

[パネル討論招待講演]

## VR の歴史と AI との関係

加納裕<sup>†1</sup>

**概要** : 本稿では, VR と AI の関係について述べる. 最初に VR と AI のそれぞれについて概観したのち, VR と AI がいかなる関係を持つかについて考察する.

**キーワード** : Virtual Reality, Artificial Intelligence

## The history of VR and its relation with AI

YUTAKA KANOU<sup>†1</sup>

**Abstract**: This article describes the relation between VR and AI. Firstly VR and AI are explained independently in a historical manner based on the author's experience. Next the relation between VR and AI is examined. Especially how the elements of VR technologies are interpreted by emerging AI technologies is considered.

**Keywords**: Virtual Reality, Artificial Intelligence

### 1. はじめに

本稿では, VR (Virtual Reality) と AI (Artificial Intelligence) のそれぞれについて, 筆者の経験をもとに, ここ約 30 年の歴史を概観する. そのあと, VR と AI がいかなる関係を持つかについて, その可能性を論じる.

### 2. Virtual Reality (VR)

筆者は, いまから約 30 年前の 1987 年, ソリッドレイ研究所という, 立体視機能を有するコンピュータ・グラフィックス・システム構築を主たる業務とする会社の設立に参加した[1].

当時はまだ VR という言葉はなく, 代わりに AR (Artificial Reality, 人工現実感) という単語が存在した. VR は VPL (Visual Programming Language) Research 社創業者 Jaron Lanior が, 1989 年に考案したとされる[2].

VPL Research 社は, 最初の商用 VR システムと言われる, RB2 (Reality Built For Two) を開発した. 本システムは, 当時の技術で開発可能な, VR に必要な要素のほぼすべてを含んでおり (HMD (Head Mounted Display) / 立体音響 / 入力用グローブ / 三次元センサ / シナリオエディタ / 3D モデラ / リアルタイムレンダリング / 等々), 当時としては画期的かつ完成度の高いものであった[3].



図 1 VPL Research 社の RB2  
Figure 1 VPL Research RB2.

VR という単語が登場して以来, 筆者らは VR システムと銘打って, さまざまなシステムを開発・納入してきた[4][5]. 特に, 立体視の品質向上に注力し, 立体感を忠実に再現できる CG 生成法を確立した[6][7]. その結果から, 箱庭効果の原因に対し, 従来と異なる要因を示した.

この時期におけるターニングポイントは, CAVE (CAVE Automatic Virtual Environment) という, プロジェクタを複数台使った, 包囲型タイプの VR 環境が現れたことである[8]. それ以前は, VR の視覚デバイスは, HMD か単一の立体視プロジェクタであった. この CAVE 的形態はひとつの方向性を確立し, さまざまな派生形が生まれることとなった.

筆者は, 1994 年にソリッドレイ研究所を退社し, スリー

<sup>†1</sup> ソフトキューブ(株)  
SoftCube

ディーという会社に入社したが、ここでも引き続き、多数の VR システムを開発・納入した[9][10][11].

1996 年、Thomas Massie が SensAble Technologies 社を設立し、Phantom という力覚を呈示する装置を開発した。この装置は、その品質の高さから、市場に受け入れられた。筆者もこれに感銘を受け、Phantom を用いたさまざまな VR システムを構築した[12][13].



図 2 SensAble 社の Phantom  
Figure 2 SensAble Technologies Phantom.

2000 年代初頭、NVIDIA 社が、GPU (Graphics Processing Unit) 上で直接プログラミング可能なシェーダ (shader) という技術を公開した。これにより、以前ではオフラインでしか生成できなかったような品質の CG がリアルタイムで生成可能となり、筆者らもシェーダを中心に据えた製品開発を行った[14][15].



図 3 シェーダを用いたソフトウェア  
Figure 3 Programmable shader.

