

计算机视觉——绪论

2023年春季

桑 农



教学安排

- 教师：桑农，王岳环
 - 办公室：科技楼1001、1003
 - Email: nsang@hust.edu.cn,
yuehwang@hust.edu.cn
- 上课时间与上课地点
 - 见课表



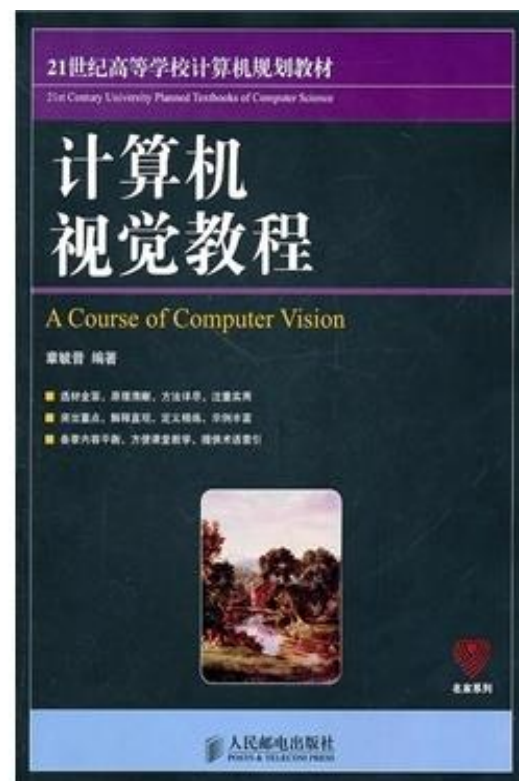
教学安排

- 教学方式
 - 课堂教学
- 考核方式
 - 平时作业（30%）
 - 每周一 交前一周作业
 - 笔试考试（70%）



教学安排

- 教材
 - 章毓晋，计算机视觉教程（第一版）
 - 课件
- 课件、资料下载
 - QQ群





第1章 绪论

1.1 计算机视觉

1.2 图像基础

1.3 像素间联系



1.1 计算机视觉

1.1.1 视觉

1.1.2 计算机视觉概述

1.1.3 相关学科

1.1.4 应用领域



1.1.1 视觉

视觉是人类观察世界、认知世界的重要功能手段。人类从外界获得的信息约有75%来自视觉系统

视觉进一步可分为视感觉和视知觉

- 视感觉主要与成像相关
- 视知觉主要与成像后得到的图像或图像序列（视频）的处理、分析与理解相关



1.1.2 计算机视觉概述

计算机视觉是指用计算机实现人类的视觉功能

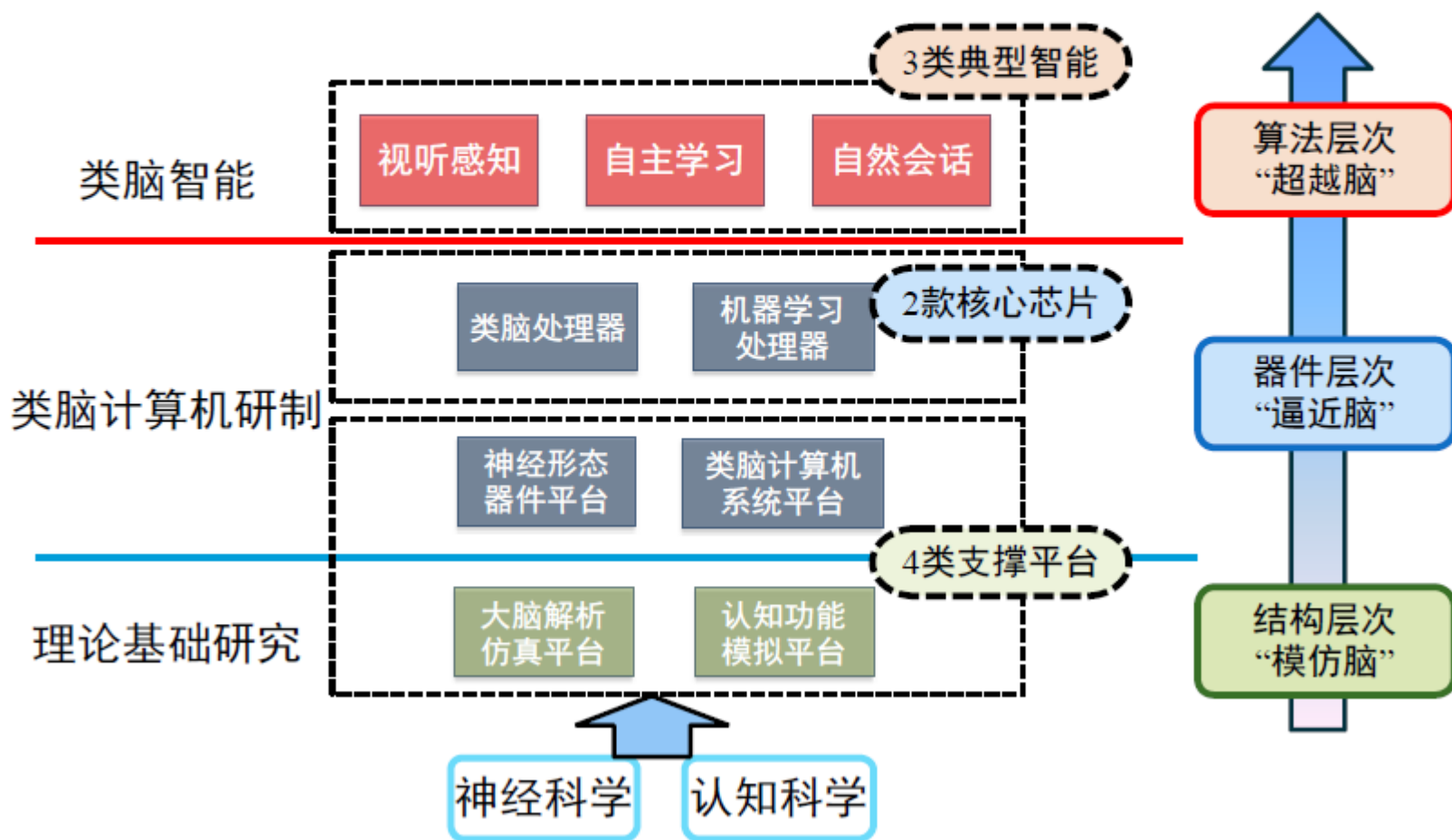
- 计算机视觉的研究方法目前主要有两种：
仿生的方法：参照人类视觉系统的结构原理
工程的方法：实现系统的功能
- 计算机视觉的主要研究目标可归纳成两个：
建立计算机视觉系统来完成各种视觉任务
加深对人脑视觉机理的掌握和理解



1.1.2 计算机视觉概述



“脑认知与类脑计算”任务规划





1.1.2 计算机视觉概述

计算机视觉是指用计算机实现人类的视觉功能

- 计算机视觉的研究方法目前主要有两种：
仿生学的方法：参照人类视觉系统的结构原理
工程的方法：实现系统的功能
- 计算机视觉的主要研究目标可归纳成两个：
建立计算机视觉系统来完成各种视觉任务
加深对人脑视觉机理的掌握和理解



1.1.3 相关学科

- (1) 图像理解：与计算机视觉有相同的目标
- (2) 机器视觉：更关注通过视觉传感器获取环境的图像，构建具有视觉感知功能的系统以及实现检测和辨识物体的算法
- (3) 模式识别：图像就是模式的一种
- (4) 人工智能：视觉功能是人类智能的体现
- (5) 计算机图形学：计算机视觉的反/逆问题



