

结合用户评分和项目标签的协同过滤算法

程高伟,丁亦喆,吴振强

(陕西师范大学 计算机科学学院,陕西 西安 710062)

摘要:在信息过载时代,推荐系统能够帮助用户发现感兴趣的内容。协同过滤是推荐系统中最常用的技术,然而传统的协同过滤算法未能充分考虑项目标签对相似度的影响,因而推荐质量不高。文中提出了一种结合用户评分和项目标签的协同过滤算法,算法中关键的相似度计算是对评分相似度和标签相似度的加权,通过加权降低了相似度矩阵的稀疏性,并且保证项目之间只有在共同评分较多且标签相似时才具有较高的相似度,从而使相似度计算更加准确。通过对比实验得出加权系数在 0.3~0.5 时推荐质量较高,在公开数据集上与传统协同过滤算法的比较结果表明,文中的算法在平均绝对误差上降低了约 3%。

关键词:协同过滤;推荐系统;标签;稀疏性

中图分类号:TP39

文献标识码:A

文章编号:1673-629X(2015)03-0071-05

doi:10.3969/j.issn.1673-629X.2015.03.017

Collaborative Filtering Algorithm Combined User Ratings with Item Tags

CHENG Gao-wei, DING Yi-zhe, WU Zhen-qiang

(School of Computer Science, Shaanxi Normal University, Xi'an 710062, China)

Abstract: Recommender system can help users find their interests in the era of information overload. Collaborative filtering is one of the most widely used techniques, while the traditional collaborative filtering algorithm has a lower effectiveness because of the seldom consideration about the similarity of item tags. A collaborative filtering algorithm combined ratings with tags is proposed in this paper, the most important procedure of the algorithm is to calculate the similarity of ratings and tags. The two kinds of similarity is combined by weighting, through this method the sparsity of the similarity matrix is greatly reduced and ensures that only when the items have more common ratings and in the same tags can reach a high similarity, so that the accuracy of the calculation of similarity is highly increased. The optimal range of the weighting coefficient is 0.3 to 0.5 by comparative experiment. Compared with the traditional collaborative filtering algorithm on public data set, the algorithm proposed in this paper improves the accuracy of recommendation to 3% in MAE.

Key words: collaborative filtering; recommendation system; tags; sparsity

0 引言

推荐系统^[1-2]根据用户的兴趣特点和历史记录向用户推荐感兴趣的内容,能有效解决信息过载问题,协同过滤^[3](Collaborative Filtering, CF)技术是推荐系统中应用最成功的技术之一。在协同过滤推荐技术中一般采用近邻关系模型,它通过分析用户的历史记录计算出用户和用户或者项目和项目之间的相似度,建立起近邻关系,然后利用近邻关系中目标用户距离最近的用户对某些项目的评分来预测目标用户对特定项目的评分,最后产生相应的推荐。在近邻关系模型中

K 近邻^[4](K -Nearest Neighbor, KNN)是应用最为广泛的,它根据用户或者项目之间的相似度来生成用户或者项目的 K 近邻集合,再利用 K 近邻集合进行预测并产生推荐。随着电子商务规模的不断扩大,用户和项目数据都在不断膨胀,然而一般在系统中每一个用户只对少数项目做出评价,因此在推荐系统中数据稀疏性成为影响推荐质量的关键因素^[5]。在用户评分数据极端稀疏的情况下,使用传统的相似性度量方法计算存在一定弊端,使得推荐算法在计算最近邻居时存在误差,进而导致推荐质量的下降。为了提高推荐

收稿日期:2014-04-21

修回日期:2014-07-23

网络出版时间:2015-01-20

基金项目:国家自然科学基金面上项目(61173190);中央高校基本科研业务费专项资金(GK261001236)

作者简介:程高伟(1988-),男,硕士研究生,研究方向为计算机网络与教育;吴振强,博士,教授,研究方向为可信服务、网络编码、计算机网络与教育。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20150120.2155.001.html>

