

立石賞功績賞の受賞記念講演概要

サンプル値制御からデジタル信号処理へ

—— 人にとって自然な情報処理とは ——

京都大学名誉教授 山 本 裕

立石科学技術振興財団より 1992 年度に「デジタル制御系の最適化・知能化の研究」というテーマで研究助成を頂きました。それ以後主としてデジタル制御の手法を用いて、人間にとってより自然な信号の処理方法、ことに音響と画像の処理について研究して参りましたが、このたび立石賞功績賞をいただくことになりました。立石科学技術振興財団他関係者の皆様に篤く御礼申し上げます。

デジタルによる音響処理

1982 年、ソニーとフィリップスの共同開発による、コンパクトディスク (CD) が導入、発売されました。アナログレコードに取って代わる初めての商用デジタル音楽メディアとして、CD は様々な規格を導入し、その後のデジタル信号処理の規格に大きな影響を与えたといえます。代表的なものとして、サンプリング周波数 44.1 kHz、記録ダイナミックレンジ 16 bit が上げられます。

このうち、サンプリング周波数 44.1 kHz というのは、1 秒間に 44,100 回、音の信号を計測し記録するということです。詳しいことは省きますが、Shannon が文献 [1] においてサンプリング定理に基づいてデジタル信号の伝送を問題にして以来、サンプリング周波数の半分の周波数、いわゆるナイキスト周波数、が伝送できるデータの高域（この場合は高音）の限界とされてきました。以後、この理論、仕組みは広くデジタル機器に採用され、いわば事実上のスタンダードとなっています。お手持ちのデジタル音響機器の殆どはこれによっていると

言って過言ではありません。

CD のその後

このデジタルメディアの登場によって、音楽メディアは一挙にデジタル処理に方向を変え、それまでのアナログレコードに急速に取って代わるようになりました。

ただし CD の音質、品質については、取り扱いやすさ、音の安定性（音程の揺れのなさなど）などの肯定的な評価の反面、音が硬い、金属的、あるいはシャープにすぎる、冷たい、雰囲気や質感に乏しい等といった批判的あるいは否定的な意見も主としてオーディオマニアと言われる人々から数多く寄せられことになりました。本来目指したはずの音楽の感動を伝えるのに適していないのではないか、その原因は何かといった議論も数多く聞かれました。

そもそもこの規格は人間の聴くことのできる周波数の上限が 20 kHz であるという知見に基づいており、そこにサンプリング定理を持ち込んで、その倍の 40 kHz、余裕を見てその 10% 増しの 44.1 kHz をサンプリング周波数に選んだところから始まります。しかしその後人間の聴くことのできる周波数上限は本当はどこまであるのか、上記サンプリング周波数は妥当だったのかという疑問が提出されるようになります。通常の実験ですと一定周波数の正弦波を被験者に聴かせて聞こえたかどうかを答えさせるのですが、これですと定常状態におけるいわゆる定



常応答のみの試験になってしまいます。一方音の立ち上がり特性を反映する過渡応答についてはいろいろな意見があり、一定していません。一つには過渡応答は瞬時に消え去るものであり、その一対同時比較は困難であることが挙げられるでしょう。また 20 kHz 以下の超高域が人にとって心地良いかどうかについても、大橋氏の研究 [2] などがあり、その効果の一部が解明されつつあると思えますが、詳しいことは未だ明らかにはなっておらず、諸説あります。一つには音と音響の評価というものは、究極は聴取者の脳の中にしかないものであり、客観的なデータとして表すことが困難であるという事情が影響しているといえるでしょう。この点は画像などに比べて随分困難なところ です。

しかし超高域が実際に「聴こえる」のかどうかはひとまず措くとしても、現実の音、自然音は 20 kHz よりはるかに高いところまで周波数成分を持っていることが多いのは事実です。図 1 はあるアナログレコードの周波数成分を取り出してみたものですが、20 kHz 以上の高域まで分布しているらしいことがわかります。

