



卒業研究論文

論文題目

STEM 教育を目的とした
動画像処理による二重振り子の軌跡推定

提出年月日	令和 5 年 2 月 22 日	
学 科	電気情報工学科	
氏 名	岸本 麗央	印
指導教員（主査）	北村 大地 講師	印
副 査	重田 和弘 教授	印
学 科 長	辻 正敏 教授	印

香川高等専門学校

Trajectory Estimation of Double Pendulum Using Dynamic Image Processing for STEM Education

Reo Kishimoto

Department of Electrical and Computer Engineering
National Institute of Technology, Kagawa College

Abstract

STEM education is an educational method that aims at motivating young students' interest in mathematical science. A physical system called the double pendulum is often used as a subject for this education. In this subject, it will be demonstrated and explained that the double pendulum causes an interesting physical phenomenon known as initial value sensitivities and chaotic systems. Specifically, the phenomenon of unpredictable and complex trajectory of the motion of a double pendulum will be introduced. This motion is generally visualised by taking long-exposure photographs of a light source that is attached to the tip of a double pendulum. However, this method requires a dark environment, which may be demanding for the demonstration at a STEM-educational event. Furthermore, this method cannot be utilized when a double pendulum does not have a light source. In this thesis, I propose another method to visualise the trajectory of the tip of a double pendulum. In the proposed method, two situations are considered; with and without the light source at the tip of the double pendulum. The motion of a double pendulum is captured as a video, and its trajectory is estimated by applying dynamic image processing. First, a median filter is utilized to extract only the pixels of moving parts of the double pendulum. This preprocessing provides robust estimation of the trajectory. When the pendulum has a light source at the tip, the pixels of the light source are estimated based on its color information. On the other hand, when the pendulum does not have a light source, the tip location is estimated by calculating the direction of the pendulums. In the experiments, the estimation accuracy of each method is evaluated using objective criteria.

Keywords: STEM education, double pendulum, trajectory tracking

(和訳)

STEM 教育とは、若年層の数理科学系への興味の誘発を目的とした教育手法である。その題材として、二重振り子という物理系が題材に挙げられることがある。これは、二重振り子が初期値鋭敏性やカオス系と呼ばれる興味深い物理現象を引き起こすことを実演し、解説する内容である。具体的には、二重振り子の先端が予測できない複雑な軌跡を描くこととして紹介される。二重振り子の先端の軌跡の可視化には、先端につけた光源を長時間露光で撮影する方法が一般的である。しかしながら、この撮影は二重振り子の周辺を暗くする必要があり、STEM 教育としてこの二重振り子を実演する場所ではこのような撮影は難しい場合も多い。さらに、光源のついていない二重振り子の場合は、前述の方法による可視化は困難である。そこで本研究では、長時間露光による撮影以外の方法で二重振り子の先端の軌跡を可視化する新しい方法について検討する。具体的には、二重振り子の挙動を動画として撮影し、動画像処理を適用することで先端の軌跡を推定する。特に、二重振り子の先端に光源がある場合とない場合のそれぞれに適切な手法について検討する。いずれの場合においても、二重振り子の背景に対して頑健な推定を実現するために、メディアンフィルタを用いた背景と二重振り子の可動部の分離を施す。その後、二重振り子の先端に光源がある場合は、光源の色情報を元に先端位置を推定する。二重振り子の先端に光源がない場合は、二重振り子の回転軸及び連結部の位置を推定した結果から先端の位置を求める。実験では、これらの手法の推定精度について、客観尺度を用いて評価する。

目次

第 1 章 緒言	1
1.1 本論文の背景	1
1.2 本論文の目的	1
1.3 本論文の構成	4
第 2 章 基礎理論	5
2.1 はじめに	5
2.2 二重振り子の物理モデルと非線形微分方程式	5
2.3 移動メディアンフィルタ	7
2.4 移動合計値フィルタ	8
2.5 2 次元 3 次スプライン補間	9
2.6 本章のまとめ	12
第 3 章 撮影した動画の前処理	13
3.1 はじめに	13
3.2 動画の行列化	13
3.3 入力動画の背景の推定	14
3.4 バイナリマスクによる動体の分離	16
3.5 本章のまとめ	19
第 4 章 色情報を用いた軌跡推定法	20
4.1 はじめに	20
4.2 動機と処理の概要	20
4.3 HSV 色空間上で画素分類	21
4.4 振り子の先端位置の推定	22
4.5 本章のまとめ	23
第 5 章 色情報を用いた軌跡推定法の実験	24
5.1 はじめに	24
5.2 実験条件	24
5.3 実験結果	25

5.4	本章のまとめ	26
第 6 章	概形情報を用いた軌跡推定法	29
6.1	はじめに	29
6.2	動機と処理の概要	29
6.3	回転軸と連結部の位置推定	30
6.4	先端の位置推定	32
6.5	本章のまとめ	36
第 7 章	概形情報を用いた軌跡推定法の実験	37
7.1	はじめに	37
7.2	実験条件	37
7.3	実験結果	37
7.4	本章のまとめ	39
第 8 章	結言	40
	謝辞	41
	参考文献	41

第1章

緒言

1.1 本論文の背景

Science, technology, engineering, and mathematics (STEM) 教育とは、科学、技術、工学、及び数学の分野に力を入れた教育方針である。本教育方針は、科学技術の発展に役立つ人材を育てることを目的としている。また、自発的な行動や、独自性のある創造が可能な人材を増やす狙いもあり [1, 2]、代表例が高等専門学校（高専）の設立である。高専は、より若い年齢層に対する数理科学分野の専門的な教育と技術者の育成を目的として設立された高等教育機関であり、これまで多くの実践的な技術者を輩出している。高専出身者の多様な分野での活躍は、理系科目に強い人材及び自発性や創造性に富んだ人材の育成という、STEM 教育の目的の達成を表しており、STEM 教育の成功例の一つといえる。

前述のように、STEM 教育の実施は、理系分野に長けた人材だけでなく、世の中をリードする人材の育成に貢献している。近年では、初等・中等教育機関におけるプログラミング教育の必修化のように、より若い年齢層に対する STEM 教育も計画・実施されており、STEM 教育を目的としたイベントも多く実施されている [3, 4, 5, 6]。例えば、ペットボトルロケットの飛翔、ドライアイスを用いた冷却実験、空気砲の製作、ロボットの簡単な制御等が STEM 教育を目的としたイベントの題材によく挙げられる。

1.2 本論文の目的

STEM 教育を目的としたイベントは、その対象となる年齢層（例えば、小学生及び中学生）の数理科学分野への興味の誘発を目的として実施される。より効果的な STEM 教育を達成するため、簡単かつ理解しやすい教材や教育手法の開発が検討されている。

STEM 教育で扱われる題材の 1 つに、Fig. 1.1 に示すような二重振り子がある [7]。これは、物理学における振り子運動と数学における非線形微分方程式をテーマとしたものであり、不思議な物理現象を持つことで有名である。具体的には、二重振り子の運動の物理モデルが非線形微分方程式で表されることに起因して、（確率的ではなく）決定的な系であるにもかかわらず、複雑・不規則・不安定なふるまいを示し、運動の長期的な予測が全く困難となる現象である。

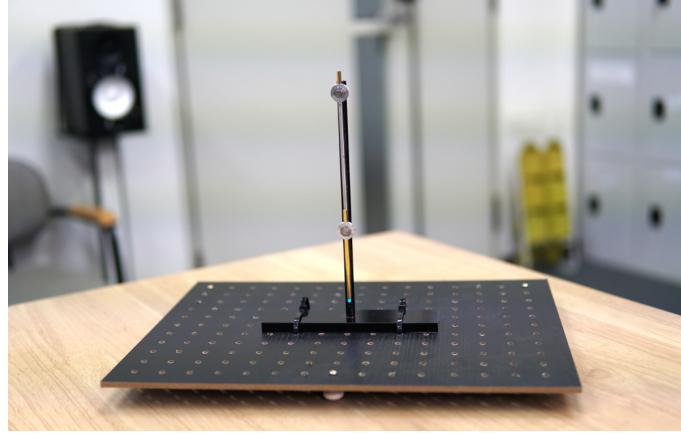
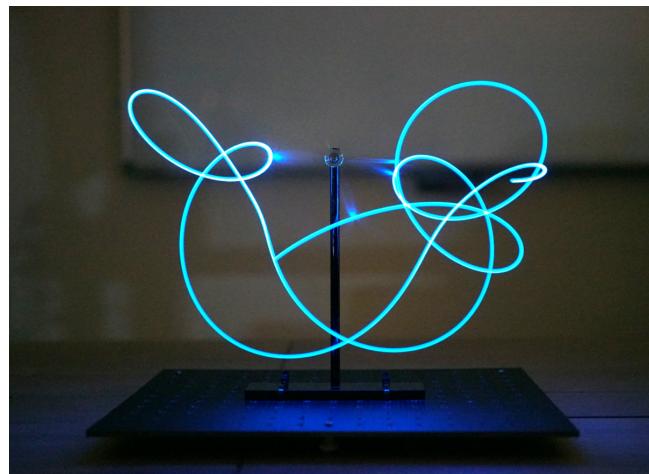


Fig. 1.1: Double pendulum.

例えば、二重振り子の初期条件（初期位置）をごくわずかに変化させただけで、その後の運動の結果（二重振り子の先端の軌跡）は劇的に変化する。そして、このような現象はカオス系や初期値鋭敏性等と呼ばれ、天体の運動や気象予測等、いたるところに存在する。この初期値鋭敏性を寓意的に表した言葉として「バタフライ効果」がある。これは、気象が非線形微分方程式でモデル化されるカオス系であることから、「ブラジルの1匹の蝶のはばたきによる小さな気流の攪乱が、やがてはテキサスでの竜巻を引き起こす結果になり得る」ということを述べている[8]。この寓意の妥当性は不明だが、内容は初期値鋭敏性の本質を表している。そしてこのような興味深い現象を、若年層の数理科学分野への興味の誘発を目的とした STEM 教育を通して分かりやすく伝えるために、二重振り子は説明や実演がしやすい題材として頻繁に選ばれるものである。一度揺らした二重振り子の先端の軌跡は、その高い初期値鋭敏性から、（現実的には）二度と再現することのできない一期一会の現象となることを、理論的な背景とともに説明することが目的となる。

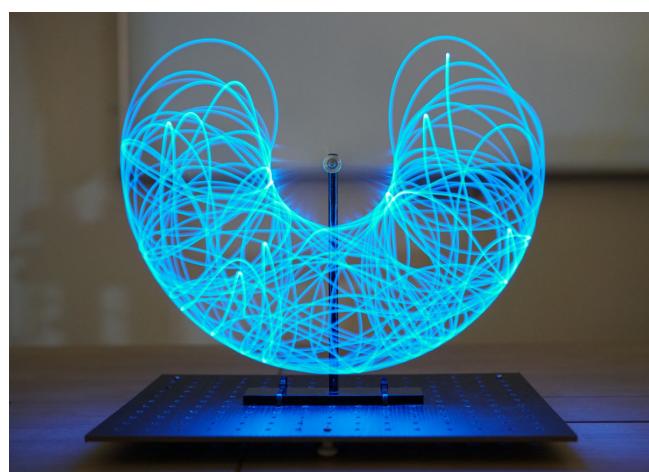
前述の二重振り子における初期値鋭敏性を STEM 教育の題材として扱う際は、初期値鋭敏性による振り子先端の運動（軌跡）の変化を可視化するため、振り子先端の軌跡を何らかの形で記録する方法が一般的である。多くの場合、二重振り子の先端に発光ダイオード (light-emitting diode: LED) を取り付けて発光させ、カメラなどの撮像機器で長時間露光させて撮影する方法がとられる。この方法で得られた軌跡の写真を Fig. 1.2 に示す。



(a) One-s-long exposure



(b) Eight-s-long exposure



(c) 20-s-long exposure

Fig. 1.2: Double pendulum trajectory photographs taken by long exposure.

しかしながら、このような長時間露光による撮影は、可能な限り周囲を暗くする必要があるため、STEM 教育の現場では実施が難しい場合がある。その他、先端につけたレーザーを感光紙に感光させる方法、ジャイロセンサや加速度センサを二重振り子に接続する方法等が考えられるが、いずれの方法も機器の大型化を招き、また運動系自体に影響を与える問題もある。

本論文では、STEM 教育のイベント等の一般的な環境でも実施しやすく、可能な限り簡便な装置で二重振り子の先端の軌跡を可視化する方法について検討する。また、その可視化の精度は、長時間露光による撮像と同程度を目指す。具体的には、(a) 二重振り子の先端に LED が装着されており何らかの色で発光させることができる場合及び (b) 二重振り子の先端を発光させることができない場合の 2 条件のそれぞれにおいて、動画を撮影した後に動画像処理で二重振り子の先端を追跡し、軌跡を可視化する方法を提案する。

1.3 本論文の構成

2 章では、提案手法を理解するために重要な基礎知識として、物理モデルである非線形微分方程式と初期値鋭敏性について説明する。また、動画の背景と二重振り子の可動部を分離するために用いる基礎技術として、メディアンフィルタについて説明する。さらに、2 次元座標上の複数の時系列点を滑らかに補間する手法として、2 次元スプライン補間を説明する。3 章では、2 章で説明した基礎技術を用いて動画の背景と二重振り子の可動部を分離する具体的な方法について述べる。4 章では、二重振り子の先端に LED が装着されており、その色情報を先端の軌跡推定に利用できる場合の提案手法について説明する。5 章では、4 章で述べた提案手法を用いて実験を行い、推定された軌跡を正解データと比較することで、手法の性能を評価する。6 章では、色情報を利用できない場合、すなわち 4 章で説明する手法が適用できない場合でも軌跡推定が可能な提案手法について説明する。7 章では、6 章で述べた提案手法を用いて実験を行い、推定された軌跡を色情報を利用して推定した軌跡と比較することで、手法の性能を評価する。8 章では、本論文の結果についてまとめ、今後の課題を述べる。

第 2 章

基礎理論

2.1 はじめに

本章では、最初に 2.2 節で二重振り子の運動系そのものについて説明し、非線形微分方程式や初期値鋭敏性について述べる。2.3 節では、入力動画の背景と動体の分離のための要素技術であるメディアン処理と移動メディアンフィルタを説明する。2.4 節では、6 章で述べる二重振り子の概形情報を用いた軌跡推定法で用いる移動合計値フィルタについて説明する。2.5 節では、動画の各時間フレームにおける二重振り子の先端位置の推定結果をプロットした時系列の座標に対して、滑らかな 2 次元曲線を推定する 2 次元スプライン補間の理論について述べる。2.6 節で本章についてまとめる。

2.2 二重振り子の物理モデルと非線形微分方程式

微分方程式等での記述が可能で、解の存在と一意性が保証された力学系のことを決定論的システムという。これは、初期値が与えられれば、そこを通る解がただ 1 つに定まるということである。これにより、決定論的システムが行う運動は原理的には予測可能となるが、この主張は現実的には困難性を含むことが、カオス現象の発見によって示された [9]。決定論的システムであるにもかかわらず、初期値鋭敏性によって、運動の長期的な予測ができない系をカオス力学系といい、二重振り子の運動はこれに該当する [10, 11]。

本節では、二重振り子の運動方程式について紹介し、初期値鋭敏性について説明する。Fig. 2.1 に示すような、第 1 の腕の長さを l_1 、第 2 の腕の回転軸（ベアリング）の質量を m_1 、第 2 の腕の長さを l_2 、及び二重振り子先端部の質量を m_2 とする二重振り子を考える。鉛直下向きと腕がなす角をそれぞれ θ_1 及び θ_2 とした際の、第 2 の腕の回転軸と二重振り子先端のカルト座標をそれぞれ (x_1, y_1) 及び (x_2, y_2) とすると、これらは次式で表される。

$$(x_1, y_1) = (l_1 \sin \theta_1, -l_1 \cos \theta_1) \quad (2.1)$$

$$(x_2, y_2) = (l_1 \sin \theta_1 + l_2 \sin \theta_2, -l_1 \cos \theta_1 - l_2 \cos \theta_2) \quad (2.2)$$

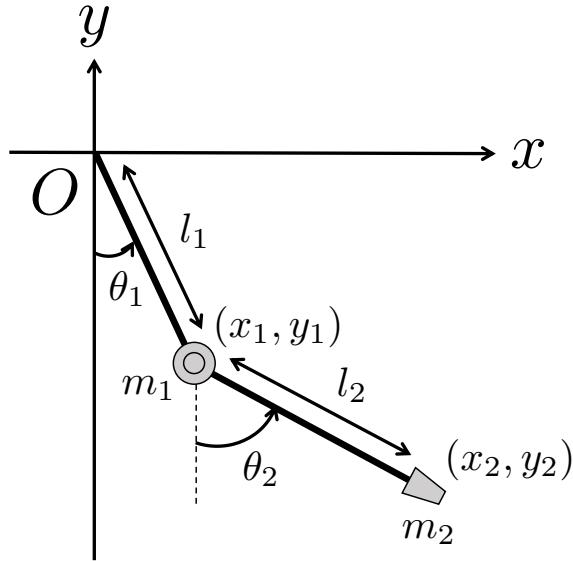


Fig. 2.1: Definition of variables in double pendulum.

