

最新

CHAPTER 08 圖像識別

人工智慧概論



最新 人工智慧概論



- 8-1 什麼是圖像識別
- 8-2 圖像識別的應用案例簡介



Chapter 08 圖像識別

【教學影片】置身於優美變化的場景中
影像處理和移動通信技術的迅速發展，給人們帶來了前所未有的多媒體感覺盛宴。隨著5G 時代的開啟，我們可期待著更好的多媒體使用者體驗。請觀看「影片案例：置身於優美變化的場景中」。



- 置身於優美變化的場景中之示例
(圖片來源：123RF 圖庫)



問題思考

你使用的手機或電腦的首頁，是否放置優美的場景圖片？請分享一下對於曾觀看過的、具有優美的場景影片感覺？



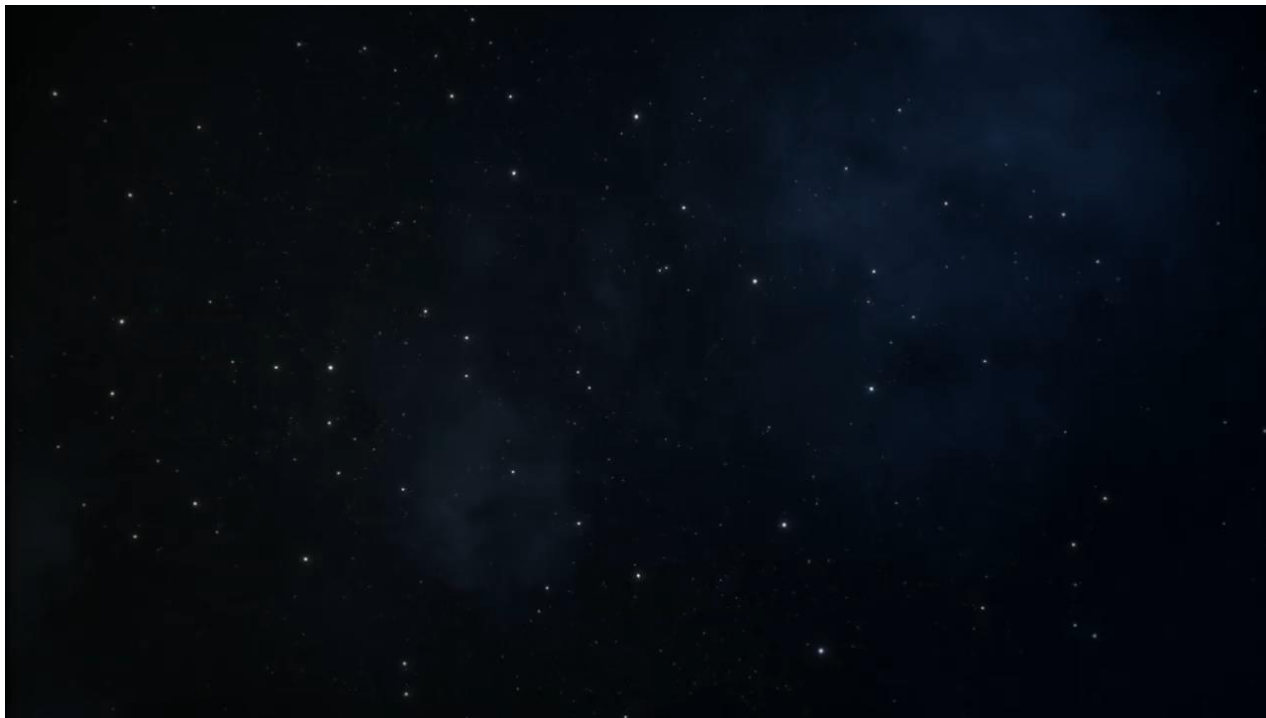
只要一分鐘，人臉識別瞭解一下！

P. 191, 01:37 20180722



How we teach computers to understand pictures

P. 191, 18:02 20150323





AI 深度學習 | AI 影像辨識

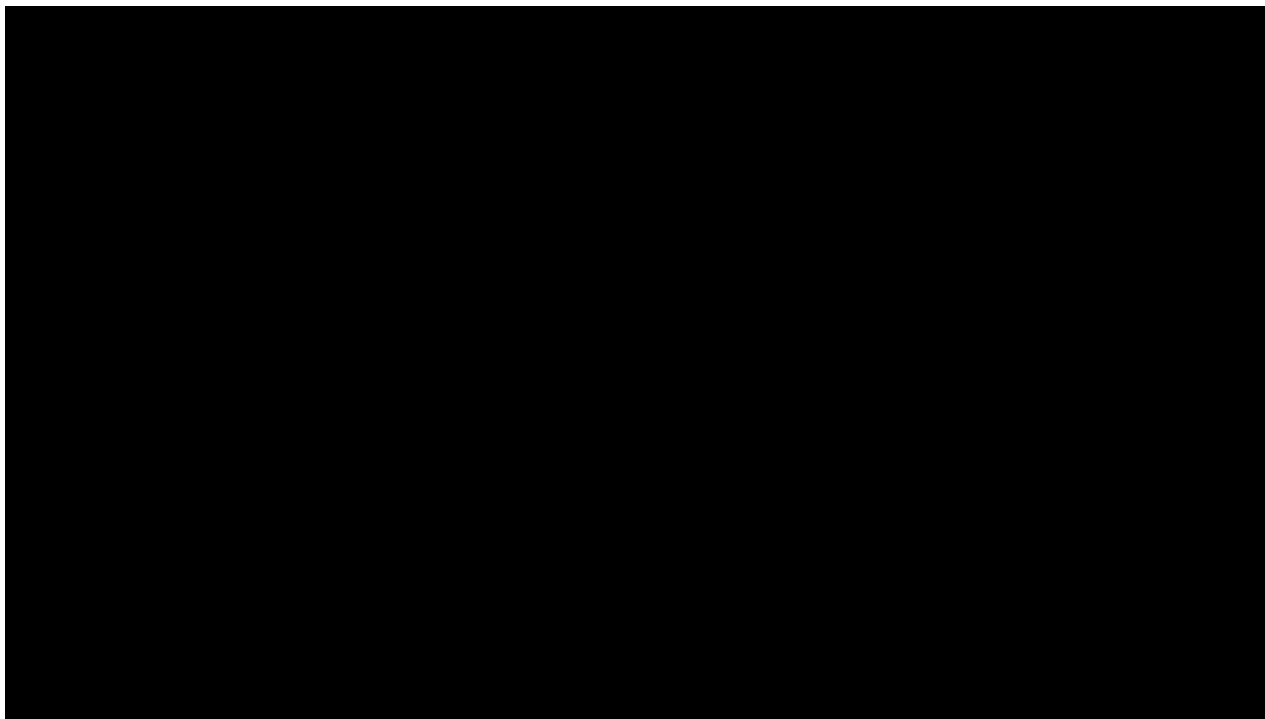
p. 191, 02:05 20190524





國情影片—美麗台灣

P191, 03:09 20180124



8-1 什麼是圖像識別

圖像識別是AI 的相關基礎技術與知識之一，為了避免對人工智慧技術群產生「見樹不見林」的迷失，在此重複呈現人工智慧大腦框架圖。請在閱讀本章內文時，仔細對照圖8-1，其中「感知層」的圖像識別，除了基本觀念與技術的說明外，並配合案例做介紹。

應 用				應用層
自然語言處理	知識圖譜	用戶畫像		認知層
視訊追蹤	AR / VR	圖像識別	語音辨識	感知層
深度學習		機器學習		演算法層
大數據分析/挖掘	數據標註	資料獲取		大數據層
雲儲存		雲計算		雲計算層

● 圖 8-1 人工智慧的基本框架——以百度大腦為例
(資料來源：百度 (2017))



