AIによる新たな映像制作を NVIDIA DGX A100が後押し

EmbodyMe様

AIによって革新的な映像サービスを生み出すEmbodyMeは、
映像の未来を切り開くためにチャレンジし続けるスタートアップ。
優れた性能と柔軟性を兼ね備えた「NVIDIA DGX A100」が、
AIの進化を支えることでエンジニアの研究をさらに加速させる。

動画配信サービスやライブストリーミングサービスが一般的になりつつあるとともに、コロナ禍の影響によってビデオチャットの需要が急速に高まっている昨今。「AIによる映像の再発明」を掲げ、AIを用いた次世代コンピュータグラフィックスの基盤技術やその応用アプリケーションを開発するスタートアップがEmbodyMeだ。

同社の代表的なプロダクトの1つとなるのが、AI技術の1つであるディープラーニングを用いた映像生成技術と、その技術を活用したスマホアプリ「xpression」である。xpressionは、1枚の画像や動画に映っている人の顔を認識し、その顔に対して、スマホカメラで映っている自分の顔の表情や頭の動きを“自然な形で”反映させることが可能。有名人の動画や写真に自分の表情を乗せることで、誰でも簡単にフェイクビデオや面白ビデオを作成できる。

このxpressionの技術を活用し、EmbodyMeの新たなメインプロダクトとして2020年9月にリリースされたのが、バーチャルカメラアプリ「xpression camera」である。xpression cameraは、「写真や動画を自分の顔の表情や頭の動きで動かす」というxpressionの機能を、ZoomやGoogle Meetといったビデオチャットや、TwitchやYouTubeなどでのライブストリーミング配信に対応させたもの。映像に対してリアルタイムに自分の顔の表情や頭の動きを乗せられることから、出社時のスーツ姿の画像を用意すれば、「寝間着姿やすっぴん顔のままであっても、オンライン会議に参加できる」と、同社の代表取締役CEOを務める吉田一星氏は説明する。

自社の強みはAIによるリアルタイムでの映像づくり

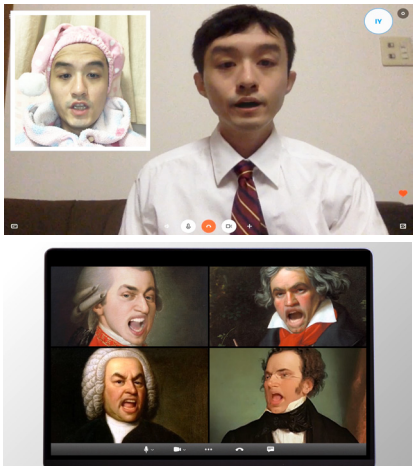
xpression cameraの開発にあたって吉田氏が注目したのは、最近のオンライン会議で「カメラをオフにしているケースが多い」という状況だ。「Zoom疲れ」とも呼ばれるオンライン会議での疲労感をはじめとしてその理由はさまざまだが、その一方で「相手の表情がわからないため、話しにくい」という意見もある。このようなニーズあるいは課題感に対する解決策としていち早く開発されたのが、まさに「xpression camera」なのである。

「新型コロナウイルス感染症が流行し始めた2020年2月当初、世の中が現在のように変化するとはもちろん想像しませんでした。しかし、時が経つにつれて『これは世の中のすべてが変わるかもしれない』と感じ、ビデオチャットやライブストリーミング配信は『次の主戦場になる』と確信しました」(吉田氏)。

また、吉田氏をサポートする同社エンジニアの柳川光理氏は、「AIによる映像づくりを“リアルタイム”で実現できる」という点を自社の強みとして強調。さらに、「現在は顔のみにフォーカスすることで、その性能を高めています。ハードウェアの性能をクリアできれば、同様の機能を全身に拡張させることも技術的には可能です」と補足する。

これに加えて、コスト面も注目すべきポイントとなる。例えば、最近では従来の3D技術を使ってリアルな人間を再現する映画もあるが、このようなケースでは数秒

EmbodyMe 代表取締役CEO 吉田一星氏



xpression cameraを利用すれば、寝間着姿であってもスーツ姿でいるかのようにビデオチャットができる(上)。また、偉人などの画像に自分の表情を乗せることも可能だ(下)

程度の短い映像であっても、その制作には膨大なコストと人手がかかっているうえに、必ずしも完ぺきな再現度とは言い難い。吉田氏としては、「GAN(Generative Adversarial Network)などのAI技術は進歩が速いため、それが従来のCG制作手法を追い越す可能性は高い」と見越し、現在のやり方で研究開発に取り組んでいるそうだ。柳川氏も「コストのかかる従来の3D技術をAIに置き換えることで、リアルなCGをより安価でより手の届きやすいものにしたい」と目標を掲げる。

3日になった学習時間のさらなる短縮を目指す

AIをさらに進化させるうえで必要不可欠となるのが、短いサイクルでの実験を実現する高速なGPUである。EmbodyMeでは、以前はAWSのGPUインタンスを利用していたが、2019年にGPU「Tesla V100」を8基搭載する「NVIDIA DGX-1」、2020年10月にGPU「NVIDIA A100」を8基搭載する「NVIDIA DGX A100」を導入した。