CAVE に代表される包囲型 VR 環境もいくつか構築した。前世紀 (20 世紀) に手がけたものは、Silicon Graphics 社 (sgi) のワークステーションを使っていたが[16], 今世紀 (21 世紀) に入ると、PC を複数台用いた、いわゆる PC クラスタ

を用いて、同様のことが実現可能となった[17]. この頃になると、HMD はほとんど使用されることはなくなり、プロジェクタを用いた VR システムが主流となった。スリーディーにおいて最後に構築した VR システムは、上海に設置された、プロジェクタと PC を 12 台使用した 360 度球体映像システムであった[18].



図 4 上海の CAVE (1999 年設置)  
Figure 4 CAVE in Shanghai set-up in 1999.



図 5 産業技術総合研究所の CAVE (2003 年設置)  
Figure 5 CAVE in AIST set-up in 2003.

筆者は、2009 年にスリーディーを退社 (代表を退任) し、2010 年にソフトキューブという会社に入社した。ここでも、当初は VR システムを開発・納入していた[19][20].

2012 年に Oculus VR 社が設立され、安価で高性能の HMD が販売されるようになり、誰もが手軽に VR システムを構築できるようになった。Unity のようなソフトウェアが簡単に使えるようになり、VR ソフトウェアの技術開発を敢えて行う必要性はなくなることとなった。

この頃、画像認識のニーズが高まってきたことから、筆者は徐々にこちらの分野にシフトするようになり、現在に至っている。

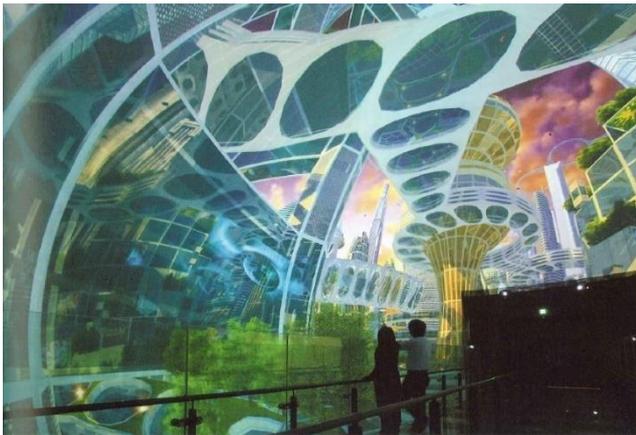


図 6 上海の 360 度球体映像

Figure 6 360 degree spherical image in Shanghai.

### 3. Artificial Intelligence (AI)

VR と AI の歴史は、時系列的に被るところがある。

1980 年代後半、ニューラル・ネットワーク（以下 NN）のブームが起こった。これは、back-propagation という、NN の学習法が開発されたことによる[21]。

その後、NN は期待されたが、さまざまな要因によりその期待に応えることができず（特に、層数を増やすと学習がうまく進まないという問題）、いわゆる「冬の時代」を迎えることとなった。1990 年代後半～2000 年代前半のことである。

この「冬の時代」には、NN ではなく、別のタイプの AI 技術の応用が推進された。そのひとつに、推論に基づくシステムがある。筆者らが、SOAR という推論エンジンを用いた、自律エージェント（Avatar Toolkit）の製品開発を試みたのもこの時期である[22][23]。

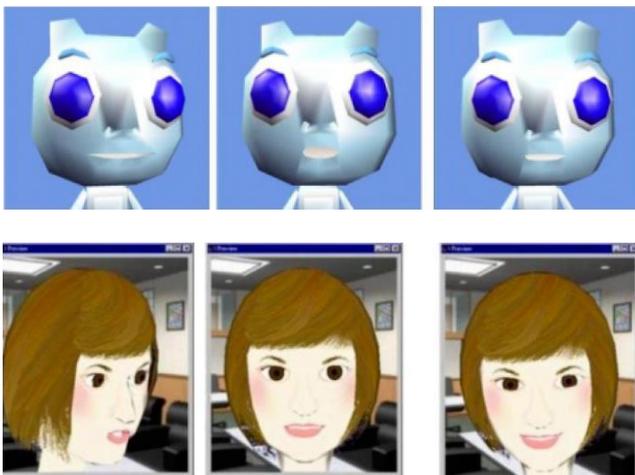


図 7 Avatar ツールキット

Figure 7 Avatar Toolkit.

