







# 中國大陸新聞傳播赴台交流陸生對台態度 與涉台新聞媒體識讀之影響分析<sup>1</sup>

王毓莉<sup>2</sup>

## 《摘要》

過去缺乏對於陸生返回中國大陸之後的後續研究，本研究從接觸理論視角切入，深度訪談 2007~2019 年間來台從事新聞傳播深度交流的陸生，研究發現參與交流活動後對台態度與印象，均有正面的提升，在台生活接觸體驗，印證了台灣言論與文化藝術的自由度。一個月的「新聞媒體實習」，提升了對於台灣媒體新聞框架實踐操作的了解，也深化對台灣新聞記者工作場域自由度與專業化表現的體驗和認識，對於日後疑似台灣假新聞的識讀能力，亦有明顯助益。

關鍵字：中國大陸交流陸生、兩岸交流、接觸理論、媒體識讀、對台態度變遷

---

<sup>1</sup> 感謝科技部贊助，本文資料為「凡走過的必留下痕跡？新聞傳播交流陸生涉入程度與對台態度之研究」畫（編號 MOST 108-2410-H-034 -040 -MY2）的部份研究成果。更感謝匿名評審，給予寶貴意見，使得本文更加完善

<sup>2</sup> 中國文化大學大眾傳播學系教授，e-mail: ylwang1022@gmail.com

## 壹、研究背景與研究意識

兩岸交流始自1987年，前後開放國人至中國大陸探親、台商赴陸投資，陸客來台觀光、以及陸生來台就讀。兩岸關係歷經了1990年代的熱絡發展，儘管到2000年民進黨陳水扁總統首次執政後，有所減退，到了2008年後國民黨馬英九重新執政後，兩岸經貿交流頻繁，再創兩岸關係的高峰。然而，2012年底中國大陸第五代領導人習近平上任，到2016年台灣民進黨蔡英文執政後，兩岸關係發展進入急凍期，兩岸官方接觸管道全面中斷，兩岸在國際社會中和解休兵現象淡化，陸客來台觀光人數明顯下降（邵宗海，2017；中華民國大陸委員會，2018）。

一般文教交流、大眾傳播、科技研究、宗教活動皆呈現逐年增加趨勢。其中對於青年學子的交流活動舉辦和招收學位生政策，均隱含了台灣政府制定政策背後若干目的，除了期望促進台灣教育多元化外，也希望透過兩岸青年學生交流與學習，相互競爭求取進步，提升台灣學生競爭力和擴展視野，期使陸生在台學期間，可以體驗台灣多元文化發展，並分享中國大陸的經驗，增進兩岸青年學生彼此認識與瞭解，有助於兩岸關係長期的和平與穩定（中華民國大陸委員會，2018年4月16日）。

然而自二次政黨輪替後，兩岸學術交流方面，也產生極大的變化。據陸委會統計2016年來台陸籍研修生人數共有32,648人，然而，自2017年起，來台陸籍研修生人數逐年驟減，2017年共25,824人，2018年共20,597人，2019年共16,696人，2019年來台陸籍研修生人數較2016年驟減約49%，另外，2020年由於新冠疫情暫緩研修生入境，故只有4位陸籍研修生入境（中華民國大陸委員會，2021年2月25日），顯見兩岸交流的大倒退。

一般而言，來台的陸生種類可區分為兩大類：第一類為「學位生」，依照政府陸生來台政策即可來台就讀學士班、碩士班或博士班，並取得台灣大專院校的學歷。第二類則為「短期研修生」，相較於學位生受到「三限六不」的政策管制，短期研修生屬於兩岸校際間交流，由大專院校向教育部報備即可，因此短期研修生並非以獲取台灣學位為目的，而是修習學分為主的遊學性質，來台期間約為一學期或二學期（高順德，2015年9月8日）。

除上述正式學制內交流外，其他學術交流的活動類型不少，但多為短期營隊交流，例如：世新大學和中國新聞學會合作，舉辦「海峽兩岸大學生新聞營」，藉由兩岸大學生共同參與交流、座談、採訪名勝、分組編報等活動，盼能增進兩

岸教育的直接交流，行程大約7-12天。較為特別的是，自2006年由陸委會出資，以「中華發展基金」為名義開始舉辦的「兩岸大眾傳播研究生交流活動」，以邀請中國大陸新聞傳播研究生為參與對象，活動最大特色是，研究生來台時間為期兩個月，除了一個月的專業講座課程、參訪與觀光外，另一個月則安排到台灣的新聞媒體從事實習工作，屬於一種相對深度的實習交流。陸委會委辦此一活動主要目的，是為了促進中國大陸大眾傳播研究生瞭解台灣民主社會新聞自由，媒體自主價值，傳遞台灣軟實力。

近幾年兩岸關係進入冰河時期，中國大陸青年學子接收訊息方式，亦有了大的變化。陸生對於台灣的印象，除了親身體驗的交流互動外，長期以來兩岸對於對岸的形象認知研究顯示，新聞資訊扮演建構對岸形象的重要框架角色。檢視過去陸生交流相關研究，大多集中探討短期研習陸生或學位生，對於活動類型的交流，缺乏長期和深度的影響性研究。

因此，本文最大的貢獻在於，期望針對前述為期兩個月深度交流的「兩岸大眾傳播研究生交流活動」對象進行跨度12年的研究，企圖了解該活動如何影響不同時期參與者對台灣的認知與態度，而這樣的影響，對於其回到中國大陸後，再度接觸到疑似台灣假新聞時，其判讀能力能否有所提昇？

## 貳、文獻探討

檢視過去陸生交流相關研究，大多集中探討來台交流期間的來台交換陸生的就學動機與生活適應議題（蕭伊婷，2009；韓宜娟，2011；Chou & Ching, 2015；楊宏文，2015；周宛青，2017）、大陸留學生在台的社會認同與族群國家認同問題（李佩雯，2014；張可，2019）和對短期陸生交換生的政治態度、統一及統獨立場的變遷研究（王嘉州，2015，2016）。

本文無意探討來台陸生認同或統獨等政治議題，而是期望了解此一特殊活動設計，自2006年以來對於參與者產生的影響。以下檢視本文涉及相關理論與文獻，並發展研究假設與問題。

### 一、框架、接觸和涉入理論與對台態度

Lippmann（1922，林珊譯，1989）在其著作《民意》一書中指出，此一世界有太多繁複及細微的事物，是人們無法親身體驗，但為了因應外在環境變化，大部分的時候，我們是先下定義，再做觀察，而此觀察會引用認知文化中已經為我



其次，心理學者Allport（1954）提出「接觸理論」（contact hypothesis），指出兩個不同團體成員的相互接觸，將帶來正面效應，有助降低對彼此既存的刻板印象與偏見，當然也助於減少衝突發生機率。Pettigrew et al.（2011）以515份相關的研究、以及總數超過250,000受試者的資料，證實在一般狀況下，「團體接觸」的確會帶來破除刻板印象、降低偏見的正面效果。這項正面效果的產生，往往是因為不同團體成員在相互接觸之後，能夠學習瞭解外團體成員的思維想法，也瞭解到外團體成員各有不同，並非都是如刻板印象所描述那般，因此也就去掉了「我群」原先對「他群」的偏狹印象，具有一種「去偏狹化」（deprovincialization）的功能；此外，持續的互動接觸，也有助於建立不同團體間的情感連結、增進友誼，發揮降低偏見、減少衝突的正面效果（Pettigrew, 1998, pp.70-73）。教育學者也認為，可從「兩岸化」概念與族群接觸理論，探討兩岸大學生的交流（周祝瑛、楊雁斐，2018）。

至於接觸的程度方面，「涉入理論」（Involve Theory）常用於行為研究，Zaichkowsky（1985）曾定義涉入為一個人基於內在的需要、價值及興趣，進而體認到廣告的關連性。這個定義值得注意的地方，在於其強調情感與認知的關連性，其意涵也指向了涉入程度對於行為決策的影響力。Zaichkowsky（1986, pp. 5-6）指出，涉入程度有三個前趨變項，包括個人的特質（person factors）、刺激物的特質（stimulus factors）與情境特質（situational factors）。

援引上述概念，就本研究探討的交流陸生而言，為期兩個月的課程及媒體實習，有了更多與台灣社會接觸的機會，在某種程度上，代表的是長時間情境涉入。相較於只參與交流而沒有媒體實習經驗的陸生，這些陸生能接觸到更多的台灣新聞工作者、政府部門、商業企業或娛樂產業，尤其部份高度政治性的部門，如行政院、立法院等，他們擁有了不一樣的體認以及更深入的情感交流。

Wang（2021, 2022）曾經以接觸理論視角，針對台灣赴中國大陸交流的學生，了解其接觸中國大陸後，是否會改變其對於統獨問題或民主的評價？由於研究對象集中在返台後的台灣學生意見，較能觸及相對敏感問題。本文嘗試補足兩岸學生交流的另一方改變，期望針對2007年至2019年來台之新聞傳播交流陸生，提出研究問題一：參與新聞傳播交流活動前後，對台資訊來源與印象的轉變為何？



事實上，兩岸的社群媒體當中，也常有假新聞流竄，例如：2018 年 9 月所發生的「關西機場事件」，當時因颱風襲擊關西機場，導致多國旅客受困於此，中國大陸微博網民「洪水猛獸 baby」最先發布貼文，意指中國派專車進入機場，在第一時間接送中國遊客離開，這則帖子馬上被中國大陸《觀察者網》製作成網路新聞，內容中還強調專車是中國領事館與日方協調的成果，報導最後提及，若台灣旅客願意「自稱為中國人」，即可搭乘專車。而後 PTT 論壇及台灣各家媒體都報導此事，但最後此事被證實為子虛烏有的事情（江旻諺、吳介民，2020 年 1 月 24 日）。

不同研究也面向新媒體中假新聞的傳播行為，或針對辨識假新聞，進行研究（Aldwairi & Alwahedi, 2018；Shu, et al., 2020；葉乃靜，2020）。哈佛大學圖書館（Harvard Library）曾提出 5 種判讀方法，分別為注意資料來源、檢查新聞的 URL、從外觀找線索、尋求其他意見、善用瀏覽器功能（Harvard Library, 2017）。國際圖書館協會聯合會（IFLA）<sup>4</sup>也針對假新聞防治整理出 8 個判別方法：考慮訊息來源、檢查文章作者、檢查文章日期、檢查自身的偏見、閱讀標題以外的內容、訊息來源的支持、是否只是一則笑話、詢問專家（International Federation of Library Associations, 2016）。黃宇弘（2019）則整理出 12 項判讀假新聞的原則考慮新聞來源，分別為檢查作者、檢查日期、自我反省、看到文字以外的意圖、參考資料、諷刺言論、網站外觀、文字排版、參考其他相關報導、檢查新聞的 URL、善用瀏覽器功能、求助專家。

綜合上述本研究擬借用假新聞判讀指標，探討這些新聞傳播專業的交流陸生，來台經過媒體實習及體驗台灣生活，回到中國大陸後對於接觸到疑似台灣相關的假新聞判讀能力是否會產生轉變？故提出研究問題二：參與新聞傳播交流活動後，對疑似台灣假新聞的媒體識讀能力影響為何？

### 參、研究問題與方法

本文企圖考察新聞傳播交流陸生涉入程度，在對台態度和識讀台灣假新聞方面的影響，研究問題如下：

---

<sup>4</sup> 國際圖書館協會聯合會（IFLA），簡稱「國際圖聯」，成立於 1927 年，總部設在荷蘭海牙。它是聯合各國圖書館協會、學會共同組成的一個機構，是世界圖書館界最具權威、最有影響的非政府的專業性國際組織。其鼓勵大眾廣泛了解圖書館提供優質資訊服務的價值。



































們對台灣的支持，因此，此交流項目在某種程度有效促進參與者對台灣的支持。

本文呼應上述對於短期來台交流生的研究發現，這些來台交流的新聞傳播陸生，不僅只在台接觸新聞媒體報導，更有機會透過一個月親赴台灣新聞媒體實習，不僅讓陸生對台灣的新聞媒體產業有更深入的了解和認識，可以更立體的理解台灣社會，對於去除對台的刻板印象，提升好感度，有所助益。對於部分受訪者原有價值觀甚至產生刺激，開始思考深層問題，進而影響他們對未來職涯規劃的選擇。

前揭文乃針對陸生在台時的調查，然而訪台的陸生終究要回到中國大陸生活，本文旨趣在於了解，此一交流活動對於其內心深遠的歷時性影響，是否會帶回中國大陸，在未來解讀疑似台灣假新聞時，能夠具備更好的抵禦能力，甚或幫助台灣澄清假訊息。

