

基于图割的矩形目标交互式分割方法

郎咸朋^{1a,1b,2,3}, 朱 枫^{1a,1b,3}

(1. 中国科学院 a. 沈阳自动化研究所; b. 光电信息处理重点实验室, 沈阳 110016; 2. 中国科学院研究生院, 北京 100049; 3. 辽宁省图像理解与视觉计算重点实验室, 沈阳 110016)

摘 要: 为得到更加精确的图像分割结果, 可以在基于图割的分割框架中引入形状先验指导分割。针对矩形目标提出了一种目标/背景交互式分割方法。分割能量用马尔可夫随机场最大后验概率描述, 通过限制有向图中的流向引入形状先验, 可以保证图割优化后的分割结果为矩形形状。对仿真图像与真实图像的实验结果证明了该算法的有效性。

关键词: 图像分割; 图割优化; 形状先验; 马尔可夫随机场

中图分类号: TP29

文献标志码: A

文章编号: 1001-3695(2010)11-4369-04

doi:10.3969/j.issn.1001-3695.2010.11.105

Interactive segmentation of rectangle object via graph cut

LANG Xian-peng^{1a,1b,2,3}, ZHU Feng^{1a,1b,3}

(1. a. Shenyang Institute of Automation, b. Key Laboratory of Optical-Electronics Information Processing, Chinese Academy of Sciences, Shenyang 110016, China; 2. Graduate School of Chinese Academy of Sciences, Beijing 100049, China; 3. Key Laboratory of Image Understanding & Vision Computation, Shenyang 110016, China)

Abstract: In order to obtain more robust result, shape prior is incorporated into segmentation framework as an efficient constrain. This paper proposed a novel object/background segmentation algorithm for rectangle shaped objects. The rectangle shape prior was imposed by restricting flows towards valid directions in the digraph. The segmentation energy was formulated by MAP-MRF. Firstly introduced the concepts of LineCut and ParaCut, and then described RectCut in detail. Experimental results demonstrate the effectiveness of the method.

Key words: image segmentation; graph cut; shape prior; MRF

0 引言

图像分割是计算机视觉领域的重要研究内容, 在过去几十年中已经提出了许多图像分割方法^[1,2], 而基于图割全局优化方法的分割方法在近几年受到了广泛的关注。将视觉问题与图割方法相结合的开创性工作是由 Greig 等人^[3]在 1989 年完成的。他们在马尔可夫随机场中建立了标号能量与图割的对应关系, 使用最大流算法求解能量极小值, 进行二值图像的复原。由于所提出的算法效率较低, 而且二值图像复原的应用范围较为狭窄, 他们的工作在很长的一段时间内并没有得到足够的重视。本世纪初, Boykov 等人^[4]在基于增广路径的最大流算法基础上提出了更为高效的求解最小割的算法, 并将其应用到视觉领域的许多方面, 如交互式分割、立体匹配、纹理合成等。在这些工作的支持下, 目前图割方法已经在视觉领域得到了广泛的应用。

为了得到更加鲁棒的分割性能, 许多学者试图将形状先验引入到图割分割过程中。这种将高层先验引入到底层分割的做法也是近年来的新趋势。特定目标形状和几何图形最先被考虑成形状先验。Freedman 等人^[5]将目标轮廓的水平集函数作为形状模板参与到图割边权值的计算中, 起到约束分割形状的目的。然而这种形状先验必须事先与分割区域仔细配准。

若要避免手动配准模板的工作, 也可以迭代估计配准参数^[6]。但这种做法可能会导致原本是全局优化的分割方法退化为局部寻优方法, 从而抵消了算法本身的优势。另外, 文献[7]介绍了一种在序列图像中使用特定形状先验分割目标的方法, 除第一帧中需要交互式操作获得初始分割结果外, 序列中其余图像都可自动分割。

