

時系列速度データからのパターン抽出法 Methods of Pattern Extraction from Velocity Data

○ 白山 晋, 東大院工, 〒113-8656 東京都文京区本郷 7-3-1, sirayama@sys.t.u-tokyo.ac.jp
Susumu SHIRAYAMA, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo, 113-8656, Japan

It can be considered that some of the validation methods are based on a comparison between the numerical solutions and some kind of reference solutions including experimental results. Since recognition of unsteady patterns of multivariate property data such as a velocity data has played important roles in understanding the flow field, direct comparison method should be used for multivariate variables especially in the case of the analysis of unsteady flow fields. However, the direct comparison among the data of unsteady flow fields is quite difficult owing to the strong unsteadiness even if using the most recent information techniques. Therefore, such a strategy that a rough pattern is extracted from data has become important. By the way, methods of pattern extraction from multivariate time series data have been studied in various fields. It is known that the performance of the method strongly depends on the characteristics of data. As far as the author knows, an adequate method for extracting a significant pattern from a fluid velocity data has not been clearly revealed in the case of unsteady flow field. In this paper, we propose methods of extracting patterns from the unsteady velocity data which obtained in both numerical and experimental methods. In addition, the advantages and disadvantages of each method will be noted.

1. はじめに

数値計算結果の有効性は、実験や観測データとの比較、異なる格子で得られた数値解同士の比較によって示されることが多い。より厳密な比較を行う際には、比較における定量性の確保が重要になる。しかしながら、異なる手法で求められたデータの定量的な比較は容易ではない。

著者らは、得られたデータの空間情報（位置）の違いが定量性の確保を難しくしている点を指摘し、比較空間という考え方を提案した。また、異なる手法で求められたデータ間の差異を直接的に明示できることを示した⁽¹⁾⁽²⁾。しかしながら、非定常速度場に対しては、非定常、およびベクトル場という点で定量的な比較は難しいことが明らかになっている。

空間情報を考慮しない場合でも、非定常性が強い場合には定量性を確保した比較は容易ではない。例えば、異なる数値計算の同じ位置での時系列速度データに注目すると、フーリエ解析やウェーブレット解析などによって成分毎の定量的記述は可能になり、原理的には定量的な比較はできる。しかしながら、非定常性が強い場合での比較は、難しいことが知られている。

一方、定量性には拘らず、「見た目」といった、分析者の判断に基づく比較が行われることも多い。例えば、時系列速度データの場合でも、明確な周期性が存在しない場合、見た目による比較が行われることも少なくない。思考を深めながら分析を進める場合、そのような比較の方が重要なことも少なくないだろう。また、低い精度のデータから有意な情報を抽出する場合にもそのような比較が行われる。認知的な研究からは、データ毎に特徴的なパターンを見出し、パターン単位で（パターンとして）比較することが知られている⁽²⁾（ただし、必ずしもパターンが明示されるわけではない）。このため、分析者の判断を支援することや、比較によって有意な情報を抽出する場合には、特徴パターン抽出法とパターンを利用した比較法が重要になる。

このような考え方は、特に実験結果との比較において重要になる。実験における計測精度は向上しているが様々な条件の違いから直接的な比較は、代表値を用いた比較などに限定的されることが多い。また、タフト法を用いた実験のような目視によって得られた情報を用いる場合には、全体の大まかなパターンを用いた比較になる。客観性を欠くこと、あるいは定量性を欠くことが指摘されるが、現在でも流体现象を理解する手がかりを得るために、そのような比較検討が数多く行われている⁽³⁾⁽⁴⁾⁽⁵⁾。この場合のパ

ーン抽出は、主として画像からのパターン抽出となる。一方、著者らはタフト法であっても定量的な計測が可能であることを示した⁽⁶⁾。文献⁽⁶⁾では、Fig.1 に示したタフト 3,4,8 の流れ方向の時間変化を求めている。Fig.2 に正規化した値のグラフを示す。



