

编号: _____



长春理工大学

Changchun University of Science and Technology

工程光学课程论文

机器视觉系统在人工智能领域的作用及发展趋势

The Role and Development Trend

of Machine Vision System in The Field of Artificial Intelligence

学生姓名: _____ 付煜文

专 业: _____ 机械电子

学 号: _____ 150321128

指导教师: _____ 丁红昌 长春理工大学

学 院: _____ 机电工程学院

二〇一八年十二月

The Role and Development Trend
of Machine Vision System in The Field of Artificial Intelligence

A thesis submitted to the
Changchun University of Science and Technology
in partial fulfillment of the requirement
for the degree of
Bachelor of Engineering
in Mechatronic

By

Fu Yuwen

Supervisor: Professor Ding Hongchang

Institute of Mechanics, Changchun University of Science and Technology

December, 2018

长春理工大学 学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。本人完全意识到本声明的法律结果由本人承担。

作者签名：

日 期：

长春理工大学 学位论文授权使用声明

本人完全了解并同意遵守长春理工大学有关保存和使用学位论文的规定，即长春理工大学有权保留送交学位论文的副本，允许该论文被查阅，可以按照学术研究公开原则和保护知识产权的原则公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密及延迟公开的学位论文在解密或延迟期后适用本声明。

作者签名：

日 期：

导师签名：

日 期：

摘 要

伴随着科学技术的不断发展进步，我们也一步步的迈入了现代化社会的发展。手机、电脑、智能电视、机器人、智能家居等等一系列高科技产品层出不穷的呈现在我们眼前。而面对这些高科技产品在我们生活中、工作中发挥着越来越重要的作用，我们不禁会对这些高技术产品的产生兴趣。而首先出现在我们脑海中的就是机器视觉和人工智能这两个在现代科技社会中不得不提到的两个关键词。而机器视觉和人工智能也是最近讨论研究颇多的热门话题。而本文主要是通过对机器视觉和人工智能发展内容的概括，阐释机器视觉系统在人工智能领域的作用及发展趋势。

关键词：机器视觉；计算机视觉；人工智能；发展前景；趋势研究

Abstract

With the continuous development of science and technology, we also step by step into the development of modern society. Mobile phone, computer, smart TV, robot, smart home and so on a series of high-tech products emerge in an endless stream in front of our eyes. And in the face of these high-tech products in our life, work is playing an increasingly important role, we can not help but have an interest in these high-tech products. And the first thing that comes to mind are machine vision and artificial intelligence, two key words that have to be mentioned in the modern scientific and technological society. Machine vision and artificial intelligence are also hot topics recently discussed and researched. This paper mainly explains the role and development trend of machine vision system in the field of artificial intelligence by summarizing the development contents of machine vision and artificial intelligence.

Keywords: MV;CV;AI;Development Prospects;Trend Research

目 录

图形列表

表格列表

符号列表

字符

Symbol	Description	Unit
R	the gas constant	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
C_v	specific heat capacity at constant volume	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
C_p	specific heat capacity at constant pressure	$\text{m}^2 \cdot \text{s}^{-2} \cdot \text{K}^{-1}$
E	specific total energy	$\text{m}^2 \cdot \text{s}^{-2}$
e	specific internal energy	$\text{m}^2 \cdot \text{s}^{-2}$
h_T	specific total enthalpy	$\text{m}^2 \cdot \text{s}^{-2}$
h	specific enthalpy	$\text{m}^2 \cdot \text{s}^{-2}$
k	thermal conductivity	$\text{kg} \cdot \text{m} \cdot \text{s}^{-3} \cdot \text{K}^{-1}$
S_{ij}	deviatoric stress tensor	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
τ_{ij}	viscous stress tensor	$\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-2}$
δ_{ij}	Kronecker tensor	1
I_{ij}	identity tensor	1

算子

Symbol	Description
Δ	difference
∇	gradient operator
δ^\pm	upwind-biased interpolation scheme

缩写

CFD	Computational Fluid Dynamics
CFL	Courant-Friedrichs-Lewy
EOS	Equation of State
JWL	Jones-Wilkins-Lee
WENO	Weighted Essentially Non-oscillatory
ZND	Zel'dovich-von Neumann-Doering

第1章 引言

1.1 研究背景

计算机视觉¹ 在国民经济, 科学研究及国防建设等领域都有着广泛的应用。视觉的最大优点是与被测的对象无接触, 因此对观测与被测者都不会产生任何损害, 这是其他感觉方式无可比拟的。另外, 视觉方式所能观察的对象十分广泛, 人眼观察不到的范围, 计算机视觉也可以观察, 例如, 红外线, 微波, 超声波等人类就观察不到, 而计算机视觉则可以利用这方面的敏感器件形成红外线, 微波, 超声波等图像。因此, 可以说是扩展了人类的视觉范围。另外人无法长时间的观察对象, 计算机是不知疲倦, 始终如一的观察, 所以计算机视觉可以广泛的用于长时间恶劣的工作环境。不过, 计算机是仍处于十分不成熟的阶段, 其发展还远远落后于人们所希望的发展水平。