式(2.1)及び式(2.2)を時間微分するとそれぞれ次式のようになる。

$$\left(\frac{dx_1}{dt}, \frac{dy_1}{dt} \right) = \left(l_1 \frac{d\theta_1}{dt} \cos \theta_1, l_1 \frac{d\theta_1}{dt} \sin \theta_1 \right) \quad (2.3)$$

$$\left(\frac{dx_2}{dt}, \frac{dy_2}{dt} \right) = \left(l_1 \frac{d\theta_1}{dt} \cos \theta_1 + l_2 \frac{d\theta_2}{dt} \cos \theta_2, l_1 \frac{d\theta_1}{dt} \sin \theta_1 + l_2 \frac{d\theta_2}{dt} \sin \theta_2 \right) \quad (2.4)$$

ここで、ラグランジアンとラグランジュ方程式を次式に示す。

$$L = K - U \quad (2.5)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} = \frac{\partial L}{\partial \theta} \quad (2.6)$$

式(2.5)のラグランジアン L は式(2.6)のラグランジュ方程式の導入にあたって定義された量であり、 K は運動エネルギー、 U は位置エネルギーである。式(2.6)のラグランジュ方程式は、扱う座標に依存することなく成り立つ運動方程式であり、デカルト座標という特別な座標系を用いる必要があるニュートンの運動方程式の一般化である。よって、ラグランジュ方程式をデカルト座標系で用いる場合、ニュートンの運動方程式と一致する[12, 13]。 $\dot{\theta}$ は θ の1階時間微分を表す。第2の腕の回転軸及び二重振り子先端の運動エネルギー及び位置エネルギーを、それぞれ K_1, U_1, K_2 及び U_2 とすると、これらは次式で表される。

$$K_1 = \frac{1}{2} m_1 \left\{ \left(\frac{dx_1}{dt} \right)^2 + \left(\frac{dy_1}{dt} \right)^2 \right\} \quad (2.7)$$

$$U_1 = m_1 g y_1 \quad (2.8)$$

$$K_2 = \frac{1}{2} m_2 \left\{ \left(\frac{dx_2}{dt} \right)^2 + \left(\frac{dy_2}{dt} \right)^2 \right\} \quad (2.9)$$

$$U_2 = m_2 g y_2 \quad (2.10)$$

ここで, g は重力加速度を表す. また, Fig. 2.1 の二重振り子のラグランジアンは次式のようになる.

$$L = K_1 + K_2 - U_1 - U_2 \quad (2.11)$$

式 (2.7)–(2.10) に式 (2.1)–式 (2.4) を代入し, さらにそれらを式 (2.11) に代入すると次式のようになる.

$$\begin{aligned} L(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2) &= \frac{1}{2}m_1 \left\{ \left(\frac{dx_1}{dt} \right)^2 + \left(\frac{dy_1}{dt} \right)^2 \right\} + \frac{1}{2}m_2 \left\{ \left(\frac{dx_2}{dt} \right)^2 + \left(\frac{dy_2}{dt} \right)^2 \right\} \\ &\quad - m_1 gy_1 - m_2 gy_2 \\ &= \frac{1}{2}m_1 l_1^2 \dot{\theta}_1^2 + \frac{1}{2}m_2 \left\{ l_1^2 \dot{\theta}_1^2 + l_2^2 \dot{\theta}_2^2 + 2l_1 l_2 \cos(\theta_1 - \theta_2) \dot{\theta}_1 \dot{\theta}_2 \right\} \\ &\quad + m_1 g l_1 \cos \theta_1 + m_2 g (l_1 \cos \theta_1 + l_2 \cos \theta_2) \end{aligned} \quad (2.12)$$

θ_1 及び θ_2 について, ラグランジュ方程式より,

$$\begin{cases} \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}_1} = \frac{\partial L}{\partial \theta_1} \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}_2} = \frac{\partial L}{\partial \theta_2} \end{cases} \quad (2.13)$$

式 (2.13) に式 (2.12) を代入すると, 連立方程式は次のように整理される.

$$\begin{cases} \frac{d}{dt} \left\{ (m_1 + m_2)l_1^2 \dot{\theta}_1 + m_2 l_1 l_2 \dot{\theta}_2 \cos(\theta_1 - \theta_2) \right\} = -m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \sin(\theta_1 - \theta_2) \\ \quad - (m_1 + m_2) g l_1 \sin \theta_1 \\ \frac{d}{dt} \left\{ m_2 l_2^2 \dot{\theta}_2 + m_2 l_1 l_2 \dot{\theta}_1 \cos(\theta_1 - \theta_2) \right\} = m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \sin(\theta_1 - \theta_2) \\ \quad - m_2 g l_2 \sin \theta_2 \end{cases} \quad (2.14)$$

式 (2.14) は, 微分方程式の変数である θ_1 及び θ_2 の積 (例えば $\dot{\theta}_1 \dot{\theta}_2$ 等) が含まれており, このような微分方程式は非線形微分方程式と呼ばれる.

一般に非線形微分方程式は解析的に解くことができないため, 数値的に解く (数値積分を用いる) 方法が一般的である. しかしながら, 数値的な計算を行う過程で, 初期条件に極めて微小な差 (倍精度浮動小数点の正の最小の正規化数である $2^{-1022} \approx 2.225 \times 10^{-308}$ よりも微小な値であっても) が生じただけでも, 求まる解は全く異なる値となる [9]. 従って, (2.14) の解は鋭敏な初期値依存性をもつシステムであり, そのようなシステムの運動が長期的には全く予測できないことを意味している.

2.3 移動メディアンフィルタ

移動メディアンフィルタは, 窓と呼ばれる信号中の特定の範囲内で中央値を返す処理 (メディアン処理) を, 窓を移動させながら適用していく非線形フィルタである. このような処理は, 信号に含まれる外れ値を除去することができる.

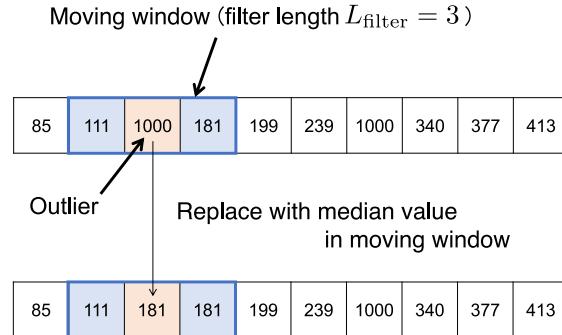


Fig. 2.2: Application of moving median filter for one-dimensional signals.

今、1次元の信号をベクトルとして $\mathbf{m} = [m_1, m_2, \dots, m_n, \dots, m_N]^T \in \mathbb{R}^N$ と定義する。ここで、 $n = 1, 2, \dots, N$ は信号の離散時間インデクスである。この信号 \mathbf{m} に対して、フィルタ長 L_{filter} の移動メディアンフィルタを適用する処理を次式のように定義する。

$$\hat{\mathbf{m}} = \text{movmedian}_{L_{\text{filter}}}(\mathbf{m}) \quad (2.15)$$

ここで、 $\hat{\mathbf{m}} = [\hat{m}_1, \hat{m}_2, \dots, \hat{m}_n, \dots, \hat{m}_N]^T \in \mathbb{R}^N$ は移動メディアンフィルタ適用後の信号であり、その要素 \hat{m}_n は次のように計算される。

$$\hat{m}_n = \text{median}(\tilde{\mathbf{m}}_n) \quad \forall n \quad (2.16)$$

また、 $\tilde{\mathbf{m}}_n \in \mathbb{R}^{L_{\text{filter}}}$ は信号 \mathbf{m} の時刻 n の時間サンプル m_n を中心とする前後の信号値を L_{filter} 点だけ切り出した短時間区間信号（但し、 L_{filter} が偶数の場合や信号のインデクス 1～ N の範囲外の要素はそれぞれ適切に処理する）であり、 $\text{median}(\cdot)$ は入力されたベクトルの全成分の中央値（スカラー）を返す関数である。すなわち式 (2.15) は、Fig. 2.2 に示すように、信号 \mathbf{m} に対して長さ L_{filter} の窓で短時間信号として切り出し中央値を返す処理を 1 点ずつ移動させるものである。このような処理は、Fig. 2.2 のように長さ L_{filter} の範囲内に含まれる外れ値を中央値に置き換えるため、外れ値の除去に高い効果を発揮する。

メディアンフィルタを 2 次元の信号（画像）に適用する場合、Fig. 2.3 のように、2 次元の窓（カーネル）を考え、カーネルが重なった領域の中心画素の輝度値を、領域内の輝度値の中央値で置き換える。カーネルを 2 次元座標上で移動させて、画像の全画素に同様の操作を行う。Fig. 2.4 に、実際に外れ値のノイズを付与した画像とメディアンフィルタを適用して外れ値除去した画像を示す。適切なカーネルサイズのメディアンフィルタにより、外れ値のノイズがある程度除去されていることが分かる。

2.4 移動合計値フィルタ

移動メディアンフィルタと同様に移動窓を用いて（中央値ではなく）合計値を返すフィルタを定義する。この処理は 6 章で述べる二重振り子の概形情報を用いた軌跡推定法で用いる。

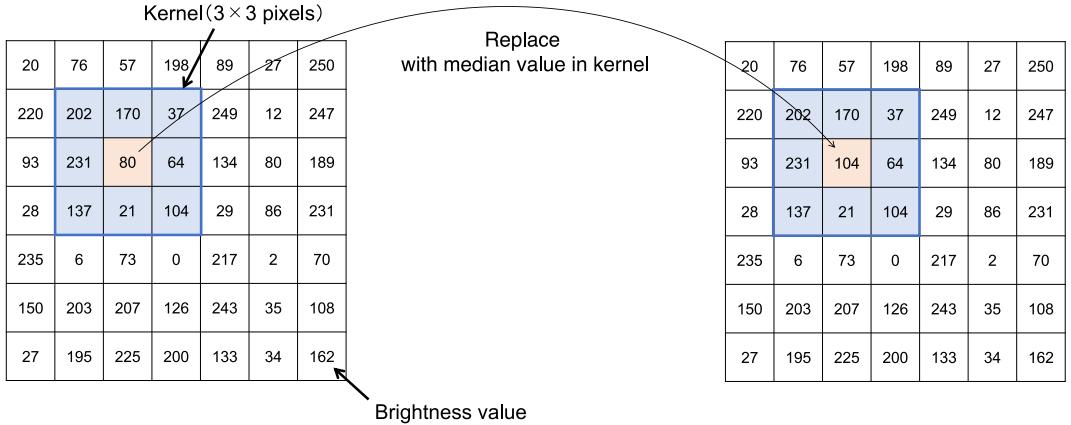


Fig. 2.3: Application of median filter for two-dimensional signals.

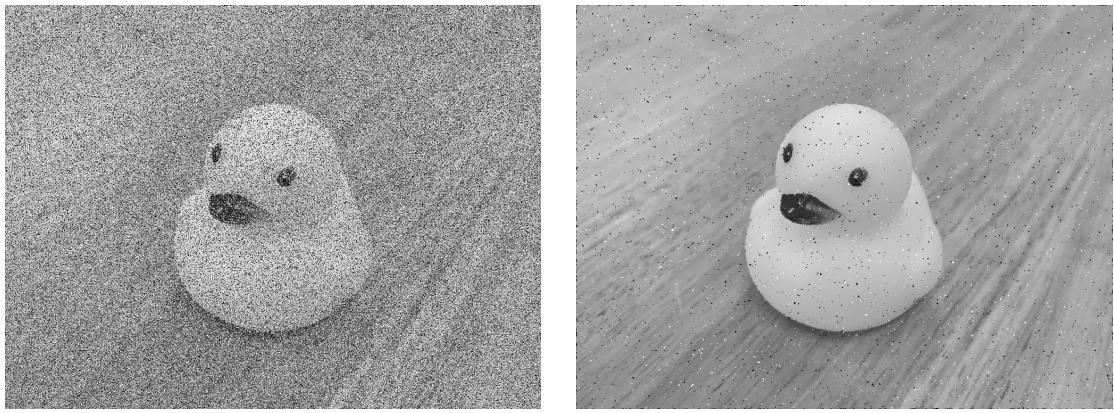


Fig. 2.4: Example of applying median filter for sparse noise removal.

信号 s に対するフィルタ長 L_{filter} の移動合計値フィルタを次式で表す.

$$\hat{s} = \text{movsum}_{L_{\text{filter}}}(s) \quad (2.17)$$

$$\hat{s}_n = \text{sum}(\tilde{s}_n) \quad \forall n \quad (2.18)$$

ここで, $\text{sum}(\cdot)$ は入力されたベクトルの成分の合計値を返す関数である. この移動合計値フィルタの適用例を Fig. 2.5 に示す.

2.5 2 次元 3 次スプライン補間

スプライン補間は、離散データ間で微分に関する連続性条件を持った区分的多項式のスプライン関数を用いて離散データの補間（内挿）を行う. これは、データ点全体を单一の多項式で補間すると、不自然な振動が発生しやすい（ルンゲ現象）ためである [14].

本節では、補間曲線の中で最良とされる 3 次自然スプライン関数について説明する [14].

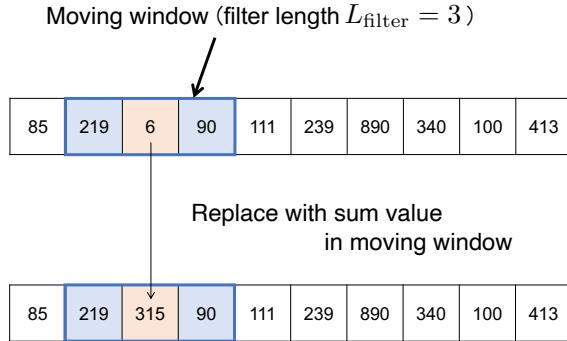


Fig. 2.5: Application of moving sum filter for one-dimensional signals.