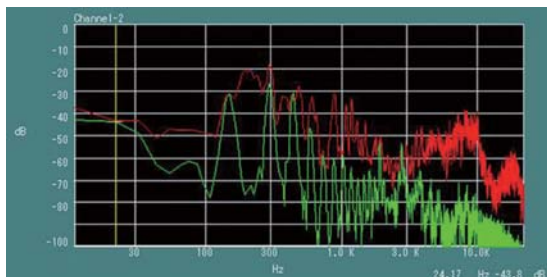


図1 アナログレコードの周波数特性例

現実の CD の録音などでは、この超高域成分があるとエイリアスという歪の原因になる（サンプリング定理に基づく）というので、この部分をすべてフィルターでカットして録音します。その結果どうなるかという、次の図 2 のような副作用が発生するのです。これはフーリエ解析という分野で Gibbs 現象として知られているものですが、周波数成分を急峻にカットすると、

このように急峻なジャンプがある場合、必ずその周りにオーバーシュートやアンダーシュートによる波打ち現象が発生します。

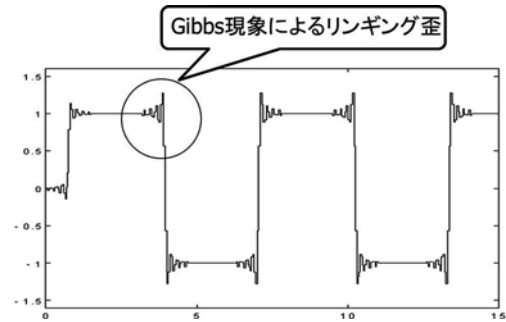


図2 Gibbs 現象による歪

音の場合には鐘のような音がするというので、リングング歪とも言われます。現実にはこのような高周波や不連続な信号は発生しないではないかという意見もあるかもしれませんが、そうではありません。例えば無音から急激に立ち上がるピアノのアタック、また画像では暗い部分から急に明るい部分になる境界部分はこのような信号となり、この歪が常に発生するのです。残念ながら、この部分はこれまで余りまともに議論されてきていませんが、CD の音が金属的であるとか、冷たいとか言う評判はこれが大きく関わっていると思われます。

自然な周波数復元に向けて

超高域が人間に聞こえるにせよ、そうでないにせよ、現在のデジタル音響処理が自然音の周波数特性を正しく反映していないのは確かです。それならそれをなるべく元に近く戻してやれば良い。最近の SACD フォーマットなどは、それを目指してサンプリング周波数を 44.1 kHz より高いところに設定し、超高域までをカバーするようになっています。しかしすでに CD フォーマットで録音されたものではそうは行きません。高周波の情報はサンプル点間の情報として、録音されずに漏れ落ちてしまっているのです。

ではその失われてしまったサンプル点間の信号をどのように復元するかということになります。もしも周波数成分がナイキスト周波数以下にしか分布していない仮定すると、サンプリング定理により、可能な限りの低周波信号で復元することが唯一正しい解であることになります。Shannon がこの論文を書いた時、元の信号の周波数分布に対して、サンプリング周波数、したがってナイキスト周波数も十分に大きく、余裕を持っている状況を想定していたと思われます。そうであれば、Shannon によるこの考え方はほぼ間違いのない結果を与えますし、実装上の問題を別にすれば（Shannon のこの方式はそのまま実装するには多くの困難があります）、完結した結果と言っても差し支え無いでしょう。

しかし現実にはしばしば我々のそうした期待を裏切るもので、多くの音楽信号ではその周波数分布は 20 kHz を超えて分布しますし、よしんば 20 kHz 以下にしかエネルギーがなかったとしても、ナイキスト周波数はその 10% 上の 22 kHz でしかありません。とても余裕を持ってサンプリング定理を適用できる、あるいは Shannon のパラダイムを実現できる状況ではないのです。