返回中國大陸的新聞傳播交流生強調，課堂中學習到「新聞框架」(news framing)理論，在台實習期間印證，發現不少台灣媒體選出的都是大陸人虐貓、虐狗等負面題材，引導台灣民眾對中國大陸產生較為負面的刻板印象，但是台灣民眾沒有資訊被封鎖問題。不過，在回到中國大陸接觸對台的報導時，由於現階段中國大陸強調媒體的黨性原則，一旦出現與政黨立場不同或是危害到國家形象等，政府宣傳部門便會採取一切手段封鎖消息。因此報導涉台時，必須符合國家政策。受訪者無奈也諷刺地說：「基本上可能《人民日報》怎麼說，你就這麼信吧！因為也找不到其他的報導了」。

值得注意的是，經過上述媒體資訊環境的洗禮，中國大陸的 90、00 後並不覺得自己言論不自由，因為多數人認為依然私下仍可暢所欲言，對於翻牆到牆外找訊息來源查證，早已不感興趣。曾來參加交流活動的 70~80 後憂心地指出：「現在的年輕人已經不相信西方那一套了，覺得 BBC 都是扯蛋，報導的準確性根本不如大陸的官方媒體。」好比這次 Covid-19 疫情過後，他們認為整個西方的國家管理體制，都不如中國。

整體而言，由於目前兩岸關係處於谷底狀況，交流陸生願意為台激烈辯論者少，但願意分享的人依然為多數，從主辦此類型活動的長遠效果角度來說，具有一定的效果。檢視這些交流陸生進入社會工作後，進入多元的工作領域，包括：新聞媒體實務工作、教職、網路電商、有的則選擇其他產業，結合其原先新聞傳播專業知識，加上來台交流之經驗，對於疑似台灣假新聞具有更高訊息辨別能力，



## 參考文獻

- ETtoday 新聞雲大陸中心 (2017 年 7 月 4 日)。因負面新聞...「茶葉蛋教授」大陸演講取消 網友繼續酸：快笑。**ETtoday 新聞雲**。取自 <https://www.ettoday.net/news/20170704/959364.htm>
- 人民日報 (2013 年 8 月 16 日)。人民日報：中國互聯網大會倡議共守“七條底線”。**人民網**。取自 <http://media.people.com.cn/BIG5/n/2013/0816/c40606-22584001.html>
- 中華民國大陸委員會 (2018)。兩岸經濟統計月報，**經貿交流統計資料**。307：2-18。取自 [https://www.mac.gov.tw/News\\_Content.aspx?n=2C28D363038C300F&sms=231F60B3498BBB19&s=88CF4AB14A7B4BBD](https://www.mac.gov.tw/News_Content.aspx?n=2C28D363038C300F&sms=231F60B3498BBB19&s=88CF4AB14A7B4BBD)
- 中華民國大陸委員會 (2018 年 4 月 16 日)。政府開放陸生來臺就學的政策目的及配套措施為何？，**網站答客問**。取自 [https://www.mac.gov.tw/News\\_Content.aspx?n=F9057F9640B28033&sms=7C8440BC86E48FD9&s=9071E9C063DBA01D](https://www.mac.gov.tw/News_Content.aspx?n=F9057F9640B28033&sms=7C8440BC86E48FD9&s=9071E9C063DBA01D)
- 中華民國大陸委員會 (2021 年 2 月 25 日)。陸生來臺研修及修讀學位統計，**陸生來臺就學專區**。取自 <https://www.mac.gov.tw/cp.aspx?n=A3C17A7A26BAB048>
- 王彥 (2021)。**媒介框架理論的前世、今生與未來：華人傳播學術社群的追古溯今**。政治大學傳播學院博士論文。
- 王泰俐 (2019)。假新聞與民主危機。**台灣民主季刊**，**16(3)**:155-161。
- 王毓莉 (2018)。高校生新聞資訊接受之研究：以上海 211 工程高校為例。**傳播管理學刊**，**19(2)**: 1-36。
- 王毓莉 (2021)。傳統媒體、網路媒體與政府治理，王信賢、寇健文 (編)，**中國大陸概論 (第二版)**，頁 267-301。五南。
- 王嘉州 (2015)。交流生共識？赴陸台生統獨立場之變遷。**東亞研究**，**46(1)**: 1-33。
- 王嘉州 (2016)。短期來臺交流陸生之社會接觸與政治態度變遷初探。**遠景基金會季刊**，**17(4)**:1-42。
- 江旻諺、吳介民 (2020 年 1 月 24 日)。「戰狼主旋律」變形入台，解析關西機場事件的中國虛假資訊鏈。**風傳媒**。取自

















syuzhet

fabula

style

Kuhn and Schmidt, 2014

Seymour Chatman

Chatman

1980

" Story and Discourse

Narrative Structure in Fiction and Film "

narrative

a story

a discourse

what

how

Chatman, 1980 19

events

action

happening

exists

character

settings

discourse

form

the structure of narrative transmission

manifestation

..... Chatman, 1980 23 1

1 Chatman

---

Expression

Content

---

Substance

---

Form

---

Chatman 1980 24

Chatman 進一步以符學觀點說明敘事架構的關係，指出敘事就是傳達 (Narratives are Communications)，被傳達的是故事 story (是敘事的內容元素形式)，而且是透過敘事論述 discourse (是敘事的表現元素形式) 來傳達，敘事論述 discourse 是用來陳述故事 (Chatman, 1980 : 31)。結合前述的敘事概念與符號學，Chatman 更新敘事架構的概念如圖 1：

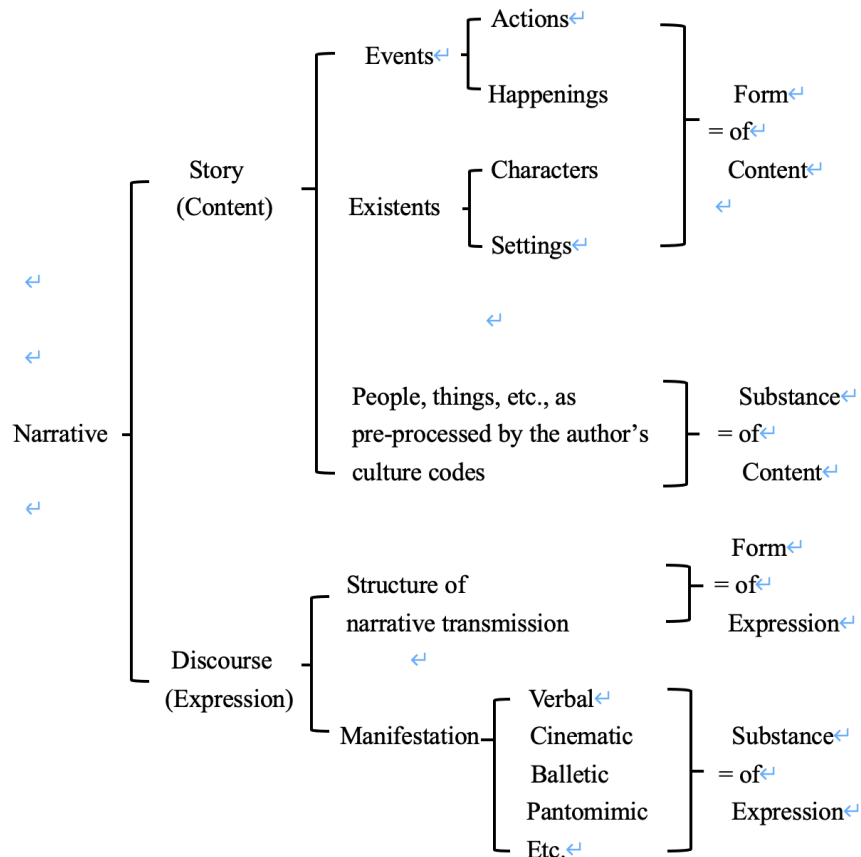


圖 1 Chatman 的敘事架構圖 (Chatman, 1980 : 26)

基於上述關於敘事理論及敘事與電影的關係初步討論，George Wilson(1976)曾說，電影中每一個鏡頭都代表一個角色的特定視覺場域……或可說電影雖然帶給我們由一系列不同視角顯現出來的世界，但這些也只是呈現出某人在某一限定狀況下的觀點。這些影像元素如故事結構與組成、事件排列順序、造型、時間長短、節奏快慢、顏色區別、聲音、敘述者(narrator)的傳達、演員的表演……等，都影響我們去理解或體會這一段文章或影像所要告訴我們的故事。

影片是一種傳達故事的媒體，動畫或偶動畫影片與這些媒介一樣，藉由設計虛擬的動畫角色，歷經各種事件後塑造影響角色成長的歷程，達到告訴觀眾一段故事的目的。敘事不只是文字本文的描述，精心安排與設計的敘事元素亦可傳達很多的意義，透過影像的靜態或動態表現、顏色渲染與氣氛營造、情節的節奏調整、角色的生動表演等都可以產生傳達的效用。本研究認為電影不論用何種敘事結構或視覺風格的表現及技術的承載也都是為了要使目標的觀眾可以接收到並理解創作者要傳達的內容。

因此，本研究以霹靂布袋戲偶動漫《奇人密碼--古羅布之謎》為例，藉由Chatman敘事理論的脈絡，從鏡頭設計探討使用定鏡在動畫影片中的敘事表達形式(form of expression)及效果。在研究定鏡之前，本研究先蒐集關於鏡頭的定義與用途，以及專業術語所表達的不同運鏡方式並據以為影片記錄相關參數，如此便能幫助了解研究的對象及各種鏡頭設計對敘事表達的差異性。

## 二、攝影機運鏡的概念

### (一) 鏡頭方向對視覺畫面之影響

拍攝方向是以被攝主體為中心，以圍繞並針對被攝體的方式，在範圍內選擇拍攝點(吳冠英、祝卉，2005)。鏡頭方向的一個普遍運用是動作線(亦可稱為180度假想線原則)，主體從第一個畫面經過後，第二個畫面也會以相同的方向出現，如此便不會造成觀眾對於銀幕方向的錯亂(井迎兆，2002)。鏡頭方向也有影響畫面氛圍的功能，例如正面方向會削弱空間感和透視，使畫面顯得呆板，但也能呈現對稱與和諧的美感。而在表現主體運動時，從側面方向拍攝的動態線條最為豐富。背面方向除了有冷漠、抗拒等情緒氛圍以外，也能製造懸疑、緊張的感受。斜側面方向能使畫面生動活潑，利於展現主體的立體型態與空間透視感，也是角色彼此交流時的常用拍攝方向，因為對話時角色有固定的方向性，藉由鏡頭切換就可以表現完整的鏡頭空間(張宇晴、鐘世凱，2009)。如表2：

表 2 鏡頭方向之定義

類型	定義
正面方向	拍攝角色或景物的正面。
背面方向	拍攝角色或景物的背面。
側面方向	拍攝角色或景物的側面。
斜側面方向	斜拍角色或景物的正面或背面。

資料來源：本研究整理

### (二) 鏡頭之景別的用途

鏡頭尺寸的大小又稱為景別，其用途在於表現空間場景與人物關係，是讓動作調度得以展現的常見重要形式，而景別的辨別、定義方法相當主觀，因為其定義來自於辨別攝影機與拍攝主體的間距（鐘世凱、張宇晴，2009）。此外，鏡頭尺寸愈大，畫面中可以傳達的環境訊息愈多；尺寸較大的景別（如極遠景、遠景）可以讓觀眾理解故事發展的環境，而鏡頭尺寸較小的景別（如特寫、極特寫）通常用以強調拍攝主體的表演（鐘世凱、張宇晴，2009）。如表 3：

表 3 鏡頭景別用途之定義

類型	定義
大極遠景	景物為鏡頭中的主要拍攝對象，角色則顯得較為不明顯。
遠景	與極遠鏡類似，但主要拍攝的對象是角色還是景物是以鏡頭中的構圖決定。
全景	鏡頭可以剛好涵括角色的全身，能明確辨識角色、空間、方位等要素。
近景	從角色的胸口以上到頭頂以下，占鏡頭總畫面的一半。

資料來源：本研究整理

### (三) 鏡頭角度對觀眾之影響

鏡頭的角度是取決於攝影機與拍攝主體的相對位置，能營造景物的透視變化，讓觀眾更能融入故事的氛圍與情感，角度的取決必須視場景與空間、光影和色彩、拍攝主體、前後連戲等畫面要素設定，在心理層面的方面可藉由鏡頭角度表現角色情緒或狀態（鐘世凱、張宇晴，2009）。妥善運用攝影機能增加故事變化和影

響力，鏡頭的角度決定了觀眾將從何種觀點去接收故事，必然會影響到觀眾對故事的理解，因此在選擇攝影機角度時，搭配故事劇情的發展和變化去設計是很重要的，如果能夠選擇適當的鏡頭角度去描述故事，則能夠增加畫面對觀眾的影響力。如表 4：