Fig.1 タフト法を用いた流れ方向の計測実験

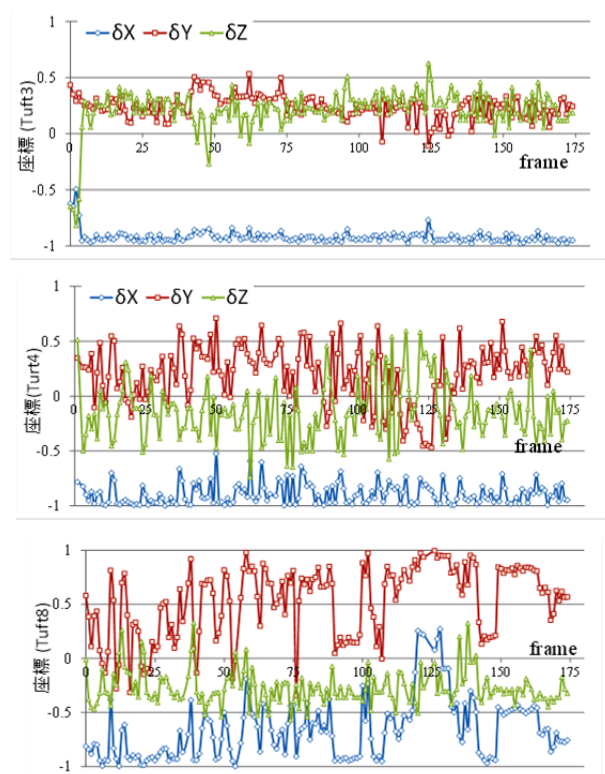


Fig.2 Fig.1のタフト3,4,8の方向の時間変化

このような計測で得られる情報は、目視から得られるものよりも多い。また、数値としてデータが取得できている。様々な条件の違いに加えて、計測精度が低いために、数値を用いた比較は難しいが、全体の大まかなパターンを用いた比較ではなく、時間区間での比較が可能になる。この点に注目すると、数値計算からのデータと同様のパターン抽出が重要になる。

本研究では、数値計算、および計測から得られた時系列速度ベクトルからのいくつかのパターン抽出法を提案する。また、実験を通して抽出法の利害得失について考察する。

2. 提案手法

時系列速度データを識別するために、速度データが取得される座標を \mathbf{x}^α とする。ここで α を下記で定義するデータ集合の識別子とする。また、速度を $\mathbf{u}^\alpha = \mathbf{u}(t, \mathbf{x}^\alpha)$ とし、その成分を $(u^\alpha, v^\alpha, w^\alpha)$ とする。 $N_\alpha + 1$ 個の時系列速度データの集合 \mathbf{S} を、

$$\mathbf{S}^\alpha \{ \mathbf{u}_0^\alpha, \dots, \mathbf{u}_n^\alpha, \dots, \mathbf{u}_{N_\alpha}^\alpha \}$$

とする。ここで、下添字は取得時間を識別するためのものとする。また、データ \mathbf{u}_n^α が取得された時間を t_n^α とする。

2.1 統計量や類似度に対するクラスタリングからのパターン抽出

成分毎の平均、分散、自己相関などを代表値としたクラスタリングによって \mathbf{S}^α を分類し、その分類に基づいて、クラスター内、およびクラスター間での類似性、あるいは相違性から特徴的なパターンを見つけるという方法が考えられる。例えば、平均と標準偏差 $(\mu_u^\alpha, \mu_v^\alpha, \mu_w^\alpha, \sigma_u^\alpha, \sigma_v^\alpha, \sigma_w^\alpha)$ を用いた、6 変量に対する K-means 法や SOM によるクラスタリングが考えられる。

しかし、この方法では、データ間の直接的な関係性を考慮できない。データ間の関係性を表すためには、時系列の相関を考えればよい。例えば、 u 成分の場合、ピアソンの相関係数を $r_{\alpha\beta}$ とすると、

$$r_{\alpha\beta} = \frac{\sum_{n=0}^N (u_n^\alpha - \mu_u^\alpha)(u_n^\beta - \mu_u^\beta)}{\sigma_u^\alpha \sigma_u^\beta} \quad (1)$$