与特定形状先验不同, 使用广义形状先验可以表示一类具有某种共同性质的形状。目前已知最早应用在图割分割中的广义形状先验是文献[8]中提出的紧凑形状先验, 它被用做晶体管门电路图像的交互式分割。作者定义了一类称为紧凑形状的先验, 使用者只要用鼠标在待分割目标内部提供一个点就可以获得令人满意的分割结果。尽管如此, 紧凑先验的定义所包含的形状还是有一定的局限性, 而且用户给出的点的位置也对分割结果有影响。最近, Veksler^[9]提出了一个新的广义形状先验, 称为星型先验。它可以表示比紧凑先验更为一般的形状, 而且同样只需提供一个内部点就满足分割需要。特别是对于凸形状, 这个点可以选在内部任意位置而不会对最终分割结果产生影响。但文献[9]中星型先验的计算方法略显复杂, 需要计算图像中过内部点的所有离散直线。Lang 等人^[10]对其提出了改进, 在简化计算步骤的同时保证了与原算法分割效果的一致性, 而且可以处理目标具有复杂纹理的情况。

收稿日期: 2010-04-26; 修回日期: 2010-06-24

作者简介: 郎咸朋 (1981-), 男, 山东青岛人, 博士研究生, 主要研究方向为图像分割、计算机视觉 (xplang@sia.ac.cn); 朱枫 (1962-), 男, 研究员, 博士, 主要研究方向为计算机视觉、图像处理、智能系统等。

上述方法虽然能够很好地解决许多分割任务,但有时会遇到特殊的交互式分割问题。待分割的目标具有简单而明确的形状(如矩形),且进行分割之前形状是已知的。在这种情况下,往往需要一个更为精确的分割结果。据笔者了解,目前还没有针对矩形形状先验设计的图割分割算法。本文提出了一种交互式分割矩形目标的方法。形状先验主要通过限制最大流在有向图中的流向引入。实验结果证明了算法的有效性。

1 基础知识

本章介绍了有向图以及图割分割算法的术语和基本概念。对这些内容更为详细的阐述请参考文献[11,12]。

1.1 图割

有向图 G 通常可表示为 $G=(V,E)$ 。其中: V 代表节点集合, E 代表有向边的集合。对于目标/背景分割而言, V 中包含两类节点:相邻节点和两个端点。相邻节点一般对应二维图像中的像素(pixel)、三维图像中的体素(voxel),或者是其他一些特征。而两个端点用符号 s 和 t 表示,对应有向图中的源点和汇点,分别代表目标和背景标号。因此,只包含 s 和 t 两个端点的有向图也叫做 $s-t$ 图。 V 中的节点对通过 E 中的边相连。对应地,集合 E 中也包含两种类型的有向边。连接端点与普通节点的边称为 $t-link$,而连接两个普通节点的边称为 $n-link$ 。与源点相连的 $t-link$ 方向一定是从源点指向普通节点的,而与汇点相连的 $t-link$ 的方向则恰好相反,是由普通节点指向汇点。 $n-link$ 方向可以是双向的。需要注意的是,没有将两个端点直接相连的边,但可以由多条 $t-link$ 与 $n-link$ 按照边的指向构成从 s 到 t 的路径。每条边都有一个非负的边权值,在最大流计算中,这个权值也称为边的容量。这样就完成了 $s-t$ 图的构建。

图的割用符号 C 表示,它是边 E 的一个子集,作用是将 V 中所有的节点分成两个不相交的子集,并且源点和汇点分别属于不同的子集。将源点所在的子集记为 S ,汇点所在的子集记为 T ,则有 $s \in S, t \in T$ 且 $S, T \subset V$ 。换言之,从图 G 中去掉 C 中所包含的边后,就找不到任何从 s 到 t 的路径了。在 $s-t$ 图上的割也称为 $s-t$ 割。本文中所涉及的割均为 $s-t$ 割。割的代价定义为 C 中所有边的权值之和,这些边通常也称为边界线。通过边界线相连的两个节点在图割后属于不同的集合。假设连接节点 p 和 q 的边是边界线,则 $p \in S, q \in T$,或者相反。需要注意的是,计算割的代价时要考虑边的方向。它只统计两个节点之间从 S 指向 T 的有向边的权值。对于一个图而言,会存在许多可能的割。这些割中总会有一个代价最小的割,称之为最小割。根据 Ford 等人^[13]的理论,最小割与最大流是等价的,因此可以用最大流方法求解最小割。