系统的推荐质量,人们提出了许多不同的推荐算法,例如 Pirasteh 等通过使用电影的类型信息和导演来降低矩阵的稀疏性和冷启动问题^[6],该方法需要使用电影类型和导演等额外信息;Pitsilis 等通过使用系统中用户评分数据建立一个假设的信任关系,通过这种方法可以解决部分冷启动问题和稀疏性问题^[7],但是该方法中的信任关系并不是真正的社会网络中的信任关系;Kumar 等采用矩阵分解技术^[8]来降低矩阵的维数,从而提高推荐系统的准确性;文献[9]中首先使用 k -means 方法对项目进行聚类,然后通过聚类之后的若干类别以及用户对项目的评分计算用户之间的相似性;Wang 等通过使用用户的评分信息对用户进行聚类,并结合 Slope One 算法对未评分项目进行预测来降低矩阵的稀疏性^[10];黄创光等对传统的 K 近邻中近邻个数 K 固定这一缺点^[11]提出了不确定近邻来进行最近邻居推荐等。

随着互联网的不断发展和完善,人们在网络中的参与度也越来越高,用户角色已经从信息的消费者转变为信息的生产者,尤其是 Web2.0 中蕴含集体智慧的力量,社会性标签^[12-13]正是集体智慧的表征之一。用户通常使用标签对资源进行描述,这种描述源自于用户对资源的主观感受,标签数据能有效反映用户对项目的喜好程度,因此在协同过滤推荐算法中,项目被用户标注的标签相似性是推荐系统中一个应该考虑的重要因素。在推荐系统中由于评分矩阵的极端稀疏性,采用传统的相似性度量方法产生的推荐质量不太高,就是由于没能充分利用标签数据,因此如何有效结合标签数据并进行相似性计算成为提高系统推荐质量的关键。在标签推荐系统中,人们主要研究如何提升系统中标签的质量,针对这一问题,研究人员对标签推荐进行了大量研究。Harvey 等通过扩展隐语义模型使用概率分布^[14]来为用户提供个性化标签推荐;张斌^[15]等通过分析标签系统中对象间关系与资源内容提出了融合关系与内容分析的社会标签推荐方法。当前的标签推荐方法大多通过分析系统中标签、用户和资源这三者之间的关系来获得系统所需要的规则,通过利用挖掘的规则进行推荐,这种基于关系的标签推荐方法仅仅利用了系统中标签、用户和资源之间的关系,并没有挖掘标签背后用户行为的动机。在已有的协同过滤推荐算法中,并没能充分考虑到标签因素对推荐质量的影响,因此文中在分析用户对项目评分的基础之上综合考虑了标签因素对项目相似度的影响,将标签相似度结合到传统的项目相似度计算之中。项目综合相似度是项目评分相似度和标签相似度的加权,通过加权系数的调整使得项目相似性矩阵的稀疏性明显降低,加权也进一步保证了只有共同评分项目

较多且标签相似的项目才会具有较高的相似度,从而提高了相似度的准确性,进而产生了更加准确的推荐结果。通过对比实验表明,该算法比传统的基于项目的协同过滤算法在 MAE 上降低了 3%。

1 基于项目的协同过滤算法

协同过滤^[16]推荐根据其他用户对项目的喜好产生对目标用户的推荐,这种推荐机制基于如下假设:如果大多数用户对某些项目的评分比较相似,则当前用户对这些项目的评分也会比较相似。协同过滤推荐利用用户的历史记录,通过分析用户兴趣爱好,使用统计技术在用户群中寻找与目标用户兴趣相似的用户,找到相似性较高的若干用户,再综合这些相似性较高用户对某些项目的评价,产生对目标用户的推荐。

基于项目的协同过滤推荐^[16-17](Item-based Collaborative Filtering, Item CF)算法根据系统中其他用户对相似项目的评价预测目标用户对特定项目的喜好,它给目标用户推荐跟他之前喜欢项目相似性较高的项目。不过,Item CF 算法的推荐并不是根据项目在内容属性上的相似性,而是通过分析用户的历史记录来计算项目之间的相似性。基于项目的协同过滤算法主要分为两步:

