

Multi Object Tracking(MOT)とは？

Multi-Object Tracking (MOT) は、動画の中で移動していく複数の人や物をそれぞれ区別して継続的に追跡する画像処理技術です。MOT ではそれぞれの追跡物体に ID を割り振りますが、同じ対象物には可能な限り同じ ID を与え続けることを目標とします。

(1) サッカー用 MOT の問題点

技術的には画像処理の基本的な技術の一つである物体認識の応用技術になりますが、静止面の物体認識とは異なり動画中の物体は常に外観が変化し続けることから、非常に難しいタスクとされています。

特に技術的に難しいのは、人々への ID の割り当てをすることで、例えば人が他の人や障害物の後ろに隠れたり、再び現れた時にどのように ID を切り替えるかというところです。

さらに試合中のサッカー選手やボールを追跡する際、下記の原因によって、頑健に追跡することはより難しいとされています。

1. 高速で、不規則な動き

サッカーの試合では、選手とボールは常に高速で不規則な動きをしていますので、それらの動きを正確に予測し、追跡することは困難を伴う。

2. 視点と距離

追跡対象の距離や角度により、選手やボールの識別が難しくなる場合があります。たとえば、距離が遠くなると角度が浅くなることで選手とボールがかなり小さく見えるので、追跡することが困難になります。

(2) MOT の代表的なアルゴリズム

MOT は近年進歩しています。その多くは、SORT とよばれる手法を改良したものであり、基本的にはカルマンフィルターを利用して、検出した物体同士に関連性を推定します。

SORT では、単純な幾何学的な特徴(位置, ボックスの形状など)を用いたカルマンフィルターにより、インスタンスの動きをモデル化しています。SORT は高いトラッキング性能をもたらしていますが、物体が①短期間では線形的な動きをとると想定し、②継続的な観察を必要とする、という特徴があります。

単純な幾何学的な特徴だけではなく視覚的な特徴量も考慮すればもっとパフォーマンスが出るのではないか？という仮説の元に研究が進みました。視覚的な特徴も使用した MOT 手法として、DeepSORT という手法が有名ですが、DeepSORT は物体検出モデルによって検出されたオブジェクトを ReID モデルにかけるため、モデル 2 回分の推論が必要で速度が遅いという欠点がありました。

(3) 最新の MOT アルゴリズム

Cascaded Buffered IoU (C-BIoU)は2023年に発表された最新で state-of-the-art なリアルタイム MOT の手法です。この手法は外観が似ている不規則な動きをするものに対して特に優秀な追跡性能を持つとされています。弊社は C-BIoU を含めて色々な最新、最先端のアルゴリズムを研究し、独自のアルゴリズムを開発し、サッカー選手とボールを高精度で追跡することを目指して努力しています。