求解最大流/最小割的算法有很多,这些算法大致可分为两类:增广路径算法(augmenting paths)和推进—重标号(push-relabel)算法。增广路径算法的思想是将水流沿着非饱和边不断地从源点推向汇点,直到图中水流值达到最大。如果每次迭代时都选择从 s 到 t 的最短路径作为增广路径,那么算法的复杂性为 $O(nm^2)$ 。其中 m 与 n 分别为图中边和节点的数目。推进—重标号算法也叫做前置流推进算法。与增广路径方法不同,它试图将尽可能多的水流一齐推向汇点 t ,最后将剩余的水流回流到源点 s 。推进—重标号算法的复杂性也为

$O(nm^2)$ 。通过控制算法的运行方向可以改进推进—重标号算法。例如,如果用先入先出(FIFO)的策略检查节点,可以得到一个 $O(n^3)$ 算法。增广路径算法在寻找从 s 到 t 的路径时,一旦所有给定长度的路径都遍历过了,就需要开始一个新的广度优先搜索。但是建立一个广度优先搜索树意味着要扫描大多数的图像像素,如果要经常执行这个操作,将是一个代价非常高的操作。重新建立图的搜索树,使得增广路径算法在实际应用中的性能非常差。针对原始算法存在的问题,Boykov 等人^[14]提出了一个新的基于增广路径的算法。新算法分别以源点和汇点为根节点,建立两个互不重叠的搜索树。由于新算法在增广路径时不能保证是最短路径,这就导致最坏情况下算法复杂性是 $O(n^2 m |C|)$ 。其中 $|C|$ 表示最小割的代价。从理论上说,这比之前算法的复杂性要高,但最坏情况并不是总会出现的,在实际应用中新算法的表现也远好于其他算法。

图 1 是一个简单的 $s-t$ 图^[14]。节点集合由九个普通节点和两个端点组成,普通节点、源点和汇点在图中分别用灰色、黑色和白色圆点表示。与源点相连的 $t-link$ 为黑色,与汇点相连的 $t-link$ 为白色,灰色圆点之间的浅色边表示 $n-link$,箭头的朝向代表每条边的方向,边的粗细表示其边权值的大小。有向图的一个割用虚线表示,它将所有节点分成了两个子集。

1.2 基于图割的图像分割

图像分割可以看做标号问题^[15]。假设有像素集合 P 和标号集合 L ,像素标号就是要找到一个映射 $f: P \rightarrow L$,使得 P 中的每个元素在标号后都能在其中找到一个对应的标号值。寻找映射 f 的过程就叫做标号过程,一般是通过最小化某个能量函数来实现的。假设有 m 个元素的离散标号集合 $L = \{l_1, l_2, \dots, l_m\}$ 和包含 n 个像素的图像 $P = \{p_1, p_2, \dots, p_n\}$,则图割分割的目的就是将 P 中的每个像素 p 赋予一个标号 l 。在目标/背景分割中,标号集合只包含两个元素:0 和 1,分别对应背景和目標。马尔可夫随机场为解决标号问题提供了一个概率的框架。令离散随机变量集合 $X = \{x_1, x_2, \dots, x_n\}$ 与 P 中像素一一对应, X 中每个元素都以某种概率从 L 中获得标号 f_x ,则 X 为一个随机场, $F = \{f_{x_1}, f_{x_2}, \dots, f_{x_n}\}$ 为它的标号组合。给定一个邻域系统 N_v ,马尔可夫性质可以描述为

$$Pr(x_v | \{x_u : u \in P - \{v\}\}) = Pr(x_v | \{x_u : u \in N_v\}) \quad \forall v, v \in P \quad (1)$$

其中: N_v 表示邻域系统,本文中使用的 4 邻域系统。这样一来,马尔可夫随机场 X 中每一种标号组合都对应一个分割结果。当获得观测数据 D 后,就可以对每一种标号组合计算后验概率,概率最大的组合就认为是最佳分割,这个过程就是求解马尔可夫场最大后验概率(MAP-MRF)的过程。根据贝叶斯法则,后验概率正比于先验概率与似然函数的乘积。将后验概率取负对数,就得到了图割分割能量的一般形式:

$$E(F) = (1 - \lambda) \sum_{p \in P} D_p(f_p) + \lambda \sum_{p \in P, q \in N_p} V_{pq}(f_p, f_q) \quad (2)$$