在信息化促进工业化的今天, 对检测的速度和精度提出了更高的要求, 检测技术方法, 精度和效率, 在一定程度上标志着企业的市场竞争力, 甚至标志的一个国家, 部门或行业的科技水平, 传统的检测大多依靠大量手工检测人眼检测来完成, 这种人工检测手段主观性强, 一致性差, 且存在安全隐患, 而且由于人眼易疲劳, 无法持续, 稳健的完成高重复性的质量检测工作, 也受到检测精度, 速度的限制, 有些特殊环境下的检测任务, 根本无法由人工来完成, 一种能够代替人工检测技术——视觉检测技术, 以其高效, 高精度, 灵活高速, 非接触式, 获取信息丰富等优点, 日益成为一种具有广泛应用潜力的检测技术, 涉及光学技术, 传感技术, 图像处理和分析技术, 人工智能, 机器学习, 自然计算, 模式识别, 智能科学, 认知科学等多学科, 结合机器人^[2]技术, 广泛应用于现代化制造业产品制造过程中。

在国外, 视觉检测技术已相对成熟, 并得到广泛应用, 而国内, 视觉检测系统应用于钢铁工业, 电子工业, 汽车工业^[2], 木材工业, 纺织工业, 食品加工业, 包装工业, 航空工业等行业, 主要引入国外的成套设备, 或者采用合作开发的方式, 国内厂商只承担低端的二次开发和机械结构设计, 而视觉检测系统的整体方案, 传感部件, 光源技术, 这样处理和分析算法被国外厂商所控制, 设备引入成本高, 维护不便且价格昂贵, 其软件风格和操作习惯都不适合本土操作人员, 因此, 视觉检测产品对大多数国内生产厂商来说可望不可及, 极大制约了视觉系统在生产领域的推广应用。因此, 研究具有自主知识产权的设计, 进行系统设计的理论方法和核心图像处理算法, 具有相当的迫切性, 战略性, 是打破国外相关技术垄断和封锁的主要手段, 对于提高企业信息化水平, 改善生产工艺, 提高劳动生产率, 降低原材料消耗和减轻劳动强度, 提高产品市场竞争力都有重要意义。

¹计算机视觉:<https://zh.wikipedia.org/wiki/%E8%AE%A1%E7%AE%97%E6%9C%BA%E8%A7%86%E8%A7%89>

第2章 概述

2.1 视觉检测技术的相关概念

视觉检测技术^[2]是一门面向特定视觉任务，建立在计算机视觉和图像处理基础之上，对目标对象进行定性检测和定量检测的一门新兴检测技术，具有高效、非接触性、获取的信息丰富等优点，涉及图像处理、图像分析和图像理解个层面的理论的方法，具有人类视觉层次好、感知的深刻相似性。图像处理是对原始图像进行一系列的操作和处理，获得改善和增强的新图像，内容涉及图像去噪、图像滤波、图像变换、图像增强的内容，图像处理过程输入的是图像，输出的是数据。图像分析是通过算法，从图像中提取初级的图像特征，是一个视觉层次的处理，主要针对图像中感兴趣的目标进行检测，例如边缘和角点检测、直线和圆检测、颜色纹理特征提取等。图像理解属于高层操作，抽象度高、数据量小，研究图像中的目标的性质和它们之间的相互关系，并通过图像内容含义的理解得出原客观场景的理解，从而指导和规划主观行为，图像理解过程中输入的数据，输出的是知识。

2.1.1 计算机视觉

计算机视觉^[2]是研究人类视觉的计算模型，利用计算机对描述景物的图像数据进行处理，以实现类似人的视觉感知功能，对客观世界的三维场景进行感知、识别和初理解，是计算机科学和自然科学的重要组成部分。计算机视觉的研究方法主要有：

一、仿生学的方法 参照人类视觉系统的机构原理，建立相应的处理模型完成类似的功能和工作。

二、工程的方法 从分析人类视觉过程的功能入手，并不刻意模拟，人类视觉内部结构。

三、仅考虑系统的输入和输出，并采取变成现有的可行手段实现系统功能。

2.1.2 广义机器视觉

广义机器视觉的概念与计算机视觉是一致的，是使用计算机图像处理技术从三维世界所感知的二维图像中研究和提取出三维景物世界的物理结构，达到对客观事物图像的识别、理解和控制。

2.1.3 狭义机器视觉

狭义机器视觉的概念是指工业视觉检测，与普通计算机视觉、模式识别、数字图像处理有明显区别，是计算机视觉最重要的应用之一。目前，最权威的机器视觉的定义是美国制造业工程师协会和机器人工业协会给出的：

“机器视觉是利用非接触式的光学传感器自动采集实景物图像并进行处理，以获得所需的信息，并控制机器和生产过程的装置。”

2.1.4 图像理解

图像理解是对图像的语义解释, 是以图像为对象, 知识为核心, 研究图像中有什么目标、目标之间的相互关系、图像是什么场景以及如何应用场景的现在, 一门科学。

2.1.5 图像工程

图像工程是全面系统研究图像理论方法、阐述图像技术原理, 推广图像技术及总结生产实践经验的新学科。计算机视觉开始作为一个人工智能问题来研究, 因此被称为图像理解, 事实上, 这两个名称也常混合使用, 本质上, 它们相互联系, 其研究的覆盖面和研究内容存在一定的交叉。机器视觉和计算机视觉有着千丝万缕的联系, 很多情况下都有作为同义词使用。计算机视觉侧重于场景的分析和图像解释的理论和算法, 计算机视觉是计算机科学的分支, 属于科学的范畴。视觉更关注图像的获取、系统的构造和算法的实现, 侧重于工程应用。

