



文章编号:1001-9081(2019)05-1466-07

DOI:10.11772/j.issn.1001-9081.2018112522

## 基于颜色四通道及空间金字塔的鱼类图像检索

张美玲<sup>1,2</sup>, 吴俊峰<sup>1,2,3</sup>, 于红<sup>1\*</sup>, 崔榛<sup>1,2</sup>, 董婉婷<sup>1,2</sup>

(1. 大连海洋大学 信息工程学院, 辽宁 大连 116023; 2. 广东省普及型高性能计算机重点实验室, 广东 深圳 518060;

3. 天津大学 计算机科学与技术学院, 天津 30072)

(\* 通信作者电子邮箱 yuhong@dlou.edu.cn)

**摘要:**随着计算机视觉技术在海洋水产领域中的应用不断加深,鱼类图像检索在渔业资源调查、鱼类行为学分析等方面发挥了巨大的作用。通过研究发现,鱼类图像的背景信息会对鱼类图像检索造成极大干扰,而且鱼类图像中颜色、纹理、形状等特征由于空间位置信息的缺乏而使检索的准确率不高。为解决以上问题,提出了一种新的基于颜色四通道及空间金字塔的鱼类图像检索算法。首先,提取视觉显著性图将鱼类图像的前景和背景分开,从而减少图像背景对检索的干扰;其次,为了使图像特征包含一定的空间位置信息,利用空间金字塔的理论对图像进行分割,在此基础上,将图像转为HSVG四通道图并提取SURF特征;最后,得到检索结果。为验证所提算法的有效性,在QUT\_fish\_data数据集和DLOU\_fish\_data数据集上对算法的查全率、查准率与经典的HSVG算法和显著性分块算法进行对比:在两个数据集上查准率分别比传统的HSVG算法最多分别提高12%和5%,查全率最多分别提高7%和22%;比传统的显著性分块算法查准率最多分别提高15%和5%,查全率最多分别提高36%和22%;从而证明所提算法是有效的,能有效提升鱼类图像的检索效果。

**关键词:**鱼类图像检索;颜色通道;空间金字塔;图像特征

**中图分类号:** TP751    **文献标志码:**A

### Fish image retrieval algorithm based on color four channels and spatial pyramid

ZHANG Meiling<sup>1,2</sup>, WU Junfeng<sup>1,2,3</sup>, YU Hong<sup>1\*</sup>, CUI Zhen<sup>1,2</sup>, DONG Wanting<sup>1,2</sup>

(1. College of Information Engineering, Dalian Ocean University, Dalian Liaoning 116023, China;

2. Guangdong Province Key Laboratory of Popular High Performance Computers, Shenzhen Guangdong 518060, China;

3. School of Computer Science and Technology, Tianjin University, Tianjin 300072, China)

**Abstract:** With the development of the application of computer vision in the field of marine fisheries, fish image retrieval has played a huge role in fishery resource survey and fish behavior analysis. It is found that the background information of fish images can greatly interfere with fish image retrieval, and the fish image retrieval results only using color, texture, shape and other characteristics of fish images are not accurate due to the lack of spatial position information. To solve the above problems, a novel fish image retrieval algorithm based on HSVG (Hue, Saturation, Value, Gray) four-channel and spatial pyramid was proposed. Firstly, a visual saliency map was extracted to separate the foreground and the background, thereby reducing the interference of the image background on the retrieval. Then, in order to contain certain spatial position information, the fish image was converted into an HSVG four-channel map, and on this basis, the theory of spatial pyramid was used to segment the image and extract the SURF (Speed Up Robust Feature). Finally, the search results were obtained. In order to verify the effectiveness of the proposed algorithm, the recall and precision of the algorithm were compared with classic HSVG algorithm and saliency block algorithm on QUT\_fish\_data dataset and DLOU\_fish\_data dataset. Compared with traditional HSVG algorithm, the precision on two datasets is increased at most by 12% and 5%, and the recall is increased at most by 7% and 22%, respectively. Compared with saliency block algorithm, the precision on two datasets is increased at most by 15% and 5%, and the recall is increased at most by 36% and 22%, respectively. So, the proposed algorithm is effective and can improve the retrieval results significantly.

**Key words:** fish image retrieval; color channel; spatial pyramid; image feature

## 0 引言

鱼类图像检索技术为鱼类知识科普、鱼类资源调查及种

群分析、鱼病诊断等提供了新思路和新方法,具有重要的研究意义。鱼类图像有前景背景复杂难以区分而且难以识别等问题,同时鱼类图像具有丰富的颜色、纹理、形状、位置等特征,

收稿日期:2018-12-04;修回日期:2019-02-20;录用日期:2019-03-08。    基金项目:国家自然科学基金资助项目(61701070,61802046);辽宁省自然科学基金资助项目(20170520327);辽宁省高等学校海洋产业技术研究院项目(2018-CY-34);广东省普及型高性能计算机重点实验室开放基金资助项目(SZU-GDPHPCL201805);大连市科技计划项目(2015A11GX022)。

**作者简介:**张美玲(1992—),女,山西吕梁人,硕士研究生,主要研究方向:计算机视觉、模式识别;吴俊峰(1983—),男,辽宁大连人,讲师,博士,CCF会员,主要研究方向:智能信息处理、计算机视觉;于红(1968—),女,辽宁大连人,教授,博士,CCF会员,主要研究方向:智能信息处理、大数据;崔榛(1995—),女,辽宁丹东人,硕士研究生,主要研究方向:计算机视觉、模式识别;董婉婷(1995—),女,辽宁鞍山人,硕士研究生,主要研究方向:模式识别、大数据。