能量函数由两项构成:第一项 D_p 是数据项,表示将标号 f_p 分配给像素 p 的惩罚;另一项 V_{pq} 是边界项,用来惩罚相邻像素之间标号不同的情况。有些文献也把 V_{pq} 称为平滑项。参数 $\lambda \geq 0$ 描述数据项与边界项的相对重要程度。当能量的交互项 V_{pq} 是一个次模(submodule)函数时,该能量就可以用图割方法优化^[16]。对于本文所研究的目标/背景分割而言,次模性质需要满足以下不等式:

$$V(0,0) + V(1,1) \leq V(0,1) + V(1,0) \quad (3)$$

通常 $V_{pq} = \exp(-(|I_p - I_q|^2 / \sigma^2)) |f_p - f_q|$ 。其中: I_p 与 I_q 分别为像素 p, q 对应的灰度值, σ 为可调参数。可以看出, 当这两个像素分配的标号 $f_p = f_q$ 时, $V_{pq} = 0$; 当 $f_p \neq f_q$ 时, $V_{pq} \geq 0$, 所以满足式(3)的次模性质。

2 算法

本章详细介绍将矩形形状先验结合到图割框架中的方法。首先定义了直线割的概念, 使分割边界为一条直线; 其次定义了平行割, 可以用一对平行线将图像分成三部分; 最后, 在前述两者与星型形状先验的基础上提出了矩形割。

2.1 直线割

使用传统的图割分割方法可以获得具有任意拓扑结构的分割结果, 所得分割结果的形状仅仅取决于有向图中 n -link 与 t -link 的边权值。考虑如图 2 所示的图, 圆点表示图的节点, 其中实心的浅色点表示源点种子点, 实心深色点表示汇点种子点。将源点与汇点种子点的 t -link 权值设为无穷大, 其他节点不计算 t -link 权值, 这样的处理等价于将式(2)中的 λ 设为 1。对 n -link 计算权值后, 分割结果可以得到一个最小割 C (虚线)。根据这个结果, 图像中的像素被分为两组: 浅色阴影点属于前景类, 而深色阴影点属于背景类。但在最小割算法结束前是无法事先得知分割边界形状的。

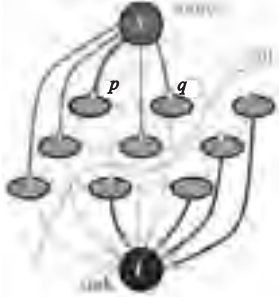


图1 s-t图



图2 传统图割分割结果

受星型先验的启发, 可以通过控制最大流的流向实现对分割边界形状的约束。图 3 展示了如何使分割边界为一条垂直线。关于种子点和节点的定义是与左边图像中相同的。但所有的垂直方向的 n -link 边权值都被强制为无穷大, 对于水平方向且朝向源点种子点的 n -link 也作相同的处理。这些被赋予无穷大边权值的 n -link 在右边图像中表示为黑色箭头。如果对这样一个图进行图割, 可以确保割的边界是一条直线。这是由于为了获得一个最小的代价, 最小割的结果必定不会包含任意一条具有无穷大边权值的 n -link, 否则就不能称为最小割。这样一来, 就可以得到一个具有直线形状边界的分割结果, 称之为直线割方法。

给定一对相邻像素 p 和 q , 根据两者的相对位置, 图 3 所示的直线形状先验 $S_L(f_p, f_q)$ 可表示为:

a) 当 p 在 q 左侧时,

$$S_L(f_p, f_q) = \begin{cases} 0 & \text{if } f_p = f_q \\ 0 & \text{if } f_p = 1 \text{ and } f_q = 0 \\ \infty & \text{if } f_p = 0 \text{ and } f_q = 1 \end{cases} \quad (4)$$

b) 当 p 在 q 下方时,

$$S_L(f_p, f_q) = \begin{cases} 0 & \text{if } f_p = f_q \\ \infty & \text{if } f_p \neq f_q \end{cases} \quad (5)$$

于是, 直线割的能量表达式可以描述为

$$E(f) = \sum_{p \in \text{seed}} D_p(f_p) + \sum_{p \in P, q \in N_p} [V_{pq}(f_p, f_q) + S_L(f_p, f_q)] \quad (6)$$

