

# 一种改进的图像相似度算法

邹承明 薛 栋 郭双双 赵广辉

(武汉理工大学计算机科学与技术学院 武汉 430070)

**摘要** 图像相似度算法在图像识别、图像搜索引擎等研究领域具有重要意义。针对传统的灰度颜色直方图算法无法准确地描述各种颜色在图像中的分布情况的问题,提出了一种改进的图像相似度算法。它融合图像的纹理特征,利用灰度共生矩阵来提取图像像素在图像各个位置的特征信息。实验表明,这种融合图像纹理特征的方法不仅保留了灰度颜色直方图算法执行效率高的特点,而且弥补了颜色直方图算法的不足,进而提高了算法的准确性。在实际的应用场景中,可以通过调整两种算法的权值,来进一步提高算法的准确性。

**关键词** 图像相似度,灰度级,颜色直方图,纹理特征,应用场景

**中图分类号** TP391.41 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.6.015

## Improved Image Similarity Algorithm

ZOU Cheng-ming XUE Dong GUO Shuang-shuang ZHAO Guang-hui

(College of Computer Science and Technology, Wuhan University of Technology, Wuhan 430070, China)

**Abstract** Image similarity algorithm has important significance in image recognition, image search engines and other research areas. Traditional gray color histogram algorithm can not accurately describe the distribution of each color in the image. To address this issue, an improved image similarity algorithm was proposed. It fuses texture feature of image, and extracts image pixels' feature information at various positions on the image by using gray level co-occurrence matrix. Experimental results show that the method of fusing image texture features not only retains the characteristics of efficient implementation of gray color histogram algorithm, but also can improve the accuracy of the algorithm. In practical application scenarios, we can adjust the weights of two kinds of algorithm to further improve the accuracy of the algorithm.

**Keywords** Image similarity, Gray-level, Color histogram, Textural feature, Application scenarios

## 1 引言

随着移动互联网的高速发展,人们可以快速方便地获取图像资源,如何从这些海量的图像资源里找到一些相似的图像显得非常重要。1991年,Swain<sup>[1]</sup>提出了颜色直方图算法,并且用它来进行图像检索。该算法的优点是计算简单,而且对图像不需要做太多的预处理,对图像的尺寸也没有严格要求,但是由于图像的颜色直方图表示的是每种颜色出现的概率,它没有指明某一种颜色在图像中的具体位置,因此当计算两幅图像的相似度时,就会造成偏差;而且该算法对颜色比较单一的两幅图像也不能做出很好的判断。针对上面提到的问题,许多学者从各个不同的角度对该算法进行了改进。文献[2]通过计算以同一颜色构成的最大联通区域及其边缘颜色粗糙度,获取了图像颜色的分布情况,但是这种方法计算量大、计算复杂。文献[3]结合纹理和颜色这两个图像的低层特征来提高图像检索精度,但是其对复杂图像的纹理描述比较

粗糙,从而导致对复杂图像的检索率不高。文献[4]通过构造HSI空间颜色直方图,然后使用累加直方图方法进行匹配,这种方法能提高图像检索的准确率,但是使算法的执行效率较低。

通过研究发现,从图像低层特征来研究图像相似度,仍然是一种非常有效的方法。所以本文结合图像最重要的两个低层特征:颜色和纹理,来计算两幅图像的相似度。通过这种方法,最大限度地提高图像检索的准确性,使得颜色直方图算法的鲁棒性更强,准确率更高。

## 2 基于灰度颜色直方图的相似度算法

颜色直方图算法具有计算简单、执行效率高、旋转不变性、对图像的大小没有限制等诸多优良特点,所以常常被用来进行图像相似度、图像检索的研究。

### 2.1 颜色直方图的定义

基于颜色直方图的图像相似度算法是从图像颜色特征和

到稿日期:2015-07-13 返修日期:2015-08-28 本文受国家自然科学基金(51179146),中央高校基本科研业务费专项基金(2014-VII-027),湖北省科技支撑计划(2015BAA120),湖北省科技支撑计划(2015BCE068)资助。

邹承明(1975-),男,博士,教授,CCF会员,主要研究方向为计算机视觉、图像处理;薛 栋(1989-),男,硕士生,主要研究方向为计算机视觉、移动互联网,E-mail:1162400912@qq.com;郭双双(1991-),女,硕士生,主要研究方向为图像处理、嵌入式系统;赵广辉(1973-),男,博士,副教授,主要研究方向为图像处理、移动互联网。

