

AI-based Spectral Analysis on Music Conducting

Fourier-Analytic Expansion of Boulez-Ito's Angular Dynamic Method for Conducting

Ken ITO, Kazuki OTSUKA and Kota HAYASHI

Interfaculty Initiative in Information Studies

The University of Tokyo

AI を用いた指揮技法のスペクトル解析

ブーレーズ=伊東“アンギュラーダイナミクス”技法の調和的拡張

伊東 乾 大塚一輝 林康太

梗概

合奏を収録した2次元の動画データから指揮の身体運動情報を抽出し、モデル身体を用いた機械学習システムによって2次元および3次元の演奏運動データを構成、そこにブーレーズ=伊東のアンギュラーダイナミクスを適用することで、指揮の基本的な運動を身体各部位関節の角速度、角加速度に分解し、そのフーリエ変換から筋骨格系の演奏運動のスペクトルを導出した。演奏の身体運動そのもののスペクトルやハーモニクスを扱う議論はかつて存在しない。ここから多様な演奏の身体技法を、新たな形で体系化することが期待される。

キーワード 指揮 身体運動 機械学習 筋骨格系 機械学習 フーリエ解析 スペクトル

1 ベルリオーズの図形からサイトウ・メソッドへ

西欧の合奏音楽における指揮者の登場は大まかに17世紀フランスのジャン・バプティスト・リュリ Jean-Baptiste LULLY (1632-87) の時期とかがえられ、近代的な指揮の原点の一つとしてエクトール・ベルリオーズ 19世紀初頭フランスの H.ベルリオーズ Hector BERLIOZ (1803-69) の名が挙げられる。一般にベルリオーズに由来するとされる以下の4拍子を示す「図形」(Fig.1)は21世紀の合奏でも利用されることがある。図で第3拍が「-2」第4拍が「-1」となっているのは次の小節の第1拍よりも各々2拍前、1拍前を示すものである。

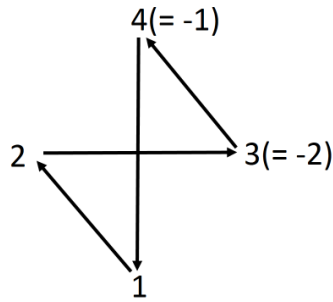


Figure.1 “Berlioz’s figure” for quadruple trajectory

19 世紀後半、主として H.フォン・ビューロー Hans von Bülow (1832–94) 以降の世代で職業的な専門指揮者が登場するとともに、さまざまな「指揮法」が提案された。4 拍子に関して考えるなら Fig.2 のような「図形」は 21 世紀の義務教育教程でも利用されることがある。

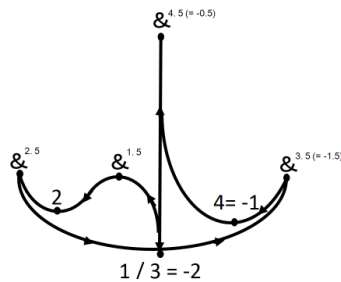


Fig.2 Typical trajectory for quadruple meter from 19th century

ユダヤ系ポーランド人指揮者ヨーゼフ・ローゼンシュトック Joseph ROSENSTOCK(1895-1985)は 20 世紀前半の新月物主義的な演奏傾向の中で、原典に忠実でシンプルな解釈と演奏のスタイルを実践していた。ナチスの惨禍を避けてローゼンシュトックが日本に滞在していた 1936-46 年にかけて、斎藤秀雄 Hideo SAITO(1902-74)は彼のバトンテクニックをモデル化し、一般にサイトウ・メソッドと呼ばれる指揮のテクニックを体系化した。Fig.3 としてサイトウ・メソッドの 4 拍子の図形を示す。

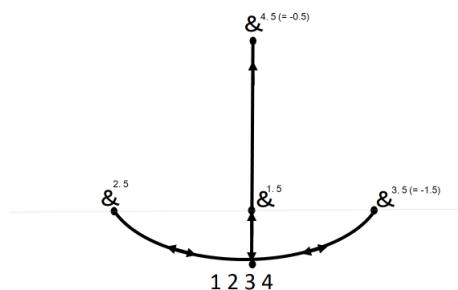


Fig.3 Figure of quadruple meter after the “SAITO method”

Fig. 1,2 および 3 を比較すると、ベルリオーズに由来するとされる Fig. 1 が主として直線で拍を示すのに対して、Fig. 2 ではそれらが曲線となると同時に、1 拍を分割したオフビー

トの点が生されているのが顕著な違いとなっている。

ただし Fig. 2 では 1 拍目、2 拍目、3 拍目・・・の各拍の打点 beat point が異なっており、奏者に正確な拍を伝える上で難点があった。

この難点を克服したのがローゼンストック＝斎藤による Fig. 3 で、すべての拍の打点は中央に集中しており、正確な拍打が可能で、より進んだ、プロフェッショナルの目的に即した優れたものと一般には考えられている。

ここで改めて軌跡の解析的な側面に注目すると、Fig. 2 には存在し、Fig. 3 で消去される要素として、Fig. 4 に示すような変曲点 inflection point の有無を指摘しておこう。

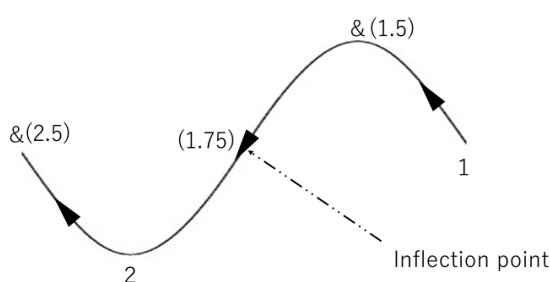


Fig.2 Inflection point in the 19th century-type trajectory

Fig.2 の 1 拍目と 2 拍目の間には、1.5 拍目に相当する極大値 maximal があり、2 拍目は極小値 minimal であることから、実質的に 1.75 拍近傍に相当する変曲点が存在する。現実の演奏ではこの点は静止して示されることはなく、最大速度に近い運動で通過するだけであるが、これによって 1 拍の 2 分割よりさらに細かい、基本拍を 4 分割した音価が生されていることに注目しておこう。