1.1.4

应用领域

(1) 工业视觉



■PCB板质量检测



■印刷质量检测



■玻璃瓶质量检测



■药品质量检测



■棉花异常纤维检测



■汽车零部件质量检测



1.1.4 应用领域

(2) 人机交互



手势遥控.avi



切水果.avi



1.1.4 应用领域

(3) 安全监控



入侵.avi



偷盗.avi



火焰检测.avi

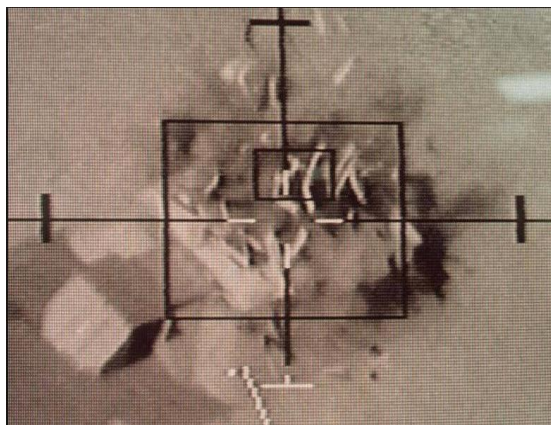


烟雾检测.avi



1.1.4 应用领域

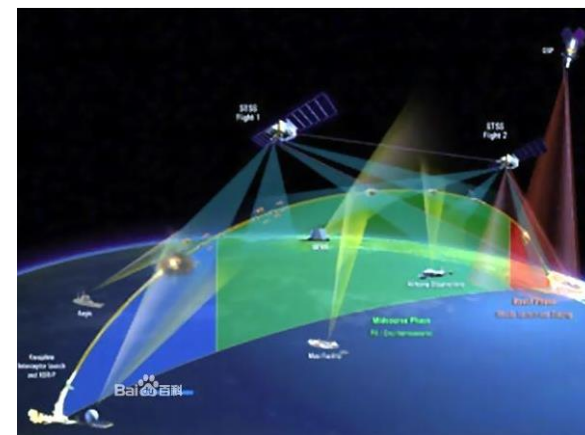
(4) 军事公安



■ 精确制导



■ 侦察



■ 预警



1.1.4 应用领域

(5) 遥感测绘



95年



98年



95年+98年

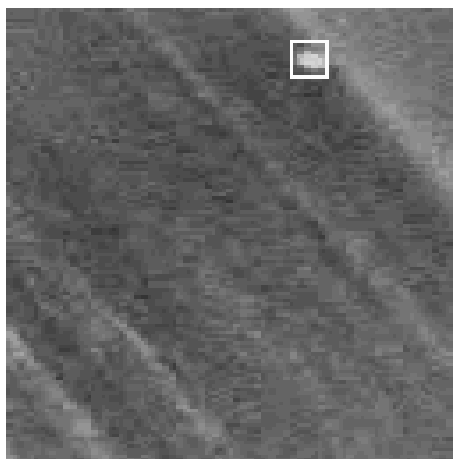


1.1.4 应用领域

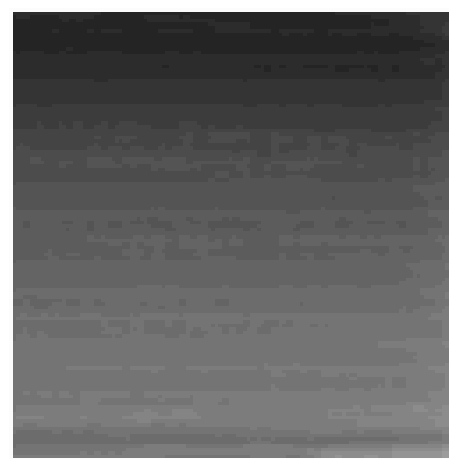
(6) 视觉导航



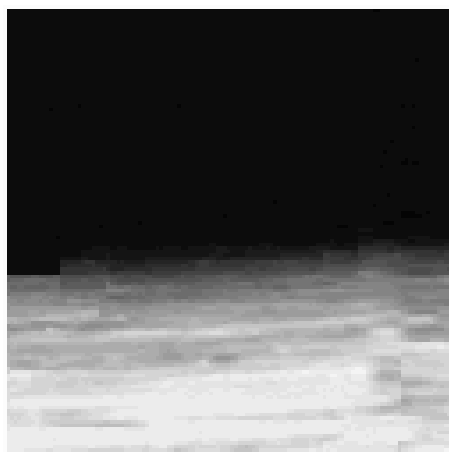
飞机检测.avi



舰船检测.avi



桥梁检测.avi



电厂检测.avi

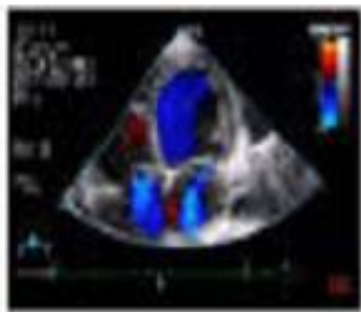


建筑物检测.avi



1.1.4 应用领域

(7) 生物医学





1.1.4 应用领域

(7) 生物医学



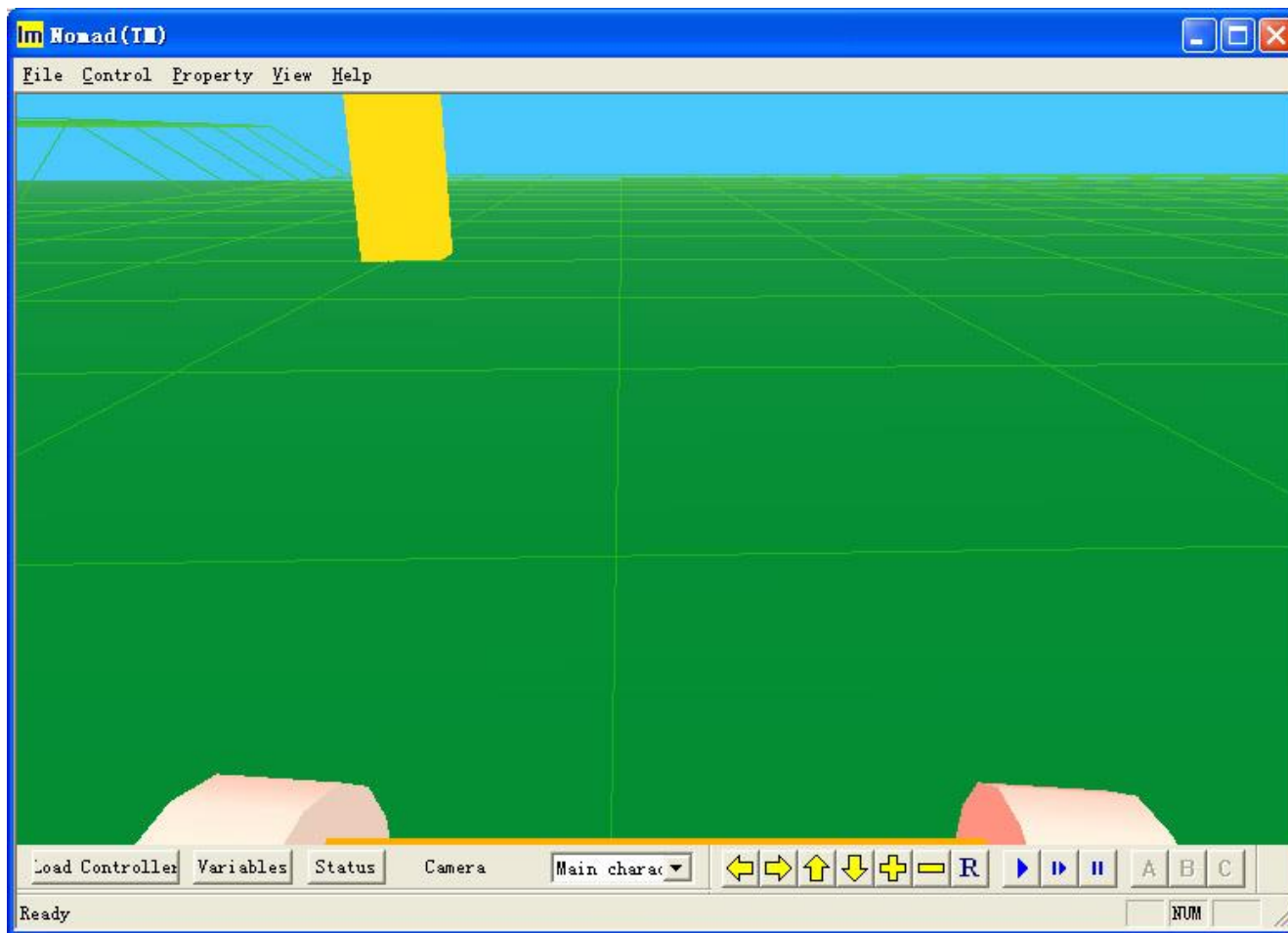
外科医生控制台 床旁机械臂系统 成像系统

达芬奇机器人



1.1.4 应用领域

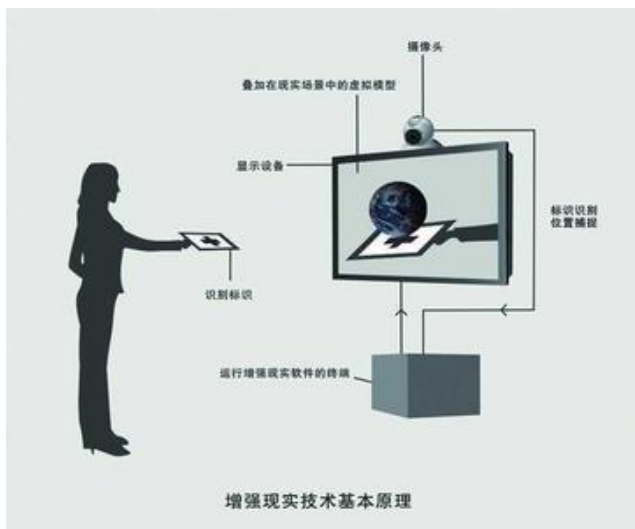
(8) 虚拟现实（Virtual Reality, VR）





1.1.4 应用领域

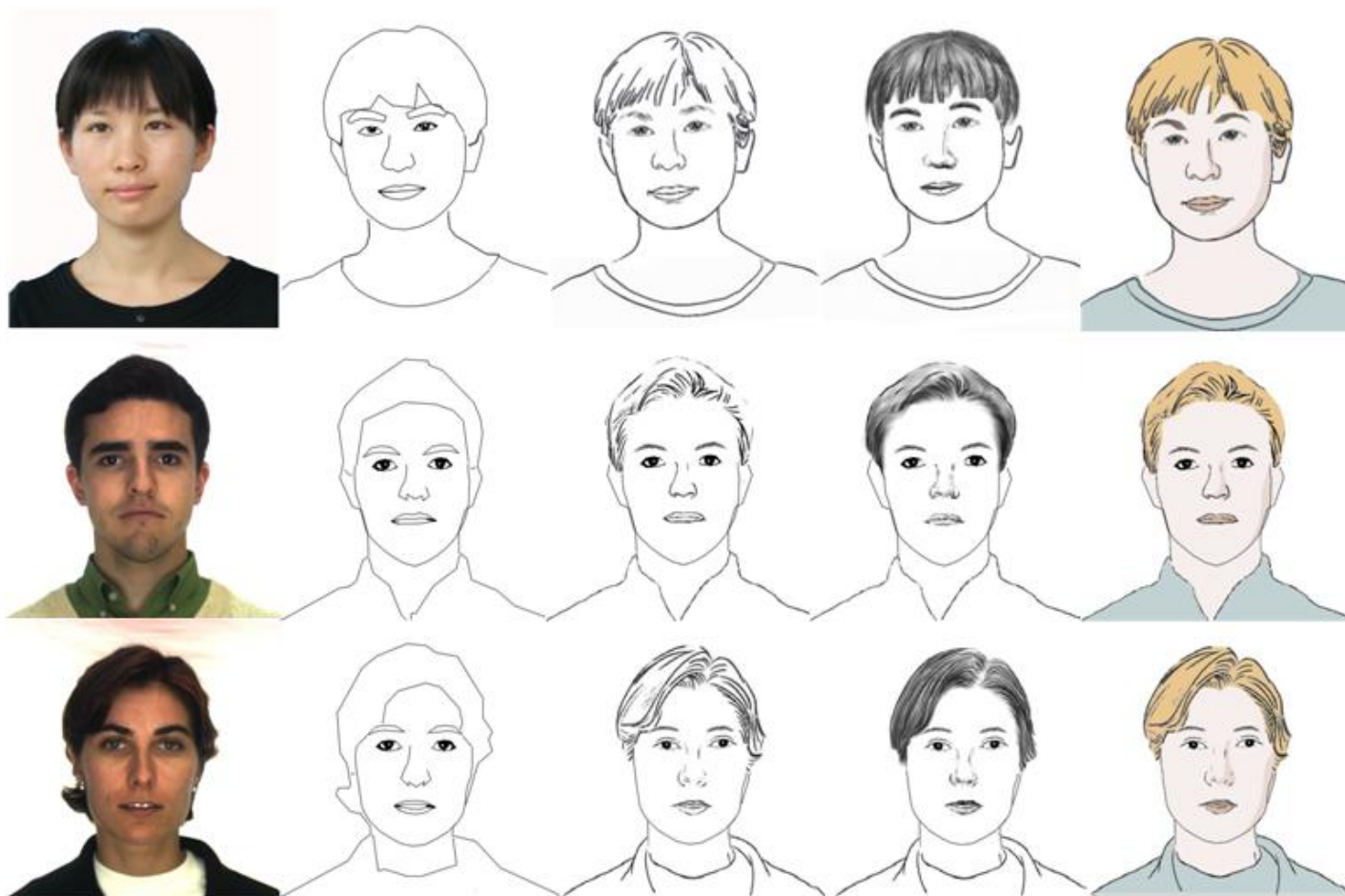
(9) 增强现实 (Augmented Reality, AR)





1.1.4 应用领域

(10) 数字艺术



■ 原始图像

检测

匹配

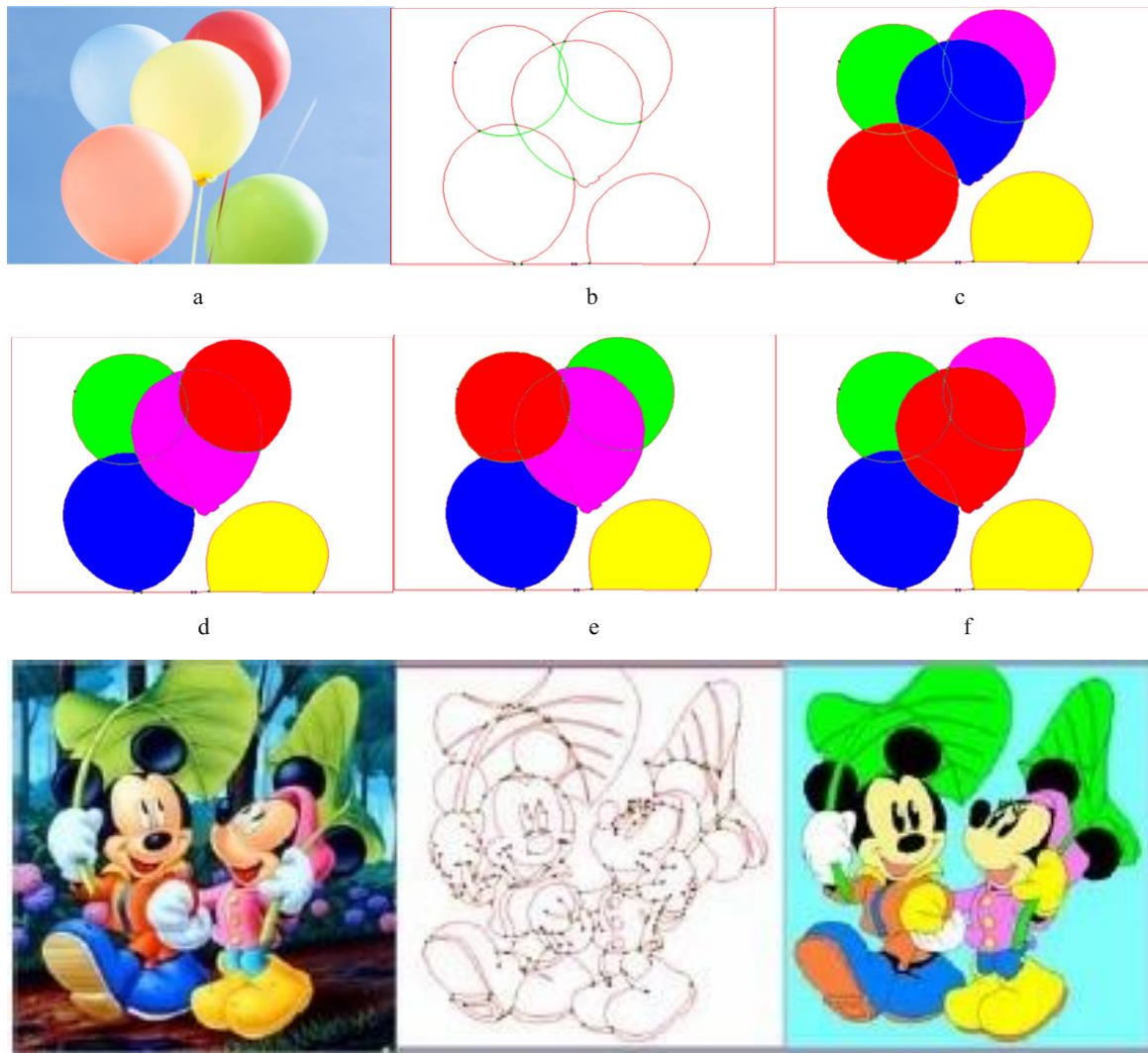
渲染

着色



1.1.4 应用领域

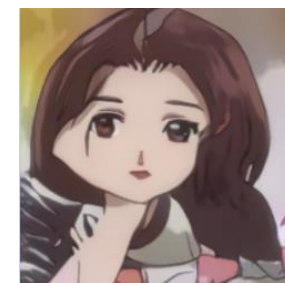
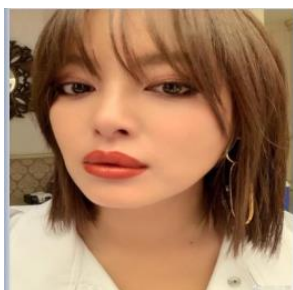
(10) 数字艺术





