

# 自駕車猜出你的下一步

 [highscope.ch.ntu.edu.tw/wordpress](https://highscope.ch.ntu.edu.tw/wordpress)



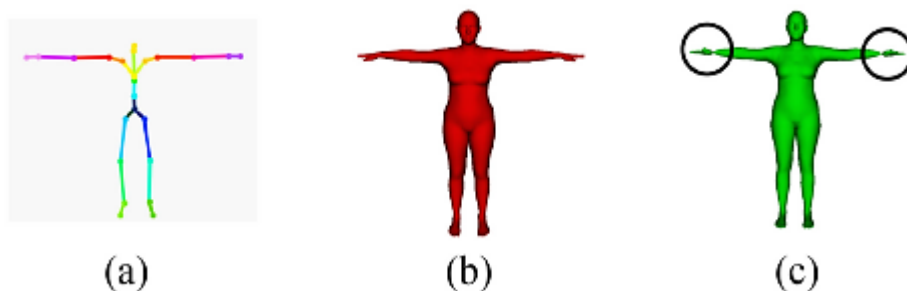
## 自駕車猜出你的下一步

撰文 / 黃柏瑋

在轉彎的路口上，司機和行人正打著心理戰。誰先停、誰先行？如果沒有達成默契，在一聲巨響之後，便會留下永久的遺憾。近年來，自動車逐漸普及，「如何預測行人的下一步」顯得更加重要。倘若AI能夠預先判斷行人的行動方向與位置，便擁有較多空檔進行更完善的策略評估，除增進人機互動的默契外，也在速度和安全間求得平衡，提升都市的交通效率。

## SMPL模型

事實上，人類姿態預測的研究不在少數。有些模型利用主要關節位置表示目標的姿勢，並以2D圖形輸出預測結果，減輕不少影像表示的麻煩；然而這個看似合理的設計，其實犧牲了不少細節，侷限了模型應用的潛力。舉例來說，圖一中 (a) 無法解釋目標的手心方向，雖說就目前的畫格而言似乎無傷大雅，但很可能影響接下來的猜測，做出完全截然不同的判斷。



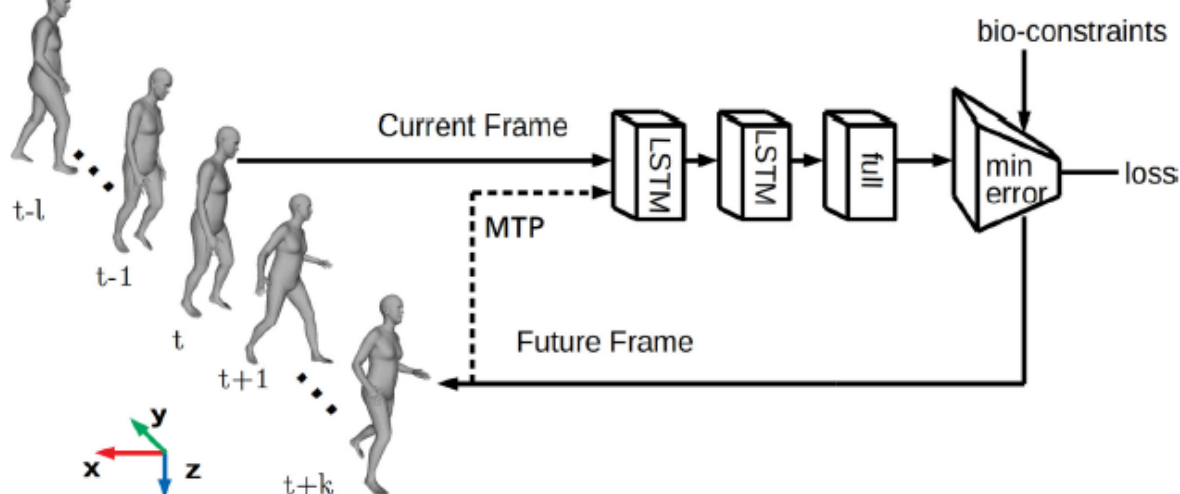
圖一、2D骨架與3D網格表示圖的差別 ( 圖片來源：X. Du, et al, 2019. )

有些模型則需要將目標置中定位，因此無法負擔多人預測。而這樣的限制會影響機器的判斷，尤其在熙來攘往的十字路口上，眾人會同時移動，我們需要注意的，不是某個特定焦點，而是整體交通狀況。

bio-LSTM便是為了這個目的而誕生，除了採用擅長序列預測（sequence prediction）的長短期記憶網絡（Long short-term memory，LSTM），還搭載SMPL模型（Skinned Multi-Person Linear Model），藉由85個參數構成表面網格（mesh），將先前的2D推廣為3D資料，其中包含3個紀錄人物與偵測系統距離的變換參數（translation parameter）、72個表示關節位置和旋轉角度的姿勢參數（post parameter）以及10個代表人體輪廓的形狀參數（shape parameter）。

