

图像工程的基本概念和理论基础

雷英杰, 郑全第, 张刚, 周创明

(空军工程大学 导弹学院, 陕西 三原 713800)

摘要:首先对图像工程领域里几个不够统一的基本概念、术语给出明确的解释或定义,接着对图像处理基本算法模型及图像运算等内容进行了概括,最后描述了图像编码系统中失真的度量方法,以期澄清图像工程的若干基本概念、揭示图像工程的理论基础。

关键词:图像工程;图像处理;图像编码;图像运算

中图分类号:TP393 **文献标识码:**A **文章编号:**1009-3516(2003)04-0060-05

图像工程是一门系统地研究各种图像理论、技术和应用的新兴交叉学科。其研究方法与数学、物理学、生理学、心理学、电子学、计算机科学等学科相互借鉴,其研究范围与模式识别、计算机视觉、计算机图形学等专业互相交叉,其研究进展与人工智能、神经网络、遗传算法、模糊逻辑等理论和技术密切相关,其发展应用与生物医学、遥感、通信、文档处理等许多领域紧密结合^[1]。

1 基本概念

数字图像(image)。把连续的图像在坐标空间 XY 和性质空间 F 都离散化了的、供计算机进行处理的图像,称为数字图像。图像中每个基本单元叫做图像元素,简称像素(pixel)。

图像技术。图像技术在广义上是与各种图像有关的技术的总称。目前人们研究的是数字图像,主要应用的是计算机图像处理技术。这包括利用计算机和其他电子设备进行和完成的一系列工作,例如图像的采集、获取、编码、存储和传输,图像的合成和产生,图像的显示和输出,图像的变换、增强、恢复(复原)和重建,图像的分割,图像目标的检测、表达和描述,图像特征的提取和测量,序列图像的校正,3-D景物的重建复原,图像数据库的建立、索引和抽取,图像的分类、表示和识别,图像模型的建立和匹配,图像和场景的解释和理解,以及基于它们的判断决策和行为规划等等。另外,图像技术还包括为完成上述功能而进行的硬件设计及制作等方面的技术。

图像工程。图像工程是将图像技术发展过程中出现的各种新理论、新方法、新算法、新手段、新设备等进行综合研究和集成应用的一个整体框架。图像工程学科则是将数学、光学等基础科学原理与图像应用过程中积累的技术经验相结合而形成和发展起来的。图像工程的内容十分丰富,根据抽象程度和研究方法等的不同可分为三个层次:图像处理、图像分析和图像理解^[2]。

图像处理。图像处理着重强调在图像之间进行的变换。

图像分析。图像分析主要针对图像中感兴趣的目标进行检测和测量,以获得他们的客观信息,从而建立对图像的描述。

图像理解。图像理解的重点是在图像分析的基础上进一步研究图像中各个目标的性质和它们之间的相互联系,并得出对图像内容含义的理解以及对原来客观场景的解释,从而指导和规划行动。

收稿日期:2002-03-16

收稿日期:国家教育部骨干教师资助计划资助项目(GG-810-90039-1003)

作者简介:雷英杰(1956-),男,陕西渭南人,教授,博士生导师;主要从事智能信息处理与模式识别研究。

2 算法模型

颜色模型。目前常用的颜色模型可分为两类,一类面向诸如彩色显示器或打印机之类的硬设备,在这一类模型中最常用的是 RGB 模型。RGB 模型是基于笛卡尔坐标系统的,三个坐标轴分别为 R、G、B。另一类是以彩色处理为目的的应用,如动画中的彩色图形,在这一类模型中最常用的是 HSI 颜色模型,其中 H 表示色调(hue),S 表示饱和度(saturation),I 表示强度(intensity,对应成亮度和图像灰度)。RGB 和 HSI 这两种颜色模型是图像技术中最常见的模型^[3]。

从 RGB 到 HSI 的转换。

$$I = \frac{1}{3}(R + G + B) \quad S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)]$$

$$H = \arccos \left\{ \frac{[(R - G) + (R - B)/2]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right\}$$