となる。ただし、簡単のため、 N を N_α と N_β の小さい方としている。また、 $t_n^\alpha = t_n^\beta$ としている。さらに、時系列データの定常性を仮定している。 $r_{\alpha\beta}$ が求められれば、ある閾値以上、あるいは以下のペアを抽出し、それらを提示することでパターン抽出を支援できる。また、 $r_{\alpha\beta}$ ($\alpha < \beta$) に対して何かしらのクラスタリングを行うことも考えられる。

統計量を用いたクラスタリングに似たものとして、時系列データ間の類似度に基づいて、クラスタリングを行うという方法が考えられる。上述の統計量は \mathbf{S}^α と \mathbf{S}^β の最も単純な類似度（非類似度）であるが、 \mathbf{S}^α と \mathbf{S}^β のユークリッド距離もそのようなものである。本稿では、 \mathbf{S}^α と \mathbf{S}^β のユークリッド距離を $D_{\alpha\beta}$ で表す。3 成分を同時に考慮する場合は、

$$D_{\alpha\beta} = \sqrt{\sum_{n=0}^N \{ (u_n^\alpha - u_n^\beta)^2 + (v_n^\alpha - v_n^\beta)^2 + (w_n^\alpha - w_n^\beta)^2 \}} \quad (2)$$

で $D_{\alpha\beta}$ を算出する。 $r_{\alpha\beta}$ の場合と同様に、 N を N_α と N_β の小さい方とし、 $t_n^\alpha = t_n^\beta$ としている。また、成分毎に算出することも考えられる。この距離から類似のものを求めるために、 $D_{\alpha\beta}$ ($\alpha < \beta$) に対して何かしらのクラスタリングを行う。

上述の方法では、データ数が多ければ類似のものをクラスターとして抽出できるだろう。クラスター内、およびクラスター間での類似性、あるいは相違性から特徴的なパターンを見つけるという方法が考えられる。

2.2 ベクトル表示からのパターン抽出

前節で述べた方法は、全体の傾向に基づいて詳細に調べる時系列速度データの組を求め、絞り込まれたデータに対して、分析者の判断に基づく比較からパターン抽出を行うというものである。

成分毎、あるいは成分をまとめてデータを絞り込むため、(i) 速度ベクトルとしてのパターン抽出が難しいこと、(ii) 部分的な類似性を推定できないこと、(iii) 位相差によって生じる非類似性の考慮が難しいこと、(iv) 一般的には、 $t_n^\alpha \neq t_n^\beta$ であることなどの問題がある。

本節では、速度ベクトルに注目したパターン抽出の方法を示す。Fig.3 に示す 2 次元円柱を過ぎる流れのある位置での速度の時間変化を例として説明する。データは間引いている。

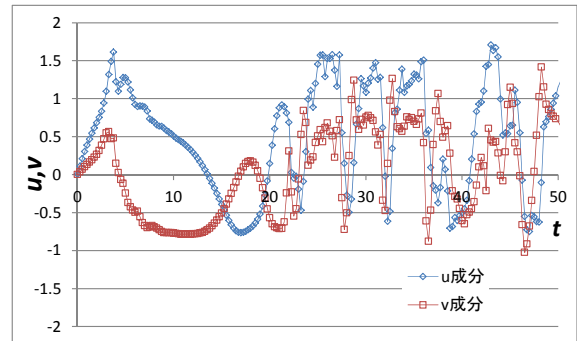


Fig. 3 速度成分の時間変化

Fig.3 に示す時系列速度データの別の表現方法としてホドグラフによるものがある。速度ベクトルの始点を一致させ、終点を結んでいくというものである。Fig.4 に Fig.3 のホドグラフ表現の一例を示す。また、Fig.5 に、成分を正規化した、方向ベクトルに対するホドグラフ表現の一例を示す。特に、Fig.5 からわかるように短時間での方向変化が捉えやすい。