2 ブーレーズ+伊東のアンギュラーダイナミクス

20 世紀後半から 21 世紀初頭にかけて活躍したフランスの作曲家・指揮者ピエール・ブーレーズ Pierre BOULEZ(1925-2016)は、指揮の運動を特定する用語の混乱とメソッドの不在を指摘し、適切な対策の必要性を訴えた。本稿の著者の一人伊東は、ブーレーズのこの要請に応えるべく、画像解析を用いたモーション・キャプチャーを用い、従来のような「図形」を一切用いない指揮技法 Angular dynamics の体系化を行った(2004-07)。

現実の指揮の運動は骨格筋が駆動する間接の「広義の回転 rotation (in broad sense)」のみで構成されるものとみなせるので、その回転角の変化、より正確には「角速度」「角加速度」だけで記述することが可能である。

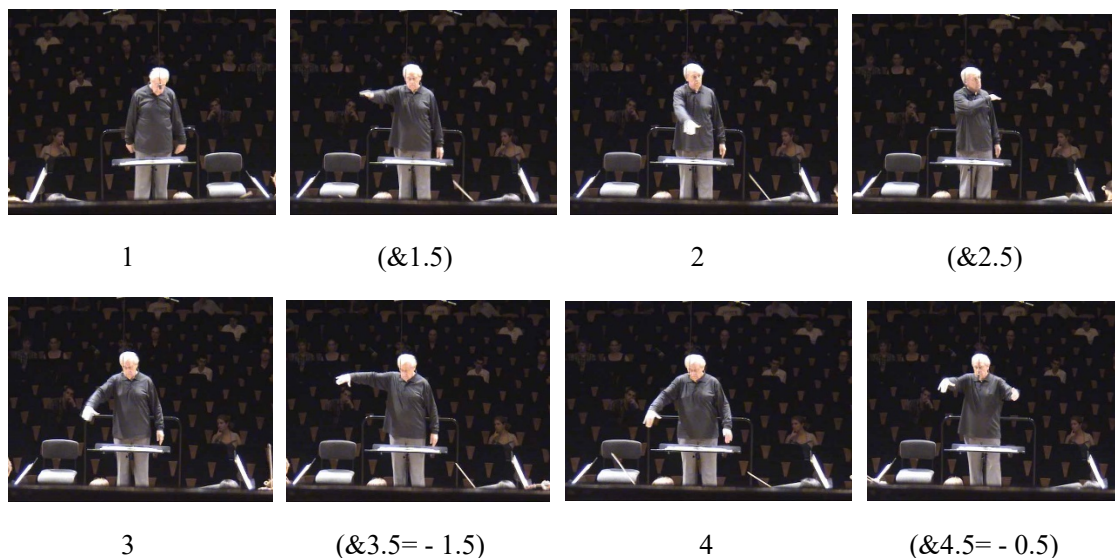


Fig. 5 Quadruple conducting gesture by Pierre Boulez

ブーレーズと伊東のメソッドを用いると、従来は2次元平面上に図示されていた「図形」は、肘関節と手首関節の2自由度に置き換えることができる。これはx-yなど2自由度の平面上に射影として描画されていた軌道を、肘関節の屈伸角 θ と手首関節の回転角 ϕ （ならびに一定の値とみなすことができる下腕長 r で表現するもので、従来「図形」が持っていたすべての情報を含む以上に、身体の中のどの部位を用いるかが特定され、さらに角速度 $\dot{\theta}, \dot{\phi}$ 角加速度 $\ddot{\theta}, \ddot{\phi}$ などを陽に扱うことで、明らかにより進んだ演奏技法を開拓する可能性をもつものである (Fig. 6)

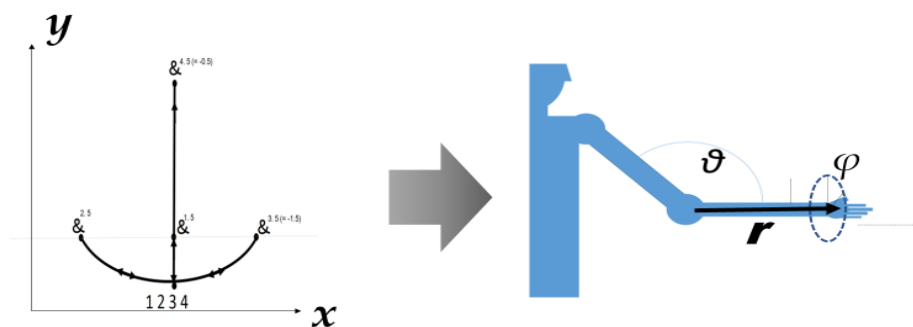


Fig. 6 From x-y plane projection to θ - ϕ - r angular dynamics

この観点からブーレーズと伊東は2005-07年にかけて系統的にモーションデータを取録、解析したが、演奏を妨げない2000年代のマーカーレス・モーションキャプチャー技術には解析可能な自由度に限界があり、システムティックな解析には限界があった。

3 機械学習を用いた演奏運動のスペクトラル・アナリシス A 2次元データ解析

2012年のGoogle社によるネコの顔の教師なし同定の成功*)以降、深層学習のニューラルネットワークを用いた画像情報解析が飛躍的に進み、自動運転の基本技術などとして機械学習の人工知能開発が産業規模で推進されるようになった[1]。

アニメーションや映画産業の要請により、モーションデータのAI解析技術も2010年代後半に入って顕著な発展を見せた。そこで、これらを応用し、新規に収録した指揮の2次元動画モーションデータから、ブレーズ=伊東の θ を2次元射影、および3次元データとして推定し、再構成を試みた。