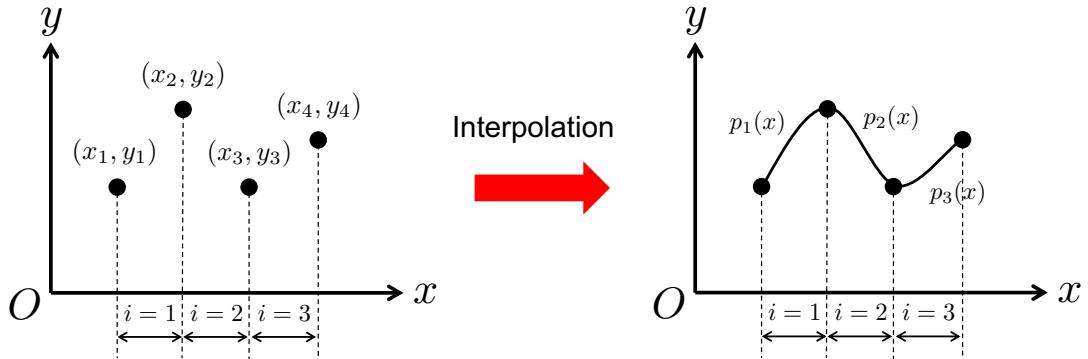


Fig. 2.6: Application of cubic spline interpolation to one-dimensional signals.

Fig. 2.6 のような 1 次元の時系列信号では、 i 番目の区間を補間する多項式 $p_i(x)$ ($i = 1, 2, \dots, I$ は区間を表すインデックス) は次式のようになる。

$$p_i(x) = x_i^3 c_3^{(i)} + x_i^2 c_2^{(i)} + x_i c_1^{(i)} + c_0^{(i)} \quad (2.19)$$

ここで、式 (2.19) の内、 $c_0^{(i)}$ 、 $c_1^{(i)}$ 、 $c_2^{(i)}$ 、及び $c_3^{(i)}$ はそれぞれ 0 次項、1 次項、2 次項、及び 3 次項の係数である。 i 番目の多項式は、式 (2.20) に示す補間条件に当てはまるよう、 i 番目のデータ点 (x_i, y_i) と $i+1$ 番目のデータ点 (x_{i+1}, y_{i+1}) を代入することで得られる連立方程式を解くことで求められる。

$$\begin{cases} y_i &= x_i^3 c_3^{(i)} + x_i^2 c_2^{(i)} + x_i c_1^{(i)} + c_0^{(i)} \\ y_{i+1} &= x_{i+1}^3 c_3^{(i)} + x_{i+1}^2 c_2^{(i)} + x_{i+1} c_1^{(i)} + c_0^{(i)} \end{cases} \quad (2.20)$$

さらに、3 次スプライン関数の定義より、次の 2 式に示すように、1 回微分と 2 回微分がともに区間の節点で連続である必要がある。

$$3x_i^2 c_3^{(i)} + 2x_i c_2^{(i)} + c_1^{(i)} = 3x_{i+1}^2 c_3^{(i+1)} + 2x_{i+1} c_2^{(i+1)} + c_1^{(i+1)} \quad (2.21)$$

$$3x_i c_3^{(i)} + c_2^{(i)} = 3x_{i+1} c_3^{(i+1)} + c_2^{(i+1)} \quad (2.22)$$

ここで、3 次スプライン関数全体で係数 $c_k^{(i)}$ を $4I$ 個 (I は区間数) 求める必要があるのに対し、条件は式 (2.20) の $2I$ 個、式 (2.21) の $I-1$ 個、式 (2.22) の $I-1$ 個の合計 $4I-2$ 個で

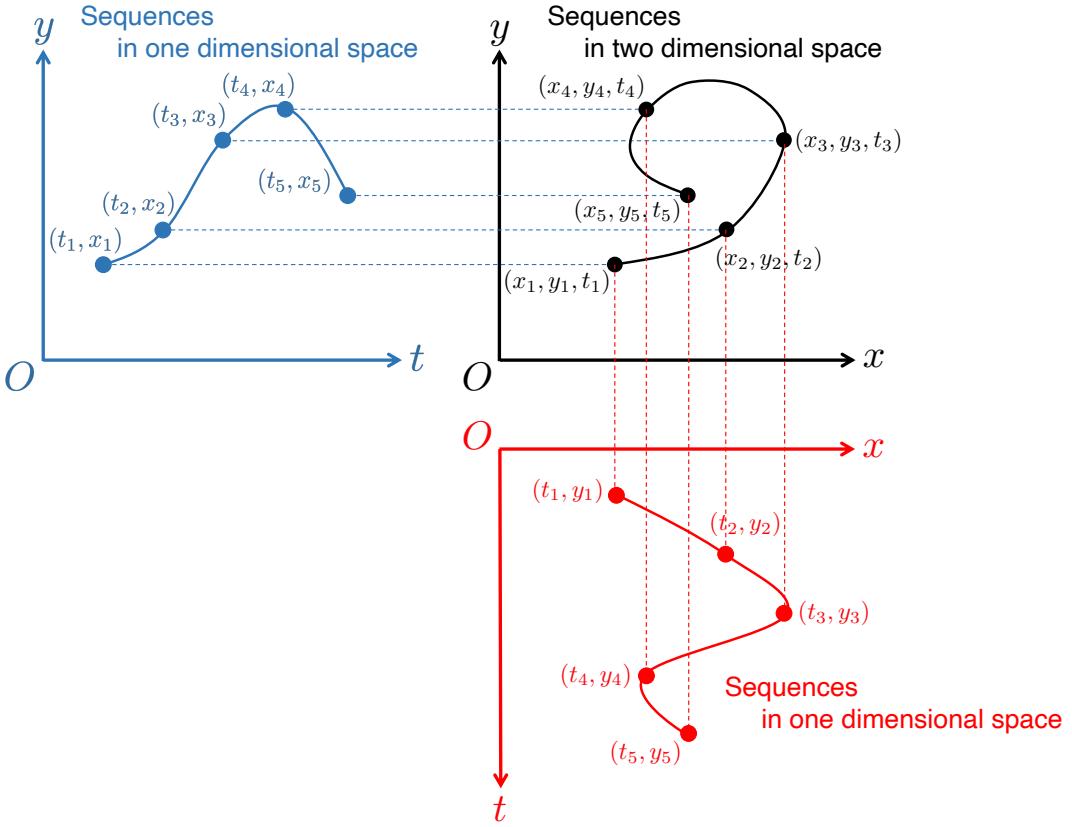


Fig. 2.7: Application of cubic spline interpolation to two-dimensional signals.

あるため、式があと 2 個必要となる。これは、両端の節点で微分値を指定する端条件によって解決できる。1 回微分の条件は両端の形状に与える影響が大きいため、式 (2.23) のような 2 回微分の条件を端条件とする。

$$\begin{cases} 3x_0c_3^{(1)} + c_2^{(1)} = 0 \\ 3x_0c_3^{(4)} + c_2^{(4)} = 0 \end{cases} \quad (2.23)$$

以上の式 (2.20)–(2.23) から成る連立方程式より、係数 $c_0^{(i)}$, $c_1^{(i)}$, $c_2^{(i)}$, 及び $c_3^{(i)}$ を解析的に求めることができ、各区間の曲線は式 (2.19) として得られる。

なお、この 3 次スプライン補間を 2 次元空間上の時系列信号に適用する場合、Fig. 2.7 のように、 x 軸及び y 軸で時間方向にそれぞれ 3 次スプライン補間を行う。得られた補間曲線を再び 2 次元空間上に戻せば、2 次元空間上の時系列信号の補間曲線が得られる。Fig. 2.8 に、実際に 2 次元空間上の時系列信号にスプライン補間を適用して得られた曲線を示す。 x 軸及び y 軸の時間方向のスプライン補間ににより、連続な 2 次元曲線が描かれたことが分かる。

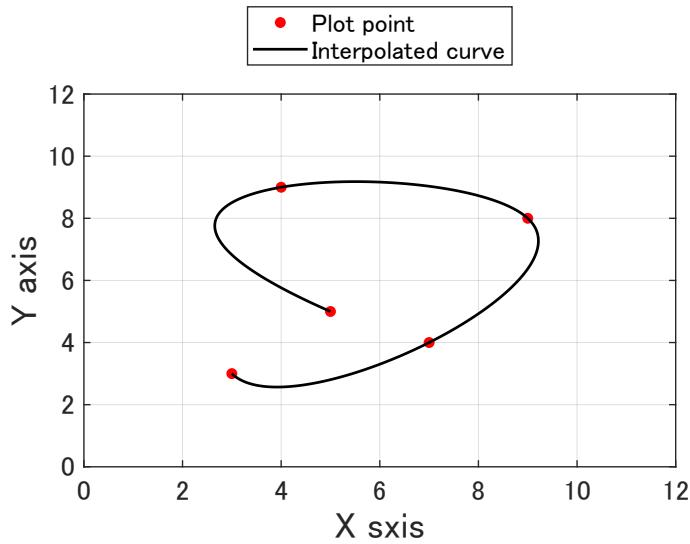


Fig. 2.8: Example of cubic spline interpolation in two-dimensional space.

2.6 本章のまとめ

本章では、二重振り子の軌跡推定が困難である原因の初期値鋭敏性や入力動画の背景と動体の分離のための要素技術であるメディアンフィルタについて説明した。また、二重振り子の概形情報を用いた軌跡推定法で適用する移動合計値フィルタや動画の各時間フレームで推定した二重振り子の先端位置のプロット点を滑らかな曲線で繋ぐための2次元3次スプライン補間について説明した。次章では、実際に撮影した入力動画での前処理について述べる。

第3章

撮影した動画の前処理

3.1 はじめに

4章や6章で述べる提案手法は、入力動画の可動部の情報（すなわち、二重振り子の動体部分の画素）のみを用いる。このとき、背景等の動体以外の画素は二重振り子先端の位置の推定精度の劣化を引き起こすため、その影響は可能な限り排除されるべきである。そこで提案手法では、入力動画に含まれる背景と動体を分離する事前処理の適用を考える。このとき、入力動画を2次元行列に整形し、時間方向にメディアンフィルタを適用することで分離を実現する。これらの手法を詳しく説明するために、本章では、入力動画のグレイスケール化と2次元行列化を3.2節、メディアンフィルタを用いた背景と動体の分離を3.3節で説明する。3.4節では、振り子の概形情報を扱う上で、各画素が振り子を表すか否かを判別するための2値化処理について述べる。3.5節で本章についてまとめる。

3.2 動画の行列化

8ビットカラーフォーマットで撮影された動画を $X^{(\text{color})} \in \{0, 1, \dots, 255\}^{W \times H \times 3 \times T}$ と定義する。ここで、 $w = 1, 2, \dots, W$ 及び $h = 1, 2, \dots, H$ はそれぞれ動画の水平方向及び垂直方向の位置インデックス（座標）である。また、 $X^{(\text{color})}$ の3の次元はそれぞれRGBのカラーチャネルを表し、 $t = 1, 2, \dots, T$ は動画の時間フレームのインデックスを表す。さらに、 $X^{(\text{color})}$ の成分値は0~255の整数値であり、輝度を表す。この動画 $X^{(\text{color})}$ をグレイスケールに変換した結果を $X \in \{0, 1, \dots, 255\}^{W \times H \times T}$ と定義する。一般にこの処理は次式で与えられる。

$$x_{w,h,t} = \text{round} \left(0.2989x_{w,h,1,t}^{(\text{color})} + 0.5870x_{w,h,2,t}^{(\text{color})} + 0.1140x_{w,h,3,t}^{(\text{color})} \right) \quad (3.1)$$

ここで、 $\text{round}(\cdot)$ は四捨五入処理を表し、 $x_{w,h,1,t}^{(\text{color})}$ 、 $x_{w,h,2,t}^{(\text{color})}$ 、及び $x_{w,h,3,t}^{(\text{color})}$ は $X^{(\text{color})}$ の要素 (R, G, 及び B)、 $x_{w,h,t}$ は X の要素を表す。従って、 $X^{(\text{color})}$ は4次元信号、 X は3次元信号となる。

以後の処理では、動画を2次元の信号（行列）として扱うため、ここではグレースケール動画 X を行列に変形する処理について説明する。Fig. 3.1 に示すように、動画の水平及び垂直の

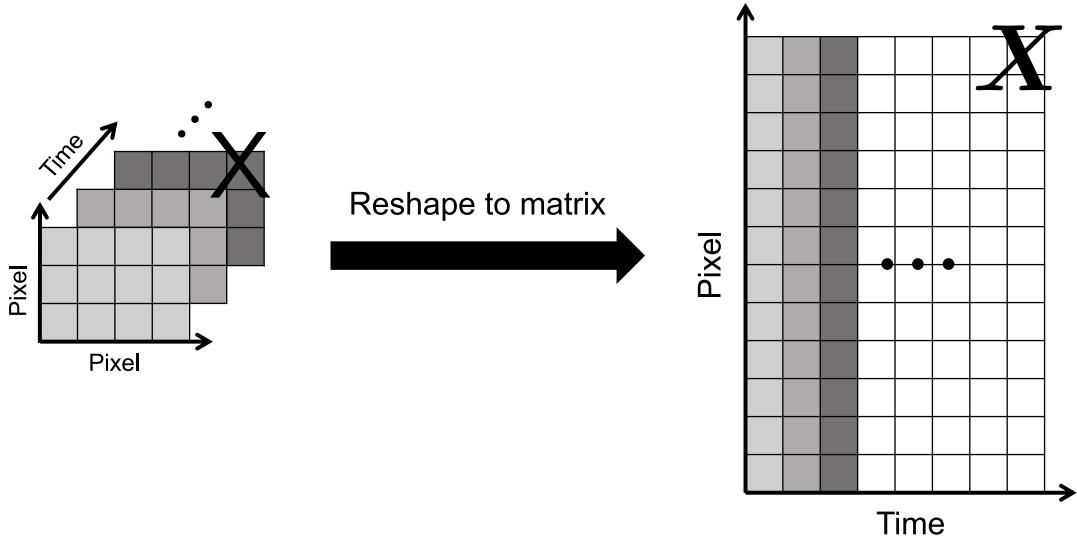


Fig. 3.1: Reshaping three-dimensional video signal to two-dimensional signal (matrix).

2次元座標を1次元に変形することで、行列に変形できる。この行列を以後動画行列と呼び、 $\mathbf{X} \in \{0, 1, \dots, 255\}^{WH \times T}$ で表す。動画行列への変形処理は次式となる。

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,T} \\ x_{2,1} & x_{2,2} & \ddots & x_{2,T} \\ \vdots & \ddots & \ddots & \vdots \\ x_{wh,1} & x_{wh,2} & \ddots & x_{wh,T} \\ \vdots & \ddots & \ddots & \vdots \\ x_{WH,1} & x_{WH,2} & \cdots & x_{WH,T} \end{bmatrix} \quad (3.2)$$

ここで、 $wh = 1, 2, \dots, WH$ は1次元に変換された画素を表すインデックスである。従って、動画行列 \mathbf{X} は行方向が水平及び垂直を含む画素、列方向が時間フレームとなる。

式 (3.2) の動画行列への変換処理は、カラー動画 $\mathbf{X}^{(\text{color})}$ の各カラーチャネルについても同様に適用できる。すなわち、 $\mathbf{X}^{(\text{color})}$ のカラーチャネルの信号をそれぞれ $\mathbf{X}^{(\text{red})}, \mathbf{X}^{(\text{green})}, \mathbf{X}^{(\text{blue})} \in \{0, 1, \dots, 255\}^{W \times H \times T}$ として \mathbf{X} と同様に構成すれば、式 (3.2) と同様の方法で $\mathbf{X}^{(\text{red})}, \mathbf{X}^{(\text{green})}, \mathbf{X}^{(\text{blue})} \in \{0, 1, \dots, 255\}^{WH \times T}$ が定義される。

3.3 入力動画の背景の推定

本節では、3.2 節で求めた動画行列 \mathbf{X} に対して動画像処理を適用し、動画中の背景と動体(二重振り子の可動部)を分離する方法について説明する。動画の中で時間的に静止している物体と動き続けている物体があった場合を考える。式 (3.2) で求めた動画行列 \mathbf{X} は、列方向が時間フレームに対応しているため、時間的に静止している物体は列方向に一定の値となって現れる。一方、動き続けている物体は画素の輝度値に時間的な変化を与えるため、列方向に変

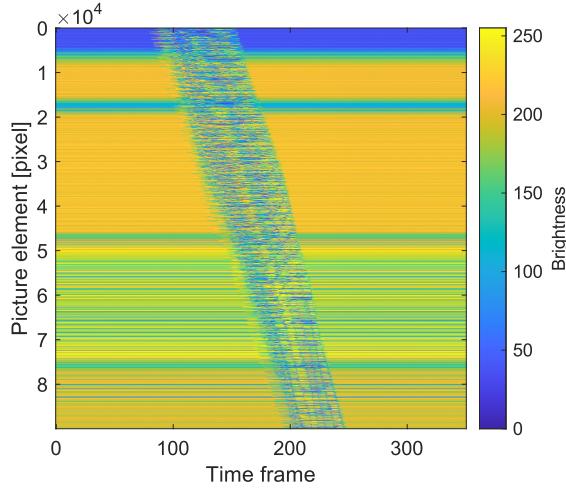


Fig. 3.2: Example of video matrix.



