



21世纪

高等学校精品规划教材

智能视感学

主 编 张秀彬

副主编 应俊豪



中国水利水电出版社
www.waterpub.com.cn



高等学校精品规划教材

智能视感学

主 编 张秀彬
副主编 应俊豪



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书从计算机视感及其信号处理的基本概念与基础理论出发,阐述基于图像信息的识别、理解和检测技术原理与方法。全书分为基础篇与应用篇两大部分,其中,基础篇系统地介绍了智能视感的基本原理、方法、关键技术及其算法;应用篇则由配合主要基础理论和方法的应用技术实例所组成。全书遵循理论知识与实用技术的紧密结合、数学方法与实用效果的相互印证等编写原则。本书涉及的教学内容,主要包括图像处理基础、摄像机数学模型、视感识别算法、视感检测原理、智能视感实用技术等。

本书可以作为检测与控制、自动化、计算机、机器人及人工智能等专业的高年级本科生和研究生的教材,同时也可作为专业技术人员的参考工具书。

图书在版编目(CIP)数据

智能视感学 / 张秀彬主编. -- 北京 : 中国水利水电出版社, 2010.3
21世纪高等学校精品规划教材
ISBN 978-7-5084-7291-1

I. ①智… II. ①张… III. ①数字图像处理—高等学校—教材 IV. ①TN911.73

中国版本图书馆CIP数据核字(2010)第037152号

书 名	21 世纪高等学校精品规划教材 智能视感学
作 者	主编 张秀彬 副主编 应俊豪
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路 1 号 D 座 100038) 网址: www.waterpub.com.cn E-mail: sales@waterpub.com.cn 电话: (010) 68367658 (营销中心)
经 售	北京科水图书销售中心 (零售) 电话: (010) 88383994、63202643 全国各地新华书店和相关出版物销售网点
排 版	中国水利水电出版社微机排版中心
印 刷	北京市兴怀印刷厂
规 格	184mm×260mm 16 开本 16.75 印张 397 千字
版 次	2010 年 3 月第 1 版 2010 年 3 月第 1 次印刷
印 数	0001—3000 册
定 价	32.00 元

凡购买我社图书,如有缺页、倒页、脱页的,本社营销中心负责调换

版权所有·侵权必究

前言

信息的获取、认识和理解是信息工程科学的核心内容。如何从图像中获取有关的信息已经成为当今世界热门的话题。

可以说，二维图像是蕴涵丰富自然信息的一种载体，它包含物体表面的纹理和粗糙度、物体的色彩特性、物体的边界形状、物体目标位置、物体颗粒大小分布和物体几何尺度等与人们生活、生产直接相关的可贵信息。

作者根据自己和博士、硕士生们的研究成果，以及多年从事本科生及研究生的教学经验与体会整理出本书，并将其定名为《智能视感学》。考虑到目前在该学科方向上尚缺乏一些浅显易懂、又能形成体系的简明教程，作者想做一次尝试，希望能用一种较为通俗和深入浅出的方式来阐述智能视感的一些深奥知识，对初学者能够起到入门和建立继续深造的起点之作用。

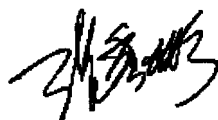
这是一本基于图像信息的非接触式传感理论与技术书。书中所阐述的内容涉及到图像识别、视差原理、计算几何原理、计算机图像图形学，乃至人类对自然界认识的诸多先验知识如何与视感检测相结合的方法和技术问题。因此是一本多学科交叉的较为前沿的大学研究型教材。

全书由张秀彬主编，应俊豪为副主编。非常感谢应俊豪博士提供的基础理论研究成果和协助组织完成诸多实用技术所做的艰苦工作，并为本书提供了丰富的例题和习题。

希望本书的出版能够给刚刚开始或者正在开展本领域技术研究的研究生和专业技术人员一定的帮助。不过总的来讲，还是希望本书能够抛砖引玉，获取诸位专家、学者对本书的批评和建议，以便再版时的修正和补充，为在“智能视感”新课程的建设中作出自己能够作出的贡献。

