り複数のソフトウェアを用いて手作業で行ってきた。そのため、データ処理に時間を要することが問題となっていた。そこで、本研究では DBD プラズマ噴流の測定データを効率的に処理することを目的としている。はじめに、測定データの処理システムを `python` ライブラリにより構築した。そして、測定データにスペクトルデータを用いて、スペクトルピークに対応する化学種の同定と窒素の振動温度の計算を行い、同条件で整理したデータを `ppt` ファイルに出力した。

2. 測定データとデータ構成

本研究では測定データに DBD プラズマ噴流のスペクトルデータを用いた。実験パラメータは印加電圧、繰り返し周波数、測定位置である。図 1 に測定したスペクトルの一例を示す。測定データはテキストファイルに波長と発光強度が 2 列 3648 行のデータで格納されている。また、測定データは図 2 に示す実験条件で整理されたファイル階層で構成されており、そのファイル数は 450 個である。

3. データ処理プログラム

高効率なデータ処理を実現するために `python` ライブラリを用いて処理システムを構築した。使用言語は

[†] E-mail address of corresponding author: m-oshima@pu-toyama.ac.jp

python 3.6 であり, python パッケージとして Anaconda を用いた. 開発には jupyter notebook ver.5.4.1 を用いた.

3.1. データ処理

測定データの処理フローを図 3 に示す. まず, 指定した電圧, 周波数を入力するとファイルの場所, 実験条件が記されたデータベースから, Pandas モジュールの query 関数, loc 関数により指定した条件に合致した測定データが逐次読み込まれる⁽³⁾. 次に scipy.signal モジュールから argrelexmax 関数を用いて, スペクトルピークの探索が行われる⁽⁴⁾. そして, そのスペクトルピークを基に対応する化学種が化学種データベースから Numpy モジュールの argmin 関数により探索され, 同定される. この結果と測定データは Matplotlib モジュールの pyplot 関数により組み合わせられ, 画像として出力される⁽⁵⁾⁽⁶⁾. この画像は画像内に OpenCV モジュールの cv2.putText 関数により実験条件が付記された後, 同条件の画像同士で cv2.concat 関数により連結され⁽⁷⁾⁽⁸⁾, python-pptx モジュールにより ppt ファイルに纏められる⁽⁹⁾.

3.2. 窒素の振動温度の算出

窒素の第二正帯に対応するスペクトルピークの放射光強度 I を探索し, 振動温度 T_{vib} を算出した. 放射光強度と振動温度の関係は以下の式で表される⁽¹⁰⁾.

$$\ln \frac{A_{02} I_{v_{02}}^4}{A_{v'v''} I_{02} v_{v'v''}^4} = -\frac{1}{kT_{vib}} (E_{v'} - E_0) \quad (1)$$