表 4 鏡頭角度之定義

類型	定義
俯視鏡頭	從水平線的上方向下拍攝。
水平鏡頭	從與被攝對象水平的位置來拍攝被攝對象。
仰式鏡頭	從水平線下方往上拍攝。

資料來源：本研究整理

#### (四) 攝影機之運動的種類

電影的敘事當中可以透過主觀的人物視角，以及客觀的攝影機視角兩種視角關係描述劇情，而此「視角關係」即是結構主義敘事學影響電影敘事學的因素之一。攝影鏡頭是在同一時間範圍內，由一整段連續畫面和延續的動作組成，為組成電影的基本單位。一個鏡頭就能傳達許多的視覺訊息，而多鏡頭的組合又可產生不同意義或說明因果關係，可見鏡頭對於電影敘事的重要地位。如表 5：

表 5 鏡頭運動的種類之定義

類型	定義
橫搖鏡頭	攝影機平行橫向移動，方向為由左向右，或者是由右向左。
升降鏡頭	攝影機由上至下或者由下至上的垂直移動。
變焦鏡頭	在定鏡的基礎下調整焦距，分為放大景物或者往前推進的（Zoom-in），縮小景物或者往後退的（Zoom-out）。
移動鏡頭	攝影機接近或遠離被攝主體，或是跟著被攝主體移動。

資料來源：本研究整理

### 參、文本分析法

國內電影學者齊隆任認為電影導演是以編碼方式進行工作，而影片的分析者

則是以解碼方式進行工作，影片分析如同文本分析的描寫之模式，以重建分鏡為主要手段，並盡可能記錄其鏡頭之各項組合內容，其包含鏡頭的編號、時間長短、角度、大小比例、運動、對白、音樂、音效等要素（齊隆任，2013）。因此，本研究擬定以下的研究方法與步驟。

#### 一、研究方法與步驟

在文獻部份提到，運用不同的鏡頭的類別，是在分析作品時記錄數據資料的明確標，因此，本研究主要的研究方法是以《奇人密碼--古羅布之謎》偶動漫畫電影之「起、承、轉、合」基本劇情作為觀察對象，觀察樣本的每個鏡頭，並以逐一檢視、分析，記錄其鏡頭的使用方式，研究每個鏡頭運用的劇情段落，出場角色的特性以及切換鏡頭的秒數、鏡頭角度和景別。以藉由觀察攝影機鏡頭的運用方式，理解個別鏡頭對劇情、角色的特性和理解劇情等要素之影響。

本研究採用內容分析法，針對《奇人密碼--古羅布之謎》偶動漫畫電影之「起、承、轉、合」基本劇情分析並以記錄個段落的各項內容參數。從 DVD 影片中採取逐段逐格的觀察方式、記錄劇情段落、鏡頭時間、景別、角度、方向、鏡頭運動等資料，這些鏡頭參數對於一部影片的劇情傳遞與敘事風格影響極大。而在傳達劇情的部份，大遠景鏡頭可以讓觀眾了解場景之空間、環境以及表現情感的昇華，近景鏡頭用途則是傳達角色行動表現的過程。構圖與景別的變化搭配是可以讓影片產生視覺特效之節奏變化。鏡頭角度則是會影響到觀眾對於被攝角色的觀點、場景呈現的視角，也會影響場景空間距離之視覺效果感受（吳冠英、祝卉，2005）。其研究流程如圖 2：

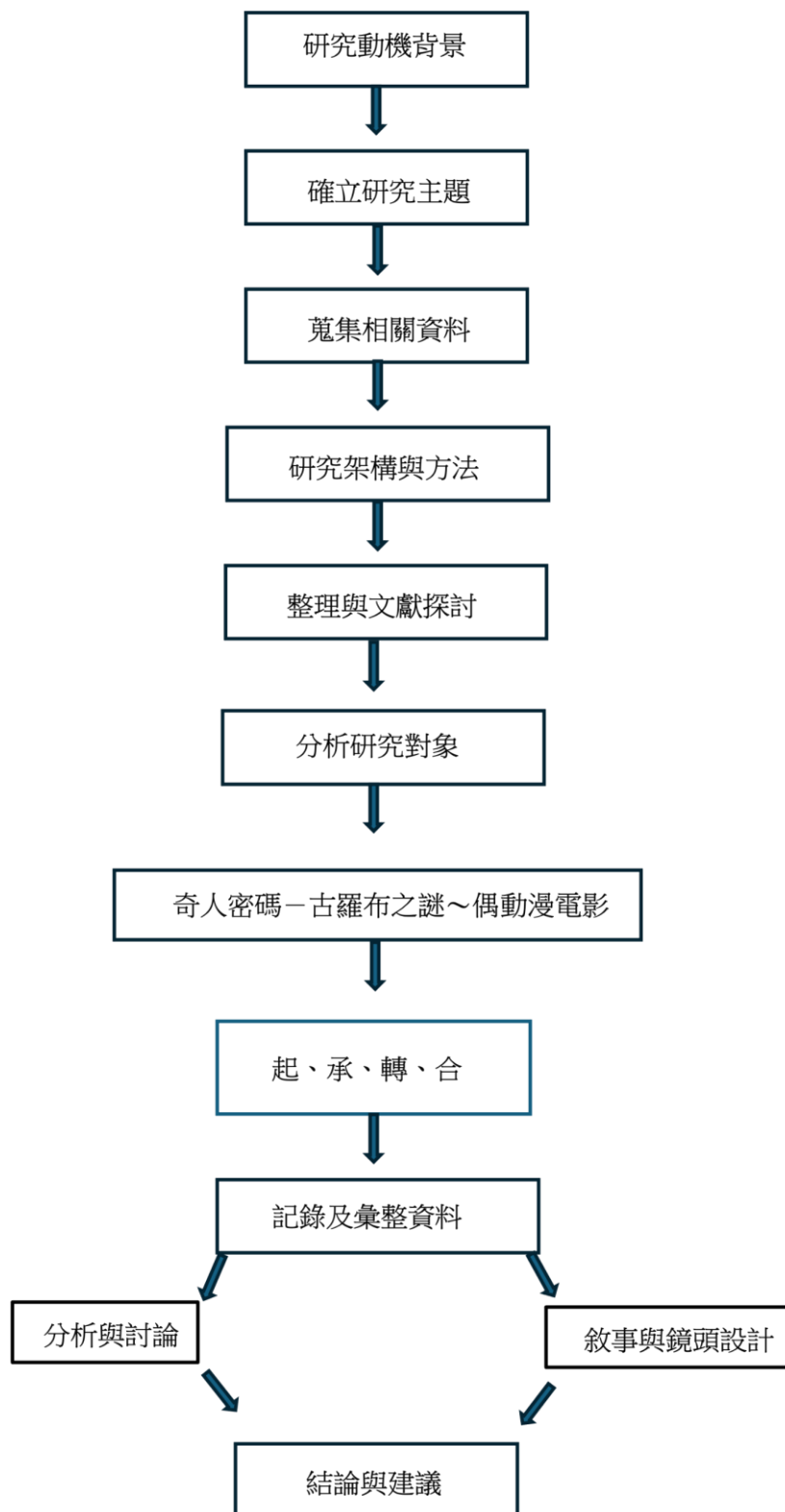


圖 2 研究流程圖

## 二、研究樣本

《奇人密碼—古羅布之謎》是霹靂集團耗時三年製作出的華人首部 3D 史詩偶動漫電影，霹靂工作團隊在製作上結合視覺特效、虛擬場景、3D 動腦動畫、木偶師操作等，並以利用多媒體動畫的概念，來呈現布袋戲。該片導演則堅持要以「偶戲」作為基本理念，因此，所有的角色，都是用人手持去拍攝的，再加上動畫及視覺特效，但也為了要保持「偶戲」的概念，人偶的眼睛、嘴唇與動作，其實是不太自然的。對於看慣了卡通動漫的觀眾來說，人偶無法對嘴每個口白，也沒有嘴唇的任何動作，眼神當然也不會移動。在傳統與創新上，為了保持「偶戲」的初衷，對於呈現大螢幕也有其限制。至今的台灣動畫產業技術已十分精良，導演堅持使用真的人偶操作，也是堅持只把背景與場景換成電腦特效。在人偶衣著的呈現和做工，都展現了霹靂精緻的工藝，因電影部分有特寫部份，觀眾很容易看見霹靂為每個角色打造的形象都那樣真實而風格突出。

由於《奇人密碼—古羅布之謎》偶動漫電影之「起、承、轉、合」基本劇情在鏡頭的表現方面，各段落皆大量使用定鏡作為敘事的主要拍攝鏡頭，符合本研究之主題，更可以提供本研究豐富的數據作為分析與鏡頭設計敘事的基礎。而這些段落角色、場景、環境等特徵是否會影響鏡頭設計時的運用考量，更是本研究希望從數據資料中探尋之目的。本研究是依據 DVD 選單段落為基準法則，將《奇人密碼—古羅布之謎》之「起、承、轉、合」基本劇情偶動漫電影分為四大段落，其說明如表 6：

表 6 《奇人密碼—古羅布之謎》之「起、承、轉、合」內容介紹

編號	段落	角色	長度	劇情內容簡介
1	起	張墨 張彤 阿西	4 分 28 秒	身為開啟絲路之門的英雄漢使-張騫後代的張墨與張彤兄妹二人，帶著以「萬鵲」為動力的-木人「阿西」，又稱「奇人」，循著西遊神書裡的古老傳說來到西域，希望能解開神秘礦石及奇人的身世之謎！一行人探索西域的奇幻冒險之旅，就此翻開序章...

編號	段落	角色	長度	劇情內容簡介
2	承	樓蘭王子 安羅伽 阿西	4 分 06 秒	與此同時，一位神秘的樓蘭王子安羅伽發現了父親遺留下來的筆記並得知了奇人的力量，他極力想要找出奇人的秘密，將奇人技術運用來強化樓蘭的軍事武力，以建立一支奇人軍隊，掌權稱王！安羅伽在得知阿西及萬鵠的存在之後，便開始一連串的野心計畫。
3	轉	安羅伽 張墨 張彤	4 分 54 秒	樓蘭王子趁機救了張彤，同時將所有的災難嫁禍於羅布族人及殺蟲的迫害，而掌握阿西生命之源的萬鵠也被羅布族人操控著，在取得了張墨張彤兄妹的信任後，安羅伽便開始一連串的野心計畫，讓張墨帶著阿西參加樓蘭王舉辦的武鬥大賽，以取得樓蘭王的幫助消滅羅布族，在此轉變。
4	合	張墨 阿西 張彤	3 分 58 秒	對於被綁走妹妹的張墨，對羅布族的誤解卻更深，決定搶回妹妹並找到萬鵠讓阿西得以延續，便引著樓蘭軍隊以及一同攻打羅布族。進入古羅布後，發現那彷彿一個人間聖地，而羅布族的居民及沙蟲為了讓萬物生命之源-萬鵠得以永續生存，世世代代的捍衛著這個人間淨土。

表 7 (續) 資料來源：本研究整理

## 肆、結果與討論

### 一、以《奇人密碼—古羅布之謎》之「起」為記錄資料範例

經過詳細閱讀與仔細記錄這四大段落偶動漫影片後，發現相關數據資料顯出鏡頭運鏡設計與敘事傳達之間的密切相關訊息。在過去關於攝影機鏡頭語言的文獻資料中，大多是僅說明鏡頭的用法與種類，少有針對故事內容以及敘事效果與鏡頭運鏡設計的關聯性探討。尤其是透過鏡頭的景別、角度與時間節奏的設計較少研究或探討。本文以每個段落的影片均以記錄各鏡頭所屬的劇情鏡號、秒數、鏡頭運動、景別、方向以及角度等資料，以折線圖表示各鏡頭時間長度的變化，同時也製作圓餅圖協助研究者可從視覺上理解每個段落內容之各種鏡頭參數的比例。下列表 7、圖 3 及表 8《奇人密碼—古羅布之謎》偶動漫電影之「起」記錄資料範例：

表 8 《奇人密碼—古羅布之謎》之「起」各鏡頭紀錄

鏡號	秒數	鏡頭運動	景別	方向	角度	鏡號	秒數	鏡頭運動	景別	方向	角度
1	3	固定鏡頭	遠景	正面	水平	28	5	固定鏡頭	遠景	正面	水平
2	6	固定鏡頭	遠景	側面	水平	29	4	固定鏡頭	遠景	背面	水平
3	3	固定鏡頭	遠景	側面	水平	30	4	固定鏡頭	遠景	正面	水平
4	4	固定鏡頭	極遠景	側面	水平	31	4	固定鏡頭	遠景	斜側面	水平
5	5	固定鏡頭	遠景	側面	仰視	32	6	固定鏡頭	遠景	正面	水平
6	4	移動鏡頭	遠景	斜側面	仰視	33	5	固定鏡頭	遠景	背面	水平
7	5	固定鏡頭	全景	正面	水平	34	4	移動鏡頭	遠景	背面	水平