从 HSI 到 RGB 的转换。

$$R = \begin{cases} I \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right] & 0^\circ \leq H \leq 120^\circ \\ I(1 - S) & 120^\circ \leq H \leq 240^\circ \\ 3I - (G + B) & 240^\circ \leq H \leq 360^\circ \end{cases} \quad G = \begin{cases} 3I - (B + R) & 0^\circ \leq H \leq 120^\circ \\ I \left[1 + \frac{S \cos(H - 120^\circ)}{\cos(180^\circ - H)} \right] & 120^\circ \leq H \leq 240^\circ \\ I(1 - S) & 240^\circ \leq H \leq 360^\circ \end{cases}$$

$$B = \begin{cases} I(1 - S) & 0^\circ \leq H \leq 120^\circ \\ 3I - (R + G) & 120^\circ \leq H \leq 240^\circ \\ I \left[1 + \frac{S \cos(H - 240^\circ)}{\cos(300^\circ - H)} \right] & 240^\circ \leq H \leq 360^\circ \end{cases}$$

成像变换。图像采集过程中将 3-D 客观场景投影到 2-D 平面所用的几何透视变换称为成像变换。我们定义透视变换矩阵 P 和空间点在对应笛卡尔坐标系 XYZ 的齐次坐标点 w_h 分别为

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/\lambda & 1 \end{bmatrix} \quad w_h = [kX \ kY \ kZ \ k]^T$$

则矢量

$$c_h = P w_h = [kX \ kY \ kZ \ -kZ/\lambda + k]^T$$

的元素给出齐次形式的摄影机坐标,这些坐标可用 c_h 的第 4 项分别去除前 3 项转换成笛卡尔形式。所以,摄影机坐标系中任意一点的笛卡尔坐标可以表示为矢量形式

$$c = [x \ y \ z]^T = \left[\frac{\lambda X}{\lambda - Z} \quad \frac{\lambda Y}{\lambda - Z} \quad \frac{\lambda Z}{\lambda - Z} \right]$$

其中 c 的前 2 项是 3-D 空间点 (X, Y, Z) 投影到图像平面后的坐标 (x, y) 。

逆投影变换将 1 个图像点反过来映射回 3-D 空间。(从略)

采样和量化的数学描述。用 Z 和 R 分别代表实整数集和实数集。采样过程可看作将图像平面化分成网络,每个网络中心点的位置由一对笛卡尔坐标所决定,他们是所有有序元素对 (a, b) 的集合,其中 a 和 b 是属于 Z 的整数。如果 x 和 y 是 Z 中的整数, $f(\cdot)$ 是给点对 (x, y) 赋予灰度值的函数,那么 $f(x, y)$ 就是一幅空间数字化的图像。这个赋值过程就是所谓的量化过程。如果灰度值也是整数,即 Z 代替了 R ,那么 $f(x, y)$ 就是一幅数字图像,且是坐标和灰度值都是整数的 2-D 函数。

图像的坐标变换^[4]。图像的坐标变换可采用统一的矩阵形式 $v' = Av$,其中 A 是一个 4×4 变换矩阵, v 是包含原坐标的矢量:

$$v = [X \ Y \ Z \ 1]^T$$

v' 是由变换后坐标组成的矢量:

$$v' = [X' \ Y' \ Z' \ 1]^T$$

平移变换。用这种表达法,平移矩阵可以写成:

$$T = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

平移过程可以表达为 $v' = Tv$ 。

缩放变换。同理,用这种表达法,用 S_x, S_y 和 S_z 沿 X, Y 和 Z 轴缩放变换矩阵可以写成:

$$S = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

缩放变换过程可以表达为 $v' = Sv$ 。

旋转变换。设旋转角是按从旋转角正向看原点而顺时针定义的,这样将一点绕 X 坐标轴旋转 α 角度可以用下列变换实现:

$$R_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

同理,将一点绕 Y 坐标轴旋转 β 角度可以用下列变换实现:

$$R_\beta = \begin{bmatrix} \cos \beta & 0 & -\sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

将一点绕 Z 坐标轴旋转 γ 角度可以用下列变换实现:

$$R_\gamma = \begin{bmatrix} \cos \gamma & \sin \gamma & 0 & 0 \\ -\sin \gamma & \cos \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

级连变换与反变换:同理可以推导出来。

3 图像运算^[4-5]

像素 p 具有坐标 (x, y) , 则其垂直和水平的 4 个邻近像素 $(x+1, y)$ 、 $(x-1, y)$ 、 $(x, y+1)$ 、 $(x, y-1)$ 称为像素 p 的 4-邻域, 记作 $N_4(p)$ 。像素 p 的 4 个对角邻近像素, 记作 $N_D(p)$ 。像素 p 的垂直、水平及对角 8 个邻近像素, 记作 $N_8(p)$ 。