柳川氏によれば、xpressionやxpression cameraなどでは、人の顔が映っている大量の映像や画像データをAIに学習させ、顔写真から正しい顔の形状を推定するような訓練を行っている。さらに、「詳細な三次元顔形状と表情を読み解く部分には、完全に自社で開発した技術を使っている」とのことだ。

また、xpressionの技術的な特徴の1つとして、詳細な顔三次元形状の推定と同時に、独自のGANを用いてリアルタイムに口内の画像を生成している点が挙げられる。例えば、xpressionでは口を閉じた顔写真を1枚入力しただけでも、口を開けて自由に喋らせることが可能だ。ただし、この場合は口の中の様子を画像から知ることができないため、口の中の画像を生成する必要があるのだが、「CGで口の中を埋める」という従来の手法では「くちびるとの境目や質感などに違和感が残ってしまう」(柳川氏)。そこでEmbodyMeでは、口の周辺画像をベースに「GANを利用して口の中の画像を推定する」という手法を採用。「従来はリアルタイムでの動作が困難だったGANを、弊社は独自の技術で改良し、スマートフォン上でも、高品位な口内画像の高速生成を可能にしました」(柳川氏)。

ただし、口の中を表現するGANの学習は、画像から画像を生成するニューラルネットワークであるためパラメータが多く、「とくに時間がかかる」とのこと。実際、AWSのGPUインタンスでは高性能なものを借りても、学習が終わるまでに3週間かかることもあったそうだ。しかし、「NVIDIA DGX-1に最適化された訓練手法では、わずか3日程度で結果を得れるようになりました。もちろん、これはこれで十分早くはなったといえます。ただ、研究者としては『もっと早く』という欲求は尽きません(笑)。処理が速くなればそれだけ研究スピードがアップし、ひいてはそれが成果に結びついてくるだけに、いまはNVIDIA DGX A100でのさらなる時間短縮に期待しています」(柳川氏)。

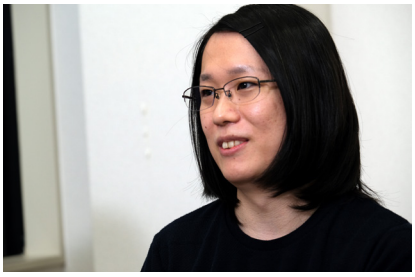
NVIDIA DGX A100の導入について、吉田氏はコスト面や効率面でのメリットを挙げる。「弊社のようなAI企業では、さまざまな研究で昼夜問わずGPUを利用しています。このような環境では、レンタルのGPUインスタンスはコストが高くついてしまうことから、昨年にNVIDIA DGX-1を導入しました。これによって状況が改善されたのは確かですが、現状ではそのNVIDIA DGX-1さえも社員が取りあうような状況になっていました。そういった意味でも、NVIDIA DGX A100の導入によって、より充実した開発環境を整えることができたと感じています」(吉田氏)。

映像の根本を作り替えて映像制作の選択肢を拡大

今後も、AIを活用して映像の未来に挑戦し続けるEmbodyMe。次の一歩として現在開発を進めているのが、音声から顔の表情や頭の動きを推定して映像に落とし込む技術である。これをビデオチャットにリアルタイムで応用できれば、「カメラの前にいる必要さえなくなり、別のことをしながらオンライン会議に参加できるようになる」(吉田氏)だろう。

また、現在は顔と頭の動きのみとなっているAIによる映像づくりについても、将来的には全身の動きにまで対応させる予定だ。これが実現できれば「例えば時代劇の殺陣のようなシーンも、写真1枚あればオンラインで映像化できるようになる」と吉田氏は考えている。

「そう遠くない未来、テレビや映画では実際の撮影が少なくなり、その映像のほとんどをAIが編集過程で作るようになるかもしれません。また、現実ではあり得ないようなバーチャルキャラクターや、すでに亡くなった有名人を出演させることも可能になるでしょう。映像の根本を作り替え、さまざまな映像制作の選択肢を拡大していくことが、我々の最終的なゴールです」(吉田氏)。



EmbodyMe リードエンジニア 柳川光理氏

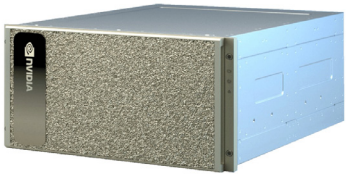
ユーザプロフィール
組織名：株式会社EmbodyMe
業界：ソフトウェア開発/ NVIDIA Inception パートナー
本社所在地：〒169-0075
東京都新宿区高田馬場3丁目23−3
設立：2016年6月
資本金：9,100万円(2020年12月現在)

導入ソリューション
NVIDIA DGX-A100/DGX-1

EmbodyMeの使用モデル

NVIDIA DGX A100

最新アーキテクチャAmpereを採用したTensorコアGPU「NVIDIA A100」を8基搭載するAIワークフローのためのユニバーサルシステム。演算性能は、倍精度で19.5TFLOPS、単精度で312TFLOPSのスループットを実現するとともに、ディープラーニング推論性能は1,248TOPSを誇る。



株式会社ジーデップ・アドバンス
〒105-0014　東京都港区芝1-12-7　芝一丁目ビル4階
TEL：03-6803-0620　URL：http://www.gdep.co.jp/