统计学方面来研究图像相似度。它描述的是每一个像素在图像中出现的频次,并没有表明某种像素在图像中的具体位置。

对于一幅数字图像,计算每一种像素在整幅图像中出现的频次,以像素出现的频次作为颜色直方图的纵坐标,以颜色值作为横坐标。通过这种方式就可以描绘出该图像的颜色直方图,可以用式(1)和式(2)表示。

$$\begin{cases} H=[(h[c_1]),(h[c_2]),\dots,(h[c_k]),\dots,(h[c_n])] \\ \sum_{k=1}^n h[c_k]=1, 0 \leq h[c_k] \leq 1 \end{cases} \quad (1)$$

其中, $h[c_k]$ 表示第  $k$  种像素在图像中出现的频次。

$$h[c_k]=\frac{\sum_{i=0}^M \sum_{j=0}^N \delta(I[i,j]=c_k)}{W} \quad (2)$$

$$\delta(x)=\begin{cases} 1, & \text{若 } x \text{ 为 true} \\ 0, & \text{若 } x \text{ 为 false} \end{cases}$$

其中, $M$ 和 $N$ 为图像的宽和高, $W$ 为图像的像素数。

## 2.2 算法描述

- 1)为了降低算法的时间复杂度,本文直接将彩色图像进行灰度化处理,得到彩色图像的灰度图像;
- 2)创建一维灰度直方图,使统计图像分布在 $[0, 255]$ 像素;
- 3)计算灰度图像的一维直方图;
- 4)对上面计算出来的一维直方图做归一化处理;
- 5)创建一张图,这张图主要是用于显示各个像素的直方图,横坐标为灰度级,纵坐标为像素个数;
- 6)将每个直方图绘制到步骤5)创建的图中;
- 7)对两张图片做步骤1)~6)的处理。

## 2.3 相似性度量

通过度量两幅图像的距离来计算图像的相似度,最终根据计算出的数值来判断图像的相似程度。文中通过度量两幅图像直方图的巴氏距离来计算图像的相似度,巴氏距离定义为:

$$S(H_1, H_2)=1-\sqrt{1-\frac{\sum_i \sqrt{H_1(i) * H_2(i)}}{\sum_i H_1(i) * \sum_i H_2(i)}} \quad (3)$$

其中, $H_1, H_2$ 为提取出来的两个颜色特征向量; $S$ 的值越靠近1,表示两张图像越相似;越靠近0,表示两张图像越不相似。

## 3 基于灰度共生矩阵的图像相似度算法

灰度共生矩阵是由文献[5]提出的一种提取图像纹理特征的算法,它反映了图像中不同距离的像素对在不同方向上的移动。

文中所提算法主要包括以下几个主要步骤:

- 1)根据灰度颜色直方图算法计算图像相似度,最后把计算出来的相似度记为 $s_1$ (算法过程在上一节中有详细描述);
- 2)采用灰度共生矩阵提取图像纹理特征,然后计算图像的相似度,最后把计算出来的相似度记为 $s_2$ ;
- 3)根据前两步计算出来的图像相似度,采用加权平均计算最终的图像相似度。

### 3.1 基于灰度颜色直方图计算图像相似度

对于输入的图像,采用第2节介绍的算法计算图像的相似度。

### 3.2 基于灰度共生矩阵计算图像相似度

#### 3.2.1 构造灰度共生矩阵

灰度共生矩阵的构造过程:假设灰度共生矩阵可以表示为从灰度值  $i$  的点经过距离  $d=(a,b)$  到达灰度值为  $j$  的点的概率。灰度共生矩阵用  $p=(i,j,d,\theta) (i,j=0,1,2,\dots,K-1)$  表示, $i,j$  分别表示像素对的灰度值, $d$  表示像素对之间的距离, $\theta$  的取值一般为: $0^\circ, 45^\circ, 90^\circ, 135^\circ$ 。距离  $d$  和角度  $\theta$  的关系如图1所示<sup>[6]</sup>。