值此，作者还要特别向计长安、赵兴勇、肖杭、曾国辉、吴浩、何斌、欧阳武（工学博士），陆冬良、马丽（博士生），刘旭涛、王卫峰、雷传华、孙济宇、门蓬涛、孙志旻、吴炯、周丛嘉、胡朝晖、王益、孙剑、华逸伦、

高翔、李君峰、蒋玉春、胡朕、莫臻（工学硕士）等表示感谢，因为他们为本教材的撰写提供了诸多素材并参与了许多实验工作。



新中国 60 周年华诞于上海交通大学
zhangxb@sjtu.edu.cn

前言

基 础 篇

第 1 章 导论	1
1.1 概述	1
1.1.1 视感概念	1
1.1.2 视感技术发展史略	2
1.1.3 视感系统分类	4
1.2 视感技术硬件基础	5
1.2.1 图像传感	5
1.2.1.1 常用图像传感器	6
1.2.1.2 CCD 图像传感器的物理特性	7
1.2.1.3 摄像机技术特点	12
1.2.1.4 摄像机选用原则与设置方法	16
1.2.2 图像采集	18
1.2.2.1 图像采集卡类型	18
1.2.2.2 图像采集卡基本性能	20
1.2.3 视感系统对计算机的配置要求	21
1.2.3.1 硬件配置	21
1.2.3.2 数据通信接口	22
习题	24
第 2 章 图像处理基础	26
2.1 灰度图像的基本处理方法	26
2.1.1 空间域增强算法	26
2.1.1.1 空间域点运算	26
2.1.1.2 图像平滑	31
2.1.1.3 图像锐化	34
2.1.2 频率域增强算法	36

2.1.2.1	低通滤波增强算法——频率域平滑法	37
2.1.2.2	高通滤波增强算法——频率域锐化法	40
2.2	灰度图像边缘检测	43
2.2.1	门限化边缘检测	43
2.2.2	基于梯度的边缘检测	45
2.2.3	边缘检测 Laplacian 算子	47
2.2.4	Canny 边缘算子	49
2.2.5	数学形态学法	54
2.2.6	其他算法的简要介绍	56
2.3	图像二值化处理与图像分割方法	57
2.3.1	概述	57
2.3.2	基于直方图谷点门限的图像二值化方法	57
2.3.3	OTSU 算法	58
2.3.4	最小误差图像分割法	59
2.4	彩色图像增强	61
2.4.1	彩色空间及其变换	61
2.4.1.1	彩色空间描述	61
2.4.1.2	彩色空间变换	63
2.4.2	彩色图像色阶直方图均衡化	64
2.5	彩色图像的边缘检测	65
2.5.1	基于梯度极值的彩色图像边缘检测	65
2.5.2	彩色图像边缘检测实用方法	67
	习题	69
第 3 章	摄像机数学模型	71
3.1	图像空间几何变换	71
3.1.1	齐次坐标	71
3.1.2	正交变换和刚体变换	72
3.1.3	相似变换和仿射变换	73
3.1.4	透视变换	73
3.2	图像坐标及其变换	74
3.2.1	图像坐标系	75
3.2.2	图像坐标变换	76
3.2.2.1	摄像机坐标线性变换	76
3.2.2.2	图像畸变描述	77
3.3	摄像机参数标定的常用方法	79
3.3.1	分步标定法	80

