

Nvidia の成功に学ぶ(その 5)

2021年3月19日 | [岡山義憲の集積回路の明日に向けて](#)
(情報統合技術研究合同会社)

本ブログでは、先ず、GPU(Graphics Processing Unit)技術に取り組む企業が近年急増していることを示す。GPU技術は、AIチップやNPU(Neural network Processing Unit)と共通項の多い技術であるが、それらを取り巻く技術の進化は、今後の集積回路技術に大きな影響を与えると思われる。

次いで、Computer Graphics(CG)処理のアルゴリズムとGPU回路の違いを見てみる。「特定のアプリケーションが要求するアルゴリズム」と「そのアルゴリズムを実装する集積回路設計」の間の違いが際立つ好例であるからである。GPUの世界では、アルゴリズムをそのままロジック回路に置き換えたスタートアップが多かったが、画像処理の世界ではそれら企業の多くが姿を消した。

アクセラレータ回路

GPUは、CPU(Central Processing Unit)にとって大きな負荷となるCG(Computer Graphics)の画像処理の多くを引き受けるアクセラレータ回路として、当初PCやワークステーション向けに開発されて来た。アクセラレータ回路としては、他に、FPU(Floating Processing Unit)や、DSP(Digital Signal Processor)が有名である。昨今のスマートフォンやタブレットPCのプロセッサ(Application Processor)には、非常に多くのアクセラレータ回路が搭載され、GPUの動作を補完している(注1)。

但し、現代のGPUは、「単にCPUの画像処理をオフロード(Off-Load)する」というよりも、「CPUとGPUの密な連携動作によって、CPUの得意なランダムな動作はCPUが行い、GPUが得意な繰り返し動作はGPUにて行う」という分担連携動作(Co-Processing)に変わったように見える。GPUは、「様々な要素データ間にも、共通な構造が有ること」や、「可動な物体には、共通の物理法則が作用する」等の多くの「共通なデータ処理」を行わせるのにも向く。過去にGPUがたどった進化の跡を見ることは、今後のロジック系集積回路の進化を考える上での参考になりうるのではないだろうか？

GPUの開発競争の過去と現況の概観

GPUの前身となるグラフィックアクセラレータ回路は、1980年代、グラフィックコントローラとも呼ばれる回路であった。日本でのゲーム機ビジネスの興隆もあり、PC用チップの開発 / 製造には、多くのスタートアップ企業がチャレンジした。但し、その多くは2000年前後までに淘汰された。

事業が行き詰まった理由は様々であったであろうが、生き残った企業には、「アルゴリズムを、そのままチップ上のカスタムロジックとして実装にする」のではなく、「プログラマブルな新回路アーキテクチャを見出した」との点が共通すると思われる。

特定のアルゴリズムは、特定用途にしか使えない。特定用途にしか使えない集積回路は、時代遅れになりやすく、知財としても商品としても価値は持続し難い。

アルゴリズムがソフトウェア技術者の知財成果物であるのと同様に、集積回路設計者も設計の過程で知財成果物を獲得する必要がある筈である。「回路構造に関する発想」であれば、細部の違いがあっても知財としての価値がありうるし、「アーキテクチャ」のレベルでの技術革新ができれば、価値を主張できる期間も長い。

2000 年前後に GPU 製品を提供した企業数は、5 社程度となっていたと見られる。筆者が調べた範囲では以下の企業名が上がった。

- ・3dfx Interactive 社(2000 年に、Nvidia が買収)、
- ・ATI Technologies 社(2006 年に、AMD 社が買収)、
- ・Silicon Graphics 社(2009 年から Silicon Graphics International 社、2016 年には HPE 社に吸収された)
- ・Matrox Electronic Systems 社
- ・Nvidia 社、

現在、PC 用の外付け GPU を供給するのは、Nvidia 社(シェア 77%)、AMD 社(23%)の 2 社である(参考資料 1)。PC 用 MPU 内蔵の GPU を供給するのは、Intel 社と AMD 社と言って良いだろう。

しかし、スマホやタブレット PC 用の APU (Application Processor Unit) 内蔵の GPU としては、

- ・Qualcomm 社(Qualcomm の自社開発。開発チームは、旧 ATI Technologies にルーツを持つ)
- ・Apple 社(Apple 自社開発)
- ・Arm 社(商標は Mali)
- ・Imagination Technologies 社(商標は Power VR)