我们用 V 表示定义连接的灰度值集合, 则有以下定义。

定义 1 4-连接 2 个像素 p 和 r 在 V 中取值且 r 在 $N_4(p)$ 中, 则称它们为 4-连接。

定义 2 8-连接 2 个像素 p 和 r 在 V 中取值且 r 在 $N_8(p)$ 中, 则称它们为 8-连接。

定义 3 m -连接(混合连接) 2 个像素 p 和 r 在 V 中取值且满足下列条件之一, 则称它们为 m -连接:

① r 在 $N_4(p)$ 中; ② r 在 $N_D(p)$ 且 $N_4(p) \cdot N_4(r)$ 是空集, 这个集合是由 p 和 r 的在 V 中取值的 4-近邻像素组成的。

同理, 可以定义 4-通路、8-通路或 m -通路。

等价关系 定义在集合 A 上的(二元)关系 R 可具有如下性质:

1) 反射性: 即对 A 中的每个 $a, a \sim R \sim a$ 成立;

2) 对称性: 即对 A 中的每个 a 和 b , 若 $a \sim R \sim b$ 成立, 则 $b \sim R \sim a$ 也成立;

3) 传递性: 即对 A 中的每个 a, b 和 c , 若 $a \sim R \sim b, b \sim R \sim c$ 成立, 则 $a \sim R \sim c$ 也成立。满足这 3 个性质的关系称为等价关系。

传递闭包 对于定义在集合 A 上的二元关系 R , 包含 R 中元素之间隐含关系的集合称为传递闭包。

距离度量函数 D 给定 3 个像素 $p(x, y)$ 、 $q(s, t)$ 、 $r(u, v)$, 如果满足下列条件, 则 D 是距离度量函数:

- ① $D(p, q) \geq 0$ ($D(p, q) = 0$ 当且仅当 $p = q$); ② $D(p, q) = D(q, p)$; ③ $D(p, r) \leq D(p, q) + D(q, r)$ 。

点 p 和 q 之间的欧氏 (Eudclidean) 距离 (也即模为 2 的距离) 定义为

$$D_E(p, q) = \sqrt{(x-s)^2 + (y-t)^2}$$

点 p 和 q 之间的 D_4 距离 (也即模为 1 的距离), 即城区 (city-block) 距离定义为

$$D_4(p, q) = |x-s| + |y-t|$$

点 p 和 q 之间的 D_8 距离 (也即模为 ∞ 的距离), 即棋盘 (chessboard) 距离定义为

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

图像的算术运算 图像处理中常用的算术运算一般用于灰度图像, 2 个像素 p 和 q 之间的算术运算有:

- 1) 加法: 记为 $p + q$;
- 2) 减法: 记为 $p - q$;
- 3) 乘法: 记为 $p * q$ (也可写为 $pq, p \times q$);
- 4) 除法: 记为 $p \div q$ 。

图像的加法主要用于图像平均以减少噪声。图像减法主要用于在医学成像中去除固定的背景信息。图像乘法和除法主要用来校正由于照明或传感器的非均匀性造成的图像灰度阴影。对于整幅图的算术运算是逐像素进行的。算术运算每次只涉及 1 个像素位置, 所以可以“原地”完成。

图像的逻辑运算 图像处理中常用的逻辑运算一般有:

- 1) 与 (AND): 记为 $p \text{ AND } q$ (也可写为 $p \cdot q$);
- 2) 或 (OR): 记为 $p \text{ OR } q$ (也可写为 $p + q$);
- 3) 补 (COMPLEMENT): 记为 $\text{NOT } q$ (也可写为 \bar{q} 或 $\sim q$)。

以上这些图像的逻辑运算是功能完备的, 即将它们组合起来可以进一步构成所有其它各种逻辑运算。与算术运算不同, 逻辑运算只用于二值图像。对于整幅图的逻辑运算是逐像素进行的。因为逻辑运算每次只涉及 1 个空间像素位置, 所以与算术运算类似地可以“原地”完成。

