

臺 北 市 立 麗 山 高 級 中 學

109 學年度學生校內科學展覽

作 品 說 明 書 封 面

參展領域：應用科學

組別：高級中等學校組

作品名稱：震動控制方法研究

關鍵詞：震動控制、強化學習

編 號 (主辦單位編列)：

目錄

摘要.....	2
壹、研究動機.....	2
貳、研究目的.....	2
參、研究設備及器材.....	3
肆、研究過程或方法.....	4
一、 研究設計.....	4
(一) 實驗裝置設計.....	4
(二) 總體流程圖.....	5
伍、研究結果.....	7
陸、討論.....	8
柒、結論.....	8
捌、參考文獻資料.....	9

摘要

目前常見震動控制器通常是用磁流變液和電流變液減震器以改變剛性或是使用液壓來改變阻尼係數，這樣的控制方式反應速度較慢且效果較差，本研究探討如何以強化學習算法來控制電訊號改變線圈產生的磁場，能讓彈簧在震盪過程中更快達到較小振幅或是停止

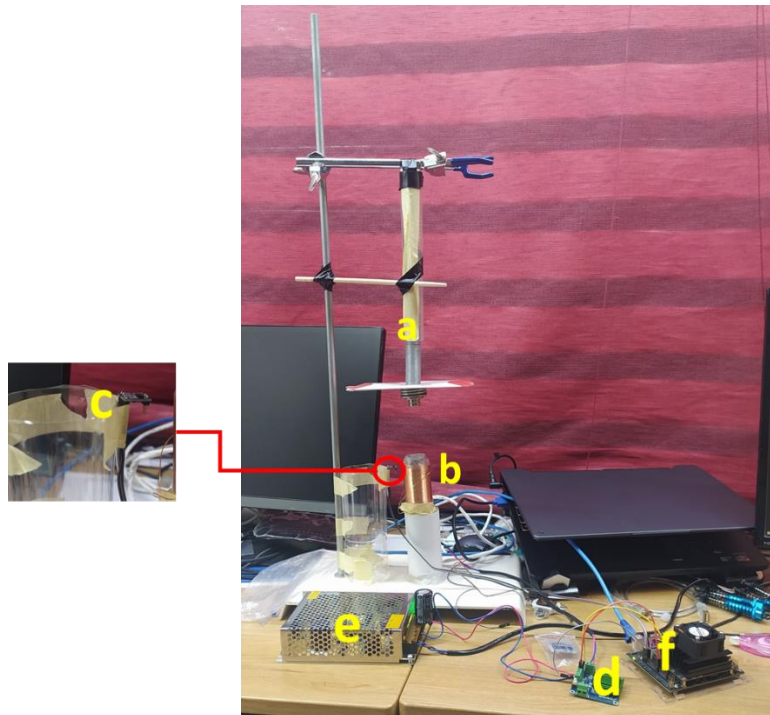
壹、研究動機

目前常見的震動控制器通常是用磁流變液或液壓或是氣動的方法來改變阻尼係數或剛性，這樣的控制方式由於流體流速會造成反應速度較慢，因此我想以改變磁場的方式控制; 利用強化學習算法控制，使震動控制更為精準快速，補足傳統機械控制理論的不足

