

第6章 評価システムの構成

評価システム、あるいは一般に何らかの道具が有効性を発揮するためには、複雑で強力な機能を持つことは必ずしも必要ではない。たとえ単純であっても、それに見合った環境・適用範囲のもとで適切に使用されれば十分有効となりうる。むしろ、あらゆる「役に立つ」道具は単純性と機能性の絶妙なバランスの上に成り立っているといえよう。評価システムについても同様で、限られた機能であっても、使い方次第で十分役に立つものになる。逆にあらゆる機能を盛り込んで汎用性を追及したり、理論的な完璧さに拘泥しても、必ずしも良いシステムができるとは限らない。

いうまでもなく、これはシステムの基礎となる理論的背景を軽視したり、単能システムをアドホックに作ってゆけばよいという意味ではない。話は逆で、ある種の単純さを達成するためには多くの試行錯誤を伴った理論的・実践的な検討が不可欠である。問題としているのは、対象領域についての基礎的研究はもちろん重要であるが、評価の目的や内容を明確にし、システムから得られる結果について、それが持つ意味や適用の限界などを評価の枠組全体の中ではっきりと位置づけることも、そのような検討と同様に評価システムの成否を決める鍵となる点である。特にここで扱う音色の評価のように、背景となる心理的な働きの実態や、音の物理的パラメタとの対応が十分に知られていない対象についての評価システムを作成してゆく上では、これらの点が顕著になる。基礎的な知識が限られている状況のもとでは、評価の目的を明確にし、その中で評価システムの果たす役割を限定しておくことが重要である。

以下本章では評価システムを構成してゆくにあたって、どのような点に留意すべきか、またシステムはどのような部分システムから構成され、それぞれがどのような機能・役割を果たすべきかを検討する。

問題を具体的にするために、以下では楽器（ピアノ）の製品検査や品質改良を例として取り上げる。製品検査の場合、直接的に評価すべき事項は製品としての合否の判定である。しかしそれとともに、音響的にどの点に問題があるのか、それは定量的にどの程度のものか、構造的に楽器のどの部分が原因になっているか、それをどのように修正したらよいかなどを判断できることが望ましい。さらに品質改良や新製品の開発などを目的とする場合には、何らかの価値基準に照らして音色の定量的評価を行ない、改良のための戦略を提示することが必要となる。

このような領域には専門のエキスパートが存在する。したがってシステムの作成にあたっては、そのエキスパートの持つ専門知識や技能をいかに取り込むかが中心的な課題になる。作成されたシステムは、エキスパートの仕事を代行あるいは補助するため、利用に供されることになる。

6. 1 評価システムの役割

6. 1. 1 評価システムの意義

評価システムを用いる意義は、評価の過程を機械的なルーチン・ワークとして形式化する点に求められる。これにより、直接的には作業の負担が量的・質的に軽減され、作業自体も迅速に行えるようになる。特に専門技能を持ったエキスパートの関与の度合を少なくし、熟練度の低い作業員でも遂行できる点は重要である。また十分形式化されていれば、作業員や状況に依らず、安定した一貫性のある評価が可能になる。

評価を形式化することには、より本質的な利点がある。評価の内容や経過が明示的に記録されるので、評価システム自身についての評価—メタ評価—が可能になるからである。評価内容に問題点がある場合、どの項目や処理が不適切であったか、あるいはシステムで考慮されていなかった項目を加える必要があるかなどを検討することにより、システムの適用限界を見極め、改良してゆくための戦略を見いだすことができる。こういった問題点には、実際にシステムを運用してみて初めて発見される性質のものも多い。したがって評価システムの開発は、プロトタイプ版を作成して試験的に運用し、問題点を検討して改良を加え、それを運用するという過程の繰り返しになる。これは場合によってはシステムの構成や背景となる理論自体の全面的な見直しを必要とすることもある。つまり評価システムの開発自体が、対象領域についての知見を得る手段となるわけである。

