

基于优化视觉词袋模型的图像分类方法

张永, 杨浩*

(兰州理工大学 计算机与通信学院, 兰州 730050)

(* 通信作者电子邮箱 kdyh123@163.com)

摘要: 针对视觉词袋(BOV)模型中过大的视觉词典会导致图像分类时间代价过大的问题,提出一种加权最大相关最小相似(W-MR-MS)视觉词典优化准则。首先,提取图像的尺度不变特征转换(SIFT)特征,并用K-Means算法对特征聚类生成原始视觉词典;然后,分别计算视觉单词与图像类别间的相关性,以及各视觉单词间的语义相似性,引入一个加权系数权衡两者对图像分类的重要程度;最后,基于权衡结果,删除视觉词典中与图像类别相关性弱、与视觉单词间语义相似性大的视觉单词,从而达到优化视觉词典的目的。实验结果表明,在视觉词典规模相同的情况下,所提方法的图像分类精度比传统基于K-Means算法的图像分类精度提高了5.30%;当图像分类精度相同的情况下,所提方法的时间代价比传统K-Means算法下的时间代价降低了32.18%,因此,所提方法具有较高的分类效率,适用于图像分类。

关键词: 图像分类;视觉词袋模型;特征提取;视觉词典

中图分类号: TP181 **文献标志码:** A

Image classification method based on optimized bag-of-visual words model

ZHANG Yong, YANG Hao*

(School of Computer and Communication, Lanzhou University of Technology, Lanzhou Gansu 730050, China)

Abstract: Concerning the problem that too large visual dictionary may increase the time cost of image classification in the Bag-Of-Visual words (BOV) model, a Weighted-Maximal Relevance-Minimal Semantic similarity (W-MR-MS) criterion was proposed to optimize visual dictionary. Firstly, the Scale Invariant Feature Transform (SIFT) features of images were extracted, and the K-Means algorithm was used to generate an original visual dictionary. Secondly, the correlation between visual words and image categories and semantic similarity among visual words were calculated, and a weighted parameter was introduced to measure the importance of the correlation and the semantic similarity in image classification. Finally, based on the weighing result, the visual word which correlation with image categories was weak and semantic similarity among visual words was high was removed, which achieved the purpose of optimizing the visual dictionary. The experimental results show that the classification precision of the proposed method is 5.30% higher than that of the traditional K-Means algorithm under the same visual dictionary scale; the time cost of the proposed method is reduced by 32.18% compared with the traditional K-Means algorithm under the same classification precision. Therefore, the proposed method has high classification efficiency and it is suitable for image classification.

Key words: image classification; Bag-Of-Visual words (BOV) model; feature extraction; visual dictionary

0 引言

近几年来,人工智能领域呈现出飞速发展的势头,图像分类技术^[1-4]也得到了人们的普遍关注。早期的词袋模型被用在文本分类中,通过统计文本词典中与顺序无关的单词的频率就可以对文本进行精确的分类。然而文本分类中的文本词典是可以被确定的,而基于词袋模型下的图像分类^[5-7]中的视觉词典则需要通过对图像进行特征的提取,并对特征进行相应处理才能获得,不同的特征提取方法和特征处理方式会产生差异较大的视觉词典,因此,如何得到适当的视觉词典是现今研究的热点。

传统视觉词袋(Bag-Of-Visual words, BOV)模型下的图像分类性能低下,因此大量对BOV模型的优化算法^[8-9]被引入

到图像分类中。Kim等^[10]利用信息熵的方法去去除掉那些图像类别与视觉单词间信息熵较小的视觉单词,减小了视觉词典的规模;Epshtein等^[11]用平均互信息的方法来计算视觉单词与图像类别间的相关性,并去除掉那些与图像类别相关性较弱的视觉单词,从而降低视觉词典的规模;Lu等^[12]应用谱聚类的思想对词袋模型中的视觉词典进行降维,提高了图像分类的效率。然而以上研究对视觉词典的优化方法只考虑了视觉词典与图像类别之间的相关性,并没有考虑单视觉词之间的冗余关系。在基于BOV模型的图像分类中,视觉词典中的视觉单词具有大小和空间分布信息,规模较大的视觉词典中含有那些大小与空间分布信息相似的视觉单词,称这种关系为视觉单词间的语义相似性,而大量这种语义相似性视觉单词的出现,导致了视觉单词间的冗余性,因此去除掉那些与

