

以手勢為互動基礎的介面設計學術研究發展— 文獻回顧與綜論

陳建雄* 王建立**

* 台灣科技大學設計系研究所

cchen@mail.ntust.edu.tw

** 台灣科技大學設計系研究所

通訊作者：15067431606@163.com

摘要

目前我們正處於傳統的 WIMP 介面轉向自然互動介面的階段。手勢作為自然互動介面中最重要的輸入模式之一已經引起了各界的關注。然而在設計研究領域，以手勢為互動基礎的介面研究仍處於起步階段。而人機互動介面易用性的提升則需要對新的輸入模式進行廣泛的研究和應用。因此，本研究使用 Web of Science 和臺灣期刊論文索引系統資料庫，藉由後設研究法收集並分析先前國際與臺灣手勢互動相關之學術研究論文，藉此了解手勢互動學術研究之現況、專業學術期刊及重要主題等。最後，本研究提出 5 個從設計研究角度切入手勢互動學術研究的重點：1.手勢如何代替滑鼠進行目標指向和選擇；2.如何選擇合適的手勢來產生指令；3.手勢輸入如何應用於多重感官互動之中；4.以手勢為互動基礎的介面和傳統 UI 介面的區別；5.以手勢為互動基礎的設計如何驗證其效用、效率和滿意度。期許，本研究的成果可以鼓勵更多的設計研究人員加入手勢互動設計研究領域之中，並為實務設計及學術研究提供重要參考。

關鍵詞：手勢互動、介面設計、人機互動、易用性測試、多重感官互動

論文引用：陳建雄、王建立（2021）。以手勢為互動基礎的介面設計學術研究發展—文獻回顧與綜論。《設計學報》，26（1），59-82。

一、前言

近年來隨著科技的發展，各種偵測體感的設備包括深度攝影機（例如 Kinect、Intel Perceptual Camera、Leap Motion 等）和感測器（例如慣性測量單元 IMU）開始商用普及，使得我們能夠透過手勢與電腦進行互動。早在 1979 年，Bolt 在 MIT 設計了一個 media room 透過手勢和語音與電腦進行互動，開啟了新一代的人機互動介面研究（Bolt, 1980）。在 20 世紀 90 年代早期，這種以手勢為互動基礎的介面開始出現在實驗室和學術研究中。Dam 於 1997 年對 GUI 的發展歷史進行了回顧，提出第三代以傳統設備（鍵盤、滑鼠）為互動基礎的介面 WIMP GUIs（windows, icons, menus, and a pointing device）和第四代面向未來

以手勢和語音為互動基礎的介面 post-WIMP (Dam, 1997)。Turk 亦針對電腦越來越普及和無處不在 (ubiquitous computer) 提出了類似的概念稱之為可知覺的使用者介面 (perceptual user interfaces) (Turk, 2001)。隨著 2010 年微軟 Kinect 問世，以機器視覺為基礎的手勢互動和語音互動開始被 HCI 領域所熟知且定義為自然互動介面 (natural user interface) (Goth, 2011)。我們正處於傳統的 WIMP 和自然互動介面混合的階段，且自然互動已經成為了人機互動研究的一個重要領域 (Turk, 2014)。而在自然互動介面中，手勢互動是最主要的輸入方式之一。

人機互動介面的易用性提升與適用於摩爾定律的電腦性能變化不同，往往在經歷一個穩定的階段後突然發生突變 (Dam, 1997)。這種突變往往來自於新的輸入模式 (input modality) 得到廣泛的研究和應用。手勢互動作為一種新的輸入模式在近幾年裡得到了多個領域 (資訊科技、心理學與人因工程等) 的關注。然而在設計研究領域，手勢互動介面的研究還處在一個起步階段，亦沒有系統性的文獻回顧。本研究希望透過對手勢互動研究的相關文獻進行回顧，來探討不同領域在手勢互動研究的現狀及其對手勢互動設計的借鑑，並提出未來手勢互動設計研究的發展重點。本研究的主要目的有三：1. 綜觀國際與臺灣手勢互動學術研究發展；2. 探討已經存在且值得借鑑到手勢互動設計研究的原則；3. 提出未來手勢互動設計研究發展重點方向。

二、手勢互動研究發展脈絡

在日常生活中，我們透過手來操作物品，也透過手勢和其他人進行交流。手是天然的互動工具，也與生俱來存在兩種不同的屬性：操作手勢和語意手勢。Quek 等人 (2002) 將伴隨著語言的手勢和具有符號語意的手勢定義為語意手勢。而那些“其預期目的是通過手的實際運動與被操作實體之間建立一個緊密關係來控制某個實體”的手勢則被 Quek 等人定義為操作手勢。操作手勢建立在人與自然的互動經驗之上，將其應用在人機互動中透過手直接控制三維模型或和虛擬世界進行互動。Vogel 等人透過手勢操控游標來和計算機進行互動 (Vogel & Balakrishnan, 2005)，他們認為透過直接指向和點擊來和傳統的使用者介面進行互動仍然是現今主要的互動模式之一。儘管近幾年出現了很多以手勢操作為基礎的新型互動模式 (例如：multimodal interface、3D、VR 等)，但是以指向和點擊為基礎的互動方式因為其簡單、自然、靈活的操作而具有獨特的優勢。Lou 等認為這種手勢互動模式不僅符合人們對傳統互動模式的認知，而且互動方式變得更為自然且脫離距離的限制 (Lou, Peng, Hansen, & Li, 2017)。除此之外，語意手勢因個人和文化差異而具有不同的語意 (Norman & Nielsen, 2010)，即使同一個人也會因為情緒、疲勞等因素而有所改變 (Morrel-Samuels, 1990)。不同於部分學者對操作手勢的推崇，Quek 等認為自然世界中的操作手勢伴隨著視覺、觸覺和動量的反饋，而虛擬世界中的手勢操作互動則缺乏這些反饋 (Quek et al., 2002; Quek, 1996)。Wexelblat 認為一對一的指令模式只會增加使用者對手勢記憶的認知負擔，而且幾乎沒有增加什麼功能。未來的人機互動應該透過人與人交流時的手勢應用來和計算機進行互動 (Wexelblat, 1995, 1998)。隨著機器學習、深度學習等演算法的成熟和電腦運算能力的提升，語意手勢將發揮越來越大的潛力。

操作手勢和語意手勢各有優缺點，兩者的結合有時候可以提供更豐富的互動形式。Rubine 設計了一個 two-phase 技術，以語意手勢來執行指令，然後以操作手勢和電腦直接互動 (Rubine, 1992)。在兩種手勢轉換中，他透過 Eager recognition 技術來平滑轉換過程，例如透過“L”手勢來繪製矩形，在該手勢的結尾點開始直接變換矩形的尺寸或旋轉。Chen 等嘗試透過手勢和觸摸來與手機進行互動。手勢用來增加互動的豐富性，觸摸則用來執行指令，例如畫一個圓 (語意手勢) 然後觸摸手機螢幕來執行選單指令

(Chen, Schwarz, Harrison, Mankoff, & Hudson, 2014)。Alkemade、Verbeek 與 Lukosch 透過語意手勢來執行抽象的指令，然後透過操作手勢對三維模型進行操作和屬性控制，進而提供一個透過手勢進行建模的系統 (Alkemade, Verbeek, & Lukosch, 2017)。Matulic 與 Vogel 透過手指投射到螢幕而組成的不同形狀來執行不同的指令，透過手的移動對目標進行操作。他們將這種操作科技稱為“多點投射”空間 (Matulic & Vogel, 2018)。在這一類的手勢互動設計中，認知和運動控制是互動成功的關鍵因素。Rempel 等人認為在一個以手勢互動為基礎的指令中，只有認知、運動控制和計算機識別三個因素緊密結合時才能提供易用的互動介面 (Rempel, Camilleri, & Lee, 2015)。

三、研究方法

本研究透過檢索國內外知名學術資料庫，收集手勢互動相關學術資料並進行嚴謹篩選後，分析國際與臺灣手勢互動研究發展現況，再進一步提出各個領域可以借鑑到手勢互動設計的原則和未來手勢互動設計研究發展重點方向。本研究選擇資料庫為 Web of Science (webofknowledge.com) 與臺灣期刊論文索引系統 (<http://readopac.ncl.edu.tw/nclJournal/>)。Web of Science 完整收錄有科學技術類期刊 (Science Citation Index Expanded, SCI-Expanded)、社會科學類期刊 (Social Sciences Citation Index, SSCI) 及人文藝術類期刊 (Arts & Humanities Citation Index, A&HCI) 等三大具影響力之資料庫，能夠較為完整的呈現相關學術研究的全貌。臺灣期刊論文索引系統建置資料已逾 40 年，收錄期刊 4800 餘種，較能完整的呈現目前相關學術研究的全貌。

針對 Web of Science 資料庫，搜索條件設定時間從 1990 年至 2018 年 (資料庫能夠提供的最大年限範圍)，針對所有字段搜索關鍵字“hand gesture interaction”。針對臺灣期刊論文索引系統，檢索條件不設定起迄時間，搜索涵蓋篇名、關鍵字及摘要有“手勢”者。搜索關鍵字之所以不採用“手勢互動”、“手勢互動”或者“hand + gesture + interaction”，是為了在手勢議題的研究中盡可能涵蓋較多的領域，並為第二階段的文獻篩選和內容分析提供足夠的基數以利一窺手勢互動研究之全貌。

文獻的搜索篩選程序如圖 1 所示，首先審閱標題，排除未符合手勢互動議題的文獻。其次審閱摘要和關鍵字，進行進一步篩選。篩選的標準是排除神經科學、嬰兒語言發展與手勢的關係、以眼睛和頭部等與電腦進行互動、以及以觸摸為基礎的手勢互動等議題。爾後，審閱全文後對文獻進行分類。分類標準為手勢識別 (包含手勢識別算法)、手勢應用 (包含手勢互動應用到某一個領域，例如螢幕控制、醫療、遊戲等)、人因工程 (包含手勢互動行為和易用性研究等) 及心理學 (主要包含手勢語言學) 等四個類別。分類可交叉重疊，最後對篩選分類後的文獻進行列表整理。紀錄文獻的篇名、發表日期、出版國家、出版期刊、作者及文獻應用領域或主題等。

本研究樣本僅限於 2 個電子資料庫，且搜索關鍵字為“hand gesture interaction”和“手勢”，使得資料來源具有一定的局限性。由於樣本限制，研究結果對於主題的推論僅限於關鍵字搜索的研究範圍，無法涵蓋手勢互動領域整體的發展現況。