- (1) 计算项目之间的相似度;
- (2) 根据项目的相似度和用户的历史记录给用户生成推荐列表。

1.1 相似度度量方法

基于项目的协同过滤推荐算法的第一步就是找出当前用户未评分项目的相似项目,目前主要的项目相似性度量方法有余弦相似度度量、皮尔逊相关系数度量和修正的余弦相似度这三种度量方法。

- (1) 余弦相似度。

通过计算两个向量夹角的余弦值作为它们之间的相似性。在推荐系统中,令 r_{ui} 代表用户 u 对项目 i 的评分, r_{uj} 代表用户 u 对项目 j 的评分,则项目 i 和项目 j 之间的相似度可以表示为^[18]:

$$\text{sim}(i, j) = \cos(\vec{i}, \vec{j}) = \frac{\vec{i} \cdot \vec{j}}{\|\vec{i}\| \times \|\vec{j}\|} = \frac{\sum_{u \in U} r_{ui} \cdot \sum_{u \in U} r_{uj}}{\sqrt{\sum_{u \in U} r_{ui}^2} \sqrt{\sum_{u \in U} r_{uj}^2}} \quad (1)$$

- (2) 皮尔逊相关系数。

皮尔逊相关系数是一种线性相关系数,广泛用于度量两个变量之间的相关程度。变量之间的皮尔逊相关系数通过两个变量之间的协方差和标准差的商来计算,其表达式为^[19]:

$$\text{sim}(i,j) = \frac{\text{cov}(i,j)}{\sigma_i \sigma_j} = \frac{\sum_{u \in U} (r_{ui} - r_i) \cdot \sum_{u \in U} (r_{uj} - r_j)}{\sqrt{\sum_{u \in U} (r_{ui} - r_i)^2} \sqrt{\sum_{u \in U} (r_{uj} - r_j)^2}} \quad (2)$$

(3)修正的余弦相似度。

余弦相似度在计算相似性时并没有考虑用户评分尺度问题,通常在推荐系统中,评分区间为1~5,对于甲用户用来说3分可能就表示喜欢,而对于乙用户可能5分才表示喜欢。修正的余弦相似度通过减去用户的平均评分来改善这一问题。其表达式为^[18]:

$$\text{sim}(i,j) = \frac{\sum_{u \in U} (r_{ui} - r_u) \cdot \sum_{u \in U} (r_{uj} - r_u)}{\sqrt{\sum_{u \in U} (r_{ui} - r_u)^2} \sqrt{\sum_{u \in U} (r_{uj} - r_u)^2}} \quad (3)$$

1.2 K近邻推荐

给定用户集合U和项目集合I,用户-项目的评分矩阵R可以表示为R=|U|×|I|。在该矩阵中,每一个行向量表示一个特定用户的评分集合,每一个列向量表示特定项目的被评分集合,每一个元素r_{ui}∈R表示用户u对项目i的评分。用户评分矩阵如表1所示。

表1 用户-项目评分矩阵

	Item ₁	Item ₂	...	Item _n
User ₁	R ₁₁	R ₁₂	...	R _{1n}
User ₂	R ₂₁	R ₂₂	...	R _{2n}
⋮	⋮	⋮	⋮	⋮
User _m	R _{m1}	R _{m2}	...	R _{mn}

得到用户-评分矩阵,按照相似度公式计算出项目之间的相似度之后,K近邻算法会记录下与项目i最相似的k个项目,记为S(i,k),当需要为目标用户u进行推荐时,通过公式(4)^[16]计算用户u对未评分项目i的预测评分。

$$\hat{r}_{ui} = r_i + \frac{\sum_{j \in S(i,k) \cap R(i)} \text{sim}(i,j) (r_{uj} - \bar{r}_j)}{\sum_{j \in S(i,k) \cap R(i)} |\text{sim}(i,j)|} \quad (4)$$