2.2 计算机视觉和机器视觉的区别

计算机视觉^[2] 的研究目标是使计算机具有通过一幅或多幅图像, 认知周围环境信息能力。这使计算机不仅能模拟人眼的功能, 而且更重要的是使计算机完成人眼所不能胜任的工作。**机器视觉**则是建立在计算机视觉基础理论上, 偏重于计算机视觉技术, 工程化应用。与计算机视觉研究的视觉模式识别、视觉理解等内容不同, 机器视觉的重点在于感知环境中物体的形状, 位置, 姿态, 运动等几何信息。**视觉检测**是建立在计算机视觉和图像处理基础上的一门新型检测技术, 用计算机来模拟人的视觉功能, 从中提取有用的信息, 通过图像处理获得被测事物的各种参数, 并最终用于实际检测、测量^[2]和控制, 具有高效, 高精度, 非接触式, 获取信息量丰富等优点。

计算机视觉 (computer vision) 和**机器视觉** (machine vision) 两个术语是不加以区分的, 在很多文献中也是如此, 但其实这两个术语是既有区别又有联系的。**计算机视觉**是采用图像处理, 模式识别, **人工智能**技术相结合的手段, 着重于一幅或多幅图像的计算机分析。图像可以由单个或者多个传感器获得, 也可以是单个传感器在不同时刻获取的图像序列。分析是对目标物体的识别, 确定目标物体的位置和姿态, 对三维景物进行符号描述和解释。在**计算机视觉**研究中, 经常使用几何模型, 复杂的知识表达, 采用基于模型的匹配和搜索技术, 搜索的策略常使用自底向上, 自顶向下, 分层和启发式控制策略。**机器视觉**则偏重于计算机视觉技术工程化, 能够自动获取和分析特定的图像, 以控制相应的行为。具体来说, 计算机视觉为机器视觉, 提供图像和景物分析的理论及算法基础, 机器视觉为计算机视觉的实现提供传感器模型、系统构造及实现手段。因此可以认为, 一个机器视觉系统就是一个能够

¹有关计算机视觉语言研究的综合性文献和分类, Keith Price 的**注解版文献** (Annotated Computer Vision Bibliography)(<http://www.visionbib.com/bibliography/contents.html>) 是极有价值的资源。

自动获取一幅或多幅目标物体图像，对获取图像的各种特征量进行处理、分析和测量，并对测量结果作出定性分析和定量解释，从而得到有关目标物体的某种认识，并作出相应决策的系统。**机器视觉系统**（图 2.1）的功能包括：物体定位、特征检测、缺陷判断、目标识别、计数和运动跟踪。

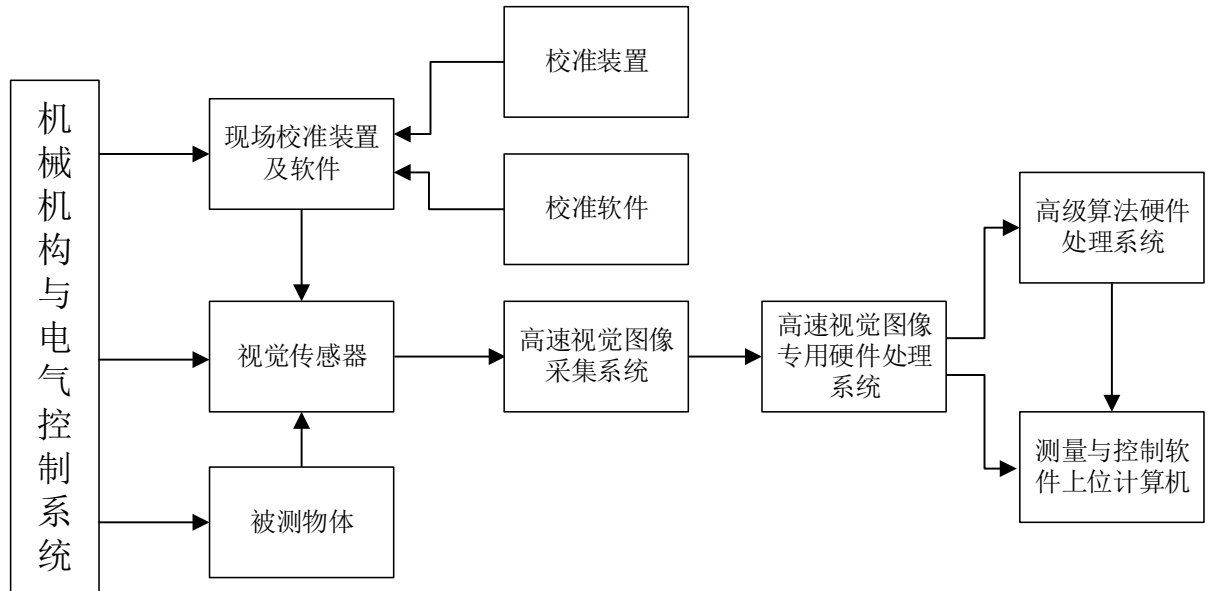


图 2.1 视觉测量系统原理框图

2.3 视觉检测涉及的图像处理方法

图像处理分析是视觉检测的基础，图像处理对获取的原始图像通过软计算来提高图像品质，为后续的图像分析和图像理解提供优质的图像数据资料，内容涉及去噪声、滤波、变换、分割、增强、特征提取、分类、识别等。