A : 遷移確率 (本研究ではフランク・コンドン係数⁽¹⁰⁾を用いた.), E : 振動エネルギー,
 k : ボルツマン定数, v : 振動数

添字 : v' : 窒素分子の C 状態の振動準位, $v'' = v' + 2$: 窒素分子の B 状態の振動準位

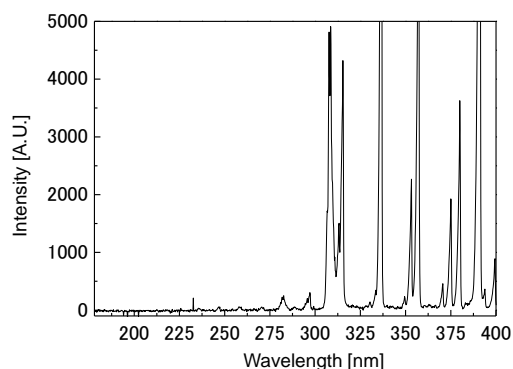


Fig.1 Example of the spectrum of DBD plasma jet

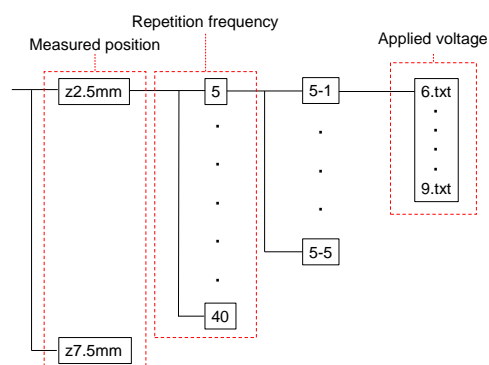


Fig.2 File tree of measured data

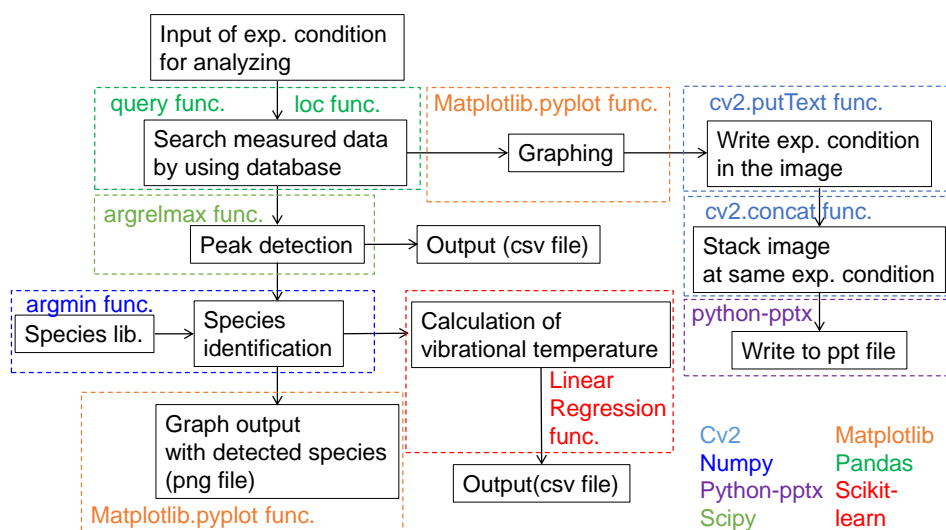


Fig.3 Flow chart of data processing

振動温度は(1)式からわかるように切片が 0 となる．本処理プログラムでは Scikit-learn モジュールの LinearRegression 関数を用いて⁽¹¹⁾，線形回帰直線の傾き $-1/kT_{vib}$ を計算し，振動温度を算出した．

4. 処理結果

図 4 は波長 178～400 nm におけるスペクトルの出力結果と同定した化学種を示した結果の一例である．同定した順番の奇数番目が青字で，偶数番目が赤字で示されている．この結果はこれまで，手作業で化学種を同定していたが，手作業を自動化することにより，ミスなく効率的にデータ処理を行うことができることを示している．図 5 は出力された csv データを基に縦軸に $\ln A_{02} I_{\nu',\nu''} \nu_{02}^4 / A_{\nu',\nu''} I_{02} \nu_{\nu',\nu''}^4$ ，横軸に $E_{\nu'} - E_0$ でまとめたボルツマンプロットを示したものである．この時の傾きは -2.91×10^{19} ，振動温度は 2490.63 K となり，手計算の計算結果と一致する．