8-1 什麼是圖像識別

隨著時代的進步，越來越多的事物逐漸依賴於越來越難以捉摸的人工智慧AI，如圖8-2。不過漸漸的，人們發現AI 也有一些缺陷需要做改進。

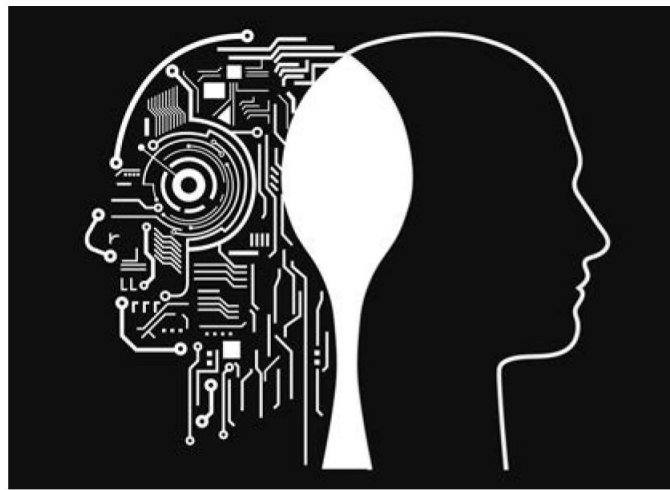


- 圖 8-2 模式識別技術應用於影視監控系統
(圖片來源：123RF 圖庫)



8-1 什麼是圖像識別

人類擁有記憶、且擁有「高明」的識別系統，假如告訴你在面前的一隻動物它是「貓(cat)」，以後你再看到貓，你一樣可以認出來。可是AI 不行，Google 實驗室做過實驗，人們需要給AI 透過識別一千萬張貓的圖片，才能讓它「認識」什麼是貓。AI 相對來說已經具備了一定的意識，但或許還是小學生而已。如果說AI 要透過那麼多張圖片才能認識什麼是貓，那麼AI 需要多少張圖片去認識這個世界呢？



● 圖 8-3 人工智慧的意識概念
(圖片來源：123RF 圖庫)



8-1 什麼是圖像識別

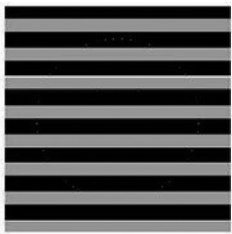
人類是透過眼睛查看到光源反射，「看」到了自己眼前的事物，但是可能很多內容元素人們並不在乎；就像你幾天前甚至更久前擦肩而過的一個人，如果你今天再次看到，你不一定會記得他，但是AI 會記住所有它見過的任何人、任何事物。



8-1 什麼是圖像識別

示例8-1 用簡單的條紋特徵識別物體

比如圖8-4，人類會覺得這是很簡單的雙色條紋。不過如果你問問最先進的人工智慧，它給出的答案也許會是校車，而且99%地肯定。



• 圖 8-4 雙色條紋

(圖片來源：123RF 圖庫)



8-1 什麼是圖像識別

以圖8-5 為例，AI 雖不能看出這是一條戴著墨西哥帽的吉娃娃狗，有的人也未必能認出，但是起碼能識別出這是一條戴著寬邊帽的狗。



- 圖 8-5 識別戴著墨西哥帽的吉娃娃狗
(圖片來源：123RF 圖庫)



8-1 什麼是圖像識別

懷俄明大學進化人工智慧實驗室的一項研究卻表明，AI 未必總是那麼靈光，AI 也會把這些隨機生成的簡單圖像當成了鸚鵡、乒乓球拍、披薩餅或者蝴蝶。當研究人員把這個研究結果提交給神經資訊處理系統大會進行討論時，專家形成了涇渭分明的兩派意見。一組人領域經驗更豐富，他們認為這個結果是完全可以理解的；另一組人則對研究結果的態度是困惑。至少他們在一開始對強大的AI 演算法，卻把結果完全弄錯感到驚訝。



8-1 什麼是圖像識別

圖像識別(image identification)，是指利用電腦對圖像進行處理、分析和理解，以識別各種不同模式的目標和物件的技術，是應用深度學習演算法的一種實踐應用。現階段圖像識別技術一般分為人臉識別與商品識別。人臉識別主要運用在安全檢查、身分核驗與行動支付中；商品識別主要運用在商品流通過程中，特別是無人貨架、智慧零售櫃等無人零售領域。

另外，在地理學中，圖像識別也是將遙感圖像進行分類的技術。

問題思考

圖像識別主要的目地或功用是對事物或對象做什麼？圖像識別與人臉識別、商品識別各自的關係是什麼？



8-1 什麼是圖像識別

一、人類的圖像識別能力

人的圖像識別能力是很強的。圖像距離的改變或圖像在感覺器官上作用位置的改變，都會造成圖像在視網膜上的大小和形狀改變。即使在這種情況下，人們仍然可以認出他們過去知覺過的圖像。甚至圖像識別可以不受感覺通道的限制。例如，人可以用眼看字，當別人在他背上寫字時，他也可認出這個字來。

模式識別（**pattern recognition**）原本是一項基本智慧，是指對表徵事物或現象的不同形式（數值的、文字的和邏輯關係的）的資訊做分析和處理，進而得到一個對事物或現象做出描述、辨認和分類等的過程。隨著電腦技術的發展和人工智慧的興起，人類本身的模式識別已經滿足不了社會發展的需要，於是人類就希望用電腦來代替或擴展人類的部分腦力勞動。



8-1 什麼是圖像識別

這樣電腦的模式識別就產生了，如圖8-6 的電腦指紋圖像識別技術就是類比人類的圖像識別過程。



- 圖 8-6 電腦類比人類的指紋圖像識別過程實例之一
(圖片來源：123RF 圖庫)



8-1 什麼是圖像識別

模式識別是資訊科學和人工智慧的重要組成部分。模式識別又常稱作模式分類，從處理問題的性質和解決問題的方法等角度，模式識別分為兩種：

- 1 有監督的分類 (supervised classification)

- 2 無監督的分類 (unsupervised classification)

模式還可分成抽象的和具體的兩種形式。抽象的形式如意識、思想、議論等，屬於概念識別研究的範疇，是人工智慧的另一研究分支。另一種具體形式模式識別主要是對語音波形、地震波、心電圖、腦電圖、圖片、照片、文字、符號、生物感測器等物件的具體模式進行辨識和分類。



8-1 什麼是圖像識別

模式識別是在圖像識別的進行過程中，不可少的步驟。要實現電腦視覺必須有影像處理的幫助，而影像處理依賴於圖像識別技術中模式識別的有效運用。

1 模式識別

模式識別是一門與數學緊密結合的科學，其中所用的思想方法大部分是概率與統計。模式識別主要分為三種：統計模式識別、結構/句法模式識別、模糊模式識別。

模式識別研究主要集中在兩方面：

1. 研究生物體（包括人）是如何感知物件的，屬於認知科學的範疇。
2. 在給定的任務下，如何用電腦實現模式識別的理論和方法。



8-1 什麼是圖像識別

2 圖像再認

圖形刺激作用於感覺器官，人們辨認出它是經歷過的某一圖形的過程，也叫圖像再認。在圖像識別中，既要有當時進入感官的資訊，也要有記憶中儲存的資訊。只有透過儲存的資訊與當前的資訊進行比較的加工過程，才能實現對圖像的再認。

問題思考

模式識別主要的目的或功用是對事物或對象做什麼？電腦視覺、影像處理、模式識別三者的關係是什麼？



8-1 什麼是圖像識別

二、圖像識別的基礎

圖像識別可能是以圖像的主要特徵作為識別的基礎。每個圖像都有它的特徵，如字母A 有個尖，B 有兩個圈、而Y 的中心有個銳角等。對圖像識別時有關眼動的研究表明，視線總是集中在圖像的主要特徵上，也就是集中在圖像輪廓曲度最大或輪廓方向突然改變的地方，這些地方的資訊量最大。而且眼睛的掃描路線也總是依次從一個特徵轉到另一個特徵上。

在圖像識別過程中，知覺機制必須排除輸入的多餘資訊，抽出關鍵的資訊。同時，在大腦裡必定有一個負責整合資訊的機制，它能把分階段獲得的資訊整理成一個完整的知覺映射。例如，有個有趣的順口溜：看人要看臉，看臉要看眼，看眼要看神。說明人們在「看人」的圖像識別關鍵點與過程。