比起單純分析骨架關節位置，SMPL在還原人物方面考慮了更多細節，有效降低預測的失誤，而變換參數也改良了原先人物置中的缺陷，解決無法進行多人分析的窘境。

## 步態生物力學

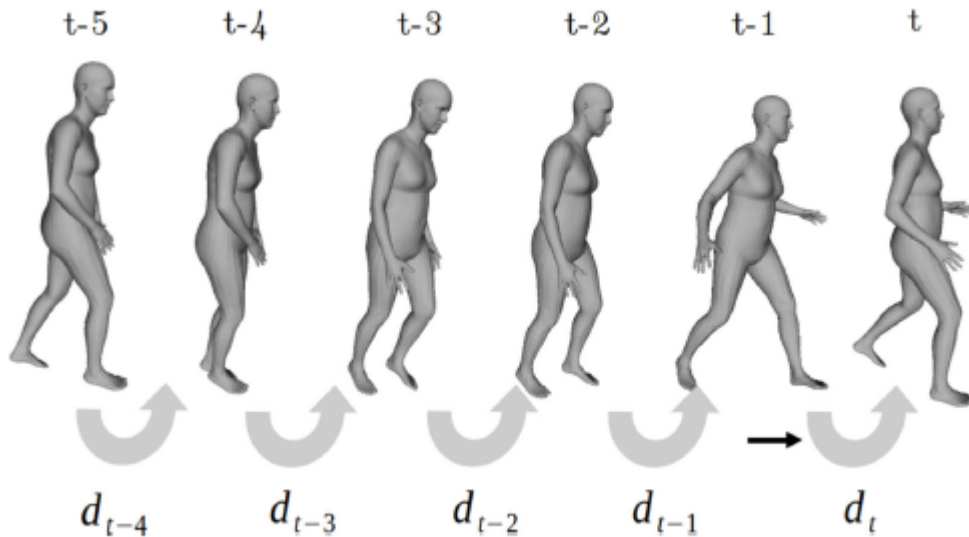


圖二、bio-LSTM的結構示意圖（圖片來源：X. Du, et al, 2019.）

bio-LSTM的結構（如圖二所示）除了以兩層式LSTM循環神經網路為主幹、SMPL參數來表示資料之外，還利用步態生物力學（gait biomechanics）來修正運動預測的不合理處。

首先，研究學者發現，人類在穩定行進時的步伐具有一定的規律，於是假設人腿會進行週期運動，其中相鄰畫格的腳速、方向和步長不會遽然大變，因此我們只需要推算畫格之間的差數即可（如圖三）也就是將預測的差數（ $d^{\wedge}$ ）和訓練資料的事實（ $dt$ ）進行比對，建立「週期損失函數」（ $L_c$ ），符合人類的實際運動狀態。

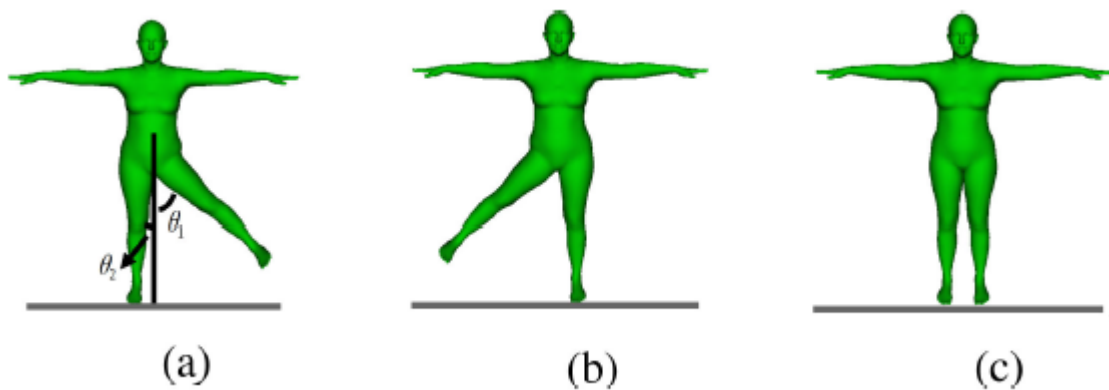
$$L_c = |dt - d^{\wedge}|$$



圖三、 $dt$ 表示時間 $t$ 和時間 $t-1$ 的資料差值。（圖片來源：X. Du, et al, 2019.）

再者，人類在站立和行走時，身體應當呈現鏡像對稱（**mirror symmetry**），維持左腳和右腳與鉛直線的夾角相等（如圖四），肩膀與身體中線的夾角亦然。因此，我們定義「對稱損失函數」（ $L_s$ ），將人類運動時的平衡感考慮在內。