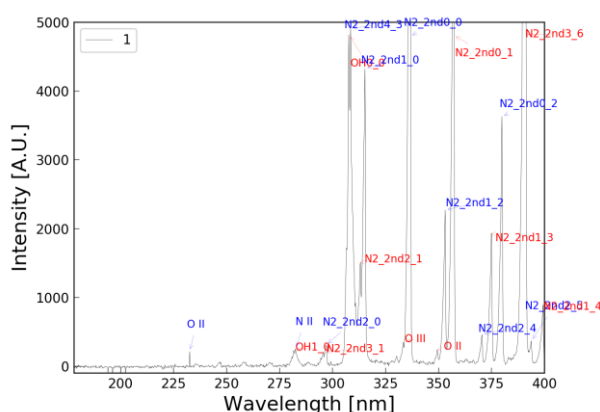


Fig. 4 Example of spectrum image output with data processing system

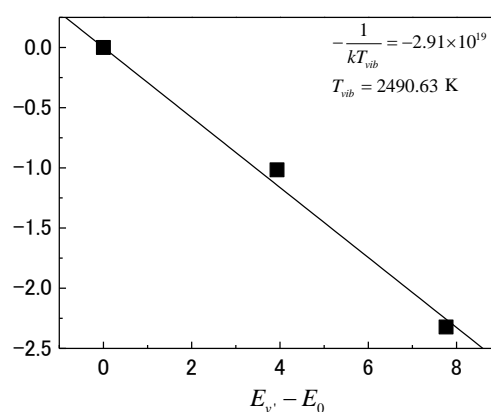


Fig. 5 Example of boltzmann plot output

5. おわりに

本研究では DBD プラズマ噴流の測定データを効率的に処理するため python ライブラリを用いて，処理システムを構築した．測定データにスペクトルデータを用いて，本処理システムによりスペクトルピークに対応する化学種の同定と窒素の振動温度を計算し，同条件にデータをまとめ ppt ファイルに出力した．その結果，本処理システムは化学種の同定の自動化を可能にし，効率的なデータ処理の実現を達成した．これらの研究成果は，python ライブラリを用いることで大規模な測定データの効率的な処理を容易に実現できると期待される．

参考文献

- [1] 大嶋元啓, 小原弘之, 中山勝之, 坂村芳孝. 大気中に噴出した誘電体バリア放電ヘリウムプラズマ流の発光特性. 日本機械学会熱工学コンファレンス 2018 講演論文集, 2018. C224.
- [2] M. Oshima, S. Kuwashima, I. Wakita, D. Ozeki, Y. Sakamura, Visualisation Study of Atmospheric Dielectric-Barrier Discharge Plasma, Jets. The 11th Pacific Symposium on Flow Visualisation and Image Processing, PSFVIP11, 2017. PSFVIP11-101.
- [3] <Python, pandas> データフレーム文字列検索, <http://nekoyukimmm.hatenablog.com/entry/2015/11/19/164317> (accessed 2018-06-04).
- [4] SciPy で離散データのピークを検出 <https://org-technology.com/posts/scipy-peak-finding.html> (accessed 2018-06-04).
- [5] 【Matplotlib】ベクトル(矢印) quiver(x,y,u,v), <https://algorithm.joho.info/programming/python/matplotlib-quiver/> (accessed 2018-09-20).
- [6] Annotation — Matplotlib 2.0.2 documentation, <https://matplotlib.org/users/annotations.html> (accessed 2018-06-08)
- [7] OpenCV 3 (core + contrib) を Windows & Python 3 の環境にインストール&OpenCV 2 と OpenCV 3 の違い&簡単な動作チェック, <https://qiita.com/hitomatagi/items/d5d475a446ec9c73261e>, (accessed 2018-06-08)

018-08-10).

- [8] OpenCV-Python チュートリアル, http://labs.eecs.tottori-u.ac.jp/sd/Member/oyamada/OpenCV/html/py_tutorials/py_tutorials.html (accessed 2018-08-10).
- [9] python-pptx まとめ, <https://qiita.com/pocket8137/items/3d8fda2c47664bf9130b> (accessed 2018-11-19).
- [10] 堤井信力, 小野茂, プラズマ気相反応工学, 内田老鶴圃, 2000, p.137.
- [11] scikit-learn で線形回帰 (単回帰分析・重回帰分析), <https://pythondatascience.plavox.info/scikit-learn/%E7%B7%9A%E5%BD%A2%E5%9B%9E%E5%B8%B0> (accessed 2018-10-15).