

# 基于内容的图像检索的相似度测量方法

王春静 许圣梅

(山东师范大学信息科学与工程学院, 济南, 250014)

**摘要:** 图像特征匹配是基于内容的图像检索(Content-based image retrieval, CBIR)实现的一个关键环节,而图像特征的匹配主要依赖于图像特征之间的相似度测量。为了提高CBIR的检索性能,本文提出了一种有效的相似度测量方法——基于图像k近邻的相似度测量(Similarity measure based on k-nearest neighbors of images, SBkNN)方法。在该方法中,查询图像与被检索图像的相似度通过计算这两幅图像属于同一语义(无论是哪种语义)种类的联合概率来衡量,而此概率可分别通过分析这两幅图像与各自近邻图像的距离得到。最后利用Corel5k数据集对本文所提出的SBkNN方法和传统的相似度测量方法进行了对比。实验结果表明,在CBIR中使用本文提出的SBkNN方法,有效地提高了CBIR的检索性能。

**关键词:** 基于内容的图像检索; k近邻; 相似度; 召回率; 查准率

**中图分类号:** TP311      **文献标志码:** A

## Effective Similarity Measure Method for Content-Based Image Retrieval

Wang Chunjing Xu Shengmei

(School of Information Science and Engineering, Shandong Normal University, Jinan, 250014, China)

**Abstract:** Image feature matching is a key link for the implementation of content-based image retrieval (CBIR), which mainly relies on the similarity measure between the features of two images. To improve the retrieval performance of CBIR, this paper proposes an effective similarity measure method—similarity measure based on k-nearest neighbors of images (SBkNN). In the proposed SBkNN method, the similarity between query image and retrieved image is obtained by calculating the probability for the two images belonging to the same semantic category (no matter what kind of semantic category), and the probability can be obtained by analyzing the distance between the two images and their k-nearest neighbors, respectively. Finally, the comparison between the proposed SBkNN method and traditional similarity measure is implemented on Corel5K dataset. Experimental results show that the proposed SBkNN method significantly improves the retrieval performance of CBIR.

**Key words:** content-based image retrieval (CBIR); k-nearest neighbors; similarity; recall; precision

## 引 言

图像检索是图像处理最重要的应用之一。随着互联网上数字图像的爆炸性增长,我们迫切地需要

一种有效的图像检索方法,以便于能够从庞大的图像数据库中迅速、精确地找出最想要的图像。目前,图像检索的方法主要有3种:基于文本的图像检索<sup>[1]</sup>(Text-based image retrieval, TBIR)、基于内容的图像检索<sup>[2]</sup>(Content-based image retrieval, CBIR)和基于语义的图像检索<sup>[3]</sup>(Semantic-based image retrieval, SBIR)。TBIR主要依赖网页中图像周围的文本来进行检索,但是图像周围的文本往往不是直接用来描述图像本身的内容,这就限制了TBIR方法的准确性。SBIR主要着重于图像高层语义的提取,由于图像内容的底层特征(颜色、形状、纹理等)与人们对图像内容的理解之间存在差异,因此在SBIR中提取的图像高层语义也不够准确,这就限制了SBIR方法的检索准确性。而CBIR就是基于图像的可视内容,例如图像的颜色、纹理及形状等,且CBIR具有计算简单、图像检索性能高的特点,因此被广泛应用于各个研究领域。

图像特征匹配是CBIR能否实现的一个关键环节,图像特征的匹配主要依赖于图像特征之间的相似度测量<sup>[4-10]</sup>,而图像特征的提取都是基于图像内容的底层特征<sup>[11-15]</sup>,如图像的颜色、形状和纹理特征等。假设图像数据集中所有图像的特征向量组成的数据集为: $\mathbf{I} = \{\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_n\}$ ,  $n$ 代表图像的个数,查询图像所对应的特征向量为: $\mathbf{I}_i = \{\mathbf{I}_{i1}, \mathbf{I}_{i2}, \dots, \mathbf{I}_{im}\}$ ;被检索图像所对应的特征向量为: $\mathbf{I}_j = \{\mathbf{I}_{j1}, \mathbf{I}_{j2}, \dots, \mathbf{I}_{jm}\}$ ,  $m$ 为图像特征向量的维数。常用的传统相似度测量方法可表示为以下3种形式:

(1) 基于欧氏距离(Euclidean distance)的相似度测量方法,表达式为

$$S(\mathbf{I}_i, \mathbf{I}_j) = 1 - \left[ \sum_{h=1}^m |\mathbf{I}_{ih} - \mathbf{I}_{jh}|^2 \right]^{\frac{1}{2}} \quad (1)$$

(2) 基于曼哈顿距离(Cityblock distance)的相似度测量方法,表达式为

$$S(\mathbf{I}_i, \mathbf{I}_j) = 1 - \sum_{h=1}^m |\mathbf{I}_{ih} - \mathbf{I}_{jh}| \quad (2)$$

(3) 基于切比雪夫距离(Chebychev distance)的相似度测量方法,表达式为

$$S(\mathbf{I}_i, \mathbf{I}_j) = 1 - \max_{h=1}^m |\mathbf{I}_{ih} - \mathbf{I}_{jh}| \quad (3)$$

在CBIR中使用这些传统的相似度测量方法时,计算简单、图像检索速度也比较快,但是CBIR的图像检索性能往往比较差,图像召回率和查准率都比较低。

为了提高CBIR的检索性能,本文提出了一种有效的相似度测量方法——基于图像 $k$ 近邻的相似度测量(Similarity measure based on  $k$ -nearest neighbors of images, SBkNN)方法。在这种方法中,查询图像与被检索图像的相似度通过计算这两幅图像属于同一语义(无论是哪种语义)种类的联合概率来衡量,而这个概率可分别通过分析这两幅图像与各自近邻图像的距离来得到。

## 1 图像特征提取

本文主要提取了图像的颜色特征和纹理特征。在提取图像的颜色特征时,首先将图像映射到 $HSV$ 空间,因为 $HSV$ 空间比其他颜色空间具有更好的感知均匀性。 $H$ 代表色调,用角度度量,取值范围为 $0^\circ \sim 360^\circ$ ,从红色开始按逆时针方向计算,红色为 $0^\circ$ ,绿色为 $120^\circ$ ,蓝色为 $240^\circ$ ,它们的补色分别是黄色为 $60^\circ$ ,青色为 $180^\circ$ ,品红为 $300^\circ$ ;  $S$ 代表饱和度,用来表示颜色的鲜艳程度, $S$ 取值范围为 $0.0 \sim 1.0$ , $S$ 值越大,饱和度越高;  $V$ 代表亮度,取值范围为 $0.0 \sim 1.0$ ,当 $V$ 取值为 $0$ 时即为黑色,当 $V$ 取值为 $1$ 时即为白色。为了实现图像颜色特征向量降维,本文将 $H$ 量化成 $16$ 级, $S$ 量化成 $4$ 级, $V$ 量化成 $4$ 级,这样就可以得到 $256$ 个不同的颜色值,然后分别统计图像内容中对应于这 $256$ 个不同颜色值的像素点的个数,就可得到 $256$ 维的颜色特征向量。

在提取图像的纹理特征时,本文利用了灰度共生矩阵的方法,主要统计了灰度共生矩阵的小梯度优势、大梯度优势、灰度分布的不均匀性、梯度分布的不均匀性、能量、灰度平均、梯度平均、灰度均方差、梯度均方差、相关性、灰度熵、梯度熵、混合熵、惯性和逆差矩共 $15$ 个特性,利用图像所对应的这 $15$ 个特性

的值,组成图像的 15 维的纹理特征向量。

假设提取的图像的颜色特征向量用  $\mathbf{C}$ (Color feature vector)来表示,纹理特征向量用  $\mathbf{T}$ (Texture feature vector)来表示,则一幅图像所对应的特征向量  $\mathbf{I}_i$  就可表示为

$$\mathbf{I}_i = [\omega_c \times \mathbf{C}, \omega_t \times \mathbf{T}]$$

$$\omega_c + \omega_t = 1 \quad (4)$$

式中:  $\omega_c$  和  $\omega_t$  分别代表颜色特征和纹理特征在图像特征向量中的权重,其取值范围为(0~1)。为了进一步提高 CBIR 图像检索的性能,可对颜色特征和纹理特征在图像特征向量中所占的权重进行适当的调整,以得到比较满意的检索结果。

## 2 基于图像 $k$ 近邻的相似度测量方法

本文提出的基于图像  $k$  近邻的相似度测量(SBkNN)方法中,衡量查询图像与被检索图像的相似度的具体步骤如下所述:

(1)提取查询图像与图像数据集中所有图像的特征向量。

(2)计算查询图像  $\mathbf{I}_i$  和图像数据集中每一幅被检索图像特征向量之间的距离(可选用传统距离计算方法:欧式距离、曼哈顿距离和切比雪夫距离中的一种),根据图像特征向量之间的距离,在图像数据集中寻找查询图像  $\mathbf{I}_i$  的  $k$ (文中取  $k=10$ )个近邻图像。

(3)根据式(5)来计算查询图像  $\mathbf{I}_i$  为某一种语义  $c$ (无论哪种语义)的概率,即

$$P(L_i = c | \mathbf{I}_i) = \frac{\sum_{\mathbf{I}_{ik} \in kNN(\mathbf{I}_i), \Delta L_{ik} = c} \sigma(D(\mathbf{I}_i, \mathbf{I}_{ik}))}{\sum_{\mathbf{I}_{ik} \in kNN(\mathbf{I}_i)} \sigma(D(\mathbf{I}_i, \mathbf{I}_{ik}))} \quad (5)$$

其中

$$\sigma(D(\mathbf{I}_i, \mathbf{I}_{ik})) = [1 + \log(1 + D(\mathbf{I}_i, \mathbf{I}_{ik}))]^{-1} \quad (6)$$

(4)计算任意一幅被检索图像  $\mathbf{I}_j$  与图像数据集中每一幅被检索图像特征向量之间的距离(与第 1 步用到的距离计算方法相同),根据图像之间的距离在图像数据集中寻找被检索图像  $\mathbf{I}_j$  的  $k$  个近邻图像(此处  $k$  的取值与第(2)步中  $k$  的取值相同)。

(5)根据式(7)计算每一幅被检索图像  $\mathbf{I}_j$  为语义  $c$ (和查询图像  $\mathbf{I}_i$  相同的语义)的概率。

$$P(L_j = c | \mathbf{I}_j) = \frac{\sum_{\mathbf{I}_{jk} \in kNN(\mathbf{I}_j), \Delta L_{jk} = c} \sigma(D(\mathbf{I}_j, \mathbf{I}_{jk}))}{\sum_{\mathbf{I}_{jk} \in kNN(\mathbf{I}_j)} \sigma(D(\mathbf{I}_j, \mathbf{I}_{jk}))} \quad (7)$$

其中

$$\sigma(D(\mathbf{I}_j, \mathbf{I}_{jk})) = [1 + \log(1 + D(\mathbf{I}_j, \mathbf{I}_{jk}))]^{-1} \quad (8)$$

(6)根据式(9)计算查询图像  $\mathbf{I}_i$  与图像数据集中任意一个被检索图像  $\mathbf{I}_j$  之间的相似度。

$$S(\mathbf{I}_i, \mathbf{I}_j) = \sum_c (P(L_i = c | \mathbf{I}_i) \cdot P(L_j = c | \mathbf{I}_j)) \quad (9)$$

式(5~9)中:  $L_i$  和  $L_j$  分别代表查询图像  $\mathbf{I}_i$  和被检索图像  $\mathbf{I}_j$  的语义标签;  $\mathbf{I}_{ik} \in kNN(\mathbf{I}_i), \Delta L_{ik} = c$  表示图像  $\mathbf{I}_{ik}$  为查询图像  $\mathbf{I}_i$  的一个近邻图像且其语义为  $c$ (无论哪种语义);  $D(\mathbf{I}_i, \mathbf{I}_{ik})$  和  $D(\mathbf{I}_j, \mathbf{I}_{jk})$  分别代表查询图像  $\mathbf{I}_i$  和被检索图像  $\mathbf{I}_j$  与其近邻图像之间的距离,  $D(\mathbf{I}_i, \mathbf{I}_{ik})$  和  $D(\mathbf{I}_j, \mathbf{I}_{jk})$  的值越小,所对应的  $\sigma(D(\mathbf{I}_i, \mathbf{I}_{ik}))$  和  $\sigma(D(\mathbf{I}_j, \mathbf{I}_{jk}))$  值就越大,两个图像属于同一种语义  $c$ (无论哪种语义)的概率也越大,两个图像的相似度也就越大。

