

# Automatic Image Annotation Based on Multi Feature Fusion

Daizhizi Liu, Xunchang Wang, Miao Zang

North China University of Technology, Beijing  
Email: [ldzz\\_ncut@sina.com](mailto:ldzz_ncut@sina.com)

Received: Apr. 4<sup>th</sup>, 2015; accepted: Apr. 24<sup>th</sup>, 2015; published: Apr. 29<sup>th</sup>, 2015

Copyright © 2015 by authors and Hans Publishers Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Automatic image annotation technology has a very important role in terms of retrieving images, and its importance is reflected in the fact that the image can be retrieved in accordance with a certain characteristic image semantic keywords, which brought unprecedented convenience for the retrieval and use of the user. This paper designs a complete set of automatic image annotation technology. Firstly, we extract multiple features from the pictures, and then do the measurements of similarity on the marked images by the method of multi-feature fusions, so that the picture will be labeled automatically.

## Keywords

Multi Feature Fusion, Automatic Annotation, Picture, KNN Algorithm

---

# 基于多特征融合的图像自动标注

刘代志子, 汪训昌, 臧 淼

北方工业大学, 北京  
Email: [ldzz\\_ncut@sina.com](mailto:ldzz_ncut@sina.com)

收稿日期: 2015年4月4日; 录用日期: 2015年4月24日; 发布日期: 2015年4月29日

---

## 摘 要

图像自动标注技术在检索图像方面具有非常重要的作用, 其重要性体现在可以根据图像语义关键词来检

索具有一定特征的图像，这为用户的检索和使用带来了前所未有的便捷。本文设计了一套完整的图像自动标注技术，首先对图片的多个特征进行提取，之后通过多特征融合的方法与已标注图像进行相似度量从而对图片进行自动标注。

## 关键词

多特征融合，自动标注，图像，KNN算法

## 1. 引言

随着计算机和现代科技的迅猛发展及广泛应用，数字图像的数量暴涨，而且此趋势将延续不衰[1]。图像自动标注技术由此发展而来，其包括图像特征提取和图像自动标注两部分。作为图像自动标注系统的关键技术，图像的特征提取对标注性能的优劣将产生重大的影响。

目前，图像的特征提取主要是针对单特征进行的。但是，用图像的单特征提取来描述图像语义存在着致命的单一性、片面性和局限性问题，效率不高。因此采用多特征融合技术，把提取到的多个特征如颜色特征、纹理特征、形状特征进行融合，将融合后的多特征作为基本数据，从而实现对图像的分类和标注。

基于多特征融合的图片自动标注的方法，首先从图像数据里提取多个特征，经过多重测试分析找到最佳的融合特征，再根据融合的多特征使用 KNN 算法进行相似度量以确定目标图像最佳自动标注特征。经过多重测试大量计算结果表明，此方法既可行又有效。

## 2. 系统框架设计

本系统共有两个模块，即图像特征提取模块和图像自动标注模块，系统框架如图 1 所示。图像特征提取是图像自动标注的基础，决定图像自动标注的准确性；图像自动标注决定图像标注的效果和效率。

图像特征包括颜色、纹理、形状和空间关系等多种特征。颜色(灰度)特征是被广泛使用的特征，提取方法较易。颜色(灰度)特征对于图像的方向和大小的依赖较少，因此具有更优质的鲁棒性。常用的颜色特征有：颜色(灰度)直方图、颜色矩、颜色集、颜色聚合向量、颜色相关图等[2]。其中，颜色(灰度)直方图是实践中最常用的图像统计特征,可使用 RGB、HSV、LAB、OPP 等颜色空间。

纹理特征是表示图像的另一种重要的视觉特征，不依赖于颜色或亮度的变化，能够反映图像亮度的空间变化情况。常用的纹理特征有基于统计方法和基于信号处理方法两大类。灰度共生矩阵、Tamura 纹理特征是常用的基于统计方法的纹理特征。灰度共生矩阵特征提取与匹配主要依赖于能量、惯量、熵和相关性四个参数。Tamura 纹理特征基于人类对纹理的视觉感知心理学研究，提出 6 种属性，即：粗糙度、对比度、方向度、线像度、规整度和粗略度[3]。Gabor 特征、Wavelet 小波特征，Haar 特征等是常用的基于信号处理的纹理特征提取方法。