2.3.1 图像增强、降噪和滤波

图像增强、降噪和滤波是提高图像质量的常用方法，图像增强就是有选择的出图像中感兴趣的部分，衰减次要的信息，而不考虑图像降质的原因，从另一个角度看，图像增强相当于对图像进行滤波处理，主要方法包括直方图均衡化，几何校正、灰度修正、图像锐化，频域增强、自适应滤波、模糊集，卡尔曼滤波等，它们主要完成对图像的平滑处理。这些方法要么会降低图像信息量，要么会使图像变模糊，从而损失其中的纹理特征和边缘特征。

近年来提出的基于小波变换的滤波和基于 Contourlet 变换的滤波可以有效的避免该问题，较好地解决了频率分辨率和时间分辨率之间的矛盾，目前使用得最频繁的滤波方法小波域方法主要分为阈值萎缩法和比例萎缩法两大类。前者主要有硬阈值法、软阈值法、VisuShrink 阈值方法、子带自适应 VisuShrink 阈值法、基于 Bayes 准则的 Bayes shrink 方法、自适应多阈值的图像滤波法以及基于主分量的小波滤波方法。后者主要有 Mihcak 等人提出的 LAWMLHE 算法、LAWMAP 算法。

图像降噪但主要目的就是提高图像的信噪比，突出图像的应用特征，常用方法有随尺度变化的接近最优化自适应阈值算法、基于 HV 视觉模型下的图像降噪的方法、非下采样 Contourlet 域高斯尺度混合模型的图像降噪算法等。

虽然，以上提出的各种图像增强，降噪滤波算法日趋完善，在一定程度上能很好的改善图像的质量，但他们都是针对某种具体情况而设计的，不能通用，具有一定的局限性。况且这些算法的计算复杂度都比较大，尤其是各种频率的算法，不能满足高速，实时的图像处理系统的要求。

2.3.2 图像分割

图像分割是指图像中有意义的特征或区域提取提取出来的过程，其目的就是 will 图像划分成若干互不相交的区域，使各区域具有一致性，而相邻区域间的属性特征有明显差别。传统的图像分割算法通常分为基于区域的、基于边缘的和两者相结合的。基于区域分割的方法就是基于物体区域内像素间的同一性，如图像灰度值或纹理统一，把图像分割直接划分为若干子区域的图像分割方法。这种分割算法主要包括区域生长、分水岭变换法和马尔科夫随机场三种。

2.3.3 特征提取

特征提取是图像模式识别的一个重要环节，其目的是针对识别目标特点提取出一组尽量精简、最有效的特征，从而提高分类器的识别效果和效率。传统的特征提取法是将图像映射为一些数字描绘子，提取模式的不变特征描述矢量，并基于统计方法进行目标分类。该方法的缺点是由于现实世界中的多样性和高度复杂的模式，很难提供一种通用的目标不变性描述技术，可能在特征提取阶段丢弃模式的某些重要特征。基本缺陷特征包括直方统计特征、小波变换特征、基于灰度共生矩阵以及不变矩特征。

2.3.4 缺陷分类器

用于图像分类的算法层出不穷。传统分类方法包括基于贝叶斯决策理论的分类方法和 Fisher 准则分类方法。Sick 系统、CSM 系统、Sipar 系统采用树分类器技术。Smartvis 系统应用机器学习方法自动设计技术分类器结构，HTS-2 系统采用人工神经网络识别技术，iS-2000 系统是树分类器结构和自学习分类器相结合的识别技术。以上分类器设计方法均采用缺陷的几何和光学统计产量作为模式特征矢量，分类器的识别精度和可辨识的缺陷类别的数量完全取决于统计参量集的复杂程度，从而导致识别算法计算工具量巨大，给分类器的结构设计工作增加了相当大的难度。

2.3.5 图像识别

图像识别^[7]的主要任务是对预处理后的图像进行图像分割和特征提取，从而对目标物体的归属作出判决，属于**人工智能**和模式识别的范畴。人工智能与模式识别，主要研究如何用计算机实现人脑的功能，常见的方法用模板匹配，统计模式识别，句法结构模式识别和神经网络模式识别。

人工神经网络由大量处理单元互联组成，是在现代神经科学的研究基础上提出来的，主要模拟人脑神经系统的工作特点，具有自适应学习、自组织、容错力强等优点。一些较好的神经元网络模式是向后传播网络、高阶网络、时延和周期性网络。

人工神经网络方法^[2]的引入极大地促进了模式识别和计算机视觉的发展。其主要的模型和学习算法有前向网络模型、分层网络模型、Hopfield 模型、竞争学习、爱玻尔兹曼机及运动控制的学习模型等。此外还有基于遗传算法的神经网络模型、基于混沌理论的神经网络模型和核匹配追踪算法。

2.3.6 人工智能

人工智能²（英语：Artificial Intelligence，缩写为 AI）亦称机器智能^[2]，指由人制造出来的机器所表现出来的智能。通常人工智能是指通过普通计算机程序的手段实现的人类智能技术。该词也指出研究这样的智能系统是否能够实现，以及如何实现。同时，人类的无数职业也逐渐被其取代。