其中,S(i,k)是和项目i最相似的项目集合;N(u)是用户u评过分的物品集合;sim(i,j)是项目之间的相似度; \bar{r}_i 是项目i的平均分。

2 结合标签的推荐算法

在传统的相似性计算方法中,项目之间的相似性是通过用户对项目的评分来计算的,这种评分的方法很难反映用户的主观感受,而标签这种由用户产生的数据不仅能够很好地反映用户的兴趣,而且可以发挥集体智慧,获得对项目比较准确的关键词描述,而准确

的内容信息是提高个性化推荐性能的关键。因此文中将项目标签信息引入到项目相似性计算中,标签的引入使得相似性计算更加准确,而且降低了评分矩阵的稀疏性。针对传统的相似性计算中用户共同评分项目较少时可能会出现相似性计算不准确的问题,通过将项目标签相似性的引入,对传统相似性度量方法进行改进,使得相似性矩阵的稀疏性明显降低,最终的项目相似度综合了项目评分相似度和项目标签相似度,相似度的准确保证推荐质量的提高。

2.1 项目相似度计算

在实际应用中,用户通常会根据自己的兴趣、爱好和习惯为项目标注上不同的标签。标签信息更能反映用户对项目的喜好程度,在传统基于项目的协同过滤算法中,推荐结果着重于维系用户的历史兴趣,并没有对用户的动态行为-标签进行分析和利用,而标签相似度则正是反映用户喜好的动态特性,因此在度量项目间的相似性时,有必要考虑项目的标签因素对相似度度量的影响。假设系统中的用户数为n,项目数为m,项目标签数为k,对于两个项目i和j而言,项目i被用户评分的集合记为R_i={r_{i1},r_{i2},⋯,r_{in}} ,项目被用户标注的标签集合记为T_i={t_{i1},t_{i2},⋯,t_{ik}} ,那么项目i和项目j的相似性就包含了基于共同评分的相似性和共同标签的相似性。如果同时对项目i和项目j进行评分的用户越多说明两者的相似性越高,如果对项目i和项目j评价被用户共同标注的标签越多说明项目i和项目j越相似。

项目综合相似度计算主要分为三部分:

- (1)基于共同评分的项目相似度;

(2)基于相似标签的项目相似度;

(3)根据前两部分计算的相似度进行权重调整加以综合。

Tanimoto系数(又称为Jaccard系数)通常应用于两个变量为布尔变量的情况下,公式为:

$$T(X,Y) = \frac{XY}{\|X\|^2 + \|Y\|^2 - XY} \quad (5)$$

由公式(3)可知,修正的余弦相似性能够很好地考虑用户的评分尺度问题,而且Sarwar利用MovieLens数据集对三种相似度进行了比较,实验结果表明修正的余弦相似度可以取得最优的结果。所以文中采用修正的余弦相似度度量项目之间在共同评分上的相似性,并记为sim_p(i,j)。由于标签信息只有1和0两种取值,分别代表有和无,且希望对一部电影在多个标签上互有重叠情况下进行度量。在此文中采用Tanimoto系数^[20]计算项目之间在标签上的相似度,记为sim_t(i,j)。

为了解决传统算法仅仅依靠用户评分数据进行相

似性计算带来的误差,文中综合相似度考虑项目标签相似性和项目在用户评分行为上的相似性。为了度量这两种相似度之间的权重,文中通过权重因子来度量这两种相似性的权重,为此定义一个权重因子 α ,用于调整 $\text{sim}_p(i,j)$ 和 $\text{sim}_l(i,j)$ 对综合相似度影响的权重。所以项目的综合相似度计算公式为:

$$\text{sim}(i,j) = \alpha \text{sim}_p(i,j) + (1 - \alpha) \text{sim}_l(i,j) \quad (6)$$

2.2 结合用户评分和项目标签的协同过滤算法描述

文中将标签相似性考虑到相似性度量中,并结合传统的基于项目的协同过滤算法,提出了一种结合用户评分和项目标签的协同过滤推荐算法。

在文中提出的推荐算法中,系统的输入为训练集历史数据、测试集和标签数据。通过训练集历史数据得到一个 $R_{m \times n}$ 的项目评分矩阵,其中行代表用户,列代表项目;通过标签数据得到一个 $R_{n \times l}$ 的项目标签矩阵,其中行代表项目,列代表标签。分别通过上面介绍的相似性计算公式计算出项目之间的评分相似度和标签相似度,在得到项目之间的评分相似度和标签相似度以后,根据上面介绍的调和因子 α 和公式(6)生成最终的项目相似性矩阵 $R_{n \times n}$,在得到相似性矩阵之后生成最近邻居集合,最后根据公式(4)产生最终的推荐。其基本流程如图 1 所示。

