

使用深度點雲技術之手勢手形比對門禁系統

A Hand Gesture and Shape Matching Access Control System Using Depth Point Cloud Technology

張厥煒

國立臺北科技大學 資訊工程系

忠孝東路3段1號

台北市106大安區

cwchang@ntut.edu.tw

林鈺皓

國立臺北科技大學 資訊工程研究所

忠孝東路3段1號

台北市106大安區

peter1205@gmail.com

摘要

本論文運用3D「手勢與手形」結合「鎖」的概念，設計並實作出一種新的以互動方式解鎖的「手勢鎖」。藉由每個人手勢手形特徵及手部習慣動作的不同，所產生出來的個別獨特生物特徵，做為解鎖時的依據。

此「手勢鎖」系統透過RGBD感測器，追蹤其手部區域的影像資訊，利用手部遮罩圖疊合點雲影像，取得完整的手部點雲資訊。再利用點雲技術的3D物件辨識演算法，進行手勢點雲的特徵抽取與比對。使用者在解鎖時，依序比劃出先前所設定儲存的手勢手形鑰匙，透過手勢點雲的3D形狀與順序排列，取得其對應的密碼序列，系統即可成功解鎖。經實驗及實測結果，此方法有其安全性、精確性與可行性。

關鍵詞: 深度感測器、手勢辨識、點雲技術、門禁系統、3D形狀比對。

Abstract

In this paper, we design and implement a new kind of interactive lock-open method for the access control, called “Hand Gesture Lock”, by using the idea of 3D hand gesture and hand shape matching approach. This lock-open method is based on the familiar behavior of hand shape change or the hand moving trajectory, or even the hand size of each person. And these familiar behavior of hand can be treated as a unique biometric feature to solve the lock-open problem.

This “Hand Gesture Lock” system is designed by four basic steps: (1) acquiring the RGB and depth images from an RGBD sensor, (2) tracking the moving hand position, (3) extracting the 3D hand shape depth image from the hand position mask, (4) matching the 3D hand shape

object with the target hand shapes in the pre-trained database by the point cloud ICP technology. When a user wants to open the lock, he needs to show both the correct hand gestures and he’s correct pre-set corresponding gesture code sequence. After several experiments and testing, this system shows very secure, accuracy and feasibility.

Keywords: Depth Sensor, Hand Gesture Recognition, Point Cloud Technology, Access Control System, 3D Shape Matching

一、緒論

早在公元前5000年，人們為了保護貴重的私有財產，而有了最原始的鎖與鑰匙的雛型。隨著科技發展的日新月異，鎖也從傳統的實體鑰匙鎖，逐漸演進成各種更便利的數位鎖的型式，如使用感應晶片卡，或是透過手機等行動裝置，搭配應用軟體的遙控電子鎖，都是已可見門禁系統採用開鎖的模式。實體鑰匙鎖是目前市面上最為廣泛應用的方式，卻仍有鑰匙遺失或被複製的風險。數位電子形式的門禁鎖，也存在使用時只認號碼、認卡，但不認使用者本人，易被側錄竊取或駭入破解等問題。

如何利用具備著使用者本人唯一性的「生物特徵」解鎖的方式，透過掃描人臉，或是指紋、虹膜等進行身分辨識，已是目前保全系統發展的趨勢。這種利用生物特徵當作解鎖媒介，是不需要再額外攜帶實體鑰匙或是其他設備，能夠較有效抑制有心人士的剽竊盜取，在安全性上較能達到更高的層級。

經本研究的歸納分析：由於每個人從小到大成長條件與生活習慣的差異，其個人手勢動作與手掌型狀就已具備有相當的獨特性。因此，本研究將此3D「手勢與手形」的獨特性，視為個人化的生物特徵，做為解鎖時的依據。基於以上想法，本論文嘗試使用各個人不同的「手勢與手形」結合「鎖」的概念，形成一種新的「手勢鎖」互動解鎖方式。藉由使用者個人手部獨特性，以及個人手勢習慣動作差異，並透過RGBD感測器 [1]及手部在空間和時間中的變化以及心智因子的組合比對，設計實作出一套基於手勢手形比對之門禁系統。

在表一中為本研究「手勢鎖」和市面上其他種類鎖的優缺點比較。其中「個人化」指的是會依據不同使用者進行客製的個別設定，即具備個人化身分識別的方式。

表一：手勢鎖與其他種類鎖優缺點比較

	鑰匙鎖	密碼鎖	感應晶片鎖	人臉辨識	指紋虹膜掃描	手勢鎖 (本研究提出)
個人化	無	有	有	有	有	有
嚴密性	中	低	高	中	高	高
盜拷複製	易	易	易	難	難	難
需要攜帶	是	否	是	否	否	否
生物特徵	無	無	無	有	有	有

二、相關研究

過去對於手勢追蹤與辨識的研究 [2-6]，多為使用RGB攝影機與電腦視覺上的影像追蹤與分類。不過隨著深度攝影機的普及化，使得追蹤與辨識方式更快速且準確。而深度攝影機其最大的優點，是可以利用不同的深度值，快速地分割 (Segmentation) 出背景與目標物體，省去濾除膚色或繁雜背景物體的步驟，迅速地找出手部區域位置。

關於使用深度攝影機手勢追蹤與分割的方式，Lee [7] 等人選取與攝影機距離50至80公分的深度影像內容，假設為手部合理工作的感興趣範圍 (Region of Interest, ROI)，因而可簡便的得到手部區域影像，如此能降低系統搜尋手部的執行時間，並保留手部影像的品質。Yang [8] 等人將深度影像中來回擺盪的物體視為手部候選區域，藉由手部的揮動找出手部可能存在的深度位置，保留所有具有相似深度值的影像，同時對影像進行加權計算，以找到手掌確切位置，接著再使用CAMSHIFT演算法持續追蹤手部。由於CAMSHIFT可自動改變追蹤區域之大小，故此方法可克服手部在空間中前後移動時，影像的縮放問題。Tang [9] 運用色彩平衡，以及Jones [10] 等人設計的膚色偵測器，對整張RGB影像像素進行膚色及非膚色之標記。同時，運用深度資訊以及既有的人體骨架追蹤方法，取得平面RGB影像中手部所在的位置，用以濾除其他雖為膚色但不屬於手部區域位置，並用深度資訊縮小感興趣範圍至手掌部分。

