

デジタル音声データの話

2016年12月7日
中田宏

音声データの大まかな種別

名称	特徴
PCM	<p>音の振幅を数値化したもの。1サンプルあたりのビット長は8,16,24,32bitなど様々なバリエーションが存在する。波形の時間変化を表現するために、一定間隔で数値を並べる。</p> <p>CD(Compact Disc)に採用されている。</p> <p>PCMのデータ量を圧縮する手法が複数存在する。圧縮から復元した時に元のPCMを完全に復元できるかどうかで</p> <ul style="list-style-type: none">• ロッシー圧縮(完全な復元は不可能)• ロスレス圧縮(完全復元が可能) <p>の2つに分類される。</p>
$\Delta\Sigma$	<p>1サンプルあたりのビット長を1(もしくはPCMよりも少ないビット長)に制限してサンプリング周期を短くした形式。技術の詳細説明は、CQ出版社『エレキ工房No.5』に記事を書いたので参照されたい。</p> <p>データの演算は面倒だが、D/A変換、A/D変換のハードウェア設計が容易なため2016年現在応用が進んでいる。</p>
DSD	<p>$\Delta\Sigma$原理に基づいたデータをSonyが商品化するときに登録した商標。</p> <p>SACD(Super Audio CD)に採用されている。</p> <p>今回は、$\Delta\Sigma$とDSDに違いがあるかどうかについて、取り上げる。</p> <p>本資料の表記では、$\Delta\Sigma$とDSDを厳密に書き分けるので、注意して欲しい。</p>

ファイル構造一般の説明

- 音声ファイルは、単純にデータが並んでいるわけではない。イマドキの音声ファイルは、以下の3段階構造を持っている。
 - 1.コンテナ(1ないし複数のコーデックデータを含む)
 - 2.コーデック(データ形式の種別を意味する)
 - 3.データ(単なるビットデータの集まり)
- 一つのコンテナの中に音声データと動画データ両方のコーデックを含むこともある。(例 MPEG2)
- 音声ファイル構造の詳細は、今まで何度もCQ出版社へ原稿を提出済み。中田の著者名で活字になった原稿もあれば、「没になった」と担当編集者に言われた後に匿名別人の記事として掲載された原稿もある。「この件の詳細調査はしないでくれ」とトラ技編集長に言われている。

『 $\Delta\Sigma$ =DSD』ではない！

疑問発生

- SACDを再生するたびに気になっていたことがあった。
『規格が公開されていないDSDは、 $\Delta\Sigma$ 変調に追加してなんらかの縛りがあるのではないか？』と予想していた。
- 『手のひら返しのTさん』のブログに書かれた内容で「1ビットオーディオ研究会が配布しているWSDファイルのDSDデータは変調率が100%を超えている」という指摘が気になっていた。当時気になった点を具体的に示すと、以下になる。
 - DSDには変調率があるのかもしれないけれど、WSDファイルに格納しているのはDSDではなくて $\Delta\Sigma$ なので、変調率は定義されていないのではないか？
 - Tさんを含めて皆が『DSDの変調率』と呼んでいるのはオーディオゲートで $\Delta\Sigma$ を再生した時のメーター表示ではないか？
 - オーディオゲートのメーター表示は、バージョンアップの時に挙動が変化しているのではないか？

『 $\Delta\Sigma$ =DSD』ではない！ 関係者にインタビュー

- 『第11回1ビット研究会』で登壇したNさんがDSDに詳しくだったので、発表後に質問してみた。
- 中田の質問1
初期のCDでは録音レベルを低くおさえていた様に、DSDもしくはSACDの規格が最大振幅を制限していないか？
- Nさんの返答
そんなことはない。SACDの録音レベルが低いと感じたのであれば、マスタリングエンジニアの趣味だろう。
- 中田の質問2
DSDやSACDの仕様書は主要なものが無料で公開されていて誰でも入手できるのか？
- Nさんの返答
その通り。
- 中田の質問3
DSD= $\Delta\Sigma$ と考えてよいか？
- Nさんの返答
Yes

