

デジタル エッセイ

「32Kの壁」

坂口 裕靖

2Kが4K、そして8Kになったとして、HDR対応とかハイフレームレート対応といった、今までとは違う次元の拡張があるわけだけでも、基本である空間解像度についてはスケールアップである。なので、この空間解像度の部分は今後いくらかでも増やしていけると思いがちだが、割とそうでもなかったりする。

まず最初に、実在する被写体を可視光線でスキャンする場合、光源波長によるサンプリングレートの上限が存在する。仮にY信号に一番近い、Gチャンネルについての限界を考えてみると、例えば555nmの波長の場合、焦点面上のセンサ密度がこの倍程度以上の間隔でない、単にオーバーサンプリングしているだけであり、解像度的には意味がない。とすると、 $555\text{nm} \times 2 = 1.110\mu\text{m} = 1.11 \mu\text{m} =$

0.00111mm あたりに1pixelが可視光を使う限りの物理的な密度限界ということになる。今、仮に対角線が1インチの16:9センサを考えると、横方向のサイズは0.8716インチ、22.138mmである。これを0.00111で割ると、19,944.2とか。つまり、1インチのセンサでは横20K解像度を超えると、オーバーサンプリングになってしまう。もちろんこの数字はレンズ系による解像度の低下とかを無視しているのだが、一つの目安であることも確かだ。36x24mmのフルサイズセンサだとして、同様の計算をしてみると32,432.4とかであり、この32Kが一つの壁となるのではないだろうか。もちろん今後レンズを使わない光学系も多々出てくるだろうし、そういったニューウェーブによって解像度の上限が突破されることは大変望ましいことで

ある。それに、そもそもエンターテインメントを前提とするのであれば、被写体のスキャンとモーションキャプチャ、バーチャルスタジオ的な組み合わせによって、光学像の制限を超える解像度も当然達成可能であろう。そこまで手間かけるべきかどうかは何とも言えないのだが。

というわけで、実写映像が32Kを上回ることは大変難しそうであることが分かったわけだが、この32Kで16:9という解像度もなかなかやっかいである。仮に控えめに考えて、マルチスペクトル方面には進化しないとしてみよう。その場合、3ch/pixelとなる。さて、それぞれのチャンネルが何ビットかだが、仮にHDRを前提に考えると、少なくとも浮動小数点フォーマットである必要があり、その意味では最低でも16bitあたりが落としどころになるだ

ワンポイント マイクロフロート

単精度浮動小数点数が32bit、倍精度浮動（以下略）が64bit、4倍精度（略）が128bitで、単精度の半分の16bitで表すのが半精度浮動小数点数、略してhalfとかbinary16とか。正の最小値は 5.96×10^{-8} 、正の最大値は65,504で、あんまり大きな数は表せない代わりに、それなりに小さい数を扱うことができる。単精度に比べると消費メモリが半分ですむため、できれば積極的に使っていきたいわけだが、DirectXだと勝手に単精度に変換されてしまい、使えなかったりする。OpenCLでは使える模様。値域が0~1とか、-1~1の係数とかを扱う場合に固定小数点よりは精度が高いので、大変便利。

こういった16bit幅より狭いデータオブジェクトで浮動小数点数を表すものを、一般的にミニフロートと呼ぶらしい。そして、16bitより狭い浮動小数点数オブジェクトのことをマイクロフロートと呼ぶ模様。例えばIBM-PCのキーボードではリピートレートを指定できるのだが、これが5ビットのデータで8

から120まで、32通り指定できるようになっていた。これが単なる整数じゃなくて、符号ゼロ、指数部2ビット、エクセス-3、仮数部3ビットの5ビットなマイクロフロートを使っているのがミソ。実際のリピートレートはマイクロフロートの値を4で割った値で、2.0、2.1、2.3と等比級数っぽい形になっていて、大変気がきいてるのであった。同じような考え方をすると、符号ゼロ、指数部4ビット、仮数部4ビットの8ビット浮動小数点で、1から507,904まで表すことができるよ、といったことが<http://www.mrob.com/pub/math/floatformats.html>に色々書いてありますな。なんかここいらへんのちまちました浮動小数点数、GPGPU等々の物量計算関係で、意外と重要になってきそうな予感があります。してみると、The Art of Computer ProgrammingのVol.2第4章に浮動小数点数を割っているのはさすがの一言ですな。

ろう。この際アルファは無視して考えると、16bit/channel、3channel/pixel、1920x1080の4:4:4非圧縮画像は約12MB/frameである。これが32Kになると、およそ縦横各16倍、全体で256倍だから、約3GB/frameとなる。アルファチャンネルも同様のビット深度が必要だから、合成系だと4GB/frameだ。これ、32bitな整数じゃ扱いきれないということすな。無理矢理符号無し整数を使ってもギリギリで、ちょっとスクロールするような素材はもうアウト。なので、最低でも64bit整数じゃないといけない。まあ今時64bitじゃないOSはほとんどないので、パソコン界隈では気にする必要もないのだが。それ以前の問題として、最終出力が16bit/channelの場合、おそらく途中の処理は32bit/channelで進めるんじゃないかしら。その場合16GB/frameである。現時点のそこらで売られてるパソコンだと、メインメモリを書き潰しても読み込みきれない。ましてやGPUをや。