这些特征可以为鱼类图像检索提供有价值的信息。因此,鱼类图像检索既要考虑有效利用颜色、纹理、形状、位置等特征又要考虑避免背景对检索效果产生过大的影响。

为解决这个难题,许多专家学者对此进行深入的研究。Bosch 等<sup>[1]</sup>提出一种基于 HSV(Hue, Saturation, Value)三通道的尺度不变特性变换(Scale-Invariant Feature Transform, SIFT)特征提取算法,该算法对图像中的像素对应的三个通道都提取 SIFT 特征并组成  $128 \times 3$  的特征描述子(标准的 SIFT 特征为 128 维特征向量),然后利用提取到的 SIFT 特征<sup>[2]</sup>实现对不同场景图像的分类,但 SIFT 算法本身的计算复杂度较高,其特征提取的计算成本很大,同时可能会在 SIFT 描述子中混入不必要的噪声。Chuang 等<sup>[3]</sup>提出 SHPC(Systematic Hierarchical Partial Classification)算法以实现对鱼类的识别,通过一种逐层带偏袒策略的分类实现对不同种类鱼在空间位置中的识别,虽然实现了空间位置的识别但是效率和精度上有一定的局限性,对自然场景中的鱼种不能很好地识别其空间位置。黄仁等<sup>[4]</sup>提出的基于颜色空间特征的图像检索能够很好地解决这一问题,该算法采用 HSV 颜色模型,在 HSV 空间中采用非等间隔量化的方式对图像中的像素逐个量化,计算并归一化图像的颜色直方图和每类颜色的中心位置从而得到图像的颜色空间特征。但是该算法是在传统的颜色直方图的基础上加入相应的位置信息,可以避免不同图像具有相同颜色直方图的情况,该方法是针对全局特征进行检索的,其检索精度不理想。胡二雷等<sup>[5]</sup>提出一种基于深度学习的图像检索系统,该算法采用 Alexnet 神经网络,前 5 个卷积层提取的是图像的低级可视化特征,后 3 层提取的是图像的高级特征。其检索库为 20 万张图片,检索出来前 10 张的平均相似度在 80% 以上,精度不够高,而且需要的训练时间较长。

经过进一步研究发现,采用鱼类图像的其中一种特征进行匹配识别往往不够全面<sup>[6]</sup>,因而要综合考虑鱼类图像的各种特征,采用多特征合并的方式进行鱼类图像的检索来提高检索效果<sup>[7-8]</sup>。传统的多特征鱼类图像检索采用多特征向量合并方式将颜色特征和纹理特征合并后进行鱼类图像检索,这种多特征合并方式对表达多种特征的能力较弱,合并后的多特征向量无法有效地将鱼类的颜色特征和纹理特征进行融合,从而影响鱼类图像检索算法的性能。其次目前的图像检索算法大多只能对背景简单的图像进行检索,对自然场景下的图像检索还没有较多的研究,针对鱼类图像以及自然场景下的鱼类图像检索则相对较少,局限于单一场景,很难将鱼类图像检索技术应用到真实的自然场景即海洋渔业领域中,其检索的结果误差较大。

基于以上分析,本文提出一种基于 HSVG(Hue, Saturation, Value, Gray)四通道及空间金字塔的鱼类图像显著区域检索算法以提高鱼类图像检索的准确性和鲁棒性。首先,根据鱼类图像的特点提取鱼类图像的视觉显著性图,并根据显著性图自适应的将鱼类图像的前景和背景有效地分开,从而避免图像背景对检索的干扰;然后,根据空间金字塔的理论对图像进行分块,在此基础上,将图像转为 HSVG 四通道

图和 SURF(Speed Up Robust Feature)的提取;最后,进行最终的匹配并得到检索结果。

## 1 本文算法

### 1.1 本文算法框架

针对自然场景下的鱼类图像检索问题,本文提出基于 HSVG 四通道及空间金字塔的鱼类图像检索算法,通过并行提取四通道下鱼头、鱼身和鱼尾的 SURF 鱼类特征使得所提特征能够融合颜色与纹理两种不同类型的特征,提高不同鱼类之间的区分度。然后对提取到的 SURF 特征进行特征匹配,最后对 SURF 特征匹配结果进行综合加权后求得最终特征的相似度,根据相似度结果进行排序得出检索结果。该算法的具体步骤如图 1 所示。

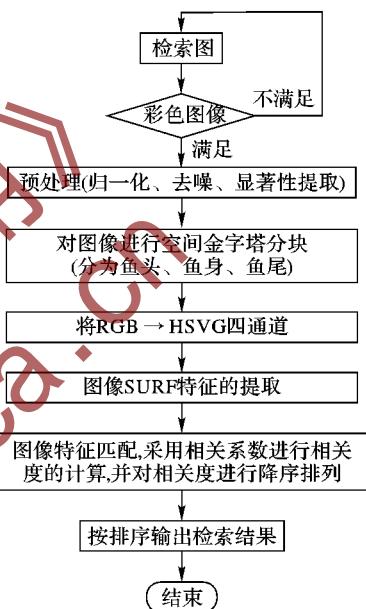


图 1 鱼类图像检索算法流程

Fig. 1 Flow chart of fish image retrieval algorithm

### 1.2 图像的预处理

在进行图像检索过程中,首先需要进行实验数据的收集及整理,由于收集图像的工具和渠道不同可能导致收集到的图像差异很大,因此需要对图像进行预处理。图像预处理的目的是最大限度地简化数据并且统一数据格式。预处理过程一般有归一化、去噪、显著性提取以及图像增强等。