3.3.2	基于多个自由平面的标定算法	81
3.3.3	非线性畸变参数标定法	84
习题		85
第4章	视感识别算法	87
4.1	图像特征提取与识别算法	87
4.1.1	决策理论方法	87
4.1.1.1	线性判别函数	88
4.1.1.2	最小距离分类函数	90
4.1.1.3	最近邻域分类函数	91
4.1.1.4	非线性判别函数	92
4.1.2	统计分类法	93
4.1.2.1	贝叶斯法则	93
4.1.2.2	最小错误率的贝叶斯分类法	94
4.1.2.3	最小风险的贝叶斯分类法	95
4.1.3	图像识别过程特征分类判别相似度	95
4.2	主分量分析	96
4.2.1	主分量分析原理	97
4.2.1.1	K-L 变换	97
4.2.1.2	K-L 变换性质	97
4.2.2	核主分量分析	99
4.2.2.1	核函数	99
4.2.2.2	核函数类型	100
4.2.2.3	核主分量	100
4.2.3	基于主分量分析的图像识别	102
4.3	支持向量机	105
4.3.1	统计学习理论的主要内容	106
4.3.1.1	学习过程一致性条件	106
4.3.1.2	VC 维理论	107
4.3.2	分类支持向量机	108
4.3.2.1	支持向量机的体系结构	109
4.3.2.2	最优分类超平面	109
4.3.2.3	最优分类超平面的求取方法	111
4.3.3	非线性回归问题求解	113
4.3.3.1	ϵ -不敏感损失函数	113
4.3.3.2	用于非线性回归的支持向量机	114
4.3.4	支持向量机的算法	116
4.3.4.1	处理大数据集的 SVM 训练算法	116

4.3.4.2	序贯最小优化算法	118
4.3.5	基于支持向量机的图像特征识别	121
4.4	不变矩与归一化转动惯量	122
4.4.1	矩理论	122
4.4.1.1	空间矩和中心矩	123
4.4.1.2	不变矩	123
4.4.1.3	矩的应用	124
4.4.2	归一化转动惯量 (NMI)	125
4.4.2.1	NMI 特征定义	125
4.4.2.2	NMI 识别算法	126
4.4.2.3	基于灰度权重的 NMI 快速匹配	127
4.5	模板匹配及其相似度	131
4.5.1	模板匹配的空间域描述	131
4.5.1.1	第一类相似度——误差	132
4.5.1.2	第二类相似度——互相关	133
4.5.1.3	空间域模板匹配的进一步讨论	136
4.5.2	模板匹配的频域描述	136
4.5.2.1	第三类相似度——相位频谱方法	136
4.5.2.2	傅里叶—梅林相似度	142
4.6	基于颜色特征的目标识别	144
4.6.1	图像色度学处理	144
4.6.1.1	偏色处理	144
4.6.1.2	彩色空间转换	145
4.6.1.3	计算彩色图像的统计特征	145
4.6.2	建立颜色特征池	146
4.6.2.1	颜色特征池基本概念及其集合特性	146
4.6.2.2	颜色特征池的初始化	146
4.6.2.3	基于颜色特征池的目标景象切割	147
4.6.3	基于颜色的目标识别	148
4.7	图像模糊识别方法	149
4.7.1	模糊内容特征及其模糊相似度	149
4.7.1.1	模糊内容特征表示方法	149
4.7.1.2	基于模糊内容特征的相似度计算	150
4.7.2	模糊结构提取	151
4.7.2.1	图像二叉树切割	151
4.7.2.2	模糊特征提取	152
4.7.2.3	基于模糊相似度的节点匹配	153