6. 1. 2 計算機の役割

実際の評価作業を人間が行うか、機械が行うかは、必ずしも本質的な問題ではない。例えばチェックリストを用意し、それに作業員が記入してゆくという形態のシステムも可能である。しかし下に示すような、計算機の持つ様々な機能を活用することにより、評価の経済性や効率性、安定性を著しく高めることができる。

計算機の利点としてまず明らかなのはデータ処理能力である。音響信号を、フーリエ解析等の様々な数理的解析法で処理したり、各種の統計的解析を行なう上で計算機は不可欠である。

次に挙げられるのがデータの収集・記録の能力である。入力データや評価結果などを自動的にデータベース化して記録しておくことにより、必要に応じて検索や分析に利用することができる。これはシステムについてのメタ評価を行うための基礎データとなる。また個々の製品に対する評価だけからはわからないような、時間遷移による、あるいは作業工程における品質のバラツキを発見するための分析にも役に立つ。

また計算機には自動制御の道具としての側面もある。評価に必要な計算処理を自動的に実行することはもちろんだが、周辺機器、例えばオーディオ機器や数値制御可能な機械（今の場合なら後述するような自動打鍵装置）の制御にも、計算機を活用することによって人間には困難な高速、長時間にわたる正確な制御も可能である。計算機制御によってシステムの動作を自動化することにより、人手による介入を大幅に減少させることができ

る。

より高度な計算機の利用形態として、知識工学的手法、すなわちエキスパート・システムなどにおける、推論に基づく処理方式がある。知識工学的手法を用いれば、統計的手法では困難な記号的データの扱い、異種データにわたる総合的な処理や判断が可能になる。またシステムは「ルール」によって制御されるため、処理の規模に比して動作を把握することは容易である。知識工学的手法は評価そのものに対してだけではなく、システム全体の制御にも有効である。さらにシステムの作成や改良自体への利用も考えられる。しかし現在の技術水準においては、これには多くの困難がある。

6. 1. 3 システムの位置づけ

評価の全体的枠組の中で評価システムが果たすべき役割には様々なレベルが考えられる。理想的には、システムが評価の全過程をカバーすることが望ましい。この場合には必要なデータをシステムに与え、結果を得ることで評価は完了する。しかしそのような「完全な」システムは実現が困難であり、現実的には人間（特にエキスパート）が最終的な判断を下すための補助としての役割が考えられる。冒頭で述べたように、そのような補助システムでも、注意深く設計・適用されれば、有効に機能しうる。

補助システムの果たす役割にはいくつかの方向が考えられる。1つは評価の判断材料となる各種のデータを、場合によっては助言を付して提示するものである。この場合、システム自身は直接には評価を下さないか、あるいは部分的な性質（例えば音の減衰率）のみについての評価を与える。それでもデータが集約され、整理された形で提示されれば、楽器音のような生のデータを直接与えられるのに比べて、非エキスパートにも正確な評価を行うことが可能になる。もう1つの方向は、「粗いレベル」の評価を与えることである。これは明らかに合格のもの、明らかに不合格のものはより分け、それ以外のものを人間の判断に委ねるといった方法である。この場合、エキスパートの関与が必要であり、作業自体は変わるわけではないが、評価すべき対象の数量や範囲は大幅に狭められることになる。また当然、この2つの方向を融合したアプローチも可能である。

6. 2 楽器音検査システムの概要

具体的な例として、楽器（ピアノ）の音色について、製品検査を行なうシステムについて考える。この場合、評価の対象となるデータは、2. 2でも述べられているように、実際にピアノで発音された放射音そのものである。そこでシステムを構成する機能単位として、以下のようなものが考えられる。

1. 測定すべき音を発音する、
2. 音響信号を採取する、

3. そこから評価に必要な特徴を抽出する、
4. 与えられた評価尺度に照らして音色の評価を行なう、
5. その結果に基づき、付随的な評価を行なう（例えば欠陥が見だされた場合、原因となる機構を特定するなど）。

