

基于空间关系的子区域最佳匹配算法

徐 丹 章毓晋

(清华大学电子工程系 北京 100084)

摘 要: 该文提出了一种基于空间关系的图像检索算法——子区域最佳匹配法。这种算法首先根据符号投影模型,将图像内目标抽象为带有多个属性的图标。然后,利用搜索邻节点的等角度区域最近原则、划分子区域的中心点法、图论中的广度优先算法进行图标匹配。最后,根据所定义的局部区域相似距离及图间距离来计算图像与图像之间的相似度,进而完成基于图像内部空间关系的检索。该文以这种算法为核心,构建了一个篮球站位分析演示系统,并用实际数据验证了所提出的基于空间关系的图像检索方法的有效性。

关键词: 基于内容的图像检索, 空间关系, 子区域划分, 遍历, 用户定制

中图分类号: TP391 **文献标识码:** A **文章编号:** 1009-5896(2005)08-1177-05

Optimal Matching Based on Spatial Relationship in Local Regions

Xu Dan Zhang Yu-Jin

(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

Abstract A new algorithm named optimal matching based on spatial relationship in local regions for content-based image retrieval is proposed in this paper. The algorithm first employs the symbol mapping model to reduce a symbolic image to a 1-D representation, and then uses the principle of nearest distance within equal angle region for searching nearest nodes, takes the center point for dividing images into sub-regions, and applies the breadth-first traversal algorithm in graph theory to perform graph matching. Finally, it uses the defined similar distances between local regions and among different graphs to compute the similarity and to retrieve images based on inner spatial relationship. Based on these algorithms, a prototype system for analyzing the position of players in basketball game is constructed. The effectiveness of image retrieval based on spatial relationship is verified in this system by using some real data.

Key words Content-Based Image Retrieval (CBIR), Spatial relationship, Sub-region division, Traversal, User customization

1 引言

近年来,基于内容的图像检索是一个受关注的研究课题^[1],它主要解决怎样从图像库中选出人们想要的图片。基于内容的含义主要是指用图像本身的视觉信息,包括颜色、纹理、形状和空间关系等。其中空间关系是图像内部目标的重要特征^[2],表达了它们之间的相对位置和联系,比较适合在图像含有较多独立部分以及检索结果强调这些部分的位置关系时使用。实际中,像篮球、足球这样体现协作能力和团队精神的体育运动中最重要的就是球员的站位^[3],室内装潢和城市规划中也涉及到各个部分的空间排列,越来越被人所关注的虚拟三维技术中也涉及到虚拟空间形体的相对位置问题。

基于空间关系的图像检索是基于内容的图像检索中的重要内容(草图查询也利用了空间关系^[4])。以前有一些图像检索系统中已涉及到了图像内部的一些空间关系。例如, QBIC 系统综合考虑了多种特征因素,包括颜色直方图、纹理、形状和空间关系^[5]。Virage 图像搜索引擎使用颜色、结构和文本信息来检索图片,其中用到了色块的空间分布特征^[6]。不过在上述系统中,空间关系主要还只是用作辅助其它特征进行检索。

在很多情况下,空间关系是图像信息的主要部分,尤其是当其中目标数目较多且目标的形状和大小相对于目标之间的相对距离而言可以忽略不计的时候^[7]。在这种场合下,对图像内部空间关系进行分析,并将空间关系直接用作检索特征就很重要。本文借助符号投影的模型,提出一种基于空

间关系的匹配算法——子区域最佳匹配算法, 直接对基于空间关系的图像检索进行研究。并以此作为核心算法, 构建了一个篮球站位分析演示系统, 验证了所提出的算法。

2 子区域最佳匹配算法

子区域最佳匹配法首先将图像抽象为图论中的连通图(graph), 将代表目标的图标抽象为图中的节点。然后按照中心点法将图划分为若干个子图, 这些子图的形状大小由且仅由节点的空间分布决定。中心点法描述的是以某一节点为中心的邻域构成子区域模式的方法。如果按广度优先法对图进行遍历, 可得到子区域顺序向量。根据待检索图和数据库图顺序向量的对应元素值可计算两个图中对应子区域的相似距离。最后通过加权求和得到两个图间的距离, 这个距离直接表示出两图的相似性。