4.7.2.4	基于模糊二叉树结构的相似度	154
4.7.3	图像匹配的模糊合成决策	155
4.7.3.1	合成决策函数——模糊合成相似度	155
4.7.3.2	相似度算法的运算简化	156
4.7.3.3	匹配过程分析	156
习题		159
第5章	视感检测原理	162
5.1	单视几何与单目视感检测原理	162
5.1.1	单视坐标系	162
5.1.2	单视检测基本算法	162
5.1.3	基于单视几何的工程方法	163
5.2	双目视感检测原理	166
5.2.1	双视几何与双目视感检测	166
5.2.1.1	双目视差测距原理	166
5.2.1.2	双视几何变换	167
5.2.1.3	双视投影与单应矩阵	168
5.2.2	外极几何学原理	169
5.2.2.1	外极线几何约束	170
5.2.2.2	外极平面本征矩阵	170
5.2.2.3	外极线解析方程	172
5.2.3	空间坐标确定方法	173
5.2.4	双目视感系统摄像机参数标定	175
5.2.4.1	“棋盘标定法”在双目视感检测中的应用	176
5.2.4.2	弱标定外极几何	176
5.2.4.3	左、右视感对空间物点坐标识别的误差评估	182
5.3	多目视感检测理论基础	183
5.3.1	张量几何学原理	184
5.3.1.1	仿射坐标系	184
5.3.1.2	n 维空间的仿射坐标	186
5.3.2	三视张量的几何属性	187
5.3.2.1	三视图像中的点线约束	189
5.3.2.2	单应变换与基本矩阵	190
5.3.3	三视张量的运算	191
5.3.3.1	从同名线计算三视张量	191
5.3.3.2	从匹配点计算三视张量	192
5.3.4	三视张量约束特征点的匹配	193
5.3.4.1	三视张量已知条件下的特征点匹配算法	193

5.3.4.2 三视张量未知条件下的特征点匹配算法	194
5.3.5 三视张量约束三视图图像特征线的匹配	196
5.3.5.1 三视张量已知条件下特征线的匹配算法	196
5.3.5.2 三视张量未知条件下特征线的匹配算法	196
习题	200

应 用 篇

第6章 智能视感实用技术	202
6.1 桥梁限载自动监控系统及其方法	202
6.1.1 系统基本组成	202
6.1.2 系统算法	204
6.1.2.1 主运算程序	204
6.1.2.2 视感检测算法	204
6.2 钢坯号智能识别系统	206
6.2.1 系统主控程序	207
6.2.2 识别算法	207
6.2.2.1 模板匹配算法	207
6.2.2.2 人工神经网络识别算法	210
6.3 基于图像信息的纸币清分技术	211
6.3.1 纸币图像预处理	212
6.3.2 纸币新旧区分	212
6.3.3 纸币面额与面向区分	213
6.3.4 纸币成色检测	214
6.4 车辆智能防撞技术	216
6.4.1 基本硬件配置	216
6.4.2 道路障碍物识别算法	217
6.4.3 车辆防撞行人智能算法	219
6.4.3.1 原理	219
6.4.3.2 阈值求取	220
6.4.3.3 相似度度量	221
6.5 交通信号灯视感智能控制	224
6.5.1 概述	224
6.5.2 交通信号灯智能控制的核心算法	224
6.5.2.1 建立坐标与车道线方程	224
6.5.2.2 车流图像处理	226
6.5.2.3 车辆排队长度检测与通行时间计算	227
习题	229

附 录

I 最小二乘及在视感检测中的常用算法	231
I.1 算法基本思想	231
I.2 视感检测中常用的最小二乘算法	232
I.2.1 线性方程组的最小二乘法	232
I.2.1.1 正则方程和伪逆	233
I.2.1.2 齐次方程最小二乘解和特征值	233
I.2.1.3 广义约束和广义特征值	234
I.2.2 非线性齐次方程组的最小二乘解	234
I.2.2.1 牛顿法 (Newton 法)	235
I.2.2.2 高斯—牛顿法 (Gauss - Newton 法)	236
I.2.2.3 Levenberg - Marquardt 方法	236
II 贝叶斯决策理论与方法	237
II.1 概述	237
II.2 贝叶斯分类决策模型	237
II.2.1 贝叶斯最小错误概率分类模型	237
II.2.1.1 两类模式集问题	237
II.2.1.2 多类问题	238
II.2.2 贝叶斯最小风险分类决策	238
III 统计学习与 VC 维理论	240
III.1 边界理论与 VC 维原理	240
III.2 推广能力边界	241
III.3 结构风险最小化归纳原理	241
IV 关于约束非线性规划问题最优解条件	243
IV.1 Kuhn - Tucker 条件	243
IV.1.1 Gordan 引理	243
IV.1.2 Fritz John 定理	243
IV.1.3 Kuhn - Tucker 条件证明	244
IV.2 Karush - Kuhn - Tucker 条件	246
术语索引	247
参考文献	251