收稿日期: 2016-12-13; 修回日期: 2017-03-11。

作者简介: 张永(1963—),男,甘肃兰州人,教授,主要研究方向: 智能信息处理、数据挖掘; 杨浩(1991—),男,甘肃陇南人,硕士研究生,主要研究方向: 图像分类、机器学习。

其他视觉单词相似性较大的视觉单词可以有效降低视觉词典的规模,提高图像分类的性能。

在优化视觉词典的过程中,不仅要考虑视觉词典与图像类别之间的相关性,还要考虑视觉词典中视觉单词之间的冗余关系。因此,本文提出了一种基于加权最大相关最小相似 (Weighted-Maximal Relevance-Minimal Semantic similarity, W-MR-MS) 准则的图像分类方法。首先分别计算视觉单词与图像类别间的相关性、视觉单词与视觉单词间的语义相似性;然后引入一个加权系数对两者进行加权计算,保留那些加权结果较大的视觉单词组成本文图像分类中最终的视觉词典。该方法不仅去除了那些与图像类别无关的噪声单词与冗余性较大的视觉单词,而且减小了视觉词典的规模。最后在两类常见数据集上的实验结果也验证了本文方法的有效性。

1 词袋模型下对图像的表达

BOV 模型下视觉词典的规模对图像的分类性能具有较大影响,而基于 BOV 模型对图像的表达就是将图像特征一一量化到视觉词典上,即用视觉单词频率直方图来表示一幅图像。视觉单词频率直方图在这里被称为视觉词汇直方图。图 1 为 BOV 下的图像表示示意图,由图 1 可知,视觉词汇直方图的好坏可以决定图像分类的精度,而且视觉词汇直方图的维度大小直接影响了图像分类中的时间复杂度,因此,适当的视觉词典规模能够提升图像分类的性能。

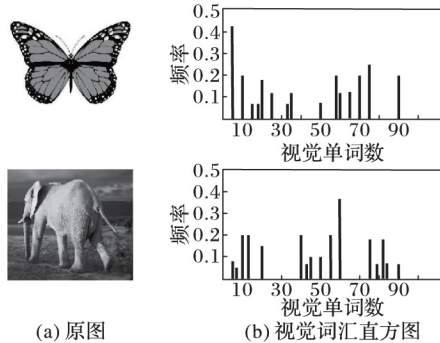


图 1 BOV 模型下对图像的表达
Fig. 1 Image representation of BOV model

2 W-MR-MS 准则的定义

2.1 视觉单词与图像类别间的相关性

本文采用信息论中平均互信息方法^[13]来计算视觉单词与图像类别间的相似性。首先定义初始视觉词典为 $D = [d_1, d_2, \dots, d_N]$ D 为 $N \times K$ 维矩阵 $K = 128$ 为尺度不变特征转换 (Scale Invariant Feature Transform, SIFT) 描述子^[14]的维数, N 为视觉词典的规模。如式 (1) 所示: $AvI(d_m, c)$ 表示视觉单词 d_m 与图像类别之间相关性的度量,其值越大表明该视觉单词对分类越重要,所以该单词在分类过程中是该被保留的;反之,那些与图像类别相关性较小的视觉单词应该被去除掉。其中: $c \in \{1, 2, \dots, C\}$ 为图像的类别, C 为图像类别总数; $d_m = 0$ 表示在分类过程中该视觉单词没有出现在对该类别分类的视觉词典中, $d_m = 1$ 表示该视觉单词出现在对该类别分类的视觉词典中。

$$AvI(d_m, c) = \frac{1}{|C|} \sum_{c \in \{1, 2, \dots, C\}} I(d_m, c) \quad (1)$$

$$I(d_m, c) = \sum_{d_m \in \{0, 1\}} p(d_m, c) \log \frac{p(d_m, c)}{p(d_m) p(c)} \quad (2)$$