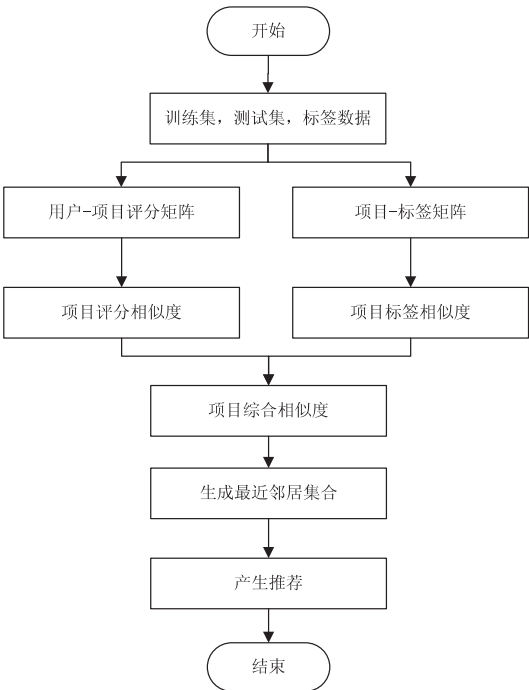


图 1 结合用户评分和项目标签的协同过滤算法基本流程

3 实验结果分析

文中采用 GroupLens 站点提供的 MovieLens 数据集来评测提出的结合用户评分和项目标签的协同过滤算法和传统的基于项目的协同过滤算法。MovieLens

数据集共有 3 个不同的版本,文中选取的是 100k 的数据集。该数据集包含 943 个用户对 1 682 部电影的 100 000 条评分,每一个用户至少对 20 部电影进行评分,评分范围为 1~5 分,在该数据集中包含项目标签 19 类,其结构如表 2 所示。

表 2 MovieLens 项目标签矩阵

项目	标签属性			
	T_1	T_2	...	T_n
1	T_{11}	T_{12}	...	T_{1n}
2	T_{21}	T_{22}	...	T_{2n}
\vdots	\vdots	\vdots	\vdots	\vdots
n	T_{n1}	T_{n2}	...	T_{nn}

3.1 推荐质量度量标准

在推荐系统中,主要通过离线实验、用户调查和在线实验这 3 种方法来评测推荐效果,这里做的是离线实验。在离线实验中预测准确度度量是最重要的评测指标,预测准确度一般通过平均绝对误差 (Mean Absolute Error, MAE) 计算,因此文中采用 MAE 作为度量标准。对于测试集中的一个用户 u 和物品 i ,令 r_{ui} 为用户 u 对物品 i 的实际评分,而 \hat{r}_{ui} 是推荐算法给出的预测评分,那么 MAE 定义^[16]如下:

$$\text{MAE} = \frac{\sum_{u,i \in T} |r_{ui} - \hat{r}_{ui}|}{|T|} \quad (7)$$

3.2 结果与分析

3.2.1 参数调整

在该实验中,由于要权衡项目评分相似度和标签相似度的权重,因此定义一个相似性调和因子 α , α 的取值范围为 $[0,1]$ 。其取值会直接影响到推荐质量。为了对改进算法做进一步的验证和分析,分别对 α 取不同的值对比推荐质量来观察 α 对推荐算法的影响程度。从图 2 中可以看到, α 能够影响到推荐质量,当 $0.3 < \alpha < 0.5$ 时推荐算法性能较优。

