

第一章 緒論

1.1 研究動機

近年來家用機器人提供的協助及服務已經愈來愈多元化，可以看得出自走式機器人正步入一般民眾的家庭當中。目前家用機器人在功能上可分為娛樂型、看護型及保全型，例如：Sony 所開發的 Qrio 雙足機器人[1]，其設計的用意是在提供家庭娛樂。它會辨別主人的樣子與聲音，並且可以與人簡單對話。三菱重工設計出可協助家庭保健和看家的家用機器人“Wakamaru”[2]，他能提醒服藥和飲食的時間，並且通過電子郵件將用戶的信息通知給異地親屬，傳遞用戶狀況；在用戶出門時，識別聲音和人，並通知給用戶，或是將住宅內的情景傳送到手機及個人電腦上。日本高科技公司歐姆龍和 Tmsuk 共同開發出的電子守衛恐龍保全型機器人[3]，能負起家居的保安工作。主人不在家時，機器人會自動巡視屋子，遇有可疑人物時，便立即發出警告，並且利用第三代手機通知主人。由以上的例子看來，家用機器人愈來愈多元化、智慧化與家庭化，將與人共同生活並且協助人們完成特定的工作，或在生活當中提供一些娛樂。

為了使家用機器人能和家庭成員產生互動、提供特定的協助及使其具有靈敏的反應，家用機器人必須快速地及持續性的追蹤人臉。但是在與家用機器互動的過程當中，可能會因為環境光源的變化（例如強光打在人臉上而造成膚色過於慘白），或者是機器人影像感測器與人臉之間的某些角度（例如當機器人的身高比人低時，則機器人的頭部必須上仰，而造成天花板上之日光燈對感測器產生的影響）會使影像感測器對某些特徵難以偵測到而造成追蹤失敗。

在人臉追蹤方面，膚色被認為是一個有利的條件[4-7]，利用尋找膚色可以在環境中快速地找出人臉的位置，但此環境必須在理想的光源及背景不是很複雜的情況之下才可以成立，例如實驗室或者光源分佈均勻地方。但是實際上環境的光源、背景是多變的，對於人臉追蹤方面會造成很大的影響。所以本論文將設計一人臉偵測與追蹤的方法，使能在光源的變化之下也可以即時地且穩定地追蹤人臉。

1.2 相關研究回顧

為了讓家用機器人可以與人持續地產生互動，我們將發展一人臉追蹤演算法，將其整合在實驗室發展之家用機器人影像平台上，完成家用機器人在光源變動之下也可達到持續追蹤人臉之效果。因此本論文可分成兩個方面來看這個問題，一方面就目前各種相關的機器人與人互動系統的現況，及其應用層面來瞭解；另一方面就本系統所需使用的技術的發展現況。

1.2.1 機器人與人互動發展現況

Christian Wengert[8]等人設計出一以人為追蹤目標的人與機器人互動系統。研究如何使機器人能藉由視覺對人的行為進行觀察和作出反應。其整個系統主要分為三個部分：1)追蹤人 (Human Tracking)，利用立體視覺進行人的追蹤。其認為立體視覺不會因為色彩、背影、光線改變而跟著改變。而且他們還利用色彩索引的方法藉以保證同一時間只有一個人被追蹤。2)辨認人 (Human Identification)，他們在人與機器人的互動是採 1 對 1 的方式。因此為了能在多人的影像中區別出目前正與何人互動，他們利用色彩索引的方法為每張影像中的人設定一個 HumanID 藉以區分出有幾個人及其位置。3)手勢偵測 (Gesture Detector)，對靜態的手勢進行辨識，藉由簡單的頭部與兩手之間幾何關係的判別，進行簡單的機器人操控。

日本早稻田大學人形機器人研究所發展出一 BUGNOID 機器人[9]。他們認為影像資訊在視覺機器人的非語言溝通扮演一個重要的角色。此視覺系統分為兩個部份，一個為全像式攝影機，此部分是為了要得到全域的資訊，例如人的移動；另一部分為立體視覺，主要是用來用色彩判定及人臉辨視。系統初始化狀態是機器人是靜止不動，並且先利用全像式攝影機來偵測影像當中移動的物體。當偵測到移動的人員之後，此機器人會向此人員移動且利用立體視覺進行人臉辨視及姿態的偵測，機器人依據人員下達頭部搖頭及點頭的命令進行反應。