鏡號	秒數	鏡頭運動	景別	方向	角度	鏡號	秒數	鏡頭運動	景別	方向	角度
8	5	固定鏡頭	全景	側面	水平	35	2	固定鏡頭	遠景	背面	水平
9	8	固定鏡頭	遠景	斜側面	仰視	36	6	固定鏡頭	特寫	背面	水平
10	3	固定鏡頭	特寫	斜側面	水平	37	6	固定鏡頭	遠景	斜側面	水平
11	9	固定鏡頭	全景	斜側面	俯視	38	4	固定鏡頭	全景	正面	水平
12	6	升降鏡頭	全景	正面	水平	39	5	固定鏡頭	遠景	側面	水平
13	5	固定鏡頭	遠景	背面	仰視	40	4	固定鏡頭	遠景	斜側面	水平
14	7	固定鏡頭	遠景	斜側面	仰視	41	2	固定鏡頭	特寫	正面	水平
15	10	橫搖鏡頭	遠景	正面	水平	42	7	固定鏡頭	遠景	側面	水平
16	5	固定鏡頭	全景	斜側面	水平	43	6	固定鏡頭	全景	斜側面	俯視
17	6	固定鏡頭	全景	正面	水平	44	5	固定鏡頭	遠景	斜側面	水平
18	5	固定鏡頭	特寫	斜側面	水平	45	6	固定鏡頭	遠景	側面	水平
19	4	固定鏡頭	遠景	正面	水平	46	4	移動鏡頭	遠景	斜側面	水平
20	5	固定鏡頭	遠景	側面	水平	47	6	固定鏡頭	極遠景	斜側面	水平
21	3	移動鏡頭	遠景	側面	水平	48	5	固定鏡頭	遠景	斜側面	水平

鏡號	秒數	鏡頭運動	景別	方向	角度	鏡號	秒數	鏡頭運動	景別	方向	角度
22	6	固定鏡頭	遠景	背面	水平	49	3	固定鏡頭	全景	斜側面	水平
23	5	橫搖鏡頭	遠景	背面	水平	50	8	橫搖鏡頭	遠景	側面	水平
24	6	固定鏡頭	遠景	斜側面	水平	51	5	固定鏡頭	遠景	斜側面	水平
25	8	固定鏡頭	遠景	斜側面	俯視	52	7	固定鏡頭	遠景	側面	水平
26	4	固定鏡頭	遠景	側面	水平	53	7	固定鏡頭	遠景	側面	水平
27	7	固定鏡頭	遠景	斜側面	仰視	54	4	固定鏡頭	遠景	側面	水平

表 9 (續) 資料來源：本研究整理

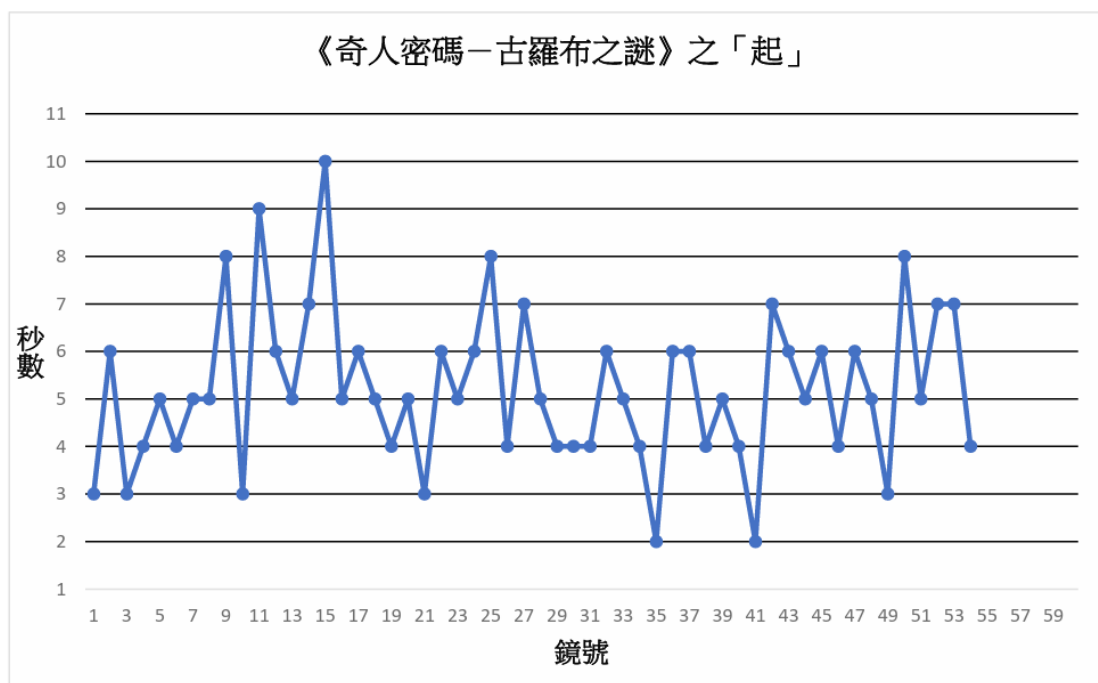
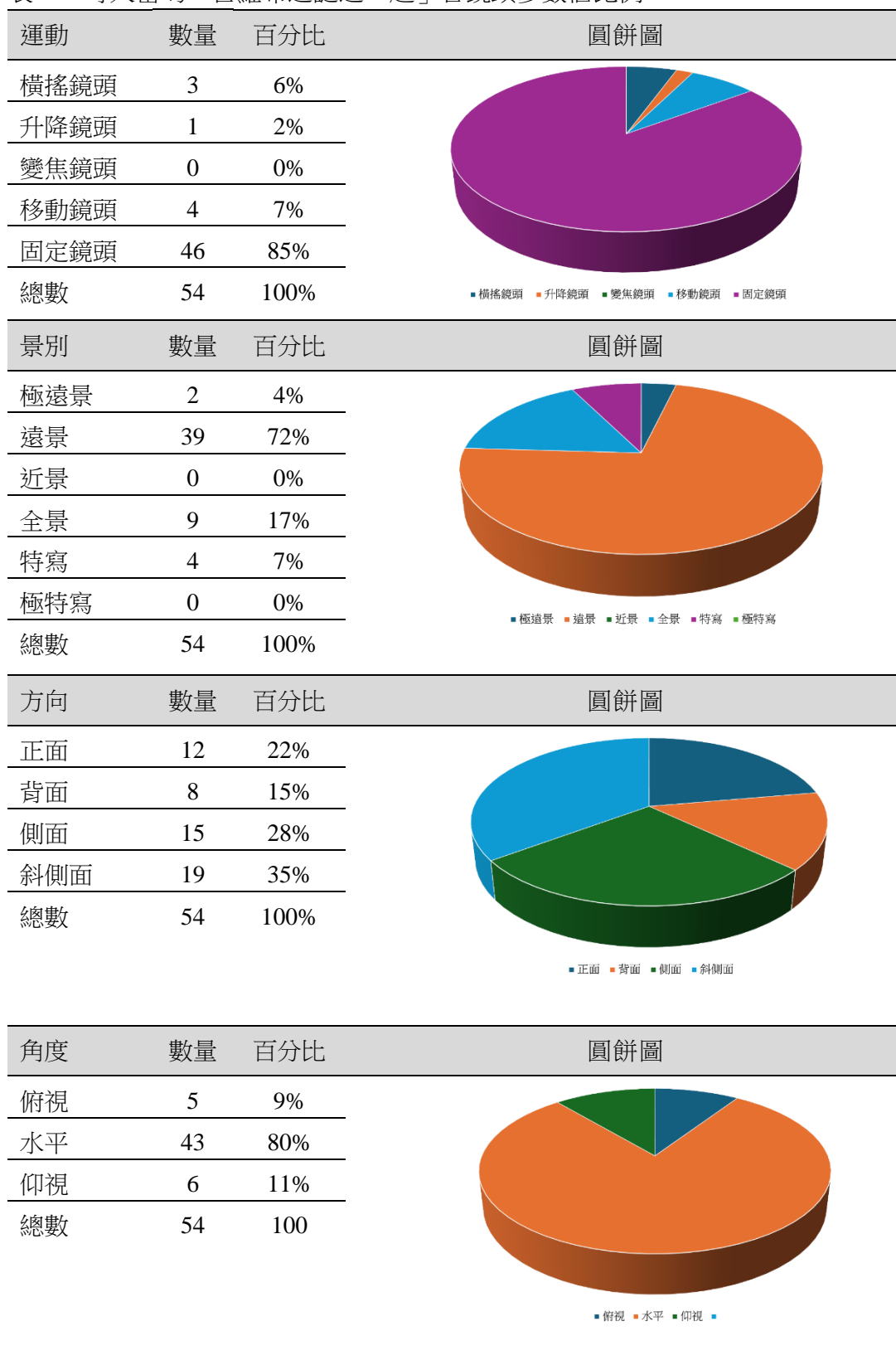


圖 3 奇人密碼—古羅布之謎之「起」各鏡秒數紀錄 (本研究整理)

表 10 奇人密碼－古羅布之謎之「起」各鏡頭參數佔比例



資料來源：本研究整理

經過分析《奇人密碼—古羅布之謎》之「起、承、轉、合」中，該影片的定鏡使用比例為 67%~98%，平均「起、承、轉、合」片段使用定鏡的比例為 87.53%，接近九成的使用比例顯示《奇人密碼—古羅布之謎》之「起、承、轉、合」是以定鏡為主所拍攝的電影偶動畫之影片。

整體而言，可以整理出導演常用的幾個敘事鏡頭設計方式。《奇人密碼—古羅布之謎》之「起、承、轉、合」時常以切換定鏡的手法片段式的帶過角色移動的路徑，縮短了角色移動耗費的時間，即使是用鏡頭跟隨角色移動時，也是以緩慢或者跟隨在角色後方的方式進行，避免橫移造成背景過度晃動造成觀者視覺上的暈眩，以及場景過快消失，使觀眾接受訊息不全而導致對劇情不理解的問題。而以切換的定鏡取代其他鏡頭表現角色的位移演出，一樣可以有生動豐富的視覺效果，重點在於場景的變換和期待劇情變化的觀眾心理。

根據資料分析也顯示，當角色有大範圍的移動動作時，定鏡會切至遠景。由於定鏡是將鏡頭位置固定，並只拍攝某一固定角度，攝影機並不跟隨著角色延續其連續性動作，如果需要表達角色連續動作的話，使用遠景拍攝能同時讓場景和角色入鏡，以此方式可以加強觀眾對於場景的了解，也能清楚表現角色大範圍的運動路徑，而如果角色的連續動作延續到下一個鏡頭，以此方式表現也可將角色動作銜接到下一個場景。

定鏡甚至可以烘托出角色的性格與情緒變化，在《奇人密碼—古羅布之謎》之「轉」當中，這是以定鏡表現角色情緒的方式之一；也可以特寫角色的表情，或用全景、近景、特寫的切換來靈活地表達想陳述的情節。

另外，《奇人密碼—古羅布之謎》之「起、承、轉、合」著重使用定鏡，較少出現其他的運鏡方式，而快速的切換定鏡在敘事上必須注意劇情連貫性，為了避免在切換鏡頭時，使觀眾對同時出現的角色產生混淆，各種角色的特徵辨識度變得極為重要，如《奇人密碼—古羅布之謎》之「承」中間片段以後，共有多個角色相繼出現，此片段就以角色作為明顯的區別。如果是不同的角色，則以形態、顏色以及動作表現做分別。除了切換定鏡之外，在靜止的定鏡片段中，此影片偶爾也會使用視覺變焦方式帶出距離遠近的感覺，再切換定鏡到被聚焦的角色身上，以此方式避免觀眾對角色產生混淆，

雖然切換定鏡的手法可以代替其他運鏡，但有時劇情中的氛圍平和，或角色

個性較為笨拙、溫吞、緩和的運鏡，就比切換定鏡這樣強烈的表現要的適合，如在《奇人密碼—古羅布之謎》之「起」中，在敘述兩兄妹情感中的主角，就以緩慢的移動鏡頭來敘事；為了配合劇情，如果想要更加突顯某一角色的特徵，也可以使用變焦鏡頭，在《奇人密碼—古羅布之謎》之「承」裡便時常用此手法，為了表現配角的渺小，特意從全景 zoom-in 至配角的背影，或者從配角 zoom-out 至全景，以及以環景方式表現配角視角顯示外在環境的龐大。在《奇人密碼—古羅布之謎》之「轉」中，當角色需要做出快速移動的表演時，背景會以模糊、橫向移動的方式表現速度感，有時也會以鏡頭定立在角色前方或後方，而角色呼嘯而過的方式加強角色的速度感。導演採用了變焦鏡頭、橫移鏡頭等的運鏡手法，表現主角在打鬥的英姿。

