

一种面向视觉效应的快速图像检索方法

宋 坤 程小平

(西南师范大学计算机与信息科学学院 重庆 400715)

摘 要 本文提出了一种面向视觉效应的图像检索方法——ACH&GCH RM,该方法在进行相似性度量时,将“面积-颜色”联合直方图(Area-Color Histogram, ACH)融合到了传统的全局颜色直方图(Global Color Histogram, GCH)之中,由于在构造 ACH 的过程中进行了区域面积的非线性量化和区域数目的非线性映射,该检索方法具有更接近人眼视觉特性的优点。而压缩技术的使用使得两种直方图融合后仍具有较快的检索速度。

关键词 人眼视觉特性,区域面积,区域数目,非线性

Image Retrieval Using Nonlinear Quantification of Region's Area and Nonlinear Mapping of the Number of Regions

SONG Kun CHENG Xiao-Ping

(Faculty of Computer and Information Science, Southwest Normal University, Chongqing 400715)

Abstract Feature descriptor is one of the key technologies for content-based image retrieval. In this paper, a Area-Color Histogram is proposed to be combined with the traditional Global Color Histogram. By using nonlinear quantification of the region's area and nonlinear mapping of the number of regions, the combination makes retrieval results more closed to human vision when compared to the case where Global Color Histogram is used only. Moreover, fast retrieval speed is available by compacting the storage of Area-Color Histogram.

Keywords Human vision, Region's area, Number of regions, Nonlinear

1 研究现状和动机

特征表示方法是基于内容的图像检索(Content-Based Image Retrieval, CBIR)的关键技术之一。已有的许多方法都用到了颜色特征^[1,2],其中,全局颜色直方图(Global Color Histogram, GCH)是描述颜色特征的最传统的方式^[3]。GCH 的计算非常简单——在颜色总数为 C 的颜色空间中, GCH 是一个 C 维的向量,形如 (h_1, h_2, \dots, h_c) , 其中, h_c 是颜色 c 在图像中所占的比例,这里假设像素的颜色用正整数 c 表示, $c = 1, 2, \dots, C$ 。对一幅尺寸为 $M \times N$ 的图像 $F(i, j)$ ($i = 1, 2, \dots, M, j = 1, 2, \dots, N$), h_c 的计算如下:

$$h_c = \frac{\sum_{i=1}^M \sum_{j=1}^N B_c(i, j)}{M \times N}$$

其中 $B_c(i, j)$ 的取值由下式确定:

$$B_c(i, j) = \begin{cases} 1, & \text{if } F(i, j) = c \\ 0, & \text{otherwise} \end{cases}$$

GCH 具有对图像旋转和缩放不变的特性,但其致命弱点在于它完全没有考虑颜色在一幅图像中的具体分布,图 1 中的例图可以说明这一点。

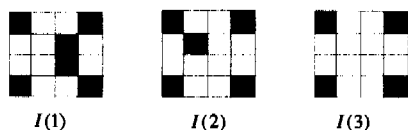


图 1

图中每一个小方格表示一个像素,每个像素共有三种可

能的颜色($C = 3, c = 1, 2, 3$); 黑色(1)、灰色(2)和白色(3)。

GCH(r)表示 $I(r)$ 的 GCH,则有:

$$\text{GCH}(1) = (4/16, 0, 12/16),$$

$$\text{GCH}(2) = (4/16, 1/16, 11/16),$$

$$\text{GCH}(3) = (4/16, 2/16, 10/16).$$

$Dgch(r, s)$ 表示 GCH(r)和 GCH(s)之间的欧几里得距离(其值越小,表明 $I(r)$ 和 $I(s)$ 越相似)。显然,在图 1 中,有:

$$Dgch(1, 2) = Dgch(2, 3) = \frac{\sqrt{2}}{16}$$

这个结果和人眼的视觉特性不相符合。从视觉上来说, $I(2)$ 和 $I(1)$ 比较相似,而和 $I(3)$ 不太相似。但当使用 GCH 作为特征向量时, $I(2)$ 被认为与 $I(1)$ 和 $I(3)$ 同样的相似(或不相似)。造成这种结果的原因之一就是 GCH 只考虑了某种颜色在图像中占有的比例,而没有考虑到该颜色在图像中具体是怎样分布的。