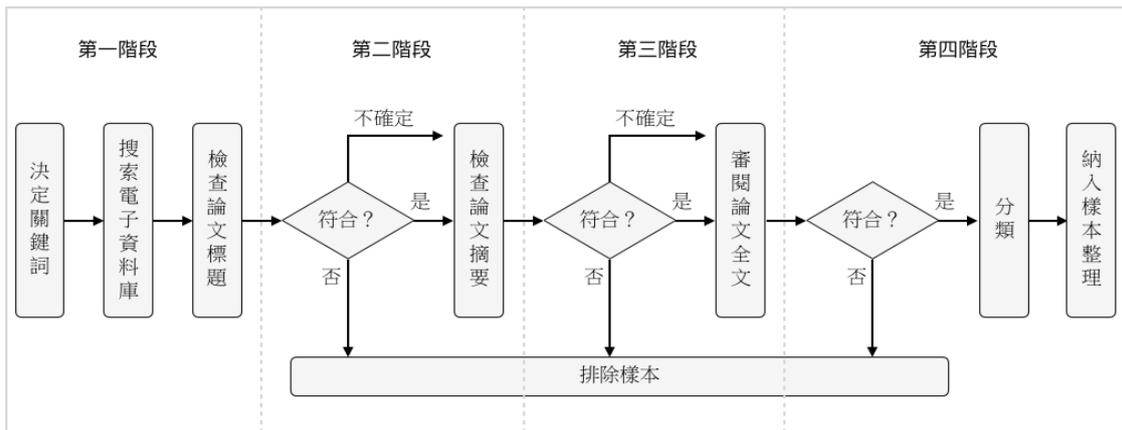


圖 1. 文獻搜索、篩選與確認過程

四、分析與發現

4-1 國際手勢互動研究發展現況

Web of Science 電子資料庫從 1990 年到 2018 年，共收錄 1004 篇手勢互動相關學術研究文獻，其中 4 篇文章被撤稿。經篩選確認後 721 篇納入樣本進行分析。其中以手勢識別的居首（338 篇，46.9%），手勢應用類的其次（312 篇，43.3%），然後是心理學類（135 篇，18.7%），後為人因工程類（63 篇，8.7%）。

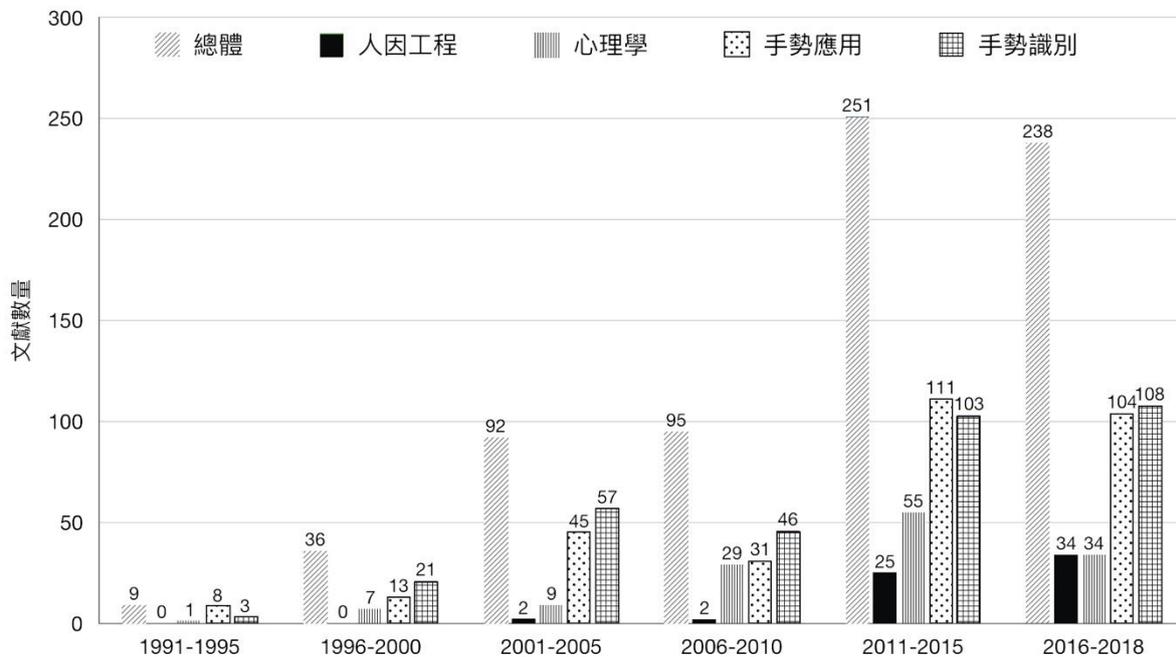


圖 2. 1991-2018 年 Web of Science 電子資料庫手勢互動研究文獻年代發表數量分析

以每 5 年為一個研究區間來分析學術研究發表篇數，手勢互動研究總體呈逐步上揚趨勢，見圖 2。尤其在 2010 年後隨著微軟 Kinect（2010 年）和 Leap Motion（2012 年）等低成本的深度攝影機推出，手勢互動研究有一個快速爆發增長的趨勢。這一趨勢在人因工程類的學術研究中尤為明顯，其次是手勢應

用和手勢識別，而在心理學方向的研究則較不明顯。另外 2001 至 2005 年和 2006 至 2010 年進行比較，手勢互動研究沒有明顯的變化，尤其在手勢應用和手勢識別方向呈輕微的下降趨勢。這和手勢互動相關技術是否普及到消費市場相關。針對國家區域而言，美國是最多發表數量的國家（114 篇），中國第 2（93 篇），南韓第 3（59 篇），臺灣有 18 篇，位居第 13 名，見表 1。在不同研究面向，多數國家區域較為均衡，但部分亞洲國家在認知心理學方向研究較為薄弱，例如中國和南韓，而臺灣在手勢應用與識別方向研究則較為領先。

表 1. 1991-2018 年 Web of Science 電子資料庫手勢互動研究文獻國家區域發表數量分析

總體			人因工程			手勢應用			手勢識別			心理學類		
國家區域	篇數	百分比	國家區域	篇數	百分比	國家區域	篇數	百分比	國家區域	篇數	百分比	國家區域	篇數	百分比
1.美國	114	15.8	1.美國	10	15.8	1.美國	48	15.3	1.中國	61	18.0	1.美國	33	24.4
2.中國	93	12.9	2.德國	7	11.1	2.中國	36	11.5	2.美國	42	12.4	2.德國	15	11.1
3.南韓	59	8.2	3.南韓	6	9.5	3.南韓	30	9.6	3.南韓	37	10.9	3.荷蘭	14	10.4
4.德國	55	7.6	4.中國	5	7.9	4.德國	27	8.6	4.德國	18	5.3	4.義大利	11	8.1
5.義大利	35	4.9	5.葡萄牙	4	6.3	5.義大利	18	5.8	5.印度	16	4.7	5.英國	10	7.4
6.英國	34	4.7	6.英國	3	4.8	6.日本	16	5.1	6.英國	14	4.1	6.法國	9	6.7
7.法國	31	4.3	7.義大利	3	4.8	7.英國	15	4.8	6.西班牙	14	4.1	7.加拿大	6	4.4
8.日本	27	3.7	8.羅馬尼亞	3	4.8	7.法國	15	4.8	7.日本	12	3.6	8.比利時	5	3.7
9.加拿大	22	3.1	7.芬蘭	2	3.2	8.臺灣	12	3.8	8.加拿大	11	3.3	9.日本	4	3.0
9.西班牙	22	3.1	7.法國	2	3.2	9.新加坡	9	2.9	9.義大利	10	3.0	9.瑞典	4	3.0
13.臺灣	18	2.5	8.臺灣	1	1.6	10.西班牙	8	2.6	11.臺灣	8	2.4	12.臺灣	1	0.7

目前 Web of Science 電子資料庫中所收錄的手勢互動研究相關學術資料期刊共 296 種。其中收錄篇數排名較前的期刊名冊見表 2 與表 3。其中在 4 個分類面向收錄篇數排名較前的期刊中有 1 本皆入選，其為《Gesture-Based Communication in Human-Computer Interaction》。這本期刊為 2 年一次的手勢互動工作坊(international gesture workshops)收錄編撰的文獻集，從 1997 年至 2011 年共 8 期。其次《Multimedia Tools and Applications》在手勢識別、應用及人因工程類別中皆收錄篇數排名較前（分別為第 2 名、第 2 名、並列第 3 名）。同時在手勢應用和人因工程類收錄篇數排名較前的期刊有：1.《International Journal of Human-Computer Interaction》、2.《Computers & Graphics》、3.《Virtual Reality》、4.《International Journal of Human-Computer Studies》、5.《Personal and Ubiquitous Computing》。同時在手勢應用和手勢識別類收錄篇數排名較前的期刊有：1.《Computer Vision and Image Understanding》、2.《Pattern Recognition Letters》、3.《Image and Vision Computing》、4.《Sensors》。認知心理學類收錄篇數較前的期刊較為獨立，其中同時在整體收錄排名中靠前的期刊有：1.《PLOS ONE》、2.《Frontiers in Psychology》。

表 2. 1991-2018 年 Web of Science 資料庫手勢研究文獻收錄篇數較前的期刊 (心理學與手勢識別應用)

心理學		手勢識別				手勢應用					
排 位	期 刊	篇 數	百 分 比	排 位	期 刊	篇 數	百 分 比	排 位	期 刊	篇 數	百 分 比
1	PLOS ONE	9	6.7	1	Gesture-Based Communication in Human-Computer Interaction	19	5.6	1	Gesture-Based Communication in Human-Computer Interaction	10	3.2
2	Frontiers in Psychology	8	5.9	2	Multimedia Tools and Applications	17	5.0	2	Multimedia Tools and Applications	9	2.9
3	Neuropsychologia	6	4.4	3	Pattern Recognition	12	3.6	3	Sensors	9	2.2
4	Journal of Nonverbal Behavior	5	3.7	4	Computer Vision and Image Understanding	9	2.7	3	Virtual reality	9	2.2
5	Gesture	4	3.0	5	IEEE Transactions on Multimedia	8	2.4	4	Computer Vision and Image Understanding	5	1.6
5	Journal of Pragmatics	4	3.0	5	Pattern Recognition Letters	8	2.4	4	Computers & Graphics-UK	5	1.6
5	Language Resources and Evaluation	4	3.0	6	Image and Vision Computing	7	2.1	4	Image and Vision Computing	5	1.6
6	Experimental Brain Research	3	2.2	6	Machine Vision and Applications	7	2.1	4	International Journal of Advanced Robotic Systems	5	1.6
6	Frontiers in Human Neuroscience	3	2.2	6	Sensors	7	2.1	4	International Journal of Human-Computer Studies	5	1.6
6	Gesture-Based Communication in Human-Computer Interaction	3	2.2	7	Neural Computing & Applications	6	1.8	4	Journal of Multimodal User Interfaces	5	1.6
6	Human Brain Mapping	3	2.2	7	Neurocomputing	6	1.8	4	Pattern Recognition Letters	5	1.6
6	Journal of Memory and Language	3	2.2	8	Engineering Applications of Artificial Intelligence	5	1.5	4	Personal and Ubiquitous Computing	5	1.6
6	Speech Communication	3	2.2	8	Gesture in Human-Computer Interaction and Simulation	5	1.5	4	Universal Access in the Information Society	5	1.6
				8	IEEE Transactions on Consumer Electronics	5	1.5				
				8	International Journal of Advanced Robotic Systems	5	1.5				
				8	Journal of Visual Languages and Computing	5	1.5				

表 3. 1991-2018 年 Web of Science 資料庫手勢互動研究文獻收錄篇數較前的期刊 (總體和人因工程)