2006 年に Christopher Bishop が、"Pattern Recognition and Machine Learning" という書籍を出版した [24]。いわゆる "PRML" である。同書は出版からすでに 10 年以上経過しているが、現在でも機械学習におけるバイブル的存在を維持している。実際、同書は多大な影響を及ぼし、AI に機械学習の手法が適用できることを広く示したものと言える。

いまの AI ブームを作ったのは、Deep Learning である。しかしながら、PRML には Deep Learning という言葉は一切出てこない。実際のところ、「冬の時代」にあった NN を復活させたのは、Geoffrey Hinton であり、2006 年のことである [25]。かれは、多層 NN でもある工夫をすれば、学習がうまくなされることを示した。その後、2012 年 ILSVRC (ImageNet Large Scale Visual Recognition Challenge) において、Deep Learning を用いたシステムが 2 位 (Support Vector Machine (以下 SVM) を用いたもの) に大差をつけて 1 位となり [26]、一躍 Deep Learning の性能が認知されるに至った。アルゴリズムの進歩と同時に、GPU による高速並列計算が手軽に使えるようになったことも大きい。現状はこの流れが続いている。

筆者は VR の出身であるが、関連のソフトやハードが安価に組みあわせ可能となったため、現職では特に VR の開発要素がなくなり、よりニーズの高い画像認識系に移行した。この分野はいまだにたくさんの課題がある。しかも AI との関連が深いので、結果的に AI の技術開発に関わるようになった。上記に説明したような、1980 年代後半の NN ブームに接していたことも影響を与えている。

なお、AI という言葉は、非常に一般的な用語である。現在 AI の代表格ともはやされているのは Deep Learning であるが、これは基本的には NN である。NN の学習というのは、データからネットの重みを繰り返して計算するので、これは機械学習である。統計的手法として以前から確立されている、主成分分析 (PCA) や各種の回帰 (regression) 技法も、いまだに AI の範疇に入っている。更には、比較的新しい手法である SVM や、起源は古い、効率的な数値計算手法が考案され、近年脚光を浴びるようになったベイズ統計も、AI に含まれる。煩雑な様相を呈しているが、これらをすべて、本稿では AI とみなす。

### 4. VR x AI

VR はもともと AI とは無関係に発展してきた。しかし近年、VR と AI がともに時代を牽引していることから、VR と AI の関係がいかなるものかを考える取り組みが各所でなされている。

本稿では、過去に VR で開発されてきたシステムの機能を、AI 的な要素として捉えた場合の位置づけについて考えてみたい。

#### 4.1 動作解析

VR における入力手段は、伝統的にヒトの動作であった。

筆者がその用途で最初に用いた装置は、磁気式の3次元センサーである。磁場発生源を近くに置き、磁気センサをヒトの位置姿勢を計測したい箇所に装着する（通常は、頭部や手首）というものである。これは、現在でも使われている定番の手法である。

そののち、リアルタイムで画像処理ができるようになり、磁気センサの代わりに、赤外線カメラで簡単に検知できるマーカをヒトに装着する、というシステムが開発された。磁気式と異なりケーブルが不要、しかも広範囲にデータが取得できることから、映画のモーション・キャプチャで多用されている。これは高額なシステムであるが、安価なVRシステムの代表例であるHTC VIVEでも、同様の技術が使われている。



図 8 HTC 社の VIVE (法政大・小池研究室にて)  
Figure 8 HTC VIVE at Koike Laboratory, Hosei Univ.

しかし、ここまではAIとは関係がない。通常の画像処理の枠組みである。

最近、ヒトになにもマーカを取り付けることなしに、カメラだけで動作が認識できる仕組みが開発されてきている。認識のための特徴量（たとえばSIFT）を事前に設計するのであれば、これは従来のコンピュータビジョン（CV）の枠組みであろうが、特徴量をあらかじめ設計することがなければ（つまり特徴量を自動的に獲得するとすれば）、これはAIの範疇であろう。ここはDeep Learningの派生である、CNN（Convolutional Neural Network）が活躍する舞台となるであろう。

VRではなく、MR（Mixed Reality、複合現実）の話になるのかもしれないが、一般に受け入れられているメガネが画像認識機能を伴うようになれば、まったく新しい未来が来るのではないかと考えている[27]。



図 9 MR の未来形  
Figure 9 Ubiquitous MR in the future.