本文所使用的数据来源由三部分组成:实际拍摄、互联网下载及由澳大利亚昆士兰大学提供。由于不同渠道得到的数据其大小不尽相同,这对于特征的提取匹配会有很大的影响。为方便后续工作,将所有的数据归一化到相同的分辨率大小。此外,收集到的鱼类图像中有可能有单通道图像,即灰度图,此时需要将这些图像去除掉,以避免影响后续的实验。同时由于图像的不同获取方式,使得图像中可能夹杂不同种类的噪声,比如高斯噪声、椒盐噪声、均匀噪声等,本文主要针对椒盐噪声和高斯噪声进行了处理,针对椒盐噪声和高斯噪声,采取了中值滤波的方法进行处理。此外,鱼类图像的背景信息会对鱼类图像检索会产生极大的干扰,许多非鱼类的特征信息被统计到鱼类特征中,由此会影响最终的检索效果。为了解决这个问题,比较理想的方法就是将图像中的鱼类与其背景分离开来。由此提出使用视觉显著性图来解决这个问题。



图像的显著性区域是最能引起人类视觉注意力的区域,因此显著性区域在绝大多数情况下为图像中的前景区域,或者说是图像中的感兴趣区,也是图像检索的目标区域。而非显著性区域一般情况下为背景,也是在特征匹配过程中需要去除的区域。一般情况下,鱼类图像本身前后背景颜色差异大,轮廓和纹理比较清晰,由此利用鱼类图像中的颜色差异性和模式差异性来提取图像的显著性图<sup>[7]</sup>:

$$S(p_x) = P(p_x) \cdot C(p_x) \cdot G(p_x) \quad (1)$$

其中: $p_x$ 是图像块, $S(p_x)$ 是图像块的显著性值, $P(p_x)$ 是图像块的模式差异性, $C(p_x)$ 是图像块的颜色差异性, $G(p_x)$ 是高斯权重公式。根据鱼类图像场景的均匀分布,但又存在一定波动性的特性,给出了判断鱼类目标或背景信息的阈值计算方法,如式(2)所示:

$$p_m = \left( \sum_{i=1}^m \sum_{j=1}^n S_{m \times n} \right) / n \quad (2)$$

其中 $S$ 为 $m * n$ 大小的显著图。利用式(2)所计算得到的阈值,将显著图二值化后对二值图像进行面积滤波从而避免图像背景中的非鱼类显著性区域对前景造成的干扰。

### 1.3 空间位置特征

一般情况下,按指定块数或指定长宽比对图像进行分块,分块是基于空间金字塔匹配原理进行的。空间金字塔通常是指将图像经过下采样处理后得到不同分辨率的图像组合,空间金字塔模型(Spatial Pyramid Model, SPM)<sup>[9]</sup>是一种利用空间金字塔进行图像匹配、识别、分类的算法<sup>[10]</sup>。

空间金字塔分块是在不同分辨率的小块上统计图像特征点分布,从而获取图像的空间信息。金字塔多分辨率生成较快,且占用存储空间少。由于鱼类图像检索中鱼类目标基本位于图像的中央主体部分,因而采用空间金字塔模型将鱼类图像分为4块、6块、8块及10块进行实验(如图2),通过大量的实验发现将鱼类图像分为3块所得到的结果更为高效并且实验结果较好。

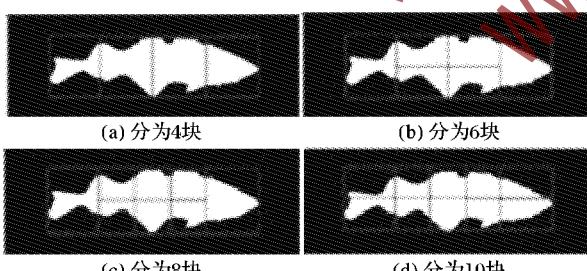


图2 不同类型的空间金字塔模型

Fig. 2 Different types of spatial pyramid model

现阶段实验所用的鱼类图库较为规范,鱼类图像中的鱼头鱼身和鱼尾的方向一致,因此按等分的方式将鱼分为3块已经初步符合其生理特性,如果进一步细分会增加其匹配的复杂度,同时因为分块区域太多容易使得鱼类图像碎片化,降低其生理特征的完整度。因此本文采用均匀分块的方式将鱼类图像分为鱼头、鱼身、鱼尾三部分<sup>[11]</sup>进行。

用空间金字塔方法表示图像是对传统方法的改进。传统方法提取图像特征时,首先提取每张检索图像的特征描述子和图库中所有图像的特征描述子,之后将检索图的特征和图库中的所有特征进行匹配,这样会丢失图像的空间分布信息即位置信息,会出现鱼头的特征点与鱼尾的特征点进行匹配

的情况,明显会有很多错误的匹配,不能够对图像进行精确的识别匹配,因而在此基础上提出空间金字塔方法,它是在不同分辨率上进行图像的特征分布统计,从而获得图像的空间信息,即将图像划分为精细的网格序列,从每个网格中导出特征并组合成为一个很大的特征向量。本文基于空间金字塔的精确匹配,提出将图像分块处理,在对鱼类图像进行预处理等一系列操作后,将图像分割成大小相等的矩形区域进行特征提取,这样能够提高鱼类图像的识别率。如图3是将鱼类原图和显著性二值图<sup>[12-13]</sup>进行分块的示意图。

