

图像的基本概念 (1) A.在韦氏 (Webster) 词典中, 图像 (image) 一词有如下定义: “物件或事物的一种表示、写真或临摹。”**B.**图像是人对视觉感知的物质再现。图像可以由光学设备获取, 如照相机、镜子、望远镜、显微镜等; 也可以人为创作, 如手工绘画。图像可以记录、保存在纸质媒介、胶片等等对光信号敏感的介质上。

图像的基本概念 (2) ——生成 A.一幅图像可以被定义为一个二维函数 $f(x,y)$, 通过取样和量化, 变成了计算机能处理的数字图像 **B.**数字图像由有限数字数值表示的像素组成, 在计算机中通常保存为二维数组。**C.**取样: 对坐标值进行量化 **D.**量化: 对幅值进行数字化 **E.**输入设备: 数码相机、扫描仪、seismographic profiling、airborne radar 等 **F.**技术合成: 数学函数、三维几何模型 (数字图像处理和计算机图形学的研究内容)

图像的基本概念 (3) ——格式 A.数字图像按照格式主要分为光栅图像和矢量图像 **B.**光栅图像格式: BMP、GIF、JPEG 等 **C.**光栅图像是通过许多像素点表示一幅图像, 每个像素具有颜色属性和位置属性。光栅图像可以从传统的相片、幻灯片上制作出来或使用数字相机得到 **D.**BMP(Windows 位图)格式: a.位图文件头 BITMAPFILEHEADER b.位图信息头 BITMAPINFOHEADER c.调色板(Palette) d.实际的图像数据: 对于用到调色板的位图, 图像数据就是该像素颜色在调色板中的索引值, 对于真彩色图像, 图像数据就是实际的 R、G、B 值 **E.**GIF(图形交换格式)格式: a. CompuServe 开发的图形交换文件格式 GIF (Graphics Interchange Format), 目的是在不同的系统平台上交流和传输图像。它是索引颜色图像, 图像最大不能超过 64 M, 颜色最多为 256 色 **F.**JPEG(Joint Photographer's Experts Group)格式: a. 由于 JPEG 的高压缩比和良好的图像质量, 使得它广泛应用于多媒体和网络程序中。JPEG 和 GIF 成为 HTML 语法选用的图像格式。b. JPEG 格式支持 24 位颜色, 并保留照片和其他连续色调图像中存在的亮度和色相的显著和细微的变化。c. JPEG 一般基于 DCT 变换的顺序型模式压缩图像, 为有损压缩。JPEG 压缩方法会降低图像中细节的清晰度, 尤其是包含文字或矢量图形的图像。**G.**矢量图像——图形 **H.**格式: WMF、SVG 等等 **I.**矢量图是用一系列绘图指令来表示一幅图, 如 AutoCAD 中的绘图语句。这种方法的本质是用数学(更准确地说是几何学)公式描述一幅图像。图像中每一个形状都是一个完整的公式, 称为一个对象。对象是一个封闭的整体, 所以定义图像上对象的变化和对象与其他对象的关系对计算机来说是简单的, 所有这些变化都不会影响到图像中的其他对象。**J.**矢量图像格式的优点: a. 公式化表示图像使得矢量图具有两个优点: 一是它的文件数据量很小; 二是图像质量与分辨率无关。这意味着无论将图像放大或缩小了多少次, 图像总是以显示设备允许的最大清晰度显示。b. 矢量图有一个明显的缺点, 就是不易制作色调丰富或色彩变化太多的图像, 而且绘出来的图像不是很逼真, 同时也不易在不同的软件间交换文件。

图像的基本概念 (3) ——分类 A.对光栅图像来说, 用来描述每个像素所采用的数据量不同, 因此存在色彩深度的差异: a.亮度图像 b.二值图像 c.索引图像 d.RGB 图像 a.二值图像 Binary Image: 图像中每个像素亮度值 Intensity 取值[0,1] b.灰度图像 Gray Scale Image: 图像中每个像素用 0~255 的亮度值表示, 0~255 表示不同的灰度级 c.彩色图像 Color Image: 通过颜色空间能够表现出更丰富的色彩