2.1 对检索图片的预处理

篮球比赛中球员的站位情况是重要的信息。这里采用图标识检索和基于联系的模式^[8], 将图片中的球员抽象为图论中图的节点。节点没有形状、大小等属性, 只有位置和球员号码属性。通过与用户交互得到图片上球员的位置和号码信息(如使用多方向拍摄的图片, 也可考虑自动检测和识别), 并以此建立起一个带权连通图。

坐标变换后, 图片信息被抽象为特征向量, 这个特征向量与原始图片的视觉信息相对应, 同时也与一个图论中的带权连通图相对应。这个连通图中两相邻节点间的权重就是它们之间的欧氏距离。相邻节点的定义将在下一小节给出。

接下来将用户指定的图片和数据库里的每一幅图片进行匹配, 即计算数据库中的每一幅图片与指定图片的距离。具体算法包括搜索邻节点、划分子区域、遍历全图和计算图与图之间距离等部分。

2.2 搜索邻节点——等角度区域最近原则

空间关系中最重要的一条是相邻节点间的位置关系, 而对于图中某一个节点来说, 它周围节点与它的关系是比较模糊并且非常复杂的。邻节点的定义很关键并且比较困难, 事实上在以某一节点为中心的附近区域内的所有点都可以看成该节点某一方向上的相邻点。那么怎样把不同邻域的节点分开但又适当考虑人对图模式识别的整体性和模糊性呢? 答案是可以借助对节点的邻域进行分区来完成。图1给出对一个节点A的邻域进行分区得到的结果, 其它节点B到I被分在了不同的分区。

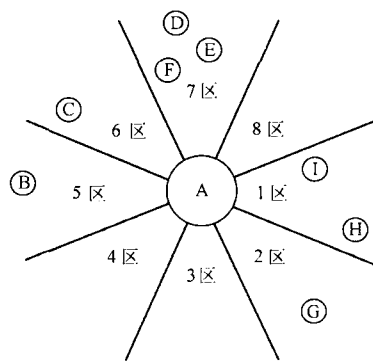


图1 节点A附近区域的划分情况

可采用以下算法来为每一个节点寻找其邻节点并建立起它与邻域内相邻节点的连接。

(1) 对于一个图中每一个节点, 将以该节点中心坐标为中心的 360° 区域分为8个扇形区, 每一个扇形区覆盖 45° 。以正东方正负 22.5° 的扇形区为1分区, 顺时针方向依次为2到8分区。

(2) 检测各分区内与该节点相邻的节点, 并用8个指针指向这些相邻的节点。值得指出, 图中每一个节点必落在8个区域中的某一个内。在每一个区域中, 找出该区域内离中心节点欧氏距离最近的点, 将它作为中心节点在这个分区中的相邻点。这样, 每一个节点附属的8个指向邻节点的指针都指向了对应分区中的邻节点(或者为空, 当某一分区内无邻节点时)。

考虑如图1所示的情况, 对A点而言, 2分区中的相邻节点应该是G, 5分区为B, 而1分区应该是I(舍去H), 7分区应该是F(舍去D和E), 3分区、4分区和8分区相邻节点为空。

2.3 划分子区域——中心点法

两幅人眼看来空间位置相似的图, 最本质的特点是他们在图的各个局部子图中有比较类似的节点分布。比如图2中, (a)和(c)在除了左上角区域外的各个子区域点的分布完全相同, 但相比(b)而言, 大部分人会认为(b)更接近于(a), 尽管(a)与(b)在各个子区域内的点分布只是类似而都不相同。这说明人眼对局部的小的差异不太敏感。在考虑用算法表示两幅图的差异时, 也要遵循人眼的这一特点。主要考虑两幅图是否在局部区域有一个近似的局部分布, 同时要对位置变化较大的点加以突出。划分子区域就是体现这一思想的方法, 它可以找到局部变化较大的点。