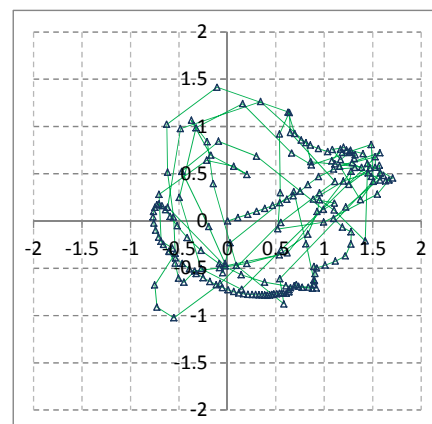


Fig. 4 ホドグラフ表現の一例（速度ベクトル）

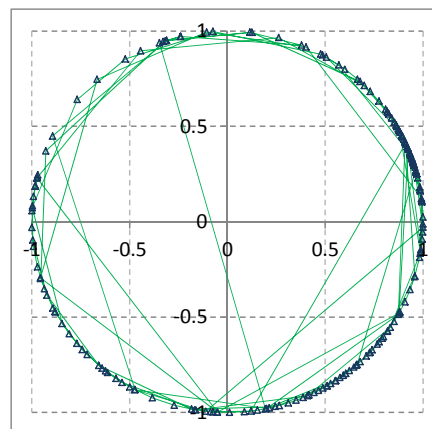


Fig. 5 ホドグラフ表現の一例（方向ベクトル）

ホドグラフ表現によって、方向の時間変化や速さの変化に対するパターン抽出が可能になる。しかし、データ量が増えるとパターンは煩雑になり、時間変化を追うことができなくなる。

そこで、時間情報の圧縮を考える。Fig.6 に示す始点を揃えたベクトルの表示が一つの方法である。左図は、Fig.3 に対応するもの ($0 \leq t \leq 50$)、右図は、 $0 \leq t \leq 300$ のものを示している。

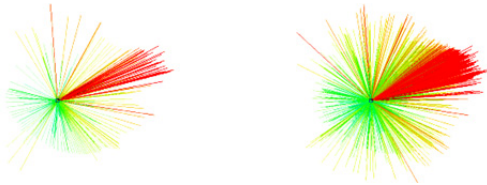


Fig. 6 時系列速度ベクトルの直接表示

色は速度の大きさを示しており、速度の大きいベクトルが偏って存在し、その方法が一つの特徴になっていることがわかる。一方、右図では重畳表示のため、時間区間全体の傾向が読み取りにくい。ここでは、文献(7)で提案したピクセル露光法の拡張(拡張ピクセル露光法⁸⁾)を利用した可視化を行う。Fig.7 は拡張ピクセル露光法の模式図である。Fig.6 に対して拡張ピクセル露光法を適用した例を Fig.8 に示す。

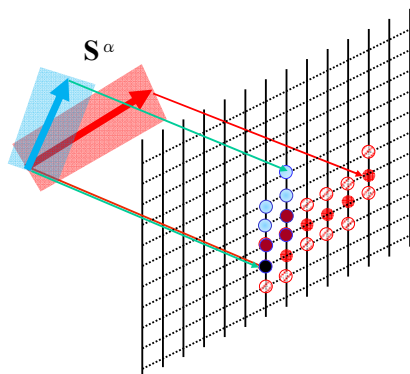


Fig. 7 拡張ピクセル露光法の模式図

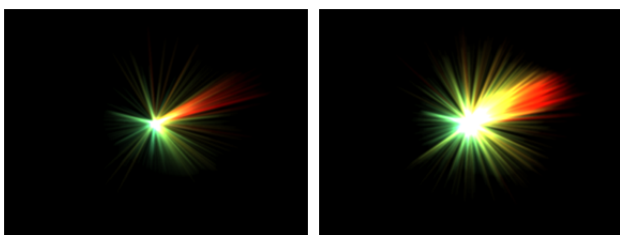


Fig. 8 拡張ピクセル露光法による時系列速度ベクトルの表示

2.3 擬似的移動軌跡によるパターン抽出

前節までの方法は、時系列速度データ全体から類似度や可視化を介して特徴的なパターンを抽出するというものである。それらの方法による部分的な類似性の推定は難しい。本節では、部分的な類似性を顕在化する方法を提案する。

はじめに、時系列速度データの集合 S^α から X_n^α を算出する。簡単のため、1次のオイラー積分を用いたものを示す。

$$X_0^\alpha = x^\alpha \quad (3)$$

$$X_n^\alpha = \sum_{l=0}^{n-1} \mathbf{u}_l^\alpha (t_{l+1}^\alpha - t_l^\alpha), \quad n=1, \dots, N_\alpha$$