## 3 实验验证

为了验证本文所提出的 SBkNN 方法的有效性,在 Corel5k 图像数据集上对基于传统相似度测量方

法的CBIR和基于SBkNN方法的CBIR进行了对比,所有的实验都是在MATLAB环境下运行。

### 3.1 图像数据集

Corel5k 图像数据集包含了科雷尔(Corel)公司收集整理的 5 000 幅图像,是图像检索实验中广泛使用的标准数据集。该数据集由 50 张 CD 组成,每一张 CD 对应一个语义主题,为简单起见,本文分别将这 50 张 CD 所对应的语义主题按顺序编码为 1,2,3,...,50。每张 CD 中包含有 100 幅图像并按照顺序进行编码,且每张 CD 中的图像起止编码都不一样,这样就可以通过每幅图像的编码来判断该幅图像所归属的语义主题。实验过程中,本文将 Corel5k 图像数据集划分为 3 部分:4 000 幅图像组成训练集,500 幅图像组成验证集,用来估计数学模型中的参数设置是否恰当;其余 500 幅图像组成测试集,用来评价数学模型所对应的算法性能的好坏。利用验证集找到数学模型的最优参数后,再将训练集中的 4 000 幅图像和验证集中的 500 幅图像混合起来组成新的训练集。

### 3.2 实验结果对比

本文主要利用图像召回率和查准率这两个标准参数对基于不同距离、不同相似度测量的 CBIR 方法进行了实验结果对比。图像召回率和查准率的计算方法分别如式(10,11)所示。

$$\text{召回率} = \frac{\text{检索到的相关图像数}}{\text{图像数据集中总的相关图像数}} \quad (10)$$

$$\text{查准率} = \frac{\text{检索到的相关图像数}}{\text{检索到的总的图像数}} \quad (11)$$

为了进行实验对比,本文分别从 25 种不同语义主题的 CD 上任意选择一幅图像作为查询图像,并利用这 25 幅图像组成查询图像数据集,将查询图像数据集输入到 CBIR 系统后,根据所检索到的对应于每一幅查询图像的前 50 幅图像来计算对应于每一幅查询图像的召回率和查准率。实验中最大召回率和查准率分别为:召回率 =  $\frac{50}{100} = 0.5$ ,查准率 =  $\frac{50}{50} = 1$ 。实验中所使用的 25 个查询图像如图 1 所示。



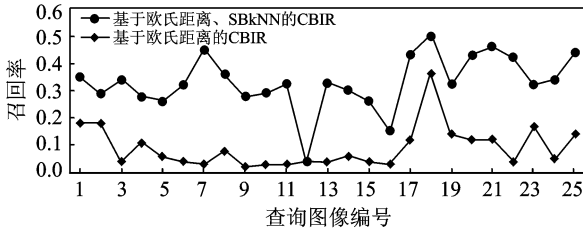
图 1 实验室中所有查询图像

Fig.1 All query images in experiment

#### 3.2.1 基于传统相似度测量方法与 SBkNN 方法的 CBIR 性能对比

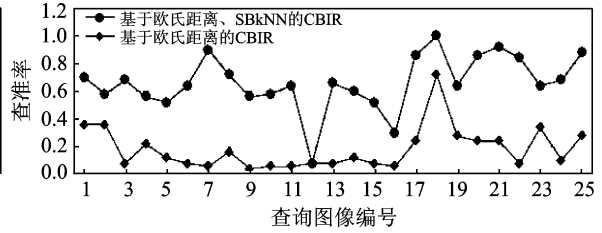
本节实验过程中取颜色特征权重  $w_c = 0.5$ ,则纹理特征权重为  $w_t = 1 - w_c = 0.5$ 。实验结果对比如图 2~4 所示。

从图 2~4 中可以看出,本文所提出的 SBkNN 方法应用欧氏距离、曼哈顿距离和切比雪夫距离中的任何一种距离来确定查询图像和每一幅被检索图像的近邻图像,基于 SBkNN 方法的 CBIR 比基于传统相似度测量方法的 CBIR 性能好。从图 5 中可以看出,基于曼哈顿距离、SBkNN 方法的 CBIR 性能最好。本文将 3 种基于不同距离、SBkNN 的 CBIR 性能进行对比,对比结果如图 5 所示。