本研究為了能對3D手形做比較，同時也應用點雲 [11] (Point Cloud) 技術在靜態手形比對上。點雲是指一個多維度空間中點的集合，通常用來表示三維資訊的資料結構。在一個三維點雲中，這些點在幾何座標中以 (X, Y, Z) 座標來定義，且通常用來描述一個物體的外部表面，正好可以搭配深度感測器所獲得的深度資訊使用。關於3D形狀的比對，過去的研究多以 ICP (Iterative Closest Point) 演算法為主。ICP 演算法為 Besl 和 McKay [12] 在1992年所提出，是一個用來最小化兩個點雲差異的演算法，使不同點雲點和點之間的距離最小化、使其重疊部分更加緊密。在 ICP 演算法中，來源與目標點雲中各有一群可以相互對應、疊合的點，稱其為對應點。ICP 會將該對應點假設為範圍內與目標點雲中最接近的點，透過最佳化方式，找出來源與目標點雲中，各個對應點之距離和為最小的平移與旋轉矩陣，再經多次迭代後使對應點距離和收斂，即完成比對。

關於使用深度資訊對手勢比對的問題，Apostol [13] 等人假設當使用者面對如微軟 Kinect [1] 的深度感測器時，手部區域是距離 Kinect 最近的物體，用以擷取出手部區域點雲資訊。接著先透過 RANSAC (Random Sample Consensus) 演算法，抽出手勢點雲中最重要點，將離群的點先做初步的濾除，以降低計算上的維度。透過 RANSAC演算法簡化後，在完整手勢點雲中數千個點中，依然有許多是屬於冗餘的點，所以必須再利用KPCA (Kernel Principal Component Analysis) 演算法，再降一次維度。最後再利用支持向量機 (Support Vector Machine, SVM) 分類器，對擷取到的點雲資訊進行訓練以及測試。

綜合以上相關研究與文獻探討，可將手勢辨識的問題，區分為手勢追蹤與手勢比對兩個部分討論。在手勢追蹤的部分，利用深度感測器可以有效解決影像是受到光影亮暗或是背景雜亂程度等影響；而在手勢比對的部分，利用點雲模型以及點雲比對演算法，也可以有效地協助本研究建立起來源點雲與目標點雲的比對相似度方法。

三、設計構想與系統架構

本研究嘗試將手形、位置、深度、時間以及心智這五種元素因子所組合建構的手勢鎖，應用於居家安全的門禁系統上，由於五種元素因子組合後變化種類繁多，本論文將主要以困難度較高的3D手形的運作部分，進行較細部的說明探討。

1. 手勢鎖設計理念

本論文的設計理念，是希望讓使用者透過自然、方便，且具高度安全嚴密的方式，達到輕鬆便利解鎖的目的。由於每個人從小到大成長條件與生活習慣的不同，手部生理的手掌及手指大、小、長、短都不相同，再加上手部關節姿態變化，具有極高自由度 (Degree of Freedom)，手部習慣動作也不同的特性，造成每個人手部生物特徵的不一樣，所產生出來的個人化手勢及手形，就具有獨特性以及唯一性。

基於前述理念，圖一為本研究「手勢鎖」系統操作概念示意圖，操作時使用者只須位於門禁系統前，操作先前所設定訓練好的個人化手勢鑰匙，即可進行解鎖的流程，無須攜帶任何實體鑰匙或是特殊設備。以下將針對此手勢鎖的特性以及手勢鎖的種類，做細部的介紹。



圖一：手勢鎖操作概念示意

本手勢鎖系統設計的運作方式，是透過一組具有彩色 (RGB) 及深度 (D) 資訊的 RGBD 感測器擷取影像串流後，再透過取得人體骨架各部位的位置，追蹤出手掌心的座標，而獲得其手部區域的 RGBD 手形影像資訊。因此，本系統也藉此手部位置及 3D 手形資訊，建立出具有多種可組合成鑰匙密碼的元素因子，分別為：(a) 3D手形、(b) 位置 (x, y)、(c) 深度 (z)、(d) 時間 (速度)，以及(e) 心智 (記憶)。

3D手形方面，即為具有個人化生物特徵的靜態手勢鑰匙。使用者在建立靜態手勢鑰匙時，應選擇自己平常習慣的手勢當作鑰匙，如果使用非平常習慣、過於複雜的手勢，很容易造成本身難以再次比出相同姿勢，導致使用者也難以準確地去解鎖的困境。

位置方面，是指使用者手勢在畫面上的位置。本系統在解鎖畫面上，會將畫面以 4×4 共16個小白點，將手部所在解鎖畫面的平面位置，切分成 3×3 的九宮格，用以提示使用

者(避免直接明顯標示出來)現在手部所在畫面上的位置,讓使用者知道目前手部位置以方便解鎖。

深度方面,為使用者手勢距深度離感應器的遠近。本系統將手勢位置與深度感應器間的解鎖空間,區分成三段範圍,而解鎖畫面的左右兩側會有提示燈號(如位於第一區時,由使用者選定的該組提示燈號中,對應的燈號會亮起),提示告知使用者現在手勢所在的深度是位於哪一個層級。再搭配上手勢所在解鎖畫面的平面位置,形成一個 3×3×3 立方體(如同魔術方塊的概念)的解鎖空間。

解鎖過程中,若使用者於限定時間內無法完成解鎖,系統則會判斷該使用者無法成功完成解鎖狀態,系統便會回到初始的上鎖狀態,此時使用者則必須重新進行解鎖。此機制主要是要避免有心人士多次嘗試,試圖破解此手勢鎖的可能。

在解鎖空間中,手部所在畫面上的平面以及深度位置這兩個變項,也將以「提示」使用者的方式呈現,也就是說在解鎖過程中,使用者以外的人,不易了解到畫面上的「提示」幾何圖形,對使用者來說此「提示」是有特殊意義的。這樣的設計是為了預防在解鎖過程中,會被有心人士從旁窺看,並且加以學習,企圖以模仿方式將鎖破解打開。這部分也是本論文所加入「心智」條件的部分。

由於本論文是將手勢鎖應用於居家安全的門禁系統上,因此需要考量到手勢鎖對於各個年齡層用戶使用上的方便性。由於本系統具有手勢、位置、深度、時間以及心智這五種元素因子,將之選項多樣組合後,可以設計出難易度由低到極為嚴密的多種門禁鎖,如:數字按鍵鎖、數字滑動鎖、手部軌跡鎖、變動手勢鎖、3D手形鎖、日期顏色鎖等,並可加入假動作的安排。由於變化種類繁多,本論文將主要以困難度較高的3D手形鎖的運作部分,進行較細部的說明探討。