2 新的特征表示方法的融入

针对全局颜色直方图(GCH)的这个弱点,笔者提出了“面积-颜色”联合直方图(Area-Color Histogram, ACH)。ACH 在一定程度上将颜色在图像中的分布纳入了考虑范围之内,其实质是统计图像中具有某种面积和某种颜色的区域的个数。

我们可以认为两个在对方的 8 邻域内,且具有相同颜色的像素属于同一区域,这是以像素为基元进行区域划分的方法,这样划分出来的区域比较细致。我们也可以将基元放大,比如放大到 4×4 个像素,这样划分出来的区域就会粗糙一

宋 坤 硕士研究生,主要研究方向为基于内容的图像检索。程小平 教授,硕士生导师,主要研究方向为计算机图形图像工程,计算机辅助设计,电子商务工程,计算机集成制造,数控及先进制造系统。

些。事实上,选择大小合适的基元,不但可以忽略那些人眼注意不到的图像细节,还可以缩短区域划分的时间。不过,这并不是本文研究的重点。

本文涉及到的区域面积是指相对面积,设区域包含的像素数为 P_c , 区域所在的图像包含的像素数为 P_i , 则区域面积 A_c 为:

$$A_c = \frac{P_c}{P_i}$$

为避免占用过大而又不必要的存储空间,应该对区域面积进行量化。笔者对区域面积进行了非线性量化,这是因为随着面积的增大,等量的面积增加给视觉带来的冲击会减小。来看图 2 中的四个正方形。

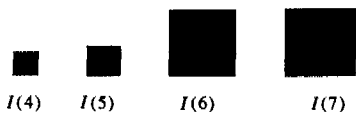


图 2

正方形 $I(4)$ 和 $I(5)$ 之间的面积之差与正方形 $I(6)$ 和 $I(7)$ 之间面积之差是相等的,但从人眼的角度来说,我们决不会认为 $I(4)$ 和 $I(5)$ 一样大,却可以接受 $I(6)$ 和 $I(7)$ 面积大致相等的说法。根据人眼的这个特性,笔者对区域面积进行了非线性量化,即使量化间隔随着区域面积的增大而增大。量化区间可以有多种不同的划分方法,表 1 中的方法将区域面积划分为了 7 个等级。

表 1

等级	区域面积	等级	区域面积
1	(0, 1/64]	5	(1/8, 1/4]
2	(1/64, 1/32]	6	(1/4, 1/2]
3	(1/32, 1/16]	7	(1/2, 1]
4	(1/16, 1/8]		

假设在颜色总数为 C 的颜色空间下,面积被量化为 A 个等级,则 ACH 是一个 $A \times C$ 维的向量,形如 $(h_1, h_{12}, \dots, h_{1C}, \dots, h_{A1}, h_{A2}, \dots, h_{AC}), h_{ac}$ 是图像中面积为 a 且颜色为 c 的区域的个数,在这里,面积和颜色均用正整数来表示,且 $a=1, 2, \dots, A, c=1, 2, \dots, C$ 。若一幅图像被划分为 K 个区域,这些区域依次被命名为 E_1, E_2, \dots, E_K , 且每个区域都具有面积和颜色两个属性,分别用 E_{k-are} 和 E_{k-col} 来表示,其中 $k=1, 2, \dots, K$, 则 h_{ac} 的计算如下:

$$h_{ac} = \sum_{k=1}^K B_{ac}(k)$$

其中, $B_{ac}(k)$ 的取值由下式决定:

$$B_{ac}(k) = \begin{cases} 1, & \text{if } E_{k-are} = a \text{ and } E_{k-col} = c \\ 0, & \text{otherwise} \end{cases}$$