该能量与式(2)表示的传统图割能量有三处不同: a) 数据项限制为只有种子点参与计算; b) 增加了直线形状先验项; c) 由于种子点的 t -link 均已设置为无穷大, 去掉了参数 λ 。

2.2 平行割

若要利用一对平行线作为分割边界将图像分割成为三个部分, 则可以同时用两个直线割来达到这个目的。如图 4 所示, 对应于图像左右边缘处像素的节点被设置为汇点种子点, 而用某些位于图像中间的像素对应的节点为源点种子点。利用垂直方向直线形状先验在源点两侧分别设置好边权值, 对这个图进行最小割, 就会得到平行的分割边界, 称之为平行割。

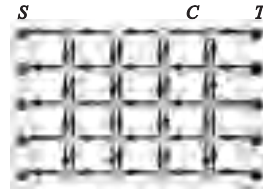


图3 直线割示意图

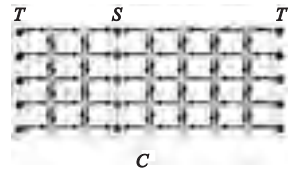


图4 平行割示意图

给定一对相邻像素 p 和 q , 根据两者的相对位置, 图 4 所示的平行形状先验 $S_P(f_p, f_q)$ 可以表示为:

a) 当 p 与 q 都在源点左侧且 p 在 q 的右侧时,

$$S_P(f_p, f_q) = \begin{cases} 0 & \text{if } f_p = f_q \\ 0 & \text{if } f_p = 1 \text{ and } f_q = 0 \\ \infty & \text{if } f_p = 0 \text{ and } f_q = 1 \end{cases} \quad (7)$$

b) 当 p 在 q 的左侧且 p 与 q 都在源点右侧时,

$$S_P(f_p, f_q) = \begin{cases} 0 & \text{if } f_p = f_q \\ 0 & \text{if } f_p = 1 \text{ and } f_q = 0 \\ \infty & \text{if } f_p = 0 \text{ and } f_q = 1 \end{cases} \quad (8)$$

c) 当 p 在 q 的下方时,

$$S_P(f_p, f_q) = \begin{cases} 0 & \text{if } f_p = f_q \\ \infty & \text{if } f_p \neq f_q \end{cases} \quad (9)$$

于是, 平行割的能量表达式为

$$E(f) = \sum_{p \in \text{seed}} D_p(f_p) + \sum_{p \in P, q \in N_p} [V_{pq}(f_p, f_q) + S_P(f_p, f_q)] \quad (10)$$

以上介绍了分割边界为垂直方向时直线割与平行割的定义, 同理也可得到分割边界为水平方向时的定义。

2.3 矩形割

本节在平行割的基础上给出矩形割算法。该算法需要使用预先在待分割目标内部提供一个矩形作为初始输入, 因此是一个交互式的分割方法。但算法对所输入的矩形没有限制, 可以是任意大小和长宽比。根据这个初始矩形, 可以将整个图像按图 5 所示进行区域划分。中间的黑色方框代表初始输入矩形, 延长这个矩形的四条边, 将图像分为了用罗马数字标记的九个区域。

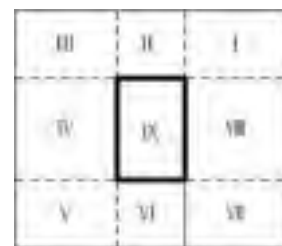


图5 矩形割区域划分

下面解释矩形形状先验的构造。首先, 将区域 IX 中的像

素作为源点种子点,图像边界作为汇点种子点。其次,对于Ⅱ、Ⅳ、Ⅵ、Ⅷ四个区域分别作平行割处理。最后,对于剩余的Ⅰ、Ⅲ、Ⅴ、Ⅶ四个区域分别作星型先验处理。这样,矩形形状先验 $S_R(f_p, f_q)$ 可以表示为

$$S_R(f_p, f_q) = \sum_{\text{region}=2,4,6,8} \sum_{p \in P, q \in N_p} S_P(f_p, f_q) + \sum_{\text{region}=1,3,5,7} \sum_{p \in P, q \in N_p} S_{\text{star}}(f_p, f_q) \tag{11}$$