基 础 篇

第 1 章 导 论

1.1 概 述

智能视感 (Intellectual Visual Perception, IVP) 体现一种非接触式传感概念, 是一种基于全新的科学概念来实现对景物图像的认识、理解的方法与技术。从总的技术范畴来讲, 视感技术属于检测技术领域, 所以在工程应用中往往直接称其为视感检测。

1.1.1 视感概念

从 20 世纪 60 年代中期开始, 图像传感及其信号处理的理论与技术已经引起世界学术界的高度重视, 因为这是一项新兴技术, 也是一门从图像所富含的信息中感知、认识与理解自然现象及其客观规律的科学。

虽然现实世界中的物体是三维的 (乃至多维的), 但是通过图像传感器所获得的景物图像却是二维的。如何从二维图像中获取三维信息, 进而感知三维世界, 正是计算机科学有关图像信息特征提取、信号分析与原形再现的研究内容, 其间存在着无穷无尽的知识需要人们去探索和研究。

图像传感及其信号处理, 与人的眼睛和大脑对外界景物图像的感知过程极为相似。作者基于以下三大方面的思考提出“视感”概念。

首先, 计算机视觉也好, 机器视觉也好, 都是沿用了人类对人体生命科学尚不发达时期所生成的生理器官术语“视觉”而形成, 显然缺乏深层次的物理和生理科学的内涵; 尽管“视感”与“视觉”仅有一字之差, 但是前者却已充分显现出人脑器官对信息获取所起到的重要且不可替代的作用。

其次, “视感”概念的提出, 将智能机器对外部景物进行认识 and 理解的整个运动过程分解为图像传感和视图感知两大部分。图像传感相当于人的眼睛及其视网膜上的杆状和椎状细胞, 直接与外界通过光线进行联系, 从而对周围环境作出最直接、最迅速的反应; 视图感知相当于人脑中枢, 由视网膜细胞产生信号传送至大脑的视感知区, 来读取、分析光的信号, 从而知道物体的远近、大小、颜色、形状等具体信息, 还可以将成像进行联想、翻转等。不难想见, 仅凭视图感知器, 而没有图像传感器就无法获取图像信息, 这好比一个人的大脑中视觉中枢完好, 但是没有眼睛及其视网膜, 或者眼睛与视网膜已经损坏, 是无法感知周围的景物图像及其信息; 同样, 仅凭图像传感器也是无法完成机器对外界景物形象的感知, 这好比一个人即使眼睛及其视网膜完好, 但是大脑的视觉中枢不健全或者视感神经缺损, 仍然致盲。换句话说, 要真正实现“基于机器视觉”的非接触式检测必须通

过图像传感器和视图感知器的完美结合才能得以完成，这就是“视感”二字的由来。

第三，从“视感”的角度出发来阐述智能机器对图像信息的认识和理解过程，能够科学地将“图像分析”与“机器视觉”有机地融合起来。这是因为，图像分析和机器视觉是相关而又不同的两个研究领域。从某种意义上说，图像分析是机器视觉的一部分；然而，从另一种意义上说，图像分析又是一门更宽广的学科。

如图 1-1 所示，这是一帧通过显微镜采集的哺乳动物大脑运作图像，它可以利用显微设备，通过接收、分析、传输和显示等功能手段对哺乳动物大脑运作的微观运动过程进行研究来探索哺乳动物脑组织细胞的临床病理学机理和表征。换句话说，“机器视觉”显然不包含如此的技术功能，而“机器视感”则已将自身的研究功能扩展至“图像分析”的领域。