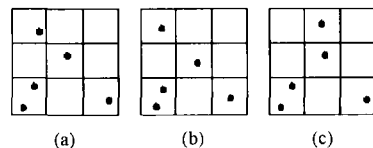


图2 三幅图的节点分布比较

实际上人在对上述篮球比赛球员站位进行分析的时候,往往是对某个球员的周围分布情况比较感兴趣,而对离这名球员较远的球员兴趣较弱(指在观察这名球员时)。所以,可以将某球员与其邻近的节点作为一个子区域,相似的两幅图在子区域内应该做到相似。一种具体的方法是,将每一个节点当作中心点,把它与其邻近节点结合组成一个子区域,对每一个子区域赋予不同的权值,最后在两幅图的各个对应子区域上分别匹配再按权值累加。

2.4 遍历全图——子区域顺序向量

按中心点法将一幅图分为各子区域后,应该对各个子区域分别匹配,这就涉及到子区域的排序。排序的目的是为了让两幅图中位置相似的两个对应子区域编号相同。只有经过子区域排序后的两幅图,才能在对应的子区域上作匹配分析。

按照连通图的理论,如果指定了一幅连通图中的起始节点,那么从这个节点起对图进行广度优先或深度优先搜索,正好能遍历所有节点且每一个节点只被访问一次,前提是从每一个节点访问其附近节点时遵循一定的指针顺序^[9]。为此可用图的遍历来确定各个节点的排列顺序,进而确定子区域的顺序。

首先要做的工作是寻找起始点。起始点的选择不是唯一的,但必须使两幅图对应起来。比如都选图的左上角点或右下角点。由于图片的坐标原点选在左上角,这里选择离左上角最近的点作为图的起始节点,它与附近的邻域构成了图的第一个子区域。对待检索图和数据库图,分别从各自的起始节点开始遍历整个图,得到的序号按顺序排列成向量,称为子区域顺序向量。考虑到让相邻的子区域编号尽可能的相邻,这里采用广度优先搜索来遍历全图。

然后对于已经访问过的节点,将对应的标志设为“已访问”。最后得到一个记录遍历顺序的向量——子区域顺序向量。具体算法为

- (1) 访问的第一个点为各自的起始节点。
- (2) 按从1区到8区的顺序访问起始节点在8个方向上相邻的节点,同时让这些节点在一个队列中排队。如果某个方向上无相邻节点,则直接跳过。
- (3) 按图论中广度优先的方法遍历余下的各点。取出当前队列中的第1个节点,按第(2)步所述的方法遍历这个节点的相邻节点(已访问过的可直接跳过),将刚访问过的节点加在队列的末尾。
- (4) 重复第(3)步的过程,直到队列为空。根据图论有关定理,这里的图都是连通图,因此队列为空时正好遍历了图中所有节点一次。

2.5 子区域相似距离和图像间距离

得到待检索图和数据库图各自的子区域顺序向量后,按

照此顺序对它们各自的节点进行排序,得到各自的节点向量。现在以这两个节点向量为依据,计算两个图的距离。这里定义两个图的距离为这两个图对应子区域的相似距离之和。这里的子区域相似距离考虑了两个图对应子区域内部对应中心节点的绝对位置差、对应中心点号码的差异、对应邻域节点与对应中心节点的距离差以及对应邻域节点号码的差异等因素。通过这样的定义,如果两个图在各个子区域空间关系都较为匹配,它们对应的图像间距离将很小,完全匹配时为零距离;如果在某个或某几个子区域上空间关系差异很大,这几个子区域相似距离将会很大,最后得到的图像间距离将随之增大。具体使用的定义式见附录。

3 篮球站位分析系统

基于上述子区域最佳匹配算法实现了一个完整的图像检索演示系统——篮球站位分析系统。系统的主要功能是从大量的篮球对阵图片中提取出相似的站位和指定的战术组合,达到分析篮球比赛和辅助球队制定战术的目的。

图3给出了系统的组成框图,这一框图包含了一般图像检索系统的主要模块^[10]。用户通过输入模块向系统提供查询图片,经过特征提取模块后得到特征向量。查询图片的特征向量与图像数据库中图片的特征向量进行匹配计算(采用本文提出的子区域最佳匹配算法),得到图片间的距离。然后按照距离递增相关度递减的顺序输出图像数据库中的图片。系统还为用户提供一个用户定制模块,用户可以改变各个队员的参数和权值从而得到不同的感兴趣的结果(见下一节)。