撮影された画角のままの画像データから、演奏身体運動データの特徴量を抽出した。解析にはカーネギー・メロン大学CMUのPerceptual Computing Labが開発した多層畳み込みニューラルネットワーク(CNN)のモデルOpenPose[2]を用いた。また解析結果はPAF(Part Affinity Field)表現[2]によって人体の25部位から成るベクトル領域のキーポイント座標で出力した。学習済みモデルとしてはMP II human multi-person dataset[3]およびCOCO[4]で訓練されたCMU提供のモデルを用いた。被写体右腕の運動から特徴量を抽出するべく、右方向から撮影された動画をを入力データとし、出力キーポイント座標の肩関節、肘関節、手根部の3点のxy2次元座標を確定、肩関節座標と肘関節座標を結ぶベクトルを v_1 、肘関節座標と手根部座標を結ぶベクトルを v_2 とすると肘角度 θ は

$$\theta = \arccos\left(\frac{\vec{v}_1 \cdot \vec{v}_2}{|\vec{v}_1||\vec{v}_2|}\right) \dots \textcircled{1}$$

として求められる。またその時間変化率すなわち角速度 $\dot{\theta}$ 、角加速度 $\ddot{\theta}$

$$\dot{\theta} = \frac{\partial \theta}{\partial t} \dots \textcircled{2} \quad \ddot{\theta} = \frac{\partial \dot{\theta}}{\partial t} \dots \textcircled{3}$$

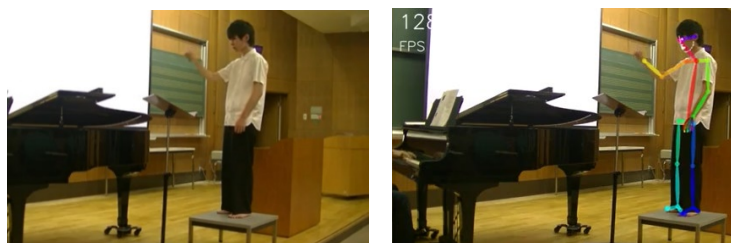


Fig. 7 Extraction of corporeal motion data from 2-D motion picture data

Fig.8-a.b.cとしてモーションデータから推定された2次元射影 $\theta, \dot{\theta}, \ddot{\theta}$ 値の例を示す。

A. Sampling Rate 30Hz

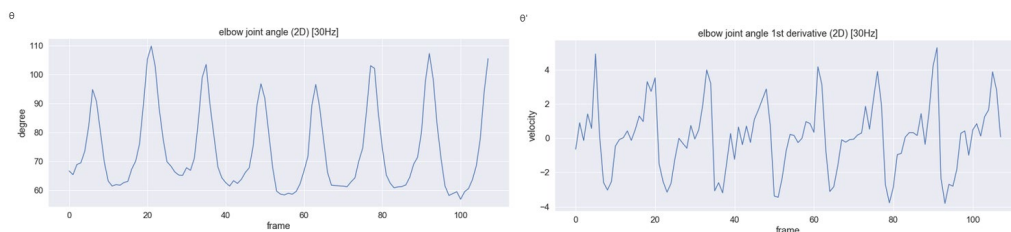
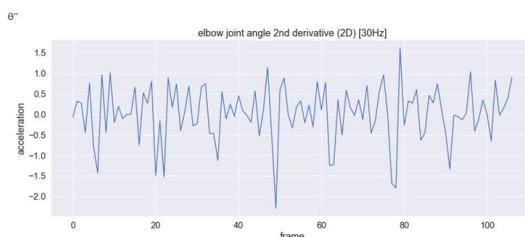


Fig.8-a. A sample of estimated θ values. b. A sample of estimated $\dot{\theta}$ values.



c. A sample of estimated $\ddot{\theta}$ values.

これらの基本的な特徴として、角度の推移は上に凸の曲線を描き、その一回微分は矩形波に近い挙動、二階微分はスパイクを伴う一定値に、被験者が初心者であるために発生するノイズ=余計な力による陽動が印加されたものとして解釈することが出来る。

次に Fig.8-a,b,c をサンプリング周波数でフーリエ分解したスペクトルを Fig.9-a,b,c として示す。

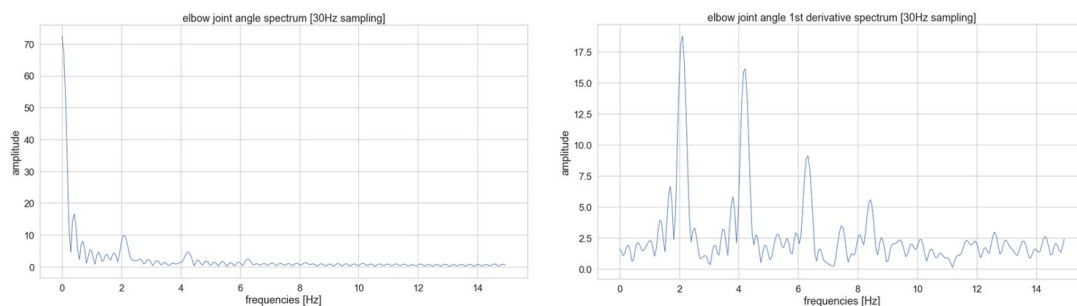
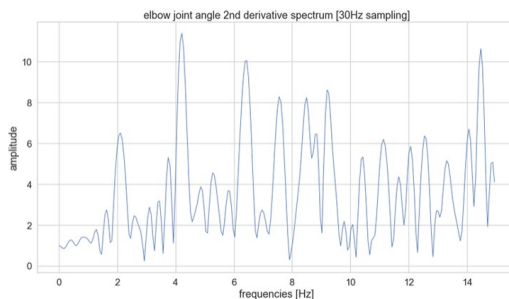


Fig.9-a. A sample of estimated θ spectra. b. A sample of estimated $\dot{\theta}$ spectra.



c. A sample of estimated $\ddot{\theta}$ spectra.