『 $\Delta\Sigma$ =DSD』ではない！ 非公開の仕様書を入手

- 2016年夏に某氏から非公開の仕様書『Super Audio CD Audio Signal Properties (Including Annex D&E) Version 1.3 March 2003』をもらった。
- 上記資料のAnnex Dに重要な記述をみつけた。以下、英語で書かれた文章を中田が要約する。
 - 項目D.1
SACDのオーディオ・レベルは、バターワース特性LPF(カットオフ50kHz、-30dB/Oct)の後で測ること
 - 項目D.2
SACDの0dBは、DSDの0dB (訳注 項目D.3の最大振幅を示すものと思われる)の50%とする
 - 項目D.3
DSDの変調において、連続する28個のデータ中に同じデータを25個以上含んではない(0データも1データも4個以上24個以下)

『 $\Delta\Sigma$ =DSD』ではない！ 仕様を知ったのでさらに疑問発生

- 第11回登壇者への質問で返ってきたコメントが間違っていたのは、オーディオ業界でよくあることとして不問にする。
- 以下の疑問が発生
 - 前記資料項目D.3の縛り(連続28サンプル中同じデータは上限24個)を真面目に実現すると、信号周波数が低域になるほど、格納できる振幅が小さくなるのでは？
 - DSD信号中のf特がフラットでないならば、A/D、D/A変換の時に補正する必要があるのでは？
 - f特フラットで仕様を守ると、低音表現のために最大振幅を小さくしなければならない。以前SACDプレーヤーのカタログに書いてあったS/N比140dBを実現できないのでは？
 - 「SACDシステムとしてはS/N比140dBを保証しないが、要素技術としての理論値上限が140dBと計算できる」などと言い訳して逃げるのかもしれない
 - 低域をブーストする目的で、 $\Delta\Sigma$ 信号のビットを並べ替えて項目D.3を無理矢理クリアする人がでてくるのではないか？データをいじるという点では、CD音源を作る時のコンプレッションみたいな発想。
 - 『2.8MHzfsで28サンプル中同一データは24個まで』の縛りは5.6MHzfsになると『56サンプル中48個まで』と緩和されるのだろうか？そうでないと、サンプリング周波数を上げて、録音レベルを下げることになる。

1ビット信号を格納する ファイルフォーマット

ファイル形式	格納するデータ	説明	仕様のライセンス
WSD	1bit $\Delta\Sigma$	みなさんご存知の1ビットオーディオ研究会が制定した、ファイルフォーマット。	無料で公開
DSDIFF	DSD	Philips/Sonyが策定したDSD格納用のファイルフォーマット。SACDのオーサリングや、DSDデータの編集に使うらしく、再生には必要ないメタデータがいろいろ格納できる。 DSDデータのロスレス圧縮も格納可能。	無料でネットから入手できるが、本来秘密だったものが流出したのかもしれない
DSF	DSD	Sonyが策定したDSD格納用のファイルフォーマット。 中田は、Sonyと秘密保持契約を結んで仕様を入手した。その後公開されたと聞くと、Sonyの公開の仕方が中途半端なので、本当に公開する気があるのか若干不透明。	非公開から公開に変更したとWEBに書かれているものの、肝心の公開情報へのリンクが切れていた

DSF仕様のライセンスについて

- 中田がSonyと秘密保持契約を交わして仕様書を取得したのは、2010年9月1日
- 2010年9月6日には、Sony宛に「DSF仕様に関わる仕事を別会社から受注する場合に秘密保持契約はどうか？」と質問したが、2016年12月に至るも返答無し。問い合わせ窓口に「返答を確約するものではない」旨が書かれていたので、想定内の対応。
- 中田が秘密保持契約に同意した時、「将来ライセンスの変更があった時に連絡するため、メールアドレスを知らせろ」とSonyに言われて教えた。本当にDSFの仕様が公開されたのであれば、メールで通知してこないことはSonyによる契約違反（個人情報保護法違反も？）になる。いかにもSonyらしい対応ではあるけれど。