8-1 什麼是圖像識別

在人類圖像識別系統中，對複雜圖像的識別往往要透過不同層次的資訊加工才能實現。對於熟悉的圖形，由於掌握了它的主要特徵，就會把它當作一個單元來識別，而不再注意它的細節了。這種由孤立的單元材料組成的整體單位叫做組塊，每一個組塊是同時被感知的。在文字材料的識別中，人們不僅可以把一個漢字的筆劃或偏旁等單元組成一個組塊，而且能把經常在一起出現的字或片語形成組塊單位來加以識別。



8-1 什麼是圖像識別

在電腦視覺識別系統中，圖像內容通常用圖像特徵進行描述，如圖8-7。事實上，基於電腦視覺的圖像檢索也可以分為類似於文本搜尋引擎的三個步驟：提取特徵、建立索引以及查詢。



- 圖 8-7 用圖像特徵進行描述
(圖片來源：123RF 圖庫)



8-1 什麼是圖像識別

三、電腦圖像識別模型

圖像識別是人工智慧的一個重要領域。為了發展能類比人類圖像識別活動的電腦程式，人們提出了不同的圖像識別模型。

1 與腦中的範本完全符合的範本匹配模型

這種模型認為，識別某個圖像，必須在過去的經驗中有這個圖像的記憶模式，又叫範本。當前的刺激如果能與大腦中的範本相匹配，這個圖像也就被識別了。



8-1 什麼是圖像識別

例如有一個字母A，如果在腦中有個A 範本，字母A 的大小、方位、形狀都與這個A 範本完全一致，字母A 就被識別了。這個模型簡單明瞭，也容易得到實際應用。但這種模型強調圖像必須與腦中的範本完全符合才能加以識別，而事實上人不僅能識別與腦中的範本完全一致的圖像，也能識別與範本不完全一致的圖像。例如，人們不僅能識別某一個具體的字母A，也能識別印刷體的、手寫體的、方向不正、大小不同的各種字母A。同時，人能識別的圖像是大量的，如果所識別的每一個圖像在腦中都有一個相應的範本，也是不可能的。



8-1 什麼是圖像識別

2 具有相似性的原型匹配模型

為了解決範本匹配模型存在的問題，格式塔心理學家又提出了一個原型匹配模型。這種模型（原型匹配模型）認為，在長時記憶中儲存的並不是所要識別的無數個範本，而是圖像的某些「相似性」。從圖像中抽象出來的「相似性」就可作為原型，拿它來檢驗所要識別的圖像。如果能找到一個相似的原型，這個圖像也就被識別了。



8-1 什麼是圖像識別

這種模型從神經上和記憶探尋的過程上來看，都比範本匹配模型更適宜，而且還能幫助對一些不規則的，但與原型相似的圖像的識別。

但是，這種模型沒有說明人是怎樣對相似的刺激進行辨別和加工的，它也難以在電腦程式中得到實現。因此又有人提出了一個更複雜的模型，即「泛魔」識別模型。這種方式在一般工業已使用中，採用工業相機拍攝圖片，然後利用軟體根據圖片灰階差做處理後識別出有用資訊。



8-1 什麼是圖像識別

四、圖像識別的發展

圖像識別的發展經歷了三個階段，如圖8-8 所示：文字識別、數位影像處理與識別、物體識別。

- 1 文字識別：是從1950 年開始研究的，一般是識別字母、數位和符號，從印刷文字識別到手寫文字識別，應用非常廣泛。



• 圖 8-8 圖像識別發展的三個階段



8-1 什麼是圖像識別

- 2 數位影像處理和識別：研究開始於1965 年。數位圖像比類比圖像更具有儲存性，傳輸方便可壓縮、傳輸過程中不易失真、處理方便等許多優勢，這些都為圖像識別技術的發展提供了強大的動力。
- 3 物體識別：主要是指對三維(3D) 世界的物體(object) 及環境的感知和認識，屬於高級的電腦視覺範疇。它是以數位影像處理與識別為基礎，並結合人工智慧、系統學等學科的研究方向，其研究成果被廣泛應用在各種工業及探測機器人上。

現代圖像識別技術有一個不足就是自我調整性能差，一旦靶心圖表像被較強的雜訊污染或是靶心圖表像有較大殘缺往往就得出理想的結果。



8-1 什麼是圖像識別

目前，在圖像識別的發展中，主要有三種識別方法：統計模式識別、結構模式識別、模糊模式識別，如圖8-9。其中，圖像分割技術是影像處理中圖像識別的一項關鍵技術。自20 世紀70 年代，圖像分割方法的研究已經有幾十年的歷史，一直都受到人們的高度重視，至今借助於各種理論提出了數以千計的分割演算法，而且這方面的研究仍然在積極地持續進行著。



● 圖 8-9 圖像識別主要的三種識別方法（統計、結構、模糊）與關鍵技術圖像分割



8-1 什麼是圖像識別

但在近二十年間，隨著基於長條圖和小波變換的圖像分割方法的研究，以及計算技術、VLSI 技術的迅速發展，有關影像處理方面的研究取得了很大的進展。圖像分割方法結合了一些特定理論、方法和工具，如基於數學形態學的圖像分割、基於小波變換的分割、基於遺傳演算法的分割等等。

圖像分割就是按照應用要求，把圖像分成各具特徵的區域，從中提取出感興趣目標。在圖像中常見的特徵有灰度、彩色、紋理、邊緣、角點等。例如，對汽車裝配流水線圖像進行分割，分成背景區域和工件區域，提供給後續處理單元對工件安裝部分的處理。



8-1 什麼是圖像識別



問題思考

圖像識別的發展經歷了三個階段：文字識別、數位影像處理與識別、物體識別。目前圖像識別技術的應用主要有人臉識別與商品識別以及最近熱門的地理地形識別，你認為三個階段的技術對上面的應用分別有哪些關聯或是有貢獻？為什麼？



8-1 什麼是圖像識別

五、機器視覺的發展

機器視覺（**machine vision**）是人工智慧領域中發展迅速的一個重要分支，正處於不斷突破、走向成熟的階段。一般認為機器視覺是透過光學裝置和非接觸感測器，自動地接受和處理一個真實場景的圖像，透過分析圖像獲得所需資訊或用於控制機器運動的裝置。由此可知，智慧影像處理技術在機器視覺中佔有舉足輕重的位置。

進入21 世紀後，機器視覺技術的發展速度更快，已經大規模地應用於多個領域，如智慧製造、智慧交通、醫療衛生、安防監控等領域。目前，隨著人工智慧浪潮的興起，機器視覺技術正處於不斷突破、走向成熟的新階段。



8-1 什麼是圖像識別

常見機器視覺系統，主要可分為兩類：

- 1 基於電腦的，如工業控制機（工控機）或PC 的機器視覺系統。
- 2 更加緊密的嵌入式機器視覺設備。

典型的基於工控機的機器視覺系統主要包括：光學系統、攝像機和工控機，還包含圖像採集、影像處理和分析、控制/ 通信等單元。機器視覺系統對核心的影像處理要求演算法準確、快捷和穩定，同時還要求系統的實現成本低，升級換代方便。

延伸學習

圖像採集就是從工作現場獲取場景圖像的過程，是機器視覺的第一步，採集工具大多為CCD 或CMOS 照相機或攝像機。



8-1 什麼是圖像識別

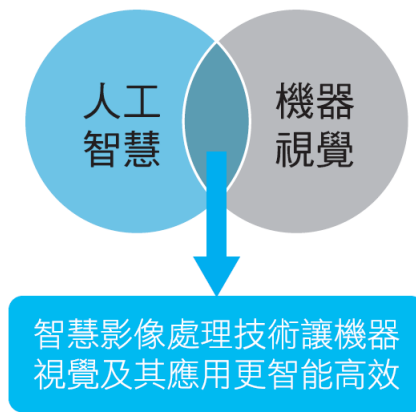
延伸學習

圖像採集就是從工作現場獲取場景圖像的過程，是機器視覺的第一步，採集工具大多為CCD 或CMOS 照相機或攝像機。



8-1 什麼是圖像識別

智慧影像處理(intelligent image processing) 是指一種可基於電腦的自我調整性，在各種應用場合的影像處理和分析技術，本身是一個獨立的理論和技術領域，但同時又是機器視覺中的一項十分重要的技術支撐。人工智慧、機器視覺和智慧影像處理技術之間的關係如圖8-10 所示。