Fig. 3.3: Example of grayscale dynamic image signal.

化のある値となる。

この現象を実際に Figs. 3.2 及び 3.3 に例示する。Fig. 3.2 は、Fig. 3.3 のような時間的変化を持つ動画 \mathbf{X} を動画行列 \mathbf{X} に変換したものである。動画は動体が画面左から右に横切る内容であるが、 \mathbf{X} では動体が横切る時間帯の列に画素値の大きな変化が確認できる。一方、動体がない画素については列方向に一定の値を維持している。従って、動画行列 \mathbf{X} の各行の全要素 $x_{wh,1}, x_{wh,2}, \dots, x_{wh,T}$ に対してメディアンフィルタを独立に適用し、動体である外れ値を除去することができれば、背景画像のみが残った動画行列 $\mathbf{X}^{(bg)} \in \{0, 1, \dots, 255\}^{WH \times T}$ を抽出することができる。この処理は次式で表される。

$$\mathbf{X}^{(bg)} = [\mathbf{b} \ \mathbf{b} \ \dots \ \mathbf{b}] \quad (3.3)$$

$$\mathbf{b} = \begin{bmatrix} \text{median}(x_{1,1}, x_{1,2}, \dots, x_{1,T}) \\ \text{median}(x_{2,1}, x_{2,2}, \dots, x_{2,T}) \\ \vdots \\ \text{median}(x_{WH,1}, x_{WH,2}, \dots, x_{WH,T}) \end{bmatrix} \in \{0, 1, \dots, 255\}^{WH} \quad (3.4)$$

この処理を実際に Fig. 3.2 の動画行列に適用した結果の背景動画行列 $\mathbf{X}^{(bg)}$ を Fig. 3.4 に示す。この結果より、メディアンフィルタを用いて背景の成分を正しく推定できることが分かる。

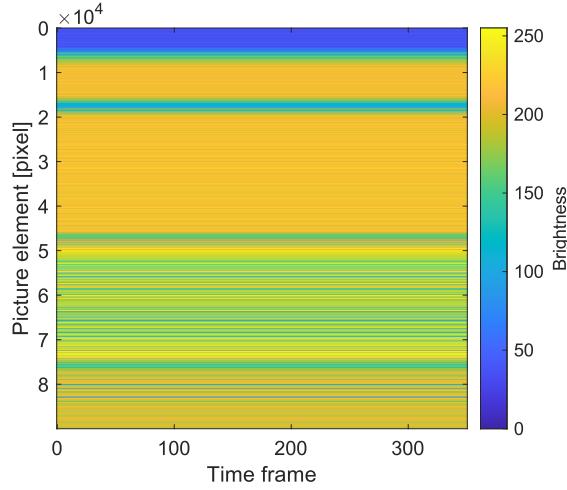


Fig. 3.4: Estimated background video matrix.

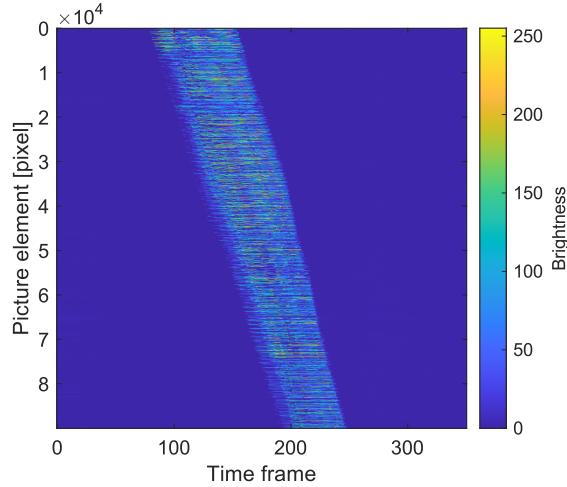


Fig. 3.5: Estimated dynamic video matrix.

3.4 バイナリマスクによる動体の分離

動画行列 \mathbf{X} から、3.3 節で推定した背景動画行列 $\mathbf{X}^{(bg)}$ を減算することで、動画中の動体に対応する画素のみが大きな値を持つ差分行列 $\mathbf{X}^{(diff)} \in \{0, 1, \dots, 255\}^{WH \times T}$ を得ることができる。

$$\mathbf{X}^{(diff)} = \text{abs}(\mathbf{X} - \mathbf{X}^{(bg)}) \quad (3.5)$$

ここで、 $\text{abs}(\cdot)$ は入力行列の各要素に対して絶対値をとった行列を返す関数である。式 (3.5) は各時刻における画像が、メディアンフィルタで抽出した背景画像 \mathbf{b} ($\mathbf{X}^{(bg)}$ の各列) とどの程度異なるかを表す。そのため、 $\mathbf{X}^{(diff)}$ の中で値の大きな要素は動体の画素を表し、値の小

さな要素は背景の画素を表している。

いま、ある動画行列 $\mathbf{A} \in \{0, 1, \dots, 255\}^{WH \times T}$ に対して、閾輝度値 λ ($0 \leq \lambda \leq 255$) に基づいて 2 値化する処理を次式で定義する。

$$\mathbf{A}' = \text{bin}_\lambda(\mathbf{A}) \in \{0, 1\}^{WH \times T} \quad (3.6)$$

このとき、2 値化された行列 \mathbf{A}' の各要素を $a'_{wh,t}$ とおくと、次式で与えられる。

$$a'_{wh,t} = \begin{cases} 1 & (\text{if } a_{wh,t} > \lambda) \\ 0 & (\text{otherwise}) \end{cases} \quad (3.7)$$

ここで、 $a_{wh,t}$ は元の行列 \mathbf{A} の各要素である。この 2 値化処理を用いて、差分行列 $\mathbf{X}^{(\text{diff})}$ を次式で 2 値化した行列を $\mathbf{M} \in \{0, 1\}^{WH \times T}$ とおく。

$$\mathbf{M} = \text{bin}_\lambda(\mathbf{X}^{(\text{diff})}) \quad (3.8)$$

差分行列 $\mathbf{X}^{(\text{diff})}$ は値の大きな要素が動体の画素、値の小さな要素は背景の画素を表すため、閾輝度値 λ を適切に設定すれば、この 2 値行列 \mathbf{M} は動画行列 \mathbf{X} から動体のみを抽出するバイナリマスク行列となる。従って、動画行列 \mathbf{X} から動体を分離して得られる動体動画行列 $\mathbf{X}^{(\text{mv})} \in \{0, 1, \dots, 255\}^{WH \times T}$ は次式のように推定される。

$$\mathbf{X}^{(\text{mv})} = \mathbf{M} \circ \mathbf{X} \quad (3.9)$$

ここで、 \circ は行列間の要素毎の積（アダマール積）を表す。

実際に Fig. 3.2 の動画行列に対して、上記を適用した結果の動体動画行列 $\mathbf{X}^{(\text{mv})}$ をそれぞれ Fig. 3.5 に示す。但し、閾輝度値の λ は 2 と設定した。さらに、分離された動体のある時間フレームの画像を Fig. 3.6 に示す。これらの結果より、差分行列を 2 値化したマスク行列による動体の推定が正しく動作することが分かる。

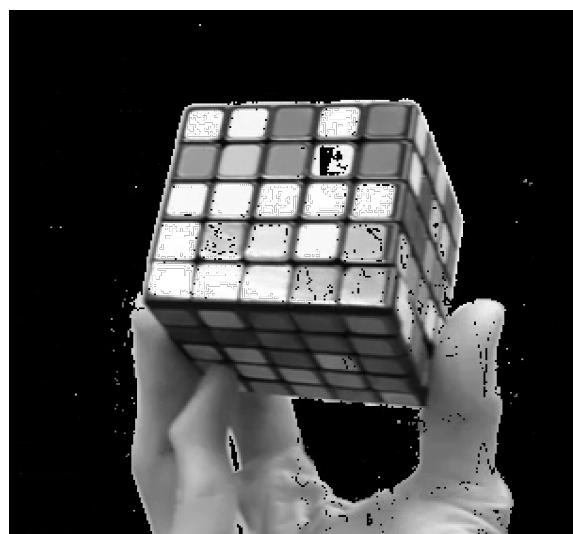


Fig. 3.6: Example of estimated dynamic image.

3.5 本章のまとめ

本章では、4章や6章で述べる手法の前処理として、入力動画の背景と動体を分離する手法について詳しく説明した。これらの前処理は、二重振り子の先端の位置を推定する精度を向上させるために必要となる。次章では、二重振り子の先端にLED等の光源が装着されている場合を想定し、本章で述べた手法で分離した二重振り子の可動部のみが含まれる動画の各時間フレームに対して、LEDの色情報を用いて先端位置を推定する提案手法について述べる。

第4章

色情報を用いた軌跡推定法

4.1 はじめに

本章では、動画撮影の対象となる二重振り子の先端に、青色や赤色等のLED等の光源が装着されており、撮影した動画の色情報を用いて先端の位置が推定できる状況における手法について述べる。まず4.2節では、本手法を提案する動機と処理の概要を述べる。本手法は光源と似た色が背景にある場合、これに対する頑健性が小さく、二重振り子先端部の位置の推定が困難になる。そこで、入力カラー動画で二重振り子の可動部以外の画素を全て0とする手法及び二重振り子先端部のピクセル位置の候補の抽出処理を4.3節で説明する。4.4節では、4.3節で得られたピクセル位置の候補からの1点の二重振り子先端のピクセル位置の決定について述べる。4.5節で本章についてまとめる。

4.2 動機と処理の概要

本研究は、二重振り子における初期値鋭敏性をSTEM教育の題材として扱うことを目的としている。初期値鋭敏性による二重振り子先端の運動(軌跡)の変化を可視化するには、二重振り子先端の軌跡を何らかの形で記録する必要がある。一般によく用いられる方法に1章で述べた二重振り子の先端にLEDを取り付けて発光させ、カメラなどの撮像機器で長時間露光させて撮影する方法があるが、これは可能な限り周囲を暗くする必要があり、STEM教育の現場では撮影環境を整えるのが難しい。明るい環境での二重振り子の先端及び軌跡の推定が可能であれば、そのような環境の制約から解放され、二重振り子における初期値鋭敏性をSTEM教育の題材として扱うことが容易になる。

これらの問題の解決を目的として、環境によらない二重振り子の先端の軌跡推定を試みる。本章で提案する手法では、二重振り子の先端に取り付けたLEDの色情報を用いる。もし、背景にLEDの発光色と同じ色の何かが存在すれば、本手法の適用は困難になる。そこで、本手法は3章の前処理で得られた入力動画の動体部分のバイナリマスクをカラー動画に適用することで、背景の画素を0とする。これは、同時に動体部分のみのカラー動画の作成を行い、背景

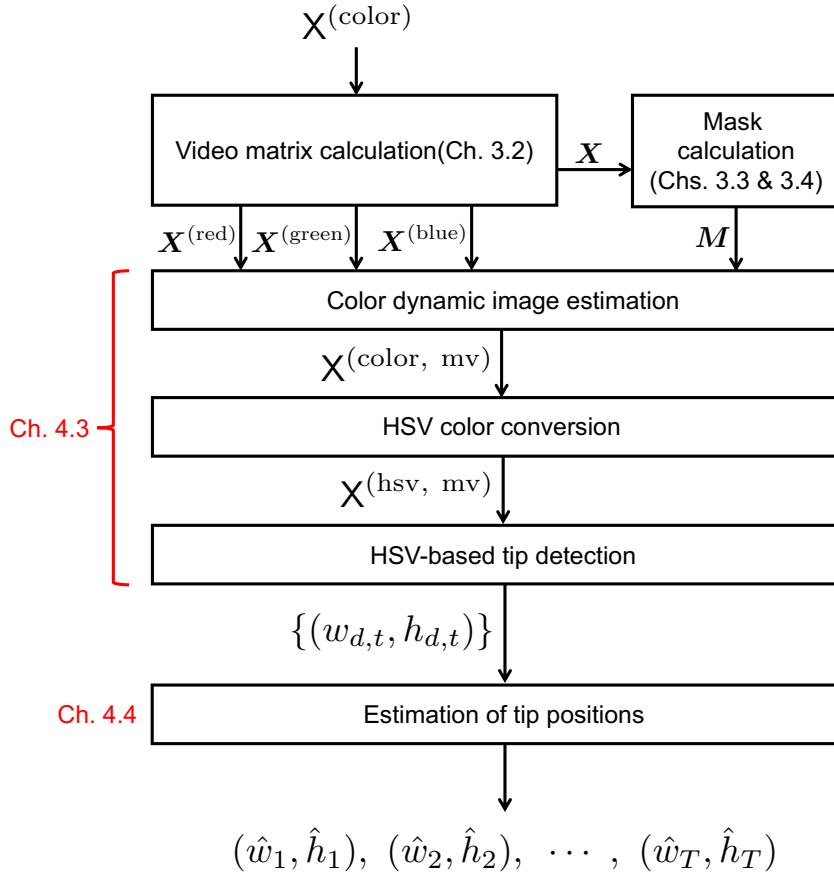


Fig. 4.1: Block diagram of trajectory estimation using color information.