貳、研究目的

研究以強化學習控制電訊號在線圈中產生磁場，控制彈簧減振的避震效果

參、研究設備及器材

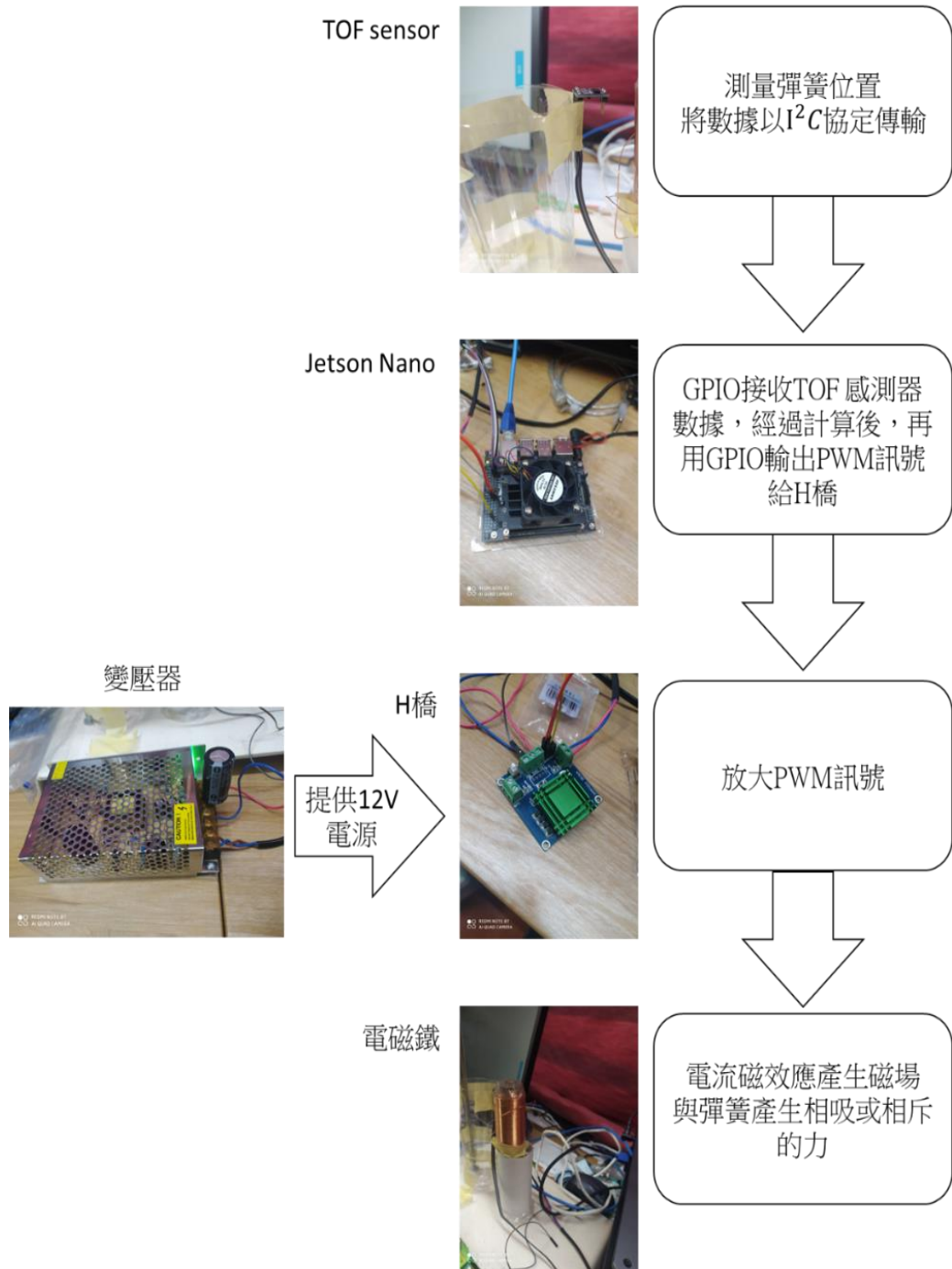


- a. 彈簧:震動本體，在其周圍加裝塑膠套管使其震動維持在一維方向
- b. 電磁鐵:用輸入的電流控制磁場方向與大小，進而和彈簧產生相吸相斥的力量
- c. TOF 感測器:測量並回傳與目標物的距離
- d. 變壓器:加上一顆電容用以穩壓
- e. H 橋:可改變其連接的負載兩端電流方向和放大電流，用以輸出控制電磁鐵的電流
- f. NVIDIA Jetson Nano:運算用的機器，可接收 TOF 感測器的資料和輸出 PWM 訊號給 H 橋