それではどうすれば良いのでしょうか。サンプル点間にどんな信号があったかを録音した後から知ることは出来ません。従って、Shannon はナイキスト周波数以下に分布すると仮定したのでした。しかし我々は実はもう少し詳細な情報を持っているのです。

次の図 3 を見てみましょう。この図 3 のような正弦波が与えられ、かつそれが丸印で与えられたような間隔でサンプルされたとします。しかしこうして得られたサンプル信号はもはや元の高周波の信号要素とは独立で、実際この図 3 の下図のようなよりゆっくりとした正弦波とも同調してしまいます。これがエイリアシングという現象で、だからこそ Shannon は周波数成分を限って、この上のような高周波が信号成分

に入っていないような条件を設定せよと言ったのでした。

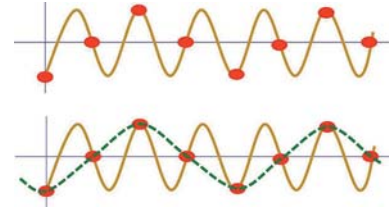


図 3 エイリアシング現象

しかしここでもし、この低周波が 80%、高周波が 20% の割合で混じっていると先験的に仮定出来たらどうでしょう。低周波のみの復元を行うのではなく、図 4 の一番下の信号が正しいということになるのではないのでしょうか。

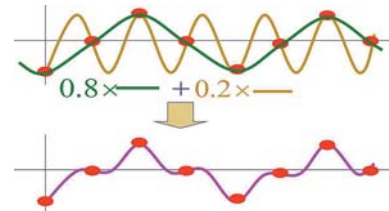


図 4 先見情報を仮定した信号復元

無論現実には与えられた信号に対してこのような情報がわかることはあり得ません。しかし平均的に、あるいは蓋然的には周波数特性というものは高域に向かって徐々に減衰していくもので、殆どの場合次の図 5 のような特性になっているものと考えられます。

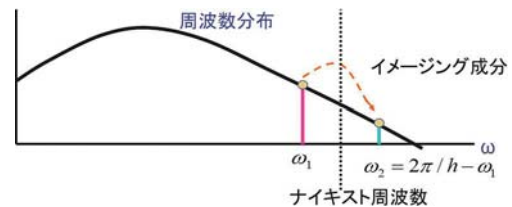


図 5 信号の周波数分布

であるならば、この周波数特性を参考にして、ナイキスト周波数以上の信号成分を上手く加えあわせてやれば、20 kHz 以上まで特性の伸びた信号復元ができるのではないかと想像されま

す。実際その通りなのですが、話はそう簡単ではありませんでした。というのは、サンプル点間をこのような周波数特性に従って、高域まで伸ばす、あるいは補間するという理論はそれまでになかったからなのです。それこそが、Shannon 以来 60 年以上の永きにわたって帯域制限された信号復元が続けられてきた理由でもあります。それをどうするかについては、少しさかのぼってデジタル制御のお話をせねばなりません。

デジタル制御またはサンプル値制御の基本問題

私達が制御したいプラントあるいは信号系は多くは連続な時間で動作します。通常これらは同じ連続時間で動作するコントローラあるいはフィルタなどで処理されてきました。これに対し、コントローラとセンサが離散的な時間で動作するものをサンプル値制御系、あるいはデジタル制御系と呼びます。これらがデジタル計算機で行われることが多いためです。このサンプル値制御系は旧くからあった概念なのですが、困った問題を抱えていました。すなわち、連続と離散の二つの時間要素を内包しているということです。この時間要素の混在は数学的取り扱いを非常に厄介な、見通しの悪いものにしていました。そのため、サンプルされた時刻の間にプラントがどのような応答をするかは、あまりうまく扱うことが出来なかった、少なくとも設計の時にはそれを上手く設計に取り込むことが出来なかったのです。それでサンプル時間を短く取れば良いだろうということになっていました。ところが実はサンプル点間を無視して設計すると、サンプル時間が短くなればなるほど応答が悪くなる例もあるのです。例えば次の図 6 です。ここで青の実線で表したのが、離散時間設計によるものですが、サンプル点の間で 0 に収束して欲しい応答が大きなリップル（波打ち）現象を起こしています。（赤の実線については後述。）