の色に対する本手法の頑健性を高める処理である。

次に、本手法の処理の概要を箇条書きで示すとともに、Fig. 4.1 に示す。Fig. 4.1 は上から下に処理の流れを示したブロック図である。

1. カラー動画 $X^{(\text{color})}$ を入力する
2. カラー動画から動画行列 \mathbf{X} , $\mathbf{X}^{(\text{red})}$, $\mathbf{X}^{(\text{green})}$, $\mathbf{X}^{(\text{blue})}$ を作る
3. 動画行列 \mathbf{X} から動体動画行列を抽出するバイナリマスク行列 \mathbf{M} を推定する
4. $\mathbf{X}^{(\text{red})}$, $\mathbf{X}^{(\text{green})}$, $\mathbf{X}^{(\text{blue})}$ にバイナリマスク行列 \mathbf{M} を適用し、動体のみが分離されたカラー動画 $\mathbf{X}^{(\text{color, mv})}$ を出力する
5. 動体のみが分離されたカラー動画 $\mathbf{X}^{(\text{color,mv})}$ を入力として、各時間フレームの二重振り子の先端の座標の推定値 (\hat{w}_t, \hat{h}_t) を出力する

4.3 HSV 色空間上での画素分類

3.3 及び 3.4 節で述べた方法により、動体動画行列を抽出するバイナリマスク行列 \mathbf{M} を得る。その後、入力カラー動画 $X^{(\text{color})}$ の各カラーチャンネルの動画行列 $\mathbf{X}^{(\text{red})}$, $\mathbf{X}^{(\text{green})}$, 及び

$\mathbf{X}^{(\text{blue})}$ にバイナリマスク行列 \mathbf{M} を次式のように適用する。

$$\mathbf{X}^{(\text{red}, \text{ mv})} = \mathbf{M} \circ \mathbf{X}^{(\text{red})} \quad (4.1)$$

$$\mathbf{X}^{(\text{green}, \text{ mv})} = \mathbf{M} \circ \mathbf{X}^{(\text{green})} \quad (4.2)$$

$$\mathbf{X}^{(\text{blue}, \text{ mv})} = \mathbf{M} \circ \mathbf{X}^{(\text{blue})} \quad (4.3)$$

この各カラーチャンネルの動画行列を $\mathbf{X}^{(\text{color})}$ と同じサイズに再び整形することで、二重振り子の可動部以外の画素が全て 0 となったカラー動画信号 $\mathbf{X}^{(\text{color}, \text{ mv})} \in \{0, 1, \dots, 255\}^{W \times H \times 3 \times T}$ を得る。

次に、カラー動画信号 $\mathbf{X}^{(\text{color}, \text{ mv})}$ の各時間フレームの画像を $\mathbf{X}_t^{(\text{color}, \text{ mv})} \in \{0, 1, \dots, 255\}^{W \times H \times 3}$ とおくとき、 $\mathbf{X}_t^{(\text{color}, \text{ mv})}$ を hue, saturation, and value (HSV) 色空間の画像に変換する。HSV 色空間は色相、彩度、及び明度から成る色の表現系であり、閾値処理による色の分類が RGB 色空間よりも容易になる。変換後の各時間フレームの画像を $\mathbf{X}_t^{(\text{hsv}, \text{ mv})} \in \{0, 1, \dots, 255\}^{W \times H \times 3}$ と定義する（3 の次元が色相、彩度、及び明度から成る HSV 色空間である）。従って、二重振り子の先端に装着されている LED の色を検出するように色相、彩度、及び明度に対する閾値を定め、その色と判断された画素のピクセル位置を $\mathbf{X}_t^{(\text{hsv}, \text{ mv})} \in \{0, 1, \dots, 255\}^{W \times H \times 3}$ から抽出する。その結果得られる、時間フレーム t における当該色の画素のピクセル位置の集合を $\{(w_{d,t}, h_{d,t})\}_d$ と定義する。ここで、 $d \in \mathbb{N}$ は抽出されたピクセルのインデックスであり、その数は時間フレーム t に依存する。例えば、時間フレーム t において 1 画素のみが当該色と判定された場合は、ピクセル位置は $d = 1$ のみ（すなわち、 $\{(w_{1,t}, h_{1,t})\}$ という 1 画素の単元集合）となり、時間フレーム t' において 2 個の画素が当該色と判定された場合は、ピクセル位置は $d = 1, 2$ （すなわち、 $\{(w_{1,t'}, h_{1,t'}), (w_{2,t'}, h_{2,t'})\}$ の 2 画素から成る集合）となる。なお、ある時刻フレームにおいて、当該色の画素が 1 個も抽出されない可能性もある。その場合は $\{(w_{d,t}, h_{d,t})\}_d = \emptyset$ （空集合）となる。

以上の手続きにより、二重振り子の先端のピクセル位置の候補として、時間フレーム毎の LED のピクセル位置の集合 $\{(w_{d,t}, h_{d,t})\}_d$ が得られる。この処理における HSV 色空間上の分類の閾値は、実際に LED の色を確認してある程度手動で設定する必要がある。しかしながら、次節で述べる処理により、複数の候補から可能な限り外れ値に対して頑健に二重振り子の先端のピクセル位置を算出できるため、当該色の判定範囲は多少広くしても大きな問題とはならないことが予想される。

4.4 振り子の先端位置の推定

前節で得られた時間フレーム毎の LED のピクセル位置の集合 $\{(w_{d,t}, h_{d,t})\}_d$ を用いて、二重振り子の先端のピクセル位置を決定する。ほとんどの時間フレームでは、複数の LED のピクセル位置が得られるため、何らかの方法でただ 1 点のピクセル位置に定める必要がある。

本論文では、推定された LED のピクセル位置がほとんど近傍に集中することを実験的に確認したことから、次式のように複数の候補点の平均値を二重振り子の先端のピクセル位置と定

める。

$$\hat{w}_t = \frac{1}{|\{(w_{d,t}, h_{d,t})\}_d|} \sum_d w_{d,t} \quad (4.4)$$

$$\hat{h}_t = \frac{1}{|\{(w_{d,t}, h_{d,t})\}_d|} \sum_d h_{d,t} \quad (4.5)$$

このようにして、時間フレーム t における二重振り子の先端のピクセル位置が (\hat{w}_t, \hat{h}_t) として推定される。なお、式 (4.4) 及び (4.5) は $|\{(w_{d,t}, h_{d,t})\}_d| > 0$ でのみ（つまり候補点が 1 個以上存在する時間フレームについてのみ）計算する。候補点が 1 個も存在しなかった時間フレームは、その時間フレームの二重振り子の先端位置は無かったものとして扱う。また、HSV 色空間での当該色の判定範囲を広くした場合等で、候補点（LED のピクセル位置）に外れ値が含まれることが懸念される場合は、より頑健な推定法として式 (4.4) 及び (4.5) を平均値ではなく中央値に変更するなどの工夫も考えられる。

以上の手続きにより、時刻フレーム毎の二重振り子の先端のピクセル位置が集合 $\{(\hat{w}_t, \hat{h}_t)\}_t$ として推定される。これらのピクセル位置から滑らかな軌跡を描く処理は、2.5 節で説明した 2 次元空間上の時系列信号に対する 3 次スプライン補間を適用することで実現できる。

4.5 本章のまとめ

本章では、色情報を用いた軌跡追跡法について説明した。本手法は、二重振り子先端部に何らかの光源が装着されている場合、その色情報を用いて先端の位置を推定するといったものである。光源と同様の色が背景にあった場合、入力カラー動画の各カラーチャンネルの動画行列に動体動画のマスクを適用することで頑健性を保障する。また、HSV 色空間での閾値を用いた二重振り子先端部のピクセル位置候補の抽出の後、それらを平均することで 1 点のピクセル位置を決定する。次章では、本手法を用いた二重振り子の軌跡推定実験について述べる。

第 5 章

色情報を用いた軌跡推定法の実験

5.1 はじめに

本章では、4章で提案した手法の性能を評価するために行った二重振り子の軌跡推定実験について示す。本手法を用いた軌跡の推定と、時間フレーム毎に先端位置を手動でマーキングした正解データとの比較を行い、客観的な評価尺度で性能を確認する。本実験の条件を 5.2 節、結果を 5.3 節に示す。最後に、5.4 節で本章をまとめるとする。

5.2 実験条件

本実験で使用した二重振り子を Fig. 5.1 に示す。この二重振り子はバッテリーと LED が内蔵されているパーツを先端に装着できるようになっている。LED は青色のものを使用した。4.3 節で述べた通り、この青色 LED が映った画素を HSV 色空間上で判定する。本実験では、あらかじめ LED の色を調べて、色相範囲を [106, 192]、彩度範囲を [64, 255]、明度範囲を [128, 255] に限定して判定した。なお、いずれのパラメータも取りうる値の範囲は [0, 255] である。本実験で用いた客観評価尺度は、二乗平均平方誤差 (root mean squared error: RMSE)、平均絶対誤差 (mean absolute error: MAE)、平均絶対率誤差 (mean absolute percentage error: MAPE)、及びピアソンの積率相関 (Pearson Correlation Coefficient: PCC) の 4 つとした。いま、正解の二重振り子の先端座標を $(\check{w}_t, \check{h}_t)$ と定義するとき、推定値 (\hat{w}_t, \hat{h}_t) と正

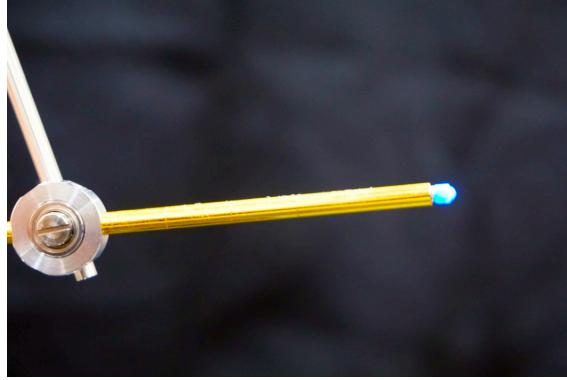


Fig. 5.1: Blue LED attached on tip of double pendulum.

解値 $(\check{w}_t, \check{h}_t)$ の間の RMSE, MAE, MAPE, 及び PCC の計算式は次式となる.

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{w}_t - \check{w}_t)^2 + \frac{1}{T} \sum_{t=1}^T (\hat{h}_t - \check{h}_t)^2} \quad (5.1)$$

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T |\hat{w}_t - \check{w}_t| + \frac{1}{T} \sum_{t=1}^T |\hat{h}_t - \check{h}_t| \quad (5.2)$$

$$\text{MAPE} = \frac{100}{T} \sum_{t=1}^T \left| \frac{\hat{w}_t - \check{w}_t}{\check{w}_t} \right| + \frac{100}{T} \sum_{t=1}^T \left| \frac{\hat{h}_t - \check{h}_t}{\check{h}_t} \right| [\%] \quad (5.3)$$

$$\text{PCC} = \frac{\sum_{t=1}^T \sqrt{(\hat{w}_t - \bar{\hat{w}}_t)^2 + (\hat{h}_t - \bar{\hat{h}}_t)^2} \sqrt{(\check{w}_t - \bar{\check{w}}_t)^2 + (\check{h}_t - \bar{\check{h}}_t)^2}}{\sqrt{\sum_{t=1}^T \{(\hat{w}_t - \bar{\hat{w}}_t)^2 + (\hat{h}_t - \bar{\hat{h}}_t)^2\}} \sqrt{\sum_{t=1}^T \{(\check{w}_t - \bar{\check{w}}_t)^2 + (\check{h}_t - \bar{\check{h}}_t)^2\}}} \quad (5.4)$$

ここで, $\bar{\hat{w}}_t$ 及び $\bar{\hat{h}}_t$ は, \hat{w}_t 及び \hat{h}_t の平均値である. なお, 本実験における正解値の座標は, 時間フレーム毎に二重振り子の先端位置を手動でマーキングして正解ラベルを作成した.

実験で用いた二重振り子の撮影動画の一部のキャプチャ画像を Fig. 5.2 に示す. この動画は, 画面サイズが $W = 1162$ pixels, $H = 720$ pixels であり, 時間は 1.8333 s であった. 動画は 240 fps のフレームレート及び 8 倍スローモーションで撮影されたため, 時間フレーム数は $T = 55$ frames であった. 二重振り子は暗幕の前に配置した状態で撮影した. 動体の分離に用いる閾輝度値 λ は 30 と設定した.

5.3 実験結果

4 章で述べた処理を適用した動画の初期フレームのキャプチャ画像を Fig. 5.3 にそれぞれ示す. Fig. 5.3(a) に示したカラー動画にバイナリマスク行列を適用すると, Fig. 5.3(b) のような動体のみが分離されたカラー動画が得られた. さらに, 得られた動画で二重振り子先端のピ

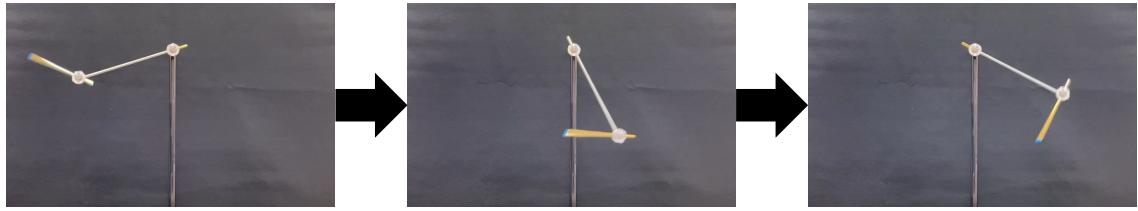


Fig. 5.2: Recorded video in this experiment.

クセル位置の候補のみを分離すると、Fig. 5.3(c) のような先端部のみが分離されたカラー動画が得られた。

次に、Fig. 5.4 に、正解ラベル、本手法による推定ラベル、及びスプライン曲線を示す。推定ラベル及び正解ラベルは近い値をとっており、描かれた曲線により、本手法が実際の二重振り子先端の軌跡を正確に推定できることが示された。

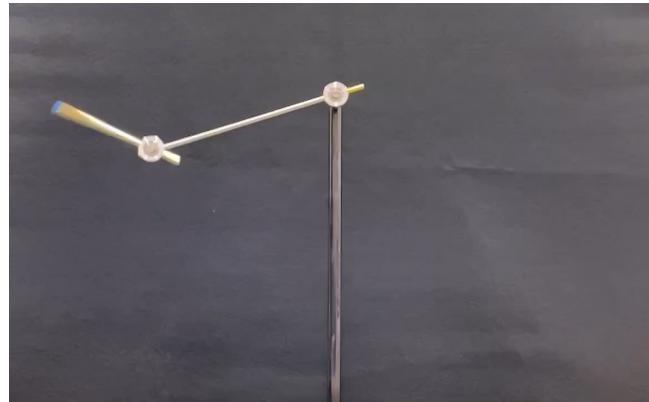
正解ラベルと推定ラベルの類似度を 4 つの方法で評価した結果を Table 5.1 に示す。RMSE, MAE, 及び MAPE の値が十分小さく、各時間フレームで二重振り子先端部の位置を正確に推定できることができることが確認された。また、PCC がほとんど 1 であり、本手法が、実際の軌跡とほとんど同じ概形の軌跡を推定する性能を持つことが数値的に評価できた。

Table 5.1: Scores of proposed trajectory estimation using color information

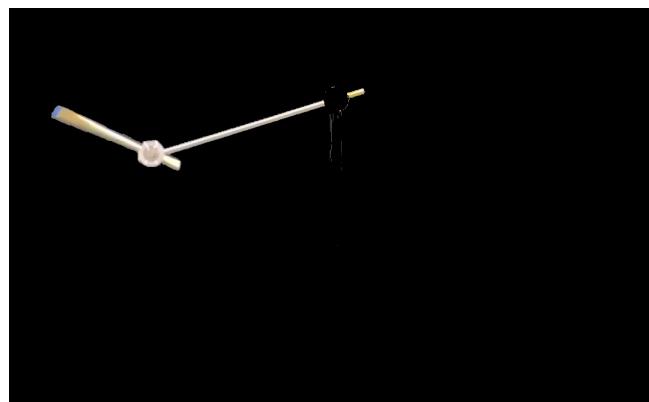
Subjective criterion	RMSE [pixel]	MAE [pixel]	MAPE [%]	PCC
Score	3.8415	1.2677	4.6512	0.9999

5.4 本章のまとめ

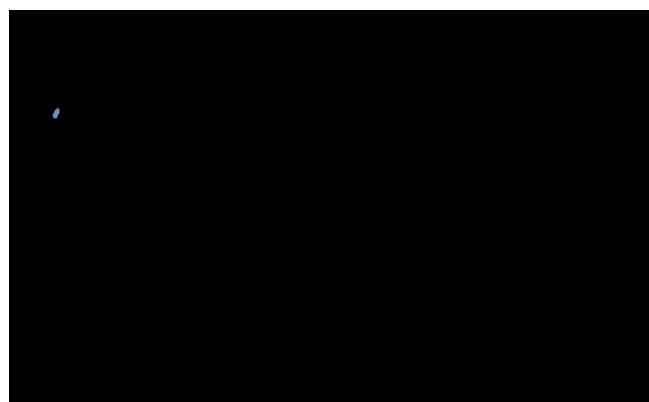
本章では、二重振り子先端部の LED の色情報を用いた手法による軌跡推定実験についての解説を行った。まず、正解ラベル、推定ラベル、及びスプライン曲線を示した。続いて、正解ラベルと推定ラベルの類似度を 4 つの尺度で評価することで、実際の軌跡とほとんど同等のものを推定できる本手法の性能の高さを示した。よって、本手法を用いることで、二重振り子の初期値鋭敏性を STEM 教育の題材とすることが可能となる。しかしながら、先端部の LED の発光色は扱う二重振り子によって異なるため、本手法では HSV 色空間上での範囲を限定するパラメータを適宜変更する必要がある。また、LED が装着されていない二重振り子も多く、本手法は二重振り子の初期値鋭敏性を STEM 教育の題材にすることを可能にする一方、教材準備の際の負担が考えられる。この問題を解決し、二重振り子を扱った STEM 教育の容易な実施を可能にするため、次章では、色情報を必要とせず、任意の二重振り子に適用できる手法を提案する。



(a) Initial frame of video used in experiment



(b) Initial frame of estimated moving object



(c) Initial frame of estimated tip

Fig. 5.3: Observed and estimated video frames used in this experiment.