## 二、定鏡運用比例低於平均值 20%的討論

經過整體的內容記錄得知，《奇人密碼—古羅布之謎》之「起、承、轉、合」影片的定鏡使用比例是 67%~98%，平均一段影片使用定鏡的比例是 87.53%，雖然有近九成的定鏡使用比例，研究亦發現其中《奇人密碼—古羅布之謎》之「合」（67%）的定鏡比例都遠低於平均值 20%或以上。現就《奇人密碼—古羅布之謎》之「合」（67%）的劇情內容做細部討論。

本《奇人密碼—古羅布之謎》之「合」劇情描述兒道茲對於被沙蟲傷到的張彤十分內疚，竟無預警的偷偷將張彤帶至族內給神女-霏兒治療。對於被綁走妹妹的張墨，對羅布族的誤解卻更深，決定搶回妹妹並找到萬鵲讓阿西得以延續，便引著樓蘭軍隊以及一同攻打羅布族。進入古羅布後，發現那彷彿一個人間聖地，而羅布族的居民及沙蟲為了讓萬物生命之源-萬鵲得以永續生存，世世代代的捍衛著這個人間淨土，也因而了解樓蘭王子的野心及真面目。

因篇幅限制就以《奇人密碼—古羅布之謎》之「合」的數據資料與劇情內容做細部討論，總長為 238 秒，分為 42 個鏡頭：如表 11：

表 11 《奇人密碼—古羅布之謎》之「合」的數據資料與劇情內容

鏡頭數	秒數	劇情內容
1~5 鏡	26 秒	兒道茲對於被沙蟲傷到的張彤十分內疚，竟無預警的偷偷將張彤帶至族內給神女-霏兒治療。

6~17 鏡	45 秒	對於被綁走妹妹的張墨，對羅布族的誤解卻更深，決定搶回妹妹並找到萬鵲讓阿西得以延續，便引著樓蘭軍隊以及一同攻打羅布族。
18~26 鏡	54 秒	進入古羅布後，發現那彷彿一個人間聖地。
27~33 鏡	61 秒	進入古羅布後，發現那彷彿一個人間聖地，而羅布族的居民及沙蟲為了讓萬物生命之源。
33~42 鏡	52 秒	萬鵲得以永續生存，世世代代的捍衛著這個人間淨土，也因而了解樓蘭王子的野心及真面目。

資料來源：本研究整理

## 伍、結論與建議

關於動態鏡頭使用的分析可以發現，導演設計鏡頭時是以定鏡為主，但仍會考量敘述表達的必要性來設計運動的鏡頭。因為劇情主題，當固定性鏡頭不足以描述的情節時，或是為了顯示的速度感時才會考慮使用跟隨或移動型態的鏡頭拍攝。而為了加強動態與靜態的對比性，有些畫面仍以固定鏡頭為主，但是會以角色鏡頭的設計來表現對比。

由文獻探討敘事的分析旨在討論故事內容及如何表達的方式。從結構主義敘事學、文學敘事學與電影敘事學，其中的關聯性即是建立於故事結構當中(孫春望、陳建雄與蔣世寶，2009)。在動畫影片中，故事結構可參考依循托多洛夫平衡-再平衡的設計原則，而妥善設計的攝影鏡頭則是表達故事的一個主要技巧。本研究透過詳細記錄並分析《奇人密碼—古羅布之謎》之「起、承、轉、合」影片每一個鏡頭的運鏡方式、景別、鏡頭方向、拍攝角度與秒數等數據資料後整理出部分關於定鏡與其他拍攝手法配合的敘事表達效果與形式，希望本研究的數據和分析結果可以成為關於動畫中定鏡敘事手法研究的參考。以下是《奇人密碼—古羅布之謎》之「起、承、轉、合」影片偶動畫影片研究數據與分析的結論。

### 一、定鏡也能營造高速移動的視覺效果

#### (一) 利用景物的變化

使用定鏡以表現出角色高速移動的感覺，可以藉由角色經過場景時景物的晃



鏡，有助於使觀眾專注在劇情上和角色身上。在《奇人密碼－古羅布之謎》之「起、承、轉、合」中出現的 6 秒以上鏡頭偏多的動畫劇情，通常是角色有長時間快速移動的表演，而且定鏡比例相較於都偏低，移動鏡頭比例偏高，這表示鏡頭與表演設計會依據角色特徵及劇情需求而變。

#### 六、類似的定鏡敘事手法可以重複使用在不同的劇情

本研究在《奇人密碼－古羅布之謎》之「起、承、轉、合」中發現許多重複形式的定鏡敘事手法，此手法的主要模式為：第 1 鏡拍攝角色、第 2 鏡拍攝角色凝望的對象、第 3 鏡再次拍攝角色。這種手法可以重複同樣順序多次，並達到不同的效果。

根據本研究結論之建議，《奇人密碼－古羅布之謎》之「起、承、轉、合」著重在影像的敘事，影片中大量使用定鏡組合的技巧作為表達故事的主要手法。成功的敘事成效不是只有靠鏡頭的設計，這其中包含角色造型、動作表演、情緒表情的掌握、色調氛圍、時間節奏……等都會影響整體影片呈現的效果。雖然 3D 電腦動畫的虛擬攝影機有非常具彈性的運動控制功能，但是本研究的結果提供一個純粹利用鏡頭設計來說故事的思考方向，建議在電腦動畫前製階段的分鏡鏡頭設計時都以固定性鏡頭為主，踏實的以鏡頭所包含的元素（景寬、景深、尺寸、方向、角度、焦距、視角觀點……），同時搭配敘事時間節奏的元素，嘗試單純的以鏡頭與節奏的組合或搭配極少數的運動鏡頭來達到敘事的效果，如此或可避免因重於呈現技巧而忽略敘事本質的結果。

## 參考文獻

- 吳冠英、祝卉(2005)。《藝術院校基礎教材-動畫分鏡頭設計》。北京：清華大學出版社。
- 孫立軍、李卓(2008)。《影視動畫鏡頭畫面設計》。北京：海洋。
- 孫春望、陳建雄、蔣世寶(2009)。情境文字的文圖轉碼過程對景別之圖像解譯差異研究。《設計學報》，14(4)，1-22。
- 張宇晴、鐘世凱(2009)。動畫的鏡位取向及表現性探討－以《玩具總動員》、《史瑞克》《冰原歷險記》三部 3D 動畫電影為例。《國立臺南大學藝術研究學報》，2(2)，1-32。
- 陳曉雲(2010)。動畫電影：敘事與意識形態。《上海大學學報(社會科學版)》，17(5)，54-60。
- Aumont, J., & Marie, M. (1996)。當代電影分析方法論(吳珮慈譯)。台北市：遠流。原著出版於 1988 年。
- Katz, S. D.(2002)。電影分鏡概論：從意念到影像(Film directing shot by shot:Visualizing from concept to screen)(井迎兆譯)。台北市：五南。原著出版於 1991 年。
- Bal, Mieke (1997). *Narratology: Introduction to the Theory of Narrative* [ 1985 ] . Toronto, Ont.: University of Toronto Press.
- Bordwell, David (1985). *Narration in the fiction film*. Madison: University of Wisconsin Press.
- Branigan, Edward R. (1984). *Point of view in the cinema: a theory of narration and subjectivity in classical film*. Berlin; New York: Mouton.
- Bruner, Jerome (1991). The Narrative Construction of Reality, *Critical Inquiry*, Vol. 18, No. 1 (Autumn, 1991), pp. 1-21.
- Casetti, Francesco (1998). *Inside the Gaze: The Fiction film and Its Spectators*. Bloomington and Indianapolis: Indiana University Press.
- Chatman, Seymour (1980). *Story and Discourse: Narrative Structure in Fiction and Film*. Ithaca, NY: Cornell University Press.
- Fludernik, Monika (2009). *An Introduction to Narratology*. London/New York: Routledge.

- Herman, David (2005). Histories of narrative theory (I): A genealogy of early developments. *A companion to narrative theory*, 19-35.
- Kuhn, Markus & Schmidt, Johann N.: Narration in Film (revised version; uploaded 22 April 2014), Paragraph 3. In: Hühn, Peter et al. (eds.): *the living handbook of narratology*. Hamburg: Hamburg University. URL=<http://www.lhn.uni-hamburg.de/article/narration-film-revised-version-uploaded-22-april-2014>[view date : 12 Aug 2016]
- Minuscule TV Series (2016), <http://www.futurikon.com/en/programme/minuscule-l-en/>, retrieved: 2016/9/30.
- Todorov, Tzvetan (1977). *The Poetics of Prose*, translated by Richard Howard. Ithaca NY: Cornell University Press.
- Wilson, George (1976). Film, Perception, and Point of View, *Modern Language Notes (MLN)*, Vol.91, No.5, Centennial Issue: Responsibilities of Critic (Oct., 1976), pp. 1026-1043.

# **A Study on the Narrative Expressiveness of Pili Movie Puppet Animations— Take the Movie " The Arti: the Adventure Begins" as an Example**

**Hsi-Bin Yang<sup>1</sup>**

Abstract

Today, in addition to the fantasy martial arts puppet show series, Pili International also develops science fiction animation. In 2012, Dahua Animation Company was established, integrating dozens of animation special effects teams and resources in Taiwan. In 2014, the second 3D three-dimensional movie " The Arti: the Adventure Begins " was launched. Pili is committed to improving the real-shooting technology and visual aesthetics of puppet animation.

This research will be based on camera lens movement techniques and narrative theory, and will record the plot and scene sequences of the Pili Puppet animation " The Arti: the Adventure Begins ". Make an in-depth analysis of the presentation and expression of its narrative, and study how to use mirror design to achieve high-quality narrative effects in " The Arti: the Adventure Begins ".

**Keywords:** Fixed lens. Anime narrative. Lens length. Movie animation. Lens movement

---

<sup>1</sup> Associate Professor, Department of Advertising, Chinese Culture University.

# 使用 Deepfake 深偽技術呈現虛擬主播系統

吳毓傑<sup>1</sup>、黃亭凱<sup>2</sup>、劉忠陽<sup>3</sup>

## 摘要

深度人工智慧 (Deep AI) 已成為近年來最重要的科技之一。本文探討了深度人工智能技術在傳播教育領域的實作應用。方法上是先以深度仿真技術(Deepfake)生成一個換臉人物，並用這個假人組成虛擬新聞主播。在傳統上，制作傳統的新聞報道需要在一個特定架好光源攝影設備且布置完善的地點、並透過專業攝影機和真實的新聞主播員錄製後，再進行後製編輯完成。所需的新聞主播也必須經由專業口條訓練、讀稿訓練、上線測試等傳統口語訓練方法而成。新聞影片拍攝後製上也非一次到位而成，多次重新拍攝、補鏡皆是時常發生。如果引入 Deepfake 技術後，大部分工作都可以簡化為一個人的工作。新聞主播無需經過嚴格培訓，就能適應不同的情況和報道工作。同時這方式也能在制作其他影片，如電影時，利用 Deepfake 人臉方式取代真人演員。在實務上，善用此法就不需讓真人去適應複雜嚴格的環境外，也能在棚內或後製上簡化許多流程，讓人力資源方面獲得大幅改善。

**關鍵詞：**深偽科技、虛擬主播、新聞視覺設計

---

<sup>1</sup> 銘傳大學新媒體暨傳播管理學系教授。E-mail：wuyc@mail.mcu.edu.tw

<sup>2</sup> 銘傳大學新聞學系助理教授。E-mail：tkhuang@mail.mcu.edu.tw

<sup>3</sup> 銘傳大學新媒體暨傳播管理學系助理教授。E-mail：cyliau@mail.mcu.edu.tw

## INTRODUCTION

With the successful introduction of deep learning technology, AI has become one of the most important techniques in recent years. Examples include, object detection, ChatGPT, facial recognition, medical object recognition, speech recognition, etc. Among them, face swap is the task of switching faces in between two different sources. It can cut out the source face and paste to the target people with almost perfect swap, though the photos or videos were derived from variant backgrounds. For different skin color, high-end face swap technique is able to fine-tune without considering the color-style which is often made manually. In tradition, people usually adopt well-known Adobe Photoshop to accomplish the task of face swap where a set of elaborated image processing techniques, such as color curve, mask, and skin recovery are required. It is not only time-consuming but also requires a lot of prior knowledge about image processing. To handle large-scale data, like videos, it is impractical which takes tens of thousand minutes by human expert. Despite the scaled high human resource efforts, there are still a lot of people with Internet tutorials available online.