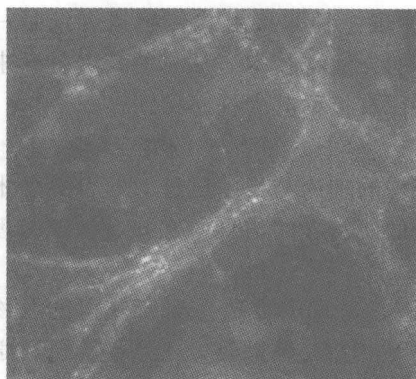


图 1-1 哺乳动物大脑运作显微图像

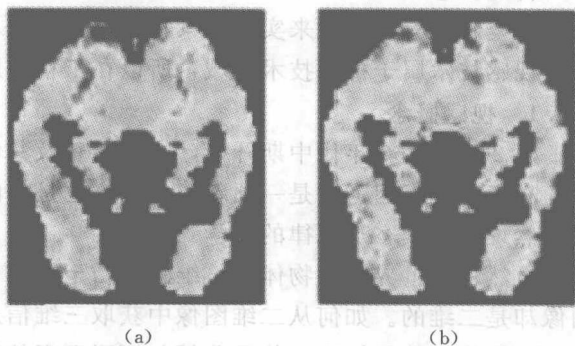


图 1-2 人脑思维显微图像

(a) 想象时；(b) 观察飞机时

再参见图 1-2，这是人脑在想象和观察某个事物（如飞机）时，通过 X 射线计算机横断层扫描（Computed Tomography, CT）所获得的两帧断层扫描图像，其中，图 1-2 (a)、(b) 分别为某人在想象和观察飞机时的大脑断层扫描图像。针对这两帧图像，依靠“机器视觉”的概念显然是无法“读懂”其中的“深层含义”，只有采用“图像分析”算法才能理解两者的区别之所在，这也就是“视感”所要蕴涵的技术内容。

1.1.2 视感技术发展史略

视感技术的开创性工作从 20 世纪 60 年代中期就已经开始，不过从开始到现在，一直沿用着视觉（Vision）术语及其概念来阐述其理论和方法。其中，Lawrence G. Roberts (1937~，美国科学家) 于 1965 年发表了著名的论文《Machine Perception of Three - Dimensional Solids》，他在该论文中提出的 Roberts 积木理论是该时期的一个典型代表。进入 20 世纪 70 年代中期，David Courtney Marr (1945~1980，英国神经学家，视感处理和计算神经学的奠基人)，在他所著的《Vision: Computational Investigation into the Human Representation and Processing of Visual Information》一书中提出了一整套视觉计算的理论用来描述视觉过程，其核心内容是一种从计算机成像来恢复物体三维形状的方法。就迄今为止较有代表性的学术发展成果来看，可以概括地将该领域的发展过程分为三个阶

段：早期视觉技术阶段、二维半简图（本征图像）阶段和三维模型阶段。

早期视觉技术阶段，着眼于景物表面的物理特性，如距离、表面方向颜色、纹理等。

二维半简图（本征图像）阶段，特点在于以观察者为中心的坐标系来描述物体表面的各种特性，根据这些描述，重建物体边界，按表面和体积分割景物。但是，在该阶段只能得到可见表面的描述，而得不到遮挡表面的描述，故称二维半简图描述。

三维模型阶段，也是当前图像信息处理的最高层次，是用二维半简图中得到的表面信息建立适用于图像识别的三维形状与形态描述。

上述三个阶段，尽管所建立的计算方法在数学上是可行的，但由于各种干扰以及逆成像问题的存在，使得识别与测量精度难以达到技术要求。习惯上称此类视觉测量方法为“视觉被动认知法”。

进入20世纪80年代中后期，立体视感技术的研究有了较大的发展，已经运用空间几何的方法以及物理知识来研究立体视觉问题。这一时期引入主动视觉的研究方法，并采用距离传感（如激光测距、微波测距、超声波测距等）与视图传感相融合的技术。由于这种研究方法可直接取得深度图或通过移动获取深度图，因而可以解决“视觉被动认知法”所难以克服的技术缺陷。