形状特征是图像的核心特征之一，它是一种稳定的特征，包括区域特征和轮廓特征。其中，前者关心的是整个形状区域，后者关心的是物体的外边界。

在图像自动标注模块中，使用的是 KNN 最近邻算法。K 最近邻(K-Nearest Neighbor, KNN)分类算法的基本思路是：如果一个样本图像在特征空间中的 K 个最相似(即特征空间中最邻近)的样本图像中的大

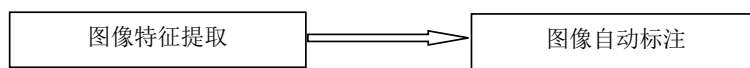


Figure 1. System design diagram summary

图 1. 系统概要设计框图

多数属于某一个类别，则该样本图像也属于这个类别，并具有这个类别上样本图像的特性。

### 3. 系统实现

系统首先提取图像库中所有图像的全部特征，并载入图像人工标注的关键词，然后用 500 幅待标注的测试图像依次计算其与所有训练图像之间的距离，从而找到距离最近、即相似度最高的图像的关键词标注给待标注图像。最后通过调用 Corel-5K 图像库来对比分析自动标注的结果与人工标注的结果，由此计算出查准率、查全率和标注正确的关键词个数。

#### 3.1. 特征提取

系统使用 RGB、HSV、OPP 颜色空间直方图作为图像颜色特征。RGB 直方图是在 RGB 颜色空间上 R、G、B 三个分量图像的一维直方图的组合。每个分量图像直方图是一个 16 维的特征向量，从而得到 48 维特征向量。HSV 和 OPP 颜色直方图首先需要进行颜色空间变换，然后再生成直方图特征向量。HSV 颜色空间变换目前存在基本转换公式以及快速近似转换公式[4]。OPP 颜色空间变换公式如式(1)所示，OPP 对每个分量图像直方图生成一个 64 维的特征向量，从而得到 192 维特征。

$$\begin{pmatrix} O1 \\ O2 \\ O3 \end{pmatrix} = \begin{pmatrix} \frac{R-G}{\sqrt{2}} \\ \frac{R+G-2B}{\sqrt{6}} \\ \frac{R+G+B}{\sqrt{3}} \end{pmatrix} \quad (1)$$

系统纹理特征使用 Gabor 和传统小波 Wavelet 特征。Gabor 特征提取是在 3 个尺度 4 个方向上对图像进行 Gabor 滤波，得到 12 幅滤波图像，对每幅图像的幅值提取均值和方差，对每幅图像的相角提取均值，共得到 36 维特征向量。Wavelet 特征是对图像进行 Daubechies4 小波对图像进行分解，然后提取一层分解的四个频带小波系数的均值和方差作为图像的 8 维纹理特征[5]。

#### 3.2. 自动标注

使用 KNN 最近邻方法对图像进行自动标注的程序流程图如图 2 所示。

## 4. 实验与分析结果

本系统测试使用的是 Corel-5k 图像库。Corel-5k 由 50 张 CD 组成，包含 50 个语义主题[6]。Corel-5k 图像库通常被分成两个部分：4500 张图像作为训练集，其余 500 张作为测试集评价算法性能。该图像库中的每张图片被标注 1~5 个标注词，训练集中总共有 374 个标注词。

为了检测本图像自动标注系统的性能，进行以下实验，自动标注的性能将使用查准率、查全率和检索正确的关键词个数来评价。查准率和查全率定义见式(2) (3)：

$$\text{查全率}(R) = \frac{\text{相关的正确标注结果}}{\text{所有相关的结果}} = \frac{A}{A+C} \quad (2)$$

$$\text{查准率}(P) = \frac{\text{相关的正确标注结果}}{\text{所有标注的结果}} = \frac{A}{A+B} \quad (3)$$