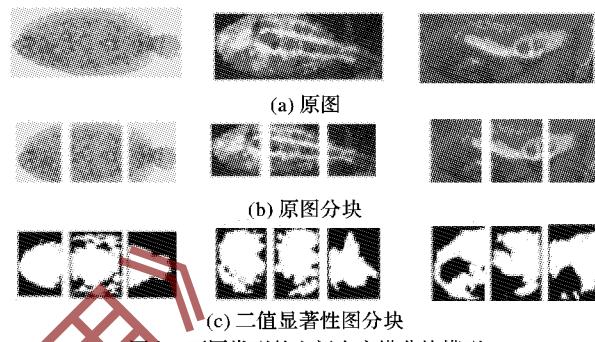


图3 不同类型的空间金字塔分块模型

Fig. 3 Different types of spatial pyramid blocking model

### 1.4 HSVG 四通道颜色特征

鱼类图像本身包含着丰富的颜色信息,这些颜色往往在色调、纯度和饱和度上各有差异,鱼类图像的花纹在单一的H,S或V通道中更为明显,这种颜色通道能够从三个方面来描述颜色的特性,在对颜色表述上能够相互补充,更有效地提取出鱼类图像所蕴藏在颜色差异中的结构信息,进而一定程度上反映颜色信息;同时图像中除了具有丰富的颜色信息之外,还会因为光照的影响在鱼类表面产生强弱程度不等的反射或产生明暗相间的花纹和轮廓,因此在HSV三通道外引入灰度图作为第四通道提取SURF特征。本文综合考虑色差和光照这两种因素的影响,提出基于HSV颜色模型和灰度Gray的四通道模型即HSVГ四通道的颜色空间模型。

### 1.5 纹理特征

SURF是图像检索中基于内容(Content-Based Image Retrieval, CBIR)的一种图像特征提取方法<sup>[14]</sup>,常用的CBIR特征有颜色、纹理等,相较于这些特征,SURF能够更加细致地表示图像的局部信息。在鱼类图像检索中,鱼的局部差异往往比较明显而全局差异不是特别大,此外由于拍摄角度等问题,图像中的鱼类即使是同一条鱼,它的整体形态也会发生变化,在这种情况下全局特征的不变性会较差。相比之下,SURF属于局部特征因而能够更好地反映出图像中的局部差异,并且具有更强的特征不变性,因此该算法使用SURF特征能够更好地对鱼类图像进行检索。

## 2 鱼类图像的特征提取和匹配

### 2.1 特征提取

为提高利用SURF特征提取和匹配的准确度,采用预处理技术得到鱼类图像显著图来提取SURF特征,这些特征集中于鱼类的关键区域。相对于直接从原始图像中提取的SURF特征,利用显著图提取到的SURF特征更加稀疏而精确,同时能够从鱼类图像中获取更为丰富的有效信息。在此基础上,将鱼类图像利用空间金字塔技术分为鱼头、鱼身、鱼



尾三部分,然后分别提取每一部分的 HSVG 四个通道下的 SURF 特征,进而分别得到鱼头、鱼身、鱼尾三个部分的 HSVG 四个颜色通道的不同的 SURF 特征,使得所提特征包含位置信息,避免在特征描述子中混入不必要的噪声信息<sup>[15]</sup>,能更好地表现鱼类的形态和局部特征;同时分四个通道提取的特征包含鱼类图像由色差和光照影响的轮廓特征和纹理特征,从而进一步提高不同种类鱼之间的区分度。这种方式能够通过通道互补使得图像具有更完整的表述能力,这为精确检索提供有效的帮助。在此基础上采用多核 CPU 并行处理的方式对四通道下鱼头、鱼身和鱼尾的 SURF 特征进行计算,从而降低检索的时间消耗。

## 2.2 特征匹配

图像特征匹配是指将图像特征进行相似度计算对图像的相似程度作出评价。特征匹配是图像检索模型的一个重要的环节,在提取出图像的特定特征之后进行相似度的匹配。图像相似度匹配方法有很多种,不同的匹配方法会影响检索图与图库图像之间的相似度,即图像特征点之间匹配的相似程度。

多通道下的图像检索能够带来比单通道检索更为丰富的特征,但是多通道的 SURF 特征会面临多种特征匹配方式的选择问题,不同特征匹配方式会影响提取到的特征向量的保存方式,同时其特征利用效率以及匹配精度都有差异。经过归纳整理,发现多通道下的 SURF 特征匹配方法主要分为两类,一类匹配方法为将不同通道中的 SURF 特征进行加权后变为列向量然后进行特征匹配,另外一类为将特征按通道分别进行特征匹配,然后将匹配后的特征得分进行加权得到最终匹配得分。第一类特征匹配方式能够降低特征本身的复杂度从而加快特征匹配的过程,特征匹配的效率较高。第二类特征匹配方式能够更为有效地利用多通道 SURF 互补性,因而从检索精度上分析,该方法的匹配结果精度要高于第一类特征匹配方法。但是该方法的特征匹配效率要低于第一类特征匹配方法。本文综合考虑图像检索的精度和效率问题,对不同的特征提取和特征匹配的耗时进行研究分析后发现,在鱼类图像检索的整个过程中,耗时主要集中在图像的特征提取环节,相比较之下图像的特征匹配的总体耗时比重不大,鱼类图像检索的整体效率依然可以保持基本不变。此外由于是采用独立匹配的方式,可以对各部分进行并行处理,从而保证算法的时间复杂度不变,因此本文选取第二类图像特征匹配方式对从多通道中提取的鱼类图像 SURF 特征进行特征的提取和匹配。

常用的图像特征匹配方式有距离度量和相关系数度量两类度量方式,其中距离度量方式能够体现特征向量之间的数值差异,相关系数度量方式能够体现出特征向量之间相关关系密切程度<sup>[16]</sup>。在 HSVG 四通道及空间金字塔下的 SURF 特征具有一定的差异性和互补性,其特征差异往往不体现在特征的数值差异上而是主要体现在特征向量的相位差异上。所以,采用相关系数的度量方式更加符合 HSVG 四通道和空间金字塔下 SURF 特征的特征匹配特性。本文首先将提取到的鱼类图像特征向量进行归一化处理,经过归一化的特征向量之间其差异性相对被压缩,因此采用相关系数度量方式对其进行相关系数的计算。