機器人視覺系統在人與機器人的互動扮演了重要的角色。但是有些研究是在光源均

勻的實驗室進行，雖然可以獲得不錯的實驗結果，但是他們沒有考慮到光源的因素，因此無法評估當受到光源影響時，其視覺系統是否可以正常運作。即使有考慮到光源的問題但在硬體可能有特別的設計及需求（例如立體視覺），或者利用複雜之演算法，但是可能難以達到即時性。

1.2.2 使用技術發展現況

在家用機器人與人互動系統方面，我們希望結合人臉偵測及人臉追蹤系統完成家用機器人在光源的變化下對人臉持續地進行追蹤。此系統分為兩個部分：1. 人臉偵測、2. 人臉追蹤。這兩大領域一直以來都有相當多的研究人員投入許多心力及研究成果。

Rein-Lien Hsu[4]等人提出一種彩色模型為基礎的方式，把 RGB 色彩空間轉換到 YCrCb 色彩空間，減少亮度對色彩的影響，並在 YCrCb 色彩空間內進行膚色判別及形態學運算，藉以找出臉部位置及臉部的眼睛與嘴巴的位置。

Sobottka and Pitas[10]使用膚色和形狀輪廓來做人臉的定位及臉部特徵的擷取。他們使用 HSV 色彩空間來搜尋膚色範圍，將膚色範圍選取出來之後，比對此範圍是否為橢圓形，如果成立的話就判定為人臉。在臉部特徵方面，他們認為眼睛和嘴巴區域的灰階值會比鄰近灰階值低，利用此特徵將臉部眼睛及嘴巴找出來，然後再利用類神經網路來偵測輸入的影像是否為臉部的特徵。

Craw[11]等人用正面臉部形狀比對法做為臉部的定位。他們使用 Sobel filter 得到邊緣影像，並且利用人臉模型定位出可能是人臉的區域，接著再反覆利用一些特徵做為比對的模型，例如眼睛、嘴巴等等。Craw 總共用了 40 種的模型來做為比對的模型，並且還用一些控制策略以增加準確率。此方法的優點在於容易實現，但是在處理臉部大小、姿態的改變上比較沒有效率。

表 1-1 為 M.H. Yang[12]等人歸納的單張影像偵測人臉法。文中提到人臉偵測主要有四大門派：

表 1-1 單張影像人臉偵測方法[12]

Face Detection in a Single Image	
Approach	Representative Works
Knowledge-based	Multiresolution rule-based method
Feature invariant - Facial Features - Skin Color - Multiple Features	Grouping of edges Mixture of Gaussian Integration of skin color, size and shape
Template matching - Predefined face templates - Deformable Templates	Shape template Active Shape Model(ASM)
Appearance-based method - Eigenface - Distribution-based - Neural Network - Support Vector Machine	Eigenvector decomposition and clustering Gaussian distribution and multilayer perceptron Neural Networks and arbitration schemes SVM with polynomial kernel

1) 基於知識法 (Knowledge-based method)

基於知識法通常是由研究者本身定義人臉有那些規則，並且根據些規則以找尋影像中人臉的位置，採用此方法的人臉偵測率容易受到研究者本身對人臉定義，而有不同的結果。

2) 基於特徵法 (Feature-based method)

利用此方法的研究者認為根據人臉特徵的不變性，可以用來偵測不同位置及不同角度的人臉。此方法面臨的主要問題為，由於影像感測器雜訊或者其他因素的影

響，而使人臉某些特徵不明顯，而造成人臉偵測失敗。

3)人臉模型法 (Template -based method)

利用一些人臉局部特徵模型，例如眼睛、嘴巴等等，在影像當中進行全域的搜尋。此方法在某些特定情況可以快速地找出人臉的位置，但是在克服人臉旋轉，大小的改變有其難度。

4)基於表象法 (Appearance-based method)

採用這種偵測方法的理論依據是：可以將人臉區域的圖像視為一高維度線性空間集合，於是可以透過統計或者類神經網路來判斷某一子視窗是否為人臉。利用此方法必須搜尋大量的人臉及非人臉樣本以訓練此系統，並且準確性會受到樣本的影響。