2. 手勢鎖系統架構

本系統的架構分為硬體層、軟體層和軟體層三個部分。硬體層部分採用如 Microsoft Kinect for Windows v2 [1] 的 RGBD 感測器,及一台計算比對結果的個人電腦;軟體層同樣採用搭配 Kinect for Windows v2 所提供的 SDK v2.0;而軟體層則包含學習模組、辨識模組和手勢鎖系統介面。

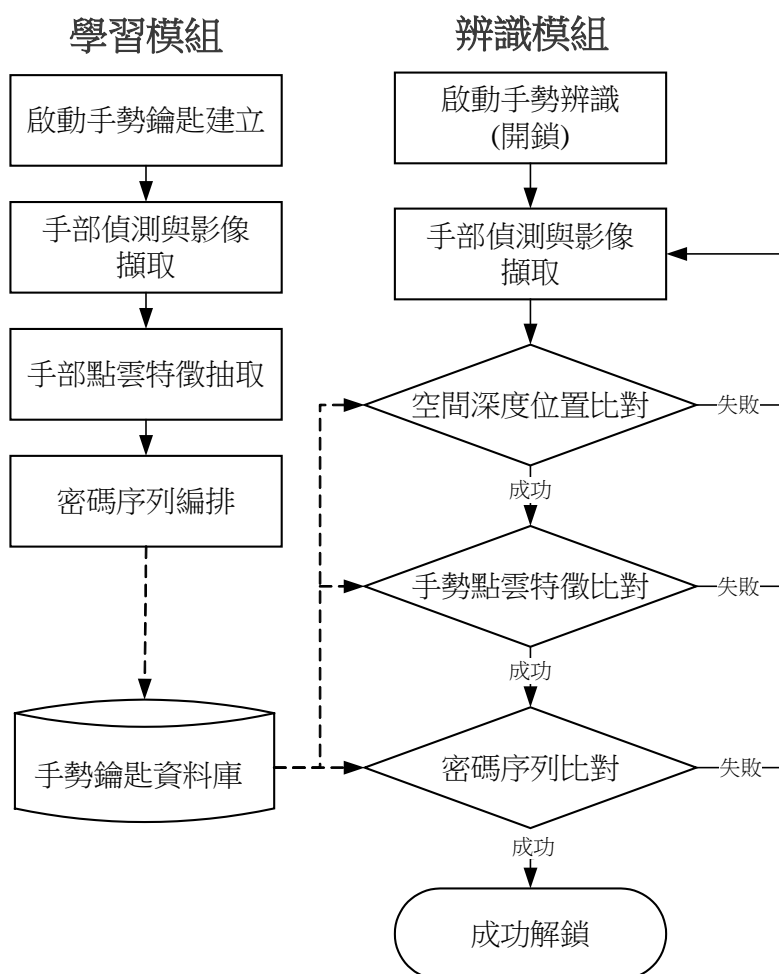
圖二為本手勢鎖系統的架構圖,架構由下到上分別是:透過感測器取得 RGBD 影像串流,再透過深度影像取得人體骨架,並追蹤其手部區域的深度影像資訊。接著區分為學習(上鎖)模組和辨識(開鎖)模組兩個部分,學習模組包含手勢特徵抽取、記錄手勢與設定密碼以及個人化手勢鑰匙建立等部分。而辨識模組則為解鎖機制,系統會進行手勢鑰匙的辨識、比對,判別該手勢是否達到解鎖的門檻,才可成功完成解鎖。最後,手勢鎖系統介面提供使用者建立個人化手勢鑰匙,完成上鎖和進行解鎖等功能之呼叫。

本系統使用時需分為:建鎖(學習模組)和解鎖(辨識模組)兩個模式,學習模組為手勢鑰匙的建立,辨識模組為解鎖機制。手勢鑰匙的建立,會依據使用者選擇鎖的種類,而有不同的設定過程,以便將使用者的手部生物特徵,習慣動作,以及所擷取到的手部區域深度影像,建立獨特且唯一的手勢鑰匙。而解鎖機制,則會將輸入進來的手部區域深度影像,依據使用者選擇鎖的種類,進行動態手勢軌跡,或靜態手勢手形辨識,並與

學習設定後資料庫所存放的手勢鑰匙，進行比對。過程需要透過與手勢鑰匙的相似度計算，給予該手勢所對應的密碼，最後再與一系列密碼序列進行比對，以完成解鎖步驟。圖三為手勢鎖(以3D手形鎖為例)系統流程，左邊為學習(建鎖)模組子流程，右邊為辨識(解鎖)模組之流程。



圖二：手勢鎖系統架構



圖三：手勢鎖系統流程 - 以3D手形鎖為例

2.1 鑰匙建立流程概述(學習模組)

(1) 啟動手勢鑰匙建立

啟動感測器，快速取得感測器所拍攝到的深度畫面及人體骨架，並追蹤其手掌心關節點的座標位置。

(2) 手部影像偵測與擷取

取得手掌心關節點座標後，透過動態調整感興趣範圍 (ROI)、手部區域 RGB 影像擷取、手部影像雜訊濾除，最後再將 RGB 影像手部區域的遮罩圖與深度點雲影像進行疊合，以取得手部區域點雲影像。

(3) 手部點雲特徵抽取

採用點雲函式庫 [11] (PCL) 所提供的抽取特徵點方法 (Uniform Sampling 函式)，在輸入的手勢點雲資料上，創造三維立體像素網格 (3D Voxel Grid)，每一個三維立體像素內當前所有的點將會被逼進 (即下採樣, Down Samples) 於它們的質心。

(4) 密碼序列編排

在建立手勢鎖時，系統會將擷取完特徵後的手勢，依據其在建立時所位於的空間深度位置，分別存進三個與距離對應的資料庫中，並給予其手勢鑰匙唯一的識別碼。每個手勢鑰匙就如同一個密碼，使用者可依據喜好及嚴密性自由，決定習慣手勢的難易度 (以相似度閾值設定)，以及手勢鑰匙的數量 (有幾個手形要依序比較)。