ここでは基本方針として、計算機等の機器を活用して処理を自動化することを考える。具体的にシステムがどのような処理を行うかは、システムの役割が評価の全体的枠組の中でどのように位置づけられ、どの程度の機能を実現しているかに応じて決まる。これ自体は、形式化するエキスパートの知識の範囲、入手しうるデータ、考慮される、あるいは抽出可能な特徴などに制約され、相互依存的に決定される。

6. 2. 1 ハードウェア構成

ピアノ音の発音については、均質な打鍵を実現するために自動的な打鍵装置（演奏装置）が必要である。計算機制御の打鍵装置を用いることには、打鍵／解放時間を特定できるという利点もある。これにより、音響信号処理の負担が大きく軽減される。精密さや装置の単純さの点では鍵をソレノイドで直接牽引する方式が優れている。しかしこれにはピアノ自体の機構改変が必要なので、機械的に打鍵を行う方式を用いることになる。いずれにせよ、微妙なタッチを実現するためには相当精密な機構や制御が必要になる。注意しなければいけないのは、制御されるのはあくまで打鍵そのものであって、実際に生じる音ではない点である。放射音を楽器の特性に合わせて調整する必要がある場合などは、入力音からフィードバックして打鍵の制御を行わなければならない。

また単に音だけではなく、弦や駒、響板の振動等、他の要因も測定することが望ましい。しかし手間や必要な機材等は著しく増大することになる。

発音の採取については、デジタル録音技術が高水準にあり、大きな問題はない。しかし2. 2でも述べられているように、録音環境については注意深く設計する必要がある。雑音、特にシステム自体から発するものはできる限り除去しなければならない。無響室の使用も考えられるが、一方で実際の使用環境に応じた残響がある方がよい場合もある。マイクの設置位置や数についても、近接位置におくのか、ある程度離れたところがよいのかなど、十分検討することが必要である。

その他の処理については汎用の計算機でソフトウェア的に実現可能である。全体を1台の計算機で賄うことも不可能ではないが、各装置の制御、全体の制御や通信、計算的処理などに別々の計算機を使用して分散的なシステムにすることが望ましい。処理内容にもよるが、これらはパーソナル・コンピュータやワークステーション程度の能力で十分処理できるだろう。高速フーリエ変換などには専用ハードウェアの使用も考えられる。

このように構成されるシステムは、検査すべきピアノまで含めると全体としてかなり大きなものになる。また技術的・経済的制約などにより実現が困難な点もあり、そのような

場合には人手の介入が必要となる。特に問題となりそうなのは打鍵装置であろう。打鍵を人間が行なうことにする場合、バラツキが生じる可能性が高いので、システムの側で補正処理を行なう必要がある。

6.2.2 システムの機能

システムがどのような機能を実現し、それをどのように使用するかは、評価の全体的枠組の中で慎重に検討しなければならない。また実現しうる機能は上述のように、利用可能な知識や資源、技術によって制約される。

具体的な評価を下すシステムとして、最も簡単なのは前述の「粗いレベル」の評価を行なうものだろう。具体的には、製品として明らかに合格のものは通し、そうでないものは判断を人間に委ねる方法である。この場合、不良品を通すという過誤を避けるために、判断基準は厳しく設定しておく必要がある。このようなシステムには、必ずしも深い、厳密な知識を盛り込む必要はない。例えば評価音と基準音とのマッチングをとり、統計的処理によって一定の偏差内に納まるものは合格とする、という方法が考えられる。ただし単純に波形のマッチングをとるとするのは無謀で、実際には抽出されるいくつかの特徴に基づいたマッチングが必要となる。

単にマッチングをとるだけでは、問題点があってもその内容を指摘することが難しい。そのためには考慮される特徴が、人間が楽器の特性を理解し、修正を加えていく上で有用なものでなければならない。一般に純粹に物理的な尺度ではこれは難しい。したがって尺度の構成にあたっては心理的要因を十分反映したものとしなければならない。また尺度やそこから得られる評価値について、その意味するところを説明するような機能も有用である。