最後 M.H. Yang 為人臉偵測下了一個結論：以上各方法都有其優缺點，故結合多種檢測方法，在提升偵測的準確率有一定的助益。另外，文中有提到光源對於人臉偵測系統會帶來很大的影響，可見無論是人臉偵測或者追蹤系統，解決光源的問題是必然的。

C. C. Han[13]提出一種利用眼睛來偵測影像點中是否有人臉的方法。眼睛為一穩定的特徵，即使受到光線的影響，眼睛灰階值會小於鄰近的灰階值，利用形態學 (Morphology) 中開啟 (Open) 與閉合 (Close) 運算元的特性，找出圖形中可能為眼睛的區域，再配合臉部特徵的幾何關係與類神經網路完成最後的人臉偵測。此方法可快速偵測人臉位置，但只使用臉部特徵(眼、口、鼻...等)作為判別是否為人臉，容易受到擁有相似特徵的物體影響而產生誤判。

Paul Viola 等人[14]提出利用多個由簡單到複雜的人臉特徵，組合成一個強大的人臉分類器，並且使用增進式串聯演算法 (Boosted Cascade) 來訓練此人臉偵測系統。雖然此方法是整張影像搜尋 (Full searching)，不過透過前幾個簡單的人臉特徵可以將不是屬於人臉的子區域 (Sub-window) 給剔除，通過初步的驗證之後再由更複雜的人臉特徵進行檢驗，如此可以有效的降低運算量。但是其訓練的過程較耗時，因為所需搜尋的人臉

資料庫數量大約為 1 萬個和非人臉資料庫數量為 1 萬，並且其人臉的準確率會受資料庫的影響。

S.H. Kim and H.G. Kim[15]利用動態偵測及膚色以實現人臉偵測。首先，利用兩張連續影像相減，並且設定一閾值，當差值大於此閾值的話則為 1，反之則為 0。得到差值影像之後，再利用膚色特徵在群集（Grouping）的範圍裡面判斷此區域是否為人臉。雖然此方法結合了動態及膚色可以快速的找出人臉的位置，不過作者是在理想的光源之下偵測膚色，如果膚色受到光源的影響而造成色偏，則此方法恐怕會失敗。為了得到差值影像，所以此系統必須使用固定式攝影機對著不變的背景，因此，此方法比較適用於特定場合，例如視訊會議，當有動態的物體，例如人體，出現在影像當中即可將移動的物體和背景分離，所以當攝影機和被觀測者都在移動的話，例如家用機器人，則無法使用此方法。

人臉追蹤系統可以視為人臉偵測的一部分，因為當偵測到人臉之後只要持續對此人臉追蹤即可。Stan Brichfield[16]提出一種利用橢圓為基礎的臉部輪廓搜尋模型，並且檢測橢圓內的色彩分佈來追蹤人臉。利用此兩種模型的交互使用，能在複雜的環境中達到穩健的追蹤效果。此系統對於物體的轉動、暫時遮蔽、接近與離開皆有穩定的追蹤效果。不過橢圓偵測的缺點在於容易受到光源及背景影響，並且需要大量的運算量而難以達到即時性人臉追蹤。

Gi-jeong Jang 等人[17] 的人臉追蹤方法是利用使用條件密度傳遞法（Conditional Density Propagation）來預測下一次人臉出現的位置。他們假設在追蹤的過程當中人臉的密度會出現最大值，所以選擇先前的人臉平均密度做為人臉追蹤的預測評估器。其方法包含兩個部份：移動預測及更新。在移動預測方面，他們將移動模型表示為 $x_k = Ax_{k-1} + Bw_{k-1}$ ，A 為動態特徵點，B 為比例因子， w_k 為高斯雜訊向量。利用此模型來預測下一次人臉的位置。更新方面主要是依據膚色的色彩分佈，其色彩空間為正規化 RGB（Normalized RGB）。他們認為膚色分佈 r 及 g 不會因為不同的膚色而有太大的改變。