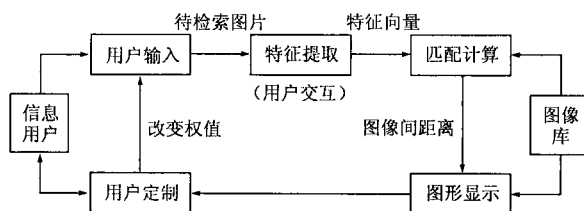


图3 检索系统组成框图

图4给出篮球站位分析系统的主界面的一个示例,从上到下依次有菜单和控制按钮,号码指定键和图片显示区。在主界面上提供了对数据进行交互标记的手段。系统输出界面的示例可见文献[7],输出界面上提供了用户交互反馈的工具。



图4 篮球站位分析系统主界面

4 实验结果及分析

下面通过一次具体的检索来讨论和分析系统的功能和性能。用户输入的查询图片如图5所示,这幅图的特点:(1)双方所有的队员都出现;(2)图片左侧有一处明显的一对一画面;(3)右下角是白色4号队员。



图5 待检索图片

图6依次是检索结果中的前8幅,可以看出检索结果与原始图片有着较好的匹配。

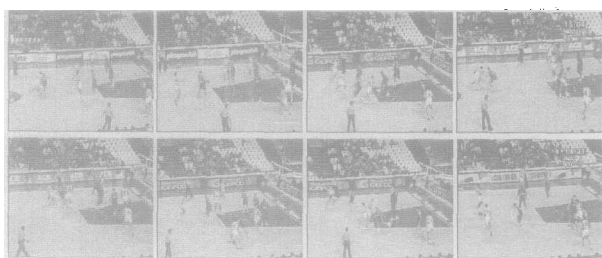


图6 对图5的第1次检索所输出的结果

另外,通过用户定制可以分别增加待检索图片左侧一对一子区域重要度和右下角白色4号球员的权值从而得到另外两次不同的检索结果。限于篇幅,仅在表1给出一些统计数据。

表1 三次检索的结果(查全率)

检索序号	总体效果		左侧一对一子区域		右下角白4号队员	
	前10幅	前30幅	前10幅	前30幅	前10幅	前30幅
第1次	100%	86.7%	70%	63.3%	80%	40%
第2次	100%	70%	100%	76.7%	60%	40%
第3次	90%	80%	70%	60%	100%	77.3%

第1次检索的结果侧重总体感观,它不是某一子区域相似匹配度很高,而是对各个子区域的匹配程度的综合考虑。这是由于系统给每一个子区域都分配了相等的权重。

第2次检索的结果侧重图片左侧一红一白队员组成的子区域,体现在序号比较靠前的图片大都在这个子区域上与原始图片非常匹配。这是由于系统给这块子区域分配了较高的权值。

第3次检索的结果侧重图片中白色4号队员的位置,体现在序号靠前的图片大都在与原始图片类似的右下角位置出现了白色4号队员。这是因为系统给白4号队员分配了较

高的权值。

由上可见,用户可以通过定制将主观意愿反映给系统,系统可考虑用户意见进行检索。

表2是本系统与以往一些系统^[11]在运算时间方面对比的统计数据(其中相对时间比是以本系统为基准,根据检索时间、图片尺寸和数目换算得到的),可见本系统所用算法的时间复杂度是比较小的(平台不同,仅供参考)。

表2 不同系统的运算时间

系统	图片尺寸	图片库规模	运行平台	检索时间(s)	相对时间比
WBIS Stanford	128×128	10,000	Sun SPARC-20	3.3	0.317
IBM QBIC	100×100	1000	IBM SP/2	2~40	3.146~62.918
Univ. of Wash. 1	128×128	1093	SGI Indy R4400	47.46	41.687
Univ. of Wash. 2	128×128	20,558	SGI Indy R4400	892.60	41.684
本系统	352×288	481	PC, Pentium III 450	3.1	1