の 4 社が知られている。加えて、韓国の Samsung は、現在 AMD 社との協業により自社向けに新規開発していると言われている(参考資料 3,4)。

更に、現在、中国での GPU 開発企業数は急増している。全社、スタートアップと言って良いだろう。

- ・兆芯社(参考資料 5)
- ・摩爾線程(Moore Threads)社(参考資料 6)
- ・Biren Technology 社(参考資料 6)
- ・Metax 社(参考資料 6)
- ・登臨科技(DengLin Technology)社(参考資料 6)
- ・天数智芯半導体(Iluvatar CoreX)社(参考資料 6)
- ・芯動科技(Innosilicon)社(参考資料 6)
- ・ZhaoXin 社(参考資料 7)
- ・ChangSha JingJia Microelectronics 社(参考資料 8)
- ・Horizon Robotics 社(参考資料 9)
- ・Black Sesame Technologies(参考資料 10)

これら企業の中には、中国製 PC や中国製ゲーム機を狙って開発を進めるケースもあるが、人工知能のような比較的新しいアプリケーションをターゲットとしている企業もある。NPU や AI チップを狙ったスタートアップは、リストアップに窮する程に増えている。

GPU の進化の過程で、「CPU と GPU の間の役割分担見直しが起こった」と同様の変化が、今後、NPU のような新たな処理回路を交えて起こるのではないだろうか？ その変化のためには、ソフトウェアにも影響が及ぶるので時間のかかる開発となるのではあろうが、今も、ロジック集積回路のアーキテクチャは、大きな進化の途中にあることは間違いない。

CG 処理のアルゴリズム

次に、CG 処理のアルゴリズムから出発し、GPU という回路アーキテクチャにたどり着いた米国 Nvidia 社の GPU 技術を概観してみる。

尚、筆者は、画像処理アルゴリズムの専門家ではなく、本内容も画像処理の解説ではない。そのように見識の乏しい分野の表現に挑戦するのは、Nvidia 社の技術が、「特定のアルゴリズムの例から、(自然界の)データを処理する集積回路の一般形を見出し、それを進化させた」という好例だと思うからである。内容は、参考資料 12~16 に基づいて、まとめている。

Computer Graphics(CG)処理は、以下のような手順(アルゴリズム)で進むとされている。(図 1)

1) 背景と物体のそれぞれをモデリング (Modeling) する。

- ・個々の物体の表面を、多数の三角形のメッシュに分解し、物体を、「番号付けされた頂点座標(Vertex Coordinate)の群」で表わす。ここで、三角形のメッシュに分解するというのは、典型的な方式の一つではない。一般的には、多角形(Polygon)、もしくは曲線のメッシュで分解する。

- ・これによって、元の物体は、「番号付けされた頂点座標(Vertex Coordinate)の群」で表されることとなる。元の物体の外形を分解した三角形(一般的には多角形)をオブジェクトと呼ぶ。

2) 物体の外形を分割した各三角形に、背景空間内に位置付けた座標と、その三角形の向きを意味する法線ベクトル情報を関連付ける。

3) 背景を照らす照明(発光体)を配置し、物体の陰影(Shading)を計算する。

- ・光源(Lighting)からの光や、各オブジェクトから反射光で照らされるオブジェクト表面の明るさと色の諧調を計算する。

- ・物体を分割した各三角形に、色、材質(マテリアル)、質感(テクスチャ)、透明度、反射率、屈折率、自己発光量、テクスチャ(表面の質感)、凹凸の設定、断層のような切れ目の設定など、多数の表現パラメータ情報を関連付ける。

陰影の補完(オブジェクトを構成する三角形の継ぎ目をなだらかにする)とか、陰面消去(視点から見るオブジェクトの前後関係を割り出し、見えない部分のデータを消去する)等の処理を行う。

4) 背景と物体を見る「視点」を定義し、背景と物体の全体像から「視点から見た 2 次元映像データ」を生成する。

5) 「視点から見た 2 次元映像データ」をスキャンして、ピクセル単位に出力する。

これらの処理にて使われる計算としては、一般に、「線形演算(もしくは行列演算)」と「逆数を得る計算」が使われる。背景と物体を分解した構成要素となるオブジェクトの圧縮・伸張(比例計算)も、回転も行列演算によって表現可能だからである。

その行列計算は、一般に、 $\{ Y=aX+b \}$ の演算を繰り返すことで求められる。従って、上記の CG のデータ処理は、いずれも、

- ・ { $aX+b$ }を計算し、返答する。
- ・ 逆数を得て、返答する
- ・ 付帯情報を関連付ける

などの比較的少ない種類の演算の組み合わせにて表現できるだろう（注 2）。

また、そのような処理ができる簡単な「プログラマブルな処理回路」を用意して、「その処理回路に、CPU からデータと計算式を送り、その返答を CPU に返してもらう」ことを繰り返すことで、アルゴリズムを進めることができるはずである。

もちろん、上記の CG 処理アルゴリズムは、図 1A のように、個別のステップ専用の処理回路を用意して進めても良い。また、下図 1B のように、「プログラマブルな処理回路にデータと計算式を送り、返答を受ける動作を繰り返して進める」という方法も有り得る。