方向ベクトルとする場合は、

$$X_0^\alpha = x^\alpha \quad (4)$$

$$X_n^\alpha = \sum_{l=0}^{n-1} \frac{\mathbf{u}_l^\alpha}{|\mathbf{u}_l^\alpha|} (t_{l+1}^\alpha - t_l^\alpha), \quad n=1, \dots, N_\alpha$$

とする。ここで、 $\mathbf{u}_l^\alpha = \mathbf{0}$ の場合は、方向ベクトルを $\mathbf{0}$ とする。

次に、式(3)、あるいは式(4)から得られた位置の時間変化を擬似的な移動軌跡とし、移動軌跡データの集合 P^α :

$$P^\alpha \{X_0^\alpha, \dots, X_n^\alpha, \dots, X_{N_\alpha}^\alpha\}$$

を形成する。Fig.3 に示す時系列速度データから求めた P^α を Fig.9 に示す。ここで、 $X_0^\alpha = (0,0)$ とした。図では、速度の大きさによって色付けている。擬似的に形成した移動軌跡データからのパターン抽出は、文献(9),(10)に示した方法によって行う。

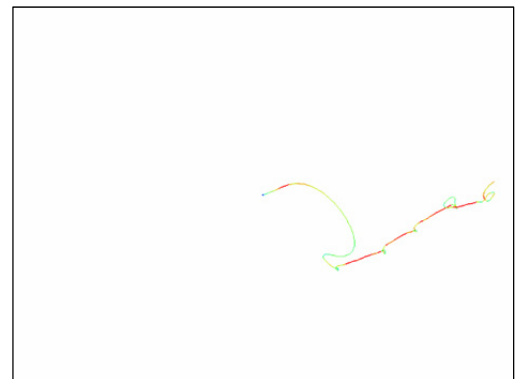


Fig. 9 擬似的移動軌跡による時系列速度ベクトルの表示

2.4 隠れマルコフモデルによる区間分割

本稿では、部分的な類似性を顕在化する方法として、2.3 節で述べた擬似的移動軌跡からのパターン抽出に加えて、隠れマルコフモデルを利用した区間分割法を提案する。

ある確率分布に従って信号を出力するいくつかの信号源(状態)の確率的な切り替わりで時系列データが生成されると仮定する。このとき、隠れマルコフモデル(HMM : Hidden Markov Model)によって、状態の時間変化が推定できる。文献(11)を参照し、Fig.10 に3状態が切り替わる場合を示す。図の下端に示すように3つの状態が赤色で示すように変化したとき、中段の状態系列を介して上段の時系列データが生成される。

HMM では、上段の時系列データを入力系列として、下段の状態の切り替わりを推定する。この切り替わりの時間を求めることで時系列速度データの区間分割を行う。

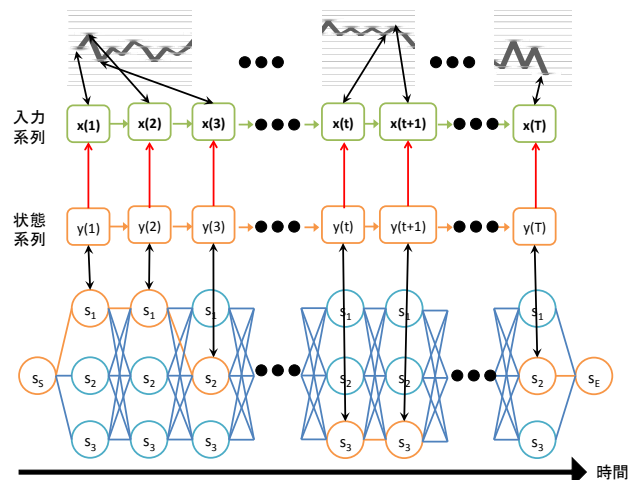


Fig. 10 マルコフモデルによる時系列データの表現⁽¹¹⁾

3. 実験結果と考察

2.1 節の方法は、特徴パターン抽出の第 1 段階でのデータのスクリーニングに用いるもので、特徴パターン抽出は分析者の判断による場所が大きい。また、多くの時系列速度データが得られたときに有効になると考えられる。本稿では、2.2 節と 2.3 節の方法を用いた結果を示す。2.4 節で提案した HMM による区間分割の結果については講演時に示す。

