

画像生成AIの新機能を試す



第3回

生成スピードを短縮できる「LCM」

湊 雄一郎

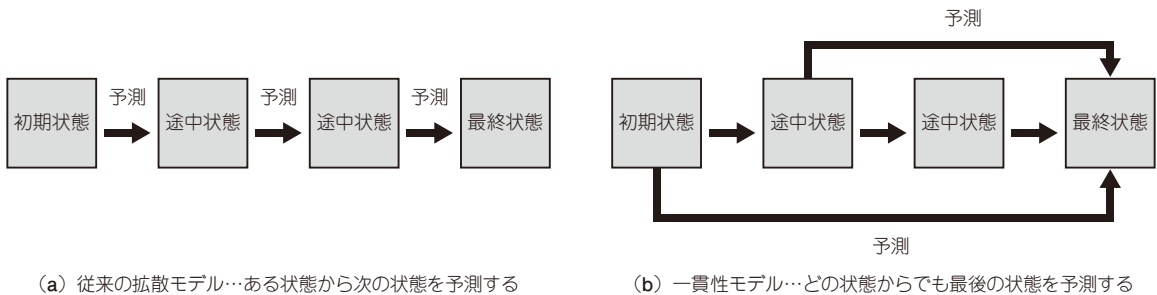


図1 画像生成のモデルのイメージ

画像生成AIは、機械学習の技術を用いて莫大な画像データを学習し、そこから新しい画像を作り出す技術です。

もう少し具体的にいうと、プロンプトと呼ばれるテキストや画像そのものをガイドとして、なるべく指示通りの画像を生成しようとします。このような新しい画像生成AIは今では気軽に個人でも利用できるようになりました。

本連載では、どんどん進化する画像生成の仕組みや利用の仕方をご紹介します。

● 今回のテーマ

今回は、画像生成の生成スピードを短縮できる、Latent Consistency Models (LCM：潜在一貫性モデル) という、Stable Diffusionなどの拡散モデルから派生したモデルをご紹介します。

従来のモデルとLCMの違い

● 拡散モデル…徐々に状態を変化させて画像を作り出す

拡散モデルはStable Diffusionなどで採用されていて、あるノイズの状態から少しずつノイズを除去して画像を復元するモデルです。初期状態から最終状態へと状態を徐々に状態を変化させてきれいな画像を作り出します。前の状態から次の状態へと移る際のノイズ

の除去の仕方を、U-Netやトランスフォーマと呼ばれるニューラル・ネットワークに学習させて利用するのが一般的です。

● 一貫性モデル…状態に自己一貫性がある

一貫性モデル (Consistency Models)⁽¹⁾ は拡散モデルの式を拡張したものです。これまで離散的 (とびとび) に捉えていた時間の概念を連続的な時間として捉える連続時間拡散モデルを提唱することで技術的な発展を実現しています。

連続時間拡散モデルは、確率微分方程式 (1つ以上の項が確率過程である微分方程式) と呼ばれる式の形で表現できます。

$$dx_t = \mu(x_t, t) + \sigma(t) dw_t$$

ここで、 μ は全体的変化の方向性を表すパラメータで、 σ はばらつき具合を表しています。また、 dw_t はウィナー過程という正規分布に従う確率分布のランダム成分です。確率という答えが定まらないイメージがありますが、条件が合うと確率微分方程式も解を一意に求めることができます。

式変形を行って解を求める方法を確認します。上記の確率微分方程式は、最終的には常微分方程式に書き換えることができ、下記のように表されます。

$$dx_t = \left[\mu(x_t, t) - \frac{1}{2} \sigma(t)^2 \nabla \log p_t(x_t) \right] dt$$

ここで、 x_t の分布を $p_t(x_t)$ としています。以前は離