1.1.4 应用领域

(10) 数字艺术





1.1.4 应用领域

(11) 图像自动解释



■ slide by Fei Fei, Fergus & Torralba



1.1.4 应用领域

(12) 视频摘要



原始视频.avi



低密度摘要视频.avi



中密度摘要视频.avi



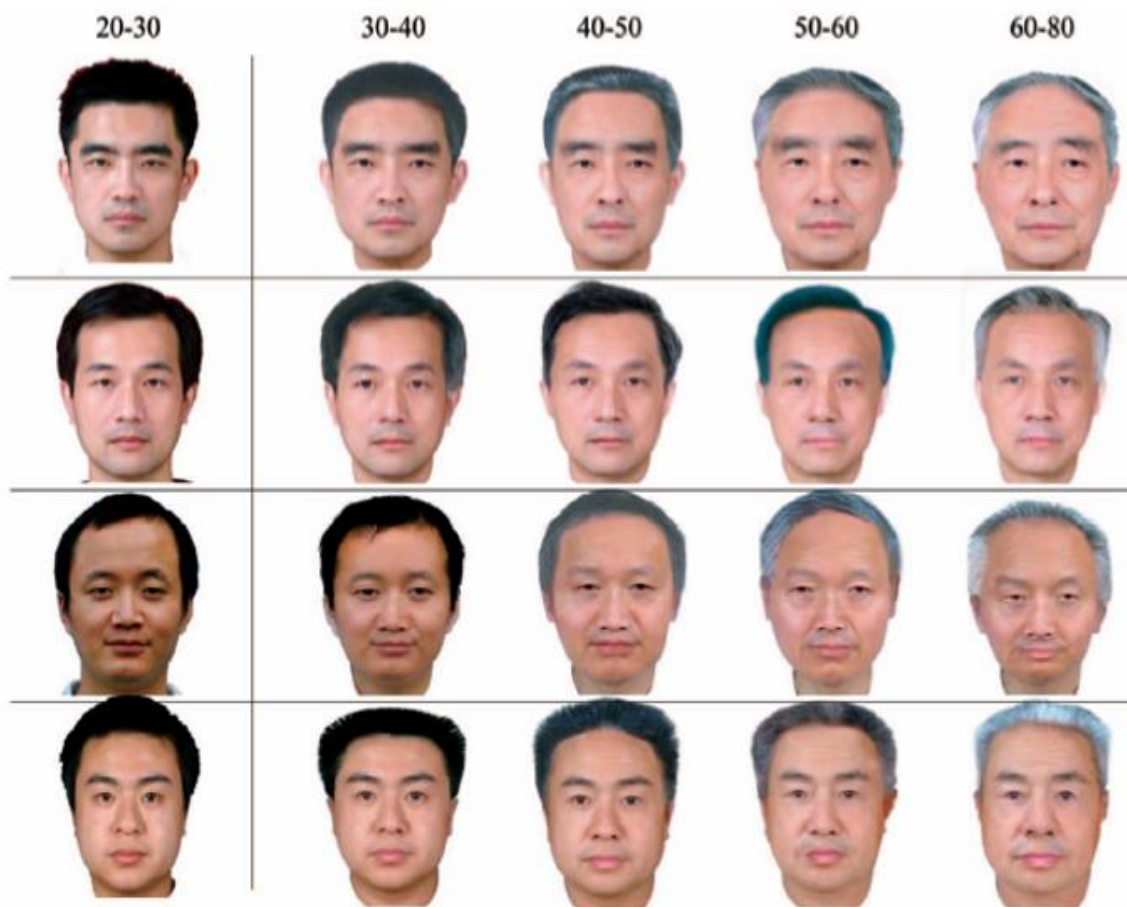
高密度摘要视频.avi



1.1.4

应用领域

(13) 其他



↑ 原照片

油画→



↑ 原照片

油画→





1.1.4

应用领域

(13) 其他

- 百度图片搜索
- Google
- 人脸、指纹识别（认证）
- 文档OCR
-



1.1.4 应用领域

(14) 对人类视觉系统和机理，以及人脑心理和生理的研究等



1.1.4 应用领域

(15) 应用实例——火焰监测

火焰检测演示软件

2009-08-03 18:00:01

平台顶层

打开视频

视频

配准区域

视频处理

结果显示

有无火焰	时间	火焰高度	报警等级
有	14:41:43	19 米	一级
有	14:41:44	19 米	一级
有	14:41:45	19 米	一级
有	14:41:46	20 米	一级
有	14:41:46	20 米	一级
有	14:41:47	20 米	一级
有	14:41:48	21 米	一级
有	14:41:49	19 米	一级
有	14:41:50	18 米	一级
有	14:41:50	19 米	一级
有	14:41:51	19 米	一级
有	14:41:52	19 米	一级
有	14:41:52	20 米	一级
有	14:41:53	20 米	一级