这样计算出来的 ACH 在一定程度上考虑了颜色在图像中的分布,这个优点是 GCH 所不具备的。

现在来计算图 1 中的例图的 ACH。为简便起见,这里将区域面积量化成 4 个等级 ($A=4, a=1, 2, 3, 4$), 详见表 2。

表 2

等级	区域面积	等级	区域面积
1	1/16	3	4/16, ..., 7/16
2	2/16, 3/16	4	8/16, ..., 16/16

这里的 ACH 是一个 $12(4 \times 3)$ 维的向量, 它的第 $(a-1)$

$\times 3 + c$ 柄表示面积为 a 且颜色为 c 的区域的个数, 其中 $a=1, 2, 3, 4, c=1, 2, 3$ 。用 $ACH(r)$ 表示 $I(r)$ 的 ACH, 有:

$$ACH(1) = (4, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1),$$

$$ACH(2) = (4, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1),$$

$$ACH(3) = (1, 2, 0, 1, 0, 0, 0, 0, 0, 0, 1).$$

用 $Dach(r, s)$ 表示 $ACH(r)$ 和 $ACH(s)$ 之间的欧几里得距离, 则:

$$Dach(1, 2) = \sqrt{1^2} = 1$$

$$Dach(2, 3) = \sqrt{3^2 + 1^2 + 1^2} = \sqrt{11}$$

显然, 就图 1 中的例图而言, 当使用 ACH 作为特征向量时, 得到了和人眼视觉特性相符合的结果。



图 3

但 ACH 也有它自身的缺陷, 比如在图 3 中, $I(9)$ 与 $I(8)$ 和 $I(10)$ 中的后者比较接近, 现在来求它们的 GCH 和 ACH:

$$GCH(8) = (4/16, 7/16, 5/16)$$

$$GCH(9) = (6/16, 7/16, 3/16)$$

$$GCH(10) = (5/16, 7/16, 4/16)$$

$$ACH(8) = (0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0)$$

$$ACH(9) = (0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0)$$

$$ACH(10) = (0, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 0)$$

可得相应的欧几里得距离:

$$Dgch(9, 10) = \sqrt{\left(\frac{2}{16}\right)^2 + \left(\frac{2}{16}\right)^2} = \frac{\sqrt{8}}{16}$$

$$Dgch(9, 10) = \sqrt{\left(\frac{1}{16}\right)^2 + \left(\frac{1}{16}\right)^2} = \frac{\sqrt{2}}{16}$$

$$Dach(8, 9) = \sqrt{1^2 + 1^2} = \sqrt{2}$$

$$Dach(9, 10) = \sqrt{1^2 + 1^2} = \sqrt{2}$$

可见, 在图 3 中, 使用 ACH 作为特征向量并没有得到和人眼视觉特性相符合的结果, 而使用 GCH 作为特征向量却得到了和人眼视觉特性相符合的结果。

造成这种结果的原因之一就是 ACH 虽然从一定程度上考虑了某种颜色的具体分布, 但它却忽视了该颜色在图像中的占有量。一幅图像中某种颜色的近似占有量可以从它的 ACH 中求出来, 比如, 由 $ACH(10)$ 可以计算出 $I(10)$ 中黑色像素的占有量大致为 $(1 \times 3) \div (1 \times 1 + 2 \times 2 + 2 \times 3 + 0 \times 4)$, 但是, 我们并不能从 $ACH(10)$ 的任意一柄中得到黑色像素的占有比例, 因此, 在仅使用 ACH 的情况下, 颜色的占有比例是无法参与相似性度量的。由此我们想到, 如果能将 ACH 融合到 GCH 之中, 既考虑颜色在图像中的占有比例, 又在一定程度上考虑颜色在图像中的分布, 从理论上来说, 应该能得到比单独使用它们中任意一个都更接近人眼特性的结果。

可以用一种简单的方式将 ACH 融合到 GCH 之中, 即使使用以下的相似性度量公式:

$$Dag = \sqrt{Dach \times Dgch}$$

对于图 1 和图 3 中的例图, 可以求得:

$$\begin{aligned} Dag(1, 2) &= \sqrt{Dach(1, 2) \times Dgch(1, 2)} \\ &= \sqrt{1 \times \frac{\sqrt{2}}{16}} = \sqrt{\frac{\sqrt{2}}{16}} \end{aligned}$$