さらに、個別の特徴だけでなく、特徴間にわたる総合的判断も必要である。これには単一音の諸特徴についてのもの（例えば高い倍音成分が多い、減衰が速いという性質は相乗的に「固い音」という性質を導く）、いくつかの音にわたるもの（例えば各音自体は許容範囲内にあっても全体のバランスが悪い、全体としてこもった音である等）、一曲の演奏全体に関するものなど、様々な種類がある。総合的判断は評価の意味を理解する上でも、また評価そのものにおいても必要である。計算機でこのような処理を実現する上では、数値的な統計的手法よりは知識工学的手法が有効と考えられる。

実用的な見地からは音色に即した問題点の指摘だけではなく、機構的に楽器のどの部分に原因があるかがわかることが望ましい。そのためには楽器の各部—例えばアクションの調整、ハンマーの打弦位置、弦や駒、響板の特性、ダンパーの作用、調律の具合などが、音色にどう寄与するかを解明・定式化しておかなければならない。これ自体は音色そのものの研究、特に心理的要因との直接的な関わりはないが、実用的なシステムを作成する上では、このような診断機能を有することは重要である。故障診断一般は、エキスパート・システムの用途として大きなウェイトを占めており、その手法はここに直接適用できる。

最後に楽器の品質改良・開発のための評価について簡単に触れておく。検査の場合との大きな相違点は、評価の基準となる指標（例えば基準音）を設定できない点である（「理

想的な音」のようなものを想定することはできるが、これを具体的に表現することは難しい）。したがって評価は（基準値に照らしての）相対的なものではなく、絶対的な指標によらなければならない、しかもそれは心理的・主観的要因に基づくものになる。評価内容も線形な序列ではなく、半順序的な（つまり単純に優劣をつけ難い）ものになると予想される。また検査の場合、考慮すべき要因は相対的に少なく、目的に応じて人為的に限定することも可能であるのに対し、改良の場合にはそのような絞り込みは困難である。このように、改良志向のシステムは検査用のシステムとは異なった性格を持つことになり、実現も難しくなる（もちろんこれは概念的な話で、後者を楽器開発に役立てることは可能ではある）。

6.2.3 評価用データの選定

検査用のデータの基本となるのが単音である。データの収集は、全鍵について、様々な強弱・長さ、また各ペダルの使用下で行なうことが理想である。しかし実際の評価においては、これらのデータすべてを採取することは必要でなかったり、困難であることもありうる。その場合でも、代表的な音域、奏法を網羅することは必要である。また和音や、音階・アルペジオなどの簡単なシーケンスは音同士の協和性や残響の様子を知るために有用である。

人間が楽器の評価を下す場合、エキスパートでなければこのような個別のデータに即して行うことは困難であり、楽曲の演奏に即した方がまとまった印象を得やすい。また楽曲の演奏には個別データからは得られない情報も含まれている。この点で、機械的な評価システムにおいてもこのような総合的データの利用は十分検討する必要がある。ただ、システムでそのような演奏データを利用するにしても、個別の音に還元して扱ってしまうのでは無意味であり、総合印象的な評価の尺度が確立されている必要がある。

いずれにせよ、エキスパートが検査に際してどのような打鍵を行なっているかは、データを選定する上で大いに参考になる。

6.2.4 特徴の抽出

システムに入力されるデータは物理的な測定値であり、そこから抽出される特徴やその上に構成される尺度は基本的には物理的なものである。心理的尺度については、直接物理的性質に翻訳することが可能なケースもあるし、評価の過程の中でその対応物が構成される場合もある。