尼尔逊教授对人工智能下了这样一个定义：

“人工智能是关于知识的学科——怎样表示知识以及怎样获得知识并使用知识的科学。”

而另一个美国麻省理工学院的温斯顿教授认为：

“人工智能就是研究如何使计算机去做过去只有人才能做的智能工作。”

这些说法反映了人工智能学科的基本思想和基本内容。即人工智能是研究人类智能活动的规律，构造具有一定智能的人工系统，研究如何让计算机去完成以往需要人的智力才能胜任的工作，也就是研究如何应用计算机的软硬件来模拟人类某些智能行为的基本理论、方法和技术。

人工智能³是研究使计算机来模拟人的某些思维过程和智能行为（如学习、推理、思考、规划等）的学科，主要包括计算机实现智能的原理、制造类似于人脑智能的计算机，使计算机能实现更高层次的应用。人工智能将涉及到计算机科学、心理学、哲学和语言学等学科。可以说几乎是自然科学和社会科学的所有学科，其范围已远远超出了计算机科学的范畴，人工智能与思维科学的关系是实践和理论的关系，人工智能是处于思维科学的技术应用层次，是它的一个应用分支。从思维观点看，人工智能不仅限于逻辑思维，要考虑形象思维、灵感思维才能促进人工智能的突破性发展，数学常被认为是多种学科的基础科学，数学也进入语言、思维领域，人工智能学科也必须借用数学工具，数学不仅在标准逻辑、模糊数学等范围发挥作用，数学进入人工智能学科，它们将互相促进而更快地发展。

用来研究人工智能的主要物质基础以及能够实现人工智能技术平台的机器就是计算机，人工智能的发展历史是和计算机科学技术的发展史联系在一起的。除了计

²<https://zh.wikipedia.org/wiki/%E4%BA%BA%E5%B7%A5%E6%99%BA%E8%83%BD>

³<https://baike.baidu.com/item/%E4%BA%BA%E5%B7%A5%E6%99%BA%E8%83%BD/9180?fr=aladdin>

计算机科学以外，人工智能还涉及信息论、控制论、自动化^[2-3]、仿生学、生物学^[2]、心理学、数理逻辑、语言学、医学和哲学等多门学科。人工智能学科研究的主要内容包括：知识表示、自动推理和搜索方法、机器学习和知识获取、知识处理系统、自然语言理解、计算机视觉、智能机器人^[2-3]、自动程序设计等方面。

第3章 实际应用场景

3.1 机器视觉的应用

3.1.1 工业类应用

- 光学字符识别 (OCR)¹: 阅读信上的手写邮政编码和自动号码牌识别 (ANPR)
- 机械检验²: 为保证质量和快速检验部件, 使用立体视觉在专门的光照下测量飞机机翼或汽车车身配件的容差, 或使用 X 光视觉检查铜铸件的缺陷。
- 零售³: 针对自动结账通道的物体识别。
- 医学成像: 注册手术前和手术中的成像, 或进行关于人老化过程中大脑形态的长期研究。
- 汽车安全⁴
- 监视和交通监控 (<http://www.honeywellvideo.com/>)
- 变形: 使用无缝的变形过度将你一个人的照片变成另一个人的。
- 3d 建模: 想你所拍摄的物体或人物的一幅或多幅快照转换成其 3d 模型^[2],
- 视频匹配运动和本地化: 通过自动跟踪附近的参照点, 将 2d 图片或 3d 模型插入视频中, 或是用视觉估计去除视频中的抖动。
- 照片预览: 通过在 3d 中自由的展示照片来浏览大型照片收藏, 比如你房子内部的照片收藏。
- 人脸识别: 用于改进照相机的聚焦以及更新相关图像的输送。
- 视觉身份认证, 当家庭成员坐在网络摄像头前时, 自动给他们在家里的计算机上登录。

3.1.2 消费类应用

- 拼图: 将有重叠的照片变成无缝拼接起来的单张全景画。
- 曝光包围: 能在有困难光照 (强烈阳光和阴影) 条件下, 拍摄的多张曝光照片, 融合成单张完美曝光的图像。
- 汽车安全: 在诸如雷达或激光雷达等主动视觉技术不好用时检测意外的障碍物, 比如街道上的行人
- 匹配运动: 通过跟踪原视频中的特征点来估计摄影机的 3D 运动和环境形状, 将计算机生成的影像与实景真人动作脚本相融合, 这样的技术在好莱坞得到广泛应用。
- 运动捕捉: 使用从多台摄像机拍摄, 反光材料标记或其他识别方法来捕捉演员的动作, 以便用于计算机动画。

¹光学字符识别 (OCR) (<http://yann.lecun.com/exdb/mnist/index.html>)

²机械检验应用

³零售应用(<http://www.evoretail.com/>)

⁴汽车安全(<https://www.mobileye.com/>)

- 监视：监控入侵者者，分析高速公路的交通状况，监控用的是以防溺水事件的发生。
- 指纹识别和生物测定学：用于自动准入身份认证以及司法应用。

3.1.3 机器视觉系统成像应用

- 伽马射线图像：同位素检查
- X 射线图像：医学、工业、农业应用
- 紫外线图像：荧光分析
- 可见光图像：彩色图像、黑白图像
- 红外线图像：遥感、夜视、测温
- 微波成像：雷达
- 无线电波成像：核磁共振
- 超声波成像