定标及等级设定

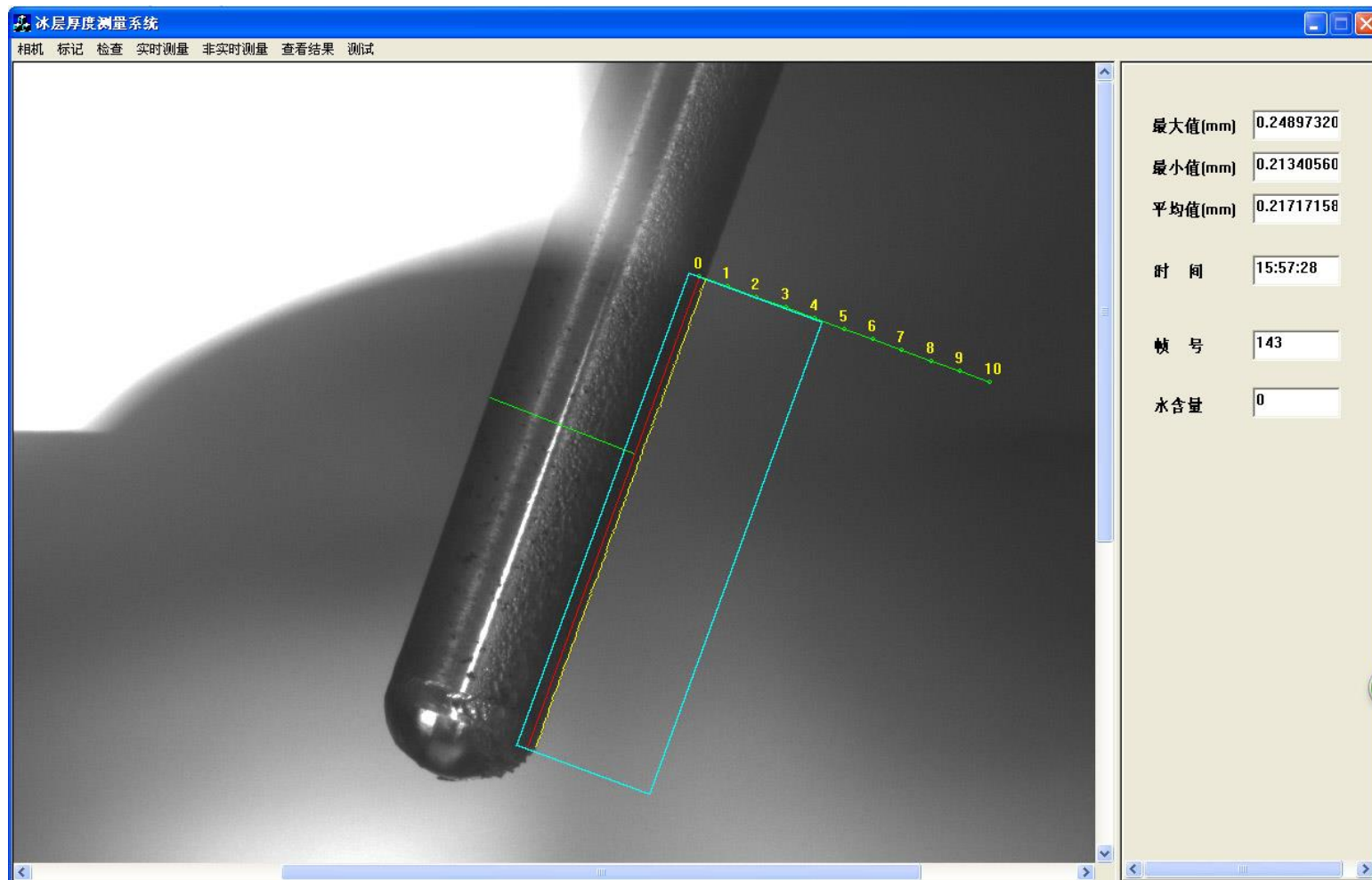
隐藏/打开悬浮窗

■ 火焰监测



1.1.4 应用领域

(15) 应用实例——结冰测量



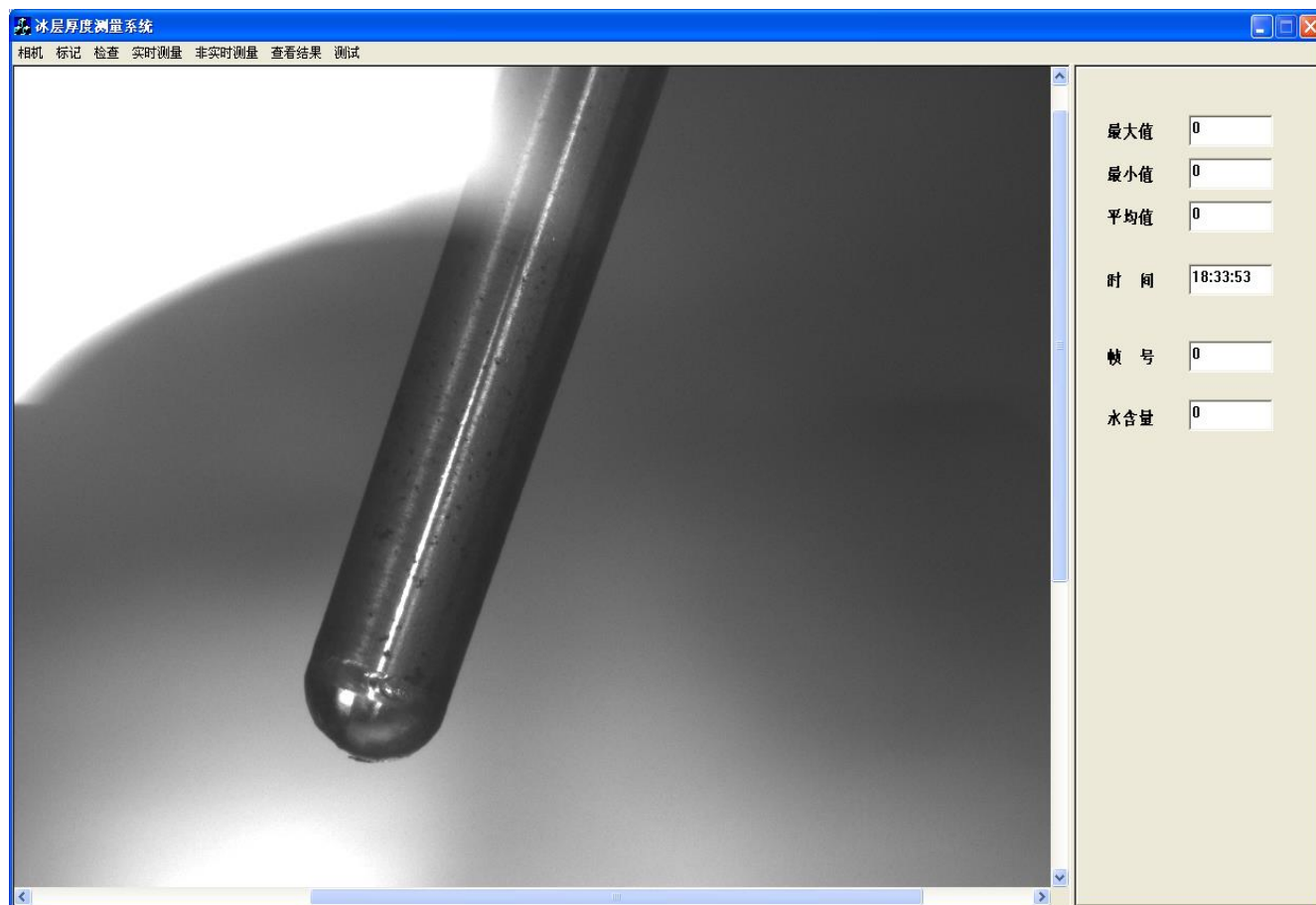
■ 结冰测量



1.1.4 应用领域

(15) 应用实例——结冰测量

- 打开第一帧图像

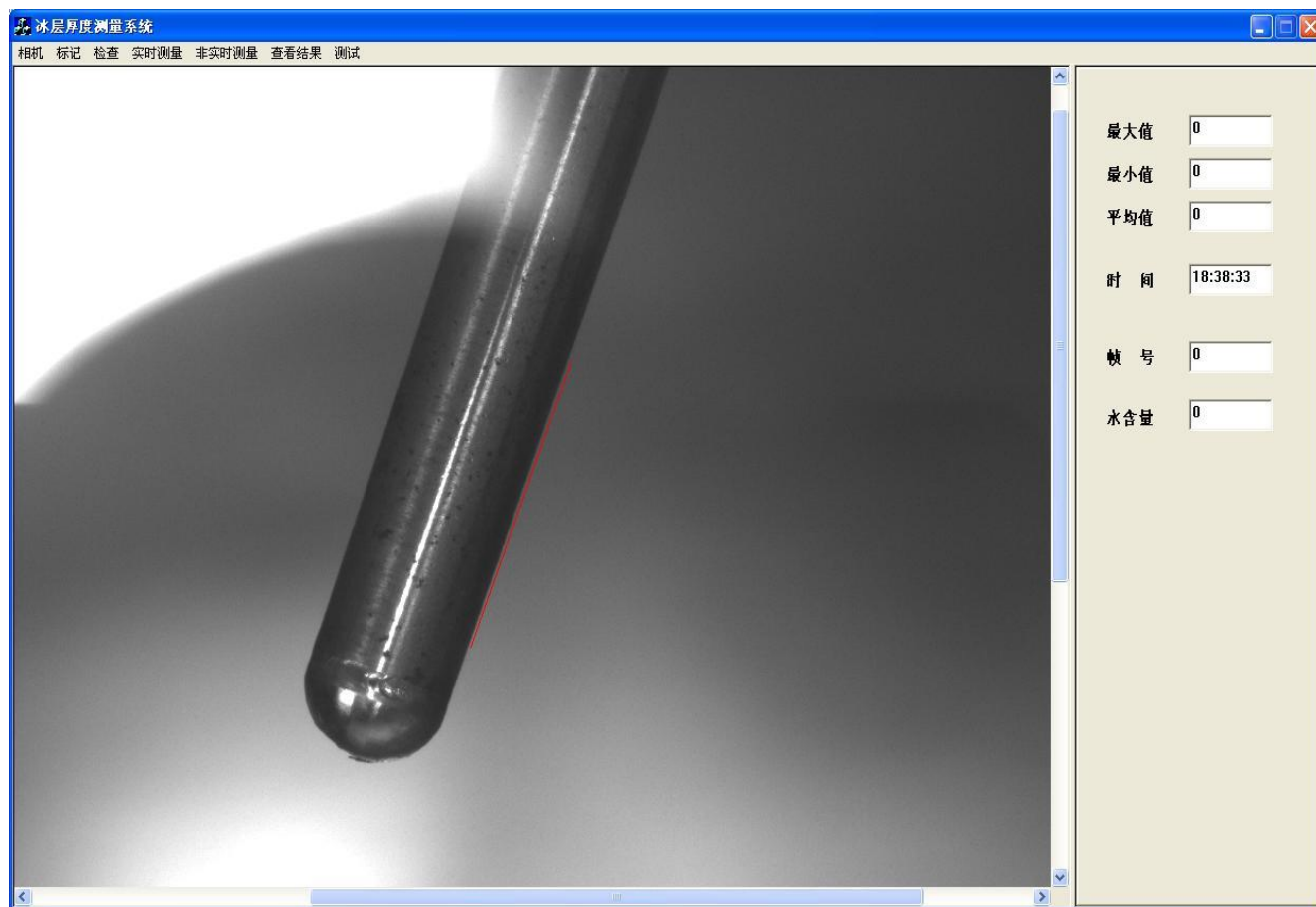




1.1.4 应用领域

(15) 应用实例——结冰测量

- 人工标记迎风面



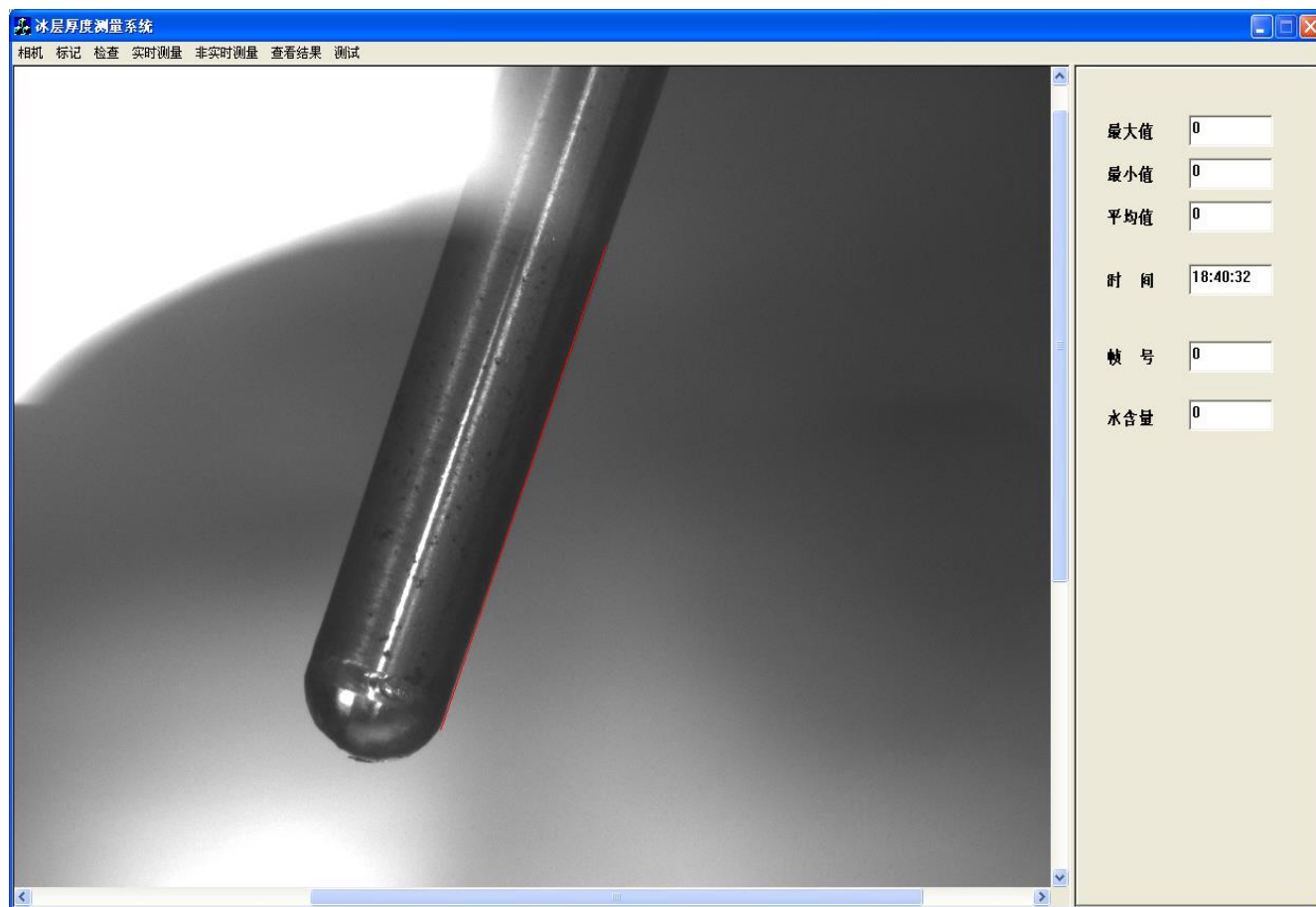


1.1.4 应用领域

(15) 应用实例——结冰测量

- 调整迎风面

精确确定迎风面

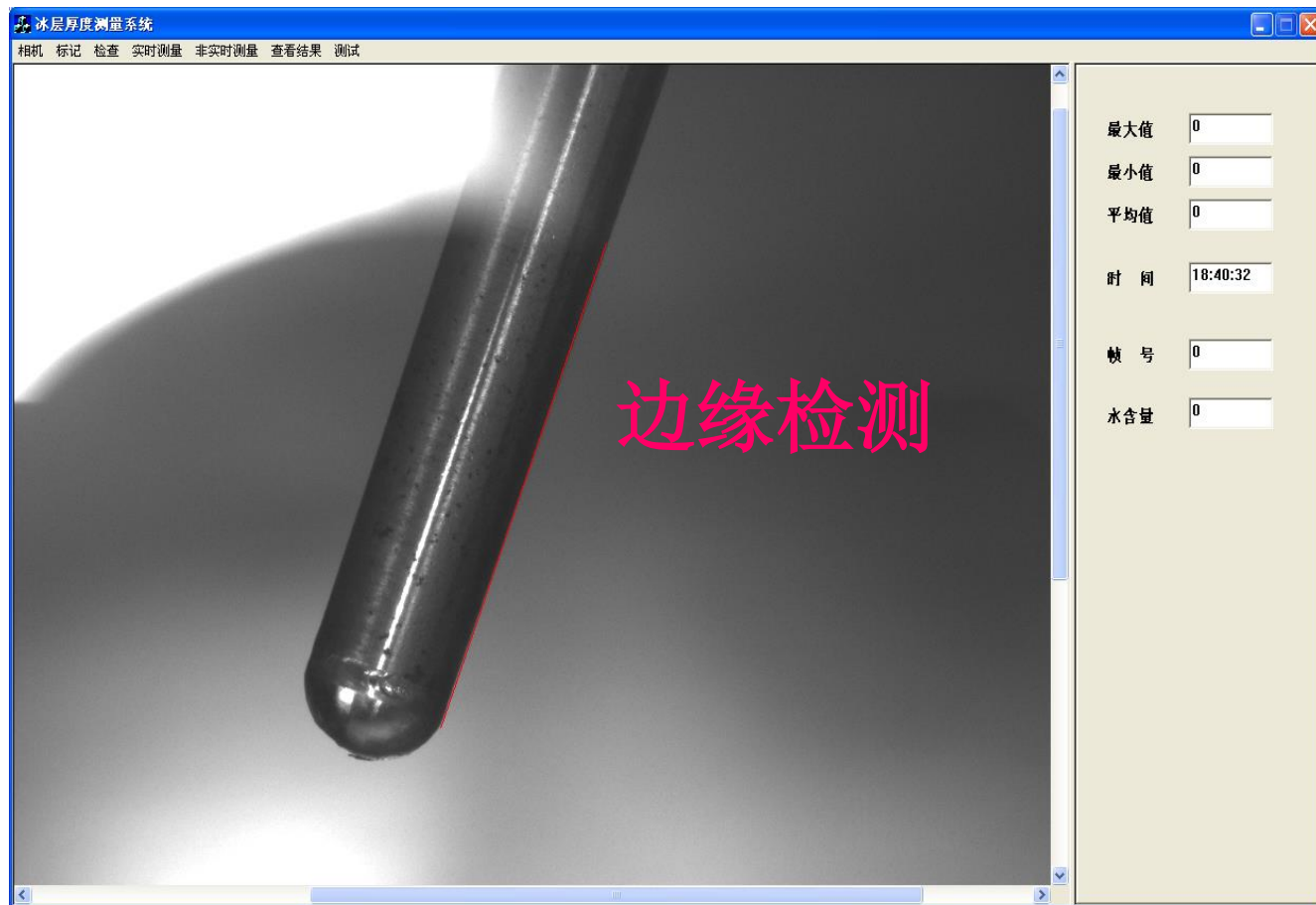




1.1.4 应用领域

(15) 应用实例——结冰测量

- 调整迎风面

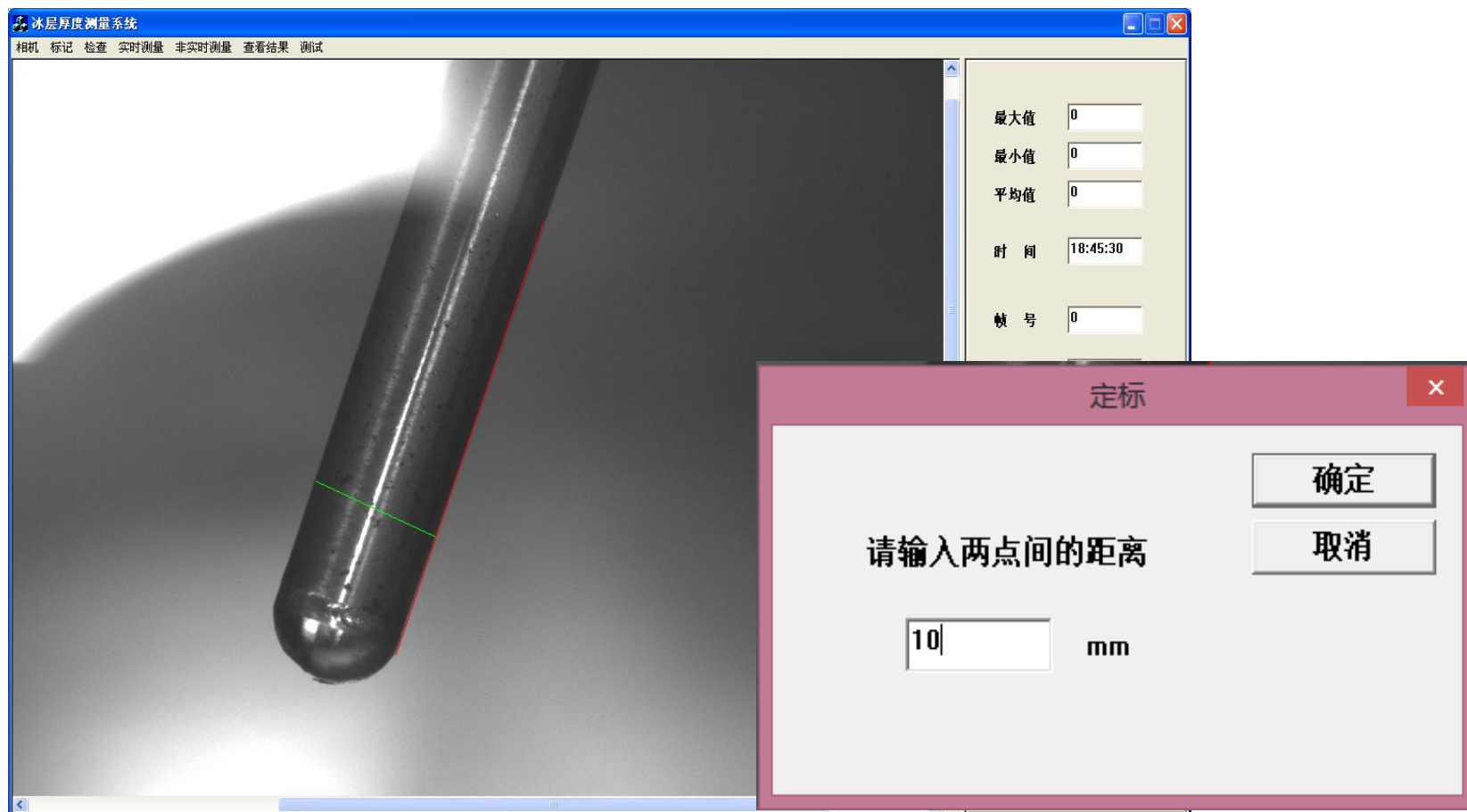




1.1.4 应用领域

(15) 应用实例——结冰测量

- 人工定标



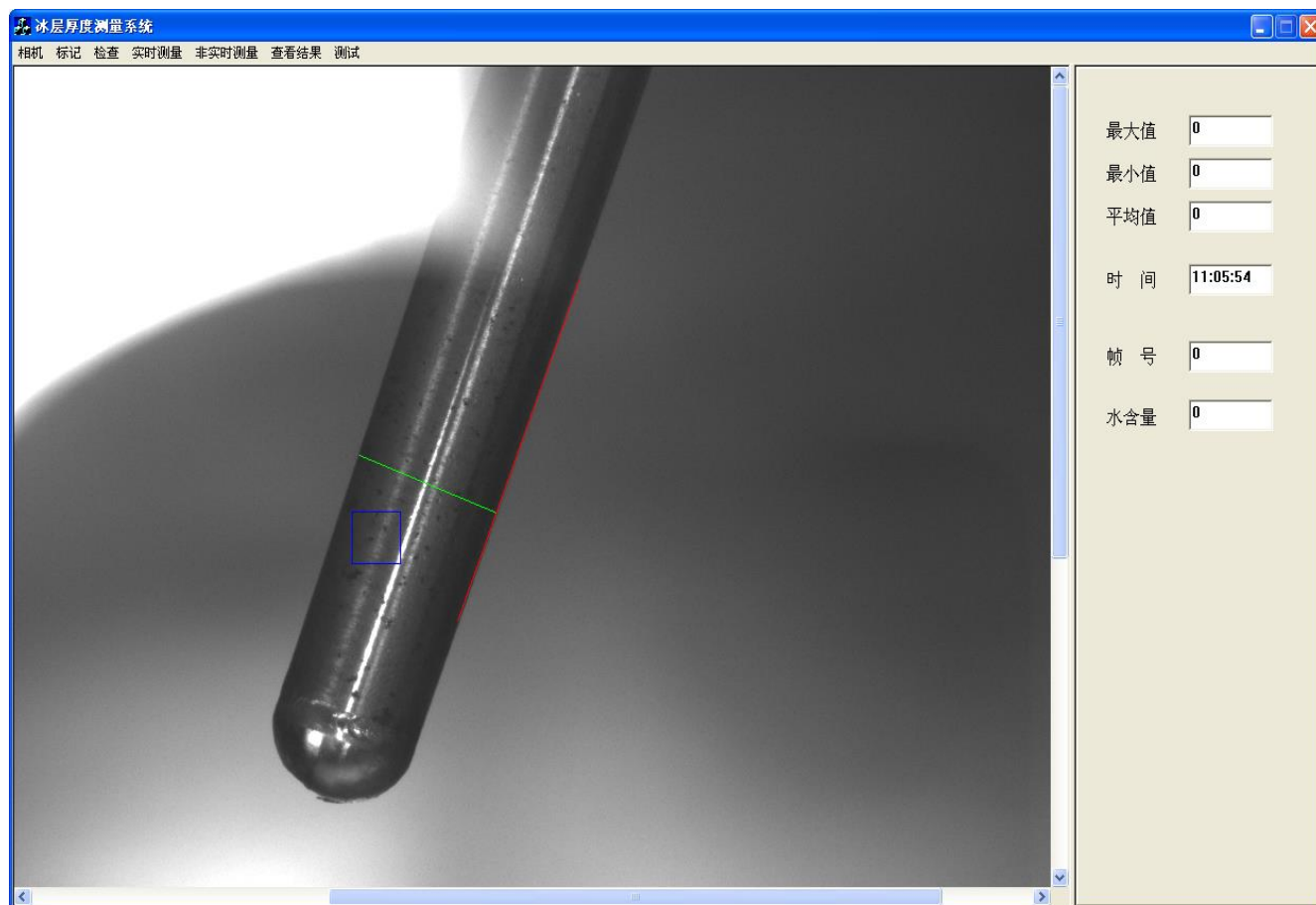


1.1.4 应用领域

(15) 应用实例——结冰测量

- 设置配准区域

克服部件运动的影响

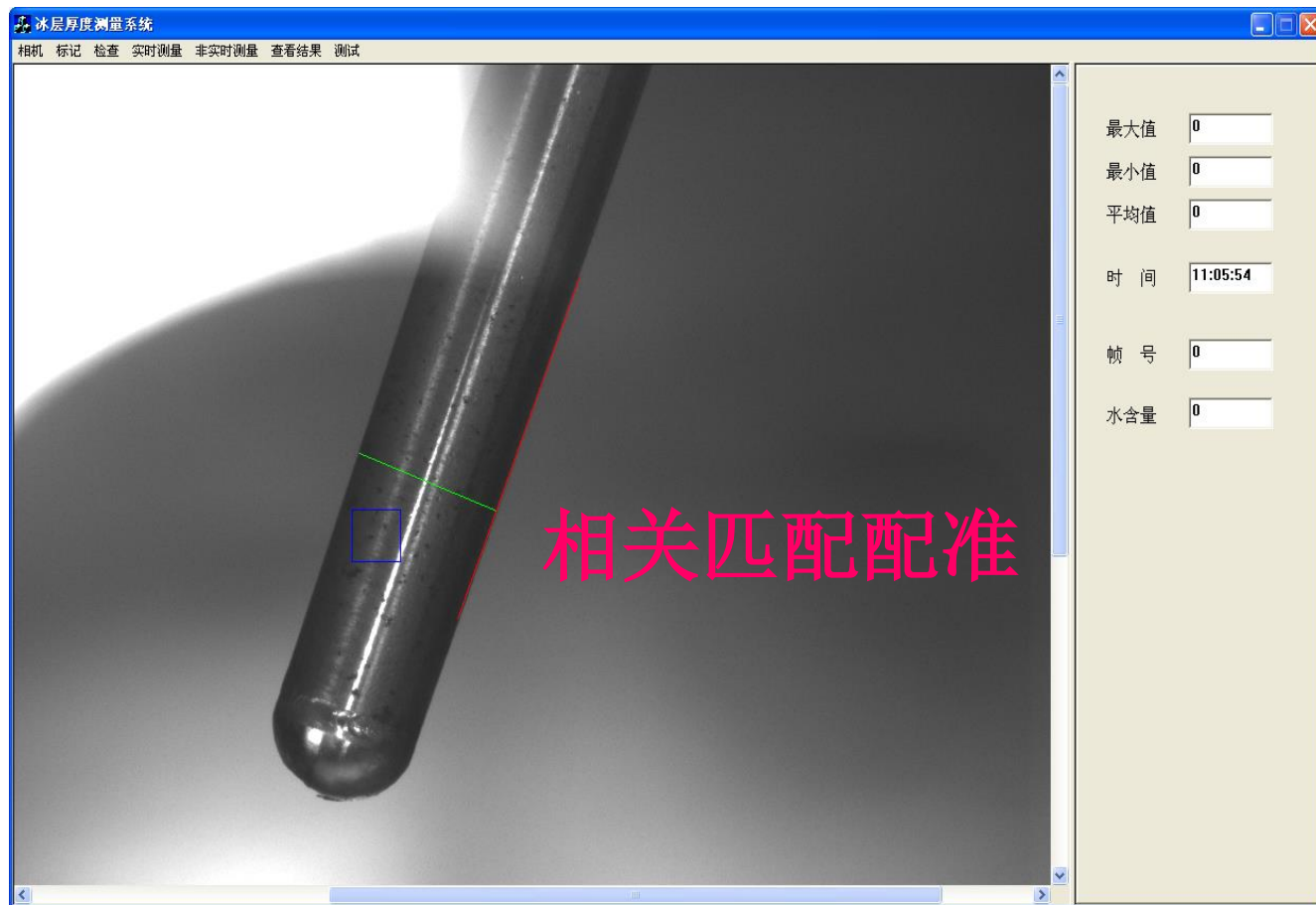




1.1.4 应用领域

(15) 应用实例——结冰测量

- 设置配准区域



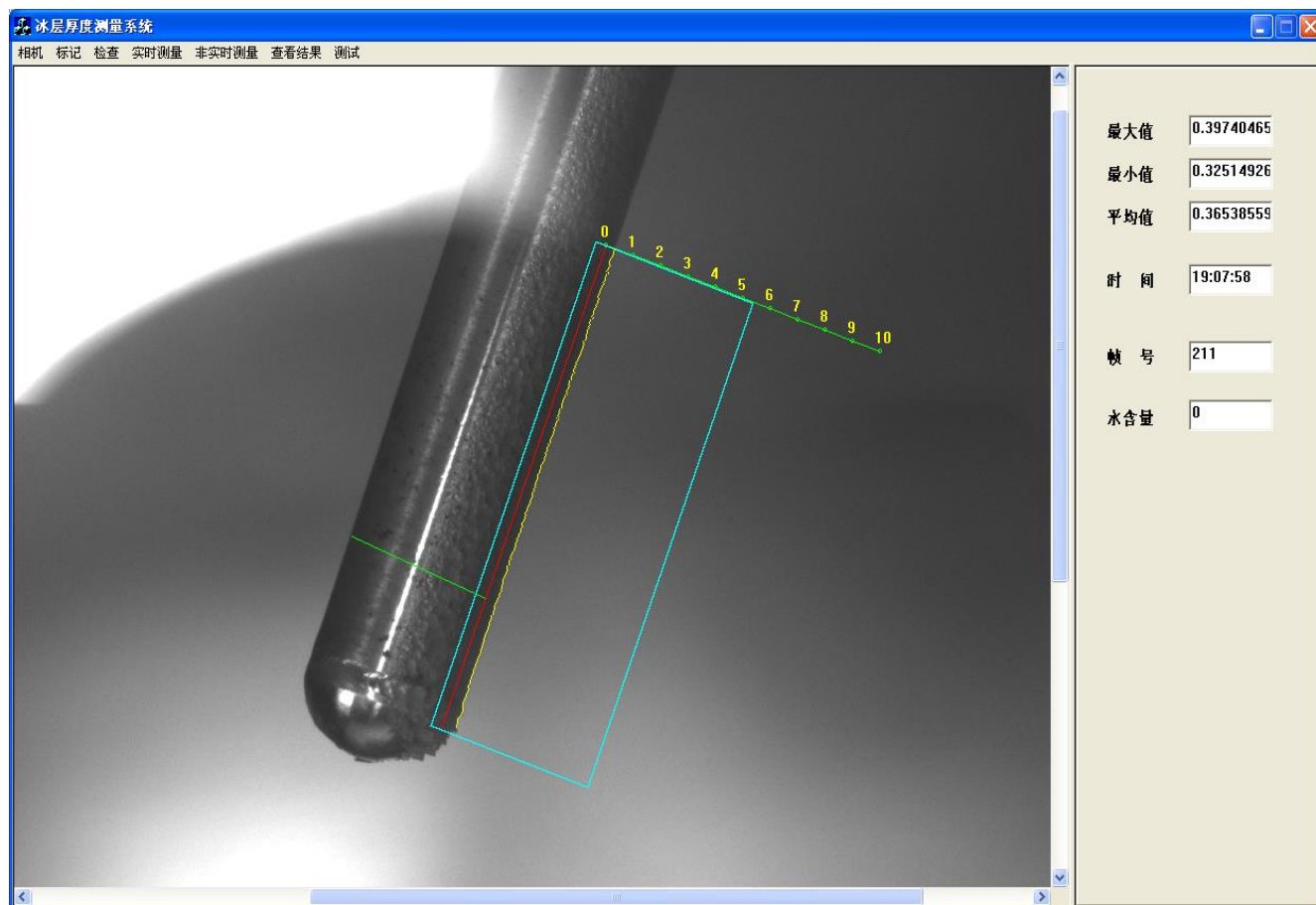


1.1.4 应用领域

(15) 应用实例——结冰测量

● 测量

冰层边缘定位？

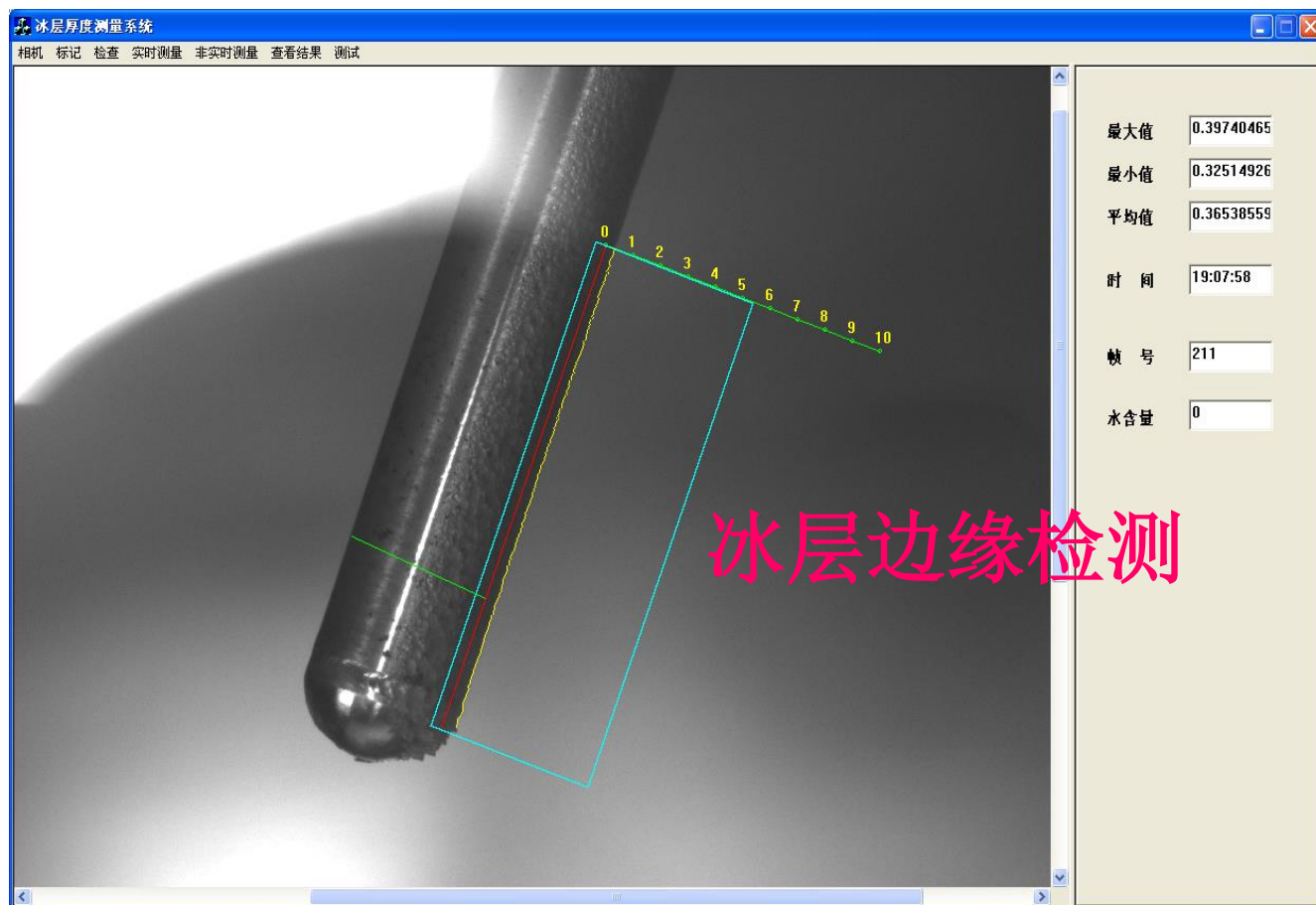




1.1.4 应用领域

(15) 应用实例——结冰测量

- 测量





1.1.4 应用领域

(15) 应用实例——公交车客流统计



■ 公交车客流统计

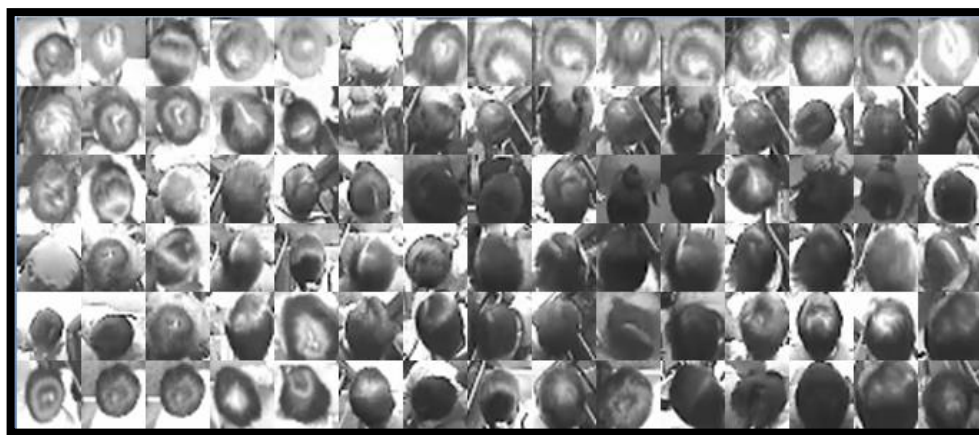


1.1.4 应用领域

(15) 应用实例——公交车客流统计

- 选择正负样本训练集

正样本集合