2.2 视觉单词之间的语义相似性

假设 SIFT 特征点 $p_i = \{r_i, u_i, s_i, \theta_i\}$, 其中: r_i 为 128 维特征描述子, u_i 为特征空间位置坐标, s_i 为特征的尺度大小, θ_i 为特征的主方向。为了计算视觉单词之间的语义相似性,首先定义该特征点的空间语义区域为 SCR_{p_i} ,而且本文认为该区域包含了 p_i 的空间语义信息,其半径为 $r_c \times s_i$,其中 r_c 为语义尺度系数,用来控制空间语义区域的大小,可设置为大于 1 的一个常数。然后定义空间语义区域中的特征点 p_j 对 p_i 的影响权重 d_{ij} 如式 (3) 所示:

$$d_{ij} = \exp(-h_{ij}^2/2) \quad (3)$$

其中: h_{ij} 的计算如式 (4) 所示, $\|\cdot\|_2$ 为向量的 2-范数。

$$h_{ij} = \|u_i - u_j\|_2 / (r_c \times s_i) \quad (4)$$

在 BOV 模型中,每个特征点都对应着一个视觉单词,定义 p_j 对应的视觉单词为 d_k 。对于空间语义区域中的所有对应于视觉单词 d_k 的特征点定义一个集合,该集合为 $H_k = \{p_j | p_j \rightarrow d_k, 1 \leq j \leq K\}$,其中 $p_j \rightarrow d_k$ 指 SIFT 点 p_j 对应的视觉单词为 d_k , K 为对应于视觉单词 d_k 的特征点数量。视觉单词 d_k 在空间语义区域中对特征点 p_i 的影响权重如式 (5) 所示:

$$cd_k(p_i) = \sum_{p_j \in SCR_{p_i}, p_j \in H_k} d_{ij} \quad (5)$$

再计算视觉词典中所有视觉单词对特征点的影响权重,得到 p_i 的空间语义信息对应的直方图如式 (6) 所示:

$$SC(p_i) = [cd_1(p_i), cd_2(p_i), \dots, cd_n(p_i), \dots, cd_N(p_i)] \quad (6)$$

接下来,对于 SIFT 特征点 p_i 所对应的视觉单词 d_m ,定义所有对应为视觉单词 d_m 的特征点集合为 $R_m = \{p_i | p_i \rightarrow d_m, 1 \leq i \leq L\}$,将 R_m 中每个 SIFT 点的空间语义信息看作是视觉单词 d_m 的空间语义信息的一种表现,则可定义视觉单词 d_m 的空间语义信息为所有 R_m 中 SIFT 点的空间语义信息的均值,如式 (7) 所示:

$$SC(d_m) = \frac{1}{|R_m|} \sum_{p_i \in R_m} SC(p_i) \quad (7)$$

其中: $|R_m|$ 为集合 R_m 的规模; $SC(d_m)$ 为综合利用了所有量化到 d_m 上的 SIFT 特征点的空间语义信息,能够较好地表征视觉单词 d_m 的空间语义特性。

视觉单词的空间语义信息是视觉单词空间分布信息的重要体现。对于两个视觉单词 d_m 与 d_n ,通过式 (8) 来计算其语义相似性:

$$sim(d_m, d_n) = \cos(SC(d_m), SC(d_n)) = \frac{SC(d_m) \cdot SC(d_n)}{\|SC(d_m)\|_2 \cdot \|SC(d_n)\|_2} \quad (8)$$

通过以上思路,可以计算出视觉单词在视觉词典中的语义相似性大小,式 (9) 为视觉单词 d_m 在视觉词典中的语义相似性的计算公式:

$$I(d_m) = \sum_{n=1, n \neq m}^N sim(d_m, d_n) \quad (9)$$

其中: N 为视觉词典规模的大小, $I(d_m)$ 为视觉单词 d_m 的语义相似性大小。

通过分别对视觉词典中所有视觉单词的语义相似性进行

计算, 可以去除掉那些语义相似性较大的视觉单词。

2.3 W-MR-MS 准则

结合上述内容, 本节给出 W-MR-MS 准则的具体内容如下: 结合 2.1 节与 2.2 节中选择视觉单词的方法, 首先分别计算视觉单词与图像类别之间的相关性和视觉单词间的语义相似性; 然后引入一个加权参数 α 对视觉单词与图像类别之间的相关性和视觉单词间的语义相似性进行权衡, 如式 (10) 所示; 最后去除掉使加权结果 $I(d_m)$ 值最小的那些视觉单词, 其中 d_m 为视觉单词, $1 \leq m \leq N$ 。

$$I(d_m) = (1 - \alpha) \times AvI(d_m, c) - \alpha \times I(d_m) \quad (10)$$

其中 $0 \leq \alpha \leq 1$, α 的值越小, 表示视觉单词与图像类别间的相关性对优化视觉词典起到了主导作用; 反之, 视觉单词间的语义相似性对优化视觉词典起到主导作用。

3 基于 W-MR-MS 准则的 BOV 模型

图 2 为本文图像分类的系统框图, 它首先提取图像的局部 SIFT 特征点, 并对该局部特征采用 K-Means 聚类算法生成视觉词典; 然后利用 W-MR-MS 准则对视觉词典进行优化, 具体优化步骤如算法 1 所示; 视觉词典优化后, 基于该优化后的视觉词典对训练图像进行视觉词汇直方图的构建; 最后采用词袋模型表示对每个图像类别训练出分类模型, 对每幅未分类的测试图像, 采用该模型可以得到其分类结果。

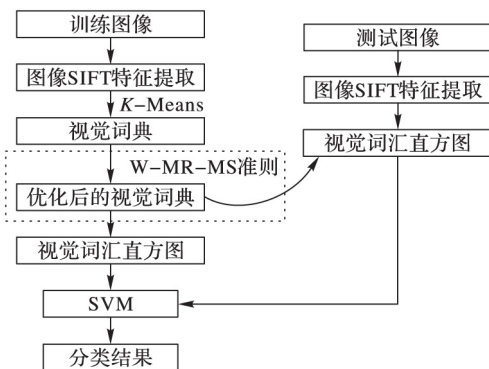


图 2 基于 W-MR-MS 准则图像分类的系统框图

Fig. 2 System diagram of image classification based on W-MR-MS criterion

算法 1 视觉词典优化算法。

- 1) 通过 K-Means 算法对局部特征聚类生成视觉词典 D , 其大小为 K 本文中 $K = 1200$ 。
- 2) 用视觉词典 D 对训练图像进行表示与分类, 得到分类精度为 P 。
- 3) 用式 (10) 选出 T 个使 $I(d_m)$ 值最小的视觉单词, 并从视觉词典中去掉这 T 个视觉单词, 得到一个大小为 $K - T$ 的视觉词典 D , 如果 $K - T$ 大于阈值 H 继续步骤 2); 否则, 停止循环。本文中 $T = 10$, $H = 400$ 。

模型训练过程中, 首先通过特征提取与特征聚类可以得到一个冗余性与规模较大的视觉词典; 然后利用 W-MR-MS 准则优化视觉词典, 可以达到去冗余与降低词典规模的效果; 最后基于优化视觉词典对每幅训练图像建立视觉词汇直方图, 并将其作为分类器的输入。

模型测试过程中, 对提取到的局部特征直接采用 W-MR-MS 准则优化后的视觉词典建立视觉词汇直方图, 将其作为模型的输入, 并得到分类结果。

4 实验结果与分析

4.1 实验设置

本文在 Caltech-101 和 COREL 图像数据集上进行实验。在 Caltech-101 数据集中, 选取其中的 12 类图像作实验, 分别为 Airplanes、Face、Watch、Motorbikes、Car、Backpack、Ketch、Bonsai、Butterfly、Crab、Revolver 和 Sunflower, 每个类别图像数目从 47 到 800 不等; COREL 数据集共有 10 类图像, 分别为 African、Beach、Buildings、Buses、Dinosaurs、Elephants、Flowers、Food、Horses 和 Mountains, 每一个类别含有 100 幅图像, 共有 1000 幅图像。实验中选取图像库中的一半图像作为训练图像, 另一半作为测试图像进行实验。为了便于实验, 将数据集图像大小调整到 300×300 像素; 然后选择一对多方式下的多类支持向量机 (Support Vector Machine, SVM) [15] 分类器对数据集进行训练, 提取 SIFT 特征的图像块大小为 16×16 像素, 步长为 8 像素; 接头使用 K-Means 算法对图像聚类生成视觉词典 $K = 1200$; 最后在每个数据集上独立进行 10 次随机实验, 并将平均分类准确率与分类时间代价作为最后判断标准。

