Pattern Recognition 47 (2014) 1777–1789

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

An efficient color quantization based on generic roughness measure

X.D. Yue [a,c,*], D.Q. Miao [b], L.B. Cao [c], Q. Wu [d], Y.F. Chen [b]

[a] School of Computer Engineering and Science, Shanghai University, Room 919, Nanchen Road 333, Baoshan District, Shanghai 200444, China
[b] School of Electronics and Information Engineering, Tongji University, Shanghai 201804, China
[c] Advanced Analytics Institute, University of Technology, Sydney, NSW 2007, Australia
[d] School of Computing and Communications, University of Technology, Sydney, NSW 2007, Australia

ARTICLE INFO

Article history:
Received 4 August 2012
Received in revised form
6 October 2013
Accepted 15 November 2013
Available online 25 November 2013

Keywords:
Color quantization
Generic roughness measure
Thresholding
Clustering

ABSTRACT

Color quantization is a process to compress image color space while minimizing visual distortion. The quantization based on preclustering has low computational complexity but cannot guarantee quantization precision. The quantization based on postclustering can produce high quality quantization results. However, it has to traverse image pixels iteratively and suffers heavy computational burden. Its computational complexity was not reduced although the revised versions have improved the precision. In the work of color quantization, balancing quantization quality and quantization complexity is always a challenging point. In this paper, a two-stage quantization framework is proposed to achieve this balance. In the first stage, high-resolution color space is initially compressed to a condensed color space by thresholding roughness indices. Instead of linear compression, we propose generic roughness measure to generate the delicate segmentation of image color. In this way, it causes less distortion to the image. In the second stage, the initially compressed colors are further clustered to a palette using Weighted Rough *K*-means to obtain final quantization results. Our objective is to design a postclustering quantization strategy at the color space level rather than the pixel level. Applying the quantization in the precisely compressed color space, the computational cost is greatly reduced; meanwhile, the quantization quality is maintained. The substantial experimental results validate the high efficiency of the proposed quantization method, which produces high quality color quantization while possessing low computational complexity.

1. Introduction

In a high quality color image, there are millions of different colors. Color quantization is a process that reduces the number of distinct colors in a digital image, usually with the intention that the reconstructed image should be as visually similar as possible to the original image. By reducing the complexity of color space, color quantization will benefit image storage and image transfer on the internet. Moreover, color quantization also simplifies feature spaces, which is helpful for image recognition and retrieval.

A color quantization algorithm generally consists of two parts. The first is color palette design and the second part is pixel mapping. There are two kinds of methods for creating color palette: image-independent methods and image-dependent methods. Image-independent methods determine a generic palette without considering any specific image contents, while image-dependent methods generate palettes based on the color distribution of images. Although it is fast, the image-independent method often produces poor results because of not considering image contents. To maintain a quality of image representation, most of the recent works on color quantization rely on image-dependent methods. For image-dependent methods, there are two different strategies to build up color palette: preclustering and postclustering [1]. The preclustering strategy partitions the original image colors into multiple subspaces based on the statistics of color distribution [7,9,11,19,26,39,47,48]. Because palette construction is an once-off procedure, the preclustering strategy usually has low computational complexity while sacrificing quantization quality in some degree. The postclustering strategy starts color quantization with an initial palette and improves it iteratively [3,13,27,35,36,38,49]. Since this strategy involves an objective function to minimize color distortion through stochastic optimization, it has better quality than preclustering strategy. However, it greatly increases the computational complexity.

From the discussion above, the challenge to color image quantization is to balance the quantization quality and computational complexity. To achieve this balance, the postclustering

* Corresponding author. Permanent address: Room 604, Unit 1, No. 136, Benxi Road, Yangpu District, 200092, Shanghai, China. Tel.: +86 13585866176.
E-mail addresses: yswantfly@gmail.com, yswantfly@hotmail.com (X.D. Yue), miaoduoqian@163.com (D.Q. Miao), longbing.cao@uts.edu.au (L.B. Cao), qiang.wu@uts.edu.au (Q. Wu), april337@163.com (Y.F. Chen).