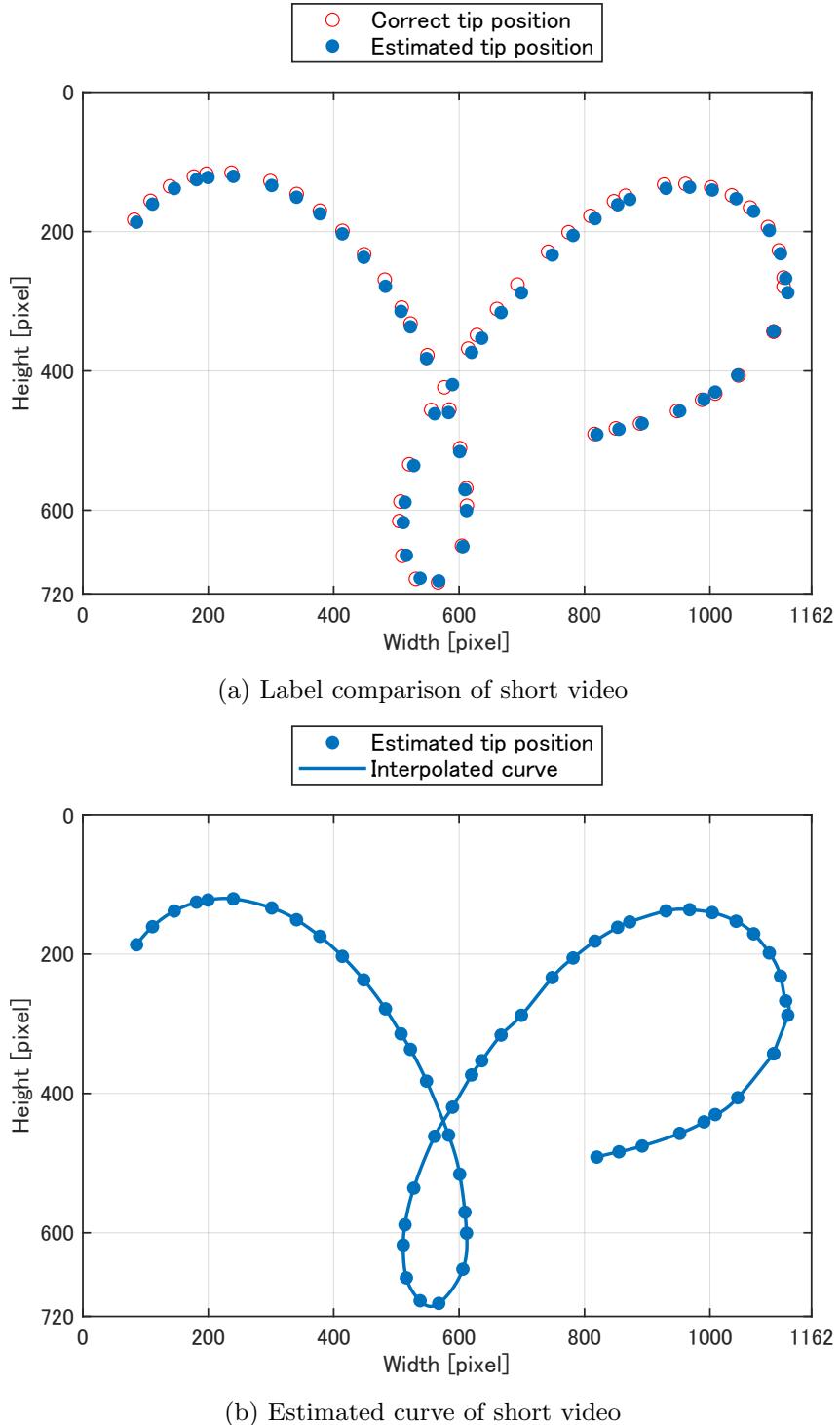


Fig. 5.4: Double pendulum trajectory estimated by proposed method using color information.

第 6 章

概形情報を用いた軌跡推定法

6.1 はじめに

本章では、二重振り子の先端に LED 等の発光体が装着されていない一般的な場合における軌跡推定法について述べる。本章で考える状況では色情報を使うことができないため、代わりに二重振り子の概形の位置推定を用いた手法を提案する。まず、6.2 節では本手法を提案する動機を述べる。本手法は、二重振り子の第 1 の腕の回転軸（以後、単に回転軸と呼ぶ）及び第 2 の腕の回転軸（以後、連結部と呼ぶ）の位置情報を推定し、先端部分の位置の推定を行う。これを説明するために、6.3 節では、振り子の回転軸及び連結部の座標探索について、述べる。6.4 節では、連結部の位置情報を用いた先端部分の位置の推定について説明する。最後に 6.5 章で本章をまとめることとする。

6.2 動機と処理の概要

4 章に記載した、色情報を用いた手法は、二重振り子の先端に LED 等の光源が装着されていない場合に適用することが難しい。より一般に、色情報を用いない二重振り子の先端及び軌跡の推定が可能であれば、そのような機材の制約から解放され、本論の提案手法や目的の適用範囲の拡大が可能となる。また、LED 等の制約に捕らわれない方法で二重振り子以外の物理現象の可視化にも適用できる方法が確立できれば、二重振り子以外の教材を用いた STEM 教育等に役立てることができる。

これらの問題の解決を目的として、色情報を用いることができないという制約付きの二重振り子の先端の軌跡推定を試みる。本章で提案する手法では、各時間フレームで、はじめから振り子先端部の位置を推定するのではなく、回転軸及び連結部の位置を推定し、それを基に二重振り子の先端部の位置を推定するという流れである。これは二重振り子以外の物理現象にただちに応用できる方法論ではないが、観測対象の物理現象に特徴的な機構や形状などがあれば、類似した手法での物体の追跡や可視化のヒントになりえるものである。

次に、本手法の処理の概要を箇条書きで示すとともに、Fig. 6.1 に示す。Fig. 6.1 は上から

下に処理の流れを示したブロック図である。

1. カラー動画 $\mathbf{X}^{(\text{color})}$ を入力する
2. グレイスケール動画 \mathbf{X} を作る
3. グレイスケール動画から動画行列 \mathbf{X} を作る
4. 動画行列 \mathbf{X} から動体動画行列を抽出するバイナリマスク行列 \mathbf{M} を推定する
5. \mathbf{X} にバイナリマスク行列 \mathbf{M} を適用し、動体のみが分離されたグレイスケール動画 $\mathbf{X}^{(\text{mv})}$ を出力する
6. 背景のみが分離されたグレイスケール動画 $\mathbf{X}^{(\text{bg})}$ を入力として、二重振り子の回転軸及び連結部の座標の推定値 $\mathbf{r}_0 = [w_0, h_0]^T$, $\mathbf{r}_t^{(1)} = [w_t^{(1)}, h_t^{(1)}]^T$ を出力する
7. \mathbf{r}_0 , \mathbf{r}_1 及び動体のみが分離されたグレイスケール動画 $\mathbf{X}^{(\text{mv})}$ を入力として、各時間フレームの二重振り子先端の座標の推定値 $\mathbf{r}_t^{(2)} = [w_t^{(2)}, h_t^{(2)}]^T$ を出力する

6.3 回転軸と連結部の位置推定

本節では、回転軸及び連結部の位置推定について順に述べる。本論文で対象としている二重振り子は、回転軸と連結部がいずれも円形で同一の大きさである。このことに着目し、3章で述べた方法で背景と動体を分離し、背景のみの動画の各時刻から回転軸、動体のみの動画の各時刻から連結部の位置を推定する。

いま、式(3.3)で推定される（グレイスケールの）背景動画行列 $\mathbf{X}^{(\text{bg})}$ を整形し、動画信号に戻した信号を $\mathbf{X}^{(\text{bg})} \in \{0, 1, \dots, 255\}^{W \times H \times T}$ と定義する。また、その動画の時間フレーム t の画像を $\mathbf{X}_t^{(\text{bg})}$ とおく。Fig. 6.2 は、 $\mathbf{X}_t^{(\text{bg})}$ を表している。回転軸は時間的にほとんど変動しないため、分離された背景である $\mathbf{X}_t^{(\text{bg})}$ には Fig. 6.2 のように回転軸が残っている。なお、図の簡単化のため、Fig. 6.2 では回転軸は円形ではなく十字形としている。この回転軸の位置を推定するために、Fig. 6.2 に示す $\mathbf{T} \in \{0, 255\}^{W_{\text{template}} \times H_{\text{template}}}$ のようなテンプレート画像を用意する。ここで、 W_{template} 及び H_{template} はテンプレート画像の横幅と高さである。このテンプレート画像は、推定したい物体の位置を模した概形でかつ画素値が 0（黒）又は 255（白）のみから成るものであり、事前に作成する必要がある。次に、 $\mathbf{X}_t^{(\text{bg})}$ 中の (w, h) の画素を中心とする部分行列 $\mathbf{X}_{w,h,t} \in \{0, 1, \dots, 255\}^{W_{\text{template}} \times H_{\text{template}}}$ を定義する。そして、テンプレート画像を Fig. 6.2 のようにスライドさせながら $\mathbf{X}_t^{(\text{bg})}$ との内積値を求めていく。この処理は、 \mathbf{T} と $\mathbf{X}_t^{(\text{bg})}$ の 2 次元畳み込みと同様である。 \mathbf{T} と $\mathbf{X}_{w,h,t}$ の内積値の計算は次式で表せる。

$$\mathcal{I}(w, h) = \|\mathbf{T} \circ \mathbf{X}_{w,h,t}\|_1 \quad (6.1)$$

ここで、 $\|\cdot\|_1$ は入力された行列の L_1 ノルム（各要素の絶対値の総和）を示す。この内積値 $\mathcal{I}(w, h)$ が最大となる画素 (w, h) が、テンプレート画像と最も類似する形状の画素を表すため、テンプレート画像に回転軸の概形を与えれば、 $\mathcal{I}(w, h)$ が最大となる画素 (w, h) が回転軸

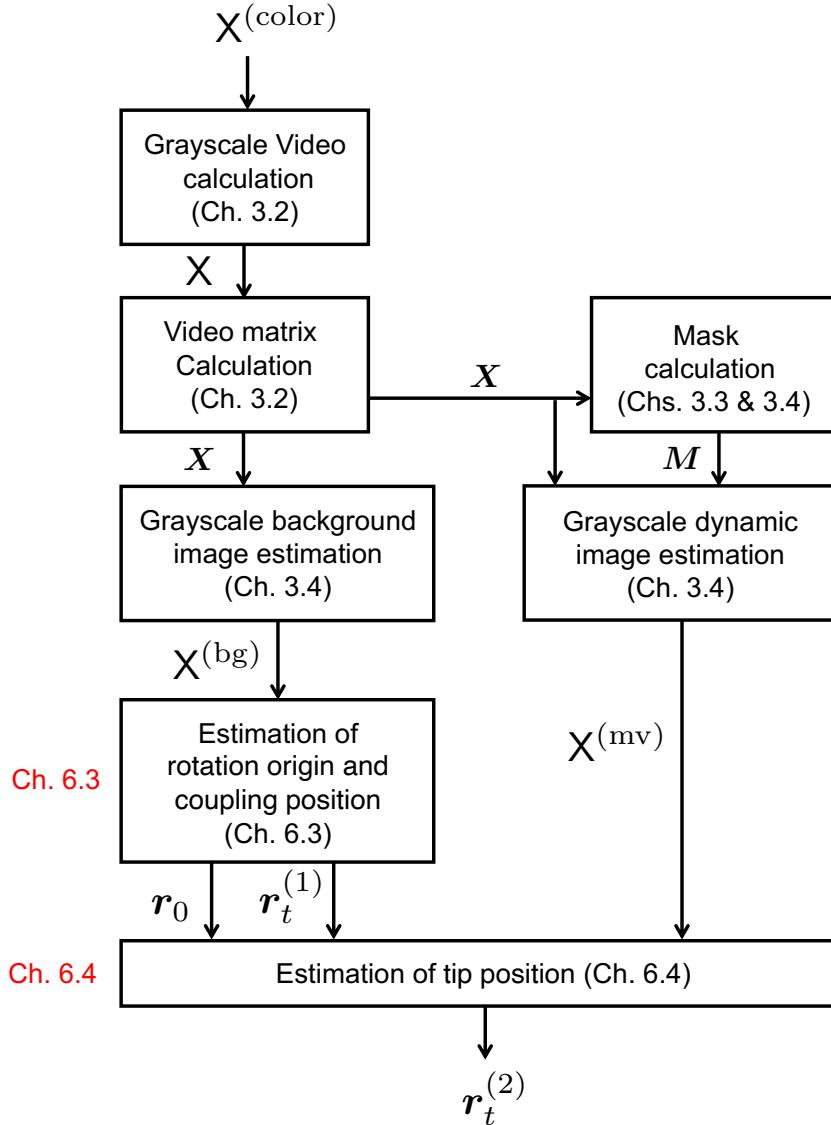


Fig. 6.1: Block diagram of trajectory estimation using outline information.