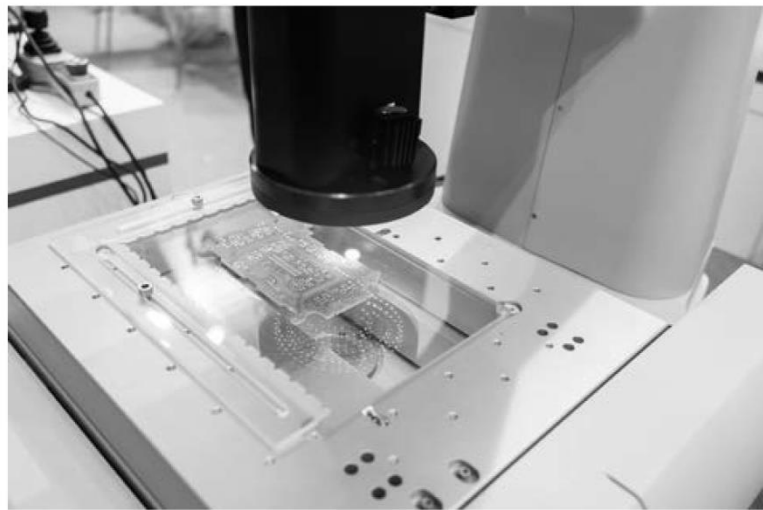


● 圖 8-10 智慧影像處理由 AI 與機器視覺組成



8-1 什麼是圖像識別

具有智慧影像處理功能的機器視覺，相當於人們在賦予機器智慧的同時，為機器裝上了眼睛，如圖8-11。讓機器能夠「看得見」、「看得準」，可替代甚至勝過人眼做測量和判斷，使得機器視覺系統可以實現高解析度和高速度的控制。而且，機器視覺系統與被檢測物件無接觸，安全可靠。



● 圖 8-11 智慧機器視覺應用於电路板的「目視」檢查
(圖片來源：123RF 圖庫)



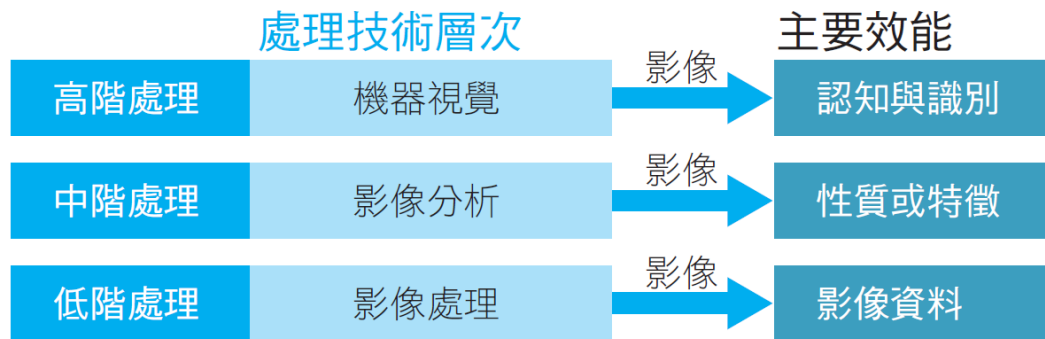
8-1 什麼是圖像識別

六、影像處理

影像處理(image processing) 又稱圖像處理，是利用電腦技術與數學方法，對圖像、視訊資料的表示、編解碼、圖像分割、圖像品質評價、目標檢測與識別，及立體視覺等方面展開科學研究。參考圖8-12，影像處理主要研究內容包括：圖像、視訊的模式識別和安全監控、醫學和材料影像處理，以及演化演算法、粗糙集(rough sets) 和資料採擷等人工智慧技術。在文字檢測和識別、人臉識別、指紋識別、語音辨識以及多個領域的資訊管理控制系統等方面均有廣泛應用。



8-1 什麼是圖像識別

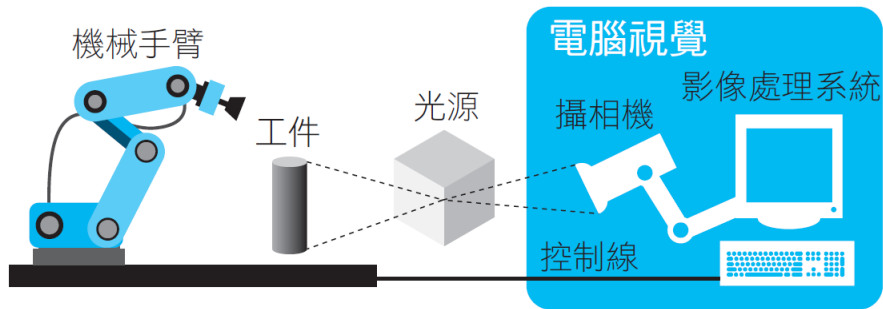


- 圖 8-12 機器視覺、影像分析與影像處理三者的關係比較



8-1 什麼是圖像識別

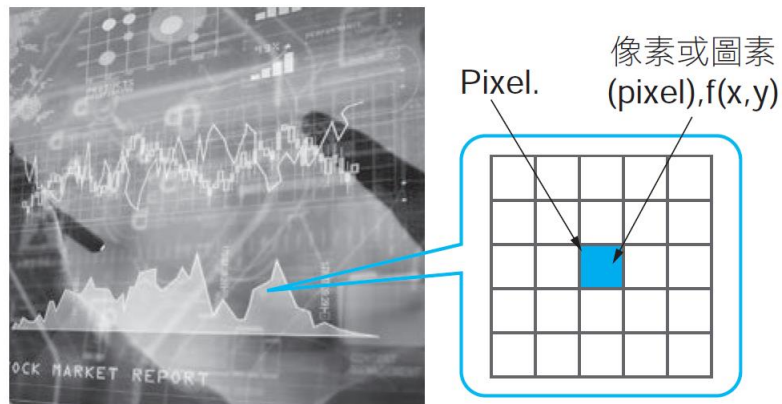
影像處理使用電腦對圖像進行分析，以達到所需結果的技術。影像處理一般指數位影像處理。數位圖像是指用數位攝像機、掃描器等設備，如圖8-13。經過採樣和數位化得到的一個大的二維陣列，該陣列的元素稱為像素(pixel, picture element)或圖素，其值為一整數，稱為灰度值如圖8-14。



- 圖 8-13 機器視覺系統實例—應用電腦視覺技術與工業機器人（機器手臂）搭配使用作為機器的眼睛



8-1 什麼是圖像識別



● 圖 8-14 影像構成與座標系統



8-1 什麼是圖像識別

◆ 表 8-1 影像處理技術與基本步驟

處理參考步驟	技術名稱	英文	備註
1	圖像（影像）擷取	Image Acquisition	
2A	圖像（影像）強化	Image Enhancement	同一或 同時處理
2B	圖像（影像）復原	Image Restoration	
3	彩色圖像（影像）處理	Color Image Processing	
3	小波理論與應用	Wavelets	
3	圖像（影像）壓縮	Image Compression	
3	形態處理	Morphological Processing	
3	圖像（影像）分割	Segmentation	
4A	表徵與描述	Representation and Description	同一或 同時處理
4B	認知與識別	Recognition	



8-1 什麼是圖像識別

影像處理技術的主要內容包括從圖像（影像）擷取、圖像壓縮、增強和復原、匹配、描述和識別等四個部分。常見的處理有圖像數位化、圖像編碼、圖像增強、圖像復原、圖像分割和圖像分析等，如表8-1。影像處理一般指對數位影像處理。