本文首先提取四通道下每个通道中的鱼头、鱼身、鱼尾三

个部分的 SURF 特征,按照特征所属通道的不同空间位置单独保存。在特征匹配过程中,为了能够继续保持各个颜色通道和各个空间位置的特征独立性,本课题组只针对检索图和图库图像的相同通道下的相同空间位置的 SURF 特征进行特征匹配。其匹配公式如下:

$$\left\{ \begin{array}{l} S_{H_1} = \text{corr}(F_{dH_1}, F_{qH_1}) \\ S_{H_2} = \text{corr}(F_{dH_2}, F_{qH_2}) \\ S_{H_3} = \text{corr}(F_{dH_3}, F_{qH_3}) \\ S_{S_1} = \text{corr}(F_{dS_1}, F_{qS_1}) \\ S_{S_2} = \text{corr}(F_{dS_2}, F_{qS_2}) \\ S_{S_3} = \text{corr}(F_{dS_3}, F_{qS_3}) \\ S_{V_1} = \text{corr}(F_{dV_1}, F_{qV_1}) \\ S_{V_2} = \text{corr}(F_{dV_2}, F_{qV_2}) \\ S_{V_3} = \text{corr}(F_{dV_3}, F_{qV_3}) \\ S_{G_1} = \text{corr}(F_{dG_1}, F_{qG_1}) \\ S_{G_2} = \text{corr}(F_{dG_2}, F_{qG_2}) \\ S_{G_3} = \text{corr}(F_{dG_3}, F_{qG_3}) \end{array} \right. \quad (3)$$

其中,corr 为图像特征相关系数的计算公式,本文采用皮尔逊相关系数作为四通道鱼类 SURF 特征的特征计算方法。式(3)计算 HSVG 四通道每个通道下检索图和图库图像的鱼头、鱼身和鱼尾的 SURF 特征的相关系数,在此基础上计算 HSVG 四通道下相同位置的 SURF 特征的相似度匹配得分  $S_1$ ,如式(4)。

$$S_1 = w_{H_1} \cdot S_{H_1} + w_{S_1} \cdot S_{S_1} + w_{V_1} \cdot S_{V_1} + w_{G_1} \cdot S_{G_1} \quad (4)$$

其中参数  $w_{H_1}$ 、 $w_{S_1}$ 、 $w_{V_1}$ 、 $w_{G_1}$  为鱼类图像检索图和图库图像对应通道提取到鱼头的 SURF 特征其相似度得分权重。本课题组是基于 HSVG 四通道及空间金字塔的鱼类图像 SURF 特征提取匹配检索过程,其通过四通道提取的 SURF 特征相互之间的互补性的重要程度相当,因而本课题组采用的四通道的权重相等,此外为了防止四通道相似度叠加后的值过大,因而  $w_{H_1} = w_{S_1} = w_{V_1} = w_{G_1} = 1/4$ 。

四通道下的 SURF 特征向量首先按通道独立匹配,然后将四通道匹配得分加权计算综合得分,采用这种方法所得到的相似度值,其特征匹配过程更加全面,特征匹配效果相较于仅从灰度图中提取 SURF 特征并匹配的效果要更加精确。

类似的,得到鱼身、鱼尾的 SURF 特征的相似度得分,最终得到检索图和图库中每条鱼的相似度得分,如式(5)所示:

$$S = W_1 \cdot S_1 + W_2 \cdot S_2 + W_3 \cdot S_3 \quad (5)$$

其中: $S_1$ 、 $S_2$ 、 $S_3$  为鱼头、鱼身和鱼尾的相似度得分,加权求和后得到这幅检索图同图库中的其中一幅鱼类图像的匹配相关度。同样的,通过空间金字塔提取的 SURF 特征之间的位置细腻的重要程度相当,因而本课题组采用的鱼头、鱼身和鱼尾的权重相等,同样地为了防止三部分相似度叠加后的得到的相似度值过大,因而其权重  $w_1 = w_2 = w_3 = 1/3$ 。经过实验验证四通道采用相同的权重同时鱼头鱼身鱼尾采用相同的权重进行特征的匹配其准确率较高,因此本文采用相同权重进行相似度的叠加。

将特征匹配加权得分进行排序,得到鱼类图像检索结果,如式(6):



$$R = \text{sort}(R) \quad (6)$$

其中:  $R = [r_1, r_2, \dots, r_i, \dots, r_n]$ ,  $n \in (1, 2, \dots, n)$ 。

下面给出特征匹配的基本步骤:

**步骤1** 提取鱼类检索图和图库图像的 HSVG 每个通道下鱼头、鱼身、鱼尾三个部分的 SURF 特征点  $F_q^i, F_d^i$ , 其中  $F_q^i$  表示检索图特征点,  $F_d^i$  表示图库图像特征点。

**步骤2** 对鱼类图像的 SURF 特征点进行点匹配, 得到单通道下检索图的每个部分每个 SURF 特征点与单幅图库图像中所有 SURF 特征点之间的最大相关度系数作为该通道该部分鱼类检索图 SURF 特征点与图库图像特征点之间的最高匹配度。

**步骤3** 将得到的该幅鱼类检索图 HSVG 四个通道下鱼头、鱼身、鱼尾部分的 SURF 特征点与单幅图库图像的 SURF 特征点的最高匹配度进行排序, 把排名前三的匹配度进行加权求和, 作为该幅检索图该通道下该部分的相似度得分。

**步骤4** 按上述步骤计算后, 对检索图和图库图像分别按 HSVG 四通道下鱼头、鱼身、鱼尾三部分对应进行特征匹配。

**步骤5** 将得到的检索图和图库图像鱼头、鱼身和鱼尾中某一部分在四通道下的相关系数进行加权计算, 得到这幅鱼类检索图与图库中所有图像该部分的相似度得分。同理计算, 得到鱼类图像检索图与图库每幅图像之间鱼头、鱼身和鱼尾三部分之间的相似度得分。