の推定位置として得られる。この処理は次式となる。

$$(w_0, h_0) = \operatorname{argmax}_{w,h} \mathcal{I}(w, h) \quad (6.2)$$

なお、式 (6.2) の右辺が単元集合にならない場合 ($\mathcal{I}(w, h)$ が最大となる画素が複数ある場合) は、 (w_0, h_0) を適當な 1 つの元として取得するものとする。このように推定された回転軸の位置を $r_0 = [w_0, h_0]^T$ と表す。なお、 r_0 は $t = 0$ の時間フレーム $X_0^{(bg)}$ から推定し、以後の $t = 1, 2, \dots, T$ では共通とする (t に依らないものとする)。

連結部の位置の推定についても、前述の回転軸の位置の推定と同様の方法で実現できる。すなわち、式 (3.3) で推定される（グレイスケールの）動体動画行列 $X^{(mv)}$ を整形し、動画信号に戻した信号を $X^{(mv)} \in \{0, 1, \dots, 255\}^{W \times H \times T}$ 、その動画の時間フレーム t の画像を

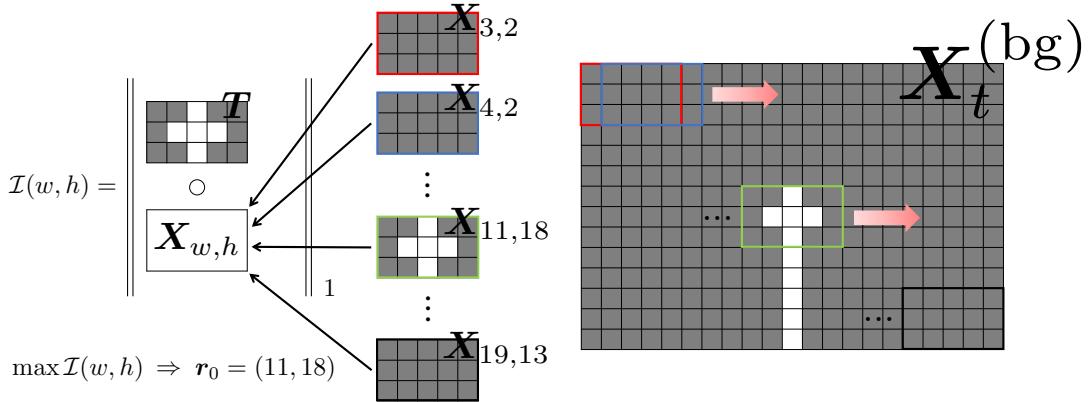


Fig. 6.2: Calculation of inner product values between template image and part of separated background image.

$\mathbf{X}_t^{(\text{mv})}$ として $\mathbf{X}_t^{(\text{bg})}$ と同様に構成すれば、あとは同じテンプレート画像 \mathbf{T} と $\mathbf{X}_t^{(\text{mv})}$ の間で式 (6.1) を計算し、これが最大となる画素を求められる。このようにして推定した連結部の位置を $\mathbf{r}_t^{(1)} = [w_t^{(1)}, h_t^{(1)}]^T$ と表す。

6.4 先端の位置推定

本節では、先端部の位置推定について述べる。二重振り子先端部の運動は複雑である。しかしながら、全体の運動を分解して考えると、先端部は回転軸に対して単振り子運動をする連結部を中心とした回転運動をするものと考えることができる。このことに着目し、6.3 節で推定した連結部を中心とした円周上にある先端部の位置を推定する。

Fig. 6.3 は、 $\mathbf{X}_t^{(\text{mv})}$ を表している。2つの腕は時間的に変化するため、分離された動体である $\mathbf{X}_t^{(\text{mv})}$ には Fig. 6.3 のように連結部が残っている。この連結部を中心とした円周上で先端部の位置を推定するために、次式に示す半径を用いる。

$$R_{\text{short}} = \begin{cases} \frac{l_1}{3} & (\text{if } t = 1) \\ l_2 - d & (\text{otherwise}) \end{cases} \quad (6.3)$$

$$R_{\text{long}} = \begin{cases} \frac{2l_1}{3} & (\text{if } t = 1) \\ l_2 + d & (\text{otherwise}) \end{cases} \quad (6.4)$$

$$l_1 = \|\mathbf{r}_0 - \mathbf{r}_1^{(1)}\|_2 \quad (6.5)$$

ここで、 l_1 は \mathbf{r}_0 及び $\mathbf{r}_1^{(1)}$ の L_2 ノルム（ユークリッドノルム）で求められる。 R_{short} 及び R_{long} は大きさが異なる 2 つの半径であり、 d は半径の大きさを調整するパラメータである。また、後に示す処理のため、 R_{short} は l_2 より小さく、 R_{long} は、 l_2 より大きい値を取る必要がある。

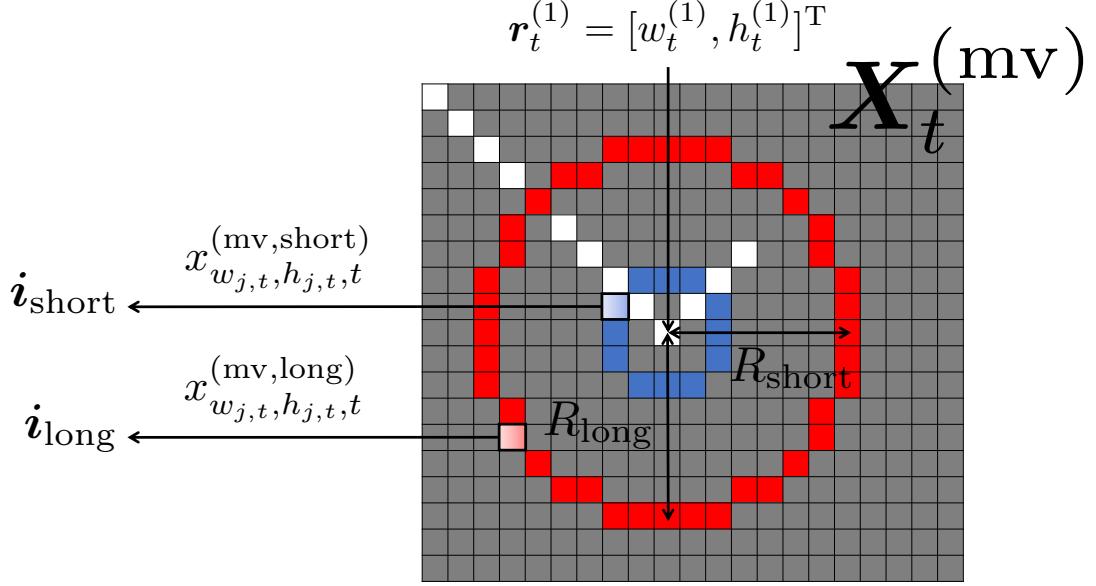


Fig. 6.3: Calculation of phase of second pendulum tip.

ある。そして、Fig. 6.3 のように、連結部を中心として R_{short} 及び R_{long} の円周上の対応する $\mathbf{X}_t^{(\text{mv})}$ の画素の輝度値を集めていく。連結部を中心とした R_{short} 及び R_{long} の円周上の対応する $\mathbf{X}_t^{(\text{mv})}$ の画素の座標は次式で表せる。

$$\begin{cases} w_{j,t}^{(\text{short})} = w_t^{(1)} + R_{\text{short}} \cos \omega_j \\ h_{j,t}^{(\text{short})} = h_t^{(1)} + R_{\text{short}} \sin \omega_j \end{cases} \quad (6.6)$$

$$\begin{cases} w_{j,t}^{(\text{long})} = w_t^{(1)} + R_{\text{long}} \cos \omega_j \\ h_{j,t}^{(\text{long})} = h_t^{(1)} + R_{\text{long}} \sin \omega_j \end{cases} \quad (6.7)$$

ここで、 ω_j は、連結部を中心とした円周上の各画素の位相ベクトル $\boldsymbol{\omega} = [-180, -179, \dots, 179]^T$ の要素であり、 $j = 1, 2, \dots, 360$ はそのインデクスである。また、 $(w_{j,t}^{(\text{short})}, h_{j,t}^{(\text{short})})$ は、 R_{short} 及び ω_i での時間フレーム上の座標であり、 $(w_{j,t}^{(\text{long})}, h_{j,t}^{(\text{long})})$ も同様に定義される。 $\mathbf{X}_t^{(\text{mv})}$ の式 (6.6) 及び (6.7) の画素の輝度値は次式で表される。

$$\mathbf{i}^{(\text{short})} = [x_{w_1,t,h_1,t,t}^{(\text{mv,short})}, x_{w_2,t,h_2,t,t}^{(\text{mv,short})}, \dots, x_{w_{360},t,h_{360},t,t}^{(\text{mv,short})}] \quad (6.8)$$

$$\mathbf{i}^{(\text{long})} = [x_{w_1,t,h_1,t,t}^{(\text{mv,long})}, x_{w_2,t,h_2,t,t}^{(\text{mv,long})}, \dots, x_{w_{360},t,h_{360},t,t}^{(\text{mv,long})}] \quad (6.9)$$

ここで、 $x_{w_j,t,h_j,t,t}^{(\text{mv,short})}$ は $\mathbf{X}_t^{(\text{mv})}$ の画素 $w_{j,t}^{(\text{short})}, h_{j,t}^{(\text{short})}$ における輝度値であり、 $x_{w_j,t,h_j,t,t}^{(\text{mv,long})}$ も同様に定義される。すなわち、 $\mathbf{i}^{(\text{short})} \in \{0, 1, \dots, 255\}^{360}$ 及び $\mathbf{i}^{(\text{long})} \in \{0, 1, \dots, 255\}^{360}$ は、 $\mathbf{X}_t^{(\text{mv})}$ の半径 R_{short} 及び R_{long} の円周上の画素の輝度値である。連結部を基準とすると、Fig. 6.3 のように、 $\mathbf{i}^{(\text{short})}$ は輝度値が大きい画素の位相が 2 つあるが、 $\mathbf{i}^{(\text{long})}$ は 1 つのみである。5 章で用いた動画の初期フレームの $\mathbf{i}^{(\text{short})}$ 及び $\mathbf{i}^{(\text{long})}$ を可視化したものを Fig. 6.4 に示す。Fig. 6.4 より、次式のように $\mathbf{i}^{(\text{short})}$ 及び $\mathbf{i}^{(\text{long})}$ の差分をとり、最大輝度値が残った画

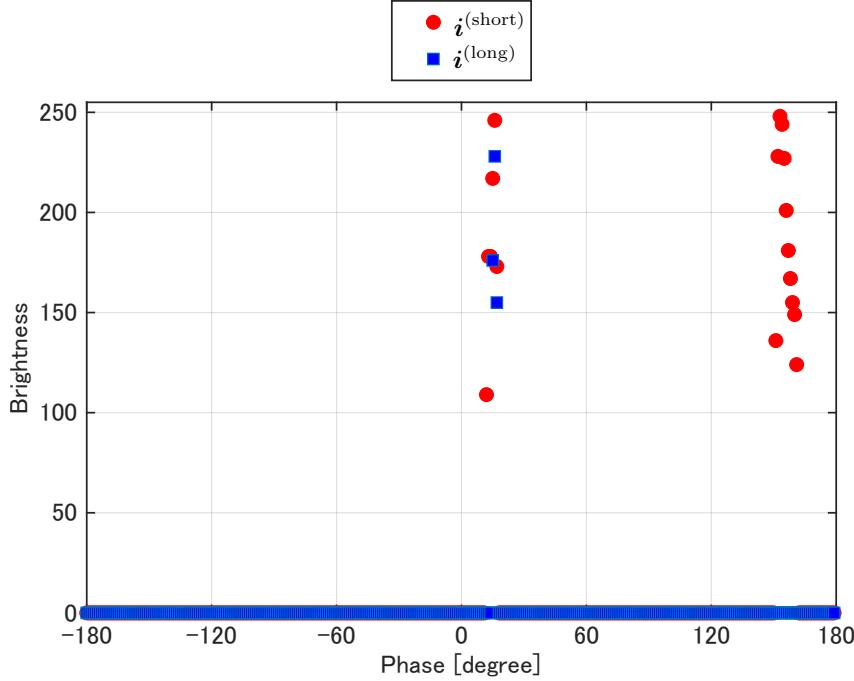


Fig. 6.4: Brightness values of two circumferential pixels of short and long radii.

素の位相が先端方向だと考えられる。

$$\mathbf{i}^{(\text{diff})} = \mathbf{i}^{(\text{short})} - \mathbf{i}^{(\text{long})} \quad (6.10)$$

ここで, $\mathbf{i}^{(\text{diff})} \in \{-255, -254, \dots, 255\}^{360}$ は各半径の円周上の画素の輝度値の差分をとったものである。

式 (6.10) のように差分を取るだけでは, 先端方向以外での輝度値が原因で正確な推定に失敗する可能性がある。そこで, $\mathbf{i}^{(\text{diff})}$ に式 (2.17) で述べた移動合計値フィルタを適用する。

$$\hat{\mathbf{i}}^{(\text{diff})} = \text{movsum}_{L_{\text{filter}}} (\mathbf{i}^{(\text{diff})}) \quad (6.11)$$

さらに, 連結部を基準とした, 前時間フレームの先端部と $\hat{\mathbf{i}}_{\text{diff}}$ の要素との位相差を外れ値除去のためのペナルティとして加える。

$$\check{\mathbf{i}}^{(\text{diff})} = \hat{\mathbf{i}}^{(\text{diff})} - a(\omega - \phi_{t-1}) \quad (6.12)$$

ここで, a は位相差の影響の大きさを決めるパラメータであり, $\phi_{t-1} = [\phi_{t-1}, \phi_{t-1}, \dots, \phi_{t-1}]^T$ は $t-1$ の先端方向の位相 ϕ_{t-1} を全要素にもつベクトルである。Fig. 6.4 の $\mathbf{i}_{\text{short}}$ 及び \mathbf{i}_{long} の差分を取り, 上記で述べた処理を行った $\check{\mathbf{i}}^{(\text{diff})}$ を可視化したものを Fig. 6.5 に示す。 $\mathbf{i}^{(\text{diff})}$ に移動合計値フィルタの適用及び位相差のペナルティの付与を行うことで, Fig. 6.5 のように最大輝度値をもつと仮定した先端の画素の位相の推定が容易になる。 $\check{\mathbf{i}}^{(\text{diff})}$ で最大輝度値をもつ要素を先端の画素と仮定し, この画素の位相を ϕ とすると, 先端部の位置 $\mathbf{r}_t^{(2)}$ は次式のよ

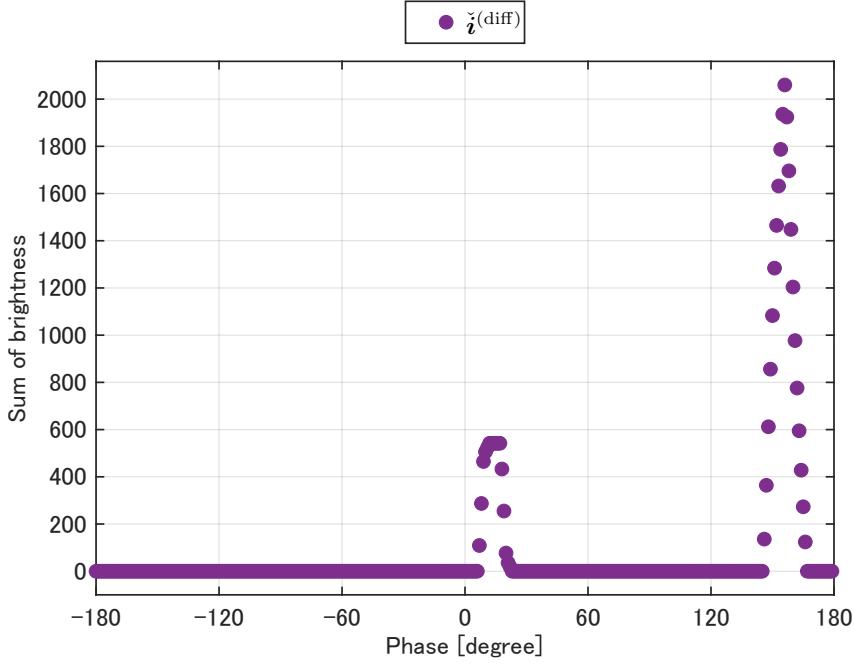


Fig. 6.5: Brightness values of pixels in each phase direction after applying moving sum filter.