8-1 什麼是圖像識別

七、電腦視覺

從圖像（影像）處理和模式識別發展起來的電腦視覺（computer vision）是用電腦來類比人的視覺機制獲取和處理資訊的能力，就是指用攝影機和電腦代替人眼對目標進行識別、追蹤和測量等機器視覺，並進一步做圖形處理，用電腦處理成為更適合人眼觀察或傳送給儀器檢測的圖像資料格式。

電腦視覺研究相關的理論和技術，試圖建立能夠從圖像或者多維度資料中獲取「資訊」的人工智慧系統。電腦視覺的挑戰是要為電腦和機器人開發具有與人類水準相當的視覺能力。機器視覺需要圖像信號，紋理和顏色建模，幾何處理和推理，以及物體建模。一個有能力的視覺系統應該把所有這些處理都緊密地整合在一起。



8-1 什麼是圖像識別

電腦視覺研究物件之一是如何利用二維2D 投影圖像恢復三維3D 景物世界。電腦視覺使用的理論方法主要是基於幾何、概率和運動學計算與三維3D重構的視覺計算理論，它的基礎包括投影幾何學、剛體運動力學、概率論與隨機過程、影像處理、人工智慧等理論。電腦視覺要達到的基本目的有以下幾個：

- 1 根據一幅或多幅二維2D 投影圖像計算出觀察點到目標物體的距離
- 2 根據一幅或多幅二維2D 投影圖像計算出目標物體的運動參數
- 3 根據一幅或多幅二維2D 投影圖像計算出目標物體的表面物理特性
- 4 根據多幅二維2D 投影圖像恢復出更大空間區域的投影圖像



8-1 什麼是圖像識別

電腦視覺要達到的最終目的是實現利用電腦對於三維3D 景物世界的理解，即實現人的視覺系統的某些功能。

在電腦視覺領域裡，醫學圖像分析、光學文字識別對模式識別的技術需要提到一定程度。例如，模式識別中的預處理和特徵抽取環節，應用了影像處理的技術；影像處理中的圖像分析，也應用模式識別的技術。在電腦視覺的大多數實際應用當中，電腦被預設為解決特定的任務，然而基於機器學習的方法正日漸普及，一旦機器學習的研究進一步發展，未來「泛用型」的電腦視覺應用或許可以成真。



8-1 什麼是圖像識別

人工智慧所研究的一個主要問題是：如何讓系統具備「計畫」和「決策能力」？進而使系統能完成特定的技術動作（例如：移動一個機器人通過某種特定環境）。這一問題便與電腦視覺問題息息相關。在這裡，電腦視覺系統作為一個感知器，為決策提供資訊。另外一些研究方向包括模式識別和機器學習，也隸屬於人工智慧領域，而且與電腦視覺應用有著重要相關性。也因此，電腦視覺時常被看作人工智慧與電腦科學的一個分支。



8-1 什麼是圖像識別

為了達到電腦視覺的目的，有兩種技術途徑可以考慮：

1 仿生學方法

即從分析人類視覺的過程入手，利用大自然提供給我們的最好參考系統——人類視覺系統，建立起視覺過程的計算模型，然後用電腦系統實現之。

2 工程方法

即脫離人類視覺系統框框的約束，利用一切可行和實用的技術手段實現視覺功能。此方法的一般做法是，將人類視覺系統作為一個黑盒子對待，實現時只關心對於某種輸入，視覺系統將給出何種輸出。



8-1 什麼是圖像識別

這兩種方法理論上都是可以使用的，但面臨的困難是，人類視覺系統對應某種輸入的輸出到底是什麼，這是無法直接測得的。而且由於人的智慧活動是一個多功能系統綜合作用的結果，即使是得到了一個輸入輸出對，也很難肯定它是僅由當前的輸入視覺刺激所產生的回應，而不是一個與歷史狀態綜合作用的結果。



8-1 什麼是圖像識別

電腦視覺研究的雙重意義：

其一，是為了滿足人工智慧應用的需要，即用電腦實現人工的視覺系統的需要。這些成果可以安裝在電腦和各種機器上，使電腦和機器人能夠具有「看」的能力。

其二，視覺計算模型的研究結果反過來對於我們進一步認識和研究人類視覺系統本身的機理（制），甚至人腦的機理（制），也同樣具有相當大的參考意義。



8-1 什麼是圖像識別

八、電腦視覺與機器視覺的區別與關係

一般認為，電腦就是機器的一種，那麼，電腦視覺與機器視覺有什麼區別與相關性呢？

1 定義不同

(1) 電腦視覺

是一門研究如何使機器「看」的科學，就是指用攝影機和電腦代替人眼對目標進行識別、追蹤和測量等機器視覺，並進一步做圖形處理，使電腦處理成為更適合人眼觀察或傳送給儀器檢測的圖像資料。



8-1 什麼是圖像識別

(2) 機器視覺

是用機器代替人眼來做測量和判斷。機器視覺系統是透過機器視覺產品，將被攝取目標轉換成圖像信號，進而根據判別的結果來控制現場的設備動作。機器視覺產品的圖像攝取裝置，又分為兩種：CMOS 和CCD 電荷耦合元件。



8-1 什麼是圖像識別

2 原理不同

(1) 電腦視覺

是用各種成像系統代替視覺器官作為輸入感應手段，由電腦來代替大腦完成處理和解釋。電腦視覺的最終研究目標就是使電腦能像人那樣透過視覺觀察和理解世界，具有自主適應環境的能力。這些都是需要人類經過長期的努力才能達到的目標。



8-1 什麼是圖像識別

在實現最終目標以前，人們努力的中期目標是建立一種視覺系統，這個系統能依據視覺感知和回饋某種程度的智慧完成一定的任務。例如，電腦視覺的一個重要應用領域，就是自主車輛的視覺導航。目前還沒有技術能力件實現像人那樣能識別和理解任何環境，完成自主導航的視覺系統。

因此，人們先努力進行的研究目標是實現在高速公路上具有道路追蹤能力，可避免與前方車輛碰撞的視覺輔助駕駛系統。這裡要指出的一點是，在電腦視覺系統中，使用電腦取代人腦的作用，但並不意味著電腦必須按人類視覺的方法完成視覺資訊的處理。



8-1 什麼是圖像識別

(2) 機器視覺

其檢測系統採用CMOS 或CCD 照相機，將被檢測的目標轉換成圖像信，傳送給專用的影像處理系統，根據像素/ 圖素分佈和亮度、顏色等資訊，轉變成數位化信號，影像處理系統對這些信號進行各種運算來抽取目標的特徵，如面積、數量、位置、長度，再根據預設的允許度和其他條件輸出結果，包括尺寸、角度、個數、合格/ 不合格、有/ 無等，實現自動識別功能。進而根據判別的結果，來控制現場的設備動作。



8-1 什麼是圖像識別

◆ 表 8-2 電腦視覺與機器視覺的區別

	項目	說明
電腦視覺 Computer vision	定義	用攝影機和電腦（代替人眼）對目標進行識別、追蹤和測量等機器視覺，並進一步做各種圖形處理、模式識別、人工智慧技術相結合的手段； 屬於中低階處理以支援更多的應用所需的基本核心技術
	原理	用各種成像系統代替視覺器官作為輸入感應手段，由電腦來代替大腦完成處理和解釋。
		使電腦能像人那樣透過視覺觀察和理解世界，具有自主適應環境的能力。



8-1 什麼是圖像識別

機器視覺 Machine vision

定義

用機器代替人眼來做測量和判斷。把被攝目標的形態資訊，轉變成數位化信號；
屬於中高階處理，以應用於工業領域的現場情景機器工作需要。

原理

將數位化信號透過影像處理系統進行各種運算、抽取目標的特徵，如面積、數量、位置、長度；再根據預設的允許度和其他條件輸出結果，如包括尺寸、角度、個數、合格 / 不合格、有 / 無等，實現自動識別功能。進而根據判別的結果來控制現場的設備動作。
計算機視覺技術工程化，能夠自動獲取和分析特定圖像，以控制相應的行為。