其中 A 为正确标注数目，B 为误标注数目，C 漏标注数目。

#### 4.1. 实验一：使用单一特征的自动标注实验

采用仅载入一种图像特征来测试自动标注结果的方法，即只载入单一的某种颜色特征或单一的某种

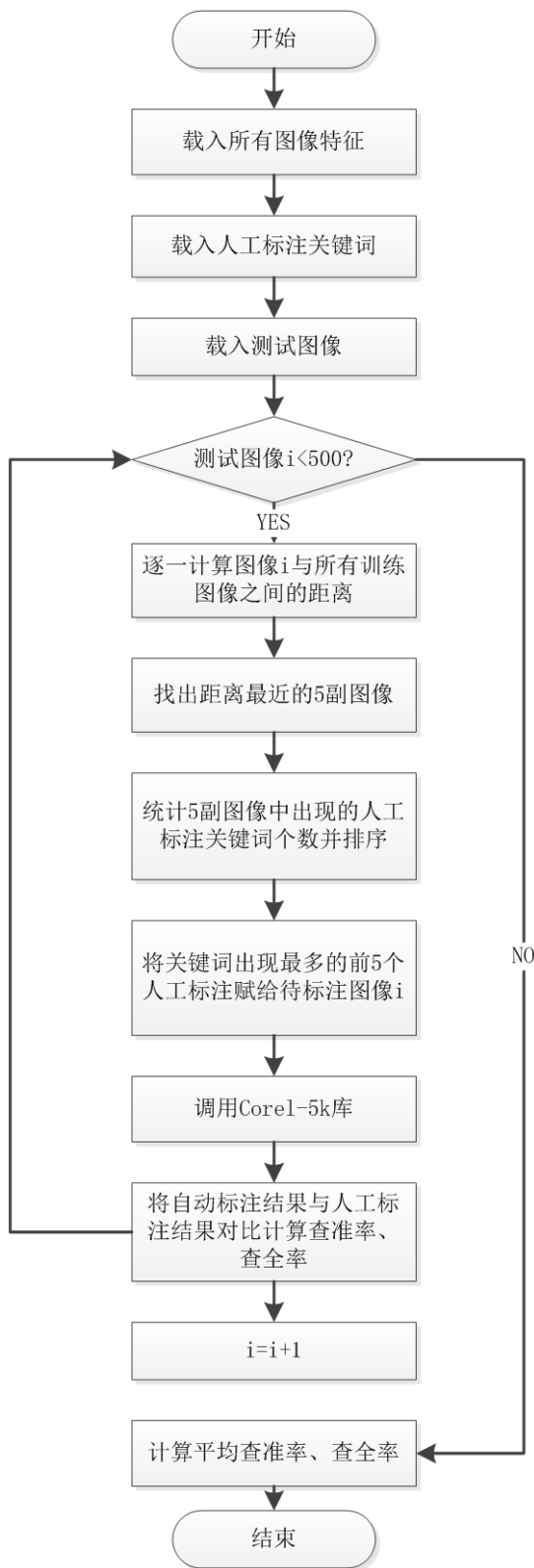


Figure 2. K-nearest neighbor method automatic annotation program flowchart

图 2. K 最近邻方法自动标注程序流程图

纹理特征或其他特征，最终得到查准率、查全率和标注正确的关键词个数，表 1 中列举了几种仅用颜色特征或纹理特征的结果。

由表 1 可见，使用单一特征时，用 HSV 颜色特征的标注效果最显著，RGB 特征与 OPP 特征效果相近，Wavelet 不太适合作为单一特征来进行图像的自动标注。

#### 4.2. 实验二：使用两种特征的自动标注实验

采用载入两种图像特征来测试自动标注结果的方法，即载入两种颜色特征，或载入两种纹理特征，或两种相同的其他特征，抑或两种不同特征的结合，表 2 列举了部分使用两种特征进行标注的结果。

由表 2 可见，使用双特征时，从查准率、查全率和标注正确的关键词个数三方面考察，HSV 与 Gabor 相结合的特征标注效果都明显优于其他双特征，尤其优于 HSV 与 Wavelet 相结合的特征组合。

#### 4.3. 实验三：使用多种特征的自动标注

采用载入本系统提出并且实现的基于多特征融合的图像自动标注方法来测试自动标注结果，即载入训练图像所有的特征，用相融合后的特征来标注待标注图像，最终测试部分结果列举如表 3、表 4 所示。

由表 3 所示，随着使用自动标注的特征数的增加，查准率、查全率准确性均提高，标注正确的关键词个数也呈上升趋势。因此，基于多特征融合的标注效果优势明显。

表 4 给出了使用全部特征的部分图像自动标注结果，尽管存在自动标注词和人工标注词有个别出入的情况，但是自动标注词仍有非常大的参考价值，能够从很大程度上体现出图像的内容和相关材料。因此多特征融合的自动标注技术应当在图像检索方面发挥其作用与优势。