3.2 计算机视觉应用领域

计算机视觉已广泛应用于工业自动化生产线、各类检验和监视^[2-3]、图像自动解释、人机交互^[2]及虚拟现实等领域。

3.2.1 工业自动化生产线应用

产品检测^[2]，工业探伤^[2]，自动流水线生产和装配，自动焊接，pcb 印制板检查以及各种危险场合工作的机器人等。将图像和视觉技术用于生产自动化，可以加快生产速度^[2]，保证质量的一致性，还可以避免人的疲劳，注意力不集中等带来的误判^[2]。

3.2.2 各种检验和监视应用

标签文字标记检查、邮政自动化、计算机辅助外科手术、显微医学操作、石油、煤矿等装炭中数据流自动监测和滤波、在纺织、印染业进行自动分色、配色、重要场所门廓自动巡视、自动跟踪报警等。

3.2.3 视觉导航应用

巡航导弹制导、无人驾驶飞机飞行、自动行驶车辆、移动机器人、精确制导及自动巡航捕捉目标和确定距离等，即可避图像自动解释应用。对方是图像显微图像，医学图像遥感，多波段图像，合成孔径雷达图像，航空行测图像等的自动判读理解、由于近年来技术的发展，图像的种类和数量飞速增长，图像的自动理解已成为解决信息膨胀问题的重要手段。

3.2.4 人机交互应用

人脸识别、智能代理等。同时让计算机可借助人的手势动作，嘴唇动作，表情测定等了解人的愿望要求而执行指令，这既符合人类的习惯，增加交互方便性和临场感等。

3.2.5 虚拟现实应用

飞机驾驶员训练，医学手术模拟，场景建模，战场环境表示等，它可帮助人们超越人的生理极限，“亲临其境”，提高工作效率。

第4章 挑战与机遇

4.1 计算机视觉面临的问题

对于人的世界来说，由于人的大脑和神经的高度发展及目标识别能力很强。但是人的视觉也同样存在障碍，例如，即使具有一双敏锐视觉和极为高度发达头脑的人，一旦至于某种特殊环境（即使曾经具备一定的先验知识），其目标识别能力也会急剧下降。事实上，人们在这种环境下对简单物体时，仍然可以有效而简便的识别，而面对复杂目标或特殊背景时则在视觉功能上发生障碍，两者共同的结果是导致目标识别的有效性和可靠性的大幅度下降。

将人的视觉引入计算机视觉中，计算机视觉也存在着这样的障碍^[2]。它主要表现在三个方面：

一、如何准确、高效（实时）地识别出目标。

二、如何有效的增大存储容量，以便容纳下足够细节的目标图像。

三、如何有效的构造和组织出可靠的识别算法，并且顺利的实现。前两者相当于人的大脑质量的物质基础，这期待着高速的阵列处理单元，以及算法（如神经网络，分类算法，小波变换算法）的新突破，用极小的计算量以及高度的并行性实现功能。

另外，由于当前对人类视觉系统的机理、人脑心理和生理的研究还不够，目前人们所建立的各种视觉系统，及大多数是只适用于某一特定环境或应用场合的专用系统，而是要建立一个可与人类的视觉系统相比拟的通用识别系统是非常困难的。

主要原因有以下几点：

(1) 图像对景物的约束不充分。首先是图像本身不能提供足够的信息来恢复景物，其次是当把三维景物投影成二维图像时丧失了深度信息。因此，需要附加的约束才能解决从图像恢复景物时的多义性。

(2) 多种因素在图像中相互混淆。物体的外表受材料性质、空气条件、光源角度、背景光照、摄像机角度和特性等因素的影响。所有这些因素都归结到一个单一的测量，即像素的灰度。要确定各种因素，对像素灰度的作用大小是很困难的。

(3) 理解自然景物，要求大量知识。例如要用到阴影、纹理、立体视觉、物体大小的知识；关于物体的专门知识或通用知识，可能还有关于物体间关系的知识等，由于所需的知识量极大，难以简单的用人工进行输入，可能要求通过自动知识获取方法来建立。

(4) 人类虽然自己就是视觉的专家，但它不同于人的问题求解过程，难以说出自己是如何看见事物，从而给计算机视觉的研究提供直接的指导。

人类视觉系统具有以下特点：具有高分辨率，有立体观察，优越的识别能力和灵活的推理能力，可灵活根据各种视觉线索进行推理。为了便于理解，现将人的视觉和计算机视觉^[2]对比于表4.1和表4.2

表 4.1 计算机视觉与人的视觉能力的比较

能力	计算机视觉	人的视觉
测距	能力有限	定量估计
定方向	定量计算	定量估计
运动分析	定量分析，但受限制	定量分析
检测边界区域	对噪声比较敏感	定量、定性分析
图像形状	受分割、噪声制约	高度发达
图像机构	需要专用软件，能力有限	高度发达
阴影	初级水平	高度发达
二维解释	对分割完善的目标能较好解释	高度发达
三维解释	较为低级	高度发达
总的能力	最适合于结构环境的定量测量	最适合于复杂的、非结构化环境的定量解释

表 4.2 计算机视觉与人的视觉性能标准比较

性能标准	计算机视觉	人的视觉
分辨率	能力有限	定量估计
处理速度	零点几秒/每帧图像	定量估计
处理方式	串行处理，部分并行处理	每只眼睛每秒处理（实时） 10^{10} 空间数据
视觉功能	二维、三维立体视觉有限	自然形式三维立体视觉
感光范围	紫外线、红外线、可见光	可见光