總體		人因工程					
排 位	期刊	篇 數	百 分 比	排 位	期刊	篇 數	百 分 比
1	Gesture-Based Communication in Human-Computer Interaction	29	4	1	International Journal of Human-Computer Interaction	5	7.9
2	Multimedia Tools and Applications	23	3.2	2	Computers & Graphics -UK	4	6.3
3	Sensors	14	1.9	2	International Journal of Human-Computer Studies	4	6.3
4	Computer Vision and Image Understanding	13	1.8	2	Universal Access in the Information Society	4	6.3
4	Pattern Recognition	13	1.8	3	Applied Ergonomics	3	4.8
5	Pattern Recognition Letters	11	1.5	3	Interacting with Computers	3	4.8
6	PLOS ONE	10	1.4	3	Multimedia Tools and Applications	3	4.8
7	Gesture in Human-Computer Interaction and Simulation	9	1.2	3	Virtual reality	3	4.8
7	IEEE Transactions on Multimedia	9	1.2	4	Gesture-Based Communication in Human-Computer Interaction	2	3.2
8	Frontiers in Psychology	8	1.1	4	Human Factors	2	3.2
8	International Journal of Advanced Robotic Systems	8	1.1	4	Personal and Ubiquitous Computing	2	3.2
8	Journal on Multimodal User Interfaces	8	1.1				
8	Machine Vision and Applications	8	1.1				
8	Personal and Ubiquitous Computing	8	1.1				
8	Virtual reality	8	1.1				
9	Computers & Graphics -UK	7	1				
9	Image and Vision Computing	7	1				
9	International Journal of Human-Computer Interaction	7	1				
9	International Journal of Human-Computer Studies	7	1				
9	Journal of Visual Languages and Computing	7	1				

最後為進一步展示手勢互動研究的全貌，本研究對收錄文獻的應用場景、採用技術及所關注的主題進行了統計分析。其中在關於應用場景的輸出介面中，探討將手勢輸入用於螢幕中 3D 物件的操作和 3D 空間探索的場景最多 (46 篇)，其次是透過手勢對傳統的桌面系統進行操作，例如導航、目標指向選擇等 (40 篇)，然後是將手勢互動應用於 AR 介面 (27 篇)、VR 介面 (21 篇)、移動設備包括手機和可穿戴設備 (8 篇)。在應用領域方面，人與機器人之間的手勢互動最受關注 (65 篇)，其次是醫療領域包括醫療輔助和健康恢復等 (30 篇)、殘疾人與計算機的互動 (30 篇)，然後是應用於智慧家庭 (8 篇)、汽車 (5 篇)、遊戲 (5 篇) 等。手勢互動研究中採用最多的是以深度攝影機為基礎的手勢識別技術 (429 篇)，其次是採用手套進行手勢識別 (33 篇)，然後是通過 IMU 和肌電圖 (EMG) 感測器進行手勢識別 (分別為 23 篇、11 篇)。在手勢互動研究所關注的主題中，排除手勢識別算法主題外，涉及最多的主

題是多重感官 (multimodal) 輸入 (146 篇)，即人們通過手勢、身體語言、聲音、眼睛等多種輸入模式與計算機進行自然互動。其次是手勢指向和選擇 (23 篇)，即人們通過手勢指向來控制游標移動並對目標進行選擇。然後是探討如何選擇合理的手勢來對應與計算機互動的指令 (18 篇)、手勢互動的反饋研究 (17 篇)、不同年齡對手勢互動的影響 (8 篇)，選單設計和多人協作下的手勢互動研究 (各 4 篇)。本研究於第五章針對這些主題中已經被充分研究且可以借鑑的原則和未來的研究重點進行進一步的探討。

4-2 臺灣手勢互動研究發展現況

臺灣期刊論文索引系統經檢索共收錄 131 篇與手勢相關學術研究文獻，經篩選確認 36 篇文獻與手勢互動研究相關，納入樣本進行分析。其中手勢應用的文獻 18 篇、手勢識別的文獻 16 篇、心理學的文獻 6 篇 (皆屬於語言學類)、易用性研究的文獻 5 篇。樣本中最早關於手勢互動的研究發表於 1997 年，主要探討手勢識別技術。2000 年到 2005 年之間發表 4 篇文獻，其中 3 篇亦探討手勢識別技術、1 篇探討語言學。2005 年到 2010 年之間共發表 3 篇文獻，其中 2 篇探討手勢識別技術、1 篇探討語言學。2010 年到 2015 年臺灣手勢互動研究有一個爆發式的增長，共發表 22 篇文獻。其中探討手勢應用研究的文獻 13 篇，主要應用於教育領域 (4 篇)、機器人互動 (3 篇)、遊戲 (3 篇，2 篇為遊戲式學習)、透過手勢進行系統解鎖 (2 篇)、傳統桌面系統類應用 (2 篇)、移動設備 (1 篇)。探討手勢識別技術的文獻 10 篇，探討語言學類的文獻 4 篇。在手勢應用類的文獻中，有 4 篇對相關互動系統進行了易用性研究，但多關注於手勢互動系統本身的優勢。例如呂俊宏等人探討體感手勢游標和滑鼠在遊戲學習時的接受度，研究顯示體感手勢游標接受度高於滑鼠 (呂俊宏等人，2015)。2016 年到 2018 年之間發表 5 篇文獻，且皆屬於手勢應用類。其中林信志等人以人機互動 GOMS 理論為基礎進行手勢互動內容創作系統設計，並通過李克特 5 點量表問卷進行滿意度的評估 (林信志、朱菀鈴、朱昱融，2016)。總體而言，目前手勢互動研究，多著眼於手勢的識別技術和手勢互動應用到具體場景中的效用和滿意度評估，而少有對以手勢為互動基礎的設計原則與方法進行研究。

五、手勢互動設計原則與未來研究重點

隨著深度攝影機等低成本可偵測手勢的設備進入消費市場，以手勢為互動基礎的新輸入模式越來越受到各界的關注和興趣。本研究通過檢索手勢互動研究的相關文獻，來檢討過去已經存在的經驗和未來發展研究的重點。本研究重點從設計的角度切入展開，因此與手勢識別技術及演算法相關的內容並不在本研究探討範圍之內。本研究根據第四章手勢互動中涉及主題的文章數量並依據 Rempel、Camilleri 和 Lee (2015) 所提出的手勢互動框架進行歸納整理。透過分析相關文獻發現以下五個方向是現今研究涉及最多的主題：1. 手勢如何替代滑鼠進行目標指向和選擇；2. 不同於滑鼠，手勢具有更豐富的表達性，如何選擇合適的手勢來產生指令；3. 手勢輸入如何應用於多重感官互動之中；4. 以手勢為互動基礎的介面和傳統的 UI 介面有什麼樣的區別；5. 以手勢為互動基礎的設計如何來驗證其效用、效率和滿意度。借助於 Rempel 的手勢互動框架，本研究將這五個主題重新建構在新的研究框架之下，如圖 3 所示。以下將分別探討各主題中已經存在的經驗和未來發展研究的重點。其中第 1、2、3 節內容重點探討手勢作為新的輸入方式所帶來的挑戰，第 4 節重點探討通過手勢輸入進行互動所帶來的對於輸出介面的設計挑戰，第 5 節則綜合描述這些挑戰下如何驗證其易用性。

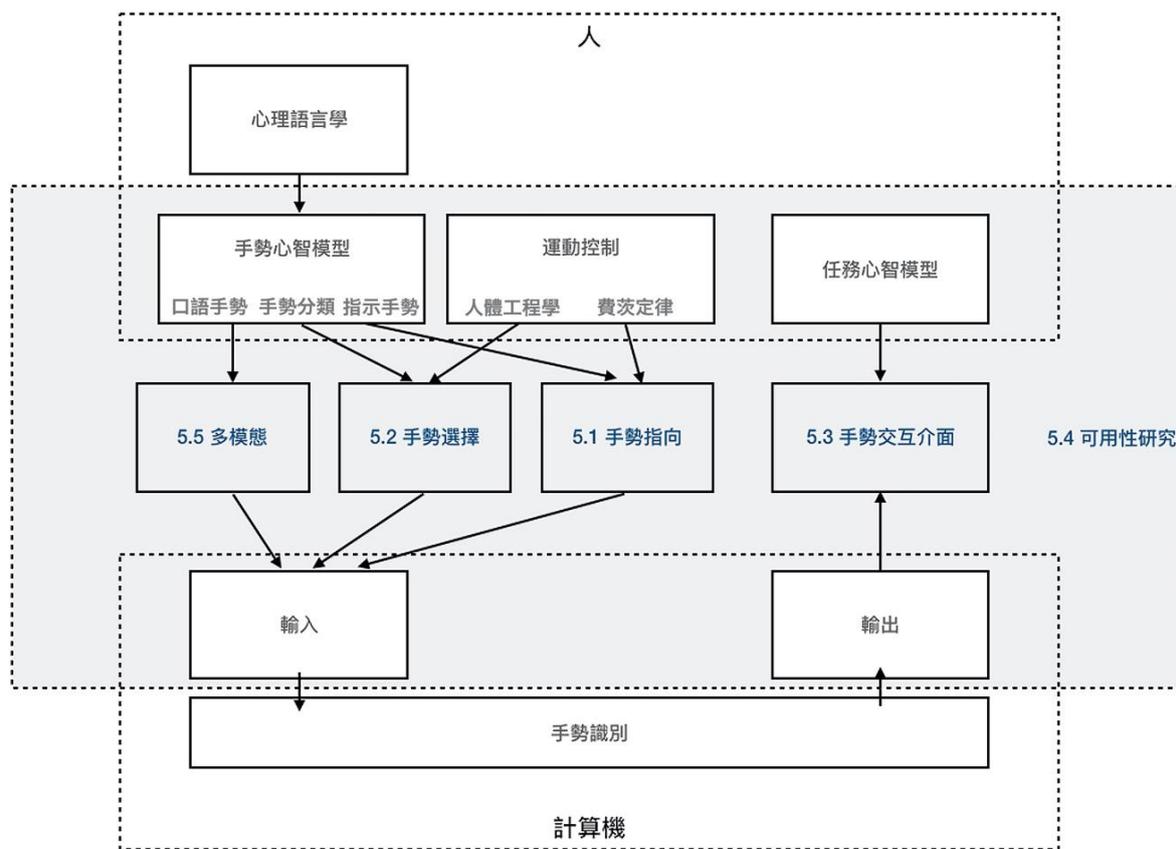


圖 3. 五個手勢互動未來研究重點在手勢互動框架中的關係和位置

5-1 手勢指向與目標選擇

5-1.1 手勢指向方法

滑鼠透過位移來控制游標的移動，但是手勢控制游標的方式卻不盡相同。其控制游標的方式可以歸納為三種方法：抓取 (grab)、投射 (raycasting) 和滑鼠 (mouse) (Jota, Pereira, & Jorge, 2009; Vogel & Balakrishnan, 2005)。抓取方法指日常生活中我們用手去抓取某個物品，再將其放到新的位置的方式，此方法從手的空間位置來確定游標在介面中的位置。投射方法則模擬我們日常生活中採用一些指示手勢，透過指定方向選擇特定的物品，再指定新的位置來放置的方式。此方法透過手轉動的角度來確定游標在介面中的位置。滑鼠方法則在與螢幕平行的平面內模擬滑鼠的工作原理，即透過手的移動距離和方向來確定游標在介面中的移動距離和方向。除此之外，手的空間位置變化和游標的移動距離之間的關係比例 (control-to-display gain) 亦會影響游標控制的性能。根據比例關係可以將其分為絕對定位和相對定位兩種方式。兩種比例關係結合三種互動方式，本研究將手勢指向歸納為 6 種指向技術或混合使用這些技術來控制游標的移動，見表 4 所示。因其中一種游標控制方式並無研究案例，未列於下表中。