### 2.3 鱼类图像检索算法

首先, 采用 HSVG 四通道颜色模型和空间金字塔按鱼头、鱼身、鱼尾进行分块; 然后, 分别进行特征提取和匹配, 最后, 加权求得特征相似度。这样既避免因外界因素光照影响和鱼类图像本身的色差产生的影响, 同时也符合鱼类的生理特性, 进而能够有效地提高检索结果的准确度。该算法的伪代码如下:

```

算法 Fish-Image-Retrieval。
Input: query fish images
Output: retrieval-results
1. Image normalization preprocessing;
1.1 Normalization preprocessing;
1.2 Denoising preprocessing;
1.3 Saliency detection;
2. Utilize spatial pyramid to segment the images into 3 blocks;
3. Color space conversion( RGB -> HSVG );
3.1 RGB convert HSV;
3.2 Histogram equalization for H channels;
3.3 Rebuild image structure HSVG.
4. Extract SURF features for each image: Parallel Computation use 4
CPU to process 4 channel of image respectively;
5. Feature matching between images;
6. Merging the scores of four channels;
7. Sort score and output the results

```

### 3 实验结果与分析

本文在 Windows7 系统下使用 Matlab2012b<sup>[17]</sup> 进行图像检索仿真实验。实验环境的硬件配置为: Intel Xeon CPU E5-250 v2 2.60 GHz 32 核处理器, 32 GB 内存。

本文采用的数据集为澳大利亚昆士兰大学提供的 QUT\_fish\_data 鱼类图像数据集, 该图像库包含 4405 张海洋鱼类静

态图像, 每张图像标有对应的种类标签。该图像库是目前已知的最大规模鱼类图像数据标注集。在此数据集上可以客观有效地评价图像检索算法的鲁棒性和准确率。此外本文还采用大连海洋大学图像组采集的鱼类场景图像数据集 DLOU\_fish\_data, 该图像库包含 200 张鱼类图像数据, 每类鱼类图像包含不同的鱼类图像场景, 采用此图像库能够在一定程度上评价鱼类图像检索算法对复杂场景下的鱼类图像的检索能力。

**评价方法** 本文采用准确率和召回率对所提算法的有效性进行评价<sup>[18]</sup>。其中准确率反映的是图像检索的精度, 召回率是图像检索算法对同类信息检索查找全面程度的评价指标。此外, 综合考虑检索算法的查准和查全能力, 采用准确率-召回率(Precision-Recall, P-R) 曲线评价图像检索算法:

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

其中:  $TP$ (True Positive) 是检索到的图像中包含有与检索图像相似的同类图的数量,  $FP$ (False Positive) 为检索到的图像中非同类图像的数量,  $FN$ (False Negative) 是图库中全部同类图中未检索到的同类图数量。

为了验证本文所提的基于 HSVG 四通道及空间金字塔的鱼类图像检索算法的有效性和鲁棒性, 采用基于 HSVG 四通道的鱼类图像检索算法(以下简称 HSVG 算法)<sup>[7]</sup> 和基于显著性及空间金字塔的鱼类图像检索算法(以下简称显著性分块算法)<sup>[8]</sup> 作基准, 在 QUT\_fish\_data 数据集和 DLOU\_fish\_data 数据集上进行仿真实验, 并采用 P-R 曲线对图像检索算法性能进行评价。

由于以上两种鱼类图像检索对比算法都有一定的局限性, 不能很好地将其应用到自然场景中, 因此提出基于 HSVG 四通道及空间金字塔的鱼类图像检索算法。该算法将 HSVG 算法和显著性分块算法结合起来, 既采用基于 HSVG 四通道鱼类图像检索算法中颜色特征和纹理特征的差异性和互补性; 同时采用基于显著性及空间金字塔的鱼类图像检索算法中的显著性提取和空间金字塔技术, 避免背景区域误检到的噪声导致的特征被误提取, 采用空间金字塔技术将空间位置信息融入到提取的特征中, 避免特征的误匹配。

为了验证该算法的有效性, 首先在 QUT\_fish\_data 数据集上进行测试, 测试集同样为数据集中 100 次随机挑选的 200 幅不同种类的鱼类图像进行 100 次实验, 检索结果如表 1 所示。

表 1 三种算法在 QUT\_fish\_data 数据集上的性能对比

Tab. 1 Performance comparison of three methods on QUT\_fish\_data

检索数	HSVG		显著性分块		本文算法	
	查准率	查全率	查准率	查全率	查准率	查全率
5	0.95	0.35	0.92	0.25	0.98	0.40
10	0.82	0.52	0.75	0.30	0.90	0.51
15	0.76	0.60	0.65	0.32	0.87	0.62
20	0.72	0.62	0.70	0.35	0.83	0.65
25	0.67	0.67	0.64	0.38	0.79	0.74

从表 1 以及图 4 可以看出, 基于本文算法在 QUT\_fish\_data 数据集上查准率和查全率高于 HSVG 算法和显著性分块



算法。但是该算法与基于 HSVG 四通道的鱼类图像检索算法差异不大,经过分析发现通过这两种算法提取到的特征点相似度较高,究其原因主要在于 QUT\_fish\_data 数据集中鱼类种类繁多,但其前后背景差异不大,因而通过显著性预处理得到前景鱼类中提取得到的特征与不采用显著性预处理提取到的特征差异不大,但是该算法中采用空间金字塔模型,使得提取到的特征中包含空间位置信息,在一定程度上提高检索精度。由图 4 中 P-R 曲线可以进一步验证本文算法在较大规模鱼类图像数据集上具有较强的鲁棒性以及较高的准确度,进而可以在一定程度上通过实验证明该算法理论在鱼类图像检索问题上具有一定的正确性和可靠性。