4.2 实验结果

4.2.1 参数 α 对图像分类性能的影响

由于参数 α 对优化视觉词典具有较大影响, 所以本节将讨论参数 α 对图像分类的影响。实验中 α 将在集合 $M = \{0, 0.05, 0.1, \dots, 1\}$ 上取值, 图 3 即为不同参数 α 下图像的平均分类精度。其中: 在 10 类 Caltech-101 数据集上, 优化视觉词典规模取为 910; 在 COREL 数据集上, 优化视觉词典规模取为 850。由图 3 可知, 当 $\alpha = 0.6$ 时图像的分类精度最高, 所以在 COREL 与 Caltech-101 数据集上视觉单词间的语义相似性在优化视觉词典中占了主导地位。这是由于以上两种数据集中的类别数较少, 所以影响了图像类别与视觉单词间的相关性对优化视觉词典的作用, 而此时原始视觉词典中具有较多语义相似性较大的视觉单词, 因此, W-MR-MS 准则在优化视觉单词时, 视觉单词间的语义相似性对优化视觉词典具有更大的作用。

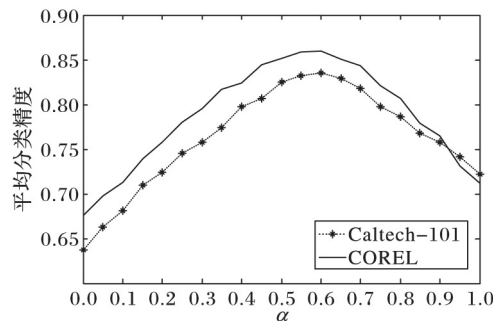


图 3 参数 α 对图像分类性能的影响

Fig. 3 Influence of α on image classification

4.2.2 视觉词典规模对图像分类性能的影响

本节将分析优化视觉词典规模对图像分类性能的影响。首先设置原始视觉词典的规模为 1200, $\alpha = 0.6$; 然后基于 W-MR-MS 准则对该规模下的视觉词典进行优化。图 4 为在 Caltech-101 与 COREL 数据集上视觉词典的规模与图像平均分类精度之间的关系示意图。由图 4 可以看出, 在 Caltech-

101 数据集上,当视觉词典的规模被优化到 910 之后,图像平均分类精度会出现明显下降;而在 COREL 数据集上,当视觉词典的规模优化到 850 之后,图像平均分类精度会出现明显下降。这是因为当视觉词典规模被缩减到一定程度时,那些对分类有用的视觉单词也会被 W-MR-MS 准则去除掉,这样会导致训练时对图像表示不足,进而图像分类精度出现明显下降。所以当原始视觉词典规模取为 1 200 时,在 Caltech-101 与 COREL 数据集上可优化视觉词典规模分别为 910 与 850,而两者大小不同,主要是由于本文所取 Caltech-101 数据集上的图像类别数比 COREL 数据集上要多,并且 Caltech-101 数据集图像结构比较复杂。