表 4. 文獻中採用的手勢控制游標移動技術

手勢指向技術	研究案例（指向技術名稱由各文獻作者定義）
相對抓取	gestural pointing (Sambrooks & Wilkinson, 2013) magnetic cursor (Mäkelä, Heimonen, & Turunen, 2014) personal space (Jude, Poor, & Guinness, 2014)
絕對投射	absolute position finger ray casting (Vogel & Balakrishnan, 2005) bubble Ray (Tse, Hancock, & Greenberg, 2007) laser Pointing (Jota, Nacenta, Jorge, Carpendale, & Greenberg, 2010)
相對投射	relative ray-casting (Bateman, Mandryk, Gutwin, & Xiao, 2013) gyro acc (Nancel, Pietriga, Chapuis, & Beaudouin-Lafon, 2015)
相對滑鼠	relative pointing with clutching (Vogel & Balakrishnan, 2005) myopoint (Haque, Nancel, & Vogel, 2015) relative pointing (Liang, Xu, & Zhou, 2017) PA based pointing (Lou et al., 2017)
混合模式	hybrid technique (Vogel & Balakrishnan, 2005) laser+gyro (Nancel et al., 2015)

絕對定位的抓取方法因為手臂的物理長度和感知度並不適合與大螢幕互動 (Lou et al., 2017)，即便是近距離的小螢幕互動也因為較大幅度的手臂運動而造成疲勞，以致於很少被應用於手勢互動之中 (Jota et al., 2009)。但是在 AR、VR 等虛擬環境的近距離指向中則偏好用這種方式來選擇目標 (Kim & Lee, 2016; Shim, Yang, Kang, Seo, & Han, 2016)。相對定位的抓取方法避免了手臂大幅度的運動，但仍受限於手的運動範圍。但因其簡單直接的位置對應符合使用者的認知而被諸多學者採用 (Jude et al., 2014; Mäkelä et al., 2014; Sambrooks & Wilkinson, 2013)。

絕對定位的投射方法一般透過手指投射出的線和螢幕的交叉點來確定游標的位置，因其小幅度的運動就能指向目標而具有最高的操作效率，但是選擇的精確度較差且操作錯誤率較高 (Jota et al., 2009; Vogel & Balakrishnan, 2005)。隨著使用者與螢幕距離的增大，該操作方式因為手的抖動而導致指向精確度降低的問題會被急遽地放大。Nancel 等人使用相對投射方法來改善手臂抖動帶來的影響，從而提高小目標選擇上的精確度和效率 (Nancel et al., 2015)。Jota 等人對投射方法進行了擴展，展示了 4 種不同的投射方法以應對不同的場景 (Jota et al., 2010)。

滑鼠本身已經成為了日常生活中人們極為熟悉的一個互動方式 (Fikkert, Vet, Veer, & Nijholt, 2010)，所以能夠很自然地延用至手勢游標控制的心智模型。該操作方式往往透過動態調整 control-to-display gains 來獲得較高的指向效率和精確度，所以基本上不存在絕對定位的滑鼠方法。不同於絕對定位，相對定位的技術包括相對滑鼠和投射方法，都面臨游標離合和 control-to-display gains 設置的問題 (Nancel et al., 2015; Vogel & Balakrishnan, 2005)。相對於單獨採用一種指向方法，部分學者則混合使用多種指向方法來平衡指向的效率和精確度 (Nancel et al., 2013; Vogel & Balakrishnan, 2005)。

不同的手勢指向方法有各自的優缺點，Jota 等人 (2009) 認為這三種方法適用於不同距離的互動，且移動到目標的距離對這三種方式的操作亦有影響。但是不同的方式下手的移動軌跡和偏好並不相同，少有文獻對如何在不同的場景下選擇合適的指向方式進行研究。這就要求未來的研究者在手勢指向研究中需要明確指出研究中採用的手勢指向方式，並盡可能採用多種比對的方式對其易用性進行驗證。而現今的諸多手勢指向應用研究中忽視了手勢互動方式的不同，也未對採用的指向方法做出明確的闡述。因此導致許多研究結果僅具有片面性，而無法看到不同場景中有關不同的指向方法的差異。

5-1.2 費茨定律

在目標指向中，費茨定律被廣泛應用於評估指向的性能 (MacKenzie, 1992)。1954 年 Fitts 模擬 Shannon 的訊息傳輸理論將其應用於描述人類的運動行為，提出人類運動的能力 (IP)、目標寬度 (W) 和移動距離 (D) 呈對數線性關係 (Fitts, 1954)，現今在 HCI 領域被大家普遍採用的公式如下：

$$MT = a + b \log_2 \left(\frac{D}{W} + 1 \right) \quad (1)$$

其中 a 和 b 是由經驗確定的常數，MT 是移動時間。對數公式 $\log_2 \left(\frac{D}{W} + 1 \right)$ 被稱為難度係數 (ID)。但是透過投射方法操控游標時，人們藉由旋轉手腕或手臂來控制游標的位置而非移動的距離。基於這點，Kopper 等人對以投射方法為基礎的遠距離指向進一步延伸了費茨定律 (Kopper, Bowman, Silva, & McMahan, 2010)。其公式如下：

$$ID = \left[\log_2 \left(\frac{\alpha}{\omega^\kappa} + 1 \right) \right]^2 \quad (2)$$

其中 α 是移動角度， ω 是目標寬度轉換成的角度，因為 α 和 ω 不成線性關係，故增加係數 κ 。同時對整個對數公式進行了平方處理，是因為隨著使用者和螢幕之間的距離增加，ID 並非呈線性增加。透過滑鼠方法來操控游標時，則因為互動方式不受角度影響，仍適用於原始的費茨定律 (公式 1)，但是相關係數相比滑鼠等傳統輸入設備仍有下降 (Burno, Wu, Doherty, Colett, & Elnaggar, 2015)。

利用費茨定律模型，游標指向可以透過減少目標距離和增加目標寬度來提高指向效率 (Balakrishnan, 2004)。這種技術被廣泛應用於以手勢為基礎的目標指向中，例如 magnetic cursor (Mäkelä et al., 2014)、sticky targets (Bateman et al., 2013) 等。在這些技術中常常會對 CD gains 進行調整。一個較高的 CD gain 會使游標的操控變得靈敏，但是過於靈敏的操控會使得小目標的指向變得極為困難。一個較低的 CD gain 則使得游標的操控過於遲鈍 (無法快速地移動到目標) 甚至產生離合問題 (輸入的最大距離無法使游標跨越整個螢幕，而需要移動多次來使游標移動到目標)。PA 技術則根據使用者移動速率的不同，定義不同的 CD gains 來優化使用者對游標的移動控制 (Casiez & Roussel, 2011; Nancel et al., 2015)。它可以讓使用者在粗略地或精確地選擇目標時對速度效率和精確度有一個較好的平衡。Nancel 等人 (2015) 在超大型螢幕手勢互動研究中定義了 S 型 PA 轉化函數。其函數定義如下：

$$CD(x) = \frac{CD_{max} - CD_{min}}{1 + e^{-\lambda(x - V_{inf})}} + CD_{min} \quad (3)$$

其中 $V_{inf} = RATIO_{inf} * (V_{max} - V_{min}) + V_{min}$ 。 V_{max} 和 V_{min} 是手能夠做出精確選擇的最大和最小速度。 $RATIO_{inf}$ 決定 S 型曲線的拐點，其值通常為 0.5。小於 0.5 則使游標的加速變得更為明顯，反之亦然。 CD_{max} 和 CD_{min} 是 CD gain 值的上限和下限漸近值。其定義為： $CD_{max} = \frac{A_{max}}{(c+1) DIM_{input}}$ ， A_{max} 是螢幕可移動的最大距離， DIM_{input} 是輸入設備能夠感知的最大距離或手能夠移動的最大距離， c 是能夠接受的離合次數 CD_{min} 則是螢幕最小距離，Nancel 認為對於大型螢幕可以用最小選擇目標寬度代替) 和手可移動最小距離 (對於傳感器而言則為傳感器能夠感受到的最小移動距離) 之比。 λ 和 S 型曲線拐點處的斜率相關，用以平衡實際操作的 CD gains 範圍和轉化快慢的關係。趨近於 0 的 λ 值會導致 CD gains 值範圍過小而產生離合和選擇精度問題，過高的 λ 值則導致 V_{inf} 位置的操作過於靈敏而很難控制游標。Nancel 認為 λ 的初始值為 $\frac{4}{V_{max} - V_{min}}$ 可以覆蓋 76% 的 CD gains 範圍。Nancel 定義 V_{max} 和 V_{min} 分別為輸入設備採集的手勢快速移動速度集的中位數和慢速速度集的 90% 位數。在 Lou 等人的實驗中，他們採用 ASUS Xtion PRO 深度攝影機進行手勢識別，通過初步測試定義 V_{max} 為 0.45m/s、 V_{min} 為 0.06m/s、 CD_{min} 為 0.2、 $RATIO_{inf}$ 為 0.5、 λ 為 20 (Lou et al., 2017)。

總體而言，相對於傳統輸入設備，以手勢為互動基礎的指向更難以預測其模型。而且手的空間移動在不同的方向存在顯著差異（Chen & Wang, 2020）。簡單的費茨定律並不能準確描述手勢移動的軌跡。除此之外，手在進行指向時手指的移動、腕部的移動以及手臂的移動亦存在不同的差異。如何改善費茨定律並結合其他方式來提高其預測精確度將是未來研究的一個重點。另外手勢指向的性能明顯低於傳統輸入設備，其不僅受限於消費級設備的精確度，同時亦受限於手的空間移動的複雜性。現今研究中僅僅透過由費茨定律而產生的目標指向技術並不能完美解決此類問題。未來的研究中需要借助手勢的語意，不同關節的指向精確度等對指向進行輔助，以此來設計更符合使用者習慣的指向方式。