4 失真的度量方法

在图像编码系统中, 评估编码系统性能的一种方法是失真度量法, 用峰值信号噪声比 PSNR (Peak Signal to Noise Ratio) 来衡量^[6], 定义为最大像素值与均方差 MSE (Mean Square Error) 之比,

$$PSNR = 10 \log_{10} \frac{(PSV)^2}{MSE} \quad (\text{dB})$$

其中 PSV 为信号峰值 (Peak Signal Value)。

对 8 位二进制图像,

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (\text{dB})$$

其中 MSE 由下式给出,

$$MSE = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |x(m, n) - \bar{x}(m, n)|^2$$

这里, $x(m, n)$ 为原始图像的像素值, $\bar{x}(m, n)$ 为解压缩之后的像素值。

在文献 [7] 中, 评估编码系统性能还使用其他方法, 这些方法包括使用规格化均方差 NMSE (Normalized Mean Square Error)、信号噪声比 SNR (Signal to Noise Ratio) 和平均绝对误差 MAE (Mean Absolute Error) 来度量, 分别定义为^[8]

$$NMSE = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [x(m,n) - \tilde{x}(m,n)]^2}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [x(m,n)]^2}, \quad SNR = 10 \log_{10} \left(\frac{1}{NMSE} \right)$$

$$MAE = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [x(m,n) - \tilde{x}(m,n)]$$

其中, $x(m, n)$ 为原始图像的像素值, $\tilde{x}(m, n)$ 为解压缩之后的像素值。

在电子工程中, 信号噪声比(SNR)一直是最流行的误差度量指标, 在大多数情况下可提供很有价值的信息, 在数学上也比较容易计算。信号噪声比虽然也用在图像编码中, 但由于它的数值与图像编码系统中高压缩比的关系不容易体现, 因此提出了其他的几种度量方法, 包括平均主观评分 MOS (Mean Opinion Score) 等。

5 结束语

图像工程作为一门新兴的综合学科, 其鲜明的特点是内容多、跨度大、覆盖面广, 涉及到图像代数——数学形态学(Morphology)及模式识别、ANN、FFT、小波理论等众多相关知识。本文仅对图像工程领域里几个不够统一的基本概念、术语给出明确的解释或定义, 还对图像处理的基本算法模型及图像运算等内容进行了概括, 最后描述了图像编码系统中失真的度量方法, 以期澄清图像工程的若干基本概念、揭示图像工程的理论基础。

参考文献:

- [1] Kastinaki V, Zervakis M, Kalaitzakis K. A survey of video processing techniques for traffic applications [J]. Image and Vision Computing, 2003, 21: 359 - 381.
- [2] Peter Bergstrom. Eye - movement - controlled transform image coders [J]. Signal Processing: Image Communication, 2003, 18: 115 - 125.
- [3] Chi - Man Pun. Rotation - invariant texture feature for image retrieval [J]. Computer Vision and Image Understanding, 2003, 89: 24 - 43.
- [4] 章毓晋. 图像工程(上册)——图像分析与处理[M]. 北京:清华大学出版社, 1999.
- [5] 章毓晋. 图像工程(下册)——图像理解与计算机视觉[M]. 北京:清华大学出版社, 2000.
- [6] Kenneth R Castleman. 数字图像处理[M]. 朱志刚. 北京:电子工业出版社, 2002.
- [7] Paul Kwan, Keisuke Kameyama, Kazuo Toraichi. On a relaxation - labeling algorithm for real - time contour - based image similarity retrieval [J]. Image and Vision Computing, 2003, 21: 285 - 294.
- [8] Ramchandran K, Vetterli M, Herley C. Wavelets, subband coding, and best bases [A]. Proceedings of the IEEE[C]. 1996, 541 - 560.

(编辑:田新华)

Fundamental Concepts and Theoretical Foundations of Image Engineering

LEI Ying - jie, ZHENG Quan - di, ZHANG Gang, ZHOU Chuang - ming

(The Missile Institute, Air Force Engineering University, Sanyuan, Shaanxi 713800, China)

Abstract: In this paper, first, the fundamental concepts and glossary of terms with vague or arbitrary meanings in the area of image engineering are defined and explained in clarification. Then, the contents such as basic algorithm models in image processing and image operations are generalized. Finally, the methods of measuring distortion in image processing are described. In this way, a clarification of several basic concepts on image engineering and an exposure to theoretical foundations of it are anticipated.

Key words: image engineering; image processing; image coding; image operations