视感技术经过三四十年的研究，已经取得了巨大的成功和进展。从当前的发展水平来看，视感技术的发展趋势可以归纳如下：

第一，以全面的观点将图像传感与其他形式的传感技术联系起来，即多传感信息融合，充分挖掘景物图像的内在信息。

第二，由传统的两视点向多视点，并从完全静态的立体视感向动态立体视感发展，通过增加图像信息输入来降低视感计算的难度。

第三，直接向智能化方向发展，开展基于知识、模型和规则的立体视感方法研究。

第四，探索并研发面向对象与用户的视感识别与检测实用技术。

必须指出，与人类自身的视感功能相比，目前的机器视感理论与技术还很不健全，还处在一个十分不成熟的阶段，研究工作仍然面临着一系列技术难点：

（1）机器视感系统在分析和理解图像时都需要立体匹配，而这一技术目前还很不完善。当存在灰度失真、几何失真，以及透视、旋转、缩放等失真和存在噪声的情况下，识别、匹配图像的过程是一项颇为困难的工作，因为目前还没有具备一个完全可靠的、通用的识别与匹配方法，尤其是在失真和噪声比较严重的情况下，识别与匹配过程显得更为困难。与此同时，现实世界中物体间的相互遮挡以及物体本身各部分间的遮挡更增加了图像识别与匹配的难度。

（2）人类借助日积月累的极为丰富的理论和先验知识对图像具有惊人的分析理解能力，但是，人是如何精确获取视觉知识的，至今尚未研究清楚。由于视觉采像具有很高的信息冗余度，从高冗余度的图像信息中选取哪些特征可以解决立体理解和复现问题，目前尚无一种通用的方法。

（3）立体视感所面对的图像信息计算量非常庞大，一个实用的技术系统对时间的要求是有一定限制的。在现阶段集成电路和特殊功能部件的发展还不能满足实时性要求的情况下，要想将一种视感系统完全变为实用的通用系统仍然存在相当大的技术难度。

1.1.3 视感系统分类

根据视感系统所采用的图像传感器数量可以将其分为单目视感系统、双目视感系统和多目视感系统，即由单摄像机、双摄像机和多摄像机构成的图像采集与处理系统，简称为单目视感系统、双目视感系统和多目视感系统。

单目视感系统是采用单个摄像机作为唯一的图像传感器来承担对景物的图像采集。所谓单视检测（单目视感检测的简称），其系统结构较为简单，仅需选用一个摄像机、单输入通道图像采集卡和一台计算机就可构成一个完整的视感检测系统。对于单摄像机数学模型（即单目图像传感器数学模型）研究也是后续多目图像传感器模型研究的基础。

许多场合下仅用单个图像传感器采集景物图像即可实现对一维移动物体的识别与测量，如生产线上物件的移动测量、目标运动状态的识别与测量等。实现单视检测的理论基础就是单目视觉几何学，它直接运用计算几何学中的仿射几何和射影几何理论，利用单幅图像所具有的场景信息来实现场景中的形态特征识别与几何尺度测量。因其所需的只是场景二维图像，所以该方法较为灵活与方便，尤其适合于交通事故现场和建筑物等方面的检测，多用于外形规则、测量范围较小和系统已知信息较多的情形。

利用单幅图像对场景进行识别与检测，已经日益引起人们的关注，但是，目前国内外在此方面还缺乏系统的研究，这是因为单幅图像所包含的信息量相对有限，实现检测过程在很大程度上需要依靠某些先验知识，如对规则被测物体几何形态的事前了解，对不规则被测景物背面形象先验知识的事先掌握等。单目视感系统多用于目标识别与跟踪、物体位移偏差检测等技术场合。

双目视感系统是直接模拟人类视感结构的一种方法。机器双目视感是用两台性能相同、位置相对固定的图像传感器，获取同一景物的两幅图像，通过计算空间点在两幅图像中的“视差”来确定场景的深度信息，进而构建场景的三维结构。