$$Dag(2,3) = \sqrt{Dach(2,3) \times Dgch(2,3)}$$

$$= \sqrt{\sqrt{11} \times \frac{\sqrt{2}}{16}} = \sqrt{\frac{\sqrt{22}}{16}}$$

$$Dag(8,9) = \sqrt{Dach(8,9) \times Dgch(8,9)}$$

$$= \sqrt{\sqrt{2} \times \frac{\sqrt{8}}{16}} = \sqrt{\frac{4}{16}}$$

$$Dag(9,10) = \sqrt{Dach(9,10) \times Dgch(9,10)}$$

$$= \sqrt{\sqrt{2} \times \frac{\sqrt{2}}{16}} = \sqrt{\frac{2}{16}}$$

显然, $Dag(1,2) < Dag(2,3)$,

且 $Dag(9,10) < Dag(8,9)$ 。

由此可见,将 ACH 融合到 GCH 中后,在图 1 和图 3 的例图中都得到了和人眼视觉特性相符合的结果。

然而,尽管融合了 ACH 能得到更好的结果,我们还是不得不考虑它所占的存储空间。到目前为止,文中出现的 ACH 的维数都很小(12),但是,随着区域面积的量化等级和颜色空间的颜色的增加,ACH 的维数会迅速增大,为了节省存储空间和查询时间,有必要对 ACH 进行压缩。观察发现,在图 1 和图 3 中,6 幅例图的 ACH 都包含大量的 0 元素。事实上,笔者在将区域面积和颜色空间分别量化为 7 个和 64 个等级的情况下,对来自众多领域的 1000 幅图像的 ACH 进行分析后发现,99.9% 的 ACH 都含有 87.7% 以上的 0 元素,这就使得 ACH 的具有较大的压缩空间。最简单的压缩方法就是只记录 ACH 中非 0 元素的位置和取值。

另外,到目前为止,文中出现的 ACH 的最大柄值为 4。事实上,在更具有一般性的例子中,ACH 的柄值变化范围会从几、几十到上百。为此,笔者提出对 ACH 的柄值进行非线性映射,这样做的理由和对区域面积进行非线性量化的理由非常相似:假设用 E_x 表示一个面积为 a 且颜色为 c 的区域,那么,当图像中的 E_x 越多的时候,等量的 E_x 的增加给视觉造成的冲击就会减小。

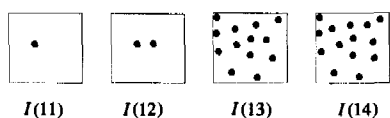


图 4

图 4 中各例图所含的小圆形面积均相等且均为黑色,各例图所含的小圆形个数依次为 1、2、13、14。由此可见,同样是一个小圆形之差,我们可以很容易看出 $I(11)$ 和 $I(12)$ 之间的不同,却较难发现 $I(13)$ 和 $I(14)$ 之间的差异。所以,有理由相信,对区域的个数进行非线性映射能使所得到的特征向量更接近人眼的特性。可以通过多种途径实现非线性映射,求对数法是其中一种:

假设 ACH 某一柄的值为 h_x ,经过非线性映射后的值为 h'_x ,则:

$$h'_x = NLM(h_x) = \begin{cases} \lfloor \log_2 h_x \rfloor + 1, & V \neq 0 \\ 0, & V = 0 \end{cases}$$

其中, NLM 是非线性映射算子。

3 实验部分

实验所用的图像数据库包含 2519 幅图像,这些图像可分

为两部分:一部分是来自各个领域的图像共 2378 张;另一部分是 12 个相似图像集 (Similar Images Set, SIS),每个 SIS 中都包含在颜色分布和语义上接近的若干图像,12 个 SIS 所含的图像总数为 141。为了减小问题的规模,笔者将库中所有图像的尺寸都按比例缩小到不大于 100×100 像素,这样做不会对实验结果造成影响,因为 GCH 和区域的相对面积对图像的缩放都具有不变性。

实验在一台 CPU 为 2GHz、硬盘为 40G、内存为 256M、操作系统为 Windows XP SP1 的计算机上进行,所有程序都在 Matlab6.5 环境中编写和执行。

实验比较了三种图像检索方法 (Retrieval Method, RM) 的检索结果和检索所需的时间。第一种是基于“面积-颜色”联合直方图的图像检索方法,简称为 ACH RM;第二种是基于全局颜色直方图的图像检索方法,简称为 GCH RM;最后一种是基于将 ACH 融合到 GCH 中的图像检索方法,简称为 ACH&GCH RM。数据库中的所有图像都在颜色空间上被量化成了 64 个等级。当计算 ACH 时,区域面积按表 1 中的方法被量化为 7 个等级,并使用求对数法对 ACH 的柄值进行非线性映射,另外,采用仅记录非 0 元素的办法对 ACH 进行压缩。图 5 所示为使用 ACH&GCH 检索方法时的检索系统流程图。