5-2 手勢選擇

與傳統的輸入設備相比，手勢輸入具有更為豐富的表達性。然而這種表達性因為社會文化和個人習慣而有所不同。這種複雜性導致手勢具有較差的自我揭示性（Baudel & Beaudouin-Lafon, 1993），導致人們無法單靠探索，而需透過記憶來發現手勢和功能之間的對應關係。Morrel-Samuels 將這種複雜性概括為三個問題：stylistic problem（符號類手勢因年齡和社會文化而不同）、situational problem（手勢因時間、環境、心情、疲勞等因素影響而不同）和 description problem（描述事物時產生的手勢因人而異）（Morrel-Samuels, 1990）。根據這點，Norman 曾提出與其研究哪一種手勢的選擇更合適，不如提供一個標準進行規範（Norman & Nielsen, 2010）。部分學者則希望通過啟發式的評估方式對手勢和指令的對應關係進行選擇。其中最早也被諸多學者採用的是 Nielsen 等提出的手勢選擇和評估方式，其共 4 個步驟：第一步，確定需要設計手勢集的場景和功能；第二步，自上而下的方式收集手勢指令；第三步，依據其提出的原則篩選手勢集；第四步，自下而上的方式測試手勢指令並計算其錯誤率。此方式測試受測者記憶手勢指令需要的次數，並利用 5 分李克特量表測試受測者執行手勢所感知的主觀壓力。這種啟發式的評估方式被應用到諸多場景進行手勢指令的選擇研究，例如音樂播放的手勢指令選擇（Löcken, Hesselmann, Pielot, Henze, & Boll, 2012）、智慧家庭系統的手勢指令選擇（Choi, Kwon, Lee, D., Lee, H., & Chung, 2014）、機器人手臂控制的手勢選擇（Wongphati, Osawa, & Imai, 2015）等。Wobbrock、Aung、Rothrock 與 Myers（2005）則在評估階段提出通過計算同意率（A）進行量化評估，其計算公式如下：

$$A = \frac{\sum_{r \in R} \sum_{Pi \subseteq Pr} \left(\frac{|Pi|}{|Pr|} \right)^2}{|R|} * 100\% \quad (4)$$

其中 R 表示需要執行的功能數， r 表示某個功能， Pr 表示執行某個功能所產生的手勢數量， i 表示執行某個功能 r 出現的手勢， Pi 表示手勢 i 的出現次數。這種量化的評估方式也被應用到諸多研究之中，例如家庭娛樂電視的手勢指令選擇（Dong, Danesh, Figueroa, & El Saddik, 2015; Vatavu, 2012, 2013）、以 Leap Motion 為基礎的電視手勢互動指令選擇（Zaiti, Pentiu, & Vatavu, 2015）、34 個常用的人機互動指令的手勢選擇（Pereira, Wachs, Park, & Rempel, 2015）、以全身動作為基礎的手勢指令選擇（Chen et al., 2018）等。除此之外，部分學者從人因工程學的角度對不同的手勢進行舒適度評估（Rempel et al., 2015; Son, Jung, & Park, 2017）。Morris 等對使用者如何提出手勢存疑，因為他們發現使用者在手勢定義時常常受傳統互動方式的影響而不能提出更多可能性的手勢（Morris et al., 2014）。

總體而言，手勢選擇的流程及評估方法已經得到了廣泛的認同，但是 Wobbrock 等人（2005）提出的手勢選擇同意率並不高。儘管啟發式的評估方式提出了易用性較高的手勢，但有鑑於較低的同意率，仍需要通過一些標準來進行規範。所以本研究認為在未來的研究中，不僅僅需要對手勢的選擇進行方法和流程上的優化，更需要對不同的手勢進行特徵的提取和分類，以利對部分手勢進行標準規範。甚至手勢的區分，不僅僅是手形，而是通過某些特徵來進行識別。這就要求未來的手勢選擇研究應該在一個合理

的框架下形成一個規範的詞彙表，例如 Gestalt 手勢詞彙表 (Jahani & Kavakli, 2018)。另外，關於手勢在不同場景下的優缺點亦缺少研究。因此，提供具標準（例如手指的張開角度、移動的關節數量等）且能於不同場景使用的手勢指令詞彙表，將是未來急需進行研究的一個重點。

5-3 多重感官輸入與以手勢為基礎的互動

Wexelblat 認為手勢互動仍停留於一對一的功能對應模式，只會增加使用者對手勢記憶的認知負擔，他提倡我們應該朝向模擬人與人交流式的多重感官互動 (Wexelblat, 1998)。因此和手勢相關的多重感官輸入研究可以明顯地被分為兩個類別：以多重輸入模式為基礎的互動應用研究和以語言學為基礎的手勢互動研究。前者著眼於解決傳統設備在某些場景中存在的問題或者增加互動的沉浸感。例如：透過手勢、眼睛和語音輸入來進行游標控制從而解決單一輸入模式的弱點 (Kim, Suh, & Lee, 2017; Tse et al., 2007)，或為對傳統輸入方式有行動障礙的人群提供多重輸入方式與機器人互動 (Gomez-Donoso et al., 2017)，或透過手勢和語音進行國際象棋遊戲 (Carbini, Viallet, Bernier, & Bascle, 2005)，或在 3D 場景中進行 CT 場景探索 (Krapichler et al., 1999)，或透過在 AR 場景中進行物體操作 (Kakez, Conan, & Bisson, 1997) 等研究來增加人機互動的自然性和沉浸感。其特點是各個輸入模式並沒有模擬人與人的交流方式進行深度融合。後者則期望能夠徹底革新以 WIMP (窗口、圖標、滑鼠、點擊) 為互動基礎的模式，更關注伴隨語言的自然手勢。在心理語言學中，手勢、單詞和句子等一樣是語言系統的一部分 (McNeil, 1992)。McNeil 將伴隨語言的手勢分為：iconic (跟講述內容語意相關的手勢)、metaphoric (用於表述抽象的想法而非具體的物體或事件時產生的手勢)、deictic (指示某個事物或方向的手勢) 和 beat (傳達內容結構的一些無含義的)。Quek 等人 (2002) 基於這個心理語言學概念，提出了對伴隨語言的手勢進行數據化的框架並將其作為一種範式應用於多重感官互動之中。Wagner、Malisz 和 Kopp (2014) 等對此類伴隨語言的手勢數據化的註釋工具進行了梳理，發現手勢註釋共有的屬性包括：形狀、位置、方向、運動軌跡以及運動前後的位置等。伴隨語言的手勢技術應用則更關注手勢和語言的融合，例如 VIENA 系統 (Wachsmuth, 1999)、用於大螢幕 2D 介面互動的 iMap 框架 (Kettebekov & Sharma, 2001) 等。而應用領域最多的則是人與機器人的互動 (Burger, Ferrané, Lerasle, & Infantes, 2012; Ghidary, Nakata, Saito, Hattori, & Takamori, 2002; Jacob & Wachs, 2016)。但是伴隨語言的自然手勢互動是否能夠明顯提升人機互動的易用性仍有存疑。Abich 與 Barber 發現在機器人控制中手勢輸入和手勢語音融合輸入比單獨語音輸入具有更高的感知工作負荷 (Abich & Barber, 2017)。Lee 等人發現以 AR 為基礎的 3D 物品操作中手勢語音融合輸入與單獨語音輸入，在效率和效用並沒有顯著區別 (Lee, Billinghamurst, Baek, Green, & Woo, 2013)。Bernardis 與 Gentilucci 認為口語和手勢並非總是互相支持，相反描述性手勢會加強口語，但是口語卻阻礙手勢的產生 (Bernardis & Gentilucci, 2006)。儘管如此，以手勢與語音為基礎的多重感官輸入在易用性上具有更高的滿意度 (Abich & Barber, 2017; Chu, Dani, & Gadh, 1997)。而手勢輸入和多重感官輸入在感知工作負荷上沒有顯著區別，則被認為手勢輸入能夠支持語音輸入 (比如增加輸入的健壯性) 而不增加認知負擔 (Abich & Barber, 2017)。除此之外，伴隨語言的手勢受社會文化習俗和個人習慣的影響，變化極大 (Baudel & Beaudouin-Lafon, 1993; Morrel-Samuels, 1990)。總體來說，以多重輸入模式為基礎的互動應用研究仍將是短期時間內的研究重點，並在可預見的未來可在多種場景下替代傳統輸入設備。而以語言學為基礎的手勢互動研究，在人與機器人的互動中將具有更好的突破。但是我們對手勢和口語的融合機制及其本質仍然缺乏充分的研究 (Wagner et al., 2014)，在未來仍有很長的一段路要走。

5-4 以手勢為互動基礎的介面設計

手勢輸入不同於傳統的輸入設備，其獨有的特徵必將帶來介面上的革新。因此諸多文獻針對以手勢為基礎的選單和反饋進行了研究。本節內容將從這兩個方面展開。

手的移動軌跡具有偏好性。Lou 等人研究發現人們的右手在右邊的操作精度要高於左邊(滑鼠方法)，反之亦然 (Lou et al., 2017)。因此 UI 的佈局應該集中在某一邊而不是橫向跨越螢幕。Ren 和 O'Neill (2013) 亦發現了這個特徵，但是以 stroke 為基礎的選擇技術(投射方法)在操作上，手向下的操作要比向上具有更好的表現。而在三維操作上，手的方向性具有更為複雜的軌跡偏好，Ren 和 O'Neill (2012) 提出了“Muffin”3D marking menu，認為手在往前及左右的方向上有較好的表現。Lubos、Bruder、Ariza 和 Steinicke (2016) 認為當一個關節(手腕、肘或肩部)固定的時候手指的運動軌跡會近似一個球的部分，他發現佈局在這個球上的選單比簡單的 2D 佈局具有更好的性能。這些研究從多個方面對手的移動偏好進行研究，並將其應用在選單上。但是正如本章第一節所述，手勢指向的方式對手的移動偏好有較大的影響。比較不同學者實驗中的結果可以發現這個問題，但是大家都沒有對其展開系統性的研究。同時不同關節的移動偏好亦不同。手勢互動的 UI 佈局需要根據不同的情景做出調整。Kopper、Bacim 和 Bowman (2011) 介紹 SQUAD 技術是通過多次選擇來確定密集編排中的目標，其選單布局呈現長方形，使用者只需要通過 4 個方向來選擇特定方向中的目標。Ni、Bowman、North 和 McMahan (2011) 設計了一種呈現圓形放射形式佈局的 rapMenu，透過手的轉向來選擇其中一個區域，然後通過掐指來確定最後的命令。Bossavit 等人則透過用手觸摸人體的各個部位來選擇選單中的指令 (Bossavit, Marzo, Ardaiz, & Pina, 2013)。總體來說，如何提高選單的選擇效率、釐清手的移動軌跡偏好以及手移動的一些固有特徵對 UI 佈局產生的影響仍將是未來研究的重點。