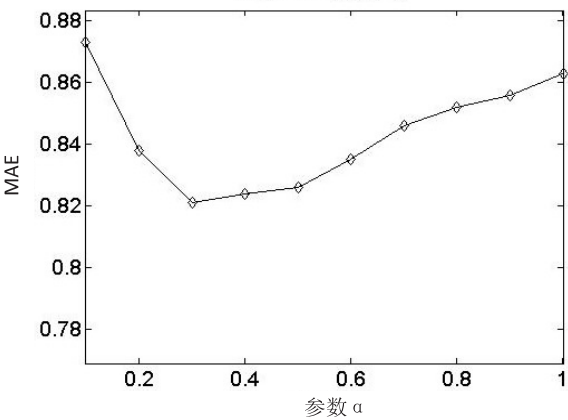


图 2 项目标签相似度影响因子 α 对 MAE 的影响实验结果表明,参数能够影响推荐算法的准确性。

由于该算法是结合用户评分和项目标签的协同过滤算法,所以参数 α 的取值会相对较小,如果 α 取值过大则会过多地偏向于基于标签的推荐算法,取值过小则体现不出标签数据对推荐质量的影响。在该实验中 α 的最优取值为0.3。

3.2.2 实验结果比较

最近邻的个数能够对推荐算法的推荐效果产生影响,因此在该实验中,将最近邻的个数从750到943进行取值,分别比较文中提出的算法与传统的基于项目和基于标签的协同过滤算法的MAE。实验参数 α 按照文中实验中的最优值来选取,实验结果如图3所示。

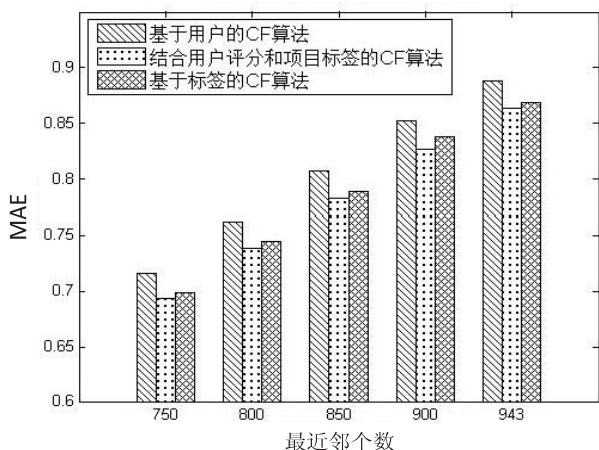


图3 不同推荐算法的MAE比较

由图3可以看出,文中提出的综合用户评分和项目标签的协同过滤算法具有较小的MAE,这是因为文中在计算项目相似性时,考虑了用户兴趣信息—标签对推荐质量的影响。由此可见,综合考虑标签信息能够有效地综合用户对项目的喜好,标签信息反映了用户对资源的主观感受,文中提出的结合用户评分和项目标签的协同过滤推荐算法能更好地满足用户的兴趣爱好,同时通过分析标签信息权重 α 对推荐算法的推荐准确性有较大的影响。从图中可以看出,综合用户评分和项目标签的协同过滤算法优于传统的基于项目的协同过滤算法和基于标签的协同过滤算法。实验数据表明,该算法在MAE上比传统的基于项目的协同过滤算法降低了3%,表明文中提出的算法确实对推荐准确性有明显的提高,取得了较好的效果。

4 结束语

文中针对传统协同过滤推荐算法仅仅依靠用户评分数据计算相似性的不足提出了改进,通过结合用户评分数据和项目标签共同计算项目之间的相似性,并将两种数据通过加权进行有机结合,设计并实现了一种结合用户评分和项目标签的协同过滤推荐算法。在该算法中,通过对评分数据和标签数据的结合使得评

分数据的稀疏性明显降低,使得计算最近邻居的准确性得到保障,进而提高了推荐准确性。从实验结果可以看出,将用户评分数据和项目标签数据进行有机结合能有效地反映用户的兴趣特点,从而将推荐准确度进一步提高。

随着推荐系统应用规模的不断扩大,推荐算法面临许多挑战,其中推荐算法的实时性问题、推荐的精确度问题、冷启动问题和推荐算法的可扩展性问题等已经得到广泛关注。通常系统中都会包含一些除了用户历史数据的额外信息,例如用户的人口统计信息、项目的分类信息、项目标签信息、项目的评分时间信息等。如何利用这些数据建立相关的模型将是一个非常有意义的研究方向。