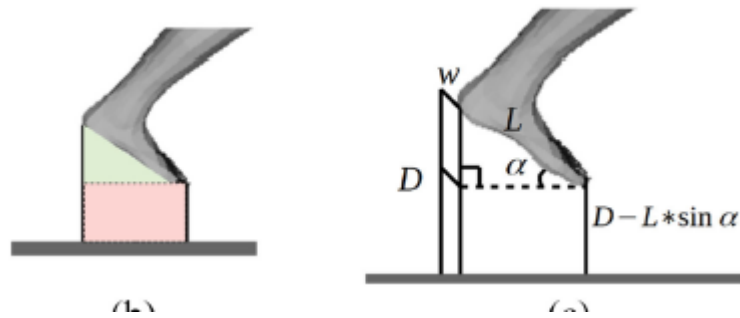
$$L_s = |\Theta_{leg1} + \Theta_{leg2}| + |\Theta_{shoulder1} + \Theta_{shoulder2}|$$



圖四、 $\Theta_1$ 和 $\Theta_2$ 分別為左腿和右腿與鉛直線的夾角，(c)為平衡穩定狀態。（圖片來源：X. Du, et al, 2019.）

最後，足夠的地面反作用力（**Ground Reaction Force**）是人類支撐自身身體十分重要的一環，而此力量主要由雙腳提供。導致人類在穩定行進時一定要與地面接觸，否則會踉蹌摔跤。於是我們需要定義「GRF損失函數」（ $L_g$ ），討論腳底離地的體積（如圖五），考慮重力帶來的運動限制：

$$L_g = wD(L\cos\alpha) - 1/2w(L\sin\alpha)(L\cos\alpha)$$



圖五，腳底板被簡化為平面，方便計算離地體積。（圖片來源：X. Du, et al, 2019.）

歸結以上三種生物力學因子，能預測人類穩定步行動作的目標函數即為：

$$\min L = L_c + \lambda_1 L_s + \lambda_2 L_g$$

兩個正規化參數 ( regularized parameter )  $\lambda_1$  和  $\lambda_2$ ，分別根據迴路測試的結果定為10和0.01，調整不同因子對人體運動的影響權重。

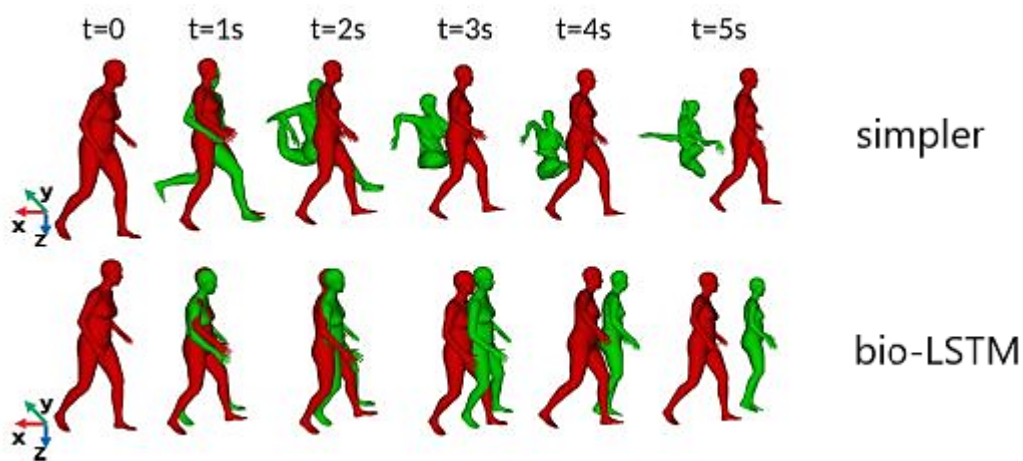
## bio-LSTM的前景



圖六、bio-LSTM的預測結果（綠人）與實際情形（紅人）的比較。（圖片來源：X. Du, et al, 2019.）

目前已可藉由過去五個畫格的表現，精確預測行人下一畫格的狀況（人類的步伐周期大約五至六個畫格），如圖六中預測結果（綠人）其實已與實際情形（紅人）相去無幾；除「一般步行」外，在「端著咖啡」、「扛揹重物」和「手機滑落」的情境下，也能有相當卓越的成績，可見bio-LSTM的學習潛力不容小覷。

但若想要一次預測數個畫格呢？這種預測稱為MTP ( Multiple-Timestep Prediction )，需要不斷將輸出餵回網路，成為新預測的參考資料（如圖二中的虛線），但也可能因此導致誤差越滾越大，波及預測的參考價值。然而，bio-LSTM的表現仍舊毫不遜色，和相對簡易的模型相比，結果明顯更貼近現實，合理還原人類穩定行進時的大致步態（如圖七）。



圖七、MTP預測結果。( 圖片來源：X. Du, et al, 2019. )

雖然bio-LSTM目前仍停留在草創階段，而只能考慮個體行為。但或許在不遠的將來，我們真能夠打造出低事故發生率又兼顧車流順暢的完美「行人友善路口」。

#### 參考資料

1. X. Du , R. Vasudevan, and M. Johnson-Roberson, “Bio-LSTM: A Biomechanically Inspired Recurrent Neural Network for 3D Pedestrian Pose and Gait Prediction”, *arXiv*, 2019.
2. Michigan IT News, “Teaching self-driving cars to predict pedestrian movement”, *Michigan IT News*, 2019.

(本文由教育部補助「AI報報—AI科普推廣計畫」執行團隊編譯)