手勢互動相較於傳統輸入設備，因為沒有力道的回饋，而更需要系統給予其他回饋來補足。其中研究較多的回饋補償為視覺、觸覺和聽覺。視覺回饋的不同主要體現在游標多變的形式上。其中圓形 (Cockburn, Quinn, Gutwin, Ramos, & Looser, 2011; Lou et al., 2017; Tse et al., 2007) 和十字形 (Mäkelä et al., 2014; Pino, Tzemis, Ioannou, & Kouroupetroglou, 2013) 是最常見的游標形式。在 3D 或虛擬現實場景中的物體操作則多採用真實的手模型進行視覺回饋 (Alkemade et al., 2017; Chen, Hu, & Yang, 2018; Lam, Arshad, Prabuwo, Tan, & Kahaki, 2018)。Moscovich 與 Hughes 則透過多個圓點對應多個手指的形式進行視覺回饋 (Moscovich & Hughes, 2006)。Matulic 與 Vogel 透過各個手指投射到螢幕的點組成的圖形來顯示手的不問輸入功能 (Matulic & Vogel, 2018)。總體來說，手勢游標可以分為具象和抽象兩種形態。具象的形式多應用於遊戲等虛擬現實場景中，而抽象的形態則多應用於選單等 2D 介面操作中。至於這兩種形態之間是否有顯著的差異，還未有系統性的研究。觸覺回饋則對手勢輸入沒有力的回饋進行直接補償，多採用震動的形式 (Buchmann, Violich, Billinghamurst, & Cockburn, 2004; Foehrenbach, König, Gerken, & Reiterer, 2009)。Park 等人為手勢操作圓形選單提供了持續 Shepard 聲調回饋和選擇操作時的深度聲調回饋，他們發現此兩種聽覺回饋可作為視覺回饋的補充，不僅提高了手勢操作的準確度並縮短了移動時間 (Park, Kim, & Lee, 2015)。總體而言，回饋能夠提高手勢互動的操作性能，但是三者之間對操作性能的影響並無顯著效應 (Cho, Lee, J.-H., Lee, B.-T., & Park, 2015; Kajastila & Lokki, 2013; Köpsel, Majaranta, Isokoski, & Huckauf, 2016)。這意味著不同於傳統輸入設備，視覺回饋並不具有最高的優先權，而在部分場景中(例如駕駛)手勢互動更依賴於其他反饋形式。本研究認為在未來的手勢互動研究中，學者應該將更多的注意力放到視覺之外的反饋形式中。除此之外，Foehrenbach 等發現觸覺回饋會阻礙其他回饋形式，降低手勢互動性能 (Foehrenbach et al., 2009)。Cho 等人 (2015) 將聲音或者震動觸覺作為前饋時(當游標接近目標邊緣時就提供回饋)發現操作性能顯著提高，他們將其原因歸於預線索效應 (precue

effect)。同時他們還發現觸覺回饋對男性和女性操作效率影響不同。增加觸覺的前饋能夠提高男性的操作效率，而對女性而言則相反(Cho et al., 2015)。總體而言，以手勢輸入為基礎的介面回饋和傳統介面有明顯的不同，這意味著以手勢為基礎的互動回饋具有更多的可能性，而不僅僅是在視覺而已。

5-5 以手勢為互動基礎的易用性研究

以手勢為互動基礎的易用性評估可分為兩個部分。對於運動控制的評估較適合透過量化的方式以總結性易用性研究的方式進行。例如手勢指向則可以採用 ISO 9241-9 制定的游標指向易用性評估標準(Burno et al., 2015; Pino et al., 2013)。對於手勢選擇則可以採用啟發式的研究形式進行並結合手勢互動的一些設計原則進行評估，見 5-2。其中易用性研究的量化和質化指標如下：

1. 涉及操作任務的效率指標，包括任務完成的時間和錯誤率。對於游標指向則可以包括操作中更細緻的量化指標包括指向能力(throughput)、目標重複進入(target re-entry)、任務軸交叉(task axis crossing)、指向方向改變(movement direction change)、正交方向改變(orthogonal direction change)、移動變化(movement variability)、移動錯誤(movement error)、移動偏移(movement offset)等(MacKenzie, Kauppinen, & Silfverberg, 2001)。

2. 任務完成後手臂感受到的疲勞度。其中測量主觀疲勞度最常用的量表是 Borg CR10(Borg, 1998)。研究表明通過 EMG 測量的肩部疲勞度和 Borg CR10 測量的肩部主觀疲勞度具有很強的相關性(Peres et al., 2009)。Hincapi-Ramos 等人通過對前臂的生物力學分析提出了 consumed endurance (CE) 指標對手臂的疲勞度進行評測(Hincapi-Ramos, Guo, Moghadasian, & Irani, 2014)。他們發現 CE 值和 Borg CR10 測量的疲勞度亦有較強的相關性。

3. 使用者對以手勢互動完成任務的滿意度評估。其中最常用的量表是 NASA-TLX 工作負荷量表(Alkemade et al., 2017; Hart & Staveland, 1988)。該量表有六個維度，通過評估受測者在任務完成過程中所感受到的心智需求、體力需求、時間需求、績效、努力程度和挫敗感等來計算最終的心智負荷指數。除此之外，亦有較多文獻通過自定義的量表來進行滿意度評估。例如 Lou 等人(2017)通過 mental effort、accuracy、speed、fatigue、comfort 與 overall easiness 等 5 個構面的 5 分李克特量表進行滿意度評估，林信志等人(2016)從畫面可視性、專業術語和系統資訊、學習因素、系統功能、系統實用性與介面設計等 5 個構面對手勢互動內容創作系統進行滿意度評估。

另外年齡及偏手性對於手勢互動具有顯著的影響，在實驗中對受測者的選擇需要充分考量這些因素(Carvalho, Bessa, Magalhães, & Carrapatoso, 2018; Hsu, Huang, Tsuang, & Sun, 1999)。Edinburgh handedness inventory (EHI) 是 Oldfield 設計的用於快速簡單地定量評估偏手性的方法，是目前實驗中測試偏手性使用最為廣泛的量表(Oldfield, 1971)。易用性測試已有一個較為規範的方法和流程，但是以手勢為基礎的輸入系統易用性研究需要從單個手形到整體系統等多個層次進行綜合評估。例如 Barclay、Wei、Lutteroth 和 Sheehan (2011) 提出以 4 個層級對手勢互動系統進行量化評估：1. 影響手勢質量的因素、2. 單獨手勢的整體質量、3. 系統中特定功能或用例的質量、4. 整體系統的質量。本研究認為有關手勢指向的研究，應該嚴格遵守 ISO 9241-9 制定的游標指向易用性評估標準或在其基礎上進行改進，方便已經存在的研究和未來的研究進行對比。而對於手勢互動的應用研究則應該在多個層次上進行易用性研究。簡單的滿意度量表無法清晰地展示系統的易用性問題存在於哪個層級，但是如何在不同的層次定義易用性評估指標以及如何進行綜合評估仍缺乏充分研究。因此，本研究認為如何透過不同層次來對以手勢為互動基礎的系統進行綜合評價將是未來研究的重點。

六、結論與建議

20 世紀 80 年代以手勢輸入為基礎的互動研究，進入學術研究的視野開始得到越來越多的關注。在 2010 年後隨著低成本的深度攝影機和 IMU 等感測器進入消費市場，以手勢為輸入基礎的互動設計的研究迅速成為一大熱點，文獻數量有一個爆炸式的增長。綜觀其發展的過程，相比其他專業學類，設計學科對手勢互動的關注明顯較少。而目前則鮮有文獻探討以手勢為輸入基礎的互動設計和傳統設計的區別。但隨著自然互動介面的到來，手勢輸入必將在各類領域扮演極為重要的角色。設計研究領域需要在這個時代的到來之前提前做好準備，提供充分的設計基礎研究。本研究檢索 Web of Science 電子資料庫自 1990 年起至 2018 年共 721 篇文獻，及臺灣期刊論文索引系統至 2018 年共 36 篇文獻展開探討，嘗試呈現手勢互動的發展脈絡和現況，並提出設計研究需要在手勢互動領域深入探討的主題、已經存在的經驗和未來研究發展的重點。

透過分析本研究發現手勢互動研究以手勢的識別研究最多，其次是以手勢為輸入基礎的應用研究，再次是以心理語言學為基礎的手勢研究，最後是人因工程類研究。部分亞洲國家地區在手勢互動研究上帶有明顯偏好，重手勢識別與應用研究，輕手勢語言學研究。收錄手勢互動研究的期刊常常跨越多個研究方向，但是心理語言學類期刊則較為獨立。通過對主要涉及的主題進行分析，本研究發現設計學科可以在五大主題下進行深入研究：1.手勢如何替代滑鼠進行目標指向和選擇，本研究整理歸納了 6 種手勢指向的方式（其中 1 種手勢指向方式鮮少被採用），但是不同的方式間的比較還有待深入探討。研究人員同時也可以進一步探索新的方式來解決現有手勢指向容易疲勞且指向精確度低的問題，同時用來更精確預測手勢指向性能的模型亦有待研究；2.如何選擇合適的手勢來產生指令，雖然已有諸多文獻探討手勢選擇的方法，但是如何對選擇後的手勢集進行定義並形成可比較的詞彙表仍然缺乏研究，同時如何得到同意率較高的手勢甚至形成一種社會共識仍是一大難題；3.人與人交流的多重感官形式如何應用於人機互動之中；4.以手勢為互動基礎的介面和傳統的 UI 介面有什麼樣的區別。UI 介面的佈局、新的控件設計以及介面的回饋值得進一步研究；5.以手勢為互動基礎的設計如何來驗證其效用、效率和滿意度可以從手形、手勢、以手勢為互動基礎的任務和系統 4 個層級來探討易用性研究的方法和評估指標，或可以從不同模式的優缺點及多重感官的深度融合 2 個方面來進行進一步的研究。限於篇幅和檢索的資料，本研究對於會議論文中關於手勢互動的最新動態及諸多本研究探討的部分理論未做進一步展開，僅著眼於探討手勢互動學術研究中已經存在的經驗供實務設計之參考，整理手勢互動設計脈絡，並歸納未來研究發展重點供學術研究之參考。

參考文獻

1. Abich, J., & Barber, D. J. (2017). The impact of human-robot multimodal communication on mental workload, usability preference, and expectations of robot behavior. *Journal on Multimodal User Interfaces*, 11(2), 211-225. doi:10.1007/s12193-016-0237-4
2. Alkemade, R., Verbeek, F. J., & Lukosch, S. G. (2017). On the efficiency of a VR hand gesture-based interface for 3D object manipulations in conceptual design. *International Journal of Human-Computer Interaction*, 33(11), 882-901. doi:10.1080/10447318.2017.1296074
3. Balakrishnan, R. (2004). "Beating" fitts' law: Virtual enhancements for pointing facilitation. *International Journal of Human-Computer Studies*, 61(6), 857-874. https://doi.org/10.1016/j.ijhcs.2004.09.002