Fig.9-a は肘角度変化のヴァリエティを示しており、運動の反復により 2 ヘルツ、4 ヘルツ 6 ヘルツと倍々のスペクトル構造が確認される。2 ヘルツ強の基本拍に対して 4 ヘルツ

の成分は上行下向の「半拍」の成分、6ヘルツの成分は1拍を3分割する陰伏の成分を明示的に示している。

Fig.9-b は肘角度運動の角速度スペクトルを示すが、明確にハーモニックな構造が確認される。

Fig.9-c は角加速度のスペクトルを示し、低域から測定限界の広域まで、さまざまな成分が確認されることから、いまだ特定の筋を訓練することで制御された運動を実現できていない初心者の運動であることが確認される。

4 機械学習を用いた演奏運動のスペクトラル・アナリシス B 3次元データ解析

同一の動画データを対象に、3次元身体運動データの抽出を試みた。3次元解析では Facebook Research 開発によるニューラルネットワークモデル VideoPose3D [5] を用いた。VideoPose3D は前節で用いた OpenPose により出力された 2D キーポイントの時間シーケンスを入力し、CNN の Dilated Temporal Convolution アーキテクチャにより時間依存関係を考慮した 2D キーポイントから 3次元空間への変換写像を学習する。学習済みモデルとして Human3.6M [6], HumanEva-I [7] で訓練された Facebook Research 提供の身体運動モデルを用いた。2次元解析の際と同様に関節位置のキーポイントの時間シーケンスから角度 θ 、角速度 $\dot{\theta}$ 、角加速度 $\ddot{\theta}$ を求め、さらに、それらのサンプリング周波数上でのフーリエ変換からスペクトルを求めた。

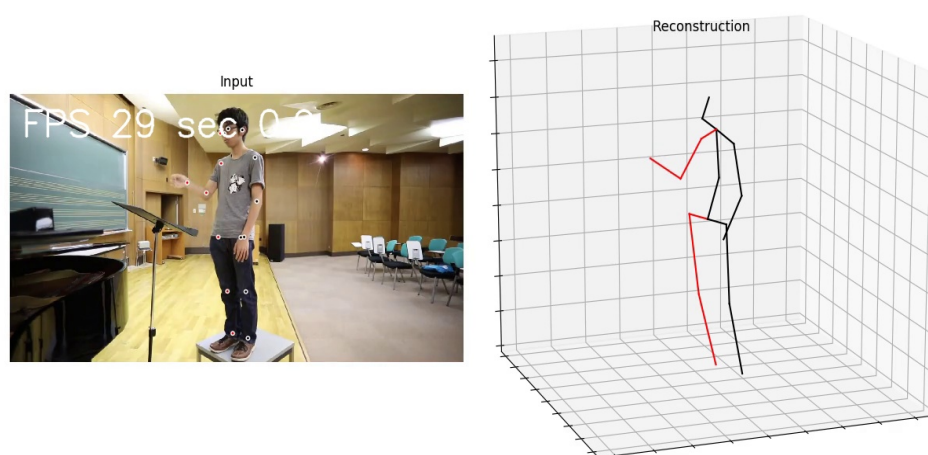


Fig.1 Estimation of 3-D corporeal motion data

前節で用いたのと同じ動画データによる θ 、 $\dot{\theta}$ 、 $\ddot{\theta}$ を 2D、3D モデル各々で解析した結果を重畳して Fig.11 に示す。

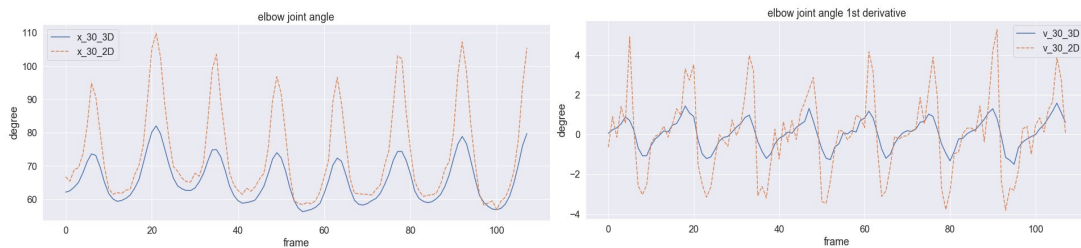
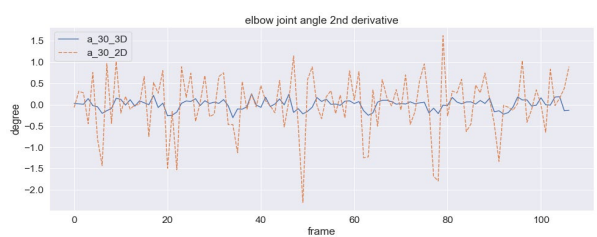


Fig.11-a. A sample of estimated 3 D- θ values. b. A sample of estimated 3 D $\dot{\theta}$ values.



c. A sample of estimated 3D- $\ddot{\theta}$ values.

3次元解析の利点は関節の屈伸角を射影ではなく絶対値で評価できることにある半面、推定演算により2次元解析で得られる情報の一部は失われる傾向も確認される。

新規にデータテイク、すなわち演奏収録を行う場合には、サンプリング周波数すなわち動画収録のフレームレートを上げることで、機械学習システムの出力精度上昇が期待できる。以下ではフレームレート 120fps=サンプリング周波数 120Hz で収録、解析した、上から初心者/中級者/上級者の同一楽曲同一部分を指揮したデータによる解析結果を示す。楽曲は P.Tchaikovsky String Serenade Op.48 II Valse: Moderato 冒頭を用いた。