ソフトウェアライセンスの話

原理原則論

- 仕様を制定した会社が、仕様を秘密にしてNDA(秘密保持契約)を結んだ相手にだけ提供するのはよくある話。
- フリーソフトのGPLと秘密保持契約のあるライセンスは同居できない。
 - GPL系列には、GPL/LGPLの2種類が存在する。それぞれバージョンアップがあったが、以前のバージョンを採用してもかまわない。
 - GPLライセンスのソフトウェアとリンク(スタティックリンク/ダイナミックリンクのどちらでも)する全てのソースコードは公開の義務がある
 - LGPLライセンスのソフトウェアとスタティックリンクする全てのソースコードは公開の義務がある(ダイナミックリンクならば非公開でもよし)
 - LGPLとダイナミックリンクするソフトウェアは、バイナリのリバース・エンジニアリングを許可しなければならない
 - 勉強のためにGPLライセンスのソースコードを読んで作成したソースコードにも、GPLの縛りが適用される。
 - LinuxカーネルのソースコードはGPLライセンス部分があるので、原則としてデバイスドライバを含む全てのソースコードはGPLで公開の義務があるはず(リナス・トーバルは別意見らしいが、中田などのGPL解釈と異なる→複数解釈ができるということは、裁判に持ち込まれたら危うい)
 - フリーソフトのソースコードを自分のソフトウェアに組み込む際は、ライセンスの衝突に気をつけなければならない。秘密保持義務のあるソースコードとGPLソースコードを間違えて組み合わせてバイナリ公開してしまうと、秘密保持義務のあるソースコードも公開する義務が発生し、損害賠償問題に発展する。LGPLと組み合わせる場合でも、リバース・エンジニアリングを許可できるかどうか検討が必要。
- 中田が作成する音声再生アプリケーションでは、 $\Delta\Sigma$ 再生部分とPCM再生部分を明確に分けている。 $\Delta\Sigma$ 再生部分にはDSFのライセンスが絡むし、圧縮PCMの展開にはLGPLライセンスのffmpegライブラリを使用しているため、両者を同居させたくないから。

その意味でもDSFのライセンスが公開となったのであれば、「以前の秘密保持契約が全部無効である」旨をSonyに宣言してもらいたい。そうしないとDSFを扱うソースコードだけいつまでも特別扱いになる。

ライセンスの話

現実編 その1

- GPLをなめてはいけない話
 - ソースコード非公開の市販ソフトウェアにGPLライセンスのソースコードを組み込んでしまったため、後からソースコードを公開するはめになったプロジェクトがいくつかあると聞く。
- 市販の情報家電
 - 中田は某社の手伝いで、オーディオに組み込むLinuxのデバイスドライバやアプリケーションやいろいろ作ったことがある。カーネルのソースコードは、だれも知らないような場所でひっそり公開されていた。しかも一部のカーネルソースコードは公開していないので、GPL違反。その旨会議で指摘したら、次回からライセンス会議に呼ばれなくなった。後からライセンス会議の結論を確認したら、「ライセンス判定は別部署に丸投げしたので、もう知らない」とのこと。社内の連絡不十分を口実に、ライセンス違反を続けているらしい。
- ああ勘違い
 - 2004年頃の話。中田がそれまでGPLで公開していたソースコードをバージョンアップの時に別ライセンスに切り替えたら、「ライセンス違反だ!」と匿名掲示板で炎上した。自作ソフトならばライセンス切替がGPL違反にならないことは、FSFのWEBサイトにも書いてある。日本人のライセンス意識の低さが露呈した。
 - ちなみにFSFのサイトには「GPLを別ライセンスに切り替えると、尊敬を失う」みたいな表記があったが、もともと日本ではフリーソフトの作者は尊敬されていないので関係ない。逆に「苦勞して作ったものをただで配るなんて、バカじゃない?」と言われることの方が多い。そういう人の生活を影で支えるLinuxは、彼らがバカと呼ぶ人たちの集団が作っている。
- フリーソフト公開の反応
 - 中田が様々なフリーソフトを公開するようになって、20年以上経つ。殆どのユーザーはサイレント・マジョリティだと思うが、時々反応が届く。好意的なものとうそでないものがほぼ同数。
 - 好意的なものは感謝の表明が多く、素直に嬉しい。ときどき、バグ報告ももらえてありがたい。
 - 好意的でないものは、先方の要求(機能追加とかソースコード公開とか)を一方的に押し付けてくることが多い。Eメールでとどいた反応は全部保存してある。
 - たまに、フリーメールのアドレスから学生を名乗って技術指導の依頼が届く。中田は将来の日本を背負って立つ若者への協力を惜しまないが、2つの指摘を返すことにしている。
 - 他人に何かを依頼するのであれば、匿名アドレスからのメールは失礼に当たると考える
 - 大学生がプロフェッショナルに依頼をするのであれば、報酬交渉を含めて指導教授を通すのが筋だと考える