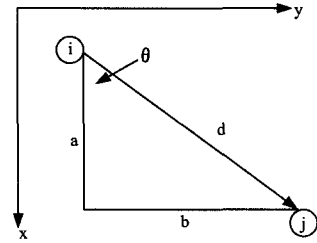


图1 距离  $d$  和角度  $\theta$  的关系图

根据图1,能清楚地通过  $a, b$  的值确定像素对之间的距离和移动的角度。

当  $a=1, b=0$  时,这个像素对是水平方向上的移动,即  $0^\circ$  扫描,并且这两个像素对之间的距离为1。

当  $a=1, b=1$  时,这个像素对是右对角线上的移动,即  $45^\circ$  扫描,并且这两个像素对之间的距离为1。

当  $a=0, b=1$  时,这个像素对是垂直方向上的移动,即  $90^\circ$  扫描,并且这两个像素对之间的距离为1。

当  $a=-1, b=1$  时,这个像素对是左对角线上的移动,即  $135^\circ$  扫描,并且这两个像素对之间的距离为1。

为了更加清楚构造灰度共生矩阵的过程,下面举例说明。假设有一个  $6 \times 6$  的矩阵,如图2所示。

3	8	9	9	2	1
0	3	11	3	11	12
9	8	1	7	8	9
10	4	9	11	0	2
1	9	6	2	4	2
1	0	8	0	4	1

图2 原图像转换为灰度图像的矩阵

图2为原图像,它的最大灰度级是16。为了表示方便,把灰度级数降低到4级,构造的灰度共生矩阵就是  $4 \times 4$  大小的矩阵。图2转化后如图3所示。

0	2	2	2	0	0
0	0	2	0	2	3
2	2	0	1	2	2
2	1	2	2	0	0
0	2	1	0	1	0
0	0	2	0	1	0

图3 灰度级为4的图像的矩阵

选择  $(a=1, b=0)$  距离为1、方向  $0^\circ$  的移动,构造的灰度共生矩阵如图4所示。

	0	1	2	3
0	4	1	5	0
1	1	0	2	0
2	5	1	5	1
3	0	0	0	0

图4  $(a=1, b=0)$  时的灰度共生矩阵

选择( $a=2, b=0$ )距离为 2、方向  $0^\circ$  的移动,构造的灰度共生矩阵如图 5 所示。

	0	1	2	3
0	3	1	4	1
1	0	1	2	0
2	6	2	2	0
3	0	0	0	0

图 5 ( $a=2, b=0$ )时的灰度共生矩阵

### 3.2.2 灰度共生矩阵的特征值

根据文献[5]提出的 14 个图像纹理特征,并结合多个文献的研究结果,最终选择 4 个来描述图像的纹理特征。根据文献[7]的实验结果可以得出,对比度、熵和逆差矩这 3 个特征具有最强的识别能力。文献[8]实验中提到这 14 个纹理特征中有 4 个特征(对比度、逆差矩、熵、二阶矩)是互不相关的。结合文献[7]、文献[8]的研究成果,最终选择 4 个(熵、二阶矩、对比度、逆差矩)最具代表性的特征来计算两幅图像的相似度<sup>[9]</sup>。

1) 熵  $G_1$

$$G_1 = - \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} P(i, j, d, \theta) \log P(i, j, d, \theta) \quad (4)$$

熵是对图像纹理复杂程度和非均匀程度的一种数学度量。

2) 二阶矩(能量)  $G_2$

$$G_2 = \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} [P(i, j, d, \theta)]^2 \quad (5)$$

二阶矩是灰度共生矩阵各个元素值的平方和,反映图像纹理的粗细程度和灰度的分布情况。

3) 对比度  $G_3$

$$G_3 = \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} (i-j)^2 P(i, j, d, \theta) \quad (6)$$

对比度反映图像的清晰程度。

4) 逆差矩  $G_4$

$$G_4 = \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} \left[ \frac{P(i, j, d, \theta)}{1+(i-j)^2} \right] \quad (7)$$

逆差矩用来度量图像局部纹理变化的多少。如果逆差矩的值大,就说明图像纹理的不同区域间变化小或者图像局部均匀。

通过式(4)~式(7),可以分别计算出图像的特征值。让像素对在这 4 个方向( $0^\circ, 45^\circ, 90^\circ, 135^\circ$ )上移动,分别计算出 4 个特征值;最后把这 4 个方向上计算出来的特征值组合成一个大小为 16 的综合向量,这个综合向量可以定义为:

$$G = [G_{1k}, G_{2k}, G_{3k}, G_{4k}], k \in (0^\circ, 45^\circ, 90^\circ, 135^\circ) \quad (8)$$