## 4.2 音声認識

VRの入力は、伝統的にヒトの動作であった。しかしながら、別の入力手段もある。それは音声である。音声認識は、VRの枠組みでは語られることがほとんどなかったが、今後は有力な入力手段となるだろう。

音声認識は、以前から研究開発がなされてきた。まず、音声を文章に変換し（speech-to-text）、その文章を解析して、意味を理解するように努める（これが可能かどうかは後述）。そして、それに応答するような文章を生成し、最終的には音声に変換する（text-to-speech）。

文章の解析においては、以前からさまざまな手法が研究されているが、ここでもDeep Learningが席卷しつつある（特にRecurrent Neural Network, RNN）。他の要素技術としては、word2vecに代表される、単語や文章の分散表現が挙げられる。これは有力な技術と見なされている。

筆者は未体験であるが、Google Homeが使われ始めている。認識精度はかなりよいといわれ、ある制限下では実用的となっている。

## 4.3 知能

上記でみたように、画像・音声（文章）の認識においては、AIにより、さまざまな進展がみられている。それでは、よりチャレンジングなAIの適用可能性はどのようなものだろうか。

VR側から考えると、AIに外界の制御をしてもらえるとありがたい。VRというのは、ヒトとのコミュニケーション技術であるから、主役はヒトである。そしてヒトはVRシステムと対話するので、対話される側には、ある程度の知性が備わっていなければならない。従来において対話される側というのは、物理的な世界であり、ニュートン力学により記述される世界であった。Unityなどに装備されている「物理エンジン」がこれである。このエンジン（＝なにかに特化した計算を専用に行うソフトウェア）は、ニュートン力学が実装されており、さまざまな物体を現実と同じように動かす。しかしこれはAIではなく、物理計算である。

では、外界にヒトが存在するとどうなるであろうか。このヒトも、実際のヒトが遠隔などで制御するものであれば特にAIは必要ないが、そうではない場合、自律性を持ったヒト（自律エージェント）が必要である。ここがAIの出番

であろう。

この問題を考えていくと、最終的には次の問題に行きあたると。AIはヒトの脳(知能)を模倣することができるだろうか。

この問題は、古くから取り上げられている。有名なのは、Roger Penrose による一連の議論であるが[28][29]、近年では、Max Tegmark も関連する議論を展開している[30][31]。国立情報学研究所の新井紀子氏は、さまざまな講演会で「コンピュータは意味を理解しない」と述べている(筆者はこの意見に賛同する)。