(5) 手勢鑰匙資料庫

系統會將設定好的手形特徵值與密碼序列存入手勢鑰匙資料庫中，以完成手勢鑰匙的建立。

2.2 解鎖機制流程概述(辨識模組)

(1) 啟動手勢辨識

如學習模組，先啟動感測器，快速取得感測器所拍攝到的深度畫面及人體骨架，並追蹤其手掌心關節點的座標位置。

(2) 手部偵測與影像擷取

取得手掌心關節點座標後，透過動態調整感興趣範圍、手部區域影像擷取、手部影像雜訊濾除，最後再將手部區域的遮罩圖與點雲影像進行疊合，以取得手部區域點雲影像。

(3) 空間中深度位置比對

會先利用空間中深度位置進行初步的篩選，當使用者的手勢位於該層深度位置時，系統才會進行手形點雲特徵的比對，以減少不必要的比對運算。

(4) 手勢點雲特徵比對

使用 Correspondence Grouping [14] 3D 物件辨識方法，將目前的手勢來源點雲與位於該層深度資料庫內的目標手勢點雲進行比對，找出相似度高於閾值的手

勢點雲。而當空間中深度位置以及手勢點雲特徵比對皆相符者，才可取得其對應的手勢鑰匙識別碼。

(5) 密碼序列比對

系統會以使用者之手勢鑰匙識別碼的數量及先後順序，決定是哪位使用者。且若是使用者在特定時間內無法成功解完鎖，系統便會回到初始的上鎖狀態，此時使用者則必須重新解鎖。

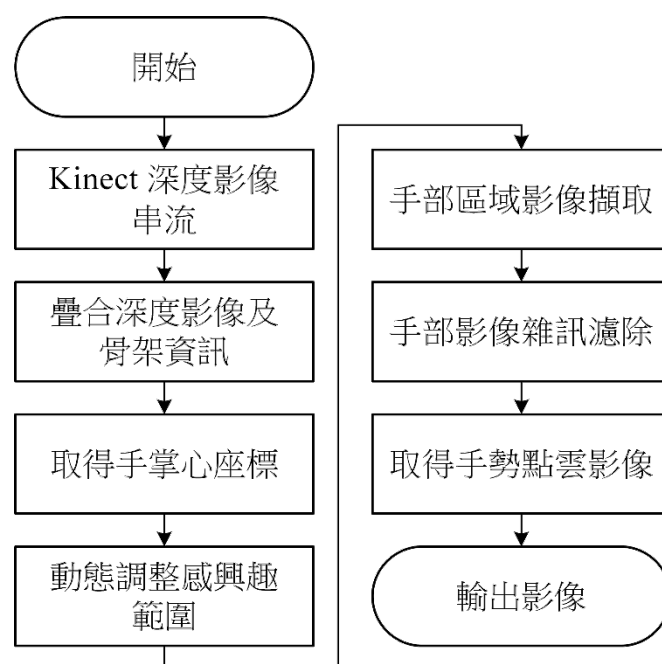
(6) 成功解鎖

若手勢鑰匙識別碼序列的順序以及數量皆符合密碼序列的設定，系統才可正確被解鎖，反之則解鎖失敗。

四、個人化手勢鑰匙建立

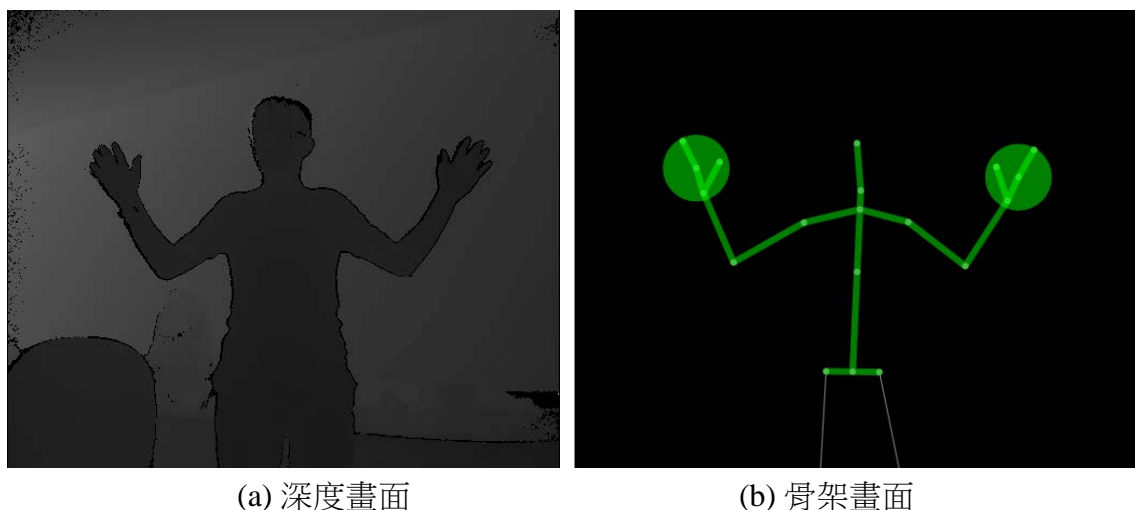
1. 手部偵測與影像擷取

手部的偵測與追蹤在手勢辨識中是相當重要的流程之一，本系統採用 Kinect v2 感測器搭配 SDK v2.0，快速取得人體骨架，並追蹤其手部區域的影像資訊。手部影像擷取大致可分為兩個部分，分別為取得手形遮罩圖，以及將手形遮罩圖與點雲影像疊合。圖四為手掌切割流程圖。



圖四：手掌切割流程

透過感測器搭配 SDK 函式庫的功能，系統可以讀取到攝影機所拍攝到的 RGB 及深度畫面，如圖五(a)所示。經由 SDK 對深度畫面的分析，便可取得骨架資訊，如圖五(b)所示。透過 SDK 的 Coordinate Mapping 雙鏡頭畫面校正 (Calibration) 程式，可以將感測器鏡頭前物體的彩色、深度畫面以及骨架資訊進行疊合。而該程式碼主要的目的，就是提供一個三維空間資訊與二維影像畫面座標相互對應及轉換的方法。