さて、空間解像度が高くなっているのだから、当然時間解像度も高くなっているに違いない。仮に現状と同じ30fpsだとしても、3GB/frameの場合は90GB/s。これが120fpsとかになると、単純に4倍で360GB/s。まあ中間ファイルを非圧縮で保存するというのは非現実的かも知れないが、仮に2時間の素材を保持するとして2.6PB。現状手軽に手に入るストレージはたかだかTBクラスなので、これがPBクラスになるようでない、32kの素材をカジュアルに扱うのは難しいだろう。というか、そもそもバスのスピードが3桁あがってくれないと、こんなものまともに扱えるようにはならないだろう。

ちなみに、PNGのフォーマット上は16bit/channel、4channel/pixelまでは許容しているし、画像解像度も4byteで持っているため、縦横各4Gピクセルまで対応可能ということになっている。もっとも、ここという16bit/channelはリニアが前提であるため、半精度浮動小数点を食わせると、多くのビューワで変な見た目になるだろうし、γ補正をしているようなシステムだとぐちゃぐちゃになりそう。そこい

らへんを割り切れれば、データの入出力は可能である。windows bmpの場合も縦横各32bitでサイズを保持している。が、縦方向については「負の数の場合は上下逆」という規約があるため、実質的には1ビット減って31bit、2Gpixまでしか対応していない、と考えることができる。というわけでサイズの方は大丈夫だが、フォーマット上32bit/pixelよりビット深度を深くすることができないので、単純に半精度小数点の画像を保存することは出来ないだろう。無理矢理収容するなら、画像上の1ピクセルをファイル上の2ピクセルに分割して、縦なり横なりに展開して保存するしかない。処理の都合を考えると横に並べるのが良さそうだが、見た目はぐちゃぐちゃになりそう。まあ、とりあえずtiffやOpenEXRなら対応できるので、そこまで無理してbmpを使う必要も無いだろう。というわけで、少なくとも交換フォーマットについては、32kどころか2Gぐらいまでは心配する必要がなさそう。

空間解像度の上限が32kで、現在8kに手が届くところにいるとすると、上限まではあと2世代しかないという話になる。これ以上空間解像度の変化が難しい場合、今後何を指すべきなのだろうか？

映画が映画として成立するには、演出的に自由に時空間を移動できる必要がある。これを実現するのがオーバーラップだったりカットつなぎだったりするわけだが、実はこれが成立するのは「画は画だから重ねちゃえ」「コマに変わらないからつなげちゃえ」という、演出空間内の物理法則持続性を無視した、天才的なヒラメキがあったらこそだ。演出空間上は時空間が重なったり飛んだりしてるのに、普通の映画でオーバーラップに違和感を感じないのは、見つめているスクリーンが平面で、奥行きがない「画」だからである。こ

れが3D映画でオーバーラップをしようとすると、OL前後のカットで同じような奥行き感、視線方向にしておかないと、違和感が盛大に発生する。二つの三次元空間が融合して違和感を持たせないような空間構成は大変難しい。うまくいくと空間が溶けるような、なかなか素晴らしい効果を生むのだが。カットつなぎも同様で、2D映画のつもりでカメラの切り返しを多用すると、目が(というか、脳が)ついていけなくて酔ってしまう。2Dの切り返しについていけるのは、空間の再構築作業が必要ない(というか、平面ということで空間部分の認識処理が発生しない)からだ。というわけなので、映像ストリームのスケールアップが難しいからという理由で奥行き方向、Zの情報を入れることは、テクニカルには可能だろうが運用上は大変難しいのではないだろうか。一方で、ご家庭用カメラはほとんどが撮って出しなので、Zの付加価値は結構ありそうな気がする。

いずれにしろカメラフォーマットは素材を左右するため重要だが、一方で出力の方はネットということになると、そもそもアスペクトすら固定されない状況になるだろう。

この世界では32Kだろうが64Kだろうが、帯域幅が許す限りは表示できるだろうし、逆に言えばフォーマットは全く重要ではなくなる。今後の進展がどうなるのか、楽しみである。

Hiroyasu Sakaguchi
(株)IMAGICA イメージワークス

祝・創刊 200号



4CH HD Live AV Mixer **A-PRO-4**
8 in Multi-format Switcher **MS-8**
VIDEO CONVERTER **AV-5S/AV-3 Series**



ATV株式会社 [営業部] 〒111-0053 東京都台東区浅草橋 2-29-11 マルケービル 9F
<http://www.atvcorporation.com/> TEL:03-5829-4919