By the introduction of deep AI technologies, face swapping can be automatically accomplished. The key step is to detect the human face in the given image precisely. To switch two different faces, face detection is the major and the first step. The goal of face detection is to identify and localize faces in the given image. In the field of face swap, it is required to restrict the search space to be only one face per image. In this step, many prior literatures can be found, for example, MTCNN (Zhang et al., 2016), S3FD(Zhang et al., 2017), Retinaface(Deng et al., 2020), and SCRFD(Guo et al., 2021). Next step is to estimate the head pose which can be automatically obtained by calculating facial landmarks (Abate et al., 2019). To get it more accurate, deep neural network-based approaches (Liu et al., 2021; Fard et al., 2021) were proposed in recent years. To reduce computing power, most face swapping techniques adopted the latter by using PFLD (Guo et al., 2019). Once the face location and landmark positions are ready, then a smoothed face cut mask is applied to replace the two faces. It is unnecessary to cut the whole face, instead, it only takes the eyes, nose, mouth into consideration. Hence, the face mask is only enabled to the central part of the face. In

the final step, the color tune refinement is needed. Color-tune refinement is to replace the color style from the source to target, and vice versa. Finally, it re-smooth the entire face by re-parameterizing the RGB value per pixel. Some black holes or abnormal pixels were also averaged by nearest neighbors. In terms of deepfake, this step is iteratively obtained to find the optimal smoothing functions between the source and target faces.

Deep AI has brought great success in the field of computer vision over the past few years. Meaning of the previously unsolved problems such as face detection, face recognition, object recognition, speech recognition, etc. By the introduction of generative neural networks, it is possible to generate more diverse and different objects. One of the biggest challenges is to automatically generate synthesized human faces. Deepfake, a well-known human face replacement system, was first still interested in the Hader video as an expertly crafted deepfake. It was a technology invented in 2014 by Ian Goodfellow. Generally speaking, deepfake technology is mainly based on generative adversarial networks (GANs) where the learning objective is to generate different output maps randomly or under some predefined conditions. More specifically, those methods are used to model target facial expressions, different headposes, movements of a person. Then, synthesize source facial image sequences of another person [9]. Users can directly control the trained face models performing specified acts by following the controller. However, in this paper, we did not model the synthesized speeches or vocals. Instead, we discuss and analyze the behavior of fake faces. Whether the audience can distinguish the artifactual and fake news. The experiment was mainly conducted by dividing the news into different groups: true news with bonifate reporters, fake news with bonafide reporters, fake news with fake reporters, and true news with fake reporters.

In this paper, we addressed the used of deep fake technology to the field of communication education especially college students. Though the Deep AI techniques and computer systems had been proposed in recent years, it is still difficult for common people without prior knowledge. To this issue, lecture should be simplified and well defined. In addition to computer science students, learners in communication college

have poor knowledge about deep AI technology. The first challenge is familiar with image processing systems and skills about image retouching. To handle this all paper gives comprehensive study of using deepfake for face swapping. In this tutorial, we start from basic steps and easy examples which allows users incrementally learn basic knowledge about deep AI especially deepfake software. the final outcome after this tutorial is to swapping faces between two different sources. In this way, one people can easily control another face by deriving face model from previous training stage. in the communication education, we also give another example how to combine with virtual news anchor which have never seen before. To generalize this idea, students can drive more and more fake face models by using the same techniques. Video authorizing and effect production techniques are the most beginning skills for most Communication students. By the introduction of deepfake techniques, most students could generate their own virtual news.

## **PRIOR ART**

Swapping faces usually involves detecting the face of the given image, then facial landmark detection is applied to find the precise position of important parts of the face, like eyes, nose, mouth, and face shape. Finally, the skin color smoothing method is used. Below, we describe the related works of the above methods.

### **Face Detection**

Before face swapping, the most fundamental step is to identify the face location in the given frame. The technique is called face detection. Over the past decade, by the introduction of deep neural networks, large and frontal faces can be identified mostly by CNN-based methods. Examples include, MTCNN (Zhang et al., 2016), S3FD(Zhang et al., 2017), Retinaface(Deng et al., 2020), and SCRFD(Guo et al., 2021).

- MTCNN showed an excellent performance in face detection tasks and is widely used for many downstream purposes, such as face swapping, facial landmark detection, and face recognition. MTCNN is a multi-task cascaded convolutional net which learns to find face location and landmark points in order under the same framework. There are three subnets under this framework, P-net (proposal network), R-net (refine network), and O-net (output network). In other words, to identify face location, there

are three stages (P-net, R-net, and O-net) that need to be filled. In the first stage, a set of small bounding boxes were identified which filter-out non-face information in this step. The input size of R-net is twice larger than P-net. In this step, face box information is refined to be more close to the real location. At the final stage, R-net takes twice the input size for CNN. However, it will finally filter-out some non-face boxes while finding 5 facial keypoints.

- S3FD was designed based on a purely single deep CNN. One good property of S3FD is the small face detection under a complex background. This framework of S3FD is an improved anchor-based algorithm where a scale-equitable anchor matching strategy is proposed. In comparison to SSD and RPN, this method showed better accuracy in mAP (using WIDER/PASCAL/AFW face dataset)

- Retinaface is a single shot with multilevel methods which identify face location, facial keypoints, and 3D vertex regression under one main CNN framework. To further enhance the effect of facial keypoints, authors also annotate 5 landmarks with 10 scalars in the WIDER FACE dataset. Several simplified versions of Retinaface were proposed in recent years. One remarkable thing is ultra-light face detection. This version is very suitable for mobile-based environments, such as RaspberryPi and Android where the trained model merely takes about 1MB which greatly reduces the computing complexity in edge devices.

- SCRFD is a very high-performance face detection framework which adopted the main features in Retinaface. They pointed out that key to success in face detection are training data sampling and computation distribution. The former (data sampling) is to enhance training data size with variant scale, poses, and angles, while the later (computation redistribution) reallocates the computational efforts between backbone, neck, and the head of the CNN. They showed the improved result (3.86% mAP) and 3X faster speed in WIDER FACE hard dataset.

### **Head pose estimation**

Datasets for facial pose include BIWI [10], AFLW2000[44] and 300W-LP [44]. Both AFLW2000 and 300W-LP use a morphable model fit to faces under large pose variation and report Euler angles. 300W-LP generates additional synthetic views to

enlarge the dataset. More recently, the UMD Faces[1] and Pandora[3] datasets provide a range of data labels, including head pose. A disadvantage for our application is that they do not cover the full range of head poses but only frontal-to-profile views.

A Wide Headpose Estimation Network (WHENet) with multi-loss approaches which was designed to predict Euler angles for head pose estimation. They showed excellent results on wide-range yaw-pitch-roll rotation of existing benchmark datasets such as BIWI and AFLW-2000. The author also provides original source code for further tracking. By manipulating the provided codes, it is not difficult to reproduce the same result on different dataset.

- They claimed that the use of CNN-based or traditional machine learning methods were not only training-based but also not efficient enough for realtime tasks. To meet the requirement, a purely statistical-based approach was proposed to estimate 3D headpose angles. The main soul of this method is to classify headposes by assigning facial keypoints into the quad-tree. For some poses, such as head left-right shake, some certain facial keypoints should occur into some area in the face map instead of centralized.

- The backbone of the ASMNet is purely MobilenetV2 (ref) which performs very efficient facial keypoints detection and pose estimation while keeping acceptable accuracy. By the introduction of the proposed ASMNet, the accuracy was even improved than traditional MobileNetV2. By conducting the experiments on both WFLW and 300W datasets, ASMNet is not only small but also accurate.

By applying 3D–2D matching keypoints, they reconstruct the 3D face map and jointly optimize both keypoint loss and Euclidean distance loss. The proposed method could be applied to depth-based camera images as input.

#### Facial landmark detection

- Khabarлак and Koriashkina (2021) gives a complete survey on the efficient and light-weight facial landmark detection. The problem definition of this task is clear, given a three-dimensional tensor of size  $W \times H \times C$  (width x height x channels) for input image  $I$ , a learned mapping function  $\Phi: I \rightarrow Y$ , where the given image  $I$  learn to find the landmark matrix  $Y$  belongs to  $\mathbb{R}$ . The  $\mathbb{R}$  is a two-dimensional coordinate sequence  $(x,$

y). However, this paper does not mention the 106 facial points which were widely evaluated in recent competitions.

- Both MobileNetV2 and MobileNetV3 are very efficient deep neural network backbones. The proposed PFLD is just a simplified version of the backbone. The name mobilenet is for mobile devices, such as cell phones. PFLD is very suitable and widely tested in mobile devices. To find the facial keypoints, this method can handle the work with a few milli-seconds. It is worth to note that there is a set of PFLD-based facial landmark detection methods trained with 106 facial keypoint dataset and received great attention.

- Other than heatmap-based and anchor-based methods, the designed Pipnet integrates both features. Pipnet computes heatmap regression scores on low-resolution feature maps. In this way, there is no need for an upsampling layer while reducing substantial computing efforts during inference time. To further enhance the result, the authors also present a self-training strategy benefiting from large unlabeled data. The result shows that the Pipnet is not only fast ( $>30$  fps) but also achieves state-of-the-art accuracy for most 68-points facial landmark detection benchmarks (300W, COFW-68, and WFLW-68). Pipnet is also ported to train with 106 facial landmark tasks (JADA/Lapa) and also showed excellent performance.

- This paper adopted well-known heatmap regression framework which is robust to different poses and its variations, illuminance, and occlusion under difficult environments. They propose a quantization error function for unbiased regression. In the experiments, they conduct their method with several well-known benchmarks, such as WFLW, AFLW, 300W, etc. The result is very convincing and also widely used in most face swapping tasks.

## **Face Swap**

Face swapping is a technique in computer vision that involves replacing the face of one person in a photograph or video with the face of another person. This can be done using a variety of techniques, including image processing algorithms and machine learning models. Originally, the face swap technique was mainly used to swap faces between two images. With the rapid development of the computing units, the run time

of swapping faces is ~30fps which can be applied to real time purposes. One common approach to face swapping is to use a machine learning model that has been trained to recognize and detect faces in images. The model can then be used to locate face positions in both source and target images, and the faces can be swapped by replacing the pixels in the source image with the pixels from the target image. Another approach is to use an image processing algorithm to analyze the features of the faces in the source and target images, such as the shapes of the eyes, nose, and mouth, and use this information to synthesize a new image that combines the features of both faces. There are also software tools and apps available that can be used to perform face swapping, which may use a combination of these techniques. However, it is important to note that face swapping can have ethical and legal implications, as it can be used to create fake or misleading images or videos.

The most remarkable integratable face swap tool so far. It contains most recent face processing techniques, such as S3FD/MTCNN/SSD for face detection and 2DFAN/PRNet for facial landmark detection. In the face segmentation stage, the TerausNet was adopted which allows cutting face features except for hair, fingers, and glasses exactly.

GAN is a two-step framework: G is to gen a latent code for sample generation whereas the D is to verify the generated sample is fake or bonafide. This paper showed earlier work on the GAN network. The generated non-real person face can be very high-resolution in comparison to previous literatures.

### **TASK DEFINITION**

In this paper, the final target is to Present virtual news report which combines face swap technology. In comparison to previous literature, this paper presents a pilot start and is well suitable for the beginners that have never prior experience in deep AI technique before. We conduct our experiment by integrating this technique to the communication college students. Obviously, communication students are totally different from engineering students who are very familiar with computer operating techniques. The detailed goals are listed below.

1. swapping two faces between two different people

2.refine swapped faces

3.representing the result to Virtual news report

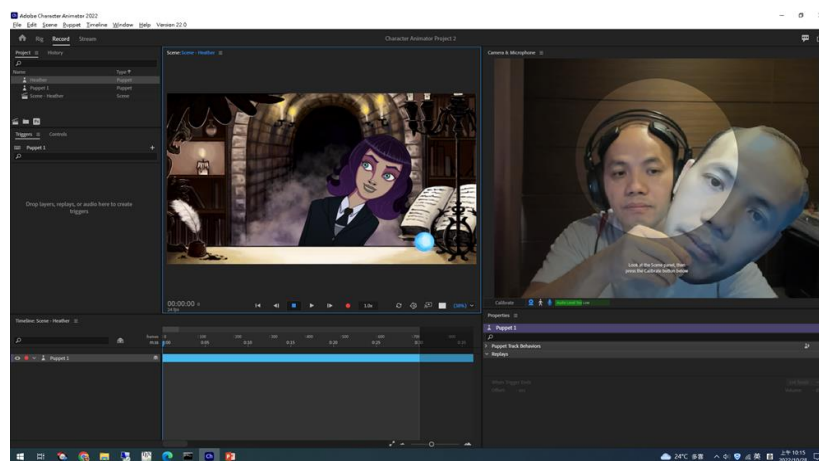
The first target is to utilize Deepfake deepfacelive software for handling face swapping. We restrict the users to record their own videos and then try to find another well-known people to be replaced. In this way, user can easily control the target and can be used to report news without considering the real people and real scene.