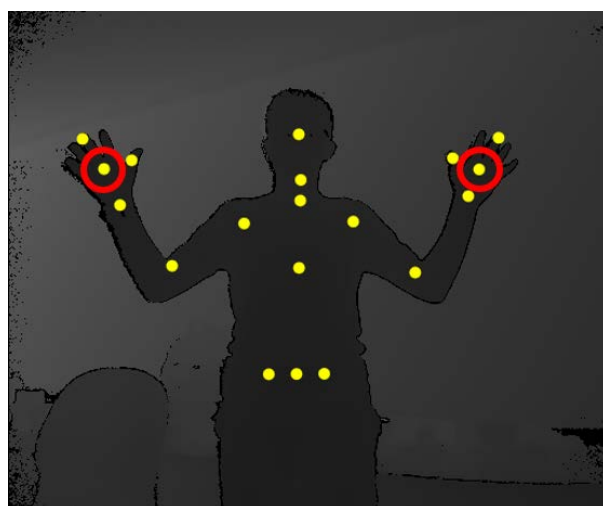


(a) 深度畫面

(b) 骨架畫面

圖五：深度及骨架畫面

為了要取得手部區域的影像資訊，需要先將深度畫面、點雲影像以及骨架資訊進行疊合，如圖六所示。而疊合後的手掌心關節點（如圖六紅圈處），即為系統所需要的手掌心座標位置。



圖六：深度畫面及骨架資訊疊合

由於手掌形狀面積較小，仍然會因為距離的遠近關係而影響其呈現的完整度。本系統在手部區域感興趣範圍 (ROI) 的動態調整上，為了使手部區域影像不會因為距離的遠近而影響其擷取後的大小，因此 ROI 範圍的大小就必須隨著與感測器的距離遠近而改變。取得骨架後，擷取手部骨架關節點的位置，透過手部骨架關節點得到當前骨架與感測器之間的距離，再透過正切函數 (Tangent) 計算當下距離、感測器視野 (Field Of View) 與手掌寬度之間的比例，計算上可分為三個步驟：(1)計算感測器視野寬度，(2)計算手掌寬度所佔視野比例，(3)設定 ROI 擷取位置。

由於僅有手部區域的資訊對系統才有意義，因此需要藉由二值化運算 (Threshold)，以取得手部區域的遮罩圖，將手部區域範圍外的部分，包含手臂以及不相關的背景皆濾














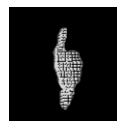
除掉，僅留下需要的手勢資訊，定義如公式 (1)。其中， T 為經二值化運算後的影像， F 為原始的手部區域輸入影像， (i, j) 為影像上的每個像素點， t_1 和 t_2 則為深度門檻值。本系統經實驗結果，在二值化門檻值範圍的設定上，採用面對感測器，手掌心前的15公分以及後的5公分，當作門檻值。

$$T(i, j) = \begin{cases} 255, & \text{if } F(x, y) \leq t_1 \text{ and } F(x, y) \geq t_2 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

由於手部區域影像經過二值化運算處理後，可能會產生邊緣不連續或是些微的雜點，這些雜訊會影響手形外型的完整度，因此需要將經二值化處理後的手部遮罩圖，再藉由型態學運算 (Morphology) 去除掉手形邊緣的雜訊，接下來才能與點雲影像進行疊合，以保留系統所需要完整的手部區域點雲資訊。而本系統所使用濾除雜點的方式為斷開(Open)，亦為先做侵蝕處理後再做膨脹處理，目的是透過侵蝕先將手部遮罩圖中的雜點濾除，使手形輪廓平滑後，再用膨脹修補過度侵蝕的手形部分。

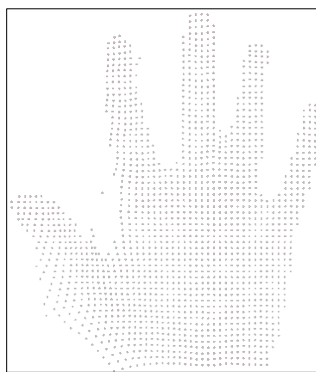
取得手部區域的遮罩圖後，即可與點雲影像進行疊合，以取得完整的手部點雲資訊，將手臂以及繁雜的背景皆濾除掉，表二為手部區域遮罩圖與點雲影像疊合的結果。對於點雲影像及骨架資訊對應的方法，由於深度及點雲影像解析度均為 512×424 像素，因此座標系統無需另行轉換。

表二：手部區域遮罩圖與點雲影像疊合的結果

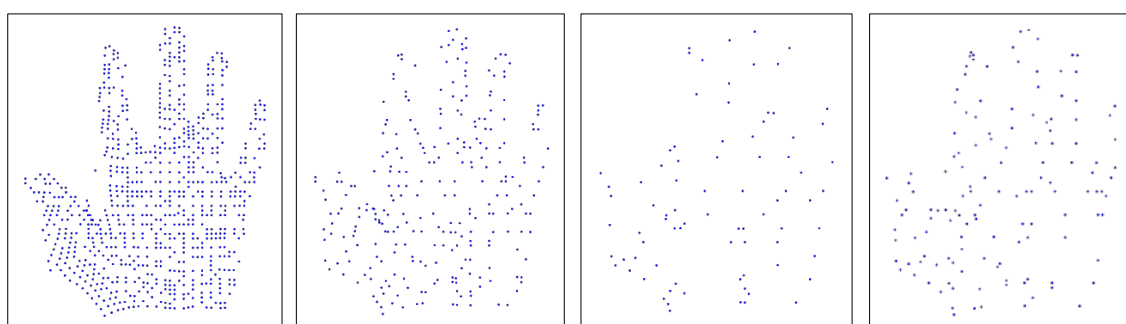
手勢編號	手勢1	手勢2	手勢3	手勢4	手勢5	手勢6	手勢7
手部區域遮罩圖							
疊合後點雲影像							

2. 手勢點雲特徵抽取

手勢點雲的特徵抽取上，本系統採用 Uniform Sampling 特徵點抽取法。Uniform Sampling 會在輸入的手勢點雲資料上，如圖七所示，創造三維立體像素網格 (3D Voxel Grid)，成為空間中極小的三維方盒 (Tiny 3D Boxes) 的集合。接著，每一個三維立體像素內當前所有的點將會被逼進 (下採樣) 於它們的質心。圖八為不同均勻取樣半徑 (Uniform Sampling Radius) 下的手勢點雲特徵點，藍色的點即代表經過 Uniform Sampling 取樣後，所抽取出來的手勢點雲特徵點。由圖八中可以發現，當均勻取樣半徑越大時，所抽取到的手勢點雲特徵點也會越少；反之，則會越多。而本系統所採用的均勻取樣半徑為10，如圖八(b)所示。最後，再將該經過特徵點抽取的手勢點雲，存進手勢鑰匙資料庫中，並給予其唯一的手勢鑰匙識別碼。