再根据得到的两个综合特征向量计算两幅图像的相似度,记为  $s_2$ 。

### 3.2.3 计算图像的相似度

通过上文提到的颜色直方图算法和灰度共生矩阵算法可以计算出图像的两个相似度,然后对这两个相似度进行加权平均,进一步得到平均后的图像相似度。最终的相似度定义为:

$$\begin{cases} s = k_1 * s_1 + k_2 * s_2 \\ k_1 + k_2 = 1 \end{cases} \quad (9)$$

其中,  $s_1$  为颜色直方图算法计算出来的图像相似度,  $s_2$  为灰度共生矩阵计算出来的图像相似度。  $k_1, k_2$  分别为两种相似度算法的权重,可以根据具体的场景通过大量的学习训练来取得合适的值。

## 4 实验结果分析

为了验证上述算法的准确性和有效性,在 Intel Core i3 CPU, 2.26GHz, 4GB RAM 环境下利用 opencv 库,基于多对不同图像数据,进行仿真对比实验。实验图像来源于 INRIA (法国国家信息与自动化研究所)。

### 4.1 实验一

实验选用一组(3 张)不同的  $2816 \times 2112$  图像进行多次实验(最终数据取的是多次实验的平均值),对改进后的算法和直方图算法、灰度共生矩阵算法分别进行比较。

实验图像如图 6 所示。

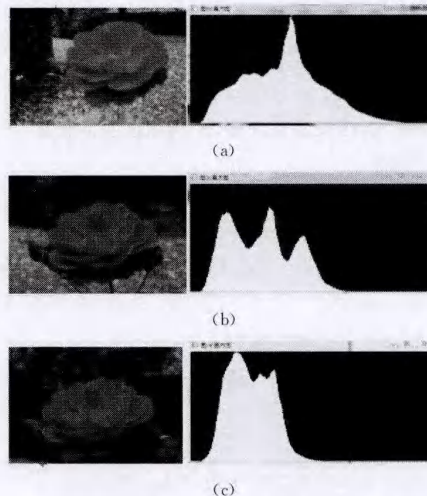


图 6 实验一图像及其对应的颜色直方图

图 6 是 3 张相似的图像,其中左边为原图像,右边的图像是左边图像对应的颜色直方图。这个颜色直方图横坐标是像素的范围,本文实验采用 256 种灰度值;纵坐标表示的是各个像素出现的次数,也就是各个像素占整幅图像的比重。观察这 3 幅图像的颜色直方图,发现这 3 幅图像的像素大部分都分布在前部,这与肉眼观察的结果是一致的。肉眼观察到 3 幅图像颜色比较单一,而且图像都比较偏暗,所以说明上面 3 幅图像是颜色比较单一的图像。还可以对 3 幅图像两两进行比较,不难发现图 6(b)与图 6(c)最相似,其它次之。接下来根据实验数据分析这 3 张图像,实验结果如表 1 所列。实验参数设定:灰度颜色直方图采用的是 256 种颜色的灰度图像;采用  $4 \times 4$  ( $k=16$ ) 大小的灰度共生矩阵,像素对之间的距离  $d=8$ ;像素对移动角度  $\theta=(0^\circ, 45^\circ, 90^\circ, 135^\circ)$ ;改进后算法的权重:  $k_1=0.5, k_2=0.5$ 。

表 1 实验结果(一)

两两比较		相似度(%)
图 6(a)	颜色直方图算法	73.3
图 6(b)	灰度共生矩阵算法	77.1
	改进后算法	75.2
图 6(b)	颜色直方图算法	71.9
图 6(c)	灰度共生矩阵算法	88.2
	改进后算法	80.05
图 6(a)	颜色直方图算法	54.8
图 6(c)	灰度共生矩阵算法	45.2
	改进后算法	50.0

肉眼观察图 6(b)与图 6(c)最相似,但是根据表 1 的实验数据分析,通过灰度颜色直方图算法计算图 6(b)与图 6(c)的

相似度却没有图 6(a)与图 6(b)的高。再结合 3 张图各自的灰度直方图可知,图 6(c)的颜色分布确实太单一,而图 6(a)比图 6(c)的颜色分布得更均匀,实验结果进一步验证了颜色直方图算法确实不能很好地描述这种颜色单一的图像的特征。然而观察图 6(a)与图 6(c),通过灰度共生矩阵计算出来的相似度偏低,将在实验二中对其进行进一步说明。再观察两两图像通过灰度共生矩阵算法计算出的图像的相似度,发现图 6(b)与图 6(c)计算出的相似度为 88.2%,相似度提高了 16.3%,改进后的算法计算出的最终相似度为 80.05%,这与肉眼观察两幅图像的相似程度比较吻合。实验表明,融合图像纹理特征的方法确实能够提高灰度颜色直方图算法的准确性,而且最大限度地保留了灰度颜色直方图算法的优良特点。