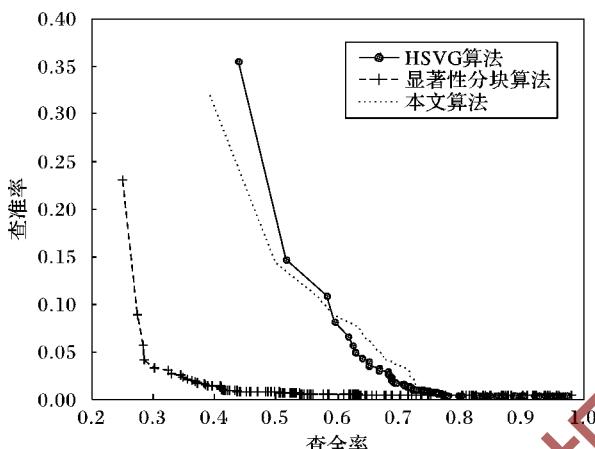


图 4 3 种算法在 QUT\_fish\_data 数据集的 P-R 曲线对比

Fig. 4 P-R curves comparison of three algorithms on QUT\_fish\_data dataset

同时,本文在 DLOU\_fish\_data 数据集上进行扩充实验,从实验结果表 2 来看,在小数据集上,本文算法检索效果比较明显,查全率随着检索数量的增加而显著增加。其原因在于,DLOU\_fish\_data 数据集中的鱼类图像处于自然场景中,背景比较复杂,对特征提取的干扰较为严重;而本文算法采用显著性提取技术,能很好地将鱼类图像的前后背景区分开来,有效地避免背景对前景的干扰,使得提取到的特征更加准确。通过图 5 中的 P-R 曲线可以看出本文检索算法在 DLOU\_fish\_data 上的总体性能要好于对比算法,这证明本文算法对于复杂场景的鱼类图像检索具有一定的鲁棒性和准确度。

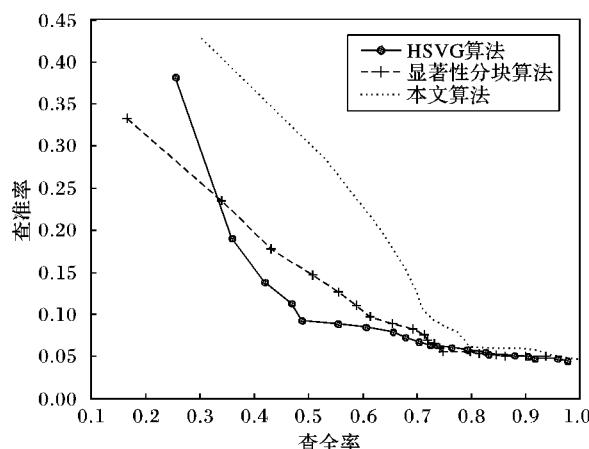


图 5 3 种算法在 DLOU\_fish\_data 数据集的 P-R 曲线对比

Fig. 5 P-R curves comparison of three algorithms on DLOU\_fish\_data

通过以上两组实验可以看出,由于这两种数据集中同类

图的数量较少,因此其准确率随着检索数量的增加而显著降低,但是其召回率显著增高。因此本文所提出的基于 HSVG 四通道及空间金字塔的鱼类图像检索算法在较大规模图像集中具有较强鲁棒性和准确率,并且对复杂的自然场景图像具有一定的适应能力,通过以上实验可以基本证明本文所提理论的客观有效性。

表 2 三种算法在 DLOU\_fish\_data 数据集上的性能对比

Tab. 2 Performance comparison of three methods on DLOU\_fish\_data

检索数	HSVG		显著性分块		本文算法	
	查准率	查全率	查准率	查全率	查准率	查全率
5	0.13	0.40	0.17	0.40	0.22	0.62
10	0.09	0.50	0.12	0.55	0.15	0.68
15	0.08	0.61	0.10	0.61	0.11	0.70
20	0.08	0.65	0.09	0.70	0.08	0.72
25	0.07	0.75	0.08	0.70	0.09	0.82

#### 4 结语

本文提出基于 HSVG 四通道及空间金字塔的鱼类图像检索算法。该算法先利用图像显著性提取算法去除背景对鱼的干扰;再利用空间金字塔模型将鱼类图像分为鱼头、鱼身、鱼尾三部分;然后对鱼类图像按 HSVG 四个通道分别提取 SURF 特征,使得提取到的特征融合形状特征、空间位置特征、颜色特征和纹理特征,利用所提取到的 SURF 特征按通道进行匹配求取检索图与图库图像之间的相关度;最后经过加权求得图像之间的相似度,按相似度由高到低顺序输出检索结果,提高检索准确度,增强检索算法的鲁棒性。为验证所提算法的有效性,在 QUT\_fish\_data 数据集和 DLOU\_fish\_data 数据集上对算法的查全率、查准率与一些经典的鱼类图像检索算法进行了对比,在两个数据集上查准率分别比传统的 HSVG 算法最多分别提高 12% 和 5%,查全率最多分别提高 7% 和 22%,从而证明本文算法是有效的,可有效提升鱼类图像的检索效果。

随着深度学习的火热发展,未来的工作应逐步扩充数据集,设计网络结构来提取到更为有效的鱼类图像特征,以增强鱼类图像检索效果。

致谢 感谢国家科技资源共享服务平台——国家海洋科学数据中心大连分中心(<http://odc.dlou.edu.cn/>)提供数据支撑。

#### 参考文献 (References)

- [1] BOSCH A, ZISSERMAN A, MUÑOZ X. Scene classification using a hybrid generative/discriminative approach[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 30(4): 712 – 727.
- [2] LIAO K, LIU G, HUI Y. An improvement to the SIFT descriptor for image representation and matching[J]. Pattern Recognition Letters, 2013, 34(11): 1211 – 1220.
- [3] CHUANG M C, HWANG J N, KUO F F, et al. Recognizing live fish species by hierarchical partial classification based on the exponential benefit[C]// Proceedings of the 2014 IEEE International Conference on Image Processing. Piscataway, NJ: IEEE, 2014: 5232 – 5236.