圖七：輸入的手勢點雲資料



(a) 半徑為5

(b) 半徑為10

(c) 半徑為15

(d) 半徑為20

圖八：不同均勻取樣半徑下點雲特徵點

五、解鎖機制

1. 手部偵測與影像擷取

在解鎖機制中的手部影像偵測與擷取，與個人化手勢鑰匙建立中的部分相同，透過深度影像串流取得深度影像後，將深度影像及骨架資訊進行疊合，接著取得手掌心關節點座標，透過動態調整感興趣範圍、手部區域影像擷取、手部影像雜訊濾除，最後將手部區域的遮罩圖與點雲影像進行疊合，以取得手部區域點雲影像。

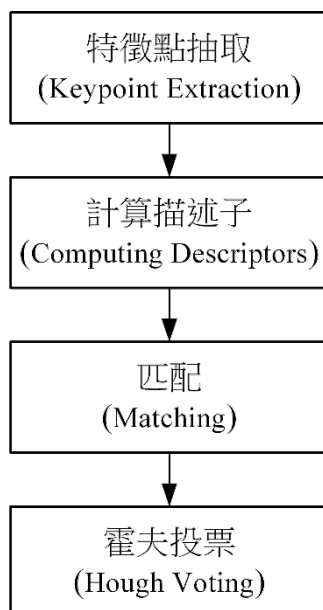
2. 空間中深度位置比對

在解鎖過程中，會先利用空間中深度位置進行初步的篩選，當使用者的手勢位於該層深度位置時，系統才會進行手勢點雲特徵的比對，以減少不必要的比對運算，而當空間中深度位置以及手勢點雲特徵比對皆相符者，系統才會給予其對應的密碼。本系統的第一層定義為 0.8~0.95公尺、第二層為 0.95~1.1公尺、而第三層為 1.1~1.25公尺。

3. 手勢點雲特徵比對

在手勢點雲特徵比對的部分，本系統使用對應分群演算法[15]，與手勢所位於該層深度資料庫內的目標手勢點雲進行比對，確認是否存在著相似度高於閾值的手勢點雲。若該目標手勢點雲存在，系統則會給予該手勢鑰匙其所代表的對應密碼。圖九為對應分群演算法流程，可分為以下幾個步驟，分別為：(1)特徵點抽取 (Keypoint Extraction)、(2)

計算描述子 (Computing Descriptors)、(3) 匹配 (Matching)、以及 (4) 霍夫投票 (Hough Voting)，進行手勢點雲在特徵比對時，所使用的方法與相似度的判定。



圖九：對應分群演算法流程

(1) 特徵點抽取

在對點雲下採樣以抽取特徵點之前，需要先計算手勢點雲與目標手勢點雲中，每個點的法線 (Normals)，本系統使用點雲函式庫中的 Normal Estimation OMP 函式，而每個點是採用 10 近鄰 (10 Nearest Neighbors) 法。接著針對每個點雲下採樣 (Down Samples)，以找到較少數量的特徵點，為了使稍後在做特徵點匹配及決定點對點對應 (Point-to-Point Correspondences) 時，能與三維描述子 (3D Descriptor) 產生關聯。而一個好的特徵點，應該要具有以下兩個特性，分別是重複性以及區分性。重複性是指，無論經過幾次抽取，即使是經過翻轉(角度不同)後，都可以抽取到相同的特徵點；區分性是指，被抽取到的特徵點，必須具有高度的獨特性及描述性，很容易去描述或是匹配它們。

(2) 計算描述子

從手勢點雲以及目標手勢點雲中，抽取並計算出每一個特徵點緊密且具有代表性的三維局部描述子。計算描述子主要在於，使三維描述子，與手勢點雲、目標手勢點雲的特徵點之間產生關聯，本系統使用點雲函式庫中的 SHOT Estimation OMP 函式，計算 SHOT 描述子。

(3) 匹配

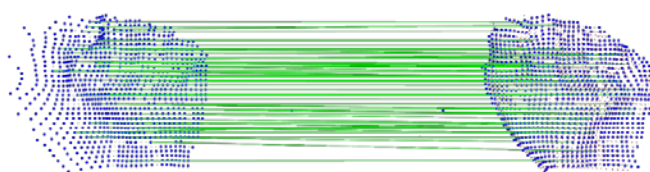
為了要決定在手勢點雲與目標手勢點雲的描述子之間，點對點的對應，本系統使用 Kd Tree FLANN (Fast Library for Approximate Nearest Neighbors) 方法。Kd Tree 是一個空間分隔的資料結構，用於將一系列 K 維的點儲存於樹的結構，可以有

效地進行範圍與近鄰搜尋。Kd Tree FLANN 會利用歐式距離 (Euclidean Distance)，有效地對每一個關聯到手勢點雲特徵點的描述子之間，找到最相近的描述子對應，並且將這個配對加入對應的陣列 (Correspondences Vector) 中。經過匹配運算後，即可得到手勢點雲間相似度 (Similarity) 判定的公式，如公式 (2)， $key_keypoints_size$ 為資料庫內目標手勢點雲的全部特徵點總數， $gesture_key_corrs_size$ 則為使用者當前手勢點雲與資料庫內目標手勢點雲一致 (有對應到) 的最佳匹配特徵點總數， $Similarity$ 則為經相似度計算後的總分。而成功解鎖定義為：(1) 相似度分數達一個閾值 S 以上，(2) 分數為資料庫中手勢鑰匙得分最高者。兩個條件皆滿足，系統才會判定為成功解鎖。

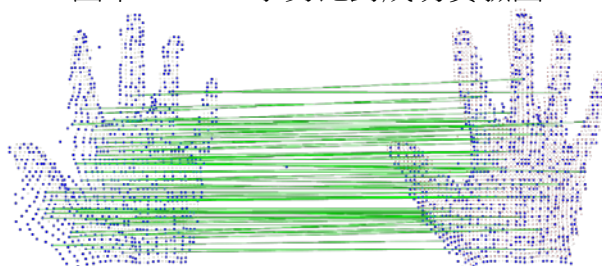
$$Similarity = \frac{gesture_key_corrs_size}{key_keypoints_size} \times 100\% \quad (2)$$