## 5. おわりに

本稿では、筆者の30年以上の経験から、VRとAIの関わりについて述べた。何かしらの示唆(または視差?)を与えることができたとすれば、望外の喜びである。

**謝辞** 本稿を執筆するにあたり、貴重なコメントをいただきました。URCF-URI-WG メンバのみなさまに感謝いたします。

## 参考文献

- [1] 加納, 齊藤, “立体映像とコンピュータグラフィックス”, 情報処理学会誌, Vol.31, No.2, pp.265-273, 1990.
- [2] <https://www.vrs.org.uk/virtual-reality-profiles/vpl-research.html>
- [3] C. Blanchard, S. Burgess, Y. Halvill, J. Lanier, A. Lasko, M. Oberman, M. Teitel, “Reality Built for Two: A Virtual Reality Tool”, Symposium on Interactive 3D graphics, pp 35-36, 1990.
- [4] 齊藤, 加納, “3次元処理のための機器, ソフトウェアの流れの可視化への応用”, 流れの可視化, Vol.9, No.35, pp.11-15, 1989.
- [5] 加納, “最近のVRの動向”, 画像ラボ, Vol.41, No.3, pp.41-44, 1993.
- [6] 加納, “立体視によるCG表現”, カラー表現による可視化技術, pp.429-433, 1996. ISBN-13: 978-4938555542
- [7] 加納, “3D-CGによる立体映像の生成”, 立体映像技術—空間表現メディアの最新動向, pp.50-60, 2008. ISBN-13: 978-4781300382
- [8] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, “Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE”, SIGGRAPH 1993.
- [9] Y. Kanou, “Virtual Reality Training System”, Review of Automotive Engineering, Vol.25, No.3, pp.265-270, 2004.
- [10] 加納, “VRの応用事例”, 可視化情報学会誌, Vol.27, No.106, pp.181-186, 2007.
- [11] 加納, “工業塗装における産業用VR技術の応用”, 塗装工学, Vol.43, No.3, pp.84-89, 2008.
- [12] S. Hirose, K. Mori, R. M. Y. Lee, Y. Kanou, “A VR Three-dimensional Pottery Design System Using PHANToM Haptic Devices”, 4th PHATNoM Users Group, 1999.
- [13] I. Goncharenko, M. Svinin, Y. Kanou, and S. Hosoe, “Predictability of Rest-to-Rest Movements in Haptic Environments with 3D Constraints”, Journal of Robotics and Mechatronics, Vol.18, No.4, pp. 458-466, 2006.
- [14] 松本(大), 加納, “工業デザインにおける実時間コンピュータ・グラフィックスの有効性と導入方法について”, 日本設計工学会誌, Vol.41, No.1, pp.28-35, 2006.
- [15] Y. Kanou, “GPU Rendering”, International Workshop on 3D Technology, Seoul, Nov. 2007.
- [16] 竹田, 加納, 伊藤, “4面立体空間構成によるバーチャル水族館の試み”, 計測自動制御学会, ヒューマンインタフェース部会, 第50回ヒューマンインタフェース研究会, Vol.13, No.3, pp.451-454, 1998.
- [17] [http://www.chusho.meti.go.jp/keiei/gijut/2001/h13\\_kaihatsujigyo.htm](http://www.chusho.meti.go.jp/keiei/gijut/2001/h13_kaihatsujigyo.htm)
- [18] “公元2050 体験館 SMART HOME/公元2500 体験館 FUTURE COSMOS”, エンターテインメントビジネス, No.21, 2008.
- [19] 加納, “OpenGLを使って立体視の絵を作ろう!”, Interface, 1月, 2011.
- [20] 加納, “力触覚技術の実用化への期待と課題”, 日本バーチャルリアリティ学会・力触覚の提示と計算研究会, 11月, 2014.
- [21] C. M. Bishop, “Neural Networks for Pattern Recognition”, Clarendon Press, 1996.
- [22] T. Tamisier, Y. Kanou, “Artificial Brain for Conversational Avatar”, 4th International Conference on Human and Computer (HC2001).
- [23] T. Tamisier, C. Gautier, Y. Kanou, “The Brain Simulation Language, A toolkit for conversational and intelligent knowledge-based systems”, 5th International Conference on Human and Computer (HC2002).
- [24] C. M. Bishop, “Pattern Recognition and Machine Learning (PRML)”, Springer, 2006.
- [25] 岡谷, 深層学習, 講談社, 2015.
- [26] A. Krizhevsky, I. Sutskever, G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks”, NIPS, 2012.
- [27] 加納, “仮想現実(VR)”, テクノロジー・ロードマップ2016-2025<ICT融合新産業編>, 日経BP社, 2016. ISBN-13: 978-4822239831
- [28] R. Penrose, “The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics”, Oxford University Press, 1989. (邦訳: 皇帝の新しい心, みすず書房, 1994)
- [29] R. Penrose, “Shadows of the Mind: A Search for the Missing Science of Consciousness”, Oxford University Press, 1994. (邦訳: 心の影, みすず書房, 2001)
- [30] M. Tegmark, “Our Mathematical Universe: My Quest for the Ultimate Nature of Reality”, Vintage, 2015.
- [31] M. Tegmark, “Life 3.0: Being Human in the Age of Artificial Intelligence”, Knopf, 2017.