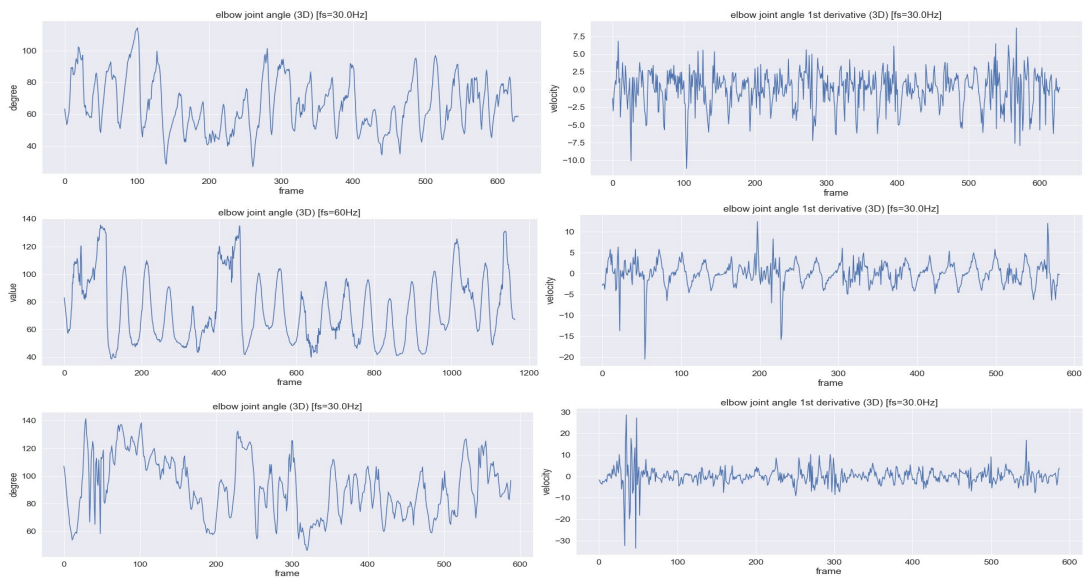
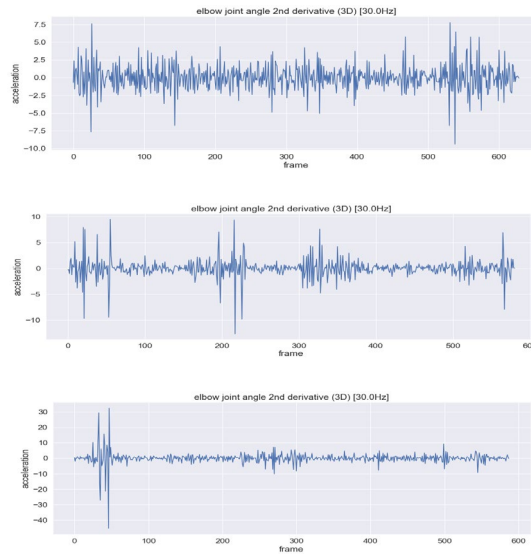


Fig.12-a. 3 D- θ values of different level conductors. b. 3 D $\dot{\theta}$ values of them.



c. 3D- $\ddot{\theta}$ values of top: beginner middle: student bottom professor conductors respectively

3つのデータのうち最も下に示したのは40年の経験を持つ被験者で、具体的には筆者自身のものである。肘角度は様々な値を取るが、その角速度は揺れが少なく、角加速度すなわち力の入りかたは常時ほとんど極小値に留まる。初学者、中級者の場合、指揮の運動自身は単調ながら、速度のブレが大きく、加速度の推移から余計な力が入っていることを明確に確認することが出来る。より進んだ解析と、それらに基づくブーレーズ=伊東のアンギュラーダイナミクス・メソッドの展開については、続稿で展開する予定である。

5 機械学習を用いた演奏運動のスペクトラル・アナリシス C 歴史的画像データの解析

このような解析は、新規に音声動画収録を行って実行するのみならず、過去に収録されたデータ、とりわけ歴史的演奏の記録動画を基に行うことが可能である。2004年から07年にかけて筆者とP.Boulezが行ったAngular Dynamicsの確立初期の音声動画データを用いて、同様の解析を行った。これにあたってはIRCAMのCarlos Agonによる示唆が大きく影響している。同氏に心から感謝したい。

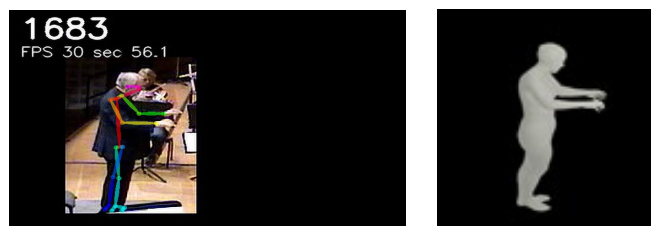


Fig.13 Motion data estimation of P. Boulez's conducting and polygon synthesis

身体運動データからポリゴン・メッシュを推定し、モデル身体にマッピングすることで、すでに地上を去った歴史的巨匠の身体運動を、骨や筋の個別運動から解析、再構成し、次世代のオーソドックスな演奏技法の確立を企図することが可能であり、そうした方向での模索も並行して行っている。以下では前節までと同様、3次元解析で得られたブーレーズの演奏例を検討してみよう。演奏している楽曲は Luciano Berio “Chmin”冒頭部である。

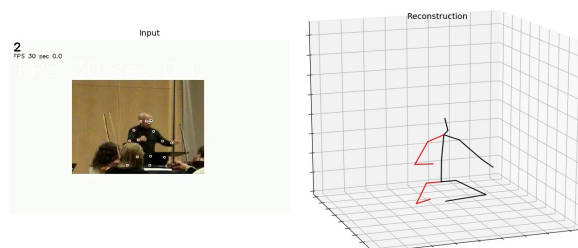


Fig.14 3D Motion data estimation of P. Boulez's conducting

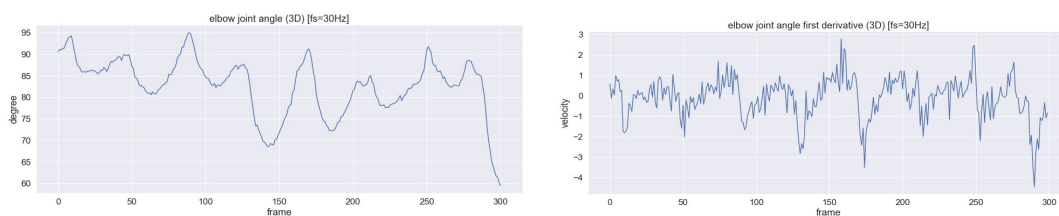
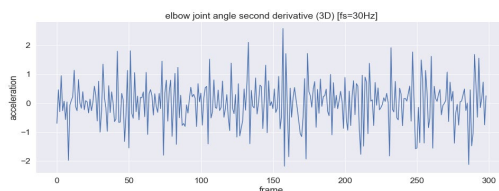


Fig.15-a. Boulez's sample of estimated 3 D- θ . b. Boulez's sample of estimated 3 D $\dot{\theta}$.



c. Boulez's sample of estimated 3D- $\ddot{\theta}$.