(a) 召回率对比

(a) Comparison of recall

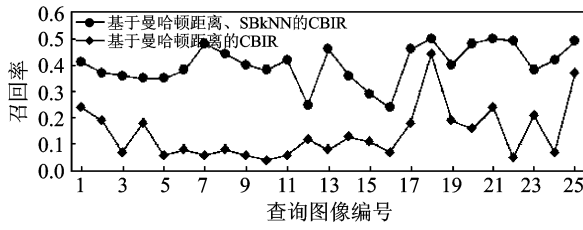


(b) 查准率对比

(b) Comparison of precision

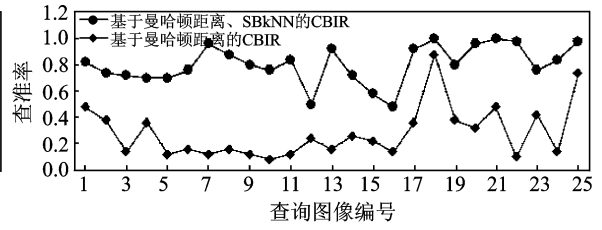
图2 基于欧式距离与基于欧式距离、SBkNN 的CBIR 的召回率和查准率对比

Fig. 2 Comparison of recall and precision between CBIRs based on Euclidean distance and based on Euclidean distance and SBkNN



(a) 召回率对比

(a) Comparison of recall

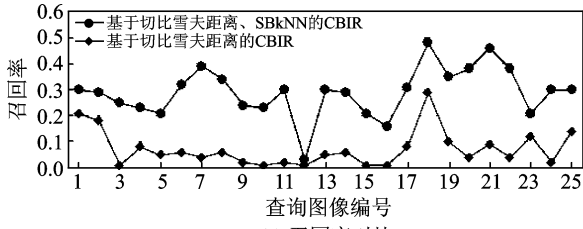


(b) 查准率对比

(b) Comparison of precision

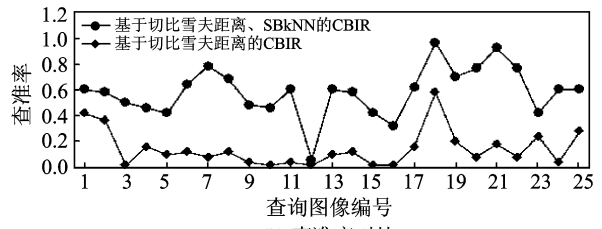
图3 基于曼哈顿距离与基于曼哈顿距离、SBkNN 的CBIR 的召回率和查准率对比

Fig. 3 Comparison of recall and precision between CBIRs based on Cityblock distance and based on Cityblock distance and SBkNN



(a) 召回率对比

(a) Comparison of recall

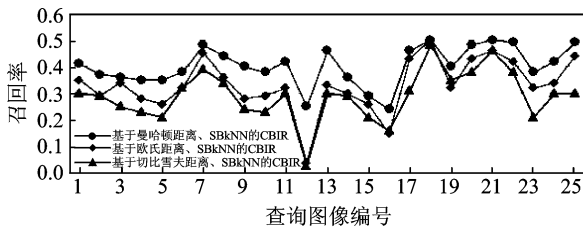


(b) 查准率对比

(b) Comparison of precision

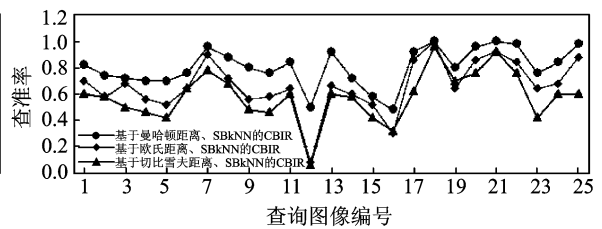
图4 基于切比雪夫距离与基于切比雪夫距离、SBkNN 的CBIR 的召回率和查准率对比

Fig. 4 Comparison of recall and precision between CBIRs based on Chebychev distance and based on Chebychev distance and SBkNN



(a) 召回率对比

(a) Comparison of recall



(b) 查准率对比

(b) Comparison of precision

图5 基于不同距离、SBkNN 的CBIR 的召回率和查准率对比

Fig. 5 Comparison of recall and precision between CBIRs based on different distance and SBkNN