Helman[6]等人提出一個新的觀點，他們認為一般研究者都使用單一的彩色空間模型(Color Space Model, CSM)在影像當中偵測膚色，並且使用特定的色彩分佈模型(Color Distribution Model, CDM)來進行膚色分割，雖然在特定的條件之下可以獲得良好的結果，不過不能完全解決環境光源對人臉膚色所產生的影響。他們利用五種不同的色彩空間，RGB、正規化 RGB、HSI、HSV 及 YCrCb 來偵測膚色，然後各個色彩空間只使用二維的色彩分佈資訊，分別是 RG、rg、HS、YQ 及 CrCb，並且其色彩轉換過程是利用 CAMSHIFT 追蹤演算法。

Yong-Beom Lee[7]等人利用 HSI 色彩空間模型做為其偵測人臉膚色的資訊。他們認為以往的許多研究都是利用二維資訊來追蹤人臉是不足的，例如 HSI 就利用 H(Hue, 色相)及 S(Saturation, 飽和度)，YCrCb 則利用 Cr 及 Cb。他們提出雖然轉換到不同的色彩空間是為了將亮度(Intensity)及色度(Hue, Saturation)分離，以減少光源對顏色的影響，但是經由他們觀察的結果，亮度還是會對色度造成影響，所以其人臉追蹤法主要是利用 HSI 三維的色彩分佈。他們先統計了四種不同亮度之下色度的分佈，並且利用 B 型曲線(B-spline)完成在不同亮度之下色度分佈變化的情況，最後利用此色彩分佈模型和移動偵測來追蹤人臉。

Lei Zhang and Dimitris Samaras[18]利用一影像集合，將在不同光源下的影像還原到原始不受光源影響的人臉。其系統分為三個部分：計算統計模型、訓練和測試。一開始



圖 1-1 利用統計模型還原之人臉集合[18]

先替不同的人臉影像統計其臉部光源分佈模型。然後在訓練的階段時送入一個新的人臉並且此臉部是受到任意光源照射，然後他們利用此方法將受到光源影響的人臉還原成不受光源影響的人臉。如圖 1-1 所示，輸入的人臉影像受到光源的影響，但是經由此方法可以看得出來處理過的人臉都會呈現相同的灰階值分佈，如此以減少光源的影響。但是環境中光源的變化是多變化的，無法將全部可能發生的情況全部事先定義好，另一個缺點則此方法只適用於正面的人臉而已。

Carlos[19]提出利用人的瞳孔做為人臉追蹤時的特徵。他們使用主動式紅外線 (Infra-ray) 照射特定的區域，當此區域出現人的臉時，人眼的瞳孔受到紅外線的照射而產生反射，所以當攝影機照到人臉時，瞳孔相對之下會變亮許多，他們也就是利用此特性並且加上啟發式規則來偵測及追蹤人臉。此方法的缺點只能偵測正面的人臉，因為當眼睛瞳孔特徵不存在時，則此系統就無法進行人臉追蹤；另外，當背景光源很亮時此方法恐亦不適用。



1.3 問題描述

在人與機器人互動介面 (Human-robot interaction) 的應用當中，需要有一具有強健性及即時性的追蹤系統可以幫助臉部特徵的擷取，例如人臉辨識、表情及情緒的分析與唇形的辨識等等。我們可以將人臉追蹤分成以下幾個方法並且討論其可行性。

1) 動態偵測

利用兩張連續影像相減以偵測出影像當中移動的物體。此方法雖然可以快速地找出移動的物體，不過容易受到身體及其他物體的影響，而且對於人與機器人互動方面，其缺點在於攝影機必須是固定不動的。

2) 臉部模型法

利用此方法的研究者認為臉部 3D 影像的資訊不會受到光源的影響，所以每一個輸入的影像即被投影到另一個人臉模型以進行後續的處理。此方法的缺點在於有許

多模型參數必須設定，例如大小、正規化、旋轉、變形及光源之變化，不僅要很多的運算量，而且也很難將現實所有的情況考慮進去。

3) 外形搜尋法

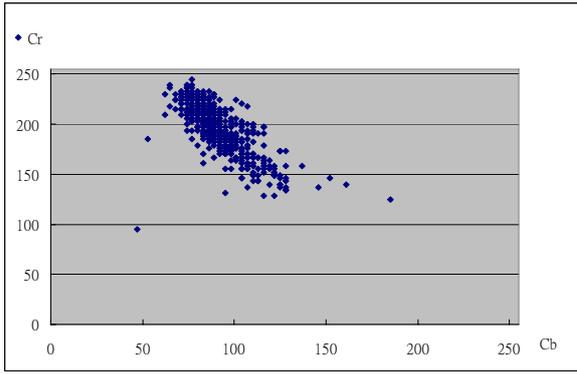
利用橢圓為基礎來搜尋影像當中符合此條件的物體，利用此方法的研究是根據頭部的形狀輪廓為橢圓形。此方法對於物體的轉動、暫時遮蔽、接近與離開皆有穩定的追蹤效果。不過橢圓偵測容易受而光源及背景影響而造成人臉追蹤失敗，而且也需要大量的運算量而難以達到即時人臉追蹤。