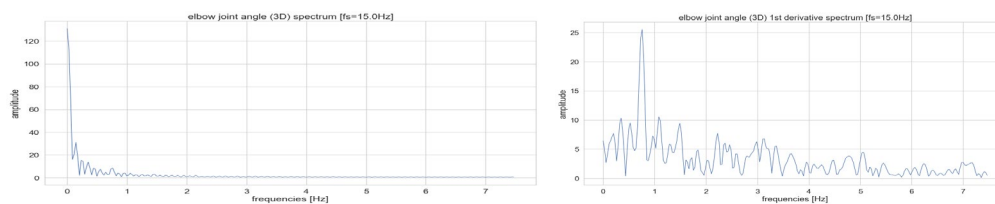
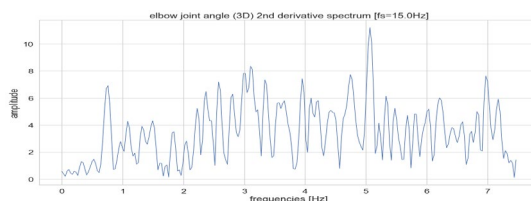


Fig.16-a. Boulez's sample of estimated 3 D- θ spectra. b. Boulez's sample of estimated 3 D $\dot{\theta}$ spectra.



c. Boulez's sample of estimated 3D- $\ddot{\theta}$ spectra.

Fig.15 として θ とその時間変化率の挙動を、Fig.16 にはそのスペクトルを示した。Fig.15 の各グラフから指揮のジェスチャーそのものは緩やかであることが知れる。Fig.16-b から角速度は 0.8 ヘルツ付近すなわちメトロノーム・テンポ MM=75 付近に明確なピークを持ち、かつその運動を構成する加速度スペクトルの分布は Fig.16-c から 5 ヘルツ強すなわち MM=300 すなわち基本拍を 4 分割するカウントで厳しく律せられていること、かつ多様なキューも出されていることが知れる。生前のブーレーズが内的に感じているテンポが、これらの解析を通じて如実に確認出来る。

こうした多様なキューと奏者の具体的な演奏との相関は、多体解析を通じて分析することが可能である。継続的な課題として現在、これらにも取り組んでいる。

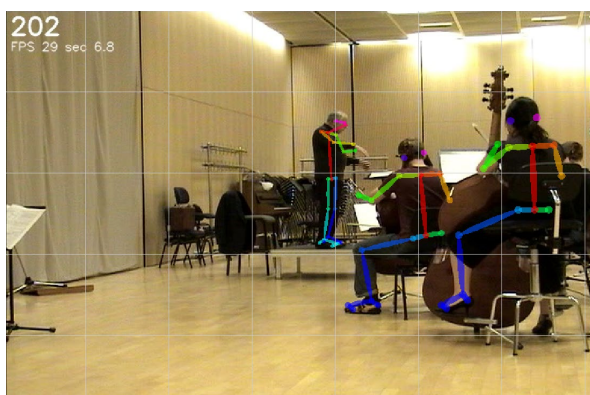


Fig.16 Example of many-body machine learning analysis and corporeal interaction study

6 Angular dynamics and Spectral methods : 新しい指揮テクニック創出のために

誤解のないように強調しておくが、このような解析を通じて私達は過去の巨匠と同様に動作するロボットを作り出したい訳ではない。現在、そして未来に音楽を造形する若い世代に、適切な演奏のメソッドを提供することに専ら関心を持つ。翻っていうなら、レッスン室やリハーサルで用いられる音楽の言葉は、科学の観点に立つなら極めて多くの初歩的な誤謬を含んでいる。

そうした誤りを正すうえで、演奏者自身が主観的に感知できない自身の演奏の特徴を、客観指標として取り出せることに、こうした AI 解析の大きな利点がある。同様の利点はスポーツ技能の向上や、リハビリテーションなどにも活用が可能で普遍的なものであるが、音楽のテクニックは最も繊細に、そのメリットを示すことが出来る例となっている。

Fig. 9-b,c のように、初心者のバトン運動に顕著なスペクトルのピークが観察されることを、演奏している本人も、またそれを見ている奏者も、まったく気づくことはない。だが、このような解析を通じて検討することで、より進んだ演奏技術を容易に、かつ高精度で獲得することが可能になる。

Fig.9-c のように肘角度変化の角加速度が、明確な周波数成分のピークに分解可能であることは、音楽家一般には周知でなく、演奏している本人も自覚することが出来ない。

しかしこれらは、筋骨とそのニューラルな制御を考えれば自然に理解することが出来る。

私達の演奏運動は、骨格系を駆動する個々の筋の収縮によって実現されており、おのおの筋の収縮は一定の時定数を持っている。肘関節は上腕側の上腕骨と下腕側の尺骨・橈骨、三つの骨から構成され、多数の筋の起始 muscle start と停止 muscle stop が結合している。

肘の進展に関しては、互いに拮抗筋の関係にある上腕三頭筋 triceps brachial muscle (伸展 extension) ⇔ 上腕二頭筋 biceps brachial muscle (屈曲 flexion)、肘筋 anconeus muscle (伸展) が協調して作動している。また肘の屈曲に関してはさらに多数の筋群、すなわち上腕二頭筋 biceps brachial muscle、上腕筋 brachial muscle、腕橈骨筋 brachioradial muscle、円回内筋 Circular prosthesis、橈側手根屈筋 Radial carpal flexor、尺側手根屈筋 Ulnar carpal flexor

