

楽器演奏コンピュータグラフィクス制作技術

巳波 弘佳

キーワード：音楽，楽器演奏，コンピュータグラフィクス，モーションキャプチャ，最適化

本稿は、関西学院大学の土井 智博さん（2014年度修士論文，2012年度卒業論文），平田 純也さん（2011年度修士論文，2009年度卒業論文），藤村 武史さん（2008年度卒業論文）をもとに再構成したものです。

1. アニメーションとOR

アニメーションとORとの間には特に関係はなさそうに思えるかもしれませんが，現実世界に存在するさまざまな問題を解決するための実学であるORは，アニメーションの制作にも「役立てる」ことができます。

たとえば，楽器演奏シーン，特にピアノ演奏シーンが含まれるアニメーションの制作です。実は，作画に膨大な労力と正確性が必要なアニメーションでは，ピアノ演奏シーンは実質的に制作不可能でした。しかし数年前，クラシック音楽がテーマのコミック・映画・アニメ「のだめカンタービレ」（二ノ宮知子原作）が人気を博し，楽器演奏がストーリー上重要であるアニメーションが注目されたこともあり，実際に楽器演奏シーンを制作することが緊急の課題となっていました。

楽器演奏場面作成に対する一つのアプローチとして，モーションキャプチャシステムを利用して演奏者の3次元的な動きを記録し，それに基づいてCG（コンピュータグラフィクス）を生成するというものがあります。

本稿では，筆者の研究室が関わった，テレビアニメ「のだめカンタービレ」の巴里編（第10話，第11話）およびフィナーレ編（全編）における，楽器演奏CG制作に関する技術の一部を紹介します。

2. モーションキャプチャを用いたCG生成

モーションキャプチャは，対象に取り付けた（一般には多数の）マーカという小球を複数のカメラで撮影

することにより，当該マーカに対応する点の3次元座標を取り込むシステムです。動画の一コマに対応するフレームごとに，各点の3次元座標値を記録し，各点に対応するマーカ名を特定する処理を行います。モーションキャプチャには，マーカを利用せず映像のみに基づく方式もありますが，いずれの方式にも長所・短所があります。

楽器の演奏動作データ収録は，どの方式であれ容易ではありません。特にピアノ演奏では，手指が狭い領域内を高速に動くうえ，重なりなどによって死角が頻繁に発生するため，誤認識や欠落が頻繁に発生します。なお，死角はピアノだけでなくサクソフォンのような管楽器にさえも生じます。このような誤認識や欠落をアニメーターの手作業で修復することは作業量の観点からも非現実的です。

もう一つ，解決すべき課題として，演奏動作データと音楽データの同期があります。一般に，演奏動作データ収録と音楽データ収録は別々に行われます。モーションキャプチャで演奏動作データを収録する際には手指や楽器に付加されるマーカが演奏の妨げとなるため，高い品質の演奏が必要な音楽データの収録を同時に行えないからです。また，音楽データの収録には高品質の楽器が用いられ，電子ピアノなどが使用されることはありません。つまり，発音時刻が得られるMidi形式でデータが収録されることはなく，発音時刻や周波数情報を含まない無圧縮のWave形式データとして収録されます。別々に収録される両データ間のずれは不可避であるため，Wave形式の音楽データにおいて発音時刻を検出し，発音時刻に合わせて動作データを延長・短縮し，発音時刻と同期した演奏動作データを生成することが必要となってきます。

以上のように，モーションキャプチャによって楽器演奏データを収録してCG制作に用いることは，実際にはまったく容易ではなかったのです。

2.1 モーションキャプチャデータの補正

楽器演奏は一般的に手指と腕で行われるため，これらにマーカが付けられます。腕から指先にかけてのマーカの並びの位置関係は変わらず，また隣接したマーカ

みわ ひろよし
関西学院大学 理工学部
〒669-1337 兵庫県三田市学園2-1
miwa@kwansei.ac.jp

間の距離はほとんど変わりません。さらに、手指や腕には可動域や可動速度の制約もあります。また、基本的に楽譜どおりに演奏するため、動きの範囲や方向にも制約があります。

収録された演奏動作データに対するマーカ名の割当は、フレームの集合と各フレーム内の点集合が与えられたとき、マーカ集合が上記の制約を満たしつつ、マーカの移動距離の総和が最小となるように、点集合の各点とマーカ名の対応を決定する最適化問題として扱うことができます。