4. Barclay, K., Wei, D., Lutteroth, C., & Sheehan, R. (2011). A quantitative quality model for gesture based user interfaces. *Proceedings of the 23rd Australian Computer-Human Interaction Conference* (pp. 31-39). Canberra: University of Auckland.
5. Bateman, S., Mandryk, R. L., Gutwin, C., & Xiao, R. (2013). Analysis and comparison of target assistance techniques for relative ray-cast pointing. *International Journal of Human-Computer Studies*, 71(5), 511-532. <https://doi.org/10.1016/j.ijhcs.2012.12.006>
6. Baudel, T., & Beaudouin-Lafon, M. (1993). Charade: Remote control of objects using free-hand gestures. *Communication of the ACM*, 36(7), 28-35. doi:10.1145/159544.159562
7. Bernardis, P., & Gentilucci, M. (2006). Speech and gesture share the same communication system. *Neuropsychologia*, 44(2), 178-190. <https://doi.org/10.1016/j.neuropsychologia.2005.05.007>
8. Bolt, R. A. (1980). Put-that-there: Voice and gesture at the graphics interface. *SIGGRAPH Computer Graphics*, 14(3), 262-270. doi:10.1145/965105.807503
9. Borg, G. (1998). *Borg's perceived exertion and pain scales*. Champaign, IL: Human Kinetics.
10. Bossavit, B., Marzo, A., Ardaiz, O., & Pina, A. (2013). Hierarchical menu selection with a body-centered remote interface. *Interacting with Computers*, 26(5), 389-402. doi:10.1093/iwc/iwt043
11. Buchmann, V., Violich, S., Billinghamurst, M., & Cockburn, A. (2004). FingARtips: Gesture based direct manipulation in augmented reality. *Proceedings of the 2nd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia* (pp. 212-221). Singapore: HIT Lab NZ.
12. Burger, B., Ferrané, I., Lerasle, F., & Infantes, G. (2012). Two-handed gesture recognition and fusion with speech to command a robot. *Autonomous Robots*, 32(2), 129-147. doi:10.1007/s10514-011-9263-y
13. Burno, R. A., Wu, B., Doherty, R., Colett, H., & Elnaggar, R. (2015). Applying fitts' law to gesture based computer interactions. *Procedia Manufacturing*, 3(Supplement C), 4342-4349.
14. Carbini, S., Viallet, J. E., Bernier, O., & Bascle, B. (2005). Tracking body parts of multiple people for multi-person multimodal interface. In N. Sebe, M. Lew, T. S. Huang (Eds.), *Computer Vision in Human-Computer Interaction. Lecture Notes in Computer Science, Vol. 3766* (pp. 16-25). Berlin: Springer.
15. Carvalho, D., Bessa, M., Magalhães, L., & Carrapatoso, E. (2018). Performance evaluation of different age groups for gestural interaction: A case study with Microsoft Kinect and Leap Motion. *Universal Access in the Information Society*, 17(1), 37-50. doi:10.1007/s10209-016-0518-4
16. Casiez, G., & Roussel, N. (2011, October). No more bricolage! Methods and tools to characterize, replicate and compare pointing transfer functions. *Paper presented at the 24th Annual ACM Symposium on User Interface Software and Technology Conference*. Santa Barbara, California, USA.
17. Chen, C.-H., & Wang, J.-L. (2020). The semantic meaning of hand shapes and Z-dimension movements of freehand distal pointing on large displays. *Symmetry*, 12(3), 329.
18. Chen, P., Hu, Y., & Yang, F. (2018). A conformal geometric algebra method for virtual hand modeling and interaction. *EURASIP Journal on Image and Video Processing*, 2018(1), 72.
19. Chen, X. A., Schwarz, J., Harrison, C., Mankoff, J., & Hudson, S. E. (2014). Air+touch: Interweaving touch & in-air gestures. *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (pp. 519-525). Honolulu, HI: Carnegie Mellon University.
20. Chen, Z., Ma, X., Peng, Z., Zhou, Y., Yao, M., Ma, Z., . . . Shen, M. (2018). User-defined gestures for gestural interaction: Extending from hands to other body parts. *International Journal of Human-Computer*

- Interaction*, 34(3), 238-250. doi:10.1080/10447318.2017.1342943
21. Cho, K., Lee, J.-H., Lee, B.-T., & Park, E. (2015). Effects of feedforward in in-air remote pointing. *International Journal of Human-Computer Interaction*, 31(2), 89-100. doi:10.1080/10447318.2014.959107
 22. Choi, E., Kwon, S., Lee, D., Lee, H., & Chung, M. K. (2014). Towards successful user interaction with systems: Focusing on user-derived gestures for smart home systems. *Applied Ergonomics*, 45(4), 1196-1207.
 23. Chu, C.-C. P., Dani, T. H., & Gadh, R. (1997). Multi-sensory user interface for a virtual-reality-based computeraided design system. *Computer-Aided Design*, 29(10), 709-725.
 24. Cockburn, A., Quinn, P., Gutwin, C., Ramos, G., & Looser, J. (2011). Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback. *International Journal of Human-Computer Studies*, 69(6), 401-414. <https://doi.org/10.1016/j.ijhcs.2011.02.005>
 25. Dam, A. v. (1997). Post-WIMP user interfaces. *Communication of the ACM*, 40(2), 63-67.
 26. Dong, H. W., Danesh, A., Figueroa, N., & El Saddik, A. (2015). An elicitation study on gesture preferences and memorability toward a practical hand-gesture vocabulary for smart televisions. *IEEE Access*, 3, 543-555. doi:10.1109/access.2015.2432679
 27. Fikkert W., van der Vet P., van der Veer G., & Nijholt A. (2010). Gestures for large display control. In S. Kopp & I. Wachsmuth (Eds.), *Gesture in embodied communication and human-computer interaction. Lecture Notes in Computer Science, vol 5934* (pp. 245-256). Berlin: Springer.
 28. Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 381-391. doi:10.1037/h0055392
 29. Foehrenbach, S., König, W. A., Gerken, J., & Reiterer, H. (2009). Tactile feedback enhanced hand gesture interaction at large, high-resolution displays. *Journal of Visual Languages & Computing*, 20(5), 341-351. <https://doi.org/10.1016/j.jvlc.2009.07.005>
 30. Ghidary, S. S., Nakata, Y., Saito, H., Hattori, M., & Takamori, T. (2002). Multi-modal interaction of human and home robot in the context of room map generation. *Autonomous Robots*, 13(2), 169-184.
 31. Gomez-Donoso, F., Orts-Escolano, S., Garcia-Garcia, A., Garcia-Rodriguez, J., Castro-Vargas, J. A., Ovidiu-Oprea, S., & Cazorla, M. (2017). A robotic platform for customized and interactive rehabilitation of persons with disabilities. *Pattern Recognition Letters*, 99, 105-113.
 32. Goth, G. (2011). Brave NUI world. *Communication of the ACM*, 54(12), 14-16.
 33. Haque, F., Nancel, M., & Vogel, D. (2015, April). Myopoint: pointing and clicking using forearm mounted electromyography and inertial motion sensors. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3653-3656). Seoul: University of Waterloo.
 34. Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139-183.
 35. Hincapi-Ramos, J. D., Guo, X., Moghadasian, P., & Irani, P. (2014). Consumed endurance: A metric to quantify arm fatigue of mid-air interactions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1063-1072). Toronto: University of Manitoba.
 36. Hsu, S. H., Huang, C. C., Tsuang, Y. H., & Sun, J. S. (1999). Effects of age and gender on remote pointing performance and their design implications. *International Journal of Industrial Ergonomics*, 23(5), 461-471.
 37. Jacob, M. G., & Wachs, J. P. (2016). Optimal modality selection for cooperative human-robot task completion. *IEEE Transactions on Cybernetics*, 46(12), 3388-3400. doi:10.1109/TCYB.2015.2506985

38. Jahani, H., & Kavakli, M. (2018). Exploring a user-defined gesture vocabulary for descriptive mid-air interactions. *Cognition, Technology & Work*, 20(1), 11-22. doi:10.1007/s10111-017-0444-0
39. Jota, R., Nacenta, M. A., Jorge, J. A., Carpendale, S., & Greenberg, S. (2010, September). A comparison of ray pointing techniques for very large displays. *Paper presented at the Conference of Graphics Interface 2010*. Ottawa, Ontario, Canada.
40. Jota, R., Pereira, J. M., & Jorge, J. A. (2009). A comparative study of interaction metaphors for large-scale displays. *Proceedings of the 2009 CHI Extended Abstracts on Human Factors in Computing Systems* (pp. 4135-4140). Boston, MA: Inesc-ID .
41. Jude, A., Poor, G. M., & Guinness, D. (2014). Personal space: User defined gesture space for GUI interaction. *Proceedings of the 2014 CHI Extended Abstracts on Human Factors in Computing Systems* (pp. 1615-1620). Toronto: Baylor University.
42. Kajastila, R., & Lokki, T. (2013). Eyes-free interaction with free-hand gestures and auditory menus. *International Journal of Human-Computer Studies*, 71(5), 627-640. doi:10.1016/j.ijhcs.2012.11.003
43. Kakez, S., Conan, V., & Bisson, P. (1997). Virtually documented environments: A new interface paradigm for task-oriented access to information. *Computer Graphics Forum*, 16(3), C319-C327.
44. Kettebekov, S., & Sharma, R. (2001). Toward natural gesture/speech control of a large display. In M. R. Little & L. Nigay (Eds.), *Engineering for Human-Computer Interaction. Lecture Notes in Computer Science, Vol. 2254* (pp. 221-234). Berlin: Springer.
45. Kim, H., Suh, K., & Lee, E. (2017). Multi-modal user interface combining eye tracking and hand gesture recognition. *Journal on Multimodal User Interfaces*, 11(3), 241-250. doi:10.1007/s12193-017-0242-2
46. Kim, M., & Lee, J. Y. (2016). Touch and hand gesture-based interactions for directly manipulating 3D virtual objects in mobile augmented reality. *Multimedia Tools and Applications*, 75(23), 16529-16550.
47. Kopper, R., Bacim, F., & Bowman, D. A. (2011). Rapid and accurate 3D selection by progressive refinement. *Proceedings of the 2011 IEEE Symposium on 3D User Interfaces* (pp. 67-74). Singapore: Virginia Polytechnic Institute and State University, Department of Computer Science and Center for Human-Computer Interaction.
48. Kopper, R., Bowman, D. A., Silva, M. G., & McMahan, R. P. (2010). A human motor behavior model for distal pointing tasks. *International Journal of Human-Computer Studies*, 68(10), 603-615. <https://doi.org/10.1016/j.ijhcs.2010.05.001>
49. Köpsel, A., Majaranta, P., Isokoski, P., & Huckauf, A. (2016). Effects of auditory, haptic and visual feedback on performing gestures by gaze or by hand. *Behaviour & Information Technology*, 35(12), 1044-1062.
50. Krapichler, C., Haubner, M., Lösch, A., Schuhmann, D., Seemann, M., & Englmeier, K.-H. (1999). Physicians in virtual environments – multimodal human – computer interaction. *Interacting with Computers*, 11(4), 427-452.
51. Lam, M. C., Arshad, H., Prabuwo, A. S., Tan, S. Y., & Kahaki, S. M. M. (2018). Interaction techniques in desktop virtual environment: The study of visual feedback and precise manipulation method. *Multimedia Tools and Applications*, 77(13), 16367-16398. doi:10.1007/s11042-017-5205-9
52. Lee, M., Billinghamurst, M., Baek, W., Green, R., & Woo, W. (2013). A usability study of multimodal input in an augmented reality environment. *Virtual Reality*, 17(4), 293-305. doi:10.1007/s10055-013-0230-0