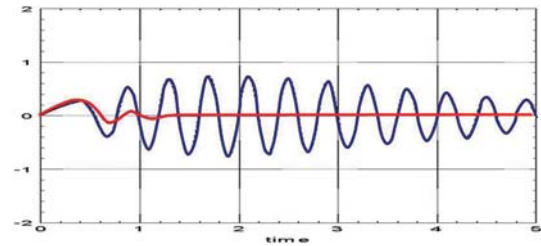


図 6 離散時間設計の応答例

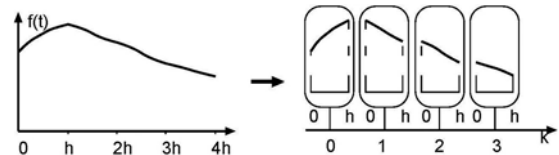


図 7 リフティング

これはそもそも、サンプリングという操作によって、連続時間の信号を落としてしまったことに原因がある。だったらせめて設計用のモデルだけでも、サンプル点間応答を落とさないようにしましょう。そう考えて提唱したのが、図 7 に示す現在リフティングと呼ばれている手法です。

専門的になるので詳しい話は省略しますが、一口に言って連続時間の信号 $f(t)$ を離散時間の関数の列 $\{f(kh)\}_{k=1}^{\infty}$ として離散時間で扱えるようにしたのがその核となるアイデアです。今となってはコロンブスの卵なのですが、当時私がこれを導入した時は、わざわざ問題を難しくしている、と非難を浴びたものでした。幸い、このアイデアが認められ、IEEE の Control Systems Society の論文賞もいただき、いまではこの業界の標準手法として確立しています。

現代的サンプル値制御系の成立

このリフティングの手法の導入とともに、現代的なサンプル値制御理論が瞬く間に展開し、成立していきました。1990 年代前半のことです。しかも同じサンプリング周期を用いていても、サンプル点間応答を最適化するコントローラを得ることができるようになりました。図 6 において赤の実線で表示したのがそうして得られた最適コントローラによる応答です。また矩形波に対して、図 2 とは対照的に次の図 8 の

ようなリングングのない応答が得られることも分かっています。

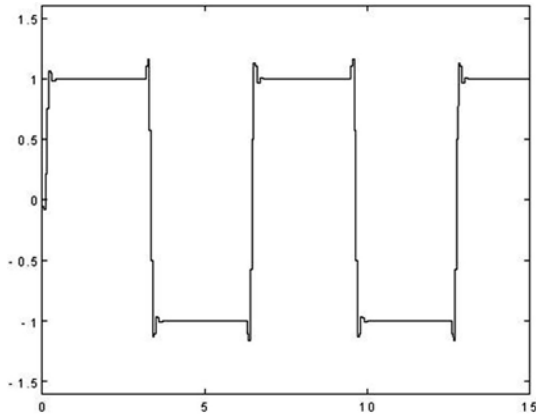


図8 サンプル値最適設計による矩形波応答

サンプル値制御から信号処理へ

上に述べた信号処理の問題、すなわちサンプル点間の（失われた）信号をどうやって復元するか、という問題に戻ります。上で述べたように、サンプル値制御理論ではサンプル点間応答を最適に補間できるということがわかりました。これを信号処理に応用するとどうなるでしょうか。これまでのデジタル信号処理の理論で、使われていなかった仮定あるいは理論は、おおまかに言って2つあります。一つはリフティングによる現代的な制御理論です。これは制御の部門ではもはやスタンダードとなっているにも関わらず、残念ながらデジタル信号処理の分野では、まだわずかししか知られておらず、理論としては広く用いられているとはいえないのが現状です。2012年に文献[3]を発表し、新しい方法論を提起しましたが、これが浸透していくにはまだ後10年単位の時間がかかるでしょう。