4.2 视觉检测面临的主要问题

视觉检测是一项多学科交叉的领域，设计传感器，处理系统结构，硬件软件系统及其评价等方面的各种理论，方法和专门知识，随着信息技术发展和信息检测理论的不断深入，视觉检测技术作为一种非接触式的检测手段，一，以其获得的信息丰富，便于利用先进的信息处理技术等优势，成为产品定性定量检测的自动化检测手段和质量控制方法，已在各行各业广泛应用，但视觉检测的发展和应用仍然存在许多问题，其应用的广度和深度不断向视觉检测理论方法和应用，提出新需求和新挑战。主要体现在以下几个方面：

(1) 面向产品定性检测和定量检测，向着高精度，高速度，高可靠性等方向发展，当高分辨率高速度的图像获取，处理相关的成像设备成本，信息处理带宽呈几何级数上升，限制这类高精度，高速检测系统的普及和应用。

(2) 目前视觉检测系统的检测任务都相对比较简单，对外界环境因素变化敏感，要求视觉检测系统能适应环境变化，具有较强的鲁棒性，即使在复杂多变的环境也能可靠工作，如何让视觉检测系统能更灵活，更广泛的适应实际复杂新环境，提高视觉检测的鲁棒性是值得深入的研究内容。

(3) 知识的利用问题。在很多视觉检测任务中，先验知识，对于消除视觉系统的不适应性、改善图像特征提取性能非常重要，充分利用特定图像的先验知识，以求得在实际应用中达到或者接近人类视觉特性方面的研究，引起了研究者的广泛关注。

(4) 在视觉检测中，目标的大小形状，出现方式都具有不确定性，况且为目标的大小，形状和出现方式也会随机出现。因此，为目标的鉴别与剔除以及识别的有效性与健壮性是目前视觉检测系统智能化的瓶颈问题。

第5章 机器视觉发展趋势与展望

5.1 机器视觉系统的发展

5.1.1 机器视觉系统的发展历程

模式识别：起源于20世纪50年代的机器视觉，早期研究主要从统计模式识别开始，工作主要集中在二维图像分析与识别上，如光学字符识别OCR（Optical Character Recognition）、工件表面图片分析、显微图片和航空图片分析与解释。

积木世界：20世纪60年代的研究前沿是以理解三维场景为目的的三维机器视觉。1965年，Roberts从数字图像中提取出诸如立方体、楔形体、棱柱体等多面体的三维结构，并对物体形状及物体的空间关系进行描述。他的研究工作开创了以理解三维场景为目的的三维机器视觉的研究。对积木世界的创造性研究给人们以极大的启发，许多人相信，一旦由白色积木玩具组成的三维世界，可以被理解，这样可以推广到更复杂的三维场景。于是，人们对积木世界进行了深入的研究。研究的范围从边缘、角点等特征提取，到线条、平面、曲面等几何要素分析，一直到图像明暗、纹理、运动以及成像几何等，并建立了各种数据结构和推理规则。

起步发展：20世纪70年代出现了一些视觉运动系统，与此同时，美国麻省理工大学的人工智能实验室正式开始“机器视觉”的课程，由国际著名学者B.K.P.Horn教授讲授。大批著名学者进入麻省理工大学，参与机器视觉理论、算法、系统设计的研究。1977年，David Marr教授在麻省理工大学的人工智能实验室领导一个以博士生为主体的研究小组，于1977年出了不同于“积木世界”分析方法的计算视觉理论，该理论在80年代成为机器视觉研究领域的一个十分重要的理论框架。

蓬勃发展：20世纪80年代到20世纪90年代中期，机器视觉获得蓬勃的发展，新概念、新方法、新理论不断涌现，如：基于感知特征群的物体识别理论框架、主动视觉理论框架，视觉集成理论框架等。到目前为止，机器视觉仍然是一个非常活跃的研究领域。

5.2 中国机器视觉系统的研究现状

随着中国企业^[2]生产自动化程度的提高，近年来，机器视觉在国内开始快速发展^[2]，中国国际机器视觉展览会每年举办，得到了行业的极大关注。近年来^[2]，国内机器视觉领域^[2]的研究机构和厂商纷纷加大投入，一致看好这一自动化领域的新技术市场。

5.2.1 机器视觉市场庞大

采用机器视觉可以完成人工很难实现的任务，特别在需要高速、高精度要求的系统中。比如，电子制造业^[2]、汽车制造业^{[2][3][4]}、包装与印刷业、化工、能源、加工机械等行业都是机器视觉的用户或者潜在用户。从国际市场来看，机器视觉目前

最大应用领域是半导体电子制造业^[7]。而中国目前已经成为全国主要的生产制造基地，全球一半以上的手机都是中国制造，很多半导体公司都在中国是有生产工厂，这些企业需要大量的机器视觉系统。

随着企业自动化程度的不断提高和对质量更加严格的控制要求，迫切需要机器视觉来代替人工检测。中国的工业生产正在从依赖廉价劳动力转向更高程度的自动化生产，这带来了对自动化设备的大量需求。另外，中国早期的工业设备自动化程度普遍较低，因此，需要大量的更新换代，这些都是构成了对包括机器视觉在内的自动化设备的庞大市场需求。