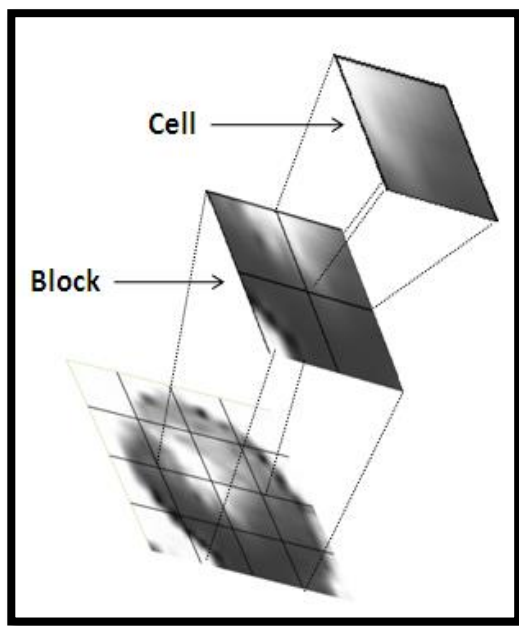
负样本集合



1.1.4 应用领域

(15) 应用实例——公交车客流统计

- 特征提取与分类训练



HOG特征

$$g(\vec{x}) = \vec{w} \cdot \vec{x} + b$$

**线性支撑向量机
(SVM)**



1.1.4 应用领域

(15) 应用实例——公交车客流统计

- 目标检测

人头目标尺度不同？





1.1.4 应用领域

(15) 应用实例——公交车客流统计

● 目标检测

多尺度分析





1.1.4 应用领域

(15) 应用实例——公交车客流统计

- 目标跟踪

分析人头目标运动轨迹



第5帧



第22帧



第30帧



第40帧



第60帧



第80帧



第90帧



第105帧



1.1.4 应用领域

(15) 应用实例——公交车客流统计

- 目标跟踪

Camshift+相关匹配



第5帧



第22帧



第30帧



第40帧



第60帧



第80帧



第90帧



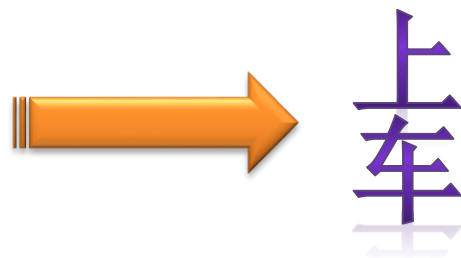
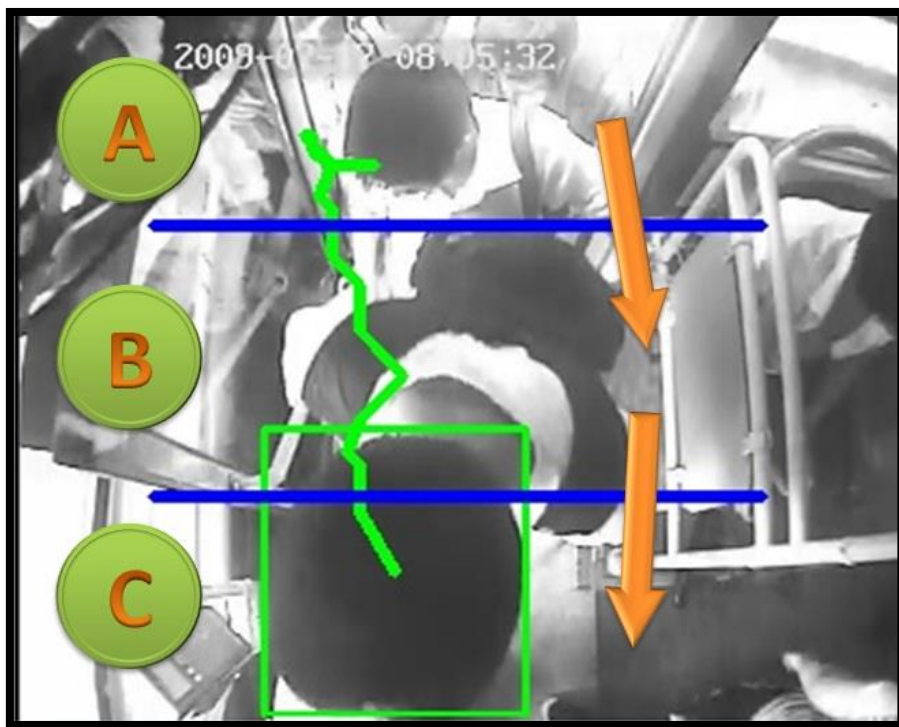
第105帧



1.1.4 应用领域

(3) 应用实例——公交车客流统计

- 计数





1.1.4 应用领域

高等学校人工智能创新行动计划

■ 重点任务

■ （一）优化高校人工智能领域科技创新体系

■ 2. 推动新一代人工智能核心关键技术创新。围绕新一代人工智能关键算法、硬件和系统等，加快机器学习、**计算机视觉**、知识计算、深度推理、群智计算、混合智能、无人系统、虚拟现实、自然语言理解、智能芯片等核心关键技术研究，在类脑智能、自主智能、混合智能和群体智能等领域取得重大突破，形成新一代人工智能技术体系；在核心算法和数据、硬件基础上，以提升跨媒体推理能力、群智智能分析能力、混合智能增强能力、自主运动体执行能力、人机交互能力为重点，构建算法和芯片协同、软件和硬件协同、终端和云端协同的人工智能标准化、开源化和成熟化的服务支撑能力。



1.2 图像基础

1.2.1 图像

1.2.2 图像表达和显示

1.2.3 图像存储



1.2.1 图像

- 图像：反射/辐射强度模式的空间分布