The second target is to refine the swapped faces and represent them in the video. Next is to use a video authoring tool, such as Adobe Premiere Pro to refine the overall color tune to match with the background. For the final target, users could directly make or create fake news by including the swap face. The generated fake news anchor is mainly derived from the deepfake. In this stage we also borrow the idea from the journeys so that audiences could not easily find whether the news is real or fake.

## DEEPPFAKE NEWS PRESENTATION

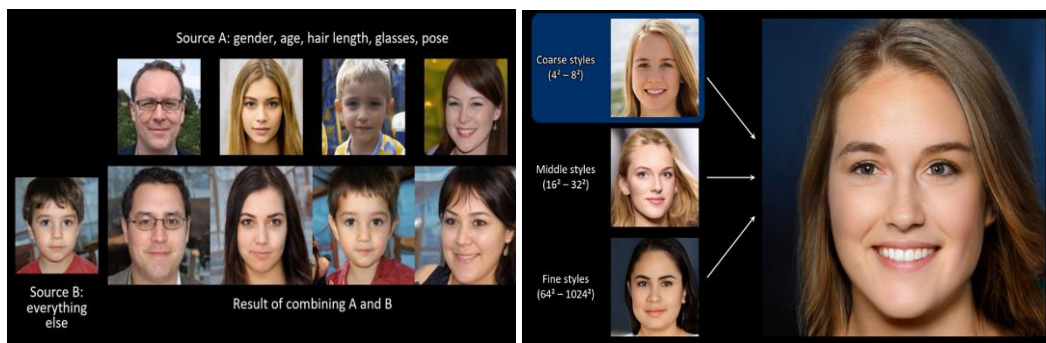
In recent years with a rapid development of motion capture, there is a set of software which supports interacting by webcam. For example, Adobe Character Animator 2022 is one of the software which can capture a human face and react by an animator. It can detect the user's head with a frontal face automatically. Users could move their head or make different head poses so that the animator can do the same action. Figure 1 illustrates an example of the adobe character animator.

However, this software is obviously restricted to the output animators. There are only a small number of actors who are not popular with common people. Another drawback is that the Avatar is far away from reality; it is still not bonafide.



**Figure 1: The motion capture software (Adobe Animator 2022)**

The second software we refer to is that GAN face generator. This online tool is free and copyright free. There's no need to care about the IP and copyright information because GAN can produce synthetic faces that are not real. However, this result is almost perfect. Directly adopting these faces to virtual news is not a good idea, the auto-generated faces were unseen for most audiences. For most people, a popular and common-seen face is very important. To make the generated face to be a real star, it is still a long way to go. Figure 2 and Figure 3 show the sample of GAN auto-generated human faces.



**Figure 2: A display of GAN face generator**

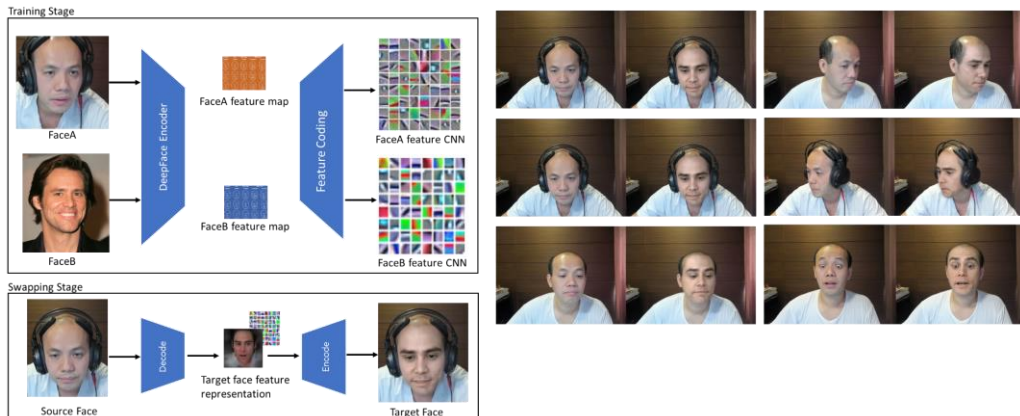


**Figure 3: Detail settings of the generated face**

**Brief review of DeepfaceLive**

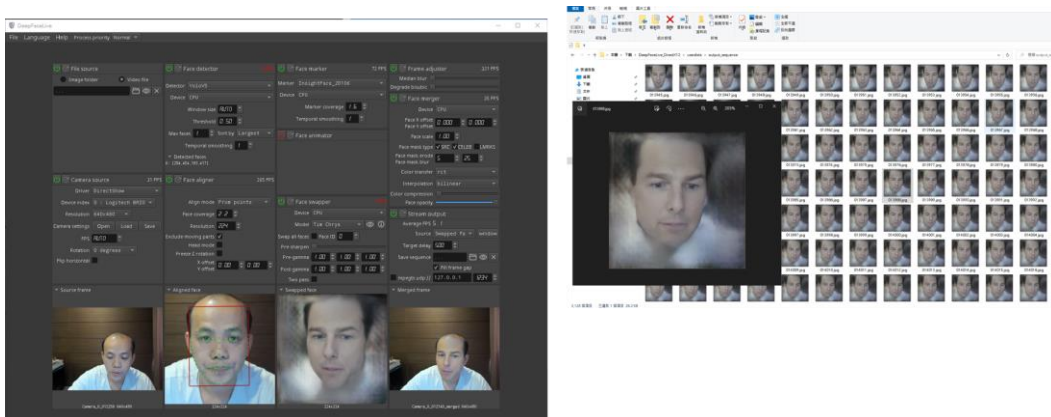
In this section we briefly review the Deepfake. The following figure illustrates the overall usage diagram of deepfake. The overall usage can be viewed as two different stages (see Figure 4). The first stage is to train to find the homogeneity between two different sources. It will interactively find the optimal solution to find the best match face position to be swapped. At the same time, it tries to find the smooth ages between

the swapped face and the target. In the training stage deepfake uses the traditional deep learning convolutional neural network solution. The feature map is generated to find common attributes or a side in different ways to different parts of the given faces. After training, the learn feature map can be generalized to different head poses, emotions, situations.



**Figure 4: Overall usage diagram of deepfake**

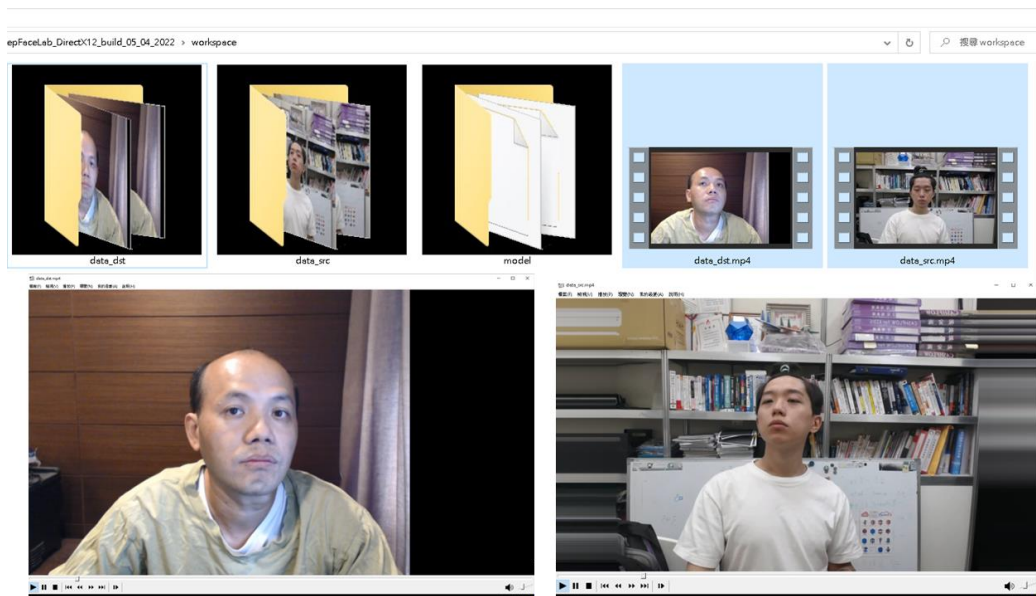
Deepfake also provides another software namely deepfacelive which supports live streaming usage. Figure 5 gives an illustration of the deepfake live streaming UI. This tool can interact with arbitrary input situations even though the input face is never seen before. the user can interact and try to control the swap of the face at this moment. the following figure shows an example of deepfacelive.



**Figure 5: Deepfake live streaming**

### Step0: Prepare source and target videos

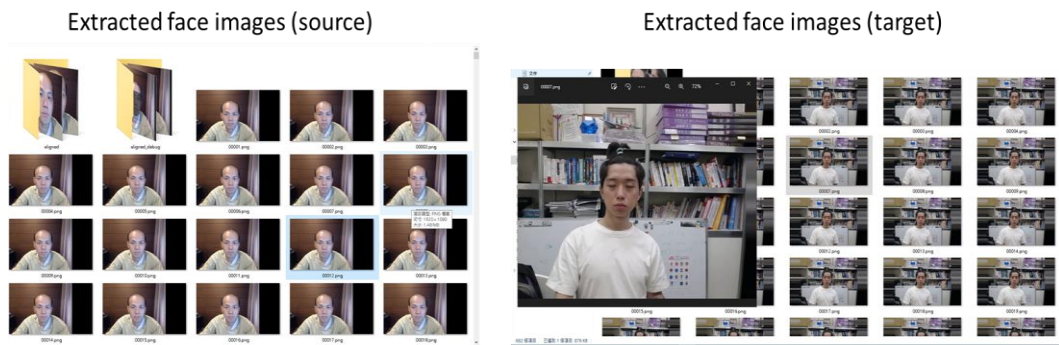
The first step of face swapping is to prepare two videos, source and target which contain two different faces. Figure shows the preparation of the source and target videos. By following the scheme of deepfake, we put the two videos into data\_src and data\_dst folders, respectively. Figure 6 shows the preprocessing allocation of the source and target folders.



**Figure 6: Preprocessing steps of the deepfake (left-hand-side, source video; right-hand-side, target video)**

### DeepFake-step1 Face detection

Next step is to identify face position and time. It is required to analyze the whole video by sequence of frames. In other words, the video is deformed by a sequence of image frames. Deepfake starts to scan each frame and preserve the frame which contains human faces. The following Figure shows the extracted face images of source video and target video. Figure 7 shows the extracted face images of the source and target folders, respectively.



**Figure 4: Illustration of the video-to-frame sequence**

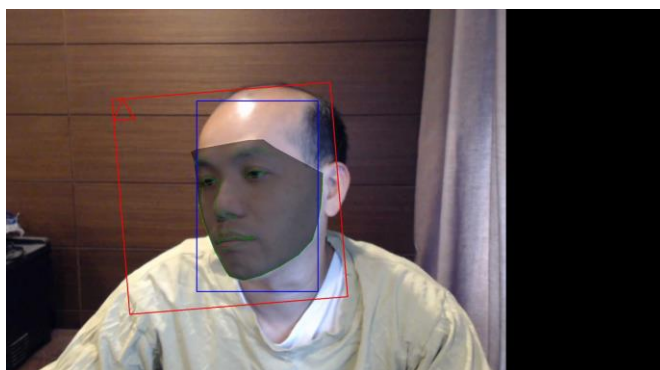
There are several face detection algorithms provided in deepfake. As mentioned in prior literature, we choose S3FD which shows excellent performance in face detection tasks to find faces in the frame image.

#### Step2: Face Alignment (512x512)

Usually, a human face could be presented in different pose angles by raw-chin-yaw axes. To classify and induct those poses, it is necessary to calculate the human head pose and facial landmarks. Once the classification of the pose is found, it can try to find similar head poses to swap instead of just using the frontal or one image. This will greatly improve the effectiveness of face swap. In tradition, face swapping is easily failed by using different poses. By the introduction of head poses and face alignment, those faults could be recovered by similar head classification. Figure 5 shows different head poses of one people. Figure 6 draws the 3D illustration of the facial alignment.



**Figure 5: Different head poses for facial alignment**



**Figure 6: 3D illustration of the facial alignment (green part is the facial landmarks of the face)**

### **Step3 Options**

In this step deepfake provides basic denoise functions that smooth the overall image by averaging similar and neighbor pixels. This step is optional. For some situations, the input images could be blurred, under/over-expose, low contrast, noise. To overcome this, Deepfake also provides basic image processing tools, like contrast tweaking, smoothing, exposure settings, color correction, etc. In this paper we did not run the experiment with this function. We have recorded our videos under pure background.