5.2.2 机器视觉系统核心技术逐步被国人掌握

机器视觉领域的厂商包括设备提供商和设备集成商。要将机器视觉系统中多个部件整合在一起，能在自动化生产线上发挥作用，还需要一个系统集成过程。现场环境的适应性、安装调试是否到位、甚至使用人员的素质，都会影响到机器视觉产品最终的质量。因此，系统集成商与提供商一样重要。2000 年以前，国内系统集成商，主要以代理国外产品为主，自主知识产权的图像算法研究是一片空白，国内企业的技术水平与国际上有很大的差距，以至于之前出现国外视觉系统以高价位占领中国整个自动化行业的片面现象；到 2000 年，国内开始陆续出现机器视觉软件包，其性能和速度能与国外软件相媲美，甚至有些图像处理工具，在应用方面也大大超过了国外产品。

5.2.3 机器视觉在国内外应用现状

在国外^[7]，机器视觉的应用普及主要体现在半导体及电子行业，其中大概 40%-50% 都集中在半导体行业。例如，各类生产印刷电路板组成设备；电子封装技术与设备；丝印印刷设备；半导体集成电路制造设备；元器件成型设备；电子工模具等。即使原系统还在质量检测的各个方面，已经得到了广泛的应用，并且其产品在应用上占据举足轻重的地位。除此之外，机器视觉还用于其他各个领域和行业。

在中国，视觉技术的应用于 20 世纪 90 年代，因为行业本身就属于新兴的领域，再加之机器视觉产品技术的普及不够，导致以上各行业的应用几乎空白。到 21 世纪，大批海外从事视觉技术行业技术人员回国创业，视觉技术开始在自动化行业成熟应用，从华中科技大学在印刷在线检测设备与浮法玻璃缺陷在线检测设备研发的成功，打破了欧美在该行业的垄断地位。国内视觉技术也日益成熟，随着配套基础设备的日益完善，技术、资金的积累，各行各业对采用图像和机器视觉技术的工业自动化、智能化需求开始广泛出现，国内有关大专院校、研究所和企业近两年在图像和机器视觉技术领域进行了积极的探索和大胆尝试，逐步开始工业现场的应用，其主要应用于制药、印刷、包装等领域，真正高端的应用也正逐步发展。

5.3 中国机器视觉系统的发展趋势

5.3.1 对机器视觉的需求呈上升趋势

积极支援发展空间较大的部分在半导体和电子行业,而据我国相关数据显示,全球集成电路产业复苏迹象明显;与此同时,全球经济衰退使我国集成电路产业获得了市场优势、成本优势、人才回流等优势;国家加大对集成电路产业这一战略领域的规划力度,“信息化带动工业化”,走“新型工业化道路”,为集成电路产业带来了巨大的发展机遇,特别是高端产品和创新产品市场空间巨大,设计环节、国家战略领域,三, 3C 应用领域、传统产业内应用领域成为集成电路产业未来几年的重点投资领域。此外,中国已成为全球集成电路的一个重要需求市场。

中国的半导体和电子市场已初具规模,而如此强大的半导体产业将需要高质量的技术做后盾。同时对于产品的高质量、高集成度的需求将越来越高。恰巧,机器视觉将能帮助解决以上的问题,因此该行业将涉及世界最好的用武之地。

5.3.2 统一开放的标准是基于世界发展的原动力

目前国内有数十家机器视觉产品厂商,与国外机器视觉产品相比,国内产品最大的差距并不单纯是在技术上,更是在品牌和知识产权上。另一现状是,目前国内的机器视觉产品主要以代理国外品牌为主,以此来逐渐朝着自主研发产品的路线靠近,起步较晚。未来,机器视觉产品的好坏不能通过单一因素来衡量,应该逐渐按照国际化的统一标准判定。依靠封闭的技术难以促进整个行业的发展,只有形成统一而开放的标准才能让更多的厂商在相同的平台上开发产品,这也是促进中国及世界朝着国际化水平发展的原动力。

标准化将成为机器视觉发展的必然趋势。机器视觉是自动化的一部分,没有自动化就不会有集视觉,机器视觉软硬件产品正逐渐成为协作生产制造过程中不同阶段的核心系统,无论是用户还是软硬件供应商都是将机器视觉系统作为生产在线信息收集的工具,这就要求机器视觉系统大量采用“标准化技术”。

5.3.3 基于嵌入式的产品将取代板卡式产品

从产品本身看,机器视觉会越来越趋于依靠 PC 技术,并且与数据采集等其他控制与测量的集成会更加紧密。基于嵌入式的产品由于体积小、成本低、低功耗等特点,将逐步取代板卡式产品,而且随着计算机技术和微电子技术的迅速发展,嵌入式系统应用领域越来越广泛。另外,嵌入式操作系统绝大部分 C 语言为基础,因此使用 C 高级语言进行嵌入式系统开发是一项带有基础性的工作,使用高级语言的优点是提高工作效率,缩短开发周期,更主要的是开发出的产品可靠性高、可维护性好、便于不断完善和升级换代等。

5.3.4 一体化解决方案是机器视觉的必经之路

用于机器视觉是自动控制的一部分,机器是软硬件产品正逐步成为协作生产制造过程中不同阶段的核心系统,无论是用户还是硬件供货商都将机器视觉产品作为