#### 4.2 实验二

实验选取两张  $1920 \times 2560$  像素的图像进行多次实验(最终数据取的是多次实验的平均值),对改进后的算法、直方图算法和灰度共生矩阵算法分别进行比较,实验图像如图 7 所示。

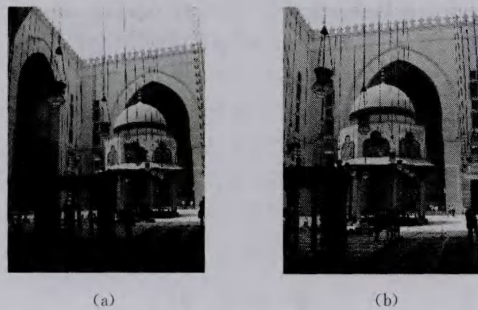


图 7 实验二所选图像

实验参数设定:颜色直方图采用的是 256 种颜色的灰度图像;采用  $k=16, 32, 64$  大小的灰度共生矩阵,像素对之间的距离  $d=3, 5, 7$ ;像素对移动角度  $\theta=(0^\circ, 45^\circ, 90^\circ, 135^\circ)$ ;改进后算法的权重为  $k_1=0.5, k_2=0.5$ 。

灰度颜色直方图记为算法 1,灰度共生矩阵算法记为算法 2,改进后的算法记为算法 3。实验结果如表 2 所列。

表 2 实验结果(二)

灰度级	距离	相似度 (%)	距离			
			1	3	5	7
16	算法 1	86.2	86.2	86.2	86.2	
	算法 2	89.6	70.6	82.6	72.6	
	算法 3	87.9	78.4	84.4	79.4	
32	算法 1	86.2	86.2	86.2	86.2	
	算法 2	97.2	93.5	91.2	86.6	
	算法 3	91.7	89.85	88.7	86.4	
64	算法 1	86.2	86.2	86.2	86.2	
	算法 2	97.3	89.7	88.1	92.2	
	算法 3	91.75	87.95	87.2	89.2	

通过上面的实验结果发现,不同的灰度级、距离会对图像相似度造成影响。分析上面的实验数据,可以得出以下几点结论:

1)通过图 8 可以看出,当图像灰度级取 32 和 64 时,在距离取 3,5,7 情况下,折线变化幅度很小,而且逐渐趋于平稳,这时计算出来的图像相似度非常稳定。在实际应用中,可以通过这种方式来获取比较稳定的图像相似度。但是图像灰度级的取值会直接影响所构造灰度共生矩阵的规模,图像灰度级取值越大,构造灰度共生矩阵所需要的时间越长,算法的执

行效率越低。考虑到算法的时间复杂度,一般需要根据自己的场景做出合适的取值。本文通过大量的实验分析比较得出:一般情况下,当灰度级为 32,距离为  $d(3 < d < 8)$  时,可以得到相对稳定的图像相似度,对于对实验精度要求高的场景,可以提高灰度级,对于对算法时间复杂度要求高的场景,可以降低灰度级。

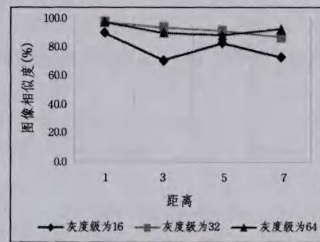


图 8 灰度共生矩阵计算的图像相似度

2)在实验一中,当像素对距离取值为 8 时,通过灰度共生矩阵计算出图 6(a)与图 6(c)的相似度为 45.2%;当像素对之间的距离取值为 3 时,图 6(a)与图 6(c)的相似度为 63.6%。这是由于像素对之间的距离取值过大,会损失像素对之间的信息,导致实验结果不准确。因此在具体的场景中,一定要通过分析比较来得到最适合的取值,才能保证算法的准确性。