うに求められる。

$$\mathbf{r}_t^{(2)} = \begin{bmatrix} w_t^{(\text{tip})} \\ h_t^{(\text{tip})} \end{bmatrix} \quad (6.13)$$

$$= \mathbf{r}_t^{(1)} + l_2 \begin{bmatrix} \cos \phi \\ \sin \phi \end{bmatrix} \quad (6.14)$$

よって、一連の処理を全時間フレームについて行い、推定される二重振り子先端の位置は次式のようになる。

$$\mathbf{w}^{(\text{tip})} = \begin{bmatrix} w_1^{(\text{tip})} \\ w_2^{(\text{tip})} \\ \vdots \\ w_T^{(\text{tip})} \end{bmatrix} \quad (6.15)$$

$$\mathbf{h}^{(\text{tip})} = \begin{bmatrix} h_1^{(\text{tip})} \\ h_2^{(\text{tip})} \\ \vdots \\ h_T^{(\text{tip})} \end{bmatrix} \quad (6.16)$$

上記で外れ値を除去するための処理をいくつか記載したが、それでも尚、先端方向以外の画素が最大輝度値をもつ誤差が発生する。そこで、式のように $\mathbf{w}^{(\text{tip})}$ 及び $\mathbf{h}^{(\text{tip})}$ に式 (2.15) で述

べた移動メディアンフィルタを適用することで、外れ値を除去をする。

$$\hat{\mathbf{w}}^{(\text{tip})} = \text{movmedian}_{L_{\text{filter}}}(\mathbf{w}^{(\text{tip})}) \quad (6.17)$$

$$\hat{\mathbf{h}}^{(\text{tip})} = \text{movmedian}_{L_{\text{filter}}}(\mathbf{h}^{(\text{tip})}) \quad (6.18)$$

最終的な推定位置を $\check{\mathbf{w}}^{(\text{tip})}, \check{\mathbf{h}}^{(\text{tip})}$ とおくと、これは式 (6.17) 及び (6.18) を間隔 μ 時間フレームごとにサンプリングしたものとする。

また、上記の処理で用いた l_2 は、動体のみが分離されたグレイスケール動画の初期フレーム $\mathbf{X}_1^{(\text{mv})}$ で、 $\mathbf{r}_t^{(1)}$ を基準として、位相 ϕ 方向の画素の輝度値を調べ、値が大きく減少した画素を先端 $\mathbf{r}_1^{(2)}$ として、次式のように求める。

$$l_2 = \|\mathbf{r}_1^{(2)} - \mathbf{r}_1^{(1)}\|_2 \quad (6.19)$$

二重振り子の先端のピクセル位置 $\check{\mathbf{w}}^{(\text{tip})}$ 及び $\check{\mathbf{h}}^{(\text{tip})}$ から滑らかな軌跡を描く処理は、2.5 節で説明した 2 次元空間上の時系列信号に対する 3 次スプライン補間を適用することで実現できる。

6.5 本章のまとめ

本章では、概形情報を用いた軌跡追跡法について説明した。本手法は、二重振り子先端が連結部を中心とした回転運動を行うといった性質を利用して先端の位置を推定するといったものであり、4章で述べた、色情報を用いた手法に対して、低コスト及び容易な準備で軌跡推定に適用できる。しかしながら、扱う二重振り子に対応したテンプレートの用意が必須という弱点がある。次章では、本手法を用いた二重振り子の軌跡推定実験について述べる。

第 7 章

概形情報を用いた軌跡推定法の実験

7.1 はじめに

本章では、6 章で提案した手法の性能を評価するために行った二重振り子の軌跡推定実験について示す。5 章と同様、本手法を用いた軌跡の推定と正解データとの比較を行う。本実験の条件を 7.2 節、結果を 7.3 節に示す。最後に、7.4 節で本章をまとめるとする。

7.2 実験条件

本実験では、5.3 節と同様、本手法の性能を評価するため、RMSE, MAE, MAPE, 及び PCC の 4 つの尺度で比較を行った。なお、5.3 節で、色情報を用いた手法は、正確に二重振り子先端部の軌跡推定を行う性能があることが確認された。そこで、本章では、色情報を用いて推定したものと正解ラベルとした。また、推定したプロット点を間引いたことで推定ラベルがない時刻では、スプライン曲線の位置で正解ラベルとの比較を行った。

本実験で用いた動画は、5 章のものと同条件下で、時間を 32 s にして撮影したものであり、時間フレーム数は $T = 960$ frames であった。動体の分離に用いる閾輝度値 λ は 35 と設定した。また、半径の大きさを調整するパラメータ d を 10, 移動メディアンフィルタ及び移動合計値フィルタのフィルタ長 L_{filter} をそれぞれ 5, 11, プロット点のサンプリングレート μ を 4 とした。

7.3 実験結果

Fig. 7.1 に、正解ラベル、本手法による推定ラベル、及びスプライン曲線を示す。ここで、推定したプロット点を間引いたことで、正解ラベルと推定ラベルの個数に差が生じている。正解ラベル及び推定ラベルは色情報を用いた手法の実験結果と比較して、それが大きくなっている。また、描かれた曲線は二重振り子先端の軌跡をおおよそ推定できた一方、推定したプロット点への移動メディアンフィルタの適用により、全体として歪みが確認された。次に、正解ラベルと推定ラベルの類似度を 4 つの方法で評価した結果を Table 7.1 に示す。推定したプロット

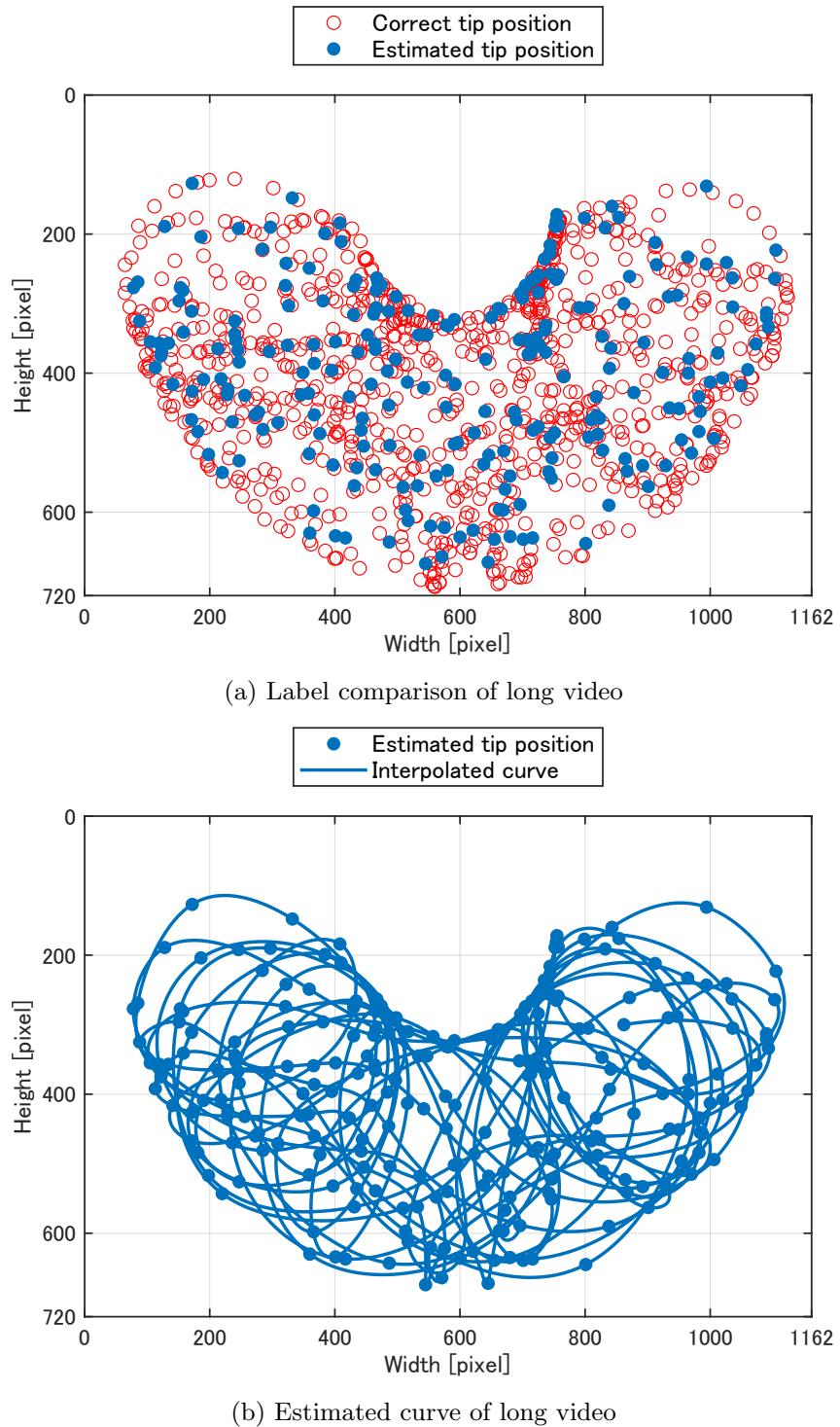


Fig. 7.1: Double pendulum trajectory estimated by proposed method using form information.

ト点への移動メディアンフィルタの適用及び間引いたことで少數となった推定ラベルに対する

3 次スプライン補間の曲線推定精度の低下により, RMSE, MAE, 及び MAPE の値は色情報を用いた手法の実験結果と比較しておよそ 10 倍の値を取っており, 本手法が外れ値に頑健でないことが確認された. また, PCC は, 色情報を用いた手法による値と比較すると低下した一方, ほとんど 1 であり, 本手法も同様に, 実際の軌跡に近い概形の軌跡を推定する性能を持つことが評価できた.

Table 7.1: Scores of proposed trajectory estimation using outline information

Subjective criterion	RMSE [px]	MAE [px]	MAPE [%]	PCC
Score	40.9429	10.3468	53.9887	0.9732

7.4 本章のまとめ

本章では, 二重振り子の概形情報を用いた手法による軌跡推定実験についての解説を行った. まず, 正解ラベル, 推定ラベル, 及びスプライン曲線を示した. 続いて, 正解ラベルと推定ラベルの類似度を 4 つの尺度で評価することで, 本手法の外れ値に対する頑健性の低さ及び推定した軌跡の概形の正確性を示した. よって, 本手法は色情報を用いた手法と比較すると, 軌跡推定の精度は低い一方, 初期値鋭敏性による二重振り子先端の軌跡の違いを示すことは可能であり, 二重振り子を STEM 教育の題材とするのに十分な性能であると評価できる. 次章では, 本論文をまとめる.

第8章

結言

本論文では、STEM 教育で初期値鋭敏性を扱うため、代表的な手法である、長時間露光を用いた撮影より容易に二重振り子先端の軌跡推定が可能な手法を 2 つ提案した。1 つ目の色情報を用いた手法では、二重振り子先端部の LED の発光色を用いた。これは実験で、実際の軌跡をほとんど正確に推定できる性能を持つことが示された。また、2 つ目の概形情報を用いた手法では、二重振り子先端が連結部を中心とした回転運動を行う性質を用いた。これは実験で、色情報を用いた手法と比較すると精度は下がる一方、初期値鋭敏性による二重振り子先端の軌跡の違いを示すのに十分な性能であることが示された。

最後に今後の課題を述べる。本論文で色情報を用いた手法より適用の一般性が高いと述べた概形情報を用いた手法は、テンプレートの作成等、扱う二重振り子の概形への依存があり、十分に一般性が高く、適用が容易である方法とは乖離がある。そこで、更に高度な動画像処理及び数学的処理を用いた手法の検討が必要であると考える。また、本論文では 240fps のフレームレート及び 8 倍スローモーション動画で各手法の性能を評価したが、より高いフレームレート及び高倍率のスローモーション動画を用いた実験による性能の推移を確認する必要がある。

謝辞

本論文は、香川高等専門学校電気情報工学科北村研究室にて行われた研究に基づくものです。まず、本研究を進めるにあたり、ご多忙のところ熱心にご指導くださいました指導教員の北村大地講師に心より感謝申し上げます。北村大地講師には、論文執筆や研究に関する議論だけではなく、サーバの増設や周辺機器の購入等の設備面の補助、研究機材の扱い方や環境構築等、細部にわたるまで丁寧にご指導いただきました。

本論の副査である重田和弘教授には、論文の構成や記述に関して大変有益な助言を頂き、大変お世話になりました。ここに厚く御礼申し上げます。

東京農工大学の矢田部浩平准教授には、共同研究ミーティングや合同ゼミを通じ多数のご支援とご助言をいただきました。心より感謝申し上げます。

北村研究室の先輩である専攻科1年の川口翔也氏、蓮池郁也氏、溝渕悠朔氏、村田佳斗氏、同期の綾野翔馬氏、唐渡昂希氏、島田優斗氏には、研究に関するアドバイスや日頃の交友関係等、1年に亘る研究室生活を強く支えていただきました。ここに感謝申し上げます。

最後になりますが、現在に至るまで私の学生生活を金銭的に支え、暖かく見守って下さった両親には感謝の念に堪えません。これまで本当にありがとうございました。

参考文献

- [1] 高阪将人, 松原憲司, “資質・能力の育成を重視する教科横断的な学習としての STEM 教育と問い合わせ,” *科学教育研究*, vol. 41, no. 2, pp. 150–160, 2017.
- [2] 石井和之, 大島まり, 川越至桜, “大学と企業の協働によるアウトリーチ活動を基盤とした STEM 教育,” *科学教育研究*, vol. 39, no. 2, pp. 59–66, 2015.
- [3] 土山紫穂, 林亮子, “画像処理を用いた折り紙飛行機の軌跡追跡,” *第 77 回情報処理学会全国大会講演論文集 2015*, vol. 2015, no. 4, pp. 61–62, 2015.
- [4] 河並崇, “プログラミングを中心とした STEM 教育教材の共有のための「STEM マップ」の提案,” *第 81 回情報処理学会全国大会講演論文集 2019*, vol. 2019, no. 1, pp. 367–368, 2019.
- [5] 石井雅樹, 御室哲志, “ロボットを題材とした産学官地域連携による創造ものづくり教育の活性化,” *秋田県立大学ウェブジャーナル A (地域貢献部門)*, vol. 5, pp. 24–31, 2018.
- [6] 石崎博基, 宇佐美仁英, 川田重夫, 佐藤佑哉, 前田太陽, 真鍋保彦, “画像処理と分散処理を用いた植物形状認識 PSE システム,” *日本計算工学会論文集*, vol. 2018, no. 1, 2018.
- [7] 荒牧義孝, 岸木敬太, 鮫島朋子, 福島和洋, 宮川英典, 米村和幸, “簡易二重振り子による教材研究,” *日本物理学会講演概要集 2004*, vol. 59, no. 1.2, pp. 417, 2004.
- [8] E. Lorenz, “Predictability: Does the flap of a butterfly’s wings in Brazil set off a tornado in Texas?,” *Proc. American Association for the Advancement of Science*, pp. 260–263, 1972.
- [9] 小室元政, “二重振り子とカオス,” *日本ロボット学会誌*, vol. 15, no. 8, pp. 1104–1109, 1997.
- [10] 鈴木三男, 増田健二, “二重振り子におけるカオス的振舞,” *物理教育*, vol. 48, no. 1, pp. 1–5, 2000.
- [11] 鈴木三男, 増田健二, “二重振り子によるカオス,” *KEK Proceedings*, pp. 23–26, 1999.
- [12] 畑浩之, “解析力学,” 東京出版, 2014.
- [13] 二間瀬敏史, 綿村哲 “解析力学と相対論,” 朝倉書店, 2010.
- [14] 北原大地, “区分的多項式とスプライン関数の基礎,” *日本音響学会誌*, 一般社団法人日本音響学, vol. 78, no. 10, pp. 570–577, 2022.