PC 用の 3D グラフィックアクセラレータが登場した当初、多くのスタートアップが、図 1A のように、アルゴリズムをそのままロジック回路に置き換えたパイプライン回路をチップにしたが、生き残った企業のいくつかは、それでは「アルゴリズムの変化への対応能力に乏しい」と考え、図 1B のようなアーキテクチャを目指した。

Nvidia 社は、図 1B の「プログラマブルな処理回路」にプログラムを駐在させているのではなく、「非常に簡単なプログラムをデータと共に受け取り付けて処理を行い、結果を返答する回路」とした（注 3）。図 1A のように、特定のアルゴリズムを論理回路とした製品は、アルゴリズムの進化に対応できなかった。

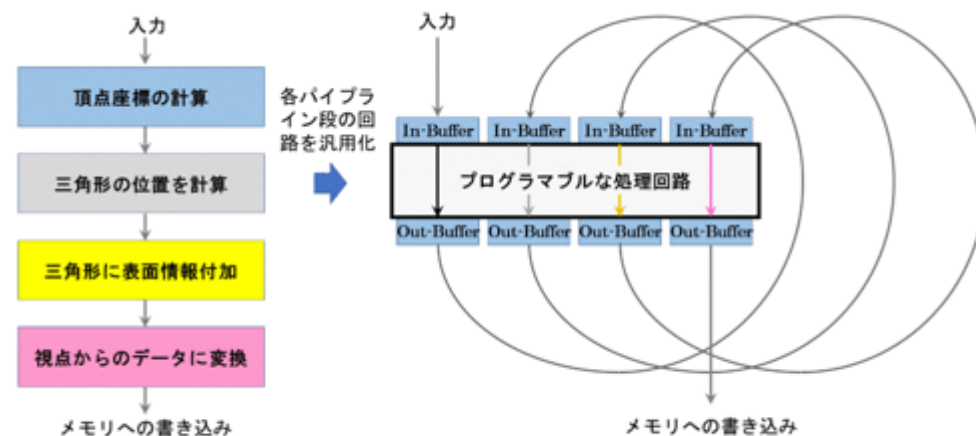


図 1 A. アルゴリズムのパイプライン化

図 1 B. GPUの進化形
プログラマブルな処理回路に、データとプログラムを一緒に送って、繰返し処理にてアルゴリズムを進める。

図 1 3D グラフィックスの画像処理アルゴリズムをそのまま論理回路化したパイプライン(図 1A)と、進化後の集積回路(図 1B)

出典: 図 1A は、参考資料 12 を元に筆者が作成。図 1B は、参考資料 13 を元に筆者が作成

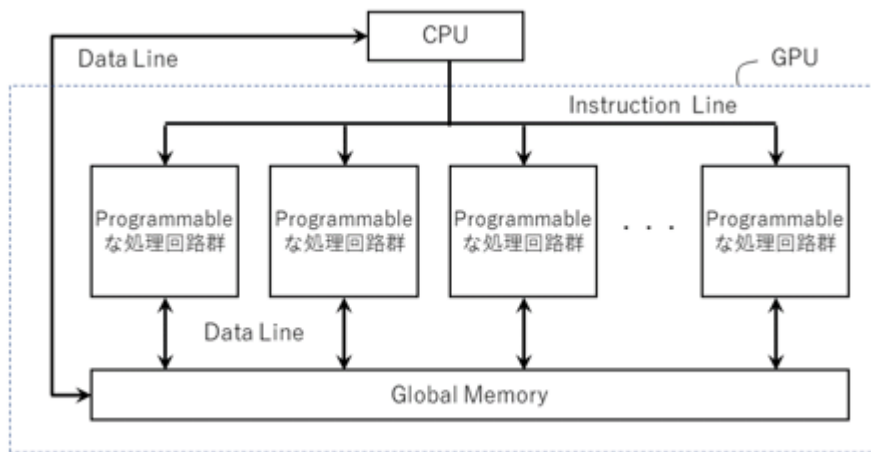


図 2 GPU の構造と CPU と接続した場合の命令と出力データの流れ

GPU は、性能がスレッドの個数に比例するスケーラブルな回路となった。一方、CPU は、GPU の各スレッドに命令を供給する回路となった。 出典:参考資料 15 を元に、筆者が作成

スケーラブルな回路

図 1B の「プログラマブルな処理回路」は、非常に小さな回路であるために、1 個のチップ内に搭載する個数を増やすことで、物体ごとに、もしくは表示画面の画素ごとに処理を進めることができる並列回路となった(図 2)。このような方式に変わると、CPU はデータを処理する回路というよりも、GPU の各スレッドにデータとプログラムを送る制御回路となる。これは、データ処理の現場が CPU から GPU に移ったともいえる大きな変革であった。