音響データから得られる特徴の基本となるのは（定常状態における）音響スペクトルである。ただしスペクトルそのものでは情報量が多過ぎるので、そこからさらに様々な特徴が抽出されることになる。そのようなものの例としては、スペクトル・エンベロープの形状（非数値的に表現されるもの）、最大値をとる周波数、ピーク値の周波数やその比、帯域ごとのエネルギー積分比、非倍音成分の量や分布、構成フォルマントの Q 値などがある。また心理的因子のいくつかはスペクトル・データから近似的に構成できることが知ら

れている⁽¹⁾及び本書の第4章を参照)。

さらにその他の特徴の例としては、周波数エンベロープの形状や減衰率、打鍵解放後(=ダンパー作用時)の残響、立ち上がりなど、非定常時の波形(あるいは経時的なスペクトル変化)、ピーク振幅の時刻や大きさなどが挙げられる。

精密な評価を行うためには、抽出される特徴はできる限り多岐にわたった方がよい。一方処理効率の点からは、代表的なものだけ選択し、不必要なもの、寄与度の低いものは排除する必要がある。また発音の物理的性質を反映した特徴も盛り込むべきである。例えば弦の振動には垂直方向、水平方向の成分があり、鍵の複数弦の間には駒を介した共鳴的な振動がある⁽²⁾。このような性質に対応した特徴は、評価の上からも有用である。

注意しなければいけないのは、どのような特徴を抽出するかという選択自体がすでに評価という側面を有している点である。与えられた特徴のセットがシステムの視野を限定するのであり、考慮されていない特徴についてはシステムは評価を下すことはできない。この点はシステム作成時や拡張性を考慮する場合、十分留意しておく必要がある。

6.2.5 評価尺度の構成

具体的にどのような評価尺度を構成すべきかについては他章で論じられているので、ここでは一般的な事項にのみ触れる。

構成にあたって大きな分岐点となるのは、参照対象となる基準音(標準音;あるいはそれを特徴的に記述したもの)が存在するか否かである。すでに触れてきたように、この場合はかなり機械的な手順によってもある程度の評価尺度を構成することは可能である。基本的な方針は、考慮すべき特徴について、基準音での値を求め、それに対する偏差の許容域を設定することである。この方法は均等な品質が求められる量産品の評価などでは有効となる。ここでのエキスパートの役割は、主に基準音を選択することであり、特徴の選択や許容域の設定にも関与することが考えられる。

この場合、鍵となるのはどのような特徴を選択するかである。しかし心理的要因は必ずしも明示的な形で考慮する必要はない。純粹に客観的(=物理的)尺度に基づいた評価でも、結果において主観的な評価と一致させることは可能である。そのような例として、Mal'ginaらによるバイオリンの(単一音の)評価システムがある⁽³⁾。ここでは前項で述べたようなスペクトルの諸特徴について、チェックリスト的な評価を行なうが、心理的要因との対応は明示的には現れていない。複数のエキスパートによる主観的評価の結果と比較してみると、個々のバイオリン(50丁)についてはバラツキが目だつものの、3つ程度のグループへの分類に関してはよい一致が得られた。基準音としてはエキスパートから最も高い評価を受けたStradivariusが選択され、それに基づいて評価尺度が構成されている。

このような方法は基本的には分析的方法であるため、総合的な評価には別の枠組を導入する必要があり、問題が複雑になる。機械的な操作でこれを行なうことも可能ではある。上のMal'ginaのシステムでは、単に各特徴の相対偏差を合計するだけで総合評価を導いている。しかしより精密で広範な評価を行なうためにはこのような単純なやり方では限界が

あり、心理的要因に即した、より深い知識の導入が必要になろう。

より大きな問題点は、評価の理由づけや正当性のチェックが困難な点にある。評価結果はあくまで基準音に対する近さという相対的なものであり、音自体の善し悪しではない（善悪の判断は基準音の選定に集約されている）。したがって評価の理由を（人間に理解しやすい形で）与えることは難しい。また評価の「正しさ」は、エキスパートの判定と一致するという、偶然的な、結果論的な形でしか保証されない。したがってシステムの使用は、作成時に考慮された範囲に注意深く限定されなければならない。