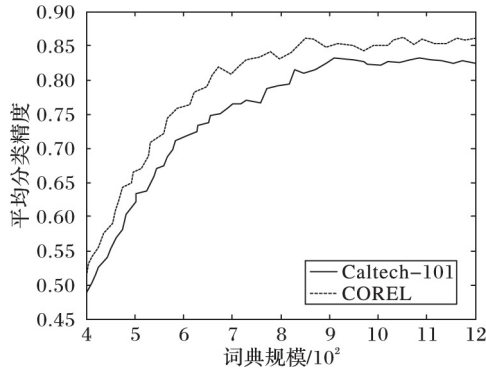


图 4 优化视觉词典的规模对图像分类性能的影响

Fig. 4 Influence of scale of visual dictionary on image classification

4.2.3 本文方法与 K-Means 算法的比较

本节将比较本文方法与 K-Means 算法下的图像分类性能,结果如表 1 所示。其中 K-Means 算法下视觉词典的大小直接由 K 值决定;而本文方法是在 K-Means 算法基础上, K = 1 200 时,用 W-MR-MS 准则对视觉词典进行了优化降维,最后的词典规模为优化后的大小。在 Caltech-101 与 COREL 数据集上,本文方法的视觉词典规模分别为 910 与 850, α 的值在两类数据集上均为 0.6。

表 1 两种算法在不同数据集上分类性能比较

Tab. 1 Image classification performance comparison of two algorithms on different datasets

数据集	方法	词典规模 (K)	平均分类精度	总分类时间 / min
Caltech-101	K-Means	1 200	0.836	43.2
	本文方法	910	0.792	24.6
COREL	K-Means	1 200	0.863	35.3
	本文方法	850	0.825	20.8

由表 1 可以看出,在 Caltech-101 数据集上,本文方法与 K-Means 算法在 K = 1 200 时的平均分类精度基本相同,但是本文方法的分类时间代价降低了 32.18%;而与 K-Means 算法在 K = 910 时的平均分类精度相比,本文方法的平均分类精度提高了 5.30%,但是时间代价本文略高,这是由于本文优化时需要优化时间代价。在 COREL 数据集上的分类性能比较结果与 Caltech-101 数据集的情况相似,这里不作过多的分析。

综合以上分析可知,本文方法提高了图像分类的性能。

图 5 为本文方法在 10 类 COREL 数据集上的分类混淆矩阵,由图 5 可知本文方法在某些图像类别上达到了较高的分类精度,所以本文方法具有较高的有效性。

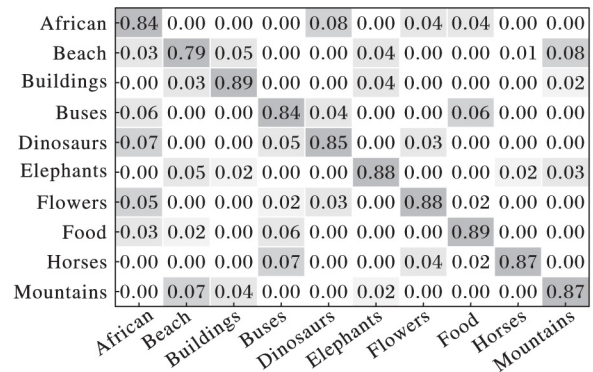


图 5 本文方法在 COREL 数据集上的混淆矩阵

Fig. 5 Confusion matrix of the proposed method on COREL dataset

5 结语

为了降低传统 BOV 模型下图像分类的时间复杂度,本文提出了 W-MR-MS 准则来优化 BOV 模型中的视觉词典。通过去除掉那些与图像类别无关、具有冗余的视觉单词,从而在不影响图像分类精度的前提下,降低了视觉词典的规模,提高了分类效率。在 Caltech-101 和 COREL 图像数据集上的实验表明,本文方法比传统 K-Means 算法具有更好的分类性能。在以后的研究中可以在此优化视觉词典的基础上,对 BOV 模型作进一步改进,比如加入图像局部特征的空间分布信息、对图像预处理等,从而实现更有效的图像分类。

参考文献 (References)

