

# 馬里蘭大學黃嘉斌教授演講： Learning to See the 3D World

文／謝宏笙 資訊科學與工程研究所甲組

黃嘉斌教授在交通大學取得學士學位，並在伊利諾大學厄巴納－香檳分校 (University of Illinois Urbana-Champaign, USA) 取得碩士學位及博士學位，目前在馬里蘭大學帕克分校 (University of Maryland College Park, USA) 擔任助理教授一職，並在 Facebook Reality Labs (FRL) 擔任 Research Scientist 一職。

黃嘉斌教授的研究領域涵蓋了電腦視覺、電腦圖學、機器學習。近年來發表超過 100 篇研究論文，其中不乏許多頂級會議論文，並且獲得超過 10000 次的引用數，除此之外，也受到了許多獎項的肯定，在學術領域中表現非常活躍。

黃嘉斌教授於 2021 年 10 月 14 日受人工智能普適中心邀請進行線上演講，演講主題為“Learning to see the 3D world”。雖然以現今 computer vision 的技術已經能夠很好地處理 2D 圖像，但這對於實際上的應用是不夠的，生活在

3D 世界的我們，所需面對的大多為 3D 動態場景，但現今的技術在構建 3D 場景上仍充滿挑戰，在此次演講中，黃嘉斌教授分享了其在此方面的研究成果。

在演講中，黃嘉斌教授分享了他們如何讓圖片中的人像換裝換姿勢，這個問題的困難點在於如何從 2D 圖像取得 3D 場景資訊，對於 2D 圖像我們通常只能看到某個特定角度及特定姿勢的畫面，所以如何從 2D 圖像取得 3D 場景資訊是個關鍵，雖然已經有相關研究能做出不錯的結果，但所生成的圖像難以保留原始圖像中的細節且會產生明顯的偽影，在黃嘉斌教授的研究中，他們透過 pose-conditioned StyleGAN network 生成不同角度的圖像，並將 surface-based method 整合於其中以達到保留圖像細節的目的，使得他們的方法所生成的圖片更為擬真，相較於先前的研究，更為適合應用於現實之中。

# Dr. Jia-Bin Huang (UMD) delivered a speech: Learning to See the 3D World

Professor Jia-Bin Huang earned a bachelor's degree in electronics engineering from National Chiao Tung University in Taiwan, as well as a Master's degree and a PhD degree in electrical and computer engineering at the University of Illinois at Urbana-Champaign. Dr. Huang is an assistant professor at the University of Maryland at College Park and a Research Scientist at Facebook Reality Labs (FRL).

Dr. Huang's research interests cover the areas of computer vision, computer graphics, and machine learning. He has published more than 100 research papers these years, which include top-tier conference papers, and has exceeded 10,000 citations. In addition, he has been recognized with numerous awards while being active in academic society.

Dr. Huang was invited by the Pervasive Artificial Intelligence Research Center at NYCU to give an online speech on the topic of "Learning to see the 3D world" on October 14, 2021. Although existing computer vision technology nowadays is able to process 2D images intelligently, it cannot meet the practical requirement of

daily life. Living in a 3D world, we constantly encounter 3D dynamic scenes; However, the construction of 3D scenes is still full of challenges for today's technology. In this speech, Professor Huang shared his research achievements in this regard.

During the speech, Professor Huang illustrated their approach to change the clothes and poses of the portraits in pictures. The real difficulty lies in retrieving the corresponding 3D models from 2D images. A 2D image only comprises the information of objects with certain angles and specific poses; therefore, the core technology is to reconstruct a 3D scene from 2D images. Although many related researches have obtained good results, the reconstructed images may lose some details of the original images and leave obvious reconstruction artifacts. In Professor Huang's research, they constructed images from different angles through the pose-conditioned StyleGAN network that also integrates the surface-based method to preserve image details so that the images generated by their method would be more realistic and applicable to the real scenarios than other research.