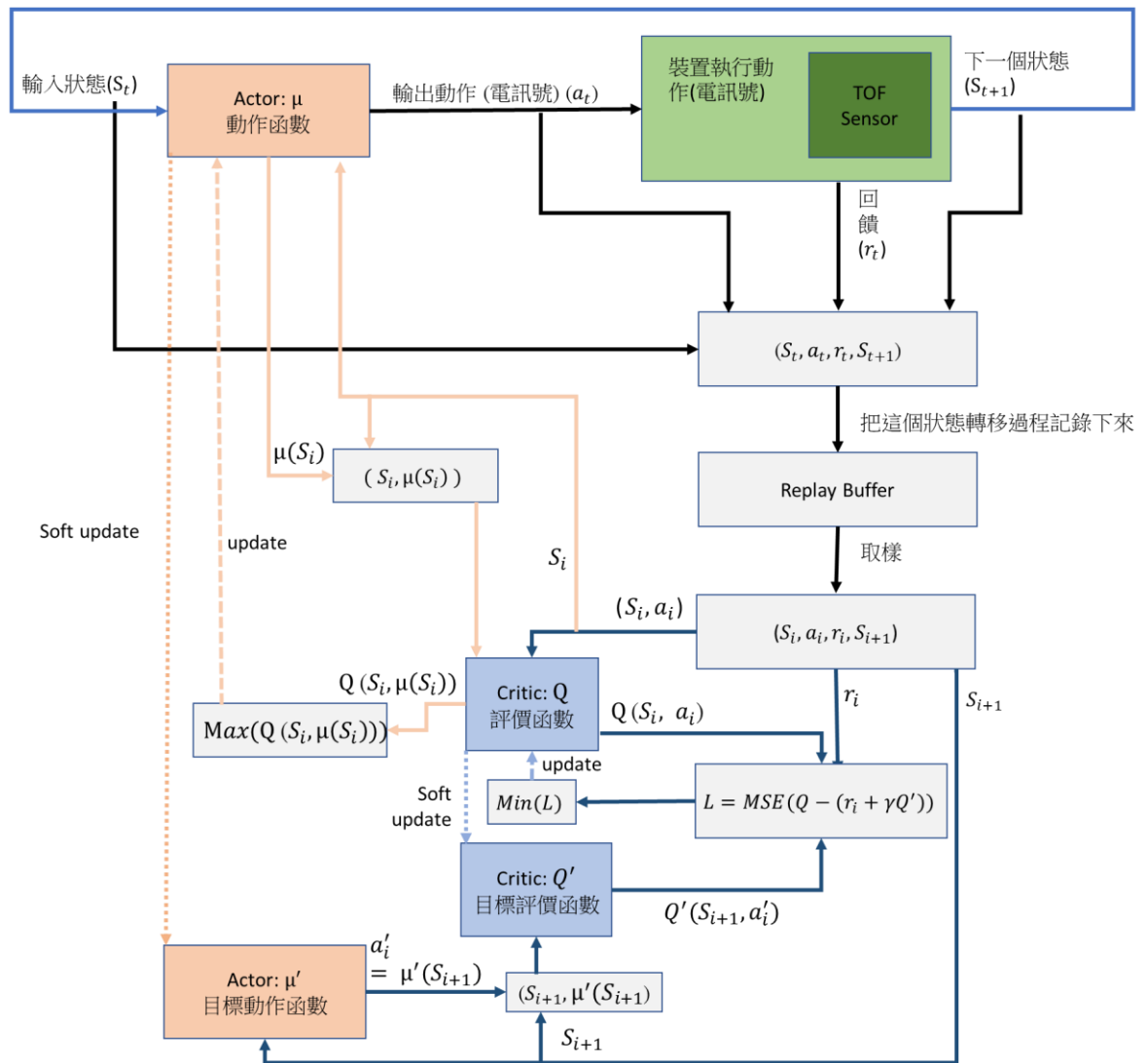
肆、研究過程或方法

一、研究設計

(一) 實驗裝置設計



(二) 總體流程圖



1. μ 定義為actor函數，每一步的行為可以通過 $a_t = \mu(s_t)$ 計算獲得，用一個深度神經網路對 μ 函數模擬，參數以 θ^μ 表示
2. Q 定義為critic函數：用一個深度神經網路對 Q 函數模擬，參數以 θ^Q 表示
3. 透過兩個相同的網路架構但是一個以梯度下降法更新參數，一個以較慢速度更新參數(soft update)，各稱為online和target網路，online網路分別表示為 Q 和 μ ，online網路分別表示為 Q' 和 μ' ，優點為target網路更新速度較慢，因此用於訓練online網路的梯度下降更新參數時，較為穩定，易收斂
4. 裝置執行 a_t ，返回reward(r_t)和新的狀態 s_{t+1}

$$r_t = -(\text{離平衡位置距離}^2 + \text{彈簧速度}^2 + \text{動作}(a_t)\text{大小}^2)$$

(1) 離平衡位置距離 = 當前彈簧與感測器距離 - 在靜止狀態的彈簧與感測器距離 (單位:mm)