これらの筋群が、単一関節の屈伸に協調しながら関係しており、かつおのおのが固有の時定数をもって作動することで、見かけ上は単純と思われる演奏動作を形成している。これらをスペクトル成分に分けることで、私達が主観的には感知困難な運動の独立、とりわけ脱力に関するテクニックを、以下ではいくつか導くことにする。

1 軌道角運動量の制御による拍分割

Fig.16 に示した、ブルーエズの指揮ジェスチャーのスペクトル成分を検討してみよう。先にも触れたとおり、軌道運動は MM=75 程度の、極めて一般的な拍を示しているが、その運動を構成している加速度成分は MM=300 すなわち基本拍の 4 倍の成分が顕著なピークを見せている。この運動は以下のように模式化して考えることが出来る。

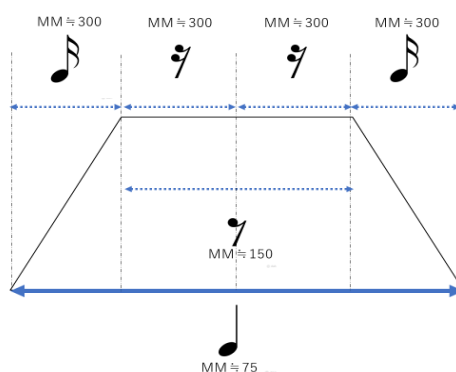


Fig.17 Example of many-body machine learning analysis and corporeal interaction study

「図形」を用いた軌道の指示では、一般には運動の速度は示されることがない。だが実際には音楽的な要請によって、おのおのの拍は様々に分割され、その角速度、角加速度は一樣ではない。一例を Fig.18 として示す。

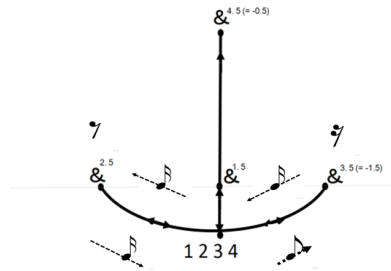


Fig.18 An example of inhomogeneous division of a quadruple meter

Fig. 18 の例では、第二拍を Fig.17 で示した分割で指示している。この指示は第三拍以降の拍の分割を示す上で有効である。また第三拍は前半を 8 分音符の音価で、後半を 16 分音符の音価で運動し、単点に 16 分音符分停留する場合を想定したものである。こうした個別の「軌道運動 orbital motion」を示す角運動量とその変化を、Fig.16 に示すようなスペクトルとして観測することが可能である。

Fig. 16, や Fig.17 の模式は、あくまで単純化したものであるが、レッスン室で用いられる言葉は、この程度の精度、あるいはさらに粗いものに、一般には留まる。ここで Fig.2 に示した模式的な運動軌道の変曲点を思い出そう。緩やかな指揮の動作に関しては、一般に下向きの運動、すなわち θ が増加する方向での有働は、十分に脱力して腕の重さと重力に任せることが基本となる。軌道運動の「変曲点」は、軌道の 2 階微分、すなわち作用する力がゼロになる点を示しており、重力と拮抗して合力がゼロになる近傍を示している。

すでに述べた通り、個別の筋収縮は時定数を持つものと考えられ、それらが音楽的な要請、とりわけ拍の分割と整合した形でコントロールされることで、自在なキューとして奏者に適切な指示を伝えている。「打点」やその「鋭さ」といったレッスン室の言葉は、実はすべて時間幅を持つ現象であり、音楽的な拍そのものもまた「点」ではなく幅を持つ。それらに対応して個別の筋収縮の時定数を制御し、全体としてシンプルで整合した軌道運動の加速や原則を実現していることに注意する必要がある。実際の演奏動作は Fig.17 のような直線的なものではなく sigmoid curve などで近似するほうが自然である。その微分やフーリエ変換を念頭におけば、拍打が複数の動作筋収縮に基づくスペクトル構造を持ち、それらが音楽的に整合した成分公正を持つことで適切な音楽的指示を与えることは、自明であろう。しかしそんな事実を演奏者の主観で捉えることは至難である。指揮動作の調和解析に基づくメソッド創成を私達は spectral conducting と呼ぶ。より多様な spectral conducting の方法と、それを実現するトレーニング手法を、私達は継続して開発してゆく考えである。

2 摂動的な取扱い・・・手根＝肘関節相互作用と高次スペクトル

ブレーズ＝伊東のアンギュラーダイナミクスでは、手首の運動とりわけ回転が決定的な役割を果たす。そこで、動画データの機械学習解析でも手首の回転運動の抽出可能性を検討しているが、これは必ずしも容易ではなく、とりわけ歴史的動画からの抽出は困難である。

だが、ここで手首の運動をつかさどる筋に注目すると、いくつかの事実が彰かとなる。

例えば「掌屈 palm flexion」と呼ばれる手首の屈曲には浅指屈筋 Superficial finger flexor muscle、深指屈筋 deep finger flexor muscle、橈側手根屈筋 radial carpal flexor、長掌筋 long palm muscle、尺側手根屈筋 Ulnar carpal flexor、長母指屈筋 long mother finger flexor などが関係するが、そのうち橈側手根屈筋 Radial carpal flexor は、先に記したように肘関節の屈曲にも関わり、浅指屈筋や深指屈筋等も前腕の肘関節に隣接して存在し、手首の運動をつかさどりながら。同時に肘関節にも影響を与える構造となっている。

このため、手首の運動は生理的に肘関節の運動を誘起する機会が多く考えられ、指揮やピアノ演奏に関しては、肘間接運動の細かな高次成分、摂動 perturbation として手首の運動を検出することが可能であると考えられる。