5 总结与展望

本文基于图像之间局部区域的相似距离和图像间距离的概念,提出了一种基于空间关系的图像检索算法——子区域最佳匹配法。这种算法使用符号投影模型,将图像内目标抽象为带有多个属性的图标,通过搜索邻节点的等角度区域最近原则和划分子区域的方法——中心点法实现了图像的匹配计算。与常用的基于2-D串(2-D string)的方法相比^[2],这里不需要那么多空间操作算子。

本文还设计实现了一个基于空间关系的图像检索演示系统,并把空间关系的算法应用到篮球战术分析中。实验结果反映出系统有较好的性能,验证了空间关系检索算法的有效性。该算法除单独使用外,还可与其他基于颜色、纹理和形状的算法结合使用。

附录 子区域相似距离和图像间距离的定义和计算

设待检索图像中队员的位置记录为 $(E_1, E_2, \dots, E_m)$,数据库里图像中队员的位置记录为 $(E'_1, E'_2, \dots, E'_n)$,经过遍历的顺序分别为 $(E'_{i_1}, E'_{i_2}, \dots, E'_{i_m})$ 和 $(E'_{j_1}, E'_{j_2}, \dots, E'_{j_n})$ 。其中 $(i_1, i_2, \dots, i_m)$ 是 $(1, 2, \dots, m)$ 的一个排列, $(j_1, j_2, \dots, j_n)$ 是 $(1, 2, \dots, n)$ 的一个排列。具体又分两种情况:

(1) $m = n$ 此时图像间距离为所有对应的子区域间的相似距离之和:

$$D_i = \sum_{k=1}^n D_s(E_{i_k}, E'_{j_k}) \quad (\text{A-1})$$

这里子区域间的相似距离定义为

$$D_s(E_{i_k}, E'_{j_k}) = F_k^{(1)} \sum_{r=1}^8 F_{k,r}^{(2)} d[L_r(E_{i_k}), L_r(E'_{j_k})] \quad (\text{A-2})$$

其中 $D_s(E_{i_k}, E'_{j_k})$ 是 E_{i_k} 和 E'_{j_k} 间的相似距离, $L_r(E_{i_k})$ 和 $L_r(E'_{j_k})$ 分别是 E_{i_k} 和 E'_{j_k} 各自周围第 r 号扇区的相邻节点, $r=1, 2, \dots, 8$, $d[L_r(E_{i_k}), L_r(E'_{j_k})]$ 是 E_{i_k} 周围第 r 号扇区的相邻节点到 E_{i_k} 的欧氏距离与 E'_{j_k} 周围第 r 号扇区的相邻节点到 E'_{j_k} 的欧氏距离差的绝对值。如果其中某个相邻节点不存在, 比如 $L_r(E_{i_k})$ 不存在, 则该距离差等于 $L_r(E'_{j_k})$ 。 $F_k^{(1)}$ 是待检索图和数据库里图对应中心节点的影响因子。 $F_{k,r}^{(2)}$ 是待检索图和数据库里图对应中心节点的相邻元素影响因子。

需要说明, 这里定义的相似距离是为了反映两幅图中的对应节点之周边节点位置的相似情况而设定的, 其中 $d[L_r(E_{i_k}), L_r(E'_{j_k})]$ 反映了 E_{i_k} 周边 r 号扇区的相邻节点到 E_{i_k} 和 E'_{j_k} 周边 r 号扇区的相邻节点到 E'_{j_k} 的距离差异, 这个差异定量反应了 E_{i_k} 和 E'_{j_k} 附近点关系的不同之处。 $F_k^{(1)}$ 和 $F_{k,r}^{(2)}$ 分别反应了中心节点和周边节点号码差异对两幅图差异性的影响。 $F_k^{(1)}$ 和 $F_{k,r}^{(2)}$ 是经验值, 两幅图对应节点号码越匹配, $F_k^{(1)}$ 和 $F_{k,r}^{(2)}$ 越接近 1.0。在最后的距离公式中, $F_k^{(1)}$ 和 $F_{k,r}^{(2)}$ 都是乘性因子。这是为了缩放已经算得的子区域相似距离以体现号码差异对距离的影响。