3.2.2 k取不同值时基于SBkNN的CBIR性能对比

本节实验中选取了基于曼哈顿距离、SBkNN方法的CBIR来对图像的邻数k的取值进行对比,如图6所示。

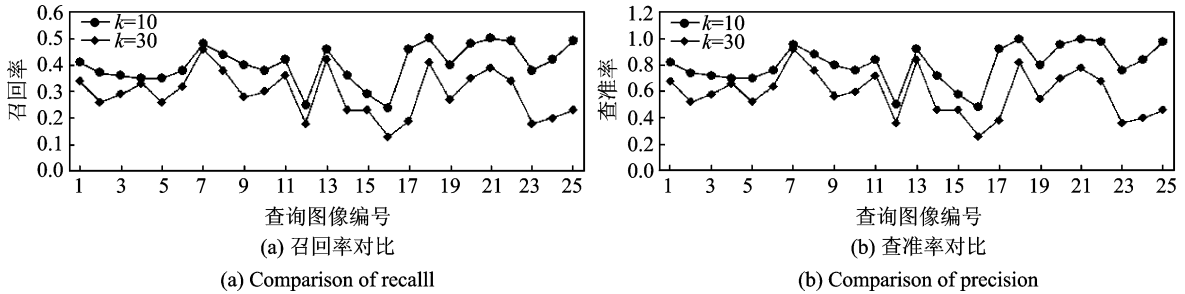


图6 k=10与k=30时基于曼哈顿距离、SBkNN的CBIR的召回率和查准率对比

Fig. 6 Comparison of recall and precision between CBIRs based on Cityblock distance and SBkNN with k=10 and k=30

从图6中可以看出,在基于曼哈顿距离、SBkNN的CBIR中,当图像邻数k=10时,要比k=30时的检索性能要好。这也说明,在基于SBkNN的CBIR中,并不是说查询图像的邻数取得越多越好,k应取适当的值。k取值越小,基于SBkNN的CBIR的计算复杂度越低,图像检索速度也就越快。

3.2.3 基于不同相似度测量方法的CBIR结果对比

本节实验中仍取k=10,wc=0.5,当输入如图7所示的查询图像时,基于曼哈顿距离的CBIR的图像检索结果如图8所示,基于曼哈顿距离、SBkNN方法的CBIR的图像检索结果如图9所示。从图8,9可以看出,基于曼哈顿距离、SBkNN方法的CBIR的图像检索结果优于基于曼哈顿距离的CBIR的图像检索结果。



图7 一幅查询图像

Fig. 7 One query image

3.2.4 基于不同相似度测量方法的CBIR的检索时间复杂度的对比

本节实验中仍取k=10,wc=0.5,当输入任意一幅查询图像时,基于曼哈顿距离的CBIR的和基于曼哈顿距离、SBkNN的CBIR的检索时间复杂度对比如图10所示。

从图10中可以看出,基于曼哈顿距离、SBkNN的CBIR的检索时间复杂度高于基于曼哈顿距离的

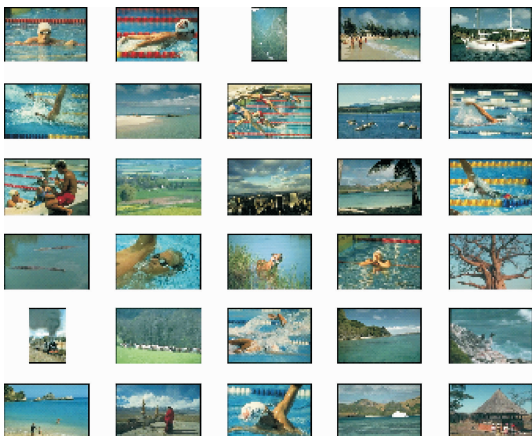


图8 基于曼哈顿距离的CBIR的检索结果

Fig. 8 Retrieval results of CBIR based on Cityblock distance



图9 基于曼哈顿距离、SBkNN的CBIR的检索结果

Fig. 9 Retrieval results of CBIR based on Cityblock distance and SBkNN

CBIR的时间复杂度,但是当检索图像数比较少时,两种CBIR的检索速度相差不是太大。因此,基于SBkNN的CBIR比较适合于小的图像数据集。当需要检索的图像数比较多时可采取分批检索的方式,以降低基于SBkNN的CBIR的时间复杂度。

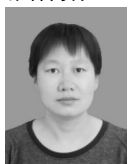
## 4 结束语

本文为CBIR系统提出一种新的、有效的相似度测量方法——基于图像 $k$ 近邻的相似度测量(SBkNN)方法。通过在Corel5k数据集上的实验可以看出,基于SBkNN的CBIR检索性能明显优于基于传统相似度测量方法的CBIR。但是基于SBkNN的CBIR计算复杂度较高、检索速度比较慢,因此它比较适合于较小的数据集,对于大数据集可以采取分批检索的方式来提高CBIR的检索速度。