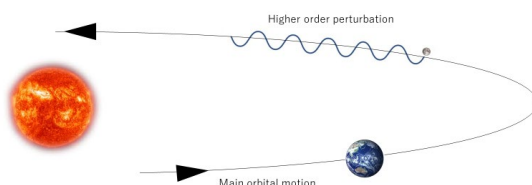


Fig.19 Main orbital motion and higher order perturbation

古典力学では太陽と地球、地球と月といった二体問題は厳密に解くことが出来るが、太陽と地球と月の運動を取り扱う三体以上の多体問題を厳密に解くことはできない。しかし第三番目以降の力が十分に弱いと考える適切な近似によって解を得ることが可能な場合がある。月の存在を無視すれば、地球は太陽の周りを楕円軌道を描いて運動すると考えられるが、月との相互作用によって現実には微小なずれを生じている。摂動法はこのずれを算出するべく古典力学で工夫され、とりわけ量子力学や場の理論の発展に主要な役割を演じている。

Fig. 16-c に示したブーレーズの演奏など、肘関節運動の角加速度スペクトルに見られる高次成分は、肘そのもののマクロな運動というより、肘関節に接続された手首や指の支配筋が作動することで、結果的に肘間接運動で観測される高次の摂動と理解することが原理的に可能である。実際の生理的詳細については継続的に体系だった研究を進める念頭である。

3 タクト、鍵盤、バイディング周波数・・・本質的に不可知の領域へ

P.ブーレーズは指揮棒を持たず、終生に渡って素手で指揮台に立った。だがより一般の指揮技法を念頭に置くとき、タクトの運動を考慮する必要がある。これにあたっては人体に外延してバトンを含む系を学習したデータベースを用いることも可能だが、摂動の考え方をいえw v バトンのない通常の身体運動データベースを用いた解析の可能性も外延される。

ブーレーズを筆頭に、すでにこの世を去った歴史的巨匠の演奏については、遺された動画

データのフレームレートに限界があり、30fps 以上の時間分解能を期待することは一般に困難である。線形補完などの方法で見かけ上の時間分解能を上昇させても、新たな微細効果、とりわけ非線形な要素を抽出することは困難、ないし不可能であろう。

この点で、新規にデータ収録が可能な演奏動画撮影に関しては、フレームレートの上昇によって別種の問題へのアプローチが可能になることを指摘しておこう。歴史的動画資料の多くが 30fps 程度のフレームレートで収録されているのは、人間の時間弁別分解能が 15Hz 程度までであること、つまり 20Hz を明らかに超える高速度で変化するパターンに、私達は前後関係を認識することが出来ない視覚生理的な特徴を踏まえて録画機材が設計されてきた事情による。

一方 120fps あるいは 240fps といったフレームレートで収録したデータは 60Hz 120Hz といった振動数での運動を抽出することが出来るが、これらは音波として考えれば可聴域に属する運動である。すでに音程として認知される 60Hz の振動を、私達はリズムとして前後関係を含めて弁別することが出来ない。トリルやトレモロ、あるいは打楽器のロールなどは前後関係を聴き取ることがすでに困難～不可能で、一種の持続音として認識される。

同様に、20 ヘルツを超えるような高速の運動を伴う指揮者のキューイングは、質的な認知を奏者や聴衆に与える。例えば、ヴィブラートを想起させるような素早い手掌の振動などを、具体例として挙げる事が出来る。

リズムの時間弁別の高域限界である 16-20Hz と、いわゆる可聴音の低域限界の間には、人間が認知しにくい振動数領域、聴覚の認知的死角の領域がある。この領域はまた人間の視覚、聴覚、触覚といった諸知覚を統合する「統覚のバインディング周波数 binding frequency of apperception」に相当する可能性があり、当該領域の周波数をもち脳全体で共有される γ 波が神経動機を司るものと考えられている。

120fps で収録された動画データは人間がリズムとしても音程としても認識することが出来ない領域での運動スペクトルを物理的に取り出すことが可能である。これにより、私達の認知をすり抜ける現象をとらえ、それを演奏の叡知に読み替える可能性が期待される、

その意味では、指揮の動作解析以上に、鍵盤その他、一般の器楽を対象とする高速度カメラを用いた解析により、生理的には認知不可能な領域での現象を捉えることが期待される。

指揮の動作解析では手掌や手指の運動に特化した学習データを活用しにくいだが、器楽演奏に関しては、手指の運動を、より高いフレームレートで収録、解析することで、新たな知見が得られることが期待される。

ピアノ、擦弦楽器、あるいはハープや琴などの撥弦楽器、多くの打楽器などで、1000fps 程度の高速度カメラを用いた収録と解析を実施することで、線形、非線形双方の可聴域での音波生成と身体運動のダイナミクスを、とりわけ認知的死角領域に注目しつつ解析することで、従来見逃されてきた現象を捉えること、またそれらを新たな演奏、創造の方法に転化することをひきつづき検討してゆきたい。

References

- [1] <https://googleblog.blogspot.com/2012/06/using-large-scale-brain-simulations-for.html>
- [2] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, Yaser Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields" <https://github.com/CMU-Perceptual-Computing-Lab/openpose>
- [3] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields" in *CVPR*, 2017.
- [4] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, "2d human pose estimation: New benchmark and state of the art analysis" in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 3686–3693.
- [5] Dario Pavllo, Christoph Feichtenhofer, David Grangier, Michael Auli, "3D human pose estimation in video with temporal convolutions and semi-supervised training" <https://github.com/facebookresearch/VideoPose3D>
- [6] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu. "Human3.6m: Large scale datasets and predictive methods for 3D human sensing in natural environments." *Transaction on Pattern Analysis and Machine Intelligence (TPAMI)*, 2014.
- [7] L. Sigal, A. O. Balan, and M. J. Black. HumanEva: "Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion." *International Journal of Computer Vision (IJCV)*, 87(1-2):4, 2010.
- [8] Federica Bogo, Angjoo Kanazawa, Christoph Lassner, Peter Gehler, Javier Romero, Michael J. Black, "Keep it SMPL: Automatic Estimation of 3D Human Pose and Shape from a Single Image" <https://smpl-x.is.tue.mpg.de>.
- [9] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, Michael J. Black, "SMPL: A Skinned Multi-Person Linear Model"