Fig.11 の青色の■で示す 2 次元円柱後方の 15 点での時系列速度データからのパターン抽出を試みる。格子は間引いて表示している。レイノルズ数は 6000、円柱直径を 1 としたときの外部領域の大きさは 50、格子点数は 500×360 である。また、時間刻み幅は 0.002 としている。計算手法は文献(12)と同じものを用いた。無次元時間で 50 までの非定常計算を行い、時系列速度データを求める。

Fig.12 と Fig.13 に、15 点における t が 0 から 50 までの u 成分と v 成分の時間変化を示す。 $N_a=4999$ ($\alpha=0, \dots, 14$) となる。また、等間隔で間引き、時間間隔を 0.01 としている。

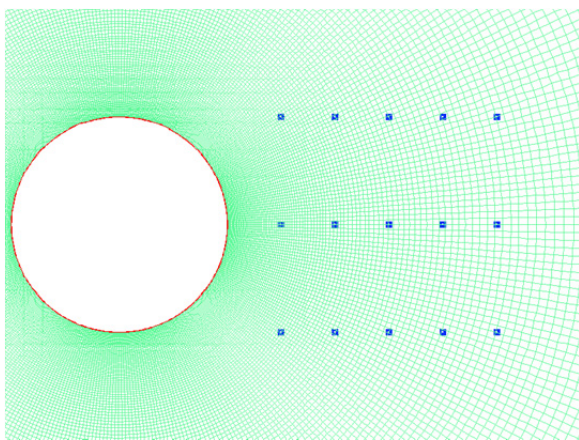


Fig. 11 計算格子と時系列データの抽出位置

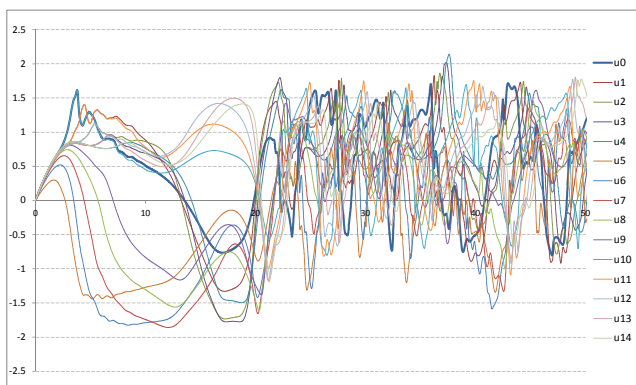


Fig. 12 u 成分の時間変化

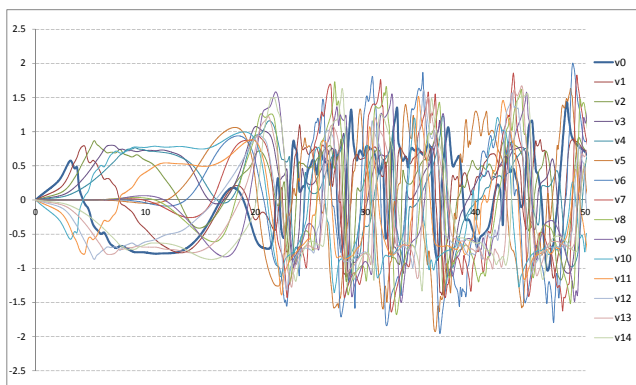
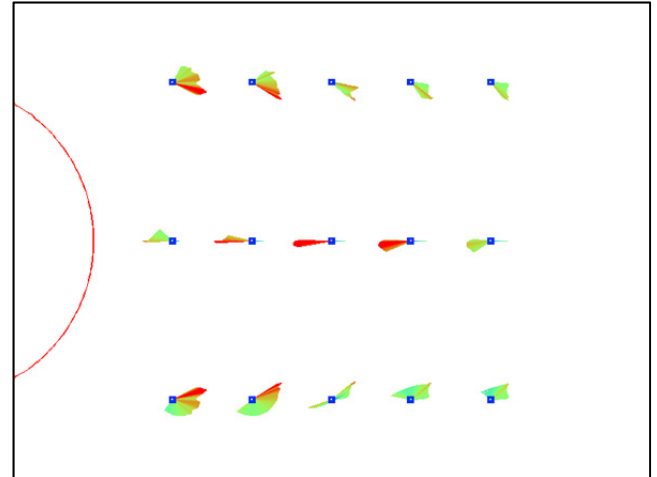