- [4] 黄仁,胡敏.综合颜色空间特征和纹理特征的图像检索[J].计算机科学,2014,41(s1):118-121.(HUANG R, HU M. Content-based image retrieval using color position and texture fused features [J]. Computer Science, 2014, 41(s1):118 - 121.)
- [5] 胡二雷,冯瑞.基于深度学习的图像检索系统[J].计算机系统应用,2017,26(3):8-19.(HU E L, FENG R. Image retrieval system based on deep learning[J]. Computer Systems & Applications, 2017, 26(3):8 - 19.)
- [6] GUO T, MOUSAVI H S, MONGA V. Deep learning based image super-resolution with coupled backpropagation[C]// Proceedings of the 2016 IEEE Global Conference on Signal and Information Processing. Piscataway, NJ: IEEE, 2017: 237 - 241.
- [7] 张美玲,吴俊峰,于红,等.一种基于HSVG-SURF特征的鱼类图像检索算法[J].小型微型计算机系统,2018,39(9):2085-2089.(ZHANG M L, WU J F, YU H, et al. Fish image retrieval algorithm based on HSVG four channel SURF feature[J]. Journal of Chinese Computer Systems, 2018, 39(9): 2085 - 2089.)
- [8] ZHANG M L, WU J F, Y H, et al. A novel fish image retrieval method based on saliency spatial pyramid[C]// Proceedings of the 2017 14th International Symposium on Pervasive Systems, Algorithms and Networks & 2017 International Conference on Frontier of Computer Science and Technology & 2017 3rd International Symposium of Creative Computing. Washington, DC: IEEE Computer Society, 2017: 312 - 317.
- [9] 程少光,何毕,布树辉,等.基于超像素空间金字塔模型的场景识别研究[J].计算机工程与应用,2014,50(7):139-143.(CHENG S G, HE B, BU S H, et al. Scene recognition research based on SSPM[J]. Computer Engineering and Applications, 2014, 50(7):139 - 143.)
- [10] 高常鑫,桑农.整合局部特征和滤波器特征的空间金字塔匹配模型[J].电子学报,2011,39(9):2034-2038.(GAO C X, SANG N. Unifying local features and filterbank features in the spatial pyramid matching model[J]. Acta Electronica Sinica, 2011, 39(9):2034 - 2038.)
- [11] YANG J, YU K, GONG Y, et al. Linear spatial pyramid matching using sparse coding for image classification[C]// Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2009: 1794 - 1801.
- [12] 陈鑫元,李筠,杨海马,等.自适应阈值图像二值化及形态学处理的FPGA实现[J].电子测量技术,2016,39(7):67-71.(CHEN X Y, LI Y, YANG H M, et al. Adaptive threshold binarization and morphological image processing based on FPGA[J]. Electronic Measurement Technology, 2016, 39(7):67 - 71.)
- [13] WU L S, CHEN J X, WEI L I. Niblack-based binaryization algorithm for palm vein image[J]. Communications Technology, 2010, 43(1):112 - 114.
- [14] BAY H, TUYTELAARS T, GOOL L V. SURF: Speeded Up Robust Features[C]// Proceedings of the 9th European Conference on Computer Vision. Berlin: Springer, 2006: 404 - 417.
- [15] PYATYKH S, HESSER J, ZHENG L. Image noise level estimation by principal component analysis[J]. IEEE Transactions on Image Processing, 2013, 22(2):687 - 699.
- [16] SALIMI-KHORSHIDI G, SMITH S M, KELTNER J R, et al. Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies[J]. Neuroimage, 2009, 45(3):10 - 23.
- [17] MURTHY A V, KARAM L J. A MATLAB-based framework for image and video quality evaluation[C]// Proceedings of the 2010 2nd International Workshop on Quality of Multimedia Experience. Piscataway, NJ: IEEE, 2010: 242 - 247.
- [18] SINGHA M, HEMACHANDRAN K. Content based image retrieval using color and texture[J]. Signal & Image Processing, 2012, 3(1):271 - 273.

This work is partially supported by the National Natural Science Foundation of China (61701070, 61802046), the Natural Science Foundation of Liaoning Province (20170520327), the Project for Institute of Marine Industry Technology of Universities in Liaoning Province (2018-CY-34), the Open fund of Guangdong Province Key Laboratory of Popular High Performance Computers (SZU-GDPHPCL201805), the Technology Program of Dalian (2015A11GX022).

**ZHANG Meiling**, born in 1992, M. S. candidate. Her research interests include computer vision, pattern recognition.

**WU Junfeng**, born in 1983, Ph. D., lecturer. His research interests include intelligent information processing, computer vision.

**YU Hong**, born in 1968, Ph. D., professor. Her research interests include intelligent information processing, big data.

**CUI Zhen**, born in 1995, M. S. candidate. Her research interests include computer vision, pattern recognition.

**DONG Wanting**, born in 1995, M. S. candidate. Her research interests include pattern recognition, big data.