参考文献:

- [1] Schafer J B, Konstan J A, Riedl J. E-commerce recommendation applications [M]//Applications of data mining to electronic commerce. US: Springer US, 2001: 115–153.
- [2] Konstan J A. Introduction to recommender systems: algorithms and evaluation [J]. ACM Transactions on Information Systems, 2004, 22(1): 1–4.
- [3] Bobadilla J, Ortega F, Hernando A, et al. Improving collaborative filtering recommender system results and performance using genetic algorithms [J]. Knowledge-based Systems, 2011, 24(8): 1310–1316.
- [4] Bell R M, Koren Y. Improved neighborhood-based collaborative filtering [C]//Proc of KDD cup and workshop at the 13th ACM SIGKDD international conference on knowledge discovery and data mining. [s. l.]: ACM, 2007.
- [5] Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions [J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(6): 734–749.
- [6] Pirasteh P, Jung J J, Hwang D. Item-based collaborative filtering with attribute correlation: a case study on movie recommendation [C]//Proc of Intelligent information and database systems. [s. l.]: [s. n.], 2014: 245–252.
- [7] Pitsilis G, Knapkag S J. Socialtrust as a solution to address sparsity – inherent problems of recommender systems [C]//Proc of ACM RecSys 2009. New York: ACM, 2009: 332–344.
- [8] Kumar R, Verma B K, Rastogi S S. Social popularity based SVD++ recommender system [J]. International Journal of Computer Applications, 2014, 87(14): 33–37.
- [9] Wei Suyun, Ye Ning, Zhang Shuo, et al. Collaborative filtering recommendation algorithm based on item clustering and global similarity [C]//Proc of 5th international conference on business intelligence and financial engineering. [s. l.]: IEEE, 2012: 69–72.

快速去除完阴影,对于后续的检测起到很好铺垫效果,在检测前期尽量减少误警率。

5 结束语

将 Vibe 算法与帧差法融合在一起,利用帧差法建立和更新的背景来消除由于太阳光照射而产生的阴影,提高了前景检测的准确率。此外,利用帧差法得到的前景与 Vibe 检测的前景进行与操作,去除“鬼影”,进一步降低了检测系统的虚警率。从实验中得到的数据计算,其中改进的 Vibe 平均耗时为 0.123 s,可以达到 8 帧每秒,能满足检测实时性要求。

参考文献:

- [1] Barnich O, van Droogenbroeck M. ViBe: a powerful random technique to estimate the background in video sequences [C]//Proc of IEEE international conference on acoustics, speech and signal processing. [s. l.]: IEEE, 2009: 945-948.
 - [2] 张伟. 基于视觉的运动车辆检测与跟踪[D]. 上海: 上海交通大学, 2007.
 - [3] Collins R T, Lipton A J, Kanade T. Introduction to the special section on video surveillance[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 745-746.
 - [4] Sen-Ching S C, Kamath C. Robust techniques for background subtraction in urban traffic video[J]. SPIE, 2004, 5308: 881-892.
 - [5] Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction [C]//Proceedings of the 17th international conference on pattern recognition. [s. l.]: IEEE, 2004: 28-31.
 - [6] Zivkovic Z, van der Heijden F. Efficient adaptive density estimation per image pixel for the task of background subtraction [J]. Pattern Recognition Letters, 2006, 27(7): 773-780.
 - [7] Power P W, Schoonees J A. Understanding background mixture models for foreground segmentation [C]//Proceedings of image and vision computing, New Zealand: [s. n.], 2002.
 - [8] Kim K, Chalidabhongse T H, Harwood D, et al. Background modeling and subtraction by codebook construction [C]//Proc of international conference on image processing. [s. l.]: IEEE, 2004: 3061-3064.
 - [9] van Droogenbroeck M, Paquot O. Background subtraction: experiments and improvements for vibe [C]//Proc of IEEE computer society conference on computer vision and pattern recognition workshop. [s. l.]: IEEE, 2012: 32-37.
 - [10] Prati A, Mikic I, Trivedi M M, et al. Detecting moving shadows: algorithms and evaluation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(7): 918-923.
 - [11] 周维. 视频监控中运动目标发现与跟踪算法研究[D]. 合肥: 中国科学技术大学, 2012.
 - [12] 王晓冬, 霍宏, 方涛. 基于快速归一化互相关函数的运动车辆阴影检测算法[J]. 计算机应用, 2006, 26(9): 2065-2067.
 - [13] 焦波, 李国辉, 汪彦明, 等. 一种基于形态学的运动车辆阴影消除方法[J]. 自动化学报, 2008, 34(7): 838-840.
 - [14] 王华伟, 李翠华, 施华, 等. 基于 HSV 空间和一阶梯度的阴影剪除算法[J]. 计算机工程与应用, 2005, 41(8): 43-44.
 - [15] 朱秀昌, 刘峰, 胡栋, 等. 数字图像处理与图像通信 [M]. 北京: 北京邮电大学出版社, 2002.
- +++++
- (上接第 75 页)
- [10] Wang Jingjin, Lin Kunhui, Li Jia. A collaborative filtering recommendation algorithm based on user clustering and slope one scheme [C]//Proc of 8th international conference on computer science & education. [s. l.]: IEEE, 2013: 1473-1476.
 - [11] 黄创光, 印鉴, 汪静, 等. 不确定近邻的协同过滤推荐算法[J]. 计算机学报, 2010, 33(8): 1369-1377.
 - [12] Symeonidis P, Nanopoulos A, Manolopoulos Y. A unified framework for providing recommendations in social tagging systems based on ternary semantic analysis [J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(2): 179-192.
 - [13] Hotho A, Jascher R, Schmitz C, et al. Information retrieval in folksonomies: search and ranking [C]//Proc of 3rd European semantic web conference on the semantic web: research and applications. Budva: Springer, 2006: 411-426.
 - [14] Harvey M, Baillie M, Ruthven I, et al. Tripartite hidden topic models for personalized tag suggestion [M]//Advances in information retrieval. Berlin: Springer, 2010: 432-443.
 - [15] 张斌, 张引, 高克宁, 等. 融合关系与内容分析的社会标签推荐[J]. 软件学报, 2012, 23(3): 476-488.
 - [16] 项亮. 推荐系统实践 [M]. 北京: 人民邮电出版社, 2012: 44-64.
 - [17] Sarwar B, Karypis G, Konstan J, et al. Item-based collaborative filtering recommendation algorithms [C]//Proceedings of the 10th international conference on World Wide Web. [s. l.]: ACM, 2001: 285-295.
 - [18] Ye J. Cosine similarity measures for intuitionistic fuzzy sets and their applications [J]. Mathematical and Computer Modeling, 2011, 53(1): 91-97.
 - [19] 罗辛, 欧阳元新, 熊璋, 等. 通过相似度支持度优化基于 K 近邻的协同过滤算法 [J]. 计算机学报, 2010, 33(8): 1437-1445.
 - [20] Toby S. 集体智慧编程 [M]. 莫映, 王开福, 译. 北京: 电子工业出版社, 2009: 7-28.

结合用户评分和项目标签的协同过滤算法

作者：[程高伟](#)，[丁亦喆](#)，[吴振强](#)，[CHENG Gao-wei](#)，[DING Yi-zhe](#)，[WU Zhen-qiang](#)
作者单位：[陕西师范大学 计算机科学学院, 陕西 西安, 710062](#)
刊名：[计算机技术与发展](#)
英文刊名：[Computer Technology and Development](#)
年，卷(期)：2015(3)

引用本文格式：[程高伟](#).[丁亦喆](#).[吴振强](#).[CHENG Gao-wei](#).[DING Yi-zhe](#).[WU Zhen-qiang](#) [结合用户评分和项目标签的协同过滤算法](#)[期刊论文]-[计算机技术与发展](#) 2015(3)