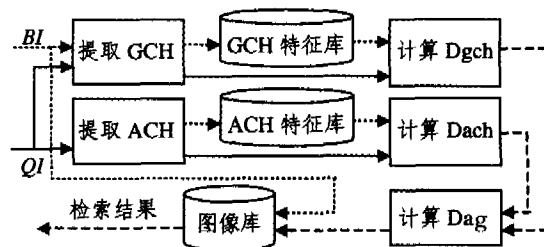


图 5

BI 是等待入库的图像,系统提取它的 GCH 和 ACH 分别存入 GCH 特征库和 ACH 特征库。实验使用相似图像集中的图像作为查询图像 (Query Image, QI),每提交一幅 QI,系统就启动一次检索。一次检索过程包含以下三个步骤:

(1) 计算 QI 的 GCH 和 ACH;

(2) 依次计算 QI 和库中所有图像之间的 Dag,并将库中图像的文件名按 Dag 的降序排列;

(3) 在图像库中查找并显示排在前面 20 位的文件名所对应的图像。

假设某相似图像集 (SIS) 中的图像数目为 N_s ,当以该集合中第 i 幅图像作为 QI 时,若返回的 20 幅图像中有 N_r 幅 (包括 QI) 和 QI 属于同一个相似图像集,则本次检索的检索率 (R_i) 为:

$$R_i = \frac{N_r}{N_s}$$

检索率的值越大,表示该特征检索方法越有效。某相似图像集的平均检索率 (R_a) 为:

$$R_a = \frac{\sum_{i=1}^{N_s} R_i}{N_s}$$

笔者通过比较 12 个相似图像集 (SIS) 的平均检索率 (R_a) 和平均检索时间 (T_a) 来衡量三种检索方法的好坏, T_a 以秒 (s) 为度量单位。表 3 显示了各相似图像集的中英文名

称对照和所含的图像数目,表4显示了比较的结果。

表3

SIS 中文名	SIS 英文名	Ns
黄色花朵	Yellow Flower, YFL	13
天空与大象	Sky and Elephant, SAE	12
恐龙	Dinosaur, DIN	9
蓝色的大海	Blue Sea, BSE	8
白玉	White Jade, WJA	9
戒指1	Ring1, RIN1	9
足球场	Football Playground, FPL	18
大红色碎花	Red Scattered Flowers, RSF	10
银色手链	Argent Bangle, ABA	9
森林与马匹	Forest and Horses, FAH	14
戒指2	Ring2, RIN2	20
山川河流	Hill and River, HAR	10

从表4中的数据中可以看出,在笔者所设计的实验环境下,当采用基于ACH&GCH的检索方法时,可以获得比GCH RM或ACH RM都更高的检索率。图6~图8给出了以相似图像集“森林和马匹”(Forest and Horses, FAH)中的图像 fah12 作为 QI 时,使用三种检索方法所得结果的图示。文件名以“fah”开头的是相似集 FAH 中的图像(当然它也是图像库中的图像),以“b”开头的是图像库中除 FAH 以外的其他的图像。图6所示的是 ACH RM 的检索结果,其检索率为 9/14;图7所示的是 GCH RM 的检索结果,其检索率为 10/14;图8所示的是 ACH&GCH RM 的检索结果,其检索率最高,达到了 13/14。

表4

SIS \ RM	ACH RM		GCH RM		GCH&ACH RM	
	Ra	Ta	Ra	Ta	Ra	Ta
YFL	21.9%	49.1	47.9%	55.6	53.9%	56.6
SAE	17.4%	49.9	41.0%	47.0	41.7%	50.1
DIN	59.3%	48.0	49.4%	40.5	61.7%	43.7
BSE	75.0%	59.8	81.3%	51.1	90.6%	60.4
WJA	48.2%	57.3	87.7%	55.7	92.6%	60.8
RIN1	39.5%	53.6	43.2%	49.4	46.9%	53.8
FPL	65.7%	62.3	76.9%	54.9	88.9%	61.5
RSF	39.0%	61.0	63.0%	55.8	76.0%	61.0
ABA	39.5%	59.5	50.6%	49.2	63.0%	59.6
FAH	44.4%	56.7	49.0%	50.7	58.2%	56.5
RIN2	24.8%	49.7	25.8%	49.4	28.8%	50.2
HAR	21.0%	56.0	35.0%	42.8	38.0%	50.2