もう一つ、Shannonが想定しなかったものに、元のアナログ信号の生成モデルがあります。実はここにこそ、サンプル値制御理論がShannonのパラダイムに対して優位性を主張できる根拠があるのです。対象が音楽であれば、その音には必ず発生源があり、それらは物理的

特性に基づく特有の周波数減衰パターンを持っているのです。ギターのような楽器を考えてみてください。例えば基音として440 Hz, Aの音を鳴らしたとしても、それに付随してその倍音、3倍高調波があり、ギターの特性に従って高周波が減衰していきます。（これ以外にも過渡特性というものがありますが、ひとまず無視します。）この特性を反映した減衰曲線を物理モデルとしてモデル化し、サンプル値制御系におけるプラントの役割をさせることによって、平均的に最適な高周波の復元ができるのではないかと考えたのが始まりでした。このフィルタによって、CDでは20 kHz以上のはるか上の高周波まで復元できるようになりました。そのための実験装置として作成したのが、図9に示す375 kHzまでの再生能力をもったDAコンバータです。



図9 試作 DA コンバータ

これと前後して、この新しいアイデアの商品化、普及に取り組みました。しかしながら多くのオーディオメーカーからの反応は芳しいものではありませんでした。まだその当時図9のようなプロトタイプが出来上がっていなかったこともありますが、一方では、信号処理の教科書をそのまま鵜呑みにして最初から結論ありきという拒否反応にも悩まされておりました。そうした中で、三洋電機（当時）からのアプローチがあり、当初は画像処理への応用からスタートしたのですが、まもなく圧縮オーディオの高品位化に目標を転換し、幸い3年の検討期間を経てチップに組み込んでいただけることになり

ました。この間、研究所におられた藤山晃司氏と二人三脚で開発に従事し、時には社内の反対に遭いながら、最終的に LSI チップとして製品化できることになったのは、藤山氏の協力無くしては考えられないことでした。深く感謝している次第です。幸いこれまでに 6,500 万石以上を出荷する成功を納めています。

おわりに

サンプル値制御理論を用いたこの新しい信号処理理論とその応用は、まだその端緒に終わったばかりともいえます。我々の処理する信号の結果は、畢竟アナログ領域で評価されるものであり、処理結果がどれだけ原信号に近いかが評価されねばならないでしょう。デジタルだからといって、安易にデータを間引いてしまうと後で問題を抱えることにもなりかねません。しかし一方では、ますます負荷の大きくなる処理に対して、データの圧縮は不可避であり、そこにこそアナログ特性を最適とするこのサンプル値処理方式の将来の発展があるべきでしょう。既

存の定説などに左右されることなく、信号処理のスタンダードが、ちょうど現代的サンプル値制御理論が旧来のサンプル値理論を書き換えたように、この新しい方式に書き換わる日を待ち望んでやみません。この新しい信号処理方式の研究を通じて、協力いただいた上記藤山氏を始め、永原正章氏、P. P. Khargonekar 氏、その他多くの大学院学生諸氏、並びに関係の各位に御礼申し上げます。

文 献

- [1] C. E. Shannon, "Communication in the Presence of Noise," *Proc. IRE*, vol. 37, pp. 10-21, 1949.
- [2] T. Oohashi et al., "Inaudible high-frequency sounds affect brain activity: hypersonic effect," *Proc. Amer. Physiological Soc.*, pp. 3548-3558, 2000.
- [3] Y. Yamamoto, M. Nagahara & P. P. Khargonekar, "Signal reconstruction via sampled-data control theory-beyond the Shannon paradigm," *IEEE Trans. Signal Processing*, vol. SP-60, pp. 613-625, 2012.