4) 色彩搜尋法

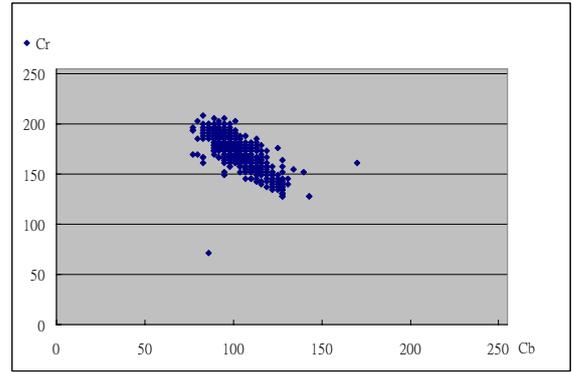
在相關的研究領域中有許多研究都認為，利用膚色可以將人臉與背景有效地分離出來，且在解決臉部旋轉及遮蔽性有一定的效果。但是色彩容易受光源的影響，而造成人臉偏暗或者偏亮，進而影響人臉追蹤的效果。雖然此派研究者提出多種色彩轉換空間，例如正規化 RGB、HSV 及 YCrCb 等等，要將亮度及色度分離開來以減少亮度對於色度的影響，不過解決的程度有限。

以上為人臉追蹤常使用的方法。另外，尚有利用紅外線偵測人臉的瞳孔等等，不過都必須在特定的情況之下才可以成立，無法在實際的情況當中實現。根據以上的討論，本論文將利用以下幾點來完成在光源變化下對人臉持續追蹤方法：

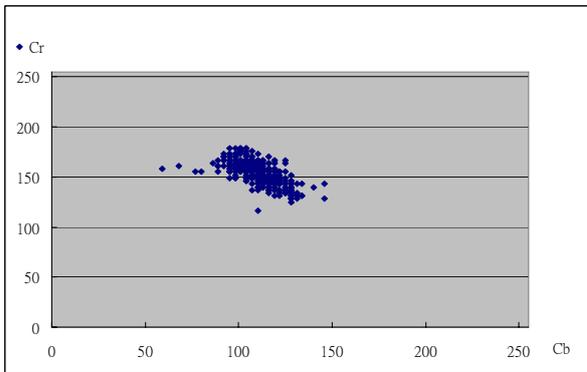
- 1) 利用連續影像的特性，將 t 時刻影像中人臉的位置及大小傳遞到下一張影像，以預測下一張影像中人臉出現的位置藉以縮小搜尋範圍。另外，利用時間性 (Temporal)，即使暫時地追蹤失敗，也可以使用最後一次人臉出現的位置及色彩分佈再進行人臉的搜尋，以增加人臉追蹤系統的穩定性。
- 2) 無論何種色彩空間必須考慮所有的資訊，如圖 1-2 所示。圖 1-2(a)、(b)、(c)、(d) 與(e)分別為臉部平均亮度 219、197、173、148 與 122 對 Cr 及 Cb 色度分佈的情況。由結果可以看得出來當亮度改變，色度的分佈會跟著改變，所以在進行膚色