(2) 彈簧速度 = (第二次測量的距離 - 第一次測量距離)/(兩次測量的時間差) (單位:mm/s) *註:這裡的第一次和第二次是指每一輪的測量，而非整個過程

(3) 動作大小指得就是 $a_t = \mu(s_t)$ ，因為 GPIO 的 PWM 的佔空比能指定的範圍只有 0 到 100，且要輸出兩種方向的數值，所以 a_t 的數值介於 [-100, 100]

5. 將這個狀態轉移過程: (s_t, a_t, r_t, s_{t+1}) 存入 replay buffer 中
6. 從 replay buffer 中，隨機採樣 32 個狀態轉移數據，作為 online actor (μ) 網路、online critic (Q) 網路的一個 mini-batch 訓練數據。用 (s_i, a_i, r_i, s_{i+1}) 表示 mini-batch 中的單個狀態轉移數據
7. 以 r_i 和平方誤差法計算 online critic 網路的梯度
8. 最小化 Loss

誤差(Loss)定義 L 為:

$$L(\theta^Q) = \text{MSE}(Q(s_i, a_i) - (r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}))))$$

以最小化 $\text{Min}(L(\theta^Q))$ 為目標梯度更新 online critic (Q) 的參數 θ^Q

這個過程有點類似監督式學習:

其中的標籤定義為:

$$r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}))$$

預測值為:

$$Q(s_i, a_i)$$

9. 更新 online actor 的參數 θ^μ :

$$\text{定義 } J(\theta^\mu) = Q(s_i, \mu(s_i))$$

最大化 $\text{Max}(J(\theta^\mu))$

10. soft update:

最後以 online critic 網路和 online actor 網路的參數，用 soft update 更新

target critic 網路和 target actor 網路的參數

θ^μ :online actor 網路的參數

$\theta^{\mu'}$:target actor 網路的參數

θ^Q :online critic 網路的參數

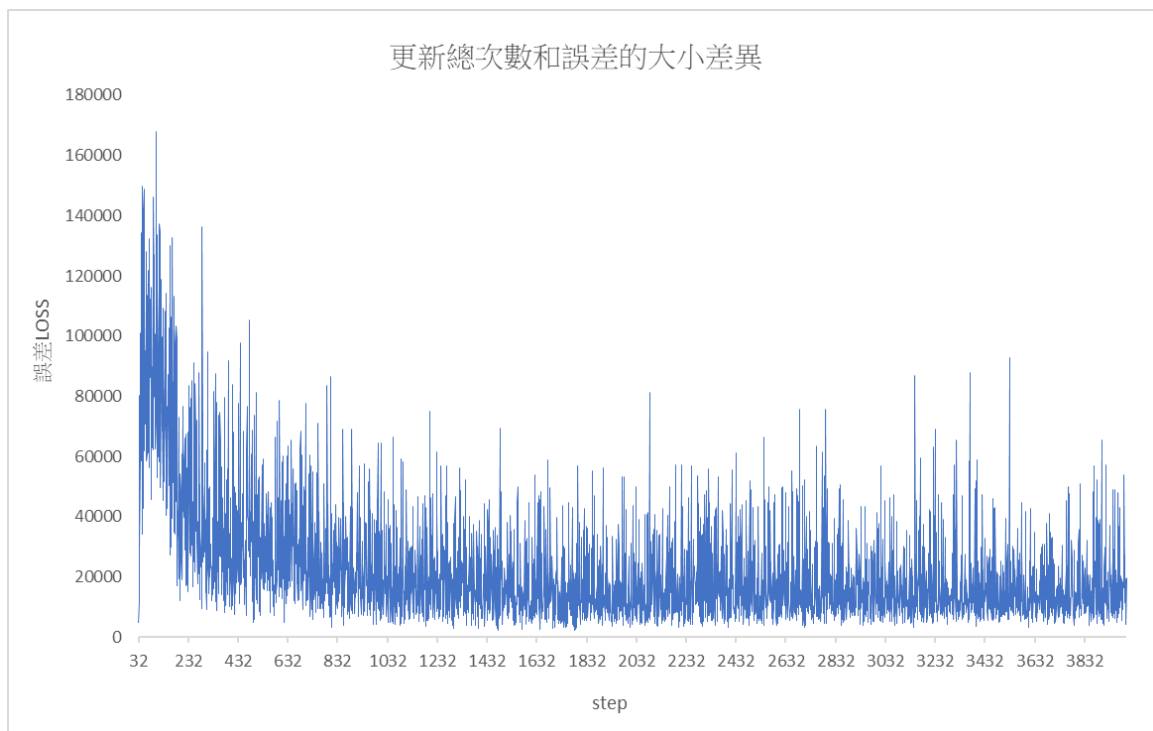
$\theta^{Q'}$:target critic 網路的參數

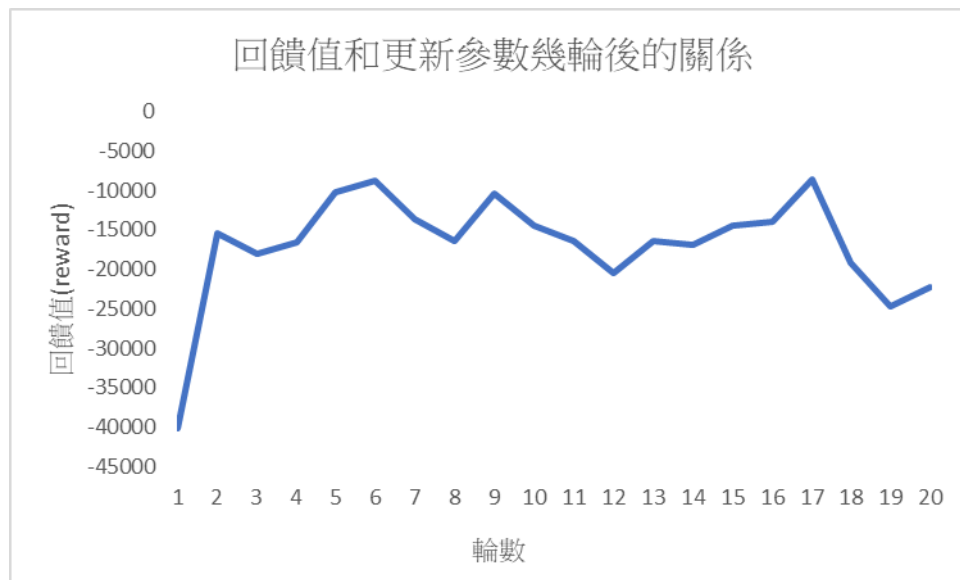
$$\text{soft update:} \begin{cases} \theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'} \\ \theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'} \end{cases}$$

τ 一般取值0.001

- (7) 以 Adam optimizer 更新 online Q 網路和 online actor 網路的參數，
soft update 更新 target Q 網路和 target actor 網路的參數

伍、研究結果





陸、討論

- 一、 橫坐標是從 32 開始是因為每次取樣的批次大小都是 32，因此必須要等到 replay buffer 裡有 32 筆以上(含)的資料才能取樣
- 二、 誤差 $L(\theta^Q) = MSE(Q(S_i, a_i) - (r_i + \gamma Q'(S_{i+1}, \mu'(S_{i+1}))))$ 有收斂的趨勢，顯示這個問題是可以以神經網路模擬控制策略，並經由重複迭代的過程更新參數找出最優解，也因為可以在線更新參數的原因，可以藉由微調大模型參數的技巧來適應不同的環境
- 三、 誤差和環境回饋值會有震盪是因為目前的 actor 網路和 critic 網路所學習到的參數應還是區域最小值的解，要真的達到全域最小值還是需要更多的狀態轉移過程的累積和超參數不斷地才有機會達到