また、GPU 方式(図 1B、図 2)は、ソフトウェア開発者にとっての開発インフラともいえる必須のハードウェアとなり、CG アルゴリズムの世界では、設定するパラメータを増やすことで質感を改善したり、また、光線や照明が醸し出す立体感を高めたり、という工夫が加速した。GPU 方式は、ソフトウェア開発者との間で進化を加速するエコシステムを形成する基礎を提供した訳である。

更に、GPU 方式は、スパコン等の HPC (High Performance Computing) にも用途を広げた。自然現象の多くは、マイクロなスケールで見た場合に線形演算の繰り返しとなるためである。科学技術計算シミュレーションへの用途の拡大は、当初、GPU 方式を発明した技術者が期待した以上の効果を事業にもたらしたのではないだろうか?図 3 はシリコンバレーにある Nvidia の本社だが、三角形のメッシュを描いた屋根は GPU を象徴している。

集積回路設計者は、回路アーキテクチャのレベルでの革新競争に日夜対面させられている。GPU の歴史は、「価値ある回路アーキテクチャの発見」が企業の命運を制したという例だったのであると思える。



図3 Nvidia 社本社の航空写真。多数の三角形の群でメッシュに切られた外観を持つ。

出典:[Nvidia Investor Presentation Q4 FY2021](#)

注

1. ビデオやオーディオデータの圧縮/伸張機能(MPEG-2 / H.264 / AAC / MP3 / WMA 等)や、セキュリティ対策を始め、様々なデータ処理 / 変換、等々。アプリケーション毎に多種多様の専用データ処理回路であるアクセラレータを搭載することで、汎用データ処理回路である CPU の動作を抑制している。
2. $\{ Y=aX+b \}$ の演算と逆数演算を繰り返すことで、級数展開された任意の関数も近似できる。従って、GPU が扱える演算子は非常に多い。尚、グラフィック用途での演算は、通常、浮動小数点数値に対する演算である。
3. Nvidia 社は、プログラマブルなパイプライン処理回路部分を複数搭載した GPU (商品名は GeForce 3) を 2001 年にリリースした。このチップで採用した方式は、データを「パイプラインの処理回路部分」に送る時に、データと共に規模の小さな一種のプログラム(シェーダーと呼ばれる)を付けて送り、パイプライン処理回路の機能をダイナミックに変えることができるアーキテクチャの先駆けとなったと思われる。

参考資料

1. Hassan Mujtaba, "[AMD Radeon RX GPUs Witnessed Market Share Gain Versus NVIDIA GeForce RTX GPUs in Q3 2020, RDNA 2 & Ampere Launches Not Yet Accounted For](#)" (2020/12/06)
2. 「[スマホの GPU とは?性能とメーカーを徹底解説](#)」(2019/08/19)
3. 「[Samsung が ARM、AMD とモバイル向けプロセッサを共同開発か?](#)」、iPhone mania(Web 記事) (2020/08/13)
4. 「[Samsung と AMD の協業で生まれる GPU、詳細は 6 月に発表?](#)」、iPhone mania (Web 記事) (2021/02/24)
5. 「[8 コア/DDR4 対応/GPU 内蔵の中国製 x86 プロセッサ『開先 KX-5000』](#)」、PC Watch (Web 記事) (2017/11/09)
6. 「[中国で GPU 開発企業への投資熱 設立 100 日の新興が数百億円も資金調達](#)」、36KrJapan (Web 記事) (2021/03/11)
7. Evan Federowicz, "[A Chinese Competitor to Intel and AMD CPUs? Zhaoxin Might be the Answer!](#)"

(2019/12/15)

8. 「[【GPU】中国産 GPU『JM9271』。GTX 1080 に匹敵か](#)」、ニッチな PC ゲーマーの環境構築 Z (Web 記事)、(2019/08/25)
9. 「[NVIDIA の中国のライバル Horizon Robotics が約 725 億円の資金調達を目指す](#)」、Tech Crunch (Web 記事)、(2020/12/23)
10. 「[アリババのクラウド事業「AliCloud」が GPU に投資する背景とは?](#)」、ビジネス+ IT(Web 記事)、2016/02/17
11. 「[レイ トレーシングとラスライズの違い](#)」、Nvidia 社のブログ記事、(2018/03/26)
12. ウィキペディアの「Graphics Processing Unit」、(更新 2021/02/11 14:30)
13. Chris McClanahan、“History and Evolution of GPU Architecture”、2010 年
14. 平野幸彦、「[GPU コンピューティングの歴史と CUDA の誕生](#)」、(2010/07/02)
15. 福井正博、林憲一、「GPU による LSI 設計の高速化技術」、2013 年 1 月、電子情報通信学会 2013
16. David E. Rogers、J. Alan Adams; “Mathematical Elements for Computer Graphics, Second Edition”, McGraw-Hill. Inc., (1989/02)

岡島 義憲

([情報統合技術研究合同会社](#))