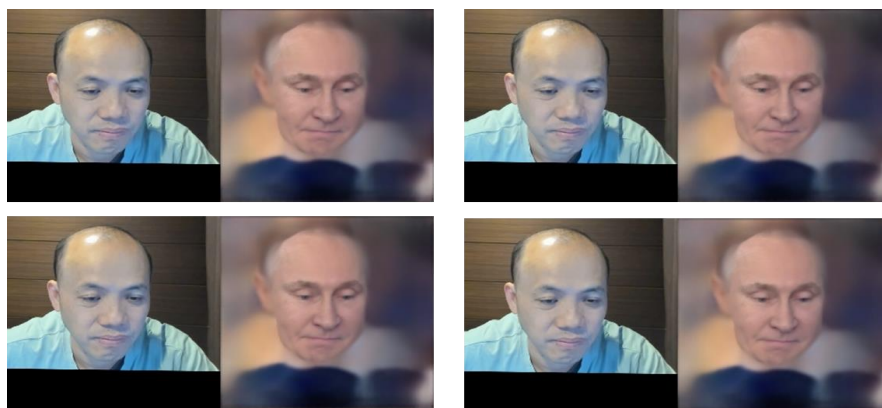
### **Step4 Training**

To swap faces between two different sources it is required to extract facial pictures and try to induct some rules by the input sequence. Those features could be generalized to different testing environments after training. This state can be viewed as the encode/decode stage which adopted the encoder and decoder for learning a compact representation of a face. In other words, this stage is used to learn the embedding of input faces. Hence in the testing stage, it can resemble the inputs. Once the feature embedding could be obtained, the choice of loss function is important. This function could guide the entire training process to be converged in a short time. Common loss functions include: L1/L2 Loss, perceptual loss, and adversarial loss. Adversarial training uses a discriminator network to differentiate between real and generated faces.

The generator aims to fool the discriminator, while the discriminator attempts to correctly distinguish between real and fake images. This process helps in producing more realistic results.

By setting up training instances, loss function, and training parameters, Deepfake training can be triggered iteratively until it meets the requirements of the agenda. The model is trained iteratively using the collected dataset. Initially, the generated faces may look unrealistic, but as training progresses and the loss functions are optimized, the quality of the generated faces improves. To prevent overfitting and improving the generalization power, the regularization techniques are used, such as dropout, batch normalization, and weight decay.

It's essential to highlight that while deepfake technology has various applications, including entertainment and visual effects, it has also raised concerns regarding misinformation and its potential misuse for malicious purposes. Ethical considerations, transparency, and responsible use are critical when developing and deploying such technology. Additionally, the use of deepfake technology to create non-consensual explicit content is illegal and unethical. Figure 10, Figure 11, Figure 12 show the sample images of the Deepfake during face swapping.



**Figure 7: Sampling image during deepfake training**



**Figure 8: Sample frame sequence of the unsmoothed swapped faces**



**Figure 9: Frame sequence of the smoothed swapped faces**

## RESULT

We merge our face to different scenarios for positive communication expression, for example, virtual anchors. Below shows the result after deepfake transform. The merged fake news is shown in this figure. The left-hand-side of the figure is the original source face while the right-hand-side people is the merged target. It is clear that the imposter is very close to real. It is quite difficult to verify whether it's a spoof face or bonafide. Figure 13 shows the final presentation of the deepfake news.



**Figure 10: Final layout of the virtual news (faces were swapped)**

In this image, three men are seen wearing headphones and clothing. The man in the center is the target face with deepfake person with wearing a white t-shirt with black stripes, while the left person is the source face. For the right man is another deepfake generated face by simply replacing the source and target faces.

The background of this image is a light gray color with standard virtual news illustrations in white. We directly adopted the available world map and made several revisions of the text to express the fake news. The three news anchors were fake and the news is also not real. In the bottom part of the Figure is merely standard news headline layout. The left-hand-side of the bottom banner is the motion visual aids which often display the news tv text; the right hand side is formed by large white text (This is totally fake news! Smile with me!) and the logo of the breaking news.

Overall, this image shows the final outcome of the face swap and merged to the virtual news layout. The result had been viewed by over 20 college students. None of them believe the news is true. All 20 testing views agree that the merged news is very close to real. Without headline text, it is quite difficult to distinguish whether the news is true or not. One student said that the news is amazing to the audience since the generated fake news is close to being real. If the face was not common, it could be heard as true news.

## CONCLUSION

Deepfake techniques had been investigated over the years. It is still a question how to utilize this tool for communication education and applications is not commonly found. This paper presents a pilot study of using deepfake generated faces to virtual news reports. We gave a complete example of deepfake usage to this field. It firstly

trains a source/target face model and generalizes the face model according to user requirements. By merging to virtual news layout, the fake people could be used to play the role of news anchor. There are still questions not solved by this paper. For example, we plan to adopt a speech analyzer to clone the target's voice. By integrating with the deepfake news, it provides a total solution of AI fake news.

## REFERENCES

- Abate, A. F., Barra, P., Bisogni, C., Nappi, M., & Ricciardi, S. (2019). Near real-time three axis head pose estimation without training. *IEEE Access*, 7, 64256-64265.
- Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018). Mesonet: A compact facial video forgery detection network. In *2018 IEEE International Workshop on Information Forensics and Security (WIFS)* (pp. 1-7). IEEE.
- Choi, Y., Choi, M., Kim, M., Ha, J. W., Kim, S., & Choo, J. (2018). Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 8789-8797).
- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1251-1258).
- Deng, J., Guo, J., Zhou, Y., Yu, J., Kotsia, I., & Zafeiriou, S. (2020). RetinaFace: Single-shot multi-level face localisation in the wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5203-5212).
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* (pp. 2672-2680).
- Guo, J., Deng, J., Lattas, A., & Zafeiriou, S. (2021). Sample and computation redistribution for efficient face detection. *arXiv preprint arXiv:2105.04714*.
- Guo, X., Li, S., Yu, J., Zhang, J., Ma, J., Ma, L., Liu, W., & Ling, H. (2019). PFLD: A practical facial landmark detector. *arXiv preprint arXiv:1902.10859*.
- Jin, H., Liao, S., & Shao, L. (2021). Pixel-in-pixel net: Towards efficient facial landmark detection in the wild. *International Journal of Computer Vision*, 129, 3174-3194.
- Khabarlak, K., & Koriashkina, L. (2021). Fast facial landmark detection and applications: A survey. *arXiv preprint arXiv:2101.10808*.

- Liu, L., Ke, Z., Huo, J., & Chen, J. (2021). Head pose estimation through keypoints matching between reconstructed 3D face model and 2D image. *Sensors*, 21(5), 1841.
- Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2017). Progressive growing of GANs for improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196*.
- Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4401-4410).
- Perov, I., Gao, D., Chervoniy, N., Liu, K., Marangonda, S., Umé, C., Facenheim, C. S., RP, L., Jiang, J., Zhang, S., Wu, P., Zhou, B., & Zhang, W. (2020). DeepFaceLab: Integrated, flexible and extensible face-swapping framework. *arXiv preprint arXiv:2005.05535*.
- Pourramezan Fard, A., Abdollahi, H., & Mahoor, M. (2021). ASMNet: A lightweight deep neural network for face alignment and pose estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1521-1530).
- Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). Faceforensics++: Learning to detect manipulated facial images. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1-11).
- Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). Deepfakes and beyond: A survey of face manipulation and fake detection. *arXiv preprint arXiv:2001.00179*.
- Yu, B., & Tao, D. (2021). Heatmap regression via randomized rounding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multi-task cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10), 1499-1503.
- Zhang, S., Zhu, X., Lei, Z., Shi, H., Wang, X., & Li, S. Z. (2017). S3FD: Single shot scale-invariant face detector. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 192-201).

Zhou, Y., & Gregson, J. (2020). WHENet: Real-time fine-grained estimation for wide range head pose. In *BMVC2020*.

# Video Virtual News Presentation by Deepfake

Yu-Chieh Wu<sup>1</sup> Ting-Kai Huang<sup>2</sup> Chung-Yang Liu<sup>3</sup>

## Abstract

Deep artificial intelligence (AI) has become one of the most important techniques in recent years. In this paper, we address the application of deep AI technology in the field of communication education. The deepfake technique is applied to generate a swapped face person and the fake person is used to form the virtual news anchor. In tradition, to produce the traditional news report requires a real place, professional camera, and real news anchor. Then, the news anchor should be trained well or refine the videos by re-camera several times. By the introduction of Deepfake, most efforts could be reduced to one-man's work. The news anchor does not need to be trained well and can adapt to different situations and reporting works. To make movie production, the real actors or actresses could be easily replaced by the deepfake faces. It is very attractive since the human resource efforts are substantially improved.

**Keywords:** deepfake technology, virtual news, news visual design

---

<sup>1</sup> Department of New Media and Communication and Management, Ming-Chuan University

<sup>2</sup> Department of Journalism, Ming-Chuan University

<sup>3</sup> Department of New Media and Communication and Management, Ming-Chuan University

# 銘傳大學 傳播學院

## 《傳播管理學刊》 徵稿啟事

《傳播管理學刊》為傳播學院之學術刊物。本刊鼓勵以嚴謹的學術研究精神，探討傳播管理相關議題，本期刊為一年二期刊，每年 6 月及 12 月出版。

### 壹、本刊內容

一、徵稿範圍：傳播管理領域之研究、理論、方法、實務分析、評論等各類學術論著，凡未發表或出版於國內外期刊，並符合本刊主題之研究論文均歡迎投稿。

二、本刊主題：

- (一)媒體產業相關議題
- (二)傳播生態相關議題
- (三)數位產業相關議題
- (四)國際傳播相關議題
- (五)整合行銷傳播相關議題
- (六)其他傳播管理相關議題

### 貳、審查制度

一、審稿及刊登：本刊設有審查制度，所有稿件均經編輯委員會送請相關領域專家學者審查通過後刊印，經錄用之稿件，不發放稿費，本學刊登載之論文，版權歸本院所有，唯文責由作者自負。

二、來稿之評審由編輯委員及傳播管理相關研究領域學者擔任。

三、送審稿件均需匿名，由兩位專家學者進行匿名審查；每位評審於評審意見上陳述意見，並於下述四項勾選其中一項：

- (一)不必修改，直接刊登。
- (二)略加修改(敬請說明建議)。
- (三)大幅修改(敬請說明建議)。
- (四)退稿(敬請說明建議)。

四、編輯委員會得就第三位評審之意見，決定是否刊登或退稿。

五、是否刊登文件，事關投稿人權益，應將評審意見等函送投稿人，說明處理方式。

## 六、處理方式

處理方式		第二位評審意見			
		刊登	略加修改	大幅修改	退稿
評審意見	刊登	刊登	寄回修改	寄回修改	第三位評審
	略加修改	寄回修改	寄回修改	寄回修改	第三位評審
	大幅修改	寄回修改	寄回修改	退稿	退稿
	退稿	第三位評審	第三位評審	退稿	退稿

## 參、投稿須知

### 一、來稿應注意事項：

(一)來稿內文可用中、英文寫作，皆須附中、英文摘要。

(二)來稿請備紙本一式三份(包括封面頁、標題、摘要、關鍵字、正文、圖表及參考文獻)，以及 Word 電子檔一份。寄至台北市士林區中山北路五段 250 號 銘傳大學傳播學院《傳播管理學刊》編輯委員會收。

二、每篇論文長度約一萬字以內。

三、語文：中、英文均可，唯英文稿須附中文題目及摘要。

四、本刊對撰寫格式擁有修改之權利，以保持編輯水準。

五、截稿日期：文稿審查(包括初審、作者修改、複審)約需三至四個月，文稿之截稿日期為該期出刊前四個月。例如：六月出刊之文稿，其截稿日期為二月底。

### 肆、稿件格式

本刊為統一文稿 APA 格式，編訂撰稿體例，來稿須合本刊徵稿格式，以便本刊作業。稿件格式請參照銘傳大學傳播學院《傳播管理學刊》網頁 <https://reurl.cc/Gj1WGy>。

# **Journal of Communications Management**

**Vol.24 No.2 December 2023**

---

- Study on the attitude change and the impact on the media literacy towards Taiwan of the exchanging journalism school students from mainland China
- A Study on the Narrative Expressiveness of Pili Movie Puppet Animations — Take the Movie " The Arti: the Adventure Begins" as an Example
- Video Virtual News Presentation by Deepfake

Published by School of Communication  
Ming Chuan University

ISSN 1993-6516