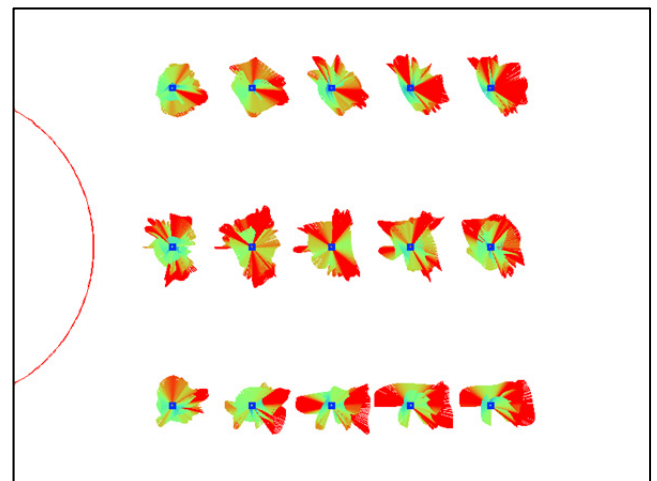
Fig. 13 v 成分の時間変化

Fig.14 に 15 点における時系列速度ベクトルの直接表示を示す。

図(a)は、 $0 \leq t \leq 15$ のデータに対するものである。図(b)は $0 \leq t \leq 50$ のものである。図(a)はほぼ対称的なパターンであるが、図(b)はそれぞれの位置で分布が偏っていることがわかる。ただし、重量のため、頻度が高い方向を読み取ることはできない。そこで、拡張ピクセル露光法を適用する。結果を Fig.15 に示す。Fig.14(b)においては、頻度が高く見える方向でも Fig.15 からはそうではないことなどがわかる。



(a) $0 \leq t \leq 15$



(b) $0 \leq t \leq 50$

Fig. 14 時系列速度ベクトルの直接表示

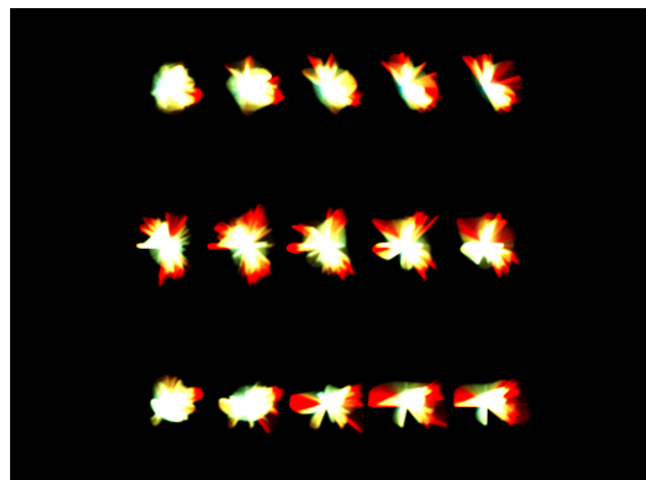


Fig. 15 拡張ピクセル露光法による時系列速度ベクトルの表示

Fig.16 は、擬似的な移動軌跡による表示である。図では、時系

列速度データ毎に色付けている。いくつかのグループに分けられること、類似度が高い部分があることなどがわかる。また、Fig.14 や Fig.15 の時系列速度ベクトルの表示よりも時間区間の特徴は捉えやすい。一方、位相差の影響が出やすいことも推察できる。