在双目视感系统的硬件结构中，通常采用两个摄像机作为视感信号的采集设备，通过双输入通道图像采集卡与计算机连接。在一个典型的双目视感系统中，两个摄像机光心沿基线放置，其光轴互成一定角度，空间点分别在左右两个摄像机成像平面上投影成像。成像点的间距即称为“视差”。由视差计算深度是非常容易的一件事，但视差本身的计算是立体视感中最困难的一步工作，它涉及到模型分析、摄像机标定、图像处理、特征选取及特征匹配等过程。双目视感最大的优势在于，能够通过外极线约束寻找景物在双目视图中的公共特征点，因此能够进一步通过坐标变换计算来获取被测景物中对应特征点的三维坐标。

双目视感系统主要应用于机器人导航、微操作系统的参数检测、三维测量和虚拟现实环境构建等技术领域。双目视感检测又简称为双视检测。

多目视感系统是对双目视感系统的一种拓展。所谓多目视感系统，就是采用多个摄像机设置于多个视点，或者由一个摄像机从多个视点观测三维景物的视感系统。对多目视感系统所采集到的景物图像进行感知、识别和理解的技术被称为多目视感系统技术。

尽管多目视感系统的硬件配置和采像方式要比双目视感系统复杂，但是，由于多目视感所获取的信息明显要比双目视感丰富得多，因此，多目视感系统用于立体检测能够弥补双目视感的技术缺陷是显而易见的。实际上，人们肉眼对物体的多视角观察就是多目视感

系统的一个生动实例，昆虫进化出的复眼具有多目视感特性正是苍天杰出造化的一项特例。

多目视感系统的主要理论基础是多视几何学。多视几何是近十多年来在计算机视感领域发展起来的一个新的热点研究方向，与传统视觉测量在很多方面有着极大的相似性，但是与传统视觉测量理论并不完全一致，两者存在很大的差异，例如，张量代数在多目视感系统中的应用为多视几何提供了一种十分便捷的表达和运算手段。如何将多视几何理论引入传统视觉测量，建立起一个新的学科发展方向，运用多视几何理论来弥补传统视觉测量理论的不足与局限性，这些工作具有十分重大的意义。

简而言之，视感是利用单目视感系统、双目视感系统或多目视感系统通过计算机（信号处理器）对所采集到的二维图像进行三维感知的一种技术，是用于对物体位置、运动速度等的识别与理解的方法。

图像传感器和信号处理器是视感检测系统的硬件基础，而视感检测系统的核心是计算软件。视感系统的软件包括计算机操作系统及其应用软件、图像处理算法软件、控制软件等。其中，图像处理算法软件是视感检测系统中最为关键的软件，因为它反映出对不同被测对象图像特征检测的核心思想方法（数学模型）。图 1-3 所示系视感系统的信号基本流程图。



图 1-3 视感系统的信号基本流程图

图 1-3 中所述图像预处理一般包括图像增强、滤波和背景淡化等；主运算程序则要面向处理对象来确定，如面向体积测量问题，一般包括透视变换、畸变矫正、坐标数据变换、对应点的寻找和匹配、物点三维坐标确定、体积计算等；数据输出则表示对图像识别与理解的结论性输出，包括显示目标参数计算结果和控制指令输出等。

实际上，图像处理算法的涉及范围十分广阔，根据应用目的的不同，可包括摄像机标定算法、图像输入处理、图像滤波、边缘检测、特征提取、图像匹配、深度识别、深度图生成、与其他类型传感信号的融合和先验知识的领会贯通等。

1.2 视感技术硬件基础

一般来说，图像传感器即摄像机（或称摄像头），实施对景物图像的采集；图像采集卡承担着对摄像机所采集图像的前置处理任务，是图像传感与主处理器之间的连接“桥梁”；计算机（即图像处理器）是图像处理、识别与理解的主运算器，是视感检测技术核心算法软件的载体。

1.2.1 图像传感

对于固体图像传感器来讲，当前主要有三大类型：CCD 图像传感器（Charge Coupled Device, CCD），又称电荷耦合图像传感器；CMOS 图像传感器（Complementary Metal-oxidesemiconductor, CMOS），又称自扫描光电二极管阵列（Self Scanned Photodiode