この最適化問題はかなり複雑ですが、シンプルなヒューリスティックアルゴリズムによって実用的な解を高速に求めることができます。基本的な考え方としては、あるフレームまで点とマーカの対応が特定されているとき、楽譜から予想される次の動き・可動範囲・速度に基づき、それ以前のマーカの動きを外挿することによって次のフレームにおける各点に対応するマーカを推定するというのを繰り返すものです。必ずしも最適解が得られるとは限りませんが、ほぼそのまま利用できる解が高速に得られます。

これにより、これまで現実的には不可能であった、CGによるリアルなピアノ演奏シーンの実現に成功したうえ、それを毎週の放映にさえも十分間に合わせるようにできるようになったのです。

2.2 演奏動作と音楽の同期

Wave形式のデータには、発音されているすべての周波数が合成された値が格納されています。ある時刻においてある周波数の音を検出したとき、その時刻以前に発音された音の残音の信号強度が弱まったものか、その時刻で発音された弱い音そのものか、その時刻で発音された音の倍音かは、各時刻の周波数の値だけでは区別できません。倍音とは、発音した音の周波数の整数倍の周波数をもつ音であり、残音とは、発音後に響くことによって残る音です。Wave形式データをフーリエ変換によって周波数分解すれば、各時刻の周波数と信号強度の情報であるスペクトログラムというものが得られますが、それを用いれば発音時刻が容易に検出できるというわけではないのです。

発音時刻検出問題とは、Wave形式の音楽データと楽譜データ（楽譜の先頭から発音される音を順番に並べたリスト）を入力として、楽譜データの各音の発音時刻を決定する問題です。スペクトログラムから、時刻 t における音 p の信号強度を微分したものを $w_p(t)$ を求めます。発音時刻では信号強度が急増するため、 $w_p(t)$ が大きければその時刻が音 p の発音時刻である可能性

が高いと推定できます。基本的な考え方は、楽譜における音の順序にしたがって $i-1$ 番目までの音の発音時刻が定まっているとき、 i 番目の音の発音可能性が高く、それ以外の音の信号強度が小さくなるような時刻を見つけて i 番目の音の発音時刻とするということを経り返すものです。ただ、音は倍音・残音などに埋もれていることも多いため、常にこれがうまくいくとは限りません。したがって、発音可能性を数値化し、曲全体でその総和を最大化するという考え方を uses。つまり、曲全体で、発音可能性の高い時刻がもっとも多く見つかるように各音の発音時刻を決定するわけです。これは、時間を離散化して動的計画法に基づいたアルゴリズムによって求められます。このようにして、Wave形式データから発音時刻を決定できます。

演奏動作データにおける打鍵時刻の検出は、モーションキャプチャ収録の際、鍵盤にマーカを付加し、マーカが一定距離・一定時間以上沈むところを打鍵時刻とすることで行えます。サクソフォンやチューバのような楽器であっても可動部分の動きを収録することで同様に行えます。次に、モーションキャプチャによって得られた演奏動作データを、打鍵時刻ごとに区切った区間に分けます。対応する音の発音時刻の区間が打鍵時刻の区間より長ければ、等間隔にフレームを挿入し、フレーム間を線形補間によって補間し、逆に短かければ、等間隔にフレームを抜き取ります。このようにして、Wave形式の演奏と演奏動作が同期したアニメーションを生成することができます。

3. まとめ

本稿では、楽器演奏CG制作に現れるORについて紹介しました。ここで述べた技術は実システムへの適用が目的でしたが、最適化問題としての定式化と効率的なアルゴリズムの設計というアプローチが有効でした。実際、ここで述べた技術の一部は、テレビアニメ「のだめカンタービレ」シリーズにも用いられ、大きな成功を収めた本作品に貢献できたのです。CGや音楽といった娯楽に関係する分野においても、オペレーションズ・リサーチのアプローチの有効性が示されたと言えるでしょう。

参考文献

- [1] N. Kugimoto, K. Takai, R. Miyazono, K. Omori, T. Fujimura, S. Furuya, H. Katayose, H. Miwa and N. Nagata, "CG Animation for piano performance," In *Proceedings of ACM SIGGRAPH 2009*, Aug. 3-7, 2009.