(2) $m \neq n$ 不失一般性, 设 $m < n$ 。算法从包含 n 个节点 ($E_1, E_2, \dots, E_n$) 的数据库图片中派生出包含 m 个节点 ($E_{i_1}, E_{i_2}, \dots, E_{i_m}$) 的子图。其中, $(E_{i_1}, E_{i_2}, \dots, E_{i_m}) \subset (E_1, E_2, \dots, E_n)$ 。这样的子图共有 C_n^m 个。对每一个子图 S_i , 按(A-1)中所述方法计算该子图与待检索图的距离 D_{si} 。最后的距离公式为

$$D_t = F^{(3)} \min_{1 \leq i \leq C_n^m} \{D_{si}\} \quad (\text{A-3})$$

其中, $\min_{1 \leq i \leq C_n^m} \{D_{si}\}$ 表示所有子区域匹配距离中的最小者, $F^{(3)}$ 表示与待检索图和数据库里图节点数目相关的影响因子。

需要说明, 两幅图中节点数目不等的时候, 应该既考虑到节点数目少的图可能与数目多的图中某个局部相似, 也应该考虑到数目不同造成的影响。因此, 这时的距离公式不是直接定义待检索图和数据库图的距离, 而是从数据库图中截取一部分子图, 看待检索图和子图的匹配情况。最后选取数据库图所有子图中与待检索图最相似(即距离最小)的子图, 把它和待检索图的距离乘上节点数目影响因子 $F^{(3)}$ 作为两图间的距离。对于 $F^{(3)}$ 的定义主要基于如下考虑: $F^{(3)}$ 既要反应

两图的节点数目差($n-m$), 又要使这个数目差在节点数不同时对两幅图差异性影响是不同的。比如节点数为 2 和节点数为 4 的图的差异应该比节点数为 8 和节点数为 10 大很多。后者几乎可以忽略节点个数对距离的影响。基于上述考虑, (A-3)中的 $F^{(3)}$ 表达式里有 $(n-m)/m$ 这一项, 在 $(n-m)$ 相同时, m 越小, $F^{(3)}$ 这个影响因子越大。

参考文献

- [1] 章毓晋. 基于内容的视觉信息检索. 北京: 科学出版社, 2003, 第1章, 第3章.
- [2] Wang Y H. A spatial relationship method supports image indexing and similarity retrieval. In: Multimedia Systems and Content-Based Image Retrieval, Deb S, ed. Hershey, London, Melbourne, Singapore: Idea Group Publishing, 2004, Chapter 12, 277 - 301.
- [3] 田福海, 等. 现代篮球训练艺术. 上海: 上海教育出版社, 1993, 第5章.
- [4] Sciascio E D, Donini F M, Mongiello M. Spatial layout representation for query-by-sketch content-based image retrieval. *Physical Review Letters*, 2002, 23(13): 1599 - 1612.
- [5] <http://www.qbic.almaden.ibm.com/>.
- [6] <http://www.ctr.columbia.edu/VisualSEEK/>.
- [7] 章毓晋. 基于内容的视觉信息检索. 北京: 科学出版社, 2003, 第7章.
- [8] Del Bimbo A. Visual Information Retrieval. San Francisco: Morgan Kaufmann Publishers, Inc., 1999, Chapter 5.
- [9] Preiss B R. Data Structures and Algorithms with Object-Oriented Design Patterns in C++. John Wiley & Sons, 1998.
- [10] 章毓晋. 图象工程(下册)——图像理解与计算机视觉. 北京: 清华大学出版社, 2000, 第8章.
- [11] Albuz E, Kocalar E, Khokhar A A. Scalable color image indexing and retrieval using vector wavelets. *IEEE Trans. on Knowledge and Data Engineering*, 2001, 13(5): 851 - 861.

徐丹: 男, 1980年生, 目前为University of Illinois at Urbana-Champaign的博士生, 研究功能磁共振成像。

章毓晋: 男, 1954年生, 教授, 博士生导师, 研究方向为图像工程(图像处理, 图像分析, 图像理解及其技术应用)和相关教学方法(<http://www.ec.tsinghua.edu.cn/~s93215>).