主要指工業領域的視覺研究與應用，例如自主機器人的視覺，用於檢測和測量的視覺。透過軟體、硬體，圖像感知與控制理論與圖像處理等技術緊密結合，實現工業領域所需高效的機器人控制或各種現場即時操作。



8-1 什麼是圖像識別

延伸學習

- (1) CMOS 代表complementary metal-oxide-semiconductor（互補金屬氧氣半導體），一般手機或PC 個人電腦的照相機較多採用。

CCD 代表charge-coupled device（電荷耦合器件；感光耦合元件），工業用照相機用較多。

- (2) 用電腦資訊處理的方法研究人類視覺的機理，建立人類視覺的計算理論，也是一項非常重要和令人感興趣的研究領域。這方面的研究被稱為電腦視覺(computer vision)或計算視覺(computational vision)。計算視覺被認為是電腦視覺中的一個研究領域。



8-1 什麼是圖像識別

九、神經網路的圖像識別技術

神經網路圖像識別技術是在傳統的圖像識別方法和基礎上，融合神經網路演算法的一種圖像識別方法。在神經網路圖像識別技術中，遺傳演算法與倒傳遞神經網路(back propagation neural network, BPNN) 相融合的神經網路圖像識別模型是非常經典的，在很多領域都有它的應用。在圖像識別系統中利用神經網路系統，一般會先提取圖像的特徵，再利用圖像所具有的特徵映射到神經網路，進行圖像識別分類。



8-1 什麼是圖像識別

以汽車拍照自動識別技術應用在車牌辨識系統（license plate recognitionsystem）為例，參考圖8-15，當汽車通過的時候，檢測設備會有所感應。此時檢測設備就會啟用圖像採集裝置來獲取汽車正反面的圖像。獲取圖像後須將圖像上傳到電腦進行保存以便識別。最後車牌定位模組就會提取車牌資訊，對車牌上的字元進行識別與比對、並顯示最終的結果。在對車牌上的字元進行識別的過程中就用到了基於範本匹配演算法和基於人工神經網路演算法等技術。



● 圖 8-15 神經網路的圖像識別技術的應用案例
（圖片來源：123RF 圖庫）



8-1 什麼是圖像識別



問題思考

請從定義不同、原理以及應用不同的觀點，說明電腦視覺與機器視覺的差異。



8-2 圖像識別的應用案例簡介

圖像是人類獲取和交換資訊的主要來源，與圖像相關的圖像識別技術是目前的研究重點之一。電腦的圖像識別技術在公共安全、生物、工業、農業、交通、醫療等很多領域都有應用。例如，交通方面的車牌識別系統，公共安全方面如圖8-16 的瞳孔生物辨識技術、人臉識別技術、指紋識別技術，農業方面的種子識別技術、食品品質檢測技術，醫學方面的心電圖識別技術等。隨著電腦技術的不斷發展，圖像識別技術也在不斷地優化，其演算法也在不斷地改進



● 圖 8-16 瞳孔的生物辨識是圖像識別技術的應用之一
(圖片來源：123RF 圖庫)



8-2 圖像識別的應用案例簡介

一、機器視覺的行業應用案例

在工業領域中，機器視覺的應用主要體現在半導體及電子行業，其中大概40%~50% 都集中在半導體行業。機器視覺在各行業應用方面，已經擴及到製藥、包裝、電子、汽車製造、半導體、紡織、煙草、交通、物流等各個行業。用機器視覺技術取代人工，可以提高生產效率和產品品質。



8-2 圖像識別的應用案例簡介

示例8-2 機器視覺在半導體及電子行業的品質檢測應用例

機器視覺系統在品質檢測的許多領域已得到了廣泛的應用，其產品在生產製造檢測應用中，佔據著舉足輕重的地位，舉例如下。

1. 具體如PCBs 印刷電路(printed circuit boards) 材料識別應用：

各類生產印刷電路板組裝技術、設備；單、雙面、多層線路板，覆銅板及所需的材料及輔料；輔助設施以及耗材、油墨、藥水藥劑、配件；電子封裝技術與設備；絲網印刷設備及絲網周邊材料等。



8-2 圖像識別的應用案例簡介

2. SMT 表面貼裝技術(surface-mount technology) 貼裝識別應用：

SMT 工藝與設備、焊接設備、測試儀器、返修設備及各種輔助工具及配件、SMT 材料、貼片劑、膠粘劑、焊劑、焊料及防氧化油、焊膏、清洗劑等；再流焊機、波峰焊機及自動化生產線設備。

3. 電子生產加工設備貼裝識別應用：

電子元件製造設備、半導體及積體電路製造設備、元器件成型設備、電子工模具。



8-2 圖像識別的應用案例簡介

示例8-3 3D機器視覺在農業、醫藥業等的品質檢測應用例

隨著經濟水準的提高，3D 機器視覺也開始進入人們的視野。3D 機器視覺大多用於水果和蔬菜、醫藥產品、電子元件、化妝品、烘焙食品和木材的評級。

它可以提高合格產品的生產能力，在生產過程的早期就報廢劣質產品，減少了浪費、節約成本。這種功能非常適合用於高度、形狀、數量甚至色彩等產品屬性的實物成像檢驗。



8-2 圖像識別的應用案例簡介



(a) 水果分級示例



(b) 藥瓶生產與品質檢測

- 圖 8-17 3D 機器視覺對 (a) 農業水果分級
(b) 藥瓶品質的生產與品質檢測示例
(圖片來源：123RF 圖庫)



8-2 圖像識別的應用案例簡介

示例8-4 機器視覺在物流行業的分揀分類應用案例

機器視覺在物流行業或現代化快遞公司，使用機器視覺技術進行快遞的分揀分類，用較少的人工進行分揀，降低物品的損壞率，大大提高分揀效率，減少人工勞動。

Video

請搜尋youtube 影片中的機器視覺與物流行業，觀賞以下影片：

《雙11 現場直擊！ Yahoo 奇摩電商全新AI 自動化物流中心首揭神秘面紗》



8-2 圖像識別的應用案例簡介

二、檢測與機器人視覺應用案例簡介

機器視覺的應用主要有檢測和機器人視覺兩個方面：

(1) 檢測

又可分為高精度定量檢測和不用量器的定性或半定量檢測兩種。前者如顯微照片的細胞分類、機械零部件的尺寸和位置測量；後者如產品的外觀檢查、裝配線上的零部件識別定位、缺陷性檢測與裝配完全性檢測。



8-2 圖像識別的應用案例簡介

(2) 機器人視覺

用於指引機器人在大範圍內的操作和行動。例如，在生產現場中，從料斗送出的雜亂工件堆中揀取工件，並按一定的方位放在傳輸帶或其他設備上，即料斗揀取問題。至於小範圍內的操作和行動，還需要借助於觸覺傳感技術。



● 圖 8-18 3D 機器人視覺、輸送帶與工件揀取
(圖片來源：123RF 圖庫)