- [1] SIVIC J, ZISSERMAN A. Video Google: a text retrieval approach to object matching in videos [C]// ICCV 2003: Proceedings of the 2003 Ninth IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2003: 1470 - 1477.
- [2] 王翔琛,汪西莉,马君亮. 基于均值漂移的半监督支持向量机图像分类[J]. 计算机应用, 2014, 34(8): 2399 - 2403. (WANG X C, WANG X L, MA J L. Semi-supervised support vector machine for image classification based on mean shift [J]. Journal of Computer Applications, 2014, 34(8): 2399 - 2403.)
- [3] 邵忻. 基于跨领域主动学习的图像分类方法[J]. 计算机应用, 2014, 34(4): 1169 - 1171. (SHAO X. Cross-domain active learning algorithm for image classification [J]. Journal of Computer Applications, 2014, 34(4): 1169 - 1171.)
- [4] TIMOFTE R, GOOL L V. Adaptive and weighted collaborative representations for image classification [J]. Pattern Recognition Letters, 2014, 43(1): 127 - 135.
- [5] ALQASRAWI Y, NEAGU D, COWLING P I. Fusing integrated visual vocabularies-based bag of visual words and weighted colour moments on spatial pyramid layout for natural scene image classification [J]. Signal Image & Video Processing, 2013, 7(4): 759 - 775.
- [6] LU Y, XIE F, LIU T, et al. No reference quality assessment for multiply-distorted images based on an improved bag-of-words model [J]. IEEE Signal Processing Letters, 2015, 22(10): 1811 - 1815.
- [7] QU Y, WU S, LIU H, et al. Evaluation of local features and classifiers in BOW model for image classification [J]. Multimedia Tools and Applications, 2014, 70(2): 605 - 624.

(下转第 2252 页)

- tificial Intelligence, 2006, 19(1): 58–56.)
- [2] SHIEH M-Y, CHIOU J-S, HU Y-S, et al. Applications of PCA and SVM-PSO based real-time face recognition system [J]. *Mathematical Problems in Engineering*, 2014, 2014: Article ID 530251.
- [3] 温浩, 卢朝阳, 高全学. 融合小波变换和张量 PCA 的人脸识别算法[J]. *西安电子科技大学学报(自然科学版)*, 2009, 36(4): 602–607. (WEN H, LU Z Y, GAO Q X. Face recognition algorithm based on wavelet preprocessing and tensor PCA [J]. *Journal of Xidian University*, 2009, 36(4): 602–607.)
- [4] 黄璞, 陈才扣. 基于局部人脸图像的 ICA 人脸识别方法[J]. *计算机工程与设计*, 2010, 31(11): 2550–2553. (HUANG P, CHEN C K. ICA face recognition method based on partial face image [J]. *Computer Engineering and Design*, 2010, 31(11): 2550–2553.)
- [5] 尹洪涛, 付平, 沙学军. 基于 DCT 和线性判别分析的人脸识别[J]. *电子学报*, 2009, 37(10): 2211–2214. (YIN H T, FU P, SHA X J. Face Recognition Based on DCT and LDA [J]. *Acta Electronica Sinica*, 2009, 37(10): 2211–2214.)
- [6] FANG L, LI S. Face recognition by exploiting local Gabor features with multitask adaptive sparse representation [J]. *IEEE Transactions on Instrumentation and Measurement*, 2015, 64(10): 2605–2615.
- [7] LEI L, KIM D-H, PARK W-J. Face recognition using LBP eigenfaces [J]. *IEICE Transactions on Information and Systems*, 2014, E97.D(7): 1930–1932.
- [8] 武伟, 江艳霞, 张喆. 基于改进的 LDP 人脸识别算法[J]. *光电工程*, 2014, 41(12): 72–77. (WU W, JIANG Y X, ZHANG Z. The face recognition algorithm based on improved local directional pattern [J]. *Opto-Electronic Engineering*, 2014, 41(12): 74–77.)
- [9] JABID T, KABIR M H, CHAE O. Robust facial expression recognition based on local directional pattern [J]. *ETRI Journal*, 2010, 32(5): 784–794.
- [10] JABID T, KABIR M H, CHAE O. Local Directional Pattern (LDP) for face recognition [J]. *International Journal of Innovative Computing Information and Control*, 2012, 8(4): 2423–2437.
- [11] 李杰, 徐汀荣, 李海彦. 基于改进 LDP 的人脸识别[J]. *计算机应用与软件*, 2014, 31(5): 148–150, 230. (LI J, XU T R, LI H Y. Face recognition based on improved LDP [J]. *Computer Applications and Software*, 2014, 31(5): 148–150, 230.)
- [12] 罗元, 张天, 张毅. 一种改进的 LDP 面部表情特征提取方法[J]. *半导体光电*, 2016, 37(1): 122–125. (LUO Y, ZHANG T, ZHANG Y. An improved LDP for facial expression feature extraction [J]. *Semiconductor Optoelectronics*, 2016, 37(1): 122–125.)
- [13] LUO Y, ZHANG T, ZHANG Y. A novel fusion method of PCA and LDP for facial expression feature extraction [J]. *Optik — International Journal for Light and Electron Optics*, 2016, 127(2): 718–721.
- [14] 刘海军, 常东超, 张凌宇. 改进的局部方向模式纹理表示方法[J]. *中国图象图形学报*, 2014, 19(4): 520–526. (LIU H J, CHANG D C, ZHANG L Y. Improved local directional pattern texture descriptor [J]. *Journal of Image and Graphics*, 2014, 19(4): 520–526.)
- [15] 高现文, 付伟. 局部定向模式在人脸识别中的应用[J]. *计算机工程与应用*, 2011, 48(12): 98–201. (GAO X W, FU W. Local directional pattern in face recognition application [J]. *Computer Engineering and Applications*, 2011, 48(12): 98–201.)
- [16] BAY H, TNYTETAARS T, VAN GOOL L. SURF: Speeded Up Robust Features [C]// *ECCV06: Proceedings of the 9th European Conference on Computer Vision*, LNCS 3951. Berlin: Springer-Verlag, 2006, 1: 404–417.
- [17] 陈兴蜀, 吴小松. 基于特征关联度的 K-means 初始聚类中心优化算法[J]. *四川大学学报(工程科学版)*, 2015, 47(1): 13–19. (CHEN X S, WU X S. An improved initial cluster centers selection algorithm for K-means based on features correlative degree [J]. *Journal of Sichuan University (Engineering Science Edition)*, 2015, 47(1): 13–19.