(4) 霍夫投票

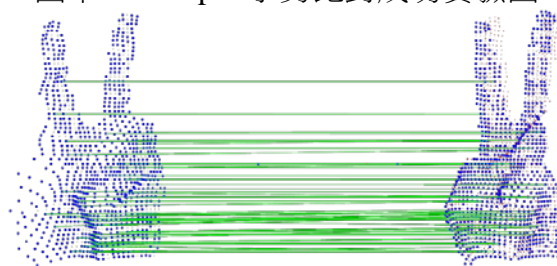
將匹配過程所找到的對應點實際地分群。使用基於霍夫投票 (Hough Voting) 的 Hough 3D Grouping 演算法，將被當作參數的點雲中的每一個特徵點，去關聯到 Local Reference Frame (LRF)。在呼叫分群演算法 (Clustering Algorithm) 之前，本系統是使用 BOARD Local Reference Frame Estimation 來計算 LRFs 的集合。對應幾何一致的特徵點，去除匹配錯誤的特徵點，若有夠多的票數累積，則代表找到該物體。此方法可以抗旋轉及平移。圖十、圖十一以及圖十二為三個靜態手勢點雲特徵比對成功的實驗結果。



圖十：Stone 手勢比對成功實驗圖



圖十一：Paper 手勢比對成功實驗圖



圖十二：Peace 手勢比對成功實驗圖

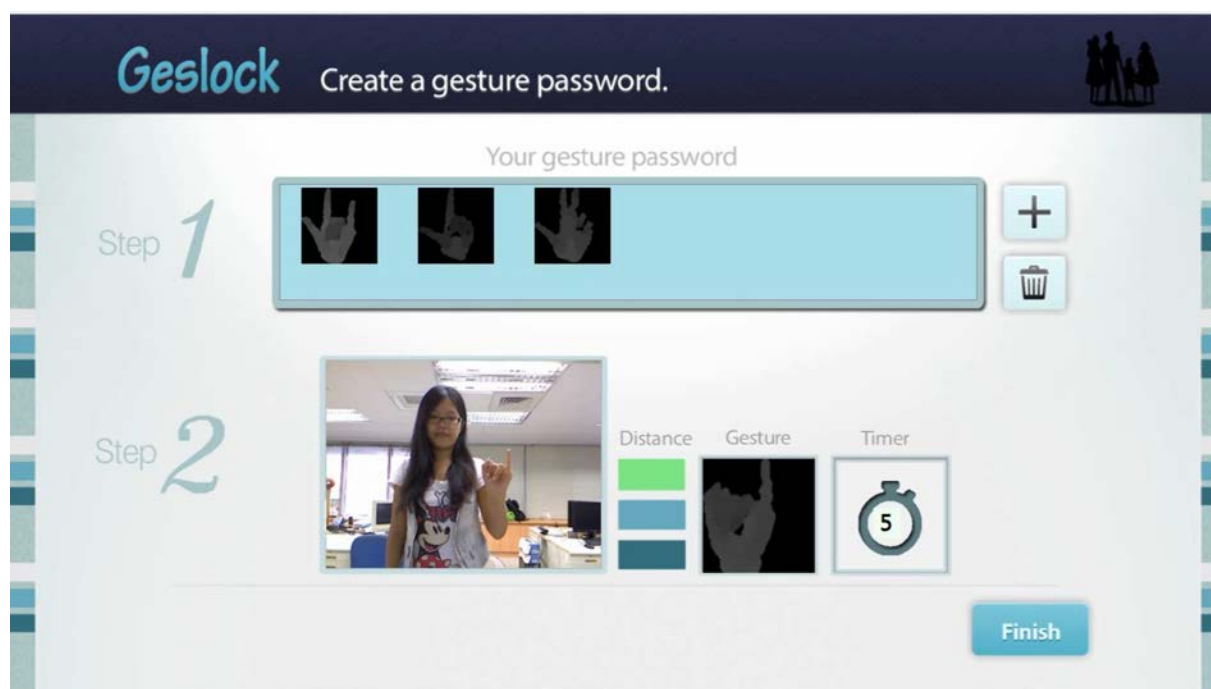
六、系統運作與實驗結果

1. 手勢鎖使用者介面

本系統進入手勢鎖建鎖模式後，會先讓使用者進行個人基本資料設定。建立各類手勢鎖時，需要先輸入姓名、選擇欲建立手勢鎖之種類、解鎖難易度，以及拍照。圖十三為「3D手形鎖」建鎖畫面，畫面中下方的「Distance」為個人顏色提示，主要是提醒使用者現在手部所在的位置，是位於哪一個層級。而在解鎖時，該手勢也要操作於相同的層級，手勢鎖系統才能成功被解開。中間的「Gesture」為使用者目前擺放至鏡頭前的手勢，可以讓使用者在建立鑰匙時，清楚自己所建立的手勢。右方的「Timer」為倒數計時器，每一道手勢鎖，系統會有10秒鐘的時間，可以將欲建立的手勢鑰匙擺放至定位，倒數完畢後則會自動儲存該手勢特徵，並給予其唯一的對應密碼與手勢鑰匙資料庫中。



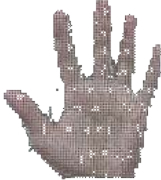
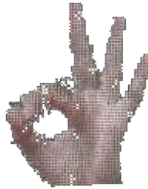

2. 測試實驗結果

在先前手勢點雲特徵比對中提到，在辨識方法上是採用點雲的對應分群3D物件辨識演算法，判斷是否存在著相似度高於閾值的3D手勢點雲。本研究在手勢點雲間特徵比對的相似度，主要是設計公式(2)進行相似度計算與判定。表三與表四分別為使用者當前(欲解鎖)手勢點雲樣本，以及資料庫內目標手勢點雲(建鎖鑰匙)樣本。樣本手勢點雲分別為特定使用者的Peace、Stone、Paper、OK及Yo手勢。