其中: S_P 为平行形状先验, S_{star} 为星型形状先验。关于星型形状先验详细介绍请参考文献[15],这里不再赘述。综上,矩形割的能量表达式为

$$E(f) = \sum_{p \in \text{seed}} D_p(f_p) + \sum_{p \in P, q \in N_p} [V_{pq}(f_p, f_q) + S_R(f_p, f_q)] \tag{12}$$

取式(12)分割结果的外接矩形,就得到了矩形割最终的分割结果。在1.2节中曾经提到,对于目标/背景分割,能够用图割优化的能量函数交互项必须满足式(3)所示的次模性质。下面就证明式(12)所示矩形割能量的交互项, $[V_{pq}(f_p, f_q) + S_R(f_p, f_q)]$ 满足这个性质,而且该过程也隐含证明了式(6)和(10)的次模性质。

证明

a)当 region = 1,3,5,7 时,
式(12)变为星型先验能量,由文献[15]可知该能量为次模函数。

b)当 region = 2,4,6,8 时,
 $S_R(f_p, f_q) = S_P(f_p, f_q)$,根据式(7)~(9),
若 $f_p = f_q = 0$ 或 $f_p = f_q = 1$, $S_R(f_p, f_q) = S_P(f_p, f_q) = 0$;
若 $f_p \neq f_q$, $S_R(f_p, f_q) = S_P(f_p, f_q) \geq 0$ 。
令 $V_{pq} = \exp(-(I_p - I_q)^2 / \sigma^2) |f_p - f_q|$,有
 $[V_{pq}(0,0) + S_R(0,0)] + [V_{pq}(1,1) + S_R(1,1)] = 0$
 $[V_{pq}(0,1) + S_R(0,1)] + [V_{pq}(1,0) + S_R(1,0)] \geq 0$
则
 $[V_{pq}(0,0) + S_R(0,0)] + [V_{pq}(1,1) + S_R(1,1)] \leq [V_{pq}(0,1) + S_R(0,1)] + [V_{pq}(1,0) + S_R(1,0)]$

因此式(12)表示的能量满足式(3)所示的次模性质,该能量为次模函数。

3 实验结果

针对第2章中提出的各个算法,本文分别进行了相应的实验。实验所用计算机配置为 Intel 2.83 GHz CPU, 2 GB RAM。使用 VC++2005 完成整个实验代码的编写,其中最大流/最小割优化算法利用了 Boykov 等人公布的函数。

图6是直线割的实验结果。实验图像包含两个含有噪声的区域,两者的边界为一条直线,但有三处断裂。左图是使用原始图割分割方法所得到的结果,右图是加入直线形状先验后得到的结果。除了是否使用形状先验之外,这两幅图在分割过程中所使用的种子点以及参数都完全相同。

图7是平行割的实验结果。与直线割实验类似,在实验图像中同样存在干扰和噪声。左图为原始图割分割方法的结果,右图为加入平行形状先验的分割结果。



图6 直线割实验结果



图7 平行割实验结果

矩形割的分割结果如图8所示。左面一列是原始图像,右边是对应的矩形割分割结果。第一行是存在噪声和干扰的矩形图像。第二行右侧图像中建筑物的整体轮廓为一矩形,但有些部分已被树木和其他背景物体遮挡住。总之,这两幅图像中的待分割目标都受到了不同程度的污染,用通常的分割方法很难得到期望的分割结果,但由于在分割之前已知这些目标都是矩形形状,应用本文提出的矩形割算法进行分割可以得到期望的结果。需要特别指出的是,使用本文所提出的矩形割方法只能保证得到的分割结果是一个规则的矩形,这适用于待分割目标具有矩形形状且没有发生旋转和强烈形变的情况。