3)观察实验二,根据灰度共生矩阵算法(特别是选取不同的灰度级、不同的距离)得出的图像相似度变化幅度比较大。通过观察实验二中的两幅图像,可以发现实验二两幅图像像素比较丰富。由此,可以进一步说明灰度颜色直方图算法非常适合颜色比较丰富的图像,而灰度共生矩阵算法则适合于图像纹理比较丰富的图像。综合这两种算法,可以提高图像检索的准确率。下一步将研究如何结合这两种算法来提高大数据下图像检索的准确率。

**结束语** 结合图像的颜色和纹理特征,通过灰度颜色直方图算法和灰度共生矩阵算法对图像进行相似性度量。算法并没有改变原来灰度颜色直方图算法,而是在它的基础上融合灰度共生矩阵算法来进行改进,尽可能多地保留了灰度颜色直方图算法快速、高效的特点,同时又通过图像的纹理特征进一步提高了算法的准确性。对于这两种算法的权重取值,可以根据自己的场景来进行调整,从而最大限度地提高算法的准确性。

#### 参考文献

- [1] Swain M J, Ballard D H. Color Indexing[J]. International Journal of Computer Vision, 1991, 7(1): 11-32
- [2] Yang Jun, Wang Ji-cheng, Xing Dan-jun. Image Retrieval Method Combing Integer and Blocks' Color Distribution [J]. Computer Applications, 2008, 28(3): 653-655, 658 (in Chinese) 杨珺,王继成,邢丹俊.融合整体与分块颜色分布的图像检索方法[J].计算机应用,2008,28(3):653-655,658
- [3] Xu Hui-ying, Yuan Jie, Zhao Jian-min, et al. Image Retrieval Based on Color and Texture [J]. Computer Science, 2009, 36(5): 282-286 (in Chinese) 徐慧英,袁杰,赵建民,等.一种基于颜色和纹理的图像检索方法[J].计算机科学,2009,36(5):282-286
- [4] Liu Zhong-wei, Zhang Yu-jin. Image Retrieval Using Both Color and Texture Features[J]. Journal of China Institute of Communications, 1999, 20(5): 37-41 (in Chinese)

刘忠伟,章毓晋. 综合利用颜色和纹理特征的图像检索[J]. 通信学报, 1999, 20(5): 37-41

- [5] Haralick R M, Shanmugam K, Dinstein I. Texture Features for Image Classification[J]. IEEE Transactions on Systemse, Man and Cybernetics, 1973, 3(6): 610-621
- [6] Gao Cheng-cheng, Hui Xiao-wei. GLCM-Based Texture Feature Extraction [J]. Computer Systems & Applications, 2010, 19(6): 195-198(in Chinese)
- 高程程, 惠晓威. 基于灰度共生矩阵的纹理特征提取[J]. 计算机系统应用, 2010, 19(6): 195-198
- [7] Equitz W, Niblack W. Retrieving Images from a Database Using Texture-algorithms from the QBIS System; Technical Report RJ, 9805[R]. 1994
- [8] Wang Zhi-zhi. Remote Sensing Object Classification Algorithm Based on the Fusion of Texture Features and Spectral Features

[D]. Xi'an; Xidian University, 2010(in Chinese)

- 王知鸞. 基于纹理及光谱信息融合的遥感图像分类方法研究[D]. 西安: 西安电子科技大学, 2010
- [9] Jiao Peng-peng, Guo Yi-zheng, Liu Li-juan, et al. Implementation of Gray Level Co-occurrence Matrix Texture Feature Extraction Using Matlab [J]. Computer Technology and Development, 2012, 22(11): 169-171, 175(in Chinese)
- 焦蓬蓬, 郭依正, 刘丽娟, 等. 灰度共生矩阵纹理特征提取的 Matlab 实现[J]. 计算机技术与发展, 2012, 22(11): 169-171, 175
- [10] Yuan Li-hong, Fu Li, Yang Yong, et al. Analysis of Texture Feature Extracted by Gray Level Co-occurrence Matrix [J]. Journal of Computer Applications, 2009, 29(4): 1018-1021(in Chinese)
- 苑丽红, 付丽, 杨勇, 等. 灰度共生矩阵提取纹理特征的分析[J]. 计算机应用, 2009, 29(4): 1018-1021

(上接第 34 页)