This work is partially supported by the Chongqing Municipal Education Commission Science and Technology Research Project (KJ130512), the Chongqing Science and Technology Commission Project (CSCT2015jcyjBX0066).

LUO Yuan, born in 1972, Ph. D., professor. Her research interests include image processing, pattern recognition, machine vision.

LI Huimin, born in 1991, M. S. candidate. Her research interests include pattern recognition, machine vision.

ZHANG Yi, born in 1966, Ph. D., professor. His research interests include robot technology, machine vision.

(上接第 2247 页)

- [8] YANG X, ZHANG T, XU C. A new discriminative coding method for image classification [J]. *Multimedia Systems*, 2015, 21(2): 133–145.
- [9] GAO S, TSANG W H, MA Y. Learning category-specific dictionary and shared dictionary for fine-grained image categorization [J]. *IEEE Transactions on Image Processing*, 2014, 23(2): 623–634.
- [10] KIM S, KWEON I S, LEE C W. Visual categorization robust to large intra-class variations using entropy-guided codebook [C]// *ICRA 2007: Proceedings of the 2007 IEEE International Conference on Robotics & Automation*. Piscataway, NJ: IEEE, 2007: 3793–3798.
- [11] EPSHTEIN B, ULLMAN S. Feature hierarchies for object classification [C]// *ICCV 2005: Proceedings of the 2005 Tenth IEEE International Conference on Computer Vision*. Piscataway, NJ: IEEE, 2005: 220–227.
- [12] LU Z, WANG L, WEN J R. Image classification by visual bag-of-words refinement and reduction [J]. *Neurocomputing*, 2016, 173: 373–384.
- [13] KELBERT M, SUHOV Y. *Information Theory and Coding by Example* [M]. Oxford: Cambridge University Press, 2013: 18–86.
- [14] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(60): 91–110.
- [15] TUIA D, VOLPI M, DALLA MURA M, et al. Automatic feature learning for spatio-spectral image classification with sparse SVM [J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2014, 52(10): 6062–6074.
- ZHANG Yong**, born in 1963, professor. His research interests include intelligent information processing, data mining.
- YANG Hao**, born in 1991, M. S. candidate. His research interests include image classification, machine learning.