53. Liang, Z., Xu, X., & Zhou, S. (2017). The smallest target size for a comfortable pointing in freehand space: Human pointing precision of freehand interaction. *Universal Access in the Information Society*, 16(2), 381-393.
54. Löcken, A., Hesselmann, T., Pielot, M., Henze, N., & Boll, S. (2012). User-centred process for the definition of free-hand gestures applied to controlling music playback. *Multimedia Systems*, 18(1), 15-31.
55. Lou, X., Peng, R., Hansen, P., & Li, X. A. (2018). Effects of user's hand orientation and spatial movements on free hand interactions with large displays. *International Journal of Human-Computer Interaction*, 34(6), 519-532. doi:10.1080/10447318.2017.1370811
56. Lubos, P., Bruder, G., Ariza, O., & Steinicke, F. (2016). Touching the sphere: Leveraging joint-centered kinespheres for spatial user interaction. *Proceedings of the 2016 Symposium on Spatial User Interaction* (pp. 13-22). Tokyo: Universität Hamburg.
57. MacKenzie, I. S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, 7(1), 91-139. doi:10.1207/s15327051hci0701_3
58. MacKenzie, I. S., Kauppinen, T., & Silfverberg, M. (2001). Accuracy measures for evaluating computer pointing devices. In the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 9-16). Washington, DC: York University.
59. Mäkelä, V., Heimonen, T., & Turunen, M. (2014). Magnetic cursor: Improving target selection in freehand pointing interfaces. *Proceedings of the International Symposium on Pervasive Displays* (pp. 112-117). Copenhagen: University of Tampere.
60. Matulic, F., & Vogel, D. (2018). Multiray: Multi-finger raycasting for large displays. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-13). Montreal: University of Waterloo & Preferred Networks.
61. McNeil, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
62. Morrel-Samuels, P. (1990). Clarifying the distinction between lexical and gestural commands. *International Journal of Man-Machine Studies*, 32(5), 581-590. [https://doi.org/10.1016/S0020-7373\(05\)80034-3](https://doi.org/10.1016/S0020-7373(05)80034-3)
63. Morris, M. R., Danielescu, A., Drucker, S., Fisher, D., Lee, B., Schraefel, M. C., & Wobbrock, J. O. (2014). Reducing legacy bias in gesture elicitation studies. *Interactions*, 21(3), 40-45. Doi:10.1145/2591689
64. Moscovich, T., & Hughes, J. F. (2006). Multi-finger cursor techniques. *Proceedings of Graphics Interface 2006* (pp. 1-7). Quebec: Brown University.
65. Nancel, M., Chapuis, O., Pietriga, E., Yang, X.-D., Irani, P. P., & Beaudouin-Lafon, M. (2013). High-precision pointing on large wall displays using small handheld devices. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 831-840). Paris: Univ Paris-Sud & INRIA.
66. Nancel, M., Pietriga, E., Chapuis, O., & Beaudouin-Lafon, M. (2015). Mid-air pointing on ultra-walls. *ACM Transactions on Computer-Human Interaction*, 22(5), 1-62. doi:10.1145/2766448
67. Ni, T., Bowman, D. A., North, C., & McMahan, R. P. (2011). Design and evaluation of freehand menu selection interfaces using tilt and pinch gestures. *International Journal of Human-Computer Studies*, 69(9), 551-562. <https://doi.org/10.1016/j.ijhcs.2011.05.001>
68. Norman, D. A., & Nielsen, J. (2010). Gestural interfaces: A step backward in usability. *Interactions*, 17(5), 46-49. doi:10.1145/1836216.1836228

69. Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
70. Park, Y., Kim, J., & Lee, K. (2015). Effects of auditory feedback on menu selection in hand-gesture interfaces. *IEEE Multimedia*, 22(1), 32-40. doi:10.1109/mmul.2015.5
71. Pereira, A., Wachs, J. P., Park, K., & Rempel, D. (2015). A user-developed 3-D hand gesture set for human-computer interaction. *Human Factors*, 57(4), 607-621. doi:10.1177/0018720814559307
72. Peres, S. C., Nguyen, V., Kortum, P. T., Akladios, M., Wood, S. B., & Muddimer, A. (2009). Software ergonomics: Relating subjective and objective measures. *Proceedings of 2009 CHI Extended Abstracts on Human Factors in Computing Systems* (pp. 3949-3954). Boston, MA: University of Houston-Clear Lake.
73. Pino, A., Tzemis, E., Ioannou, N., & Kouroupetroglou, G. (2013). Using kinect for 2D and 3D pointing tasks: Performance evaluation. *Lecture Notes in Computer Science, Vol. 8007* (pp. 358-367). Berlin: Springer.
74. Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X.-F., Kirbas, C., . . . Ansari, R. (2002). Multimodal human discourse: Gesture and speech. *ACM Transactions on Computer-Human Interaction*, 9(3), 171-193.
75. Quek, F. K. H. (1996). Unencumbered gestural interaction. *IEEE Multimedia*, 3(4), 36-47.
76. Rempel, D., Camilleri, M. J., & Lee, D. L. (2015). The design of hand gestures for human-computer interaction: Lessons from sign language interpreters. *International Journal of Human-Computer Studies*, 72(10-11), 728-735. doi:10.1016/j.ijhcs.2014.05.003
77. Ren, G., & O'Neill, E. O. (2012). 3D Marking menu selection with freehand gestures. *Proceedings of 2012 IEEE Symposium on 3D User Interfaces* (pp. 61-68). Costa Mesa, CA: University of Bath.
78. Ren, G., & O'Neill, E. (2013). 3D selection with freehand gesture. *Computers & Graphics*, 37(3), 101-120. <https://doi.org/10.1016/j.cag.2012.12.006>
79. Rubine, D. (1992). Combining gestures and direct manipulation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 659-660). Monterey, CA: Carnegie Mellon University.
80. Sambrooks, L., & Wilkinson, B. (2013). Comparison of gestural, touch, and mouse interaction with Fitts' law. *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration* (pp. 119-122). Adelaide: Flinders University.
81. Shim, J., Yang, Y., Kang, N., Seo, J., & Han, T.-D. (2016). Gesture-based interactive augmented reality content authoring system using HMD. *Virtual Reality*, 20(1), 57-69. doi:10.1007/s10055-016-0282-z
82. Son, M., Jung, J., & Park, W. (2017). Evaluating the utility of two gestural discomfort evaluation methods. *PLoS ONE*, 12(4), e0176123. doi:10.1371/journal.pone.0176123
83. Tse, E., Hancock, M., & Greenberg, S. (2007). Speech-filtered bubble ray: Improving target acquisition on display walls. *Proceedings of the 9th International Conference on Multimodal Interfaces* (pp. 307-314). Nagoya: University of Calgary.
84. Turk, M. (2001). Perceptual user interfaces. In R. A. Earnshaw, R. A. Guedj, A. v. Dam, & J. A. Vince (Eds.), *Frontiers of human-centered computing, online communities and virtual environments* (pp. 39-51). London: Springer London.
85. Turk, M. (2014). *Multimodal interaction: A review*. *Pattern Recognition Letters*, 36, 189-195. <https://doi.org/10.1016/j.patrec.2013.07.003>
86. Vatavu, R.-D. (2012). User-defined gestures for free-hand TV control. *Proceedings of the 10th European*

- Conference on Interactive TV and Video* (pp. 45-48). Berlin: University Stefan cel Mare of Suceava.
87. Vatavu, R.-D. (2013). A comparative study of user-defined handheld vs. freehand gestures for home entertainment environments. *Journal of Ambient Intelligence and Smart Environments*, 5(2), 187-211.
88. Vogel, D., & Balakrishnan, R. (2005). Distant freehand pointing and clicking on very large, high resolution displays. *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology* (pp. 33-42). Seattle, WA: University of Toronto.
89. Wachsmuth, I. (1999). *Communicative rhythm in gesture and speech* (pp. 277-289). Berlin: Springer.
90. Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, 57, 209-232. <https://doi.org/10.1016/j.specom.2013.09.008>
91. Wexelblat, A. (1995). An approach to natural gesture in virtual environments. *ACM Transactions on Computer-Human Interaction*, 2(3), 179-200. doi:10.1145/210079.210080
92. Wexelblat, A. (1998). Research challenges in gesture: Open issues and unsolved problems. In I. Wachsmuth, & M. Fröhlich (Eds.), *Gesture and Sign Language in Human-Computer Interaction. Lecture Notes in Computer Science, Vol. 1371* (pp. 1-11). Berlin: Springer.
93. Wobbrock, J. O., Aung, H. H., Rothrock, B., & Myers, B. A. (2005). Maximizing the guessability of symbolic input. *Proceedings of 2005 CHI Extended Abstracts on Human Factors in Computing Systems* (pp. 1869-1872). Portland, OR: Carnegie Mellon University.
94. Wongphati, M., Osawa, H., & Imai, M. (2015). User-defined gestures for controlling primitive motions of an end effector. *Advanced Robotics*, 29(4), 225-238. doi:10.1080/01691864.2014.978371
95. Zaiți, I.-A., Pentiu, Ș.-G., & Vatavu, R.-D. (2015). On free-hand TV control: Experimental results on user-elicited gestures with Leap Motion. *Personal and Ubiquitous Computing*, 19(5), 821-838.
96. 呂俊宏、徐俊斌、Latief, S.、羅文伶、潘文福、劉從義 (2015)。體感手勢游標融入配對遊戲學習之使用者接受度評估研究。 *教育傳播與科技研究*, 110, 63-78。
- Lu, C.-H., Hsu, C.-P., Latief, S., Lo, W.-L., Pan, W.-F., & Liu, T.-Y. (2015). The assessment study of users' acceptance on kinect magic cursor integrated into match game learning. *Research of Educational Communication and Technology*, 110, 63-78. [in Chinese, semantic translation]
97. 林信志、朱菀鈴、朱昱融 (2016)。手勢互動內容創作系統之開發與評估。 *數位學習科技期刊*, 8 (2), 1-16。
- Lin, H.-C., Chu, W.-L., & Chu Y.-J. (2016). Development and evaluation of gesture-based interactive content authoring system. *International Journal on Digital Learning Technology*, 8(2), 1-16. [in Chinese, semantic translation]

The Development of Academic Research on Gesture-based Interface Design- A Review and Synthesis

Chien-Hsiung Chen* Jianli Wang**

* Department of Design, National Taiwan University of Science and Technology
cchen@mail.ntust.edu.tw

** Department of Design, National Taiwan University of Science and Technology
15067431606@163.com

Abstract

We are in the transition from a traditional WIMP interface to a natural interaction interface. Gesture as one of the most important input modes in the natural interaction interface has attracted the attention of all circles. However, in the field of design research, gesture-based interface research is still in its infancy. Therefore, this study aims to assess the international and Taiwanese academic research development in gesture-based interface via Web of Science and Index to Taiwan Periodical Literature System of National Central Library. This study employs the meta-analysis methodology to collect and analyze academic research articles of gesture interaction and reviews the current research development, academic subject areas, professional journals, and important topics. Finally, this study proposes five emphases of academic research on gesture interaction from the perspective of design research: 1. How to point and select targets by gesture; 2. How to select appropriate gestures to generate commands; 3. Differences between gesture-based interface and traditional UI interface; 4. Usability research on the gesture-based interface; 5. How gesture input can be applied to the multimodal interface. Hopefully, the results of this study can be treated as a reference for practical design and academic research.

Keywords: Gesture Interaction, Interface Design, Usability, Multimodal Interface, HCI.