本文将负载均衡多路径路由算法(本文提出, 标记为 LBMP)和 ElasticTree 中的拓扑感知启发式算法(标记为 TAH)进行比较, 两个算法均基于胖树的拓扑特性使用一个快速启发式, 但 TAH 依赖于流划分。在这 15 次左右的实验中, LBMP 在网络吞吐量和网络延迟方面优于 TAH, 如图 3 和图 4 所示。在相同的输入负载下, 本文的算法在胖树结构变大时端到端延迟只有微小的上升, 这是因为胖树路径的平均长度变大, 而吞吐量保持基本一致, 这是因为本文算法可以均衡网络的流量负载。但是对于 TAH, 结果并不是这样的: 在更高的输入负载下, 开始形成拥塞从而导致网络吞吐量大幅度降低, 特别当网络更大时, 该现象更加明显。

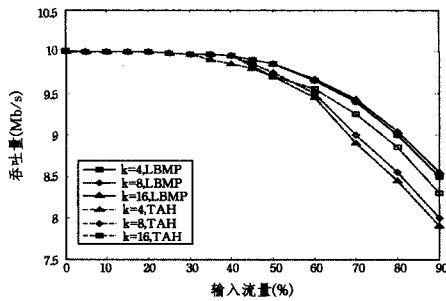


图 3 吞吐量比较

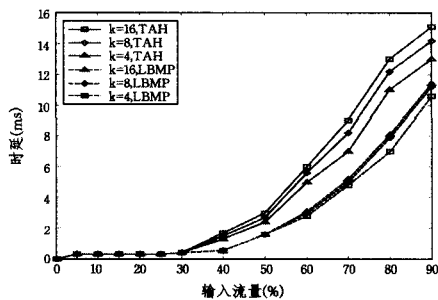


图 4 时延比较

**结束语** 具有多路径能力的胖树拓扑因其充裕带宽和高容错性, 在数据中心网络中已经成为一种很受欢迎的网络拓扑。本文提出了基于 SDN 的负载均衡多路径路由算法, 它可以在胖树拓扑中更好地分发流量负载和充分利用有效带宽。首先提出一个属于线性规划范畴的问题, 并通过整数划分问题推导证明其 NP 完全性。进而提出了一个基于 SDN 的数

据中心网络的多路径路由算法, 它是一个实用性算法。该算法利用中央控制器去收集整个网络的状态信息, 进而为每一个新进入网络的数据流做出最优的负载均衡的路由决策。最后在 Mininet 仿真平台中验证了算法的正确性和有效性。实验结果表明, 本文的算法在吞吐量和端到端延迟方面的性能均优于拓扑感知启发式多路径算法。

未来将计划做更多有效性验证实验, 特别是比较 SDN 仿真平台和真实 SDN 网络, 以及和其他多路径算法的区别; 同时使用真实流量负载进行实验。

## 参考文献

- [1] Fei Y. Introduction to the development of data center network architecture[J]. Network Security Technology & Application, 2014, 22(6): 23-28
- [2] Qiao L, Yin X H, Zhuo D I, et al. Research on SDN Network Architecture for Electric Power Big Data Platform[J]. Electric Power Information & Communication Technology, 2015, 12(6): 1-6
- [3] Sun Y, Cheng J, Shi K. Data Center Network Architecture[J]. Zte Communications, 2013, 11(5): 5-9
- [4] Zahid F, Gran E G, Bogdanski B, et al. A Weighted Fat-Tree Routing Algorithm for Efficient Load-Balancing in Infini Band Enterprise Clusters[C] // 2015 23rd Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP). IEEE, 2015: 35-42
- [5] Chmeritskiy E, Smelansky R. On QoS management in SDN by multipath routing[C] // 2014 First International Science and Technology Conference (Modern Networking Technologies) (MoNeTeC). IEEE, 2014: 1-6
- [6] Al-Fares M, Radhakrishnan S, Raghavan B, et al. Hedera: dynamic flow scheduling for data center networks[J]. Proceedings of the 7th USENIX Symposium on Networked Systems Design and Implementation, 2010, 19(1): 1-15
- [7] Garey M R, Johnson D S. Computers and Intractability: A Guide to the Theory of NP-Completeness[M]. New York, NY, USA: W. H. Freeman & Co, 1990
- [8] Ji P N, Qian D, Kanonakis K. Design and Evaluation of a Flexible-Bandwidth OFDM-Based Intra-Data Center Interconnect [J]. IEEE Journal of Selected Topics in Quantum Electronics, 2013, 19(2): 33-39