一方、このような基準音を設定しない（できない）場合には、評価基準は具体的な音を離れたメタ的なものとなる。これは基本的に心理的尺度が依り所となる。必要な心理的尺度が与えられていると仮定して（実はこれ自身、大問題ではあるが）、問題となるのは、これをどうシステムに組み込むかである。比較的単純なものには、直接物理的特徴によって数理的に記述可能なものもある。しかし一般の場合、特に音色表現語に近いレベルでは、多くの物理的要因が複雑に絡み合っている。これを数式の形で表わすことは難しく、計算機プログラム、とりわけ知識工学的なルールの形で記述し、推論によって導くことが方法としては考えられる。さらに総合印象のようなものでは、状況依存的要因や過去の記憶に基づく要因なども加わってくることになり、問題ははるかに複雑になる。このような要因まで考慮してシステムを作成することは、実質的に利用者の聴取・嗜好のモデルを組み込むことに相当する。

心理的尺度は、（心理的に表現された）音色の性質を表すものであり、それ自体は価値判断に関してはニュートラルであるのが普通である。したがってそれらを総合して価値判断を行なう処理が別個に必要となる。しかし一方、これらの尺度で得られる表現は人間にとっても理解しやすいので、判断そのものは人間に委ねてしまうことも考えられる。

こちらの方法の長所・短所はちょうど最初の（基準音を設定する）方法の裏返しになる。まずシステムの構造や動作は、人間の心理的レベルでの記述に関する限り、理解がしやすくなり、修正や検証も容易になる。またシステムの機能には一般性があり、より広い範囲への適用、あるいは全く別の分野への応用も可能である。しかしシステム自体ははるかに巨大なものになり、その点から保守や理解が困難になる。そもそも現段階では実現することは極めて困難で、音色についての諸研究の進展が待たれる。

6.3 システムの作成過程

6.1で触れたように、評価システムの作成は試行錯誤を伴う段階的な過程となる。これには基礎的な研究や個別の処理ルーチンの開発とともに、エキスパートのもつ知識を利用可能な形に抽出することが重要な作業である。エキスパートから得られる知識は領域固有知識、つまり音色の評価に関わる知識だけでなく、制御的知識、つまり評価作業をどう進めるかという知識も重要である。例えば評価にあたって鍵盤をどのような順番で、どのように打鍵するか、問題がある場合にそれをどう確認し、修正を行なうか、その他、どのような動作をするかなどは、システムの処理を構成してゆく上で参考になる。

このようなシステム作成作業自体をエキスパート・システムによって管理する、あるいは

は自動的に行なうことも考えられる。楽器音評価に関しても、そのような構想(?)が報告されている⁽⁴⁾。学習によるシステムの自己形成が強調される場合、これを適応的システムと呼ぶ。適応的システムのラディカルな形は、評価用のデータとそれに対するエキスパートの評価を与えられ、そこから自動的に必要な特徴の同定やその抽出法、評価過程を帰納的に導いてしまうというものである。しかしこれには現実的に多くの困難があり(例えば膨大なデータが必要になる、学習/帰納的推論の研究自体が十分進んでいない等)、現段階で実現することは事実上不可能である。

文献

- (1) 音の評価システムに関する調査研究報告書、サウンド技術振興財団,1986.
- (2) Weireich,G.: ピアノの弦の物理学(The Coupled Motions of Piano Strings),サイエンス3月号,1979.
- (3) Mal'gina,L.A. and Porvenkov, V.G.: Appraisal of the quality of musical instruments by the principles of pattern-recognition theory, Sov.Phys. Acoust., vol.27,no.5,1981,428-433 .
- (4) Haller,K.: Anwendung eines Expertensystems zur Modellierung akustischer Signalverarbeitungsprozesse fuer die Qualitaetsdiagnose von Musikinstrumenten,Nachrichtentech.,Elektron,1987,37,27-29.