**Table 1.** Using single feature automatic annotation  
**表 1.** 使用单一特征自动标注

特征	查准率	查全率	标注正确的关键词
RGB	10.06%	12.37%	87
HSV	14.41%	16.13%	97
OPP	9.58%	11.37%	82
Gabor	6.47%	7.73%	76
Wavelet	1.51%	2.23%	33

**Table 2.** Using a combination of two features for automatic labeling  
**表 2.** 使用两种特征相结合进行自动标注

特征	查准率	查全率	标注正确的关键词
RGB + Gabor	11.86%	14.10%	94
OPP + Gabor	13.02%	16.22%	98
HSV + Gabor	16.86%	19.67%	118
HSV + Wavelet	8.44%	8.68%	79

**Table 3.** Using a combination of a variety of features for automatic labeling  
**表 3.** 使用多种特征相结合进行自动标注

特征	查准率	查全率	标注正确的关键词
RGB + HSV + Gabor	17.33%	19.36%	113
RGB + HSV + OPP + Gabor	17.42%	20.30%	119

**Table 4.** Manual and automatic annotation results marked contrast

**表 4.** 人工标注和自动标注结果对比

					
Predicted key words	Bear polar snow tundra mountain	Formation jet plane sky smoke	Clouds mountain storm sun sky	Beach palm people tree water	Boats coast harbor water sky
Human annotation	Bear polar snow tundra	Jet plane sky smoke	Clouds sky sun	Beach palm people tree	Ships sky water
					
Predicted key words	Iguana lizard marine water beach	Deer grass hills mule tree	Garden ground people mountain	Cat forest head tiger water	Buildings people shops street water
Human annotation	Iguana marine water	Deer field grass mule	Tree	Bengal cat forest tiger	Canal house people water

#### 4.4. 对比分析

在被海量图片充斥的今天，与目前相关领域的基于单特征的标注技术对比，本文所设计的这套完整的基于多特征的图像标注算法能在一定程度上提升标注的准确性和可靠度。且随着图像类别的增加，利用常规的一对一或一对多的标注算法进行训练和分类都会越来越困难，因此本文设计的算法能帮助避免此类困难。进行融合过的多个自动标注词有非常大的参考价值，能够从很大程度上体现出图像的内容和含义以及相关信息，因此多特征融合的自动标注技术应当在图像检索方面发挥其作用与优势。

#### 5. 结论

图像特征提取是图像识别的关键步骤，图像特征提取的效果优劣如何将直接决定着图像识别的效果[7]。如何从原始图像中提取具有较强表示能力的图像特征是图像自动标注处理的一个研究热点。

本文首先引入了系统框架的两个主要模块，即图像特征提取和图像自动标注的相关含义和方法。接着仔细阐释了本系统具体实现的流程，也就是根据图像识别常用的几种特征和多特征融合所使用的算法，提出和阐述了基于多特征融合的图像自动标注技术的原理。最后根据实验数据做出了相应的测试分析，并且通过和现有技术的对比，揭示了本系统的特点和存在的必要性。从结果来看，使用多融合特征来标注图像在检索图像方面的优势明显，后续研究是针对不同特征设置不同权重的特征融合方法。

#### 基金项目

北京市大学生科研训练项目“基于多特征融合的图像自动标注”项目(项目编号：14005)；北京市教委面上项目“基于嵌入式多核处理器的图像自动标注系统研究”(项目编号：KM201410009006)。

---

## 参考文献 (References)

- [1] 付玮, 曾接贤 (2007) 基于形状特征的图像检索技术研究. *计算机技术与发展*, **11**, 228-232.
- [2] 储鑫 (2006) 基于颜色特征的图像检索系统研究与实现. 中山大学, 广州.
- [3] 汪建, 方洪鹰 (2013) 粒度计算与图像匹配技术改进. *电脑知识与技术*, **33**, 7571-7574.
- [4] 姜兰池, 沈国强, 张国煊 (2009) 基于 HSV 分块颜色直方图的图像检索算法. *机电工程*, **11**, 54-57.
- [5] 张轩, 臧淼, 李金泉 (2014) 基于 Matlab 的图像自动标注. *现代电子技术*, **3**, 73-75.
- [6] University of California Berkeley. Corel Dataset Website. <http://calphotos.berkeley.edu>
- [7] 翟俊海, 赵文秀, 王熙照 (2009) 图像特征提取研究. *河北大学学报(自然科学版)*, **1**, 106-112.