柒、結論

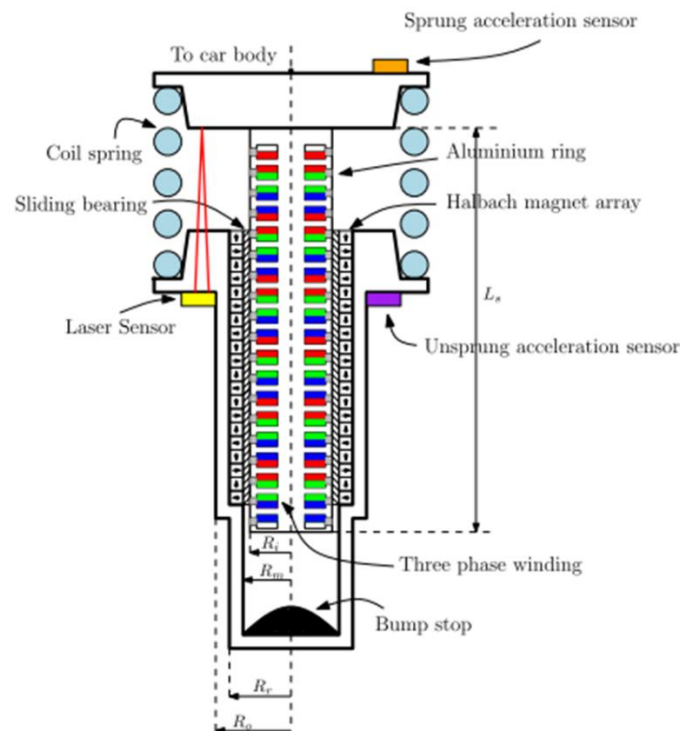
一、實驗結論:

經過在實際環境中測試發現以強化學習算法控制振盪效果有明顯效果，和許多舊的方法相同都是可以快速地使震動停下，但是相較於以往的固定算法，這種可以在線更新參數的方法的自適應能力是其所無法達成的，且在算力的加持下，計算速度越快，反應時間也就越快，因此如果把此裝置部屬在車輛的主動式懸吊上應該可以得到不錯的避震效果，或是使用在防止精密儀器震動上或許也能不錯的效果，除非有一天人類克服了重力

二、未來展望:

未來在實驗裝置的改善方面會把磁鐵以海爾貝克陣列的方式將磁場集中在軸中間，能使

電磁鐵的效率更高，就能使控制時所消耗的能量大幅減少，且因在一個海爾貝克陣列上會有很多方向的磁場，如果對於每個方向的磁場都放置一個電磁鐵，就能控制出更精細的動作，不過同時由於輸出的維度相對多了好幾倍，因此需要累積更多的資料去做參數更新，以及環境的誤差要變小，避免整個模型的參數網錯誤的方向學習



海爾貝克陣列用於避震器中示意圖

捌、參考文獻資料

- 一、 力學振盪之實驗設計與液體黏滯係數的討論
陳毓婷、廖苓卉、楊佳鳳、陳筱青
- 二、 非線性彈簧阻尼系統之振動消能分析
成功大學土木工程學系碩博士班李明彥
- 三、 Active Electromagnetic Suspension System for Improved Vehicle Dynamics
Article in IEEE Transactions on Vehicular Technology · April 2010
https://www.researchgate.net/publication/224091813_Active_Electromagnetic_Suspension_System_for_Improved_Vehicle_Dynamics?enrichId=rgreq-0295153f2b56b6241629921591741c87XXX&enrichSource=Y292ZXJQYWdlOzIyNDA5MTgxMztBUoxMDExMzMxMTY3NzIzNTIAMTQwMTEyMzQyMjUxNg%3D%3D&el=1_x_2&esc=publicationCoverPdf
- 四、 無刷電動機結構設計與特性分析
- 五、 Control of an automotive electromagnetic suspension system
T.P.J. van der Sande D&C 2011.016
<http://mate.tue.nl/mate/pdfs/12819.pdf>
- 六、 楊政瑜、蘇琨祥 國立高雄應用科技大學 電機工程學系

<http://ir.lib.kuas.edu.tw/bitstream/987654321/8881/2/18133851-200809-5-3-425-445-a.pdf>

七、 6. Deep Reinforcement Learning - 1. DDPG 原理和算法

https://blog.csdn.net/kenneth_yu/article/details/78478356