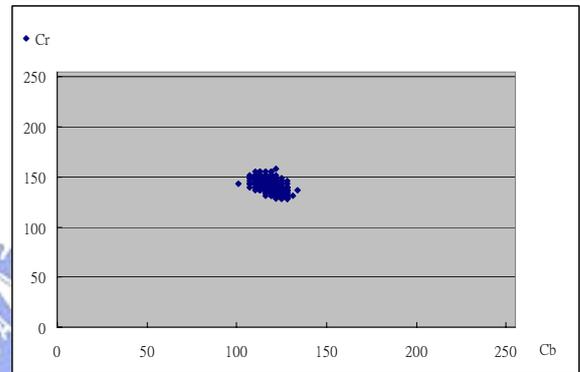
(a) 平均亮度 $Y = 219$



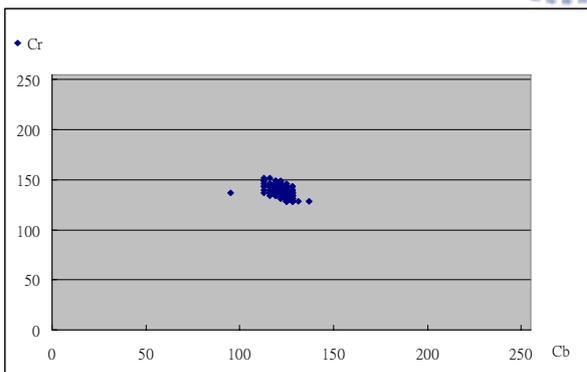
(b) 平均亮度 $Y = 197$



(c) 平均亮度 $Y = 173$



(d) 平均亮度 $Y = 148$



(e) 平均亮度 $Y = 122$

圖 1-2 亮度對色度分佈之影響

的色彩分割時應考慮不同亮度下色度分佈的情況，以準確地將膚色分割出來。本論文將利用 YCrCb 三維色彩分佈模型做為搜尋膚色及色彩分割之資訊。利用色彩分

割將影像中人臉的區域與背景分開。

3) 關閉自動白平衡 (Auto White Balancing, AWB) 及自動曝光控制 (Auto Exposure Control, AEC) 以減少對擷取影像的影響。目前在影像感測器方面，此部分最主要是針對有些特殊情況例如感測器受到強光照射或者環境太暗，感測器可以自動調整其控制參數，以維持影像一定的品質。不過由觀測的結果可以發現到，當人臉受到背光影響時，臉部的色彩分佈會有極大的變化，這對於利用膚色作為人臉追蹤的條件會產生很大的影響。另外，因為本論文是利用適應性膚色搜尋法以完成人臉追蹤，為了避免人臉追蹤系統和影像感測器之色彩分佈差異過大，造成追蹤人臉失敗，所以我們會關閉以上兩個功能以完成適應性膚色搜尋法。

4) 雖然光源會對人臉追蹤造成重大的影響，不過對於實際情況的觀測，不論是人員的移動或者光源的變化，我們可以將此現象視為連續的變化，可藉由 $t-1$ 時刻所保留下來的資訊，做為 t 時刻人臉追蹤所需之資訊。本論文將利用適應性膚色搜尋法以進行人臉追蹤以解決光源變化的問題。

總結以上討論結果，本論文將利用色彩空間模型 YCrCb 的三維色彩分佈做為人臉追蹤時的資訊。選擇此色彩空間模型是因為本論文將利用嵌入式影像平台來實現人臉追蹤，而此平台之影像感測器其輸出影像格式分為 RGB 及 YCrCb，若選擇 YCrCb 色彩空間的話，其優點在於不需要花費運算量做色彩轉換；在移動方面，藉由預測以縮小處理的範圍以免不必要的搜尋；在追蹤的穩定性方面，加入了時間性以克服短暫追蹤人臉的失敗；另外，最重要的一點，因為環境中光源對臉部的色彩分佈是一直在變動的，所以人臉追蹤系統之色彩分佈模型是隨時間一直在改變及調整的。

透過以上的方法，本論文將利用我們所提出之適應性膚色搜尋法，以達到家用機器人在不受環境光源變化，或者是機器人影像感測器與人臉之間的角度影響下，也可即時地及穩定地追蹤人臉。

1.4 章節說明

內容本論文共分六章，第一章為介紹相關研究背景及研究動機。第二章介紹人臉追蹤系統，此部分藉由適應性膚色搜尋法完成持續性的人臉追蹤。第三章介紹人臉偵測系統，包括膚色偵測及特徵判別法則。第四章介紹家用機器人之影像平台構架，並且將人臉偵測系統及人臉追蹤系統整合到家用機器人。第五章為實驗結果，使用所發展的家用機器人與人互動系統實驗的進行測試。第六章為結論與未來展望。