圖十三：「3D手形鎖」建鎖畫面

表三：使用者當前(欲解鎖)手勢點雲樣本

Gesture				
Peace	Stone	Paper	OK	Yo
				

表四：資料庫內特定使用者的目標手勢點雲(建鎖鑰匙)樣本

Key				
Peace	Stone	Paper	OK	Yo
				

手勢點雲間的相似度判定實驗結果如表五所示，表格最左方行為使用者當前（欲解鎖）手勢點雲，而表格最上方列為資料庫內目標手勢點雲（鑰匙），當比對到資料庫內的目標手勢點雲時，經由公式 (2) 所得到的相似度都有90%以上。在表五中可以發現，當欲解鎖手勢點雲為 OK 及 Yo 手勢時，所對應到資料庫內目標手勢點雲 Paper，其相似度有 75.75 % 及 77.40 %；而當欲解鎖手勢點雲為 Paper 時，所對應到資料庫內目標手勢點雲 OK 及 Yo，其相似度卻只有 66.27 % 及 59.50 %，其原因為每個手勢點雲所抽取出來的特徵點數目不同所致，這也是為什麼在成功解鎖的條件判定上，除了滿足條件 (1) 相似度分數達閾值 S (本系統中為85分) 以上，還須加上條件 (2) 分數為資料庫中手勢鑰匙得分最高者的緣故。當這兩個條件皆滿足時，系統才會判定為比對成功，成功的被解鎖。

在測試的過程中發現：由於使用者在操作上，容易因當初設定的手勢鑰匙複雜度過高，而導致解鎖時間過長或是遇到阻礙。因此，在建立手勢鑰匙時，應當選擇自己所習慣，且具有生物特徵的手勢來當作鑰匙，以達到「複雜」卻「好記」的特點。因此，本系統也必須思考該如何能夠容忍更多的使用者，並且準確地過濾掉不同使用者的手勢鑰匙，以加快系統比對速度及提升系統的辨識度。

表五：手勢點雲間的相似度

Key Gesture	Peace	Stone	Paper	OK	Yo
Peace	92.50 %	42.63 %	34.53 %	24.57 %	31.66 %
Stone	22.67 %	98.41 %	26.03 %	23.99 %	26.65 %
Paper	22.67 %	73.60 %	97.30 %	66.27 %	59.50 %
OK	25.65 %	52.87 %	75.75 %	92.84 %	43.73 %
Yo	23.55 %	55.50 %	77.40 %	45.54 %	96.27 %

七. 結論

本研究將 3D 動、靜態手勢比對理論，應用於門禁系統的機制上。此系統在使用上是不需要攜帶實體鑰匙或者是額外設備，真正的鑰匙，是放於使用者的身體或心中，也只有自己能夠透過「手勢特徵」及「心智思考」，進行唯一解鎖，成為一種利用生物特徵，結合心智的鎖。同時，可以讓使用者免去攜帶或是避免遺失的困擾，也因手勢是具有個人的生物特徵以及習慣動作，相當不易被口語描述或是影像傳達，也在安全性上更添了一層保障。

目前本系統在手勢鑰匙的建立以及解鎖機制上，解鎖時均是以單人單手的方式操作，將來也可以思考如何增加至多人單手，甚至是多人多手（協同開鎖）的方式操作解鎖，讓手勢鎖的變化夠更加多元化，也能使系統的實用性甚至是安全性上獲得提升。

參考文獻

- [1] Microsoft, "Kinect for Windows", <http://www.kinectforwindows.org>, 2014.
- [2] Gutzeit, E., Vahl, M., Zhou, Z. and Lukas, U. V., "Skin Cluster Tracking and Verification for Hand Gesture Recognition," *International Symposium on Image and Signal Processing and Analysis (ISPA)*, 2011: pp. 241-246.
- [3] Zhu, X., Yang, J. and Waibel, A., "Segmenting Hands of Arbitrary Color," *Proceedings of the 4th IEEE Intel. Conference on Automatic Face and Gesture Recognition*, Grenoble, 2000: pp. 446-453.
- [4] Zhu, Y., Xu, G. and Kriegman, D. J., "A Real-Time Approach to the Spotting, Representation, and Recognition of Hand Gestures for Human-Computer Interaction," *Computer Vision and Image Understanding*, (85:3), 2002: pp. 189-208.
- [5] Wang, R. Y. and Popovic, J., "Real-Time Hand-Tracking with a Color Glove," *ACM Transactions on Graphics*, (28:3), 2009: pp. 505-513.

- [6] Oka, K., Sato, Y. and Koike, H., "Real-Time Tracking of Multiple Fingertips and Gesture Recognition for Augmented Desk Interface Systems," *IEEE 5th International Conference on Automatic Face and Gesture Recognition*, 2002: pp. 429-434.
- [7] Lee, H. L., Hsu, C. L., Chen, C. C., Taur, J. S. and Tao, C. W., "Real-Time Hand Gesture Controlled Mouse Using Kinect," *Proceedings of the 25th CVGIP*, 2012: pp. 189-197.
- [8] Yang, C., Jang, Y., J. Beh, Han, D. and Ko, H., "Gesture Recognition Using Hand Tracking for Contactless Controller Application," *Proceedings of the 2nd IEEE Intl. Conference on Consumer Electronics (ICCE)*, 2012: pp. 297-298.
- [9] Tang, M., "Recognizing Hand Gestures with Microsoft's Kinect," online available at: http://www.stanford.edu/class/ee368/Project_11/Reports/Tang_Hand_Gesture_Recognition.pdf, 2011.
- [10] Jones, M. J. and Regh, J. M., "Statistical Color Models with Application to Skin Detection," *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1999: pp. 638-646.
- [11] PCL - Point Cloud Library (PCL), <http://pointclouds.org/about/>.
- [12] Besl, P. J. and McKay, N. D., "A Method for Registration of 3-D Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (14,2), 1992: pp. 239-256.
- [13] Apostol, B., Mihalache, C. R., and Manta, V., "Using Spin Images for Hand Gesture Recognition in 3D Point Clouds," *Proceedings of the 18th International Conference on System Theory, Control and Computing(ICSTCC)*, 2014: pp. 544-549.
- [14] 3D Object Recognition based on Correspondence Grouping, http://pointclouds.org/documentation/tutorials/correspondence_grouping.php#correspondence-grouping.
- [15] Tombari, F. and Stefano, L. Di, "Object Recognition in 3D Scenes with Occlusions and Clutter by Hough Voting," *In Proc. 4th Pacific-Rim Symposium on Image and Video Technology (PSIVT 10)*, 2010: pp. 349-355.