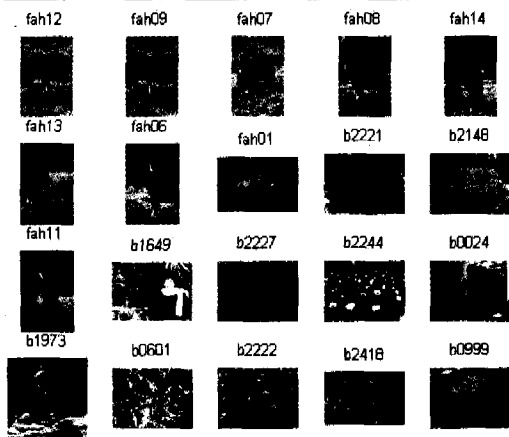


图6

结论 以上的理论分析和实验结果表明:将“面积-颜色”联合直方图(ACH)融合到全局颜色直方图(GCH)而产生的基于ACH&GCH的图像检索方法,由于其不仅考虑了颜色在图像中的占有比例,还在一定程度上考虑了颜色在图像中的具体分布,因而能取得比基于GCH的检索方法更好的检索结果。与此同时,虽然在检索时间上,GCH&ACH RM平均比GCH RM多5.2s,但由于两者的检索时间都在50s左右,使得5.2s的差异不易被人的感觉所捕捉,因此,ACH&GCH RM是一种比GCH RM更符合人眼视觉特性的快速检索方法。

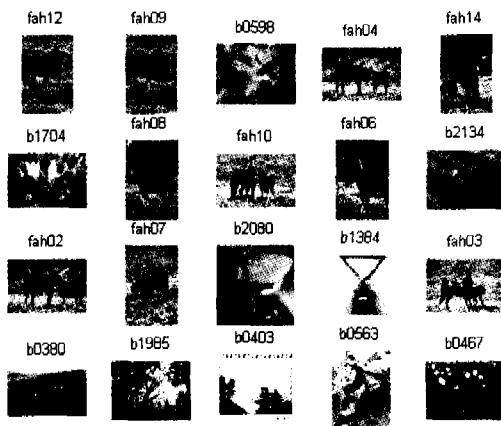


图7

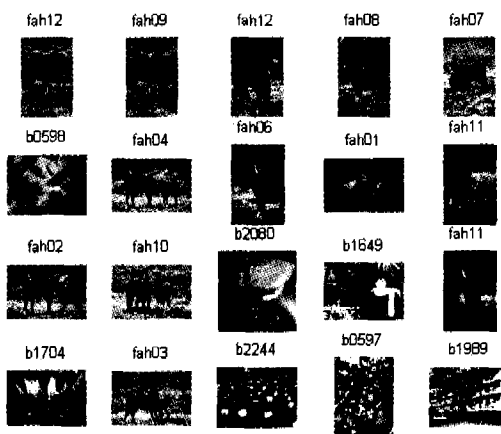


图8

参考文献

- 1 刘忠伟,章毓晋. 利用局部累加直方图进行彩色图像检索. 中国图象图形学报, 1998, 3(7): 533~537
- 2 Nascimento M A, Tousidou E, Chitkara V, Manolopoulos Y. Image indexing and retrieval using signature trees. Data & Knowledge Engineering, 2002, 43: 57~77
- 3 Bimbo D. Visual Information Retrieval. Chapter 3 and 4. Morgan Kaufmann Publishers, Inc, 1999
- 4 Smeulders A W M, Worring M, Gupta A, Jain R. Content-Based Image Retrieval at the End of the Early Years. IEEE Transaction on Pattern Analysis and Intelligence, 2000(12)
- 5 Fauqueur J, Boujemaa N. Region-based image retrieval: fast coarse segmentation and fine color description. Journal of Visual Language and Computing, 2004(12): 69~95
- 6 Chitkara V. Color-based image retrieval using compact binary signatures; [Master's thesis]. Dept. of Computing Science, University of Alberta, 2001
- 7 Gonzalez R C, Wood R E. 数字图像处理, 第二版. 阮秋琦, 阮宇智, 等译. 北京: 电子工业出版社, 2003