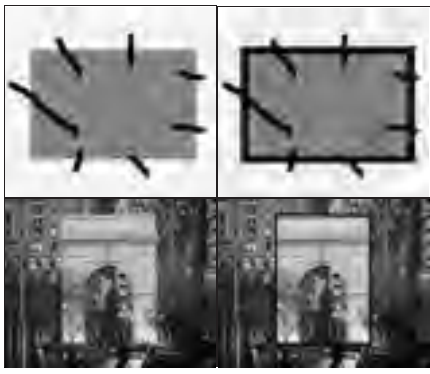


图8 矩形割实验结果

4 结束语

近年来,图割作为一种全局优化技术在计算机视觉领域受到了广泛的关注。在图割框架下,交互式分割方法得到了迅速的发展,出现了结合形状先验的分割方法,使得分割结果更加准确。本文针对矩形目标提出了一种新颖的交互式分割方法。实验结果表明,该方法能够获得期望的分割效果,并且交互操作较为简便,只需在待分割目标内部任意输入一个矩形即可。但是目前本文算法还不能处理矩形目标发生旋转的情况,未来将针对此方面开展进一步的研究工作。

参考文献:

[1] 章毓晋. 图像工程[M]. 2版. 北京: 清华大学出版社, 2005.
[2] 傅冠华,陈益强,刘军发,等. 一种自动分割人脸的方法[J]. 计算机应用研究,2008,25(10): 3060-3062.
[3] GREIG D M, PORTEOUS B T, SEHEULT A H. Exact maximum a posteriori estimation for binary images[J]. Journal of the Royal Statistical Society, 1989, 51(2): 271-279.
[4] BOYKOV Y, JOLLY M. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images[C]//Proc of the 8th IEEE International Conference on Computer Vision. 2001:105-112.
[5] FREEDMAN D, ZHANG Tao. Interactive graph cut based segmentation with shape priors[C]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2005:755-762.
[6] CHANG Hang, YANG Qing, PARVIN B. A Bayesian approach for image segmentation with shape priors[C]//Proc of the 26th IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2008:1-8.
[7] LANG X Pian-peng, ZHU Feng, HAO Ying-ming, et al. Automatic image segmentation incorporating shape priors via graph cuts[C]//Proc of IEEE International Conference on Information and Automation. Piscataway: IEEE Computer Society, 2009:192-195.

和缩放。旋转实验将图像从 0 ~ 360°依次旋转 5°,缩放从 0.5 ~ 3 倍依次增加 0.1,得到实验数据如表 1 所示。

表 1 特征的旋转缩放稳定性

类	项目	λ_1	λ_2	λ_3	λ_4
缩放	均值	-0.768	-0.471	-0.223	1.463
	均方差 E-3	3.945	4.676	1.519	0.638
	比值 E-3	5.133	9.929	6.825	0.436
	均值	-0.768	-0.471	-0.223	1.463
旋转	均方差 E-3	6.567	7.067	1.422	1.114
	比值 E-3	8.548	14.99	6.384	0.762

从表 1 可以看出,所提取的特征在旋转和缩放下保持了极好的稳定性,变异系数均在 0.1% 左右。

3.2 相似性度量对比实验

为了简化讨论而又不失本质,本节对比实验中采用了图 2 所示的四张图像(注:图像中白点表示图像四个对象的质心,为了便于读者清晰看出图像对象的位置,图片中将白点做了加粗处理以示强调,图像中每个白点代表一个对象)。



图1 算法基本原理

图2 相似性度量对比实验图像

从人的主观视觉心理上,这四幅图片中图 2(b)(c)(d)与(a)所表达的四个对象空间关系相似性(而不是图像本身相似性)越来越低。

所提取的特征实验数据如表 2 所示。

表 2 实验图像的特征向量

图像	λ_1	λ_2	λ_3	λ_4
(a)	-0.708 0	-0.516 7	-0.248 4	1.473 1
(b)	-0.719 6	-0.506 8	-0.245 1	1.471 5
(c)	-0.808 7	-0.437 3	-0.207 2	1.453 2
(d)	-0.966 5	-0.302 5	-0.130 1	1.399 1

由表 2 可以看出,对于包含若干对象的图像,当图像不相似时,其特征向量处于特征空间相对分散的位置,就是说类间相对分散,类内相对聚近。图像相似性度量矩阵如表 3 所示。

表 3 图像相似性度量矩阵

图像	d			
(a)	0	0.000 04	0.003 10	0.022 00
(b)	0.000 04	0	0.002 40	0.020 20
(c)	0.003 10	0.002 40	0	0.008 70
(d)	0.022 00	0.020 20	0.008 70	0

表 3 中的距离定义为 $d(x,y) = 1 - \cos(x,y)$ 。由表 3 可以看出,图 2(b)(c)(d)与(a)的距离逐渐增大,所以(b)与(a)最相似,这与人主观视觉心理一致。

(上接第 4372 页)

[8] DAS P, VEKSLER O, ZAVADSKY V, et al. Semiautomatic segmentation with compact shape prior[J]. *Image and Vision Computing*, 2009,27(1-2):206-219.