8-2 圖像識別的應用案例簡介

此外還有自動光學檢查、人臉識別、無人駕駛汽車、產品品質等級分類、印刷品質自動化檢測、文字識別、紋理識別、追蹤定位等等機器視覺圖像識別的應用。



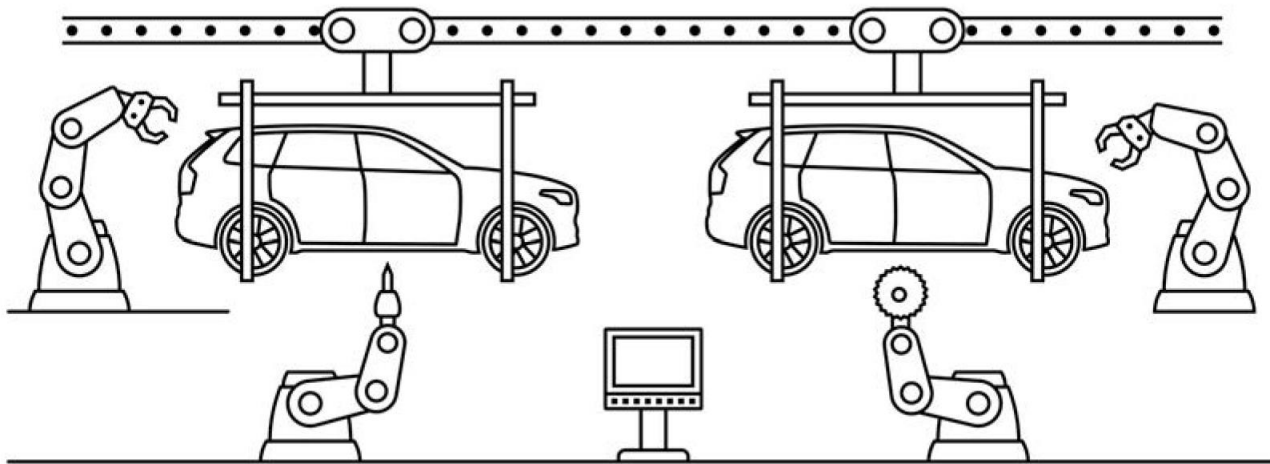
8-2 圖像識別的應用案例簡介

示例8-5 機器人視覺在汽車車身檢測系統應用案例

英國 ROVER 汽車公司800 系列汽車車身輪廓尺寸精度的100% 線上檢測，是機器視覺系統用於工業檢測中，一個較為典型的成功示例（圖8-19）。該系統由62 個測量單元組成，每個測量單元包括一台雷射器和一個CCD 攝像機，用以檢測車身外殼上288 個測量點。汽車車身置於測量框架下，透過軟體校準車身的精確位置。



8-2 圖像識別的應用案例簡介



• 圖 8-19 汽車線上檢測的示例
(圖片來源：123RF 圖庫)



8-2 圖像識別的應用案例簡介

測量單元的校準將會影響檢測精度，因而受到特別重視。每個雷射器/ 攝像機單元均在離線狀態下經過校準。同時還有一個在離線狀態下用三座標測量機校準過的校準裝置，可對攝像頭進行線上校準。

檢測系統以每40 秒檢測一個車身的速度，檢測三種類型的車身。系統將檢測結果與人、從CAD 模型中提取出來的合格尺寸相比較，測量精度為 $\pm 0.1\text{mm}$ 。

ROVER 的品質檢測人員用該系統來判別關鍵部分的尺寸一致性，如車身整體外型、門、玻璃窗口等。

實踐證明，該系統是成功的，並用於ROVER 公司其他系列汽車的車身檢測。



8-2 圖像識別的應用案例簡介

示例8-6 機器人視覺在紙幣印刷、瓶裝啤酒品質檢測系統應用案例

紙幣印刷品質檢測系統利用影像處理技術，透過對紙幣生產流水線上的紙幣20多項特徵，號碼、盲文、顏色、圖案等，進行比較分析，檢測紙幣的品質，替代傳統的人眼辨別的方法。

瓶裝啤酒生產流水線檢測系統可以檢測啤酒是否達到標準的容量、啤酒標籤是否完整。



8-2 圖像識別的應用案例簡介

示例8-7 軸承即時監控應用案例

視覺技術即時監控軸承(bearing)的負載和溫度變化，消除超載和過熱的危險。將傳統上透過測量滾珠表面保證加工品質和安全操作，把被動式測量變為主動式監控。



8-2 圖像識別的應用案例簡介

三、機器視覺的自動識別技術應用案例分析：有效改善重複性勞動、容易出錯且效率低的布匹品質檢測

在布匹生產過程中，像布匹品質檢測這種有高度重複性和高專業智慧性的工作，通常只能靠專業人工檢測來完成。在現代化流水線後面常常可看到很多的檢測工人來執行這道工序，除了給企業增加巨大的人工成本和管理成本外，卻仍然不能保證100 %的檢驗合格率（即「零缺陷」）。對布匹品質的檢測是重複性勞動，容易出錯且效率低。採用機器視覺的自動識別技術，在大批量的布匹檢測中，可以大大提高生產效率和生產的高自動化程度。請參考圖8-20。



● 圖 8-20 應用機器視覺的布匹品質檢測主要處理流程觀念圖



8-2 圖像識別的應用案例簡介

1 特徵提取辨識

一般布匹檢測要能夠進行自動識別，須先利用高清晰度、高速攝像鏡頭拍攝標準圖像，在此基礎上設定一定標準；然後拍攝被檢測的圖像，再將兩者進行比對。在進行布匹品質檢測工程前，需要建立的專業認知有：

- 1 圖像的內容不是單一的圖像，每塊被測區域存在的雜質的數量、大小、顏色、位置不一定一致。
- 2 雜質的形狀難以事先確定。
- 3 由於布匹快速運動對光線產生反射，圖像中可能會存在大量的雜訊。
- 4 在流水線上，對布匹進行檢測，有即時性的要求。



8-2 圖像識別的應用案例簡介

由於上述原因，圖像識別處理時應採取相應的演算法，提取雜質的特徵，進行模式識別，實現智慧分析。

2 色質檢測

一般而言，從彩色CCD相機中獲取的圖像都是RGB圖像。也就是說每一個圖元都由紅（R）綠（G）藍（B）三個成分組成，表示RGB色彩空間中的一個點。問題在於這些色差不同於人眼的感覺。即使很小的雜訊也會改變顏色空間中的位置。所以無論我們人眼感覺有多麼的近似，在顏色空間中也不盡相同。基於上述原因，需要將RGB圖元轉換成為另一種顏色空間CIELAB。目的就是使我們人眼的感覺盡可能的與顏色空間中的色差相近。



8-2 圖像識別的應用案例簡介

3 Blob 斑點檢測

根據上面得到的處理圖像，依需求，在純色背景下檢測雜質色斑，並且要計算出色斑的面積，以確定是否在檢測範圍之內。因此影像處理軟體要具有分離目標、檢測目標，並且計算出其面積的功能。

Blob分析 (blob analysis) 是對圖像中相同圖元 (素) 的連通域進行分析，該連通域稱為Blob。經二值化 (**binary thresholding**) 處理後的圖像中色斑可認為是blob。Blob 分析工具可以從背景中分離出目標，並可計算出目標的數量、位置、形狀、方向和大小，還可以提供相關斑點間的圖形結構。



8-2 圖像識別的應用案例簡介

在處理過程中不是採用單個的圖元逐一分析，而是對圖形的行進做操作。圖像的每一行都用遊程長度編碼(RLE)來表示相鄰的目標範圍。這種演算法與基於像素的演算法相比，大大提高處理速度。



8-2 圖像識別的應用案例簡介

4 結果處理和控制

應用程式把返回的結果存入資料庫或使用者指定的位置，並根據結果控制機械部分做相應的運動。

根據識別的結果，存入資料庫進行資訊管理。以後可以隨時對資訊進行檢索查詢，管理者可以獲知某段時間內流水線的情況紀錄，為下一步的工作作出安排；可以獲知布匹的品質情況等等。



8-2 圖像識別的應用案例簡介

最後將機器視覺技術的四種主要應用方式，用圖8-21 來表示。



● 圖 8-21 機器視覺的技術與主要應用摘要圖



8-2 圖像識別的應用案例簡介

問題思考

請舉兩個行業或實際例子，說明機器視覺系統的圖像識別技術應用在品質檢測的貢獻。



參觀與討論：校園安防的圖像識別系統

實訓目的

在開始本實訓之前，請認真閱讀課程的相關內容。

- (1) 熟悉模式識別、圖像識別等概念和相關知識。
- (2) 熟悉機器視覺、電腦視覺與影像處理的概念和相關知識



實訓內容與步驟

- (1) 請組織或自行組織，參觀校園或停車場已經建設的「車牌辨識系統」與「智慧停車管理系統」，並瞭解校園安防系統中應用的圖像識別功能，理解這些功能的安全作用。
- (2) 請組織或自行組織，參觀校園已經建設的安防系統，瞭解校園安防系統中應用的圖像識別功能，理解這些功能的安全作用。
- (3) 討論：在國內的大型商場、道路交通等重要場合建設了較為完整的「天網」影片監控系統，請瞭解這樣的系統，討論它們存在的意義和可能存在的問題。紀錄：請紀錄你們小組討論的主要觀點，推選代表在課堂上簡單闡述你們的觀點。