图像表达函数：辐射能量在空间分布的函数

通用图像表达函数： $T(x, y, z, t, \lambda)$

如： γ 射线图像、X射线图像、紫外线图像、可见光图像、红外线图像、微波图像、无线电波图像、交流电波图像，3-D图像、彩色图像、多光谱图像、立体图像和多视图图像，序列图像、深度图像、纹理图像、投影重建图像.....



1.2.1 图像

- **模拟图像：**从连续的客观场景直接观察到
用一个2-D数组 $f(x, y)$ 来表示， f , x , y 的值
可以是任意实数
- **数字图像：**把连续的模拟图像在坐标空间
 XY 和性质空间 F 都离散化了的图像
用 $f(x, y)$ 代表数字图像， f , x , y 都为整数



1.2.1 图像



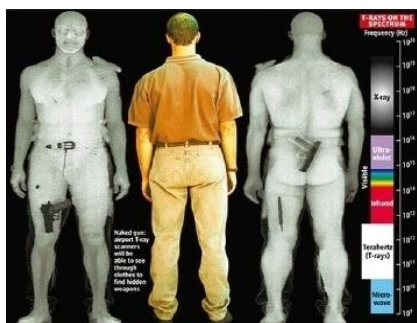
■ 可见光图像



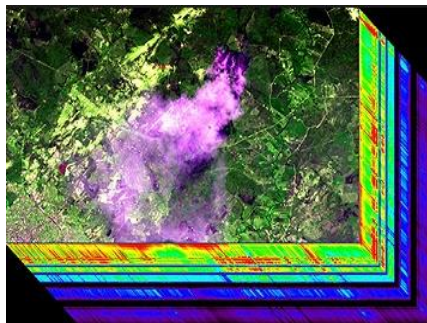
■ 近红外图像



■ 热红外图像



■ 太赫兹图像



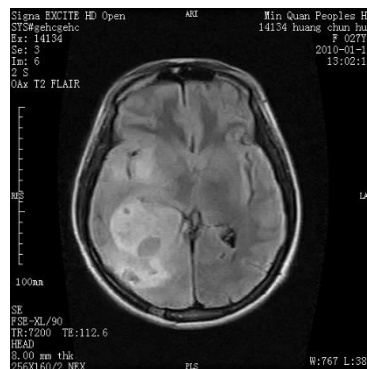
■ 高光谱图像



■ 深度(Kinect)图像



■ X射线图像



■ MRI图像



■ B超图像



1.2.2 图像表达和显示

- 图像表达:

矩阵表达

$$F = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1N} \\ f_{21} & f_{22} & \cdots & f_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ f_{M1} & f_{M2} & \cdots & f_{MN} \end{bmatrix}$$

矢量表达

$$F = [f_1 \quad f_2 \quad \cdots \quad f_N]$$

$$f_i = [f_{1i} \quad f_{2i} \quad \cdots \quad f_{Mi}]^T \quad i = 1, 2, \cdots, N$$



1.2.2 图像表达和显示

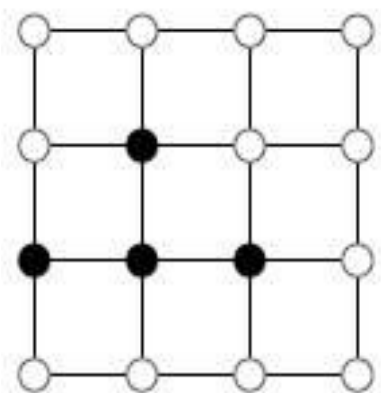
- 图像显示:

二值图像的3种不同的显示方式:

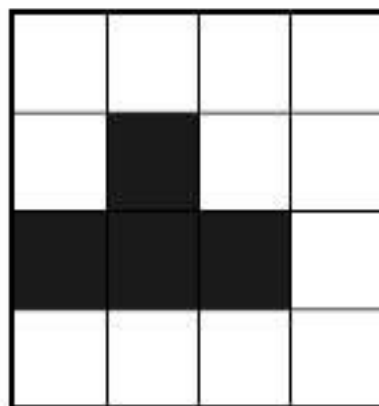
离散点集

覆盖区域

矩阵表达



(a)



(b)

0	0	0	0
0	1	0	0
1	1	1	0
0	0	0	0

(c)

图 1.2.1 3 种表达同一幅 4×4 的二值图像的方式

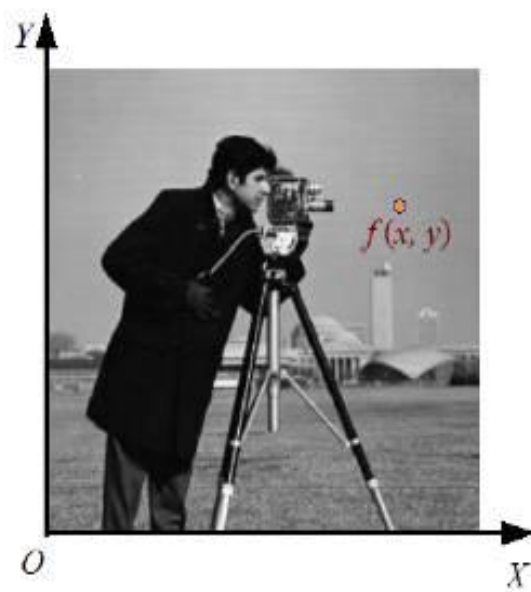


1.2.2 图像表达和显示

- 图像显示:
标准图像



(a)



(b)

图 1.2.2 标准图像显示示例



1.2.3 图像存储

- 图像文件格式:
 - (1) BMP格式
 - (2) GIF格式
 - (3) TIFF格式
 - (4) JPEG格式



1.3 像素间联系

1.3.1 像素邻域

1.3.2 像素间距离

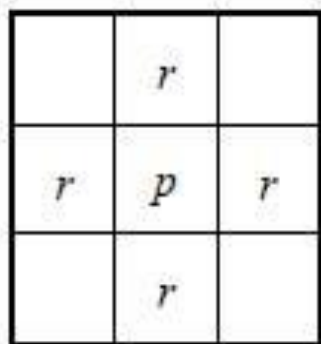


1.3.1 像素邻域

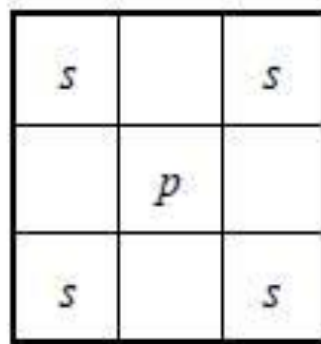
4-邻域: $N_4(p)$

对角邻域: $N_D(p)$

8-邻域: $N_8(p)$



(a)



(b)



(c)

图 1.3.1 像素的邻域



1.3.1 像素邻域

邻接：对两个像素 p 和 q 来说，如果 q 在 p 的邻域中，则称 p 和 q 满足邻接关系 **空间**

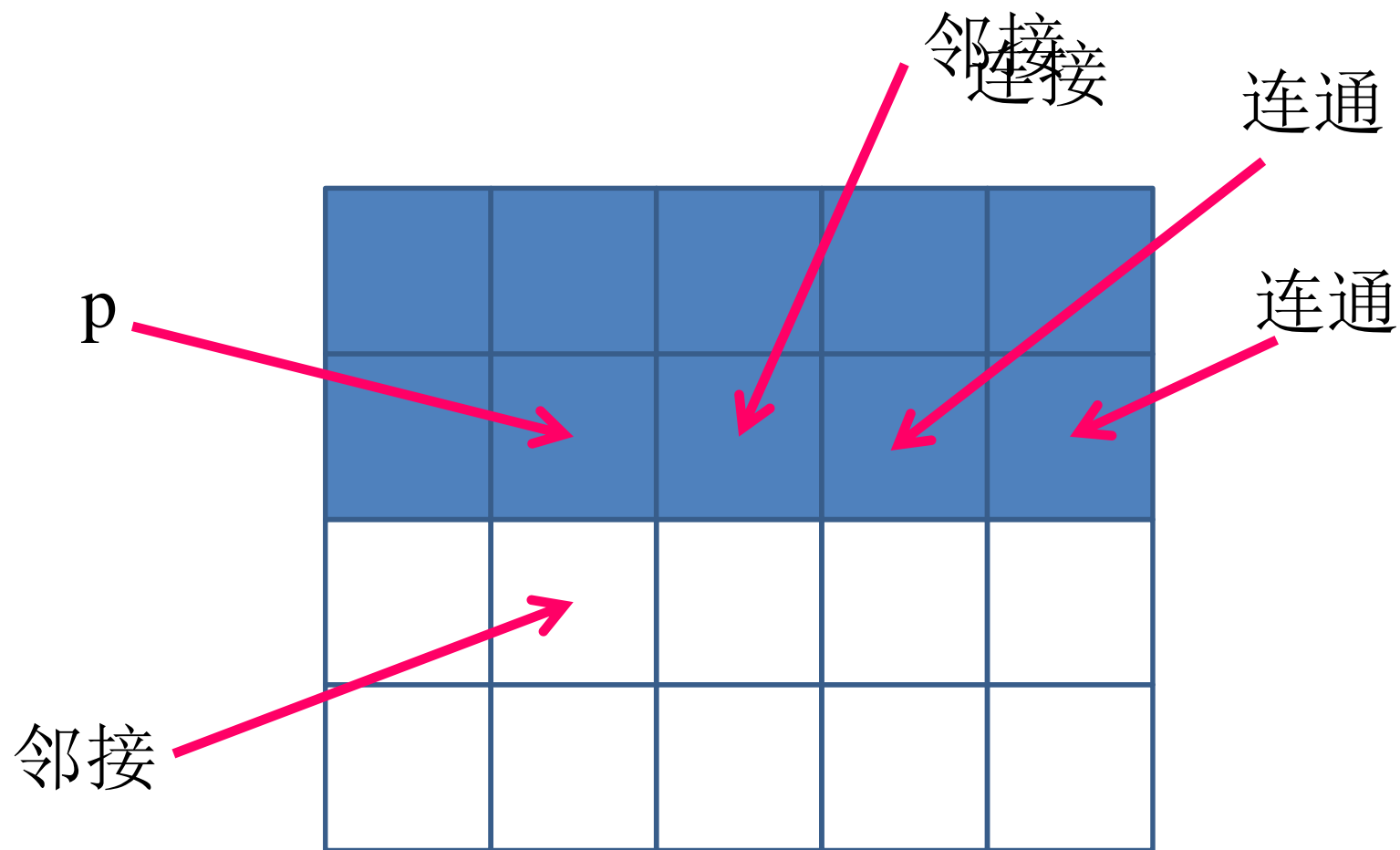
连接： p 和 q 邻接且灰度值均满足某个特定的相似准则 **空间+灰度**

连通：不（直接）邻接，但均在另一个像素的相同邻域中，且这3个像素的灰度值均满足某个特定的相似准则 **空间+灰度+空间传递**



1.3.1

像素邻域



二值图像，4-邻域



1.3.2 像素间距离

- 距离与范数

$$D_w(p, q) = \left[|x-s|^w + |y-t|^w \right]^{1/w}$$

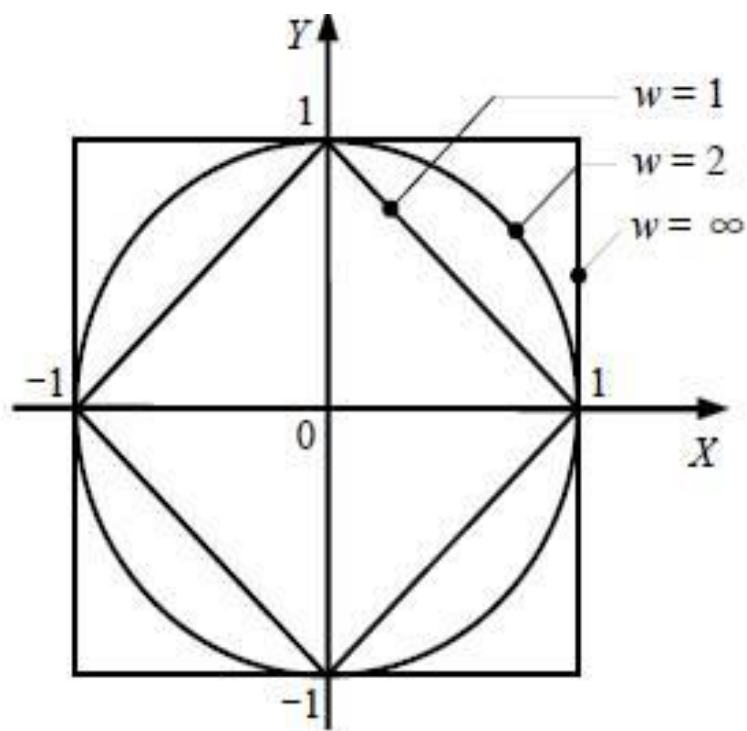


图 1.3.4 3 种范数和 3 种距离

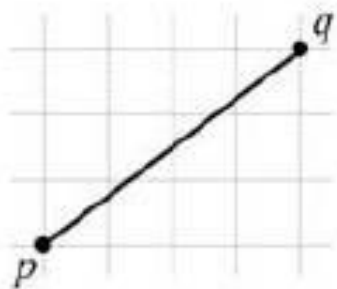


1.3.2 像素间距离

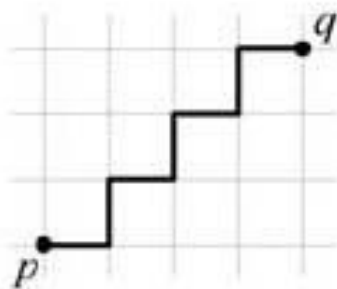
欧氏距离: $D_E(p, q) = [(x-s)^2 + (y-t)^2]^{1/2}$

城区距离: $D_4(p, q) = |x-s| + |y-t|$

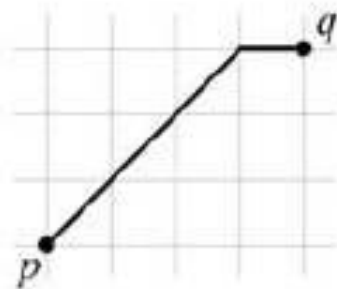
棋盘距离: $D_8(p, q) = \max(|x-s|, |y-t|)$



(a)



(b)



(c)

图 1.3.2 像素间距离的计算

欧氏距离

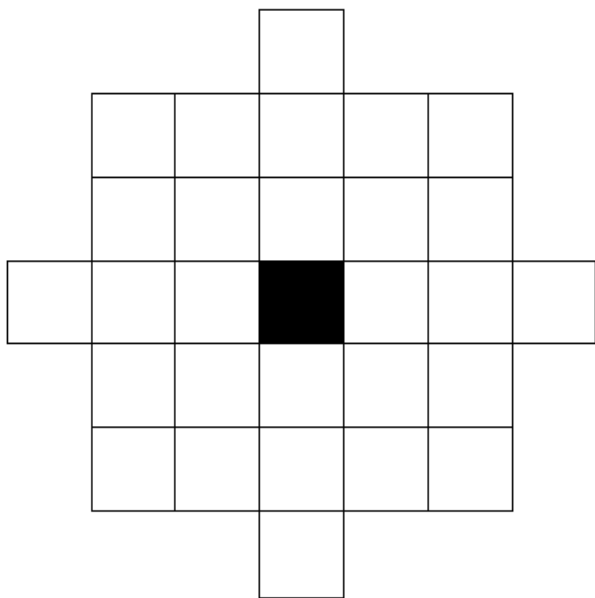
城区距离

棋盘距离

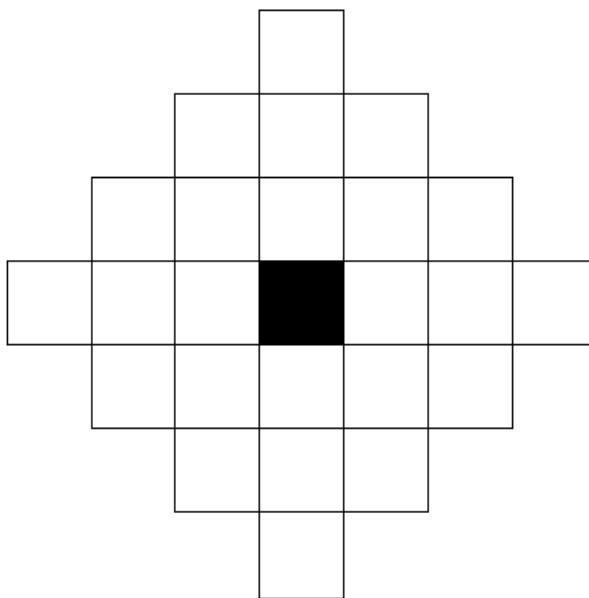


1.3.2 像素间距离

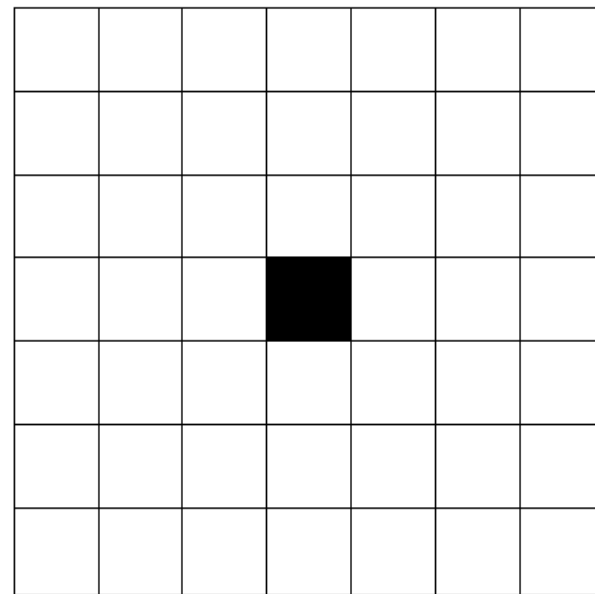
距离和邻域:



欧氏距离



城区距离



棋盘距离



1.3.2 像素间距离

距离和邻域:

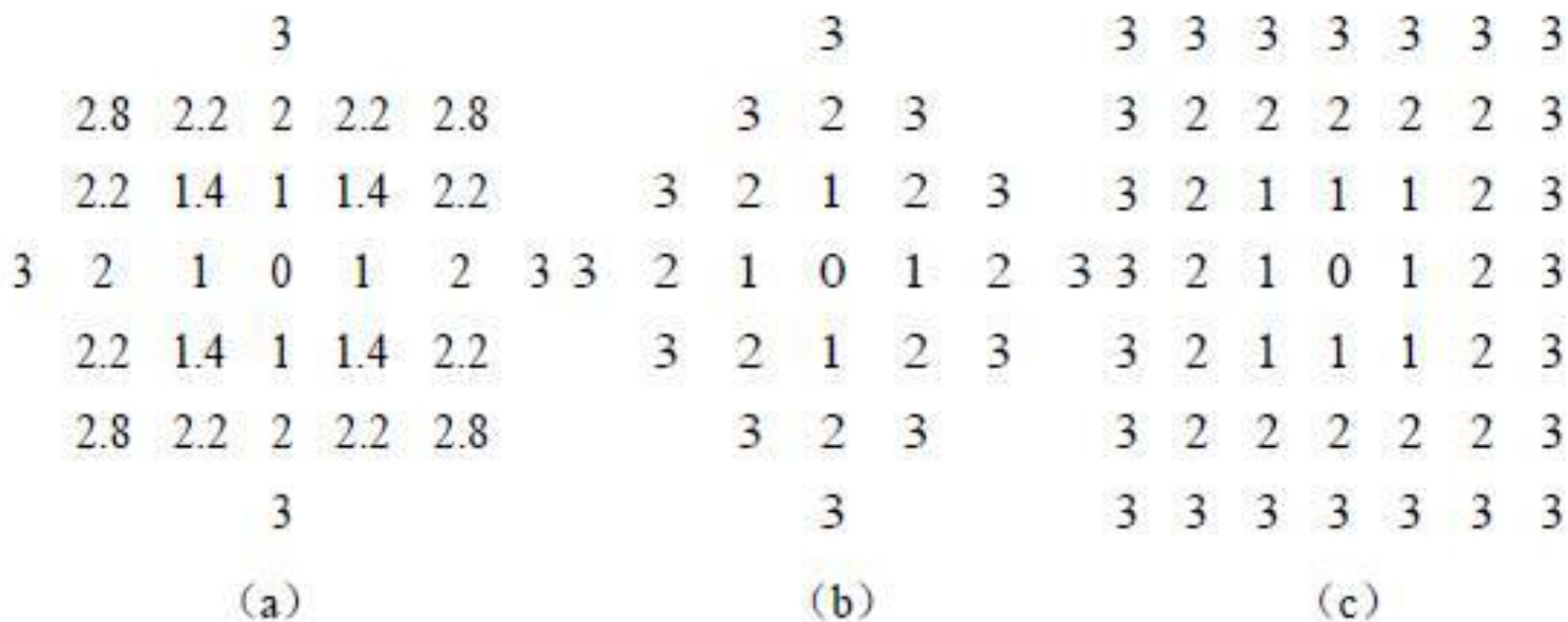


图 1.3.3 等距离轮廓示例

欧氏距离

城区距离

棋盘距离



1.3.2 像素间距离

距离和邻域:

4-邻域: $D_4=1$ 的像素集合

p的4-邻域: $N_4(p) = \{r | D_4(p, r) = 1\}$

8-邻域: $D_8=1$ 的像素集合

p的8-邻域: $N_8(p) = \{r | D_8(p, r) = 1\}$



1.3.2 像素间距离

距离和邻域:

等距离轮廓: 与中心像素的某种距离等于某个值的像素集合

$\Delta_i(r), i = 4, 8$: 与中心像素的 D_i 距离小于或等于 r 的等距离轮廓

$\#[\Delta_i(r)]$: 除中心像素外 $\Delta_i(r)$ 所包含像素个数

$$\#[\Delta_4(r)] = 4 \sum_{j=1}^r j$$

$$\#[\Delta_8(r)] = 8 \sum_{j=1}^r j$$

			3				3	3	3	3	3	3		
			3	2	3			3	2	2	2	2	2	3
	3	2	1	2	3		3	2	1	1	1	2	3	
	2	1	0	1	2	3	3	2	1	0	1	2	3	
	3	2	1	2	3		3	2	1	1	1	2	3	
			3	2	3			3	2	2	2	2	2	3
								3	3	3	3	3	3	3



1.3.2 像素间距离

距离和邻域:

距离变换：像素与最接近的区域外像素的距离可把二值图像变换为灰度图像

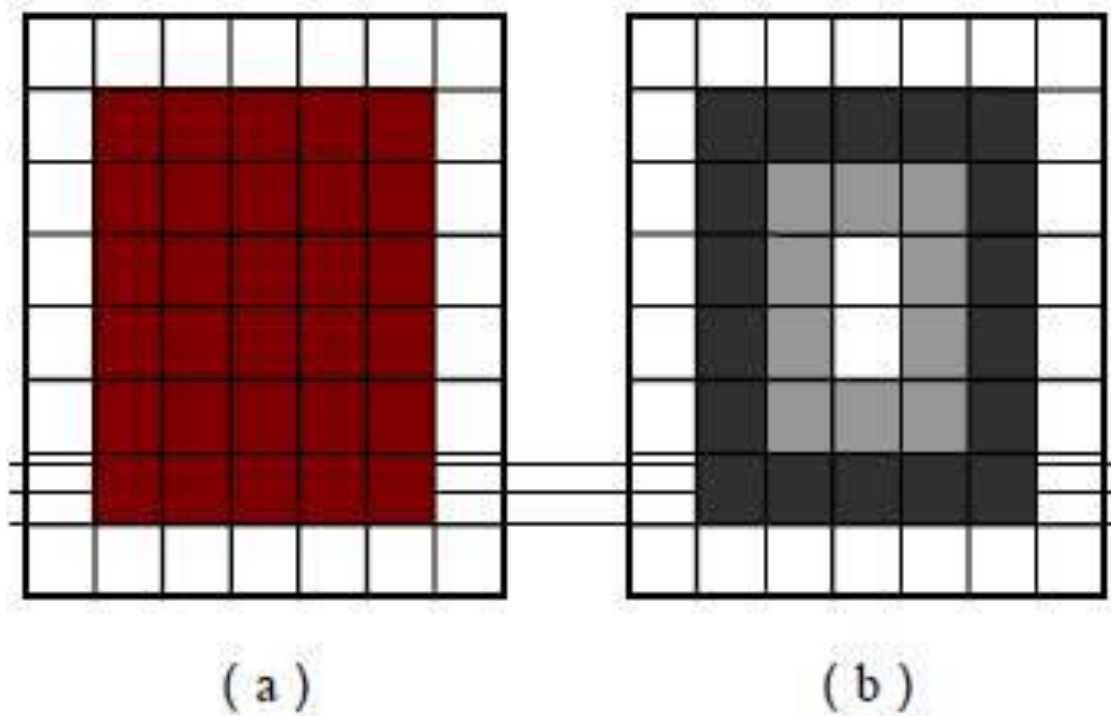


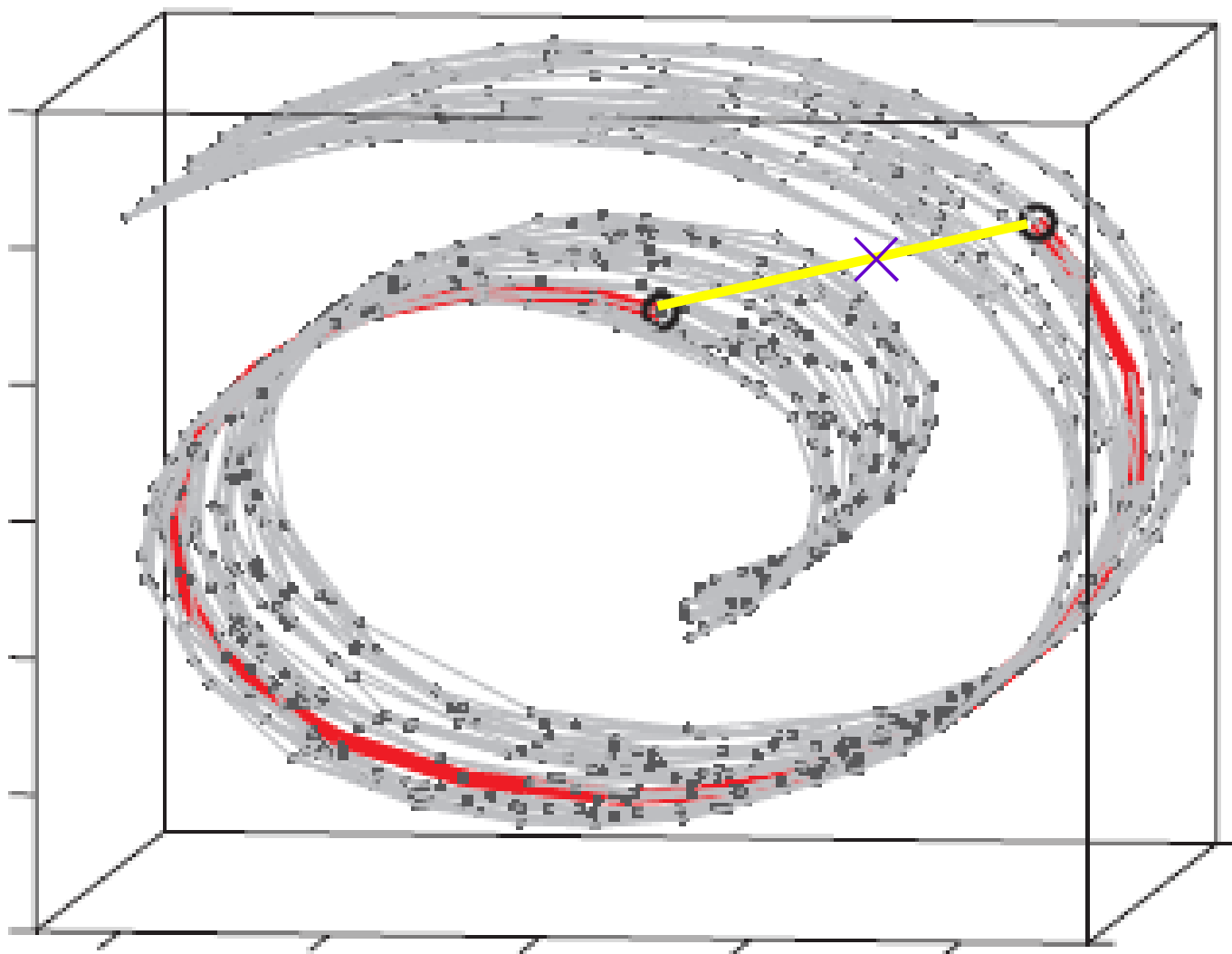
图 1.3.5 一幅二值图和它的距离变换图



1.3.2 像素间距离



测地距离 (Geodesic Distance)





1.3.2

像素间距离



度量学习 (Metric Learning)

相似度学习：如何度量图片之间的相似度使得不同类别的图像相似度小而相同类别的图像相似度大

人脸识别：构建距离函数以强化合适的特征（如发色，脸型等）

人体姿势识别：构建捕获姿势相似度的距离函数

马氏度量学习 $\|\mathbf{x} - \mathbf{y}\|^2 = (\mathbf{x} - \mathbf{y})^T \bullet (\mathbf{x} - \mathbf{y}) \Rightarrow \|\mathbf{x} - \mathbf{y}\|^2 = (\mathbf{x} - \mathbf{y})^T \mathbf{M} (\mathbf{x} - \mathbf{y})$

非线性度量学习 (Isomap, LLE)

深度度量学习



作业

- 练习题:

– 1.10

*1.10 计算图题 1.10 中两个像素 p 和 q 之间的 D_E 距离、 D_4 距离和 D_8 距离。

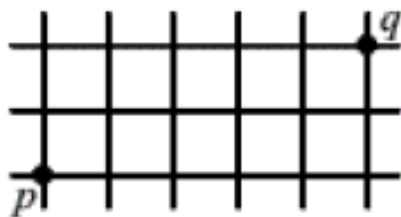


图 题 1.10

– 1.11

1.11 试分别计算 $\#[\Delta_8(4)]$ 和 $\#[\Delta_4(8)]$ 。

The end !