[9] VEKSLER O. Star shape prior for graph-cut image segmentation [C]//Proc of the 10th European Conference on Computer Vision: Part III. Berlin: Springer-Verlag, 2008:454-467.

[10] LANG Xian-peng, ZHU Feng. Image segmentation with improved star shape prior via graph cut [C]//Proc of the 9th Asian Conference on Computer Vision. 2008.

[11] BOYKOV Y, FUNKA-LEA G, ZAVADSKY V, et al. Graph cuts and efficient N-D image segmentation[J]. *Journal of Computer Vision*, 2006,70(2):109-131.

3.3 类别可分离性实验与空间关系粗分类

为了能够更直观更清晰地观察类间分散类内聚集效果,仍然以图 2 中选用的图片,将每幅图片看做一类别,将图像从 0 ~ 360°依次旋转 5°,缩放从 0.5 ~ 3 倍依次增加 0.1,以 (λ_1, λ_2) 作为特征,画出四个类别的散点图,如图 3 所示。

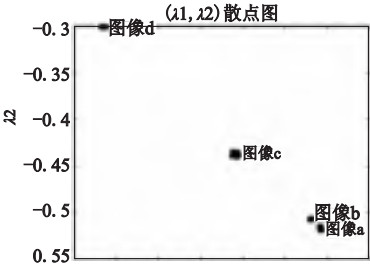


图3 表征类别分离性的散点图

由图 3 可以看出,类内相对聚合(其分散程度在 0.1% 数量级),类间相对分散,这个特点有利于图像的模式识别。图 2(a)与(b)很相似,所以它们的特征向量非常接近。

4 结束语

本文构造了基于度量矩阵特征值的对象空间关系描述子,在平移、缩放、旋转下保持了极好的稳定性,用于描述图像时与人的主观视觉心理基本一致,用于图像识别、图像检索、图像配准可以取得较好的效果。

但是本文的算法仍然存在需要改进的地方,主要体现在:算法效果与图像分割的质量密切相关,如果图像受噪声污染后造成图像分割不理想,则算法的效果就不甚理想。但这个问题是所有空间关系描述算法都必须面临的问题。而且图像分割不是本文讨论的范围,是图像分析的一个专门分支领域。

参考文献:

[1] 周明全,耿国华,韦娜. 基于内容图像检索技术[M]. 北京:清华大学出版社,2007.

[2] 章毓晋. 基于内容的视觉信息检索[M]. 北京:科学出版社,2003.

[3] CHANG Shi-kuo, SHI Qing-yun, YAN Cheng-wen. Iconic indexing by 2-D strings [J]. *IEEE Trans on Pattern Anal Mach Intell* PAMI,1987,9(3):413-428.

[4] HOU Tai-yuan, HSU A, LUI Pei-ya, et al. A content-based indexing technique using relative geometry features [C]//Proc of SPIE- The International Society for Optical Engineering. 1992:59-68.

[5] GUDIVADA V N, RAGHAVAN V V. Design and evaluation of algorithms for image retrieval by spatial similarity [J]. *ACM Trans on Information Systems*,1995,13(2):115-144.

[12] 邦詹森 J, 古廷 G. 有向图的理论、算法及其应用[M]. 姚兵,张忠辅,译. 北京:科学出版社,2009.

[13] FORD L R, FULKERSON D R. Flows in networks[M]. Princeton: Princeton University Press, 1962.

[14] BOYKOV Y, KOLMOGOROV V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*,2004,26(9):1124-1137.

[15] KOHLI P, TORR P H. Dynamic graph cuts for efficient inference in Markov random fields [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*,2007,29(12):2079-2088.

[16] KOLMOGOROV V, ZABIH R. What energy function can be minimized via graph cuts [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2004,26(2):147-159.