ANPR System, Automatic Number Plate Recognition

P. 235, 2:54 20120614



<https://www.youtube.com/watch?v=VVGFr7g0qxU>



傳統車牌辨識與AI 車牌辨識的差異

車牌辨識專題 P. 235, 2:15 20190824



<https://www.youtube.com/watch?v=nLxAg452n0g>



Chapter 8 作業

一、重要關鍵字練習：根據英文關鍵字，把適當的中文編碼寫至對應的空格中

A. 影像處理

B. 模式識別

C. 機器視覺

D. 圖像識別

E. 電腦視覺

F. 三維（簡稱）

G. 電荷耦合器件；感光耦合元件

H. 監督的分類

I. 智能影像處理

J. 表面貼裝技術（簡稱）

題號	英文	中文	題號	英文	中文
1	SMT	J	6	3D	F
2	machine vision	C	7	intelligent image processing	I
3	computer vision	E	8	image processing	A
4	supervised classification	H	9	image identification	D
5	charge-coupled device	G	10	pattern recognition	B

Chapter 8 作業

一、重要關鍵字練習：根據英文關鍵字，把適當的中文編碼寫至對應的空格中

K. 物體

L. 圖元（素）

M. 互補金氧半導體（簡稱）

N. 無監督的分類

O. 電荷耦合器件（簡稱）

P. 計算視覺

Q. 二值化

R. 印刷電路（簡稱）

S. 車牌辨識

T. 對圖像中相同圖元的連通域（斑點）進行分析

題號	英文	中文	題號	英文	中文
11	object	K	16	CCD	O
12	computational vision	P	17	license plate recognition	S
13	blob analysis	T	18	unsupervised classification	N
14	PCB	R	19	binary thresholding	Q
15	pixel	L	20	CMOS	M



Chapter 8 作業

二、是非題

- ☐ 1. 模式識別是圖像識別處理過程的其中一個處理項目。
- ☒ 2. 要實現電腦視覺必須有影像處理的幫助，而影像處理依賴於模式識別技術中的圖像識別有效運用。
- ☐ 3. 圖像再認是透過儲存的資訊與當前的資訊進行比較的加工過程，實現對圖像的再認。



Chapter 8 作業

三、選擇題（單複選）

(A) 1. 模式識別原本是_____的一項基本智慧。

(A) 人類 (B) 動物 (C) 電腦 (D) 人工智慧

(C) 2. 人工智慧領域通常所指的模式識別主要是對語音波形、地震波、心電圖、腦電圖、圖片、照片、文字、符號、生物感測器等物件的具體模式進行_____。

(A) 分類和計算 (B) 清洗和處理

(C) 辨識和分類 (D) 儲存與利用



Chapter 8 作業

- (B) 3. 要實現電腦視覺必須有影像處理的幫助，而影像處理依賴於_____的有效運用。
- (A) 輸入和輸出 (B) 模式識別 (C) 專家系統 (D) 智慧規劃
- (D) 4. 模式識別是一門與概率與統計緊密結合的科學，主要分為三種，但下列_____模式識別不屬於其中之一。
- (A) 統計 (B) 句法/ 結構 (C) 模糊 (D) 智能
- (B) 5. 圖像識別是指利用_____對圖像進行處理、分析和理解，以識別各種不同模式的目標和物件的技術。
- (A) 專家 (B) 電腦 (C) 放大鏡 (D) 工程師



Chapter 8 作業

(A) 6. 圖形刺激作用於感覺器官，人們辨認出它是經歷過的某一圖形的過程，稱為_____。

(A) 圖像再認 (B) 圖像識別 (C) 影像處理 (D) 圖像保存

(C) 7. 圖像識別是以圖像的主要_____為基礎的。

(A) 元素 (B) 圖元 (C) 特徵 (D) 部件

(D) 8. 基於電腦視覺的圖像檢索可以分為類似文本搜尋引擎的三個步驟，但下列_____不屬於其中之一。

(A) 提取特徵 (B) 建立索引 (C) 查詢 (D) 清晰



Chapter 8 作業

(B) 9. 圖像識別的發展經歷了三個階段，但下列_____不屬於其中之一。

- (A) 文字識別 (B) 圖元識別
- (C) 物體識別 (D) 數位影像處理與識別

(A) 10. 現代圖像識別技術的一個不足是_____。

- (A) 自我調整性能差 (B) 圖像圖元不足
- (C) 識別速度慢 (D) 識別結果不穩定

(C) 11. 圖像識別的主要方法有三種模式，但下列_____識別不屬於其中之一。

- (A) 統計模式 (B) 結構模式 (C) 圖元模式 (D) 模糊模式



Chapter 8 作業

- (D) 12. _____是影像處理中的一項關鍵技術，一直都受到人們的高度重視。
- (A) 數據離散 (B) 圖像聚合 (C) 圖像解析 (D) 圖像分割
- (A) 13. 具有智慧影像處理功能的_____, 相當於人們在賦予機器智慧的同時為機器裝上了眼睛。
- (A) 機器視覺 (B) 圖像識別 (C) 影像處理 (D) 資料視訊
- (B) 14. 影像處理技術的主要內容包括3 個部分，但下列_____不屬於其中之一。
- (A) 圖像壓縮 (B) 數據排序
- (C) 增強和復原 (D) 匹配、描述和識別



Chapter 8 作業

- (D) 15. 影像處理一般指數位影像處理。常見的處理有圖像數位化、圖像編碼、圖像增強、_____等。
- (A) 圖像復原 (B) 圖像分割 (C) 圖像分析 (D) A、B 和C
- (A) 16. 機器視覺需要_____, 以及物體建模。一個有能力的視覺系統應該把所有這些處理都緊密地整合在一起。
- (A) B、C 和D (B) 圖像信號
- (C) 紋理和顏色建模 (D) 幾何處理和推理



Chapter 8 作業

- (D) 17. 電腦視覺要達到的基本目的是_____，以及根據多幅二維投影圖像恢復出更大空間區域的投影圖像。
- (A) 根據一幅或多幅二維投影圖像計算出觀察點到目標物體的距離
 - (B) 根據一幅或多幅二維投影圖像計算出目標物體的運動參數
 - (C) 根據一幅或多幅二維投影圖像計算出目標物體的表面物理特性；
 - (D) A、B 和C



Chapter 8 作業

- (B) 18. 神經網路圖像識別技術是在_____的圖像識別方法和基礎上融合神經網路演算法的一種圖像識別方法。
- (A) 3D (B) 傳統 (C) 智能 (D) 先進
- (C) 19. 圖像採集就是從_____獲取場景圖像的過程，是機器視覺的第一步。
- (A) 終端設備 (B) 資料儲存 (C) 工作現場 (D) 離線終端
- (A) 20. 圖像分割就是按照應用要求，把圖像分成不同_____的區域，從中提取出感興趣目標。
- (A) 特徵 (B) 大小 (C) 色彩 (D) 圖元



Chapter 8 作業

- (B) 21. 電腦視覺根據一幅或多幅二維投影圖像計算出觀察點到目標物體的距離外，還要計算出_____等。（複選）
- C
- D (A) 目標物體的圖像聚合 (B) 更大空間區域的投影圖像
(C) 目標物體的運動參數 (D) 目標物體的表面物理特性
- (C) 22. 影像處理技術的主要內容包括幾個部分，包括圖像壓縮、_____等。（請選兩項）
- D (A) 圖像轉換 (B) 數據排序
(C) 增強和復原 (D) 匹配、描述和識別