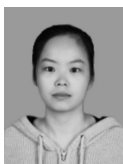
## 参考文献:

- [1] Li W, Duan L, Xu D, et al. Text-based image retrieval using progressive multi-instance learning[C]// Proceedings of the IEEE Computer Vision (ICCV 2011). Barcelona:[s. n.], 2011: 2049-2055.
- [2] Singhai N, Shandilya S. A survey on: Content based image retrieval systems[J]. International Journal of Computer Applications, 2010, 4(2):22-26.
- [3] Liu Guanghai, Zhang Lei, Hou Yingkun, et al. Image retrieval based on multi-texton histogram[J]. Pattern Recognition, 2010, 43(7): 2380-2389.
- [4] Sun Huifeng, Peng Yong, Chen Junliang, et al. A new similarity measure based on adjusted Euclidean distance for memory-based collaborative filtering[J]. Journal of Software, 2011, 6(6): 993-1000.
- [5] Huang Wei, Zhang Peng, Wan Min. A novel similarity learning method via relative comparison for content-based medical image retrieval[J]. J Digit Imaging, 2013, 26(5): 850-865.
- [6] Min R, Cheng H D. Effective image retrieval using dominant color descriptor and fuzzy support vector machine[J]. Pattern Recognition, 2009, 42(1):147-157.
- [7] Zhong Shaobo, Cheng Yimin, Qu Xilong. Divergence color histogram for content-based image retrieval[J]. Applied Mechanics and Materials, 2011, 50/51:639-643.
- [8] Wang Xiangyang, Yu Yongjian, Yang Hongying. An effective image retrieval scheme using color, texture and shape features [J]. Computer Standards & Interfaces, 2011, 33(1):59-68.
- [9] Liu Guanghai, Li Zuoyong, Zhang Lei, et al. Image retrieval based on micro-structure descriptor [J]. Pattern Recognition, 2011, 44(9):2123-2133.
- [10] Liu Guanghai, Yang Jingyu. Content based-image retrieval using color difference histogram[J]. Pattern Recognition, 2013, 46(1):188-198.
- [11] Singha M, Hemachandran K. Content based image retrieval using color and texture[J]. Signal and Image Processing: An International Journal (SIPIJ), 2012, 3(1):39-57.
- [12] Murala S, Maheshwari R P, Balasubramanian R. Local tetra patterns: A new feature descriptor for content-based image retrieval[J]. IEEE Transactions on Image Processing, 2012, 21(5): 2874-2886.
- [13] Wang Xiangyang, Zhang Beibei, Yang Hongying. Content-based image retrieval by integrating color and texture features [J]. Multimedia Tools Applications, 2014, 68(3): 545-569.
- [14] 熊伟, 龚勋, 罗俊, 等. 基于局部纹理特征的超声甲状腺节良恶性识别[J]. 数据采集与处理, 2015, 30(1): 186-191. Xiong Wei, Gong Xun, Luo Jun, et al. Ultrasound thyroid images classification based on local texture features[J]. Journal of Data Acquisition and Processing, 2015, 30(1): 186-191.
- [15] Liu Guanghai, Yang Jingyu, Li Zuoyong. Content based-image retrieval using computational visual attention model[J]. Pattern Recognition, 2015, 48(8):2554-2566.

## 作者简介:



王春静(1978-),女,讲师,研究方向:机器学习、模式识别, E-mail: 61899029@qq.com.



许圣梅(1993-),女,本科,研究方向:机器学习。

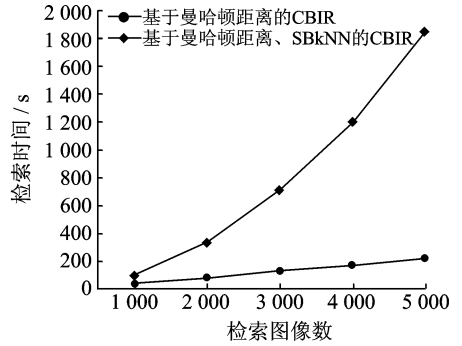


图 10 基于曼哈顿距离的CBIR和基于曼哈顿距离、SBkNN的CBIR的时间复杂度对比

Fig. 10 Comparison of time complexity between CBIRs based on Cityblock distance and CBIR based on Cityblock distance and SBkNN