内容提要 II - 图像检索基础: 应用需求, 表示模型, 检索方式, 系统框架, 视觉特征, 相似度量, 相关反馈, 演示系统 **A. 应用需求 - 1** a.传统信息检索方法缺陷: aa. 需要事先对信息按设计好格式统一加工和索引 (包括分类、标记关键词或索引词等) ab. 多媒体信息的数据量急剧增加, 人工加工的工作量太大, 成本极高 ac. 对多媒体信息, 特别是对图象和视频等视觉信息, 很难进行有效的、精确的文字描述 ad. 对信息的人工加工分散在世界各地进行, 因而所作的分类、所用的关键词等均可能很不相同 b. 必须研究全新的检索技术 - 基于内容的多媒体信息检索和查询技术 **B.应用需求 - 2** a.第一个商用的基于内容的图象搜索引擎是 IBM 的 QBIC, 该系统用颜色直方图和 Tamura 纹理表示图象内容。类似的系统有 MIT 媒体实验室的 Photobook 系统, 还有 VisualSeek、Netra、Virage、MARS 等 b.人们逐步认识到, 虽然图象可以通过颜色、纹理、形状等中间层 (metalevel) 数据描述, 但是对于普通的图象来讲图象理解仍然是一个难题! c.一个挑战性工作就是基于特征空间的相似度实现图象检索, 而不是基于图象理解和图象分析等技术 **C.应用需求 - 3** a.困难 aa.不同的人对图像有不同的理解 ab.即使同一个人不同的上下文环境下, 对图像的理解也不同。ac.这些问题刺激人们开发崭新类型的基于内容搜索技术和新的查询类型 **D. 应用需求 - 4** a. 问题: 1、在人的视觉系统中, 其实并不存在这种相似匹配过程; 2、人没有通用的相似度概念, 相反人是依据具体场合决定相应的相似度。因此, 为实现类似人的相似匹配, 必须: 1、选择一个特定的应用领域; 2、理解用户在该领域如何判断相似度的; 3、构造一个可以重复人的能力的系统。b. 因为颜色和纹理是人的视觉感知的基本方面, 主要考虑开发一组搜索和处理颜色图案的技术。在不考虑具体应用的前提下, 为了检索成功, 就非常必要弄清楚: 1、哪些颜色和纹理信息是人实际使用的? 2、人是如何组合这些信息来决定两个图案是相似的? **E.应用需求 - 5** a 文本搜索引擎得到普遍应用 aa 用户提交查询简单方便 ab 经常返回大量无关文档 ac 智能化程度有待提高 b 对海量图象视频数据来说缺乏有效搜索工具 ba 用户提交查询困难 bb 很多问题没有解决 bc 尚未真正进入应用

图像表示 - 不同层次: A 数据层 a 各种格式、各种媒体的数据对象, 可能包括文本、图片、视频、音频等, 一般无结构或半结构 **B 特征层** a 从不同角度观察数据对象所得到的结果, 是关于数据对象的抽象表示和具体描述 **C 概念 (语义) 层** a 关于数据对象更高层次的、基于语义的抽象和表示 b 特征层和概念层之间存在鸿沟, 难以跨越。 ***图像内容的层次模型 A** 第 1 层 原始数据层 图像的原始像素点, **B** 第 2 层: 物理特征层 图像的颜色、纹理、形状和轮廓等低层物理特征(此层进行特征提取), **C** 第 3 层: 语义特征层 人们对图像内容概念级的反映 ***图像检索研究的三个发展阶段 A** 基于文本的图像检索 a 人工标注信息 **B** 基于内容的图像检索 a 低层物理特征 **C** 基于语义的图像检索 a 图像自动标注 ***图像检索 - 不同方式 A** 精确匹配 a 用于传统关系数据库 b 数据对象具有严谨结构 **B** 相似检索 a 基于相似度大小返回查询结果 b 不同应用需不同相似度计算方法 **C** 查询种类 a 范围查询: 返回相似度大于给定门限的结果 b 近邻查询: 返回若干最相似的结果。 ***视觉特征-颜色 A** 颜色特征 a.一般为全局特征 b.和分辨率无关 c.不需要分割出图象对象 d.问题: 不同光照下条件对物体颜色的感知不一样 **B** 优点和缺陷 a.提取算法非常简单 b.全局特征, 和图象的旋转、平移、缩放等无关 c.语义鸿沟: 可找到完全不同内容的图象, 具有相同颜色直方图。***视觉特征-颜色:颜色直方图** SCD, CLD, DCD ... ***视觉特征-纹理:** 灰度共生矩阵, Gabor, Wavelet, Tamura, EHD... ***视觉特征-形状:** 基于轮廓 基于区域 ***相似度量-相似检索基础: A** 单一特征的相似度计算 a 基于距离度量 b 基于相关系数 c 基于关联系数 **B** 多个特征时相似度计算 ***相似度量-单一特征 A** 分别用特征矢量 X 和 Y 描述多媒体对象 OX 和 OY **B** 有三种计算对象之间相似度大小的方法: 距离度量(distance measure)、相关系数(correlation coefficients)和关联系数(association coefficients) ***距离度量 A** Minkowski 距离 a.r=1 时: Manhattan 距离 (l1 距离) b.r=2 时: Euclidean (欧氏) 距离 (l2 距离) c.r= ∞ 时: Chebyshev 距离 (l ∞ 距离) **距离度量-续** 加权欧氏距离 (Φ 为加权矩阵) ***相关系数 A** 非归一化相关系数 **B** 余弦度量(Cosine Measure) - 矢量夹角 ***关联系数 A** 假设特征矢量 X 和 Y 均为二值的, 即: X 和 Y 属于 {0,1}^n, 下面定义四个量: α : 特征矢量 X 和 Y 中同为 1 的数量 β : 特征矢量 X 中为 1、Y 中为 0 的数量 γ : 特征矢量 X 中为 0、Y 中为 1 的数量 σ : 特征矢量 X 和 Y 中同为 0 的数量 **B** 各种计算公式 Russel & Rao Jaccard & Needham Kalzinski Sokal & Mitchener Rogers & Tanimoto Yule ***相似度计算-多特征融合 A** 每个特征单独计算相似度 σ_i **B** 基于各特征的相似度复合计算, 加权系数 λ_i 根据实际应用确定 ***相关反馈-提供查询性能 A** 基于内容的图象检索比文本检索技术落后 人很难精确表达和提交其视觉查询 **B** 相关反馈是一种比较现实的解决方案, 它不要求用户输入精确的初始查询, 只希望用户指出查询结果中的正/负样例, 然后通过相关反馈估计出用户想要提交但是又无法描述的查询