ライセンスの話

現実編 その2

- ライセンス・テロの可能性

- ソースコード非公開のソフトウェアをLinuxの追加パッケージとして登録しても、利用するライブラリがLGPLまでならば問題ない。

ただしある日を堺に、利用していたライブラリがLGPL→GPLにライセンス変更されたりしたら、誰かのところへバイナリが届いた瞬間に公開の義務が発生する。こんなライセンス変更を実行するとテロリズムの一種なのでまず起きないと思うが、リスク管理として可能性は考慮すべき。

- WindowsのデバイスドライバがGPLライセンス？

- Windowsの自作デバイスドライバをソースコードで配布して、GPLライセンスだと(LGPLではなく)言っている人を見たことがある。額面通りに解釈すると、デバイスドライバを作成するときに使用するSDK(リンクするライブラリを含むので)やWindows本体のソースコードを公開する義務が発生するが、この場合誰が非公開部分のソースコードを公開するのだろうか？

- ハードウェア屋さんはライセンス問題を過小評価する

- 商品企画の段階で、「商品に組み込むライブラリのライセンス上、企画しているプランの一部は実行できない」と主張してもわかってもらえない。「N社日本事業所のAさんは優しい人だから許可してくれるよ」とか言われてしまう。許可を出す権限はヨーロッパにあるN社のライブラリ作成チームが持っていたので、商品出荷直前に揉めた。ハードウェアの世界では、買ってきたLSIをほぼ自由に使えるので、ライセンス縛りが実感できないらしい。

- CQ出版の方針

- インタフェース誌の2015年11月号111ページにはライセンスの話が通り一辺書いてあった。ここに書いたような話をさらに取り上げたらどうか？と2015年9月23日に提案したが、スルー。『技術者向けの情報発信』を標榜するなら『便利なソフトウェアがただで使えてラッキー』的なあおり記事よりもよっぽど有用だと思ったので、トラ技編集長に伝えてある。

印刷界の編集者にも、最近話題のDeNAみたいな人が増えたらいい。

WSDフォーマットのあれこれ

- 中田が『Specifications for 1-bit-coding Data File Version 1.0 September 2002』を入手したのが2011年の夏頃。
 - 1ビットオーディオ研究会が公開しているWSDファイルと仕様書を比較して気がついたことがある。メタデータを格納するTextDataエリアについて、仕様書には『Asciiコード(中略)16 進表記:20h~7Eh』と明確に書かれているのに、シフトJISの日本語コードが入っているファイルがあった。
 - 1ビット研究会に所属していた某オーディオ・メーカー経由で質問したものの、返答無し。
 - 山崎先生にも最近直接指摘したような気がする。
 - 改定された仕様では、格納コードに関する記述が削除された。でもヘッダーにVer.1.0と明記しているファイルは、旧仕様が適用されるはず。
 - マルチチャンネルのマイク配置を示すビットパターンを必ずヘッダに格納することになっているが、配布されているファイルでは全ビットがゼロ。
- あとから11MHzfsのデータ・ファイルが一つ追加公開されたが、WSDフォーマットではなくヘッダのついていない生データで、前後に0x00が大量についていた。中田は自分で0x00を削除してヘッダを付け、CQ出版社の録再キットのデバッグに利用した。
 - エレキ工房No.5の付録CD-ROM内容を作っている時に、山崎先生から「うちの11MHzファイルを再生できないプレーヤーがある」と言われて、ヘッダが無いことを指摘した。
- 早稲田大学のお手伝いをしている時、自作の再生アプリケーションとオーディオゲートVer.2.3で挙動が違っていた。調査したら、両方に問題があった。WSDフォーマットは仕様改定で、ヘッダ中のファイルサイズ項目を32bit→64bitに拡張している。拡張の仕方がトリッキーで、他のヘッダデータはBigEndianで統一されているのに、新しいファイルサイズは下位32bit、上位32bitの順で32bitの中はBigEndianという形。中田の再生アプリは、64bitのBigEndianだと解釈していたし、オーディオゲートVer.2.3は、拡張前の32bit(拡張後は下位に相当する部分)だけ読んでいた。