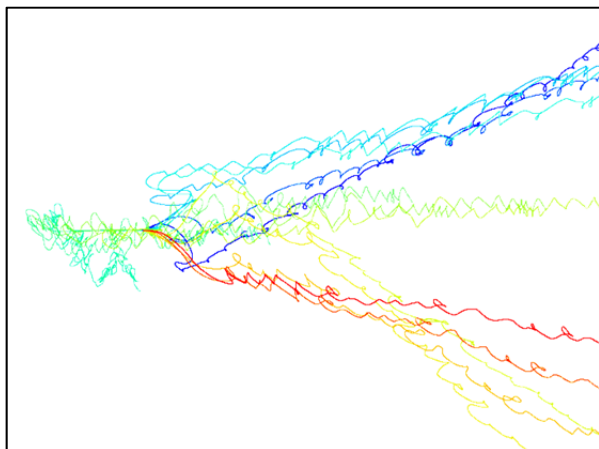


Fig. 16 擬似的移動軌跡による時系列速度ベクトルの表示

4. おわりに

計算、実験、観測において、膨大な非定常データが得られるようになってきている。一方、そのようなデータの分析においては、結果に対する定量性や客観性を確保することは未だに難しく、動画を利用するか、空間を固定し時間変動をグラフ化するという方法が用いられることが多い。前者では、定量的評価が問題となり、後者では、時間、また時空間の関連性を直感的に把握することが難しい。

本稿では、時系列速度ベクトルに対する分析結果の定量的評価と関連性の直感的把握において、特徴パターン抽出を介在させる必要性を指摘し、いくつかのパターン抽出法を提案した。

目的に応じたそれぞれのパターン抽出法の利害得失を明らかにすること、精度の低いデータの検証に役立てることが今後の課題である。

本研究は、RING!RING!プロジェクト⁽¹³⁾ 競輪の補助 (27-111) を受けて実施しました。ここに記して謝意を表します。

参考文献

- (1) 白山晋, 斎藤幸二郎, 大和裕幸, 増田宏, 安藤英幸, 比較空間を用いた CFD シミュレーションのデータ比較に関する研究, 日本造船学会論文集, 190 巻, (2001), pp.41-50.
- (2) 白山晋: 知的可視化, 丸善出版 (2006).
- (3) 新井啓之, 平岡克己: 低レイノルズ数領域でのプロペラブレード表面の圧力とタフト測定に基づく流体解析, 日本航空宇宙学会論文集, 57-662, (2009), pp.98-109.
- (4) Shawn, A., Andrzej, F., and Stephen, T., Flow separation on a high Reynolds number, high solidity vertical axis wind turbine with straight and canted blades and canted blades with fences, Renewable Energy, 41 (2012), pp.13-22.
- (5) Mizota, T., Kurogi, K., Ohya, Y., Okajima, A., Naruo, T., and Kawamura, Y., The strange flight behavior of spinning soccer balls, Scientific Reports (2013), 3:1871, DOI:10.1038/srep01871.
- (6) 白山晋, 尾亦範泰, 安達健二, 実環境下での 3 次元流れ方向の計測手法, 可視化情報学会論文集, 35-8 (2015), pp.17-27.
- (7) Shirayama, S. and Ohta, T., A Visualization of a Vector-field by a Homogenized Nascent-particles Tracking, Journal of Visualization 4.2 (2001), pp.185-196.
- (8) 白山晋, 移動軌跡データの新しい可視化法, 計算工学講演会

論文集, 14 (2009), B-3-1.

- (9) 白山晋, 移動軌跡データからのパターン形成, 第 37 回可視化情報シンポジウム講演論文集 (2009), pp.339-342.
- (10) 白山晋, 移動軌跡データのパターン分類による分析, 計算工学講演会論文集, 16 (2011), E-10-1.
- (11) 大森崇弘, 白山晋, 行動センシングデータを利用した作業負荷算出手法に関する研究, 人工知能学会第 2 種研究会資料, (2014), SIG-KST-2013-03-04.
- (12) Shirayama, S., Flow past a sphere: topological transitions of the vorticity fields, AIAA Journal, 30.2 (1992), pp.349 - 358.
- (13) RING!RING!プロジェクト, <http://ringring-keirin.jp/>