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.11.017

quantization methods have been improved in different ways. Puzicha et al. proposed a color quantization method incorporating spatial and contextual information, in which the quantization was performed by an efficient multi-scale procedure to alleviate the computational burden [34]. Mojsilović and Soljanin proposed another quantization approach based on Fibonacci numbers and spiral lattices, in which the sampling scheme was used to generate color palettes [23]. Computational intelligence theories such as neural networks [29], PSO [25], GA [37], SOM [4] and competitive learning [2,45] were also used to optimize the color quantization process. Zhou et al. proposed an algorithm to adjust the color quantization results which tuned the palette by assigning weights to pixel clusters and color distances [50]. For the postclustering quantization based on clustering strategies, the quantization efficiency was improved by reducing the computational cost of pixel clustering. The modifications focused on speeding up clustering process, in the meantime, optimizing the clustering initialization [3,12,27,42]. The above improvements of postclustering methods could produce the higher quality of quantization results and alleviate the computational burden to some degree. However, these methods needed to heavily traverse pixels iteratively thus computational complexity was still high.

To tackle the problem above, we propose a two-stage color quantization framework based on Generic Roughness measure, which is abbreviated as *GR* framework. The basic idea of this framework is to integrate the low complexity of the histogram-based color space segmentation and the high quality of the clustering-based color quantization. First, the original image color is initially compressed to a condensed color space through thresholding color components. It is very important to form precise segmentation of color space to avoid the severe color distortion in final quantization results. However, the traditional histogram-like statistics cannot guarantee the segmentation precision. We propose generic roughness measure for color segmentation in the initial compression stage. Generic roughness can represent the spatial color homogeneity and thus generate the delicate color segmentation results. In the second stage, the initially compressed colors are further merged to a palette using clustering methods. Carrying out clustering in a compressed color space rather than on the pixel level, merging color in the second stage is very fast. Thus the computational cost of the framework mainly depends on the roughness thresholding in the first stage. The efficiency of the framework is analyzed in Section 5.3. Meanwhile, because of the precise segmentation of color space through generic roughness, the proposed framework causes little color distortion in the initial compression stage and guarantees the quantization quality. Therefore, the proposed framework can well balance the quantization quality and computational cost. The contributions of this paper are summarized as follows:

- *Propose a postclustering quantization strategy at color space level*: Common postclustering quantization methods are implemented at the pixel level. The proposed strategy applies the postclustering quantization in a precisely compressed color space. At this level, the quantization efficiency is greatly improved. In the meantime, the quantization quality is maintained.
- *Propose a generic roughness measure for color space segmentation*: Generic roughness measure is the key to precise color space compression. It can tolerate the disturbance of imbalanced color distribution and thus produces the accurate segmentation of image color space.
- *Design an efficient two-stage quantization framework*: In the first stage, a self-adaptive algorithm is designed to threshold roughness on color components to compress color space. In the second stage, the Rough *K*-means algorithm [18,33] is modified by integrating the color weights to merge the compressed colors to obtain the final quantization result.

The rest of this paper is organized as follows: Section 2 reviews the related work and analyzes the existing problems. Section 3 describes the basic framework and workflow of the roughness-based quantization. Section 4 introduces the construction of generic roughness measure. Section 5 presents the specific quantization method which includes the roughness thresholding algorithm and color merging algorithm. In Section 6, the comprehensive experimental results validate the high efficiency of the proposed color quantization framework. The work is concluded in Section 7.

2. Related work

The clustering technique is a key component in postclustering color quantization. To accelerate the clustering for quantization, Celebi proposed an improved *K*-means algorithm which simplified distance calculation and comparison in the clustering procedure through sorting cluster means [3]. Using partition indices, Özdemir and Akarun proposed a variant of fuzzy *C*-means algorithm to reduce the computational cost of the quantization based on soft clustering [27]. The initialization strategies based on color histograms were also utilized to speed up the clustering-based quantization. Tan and Isa presented a strategy for locating initial cluster centers through thresholding the histograms on color components [42]. Hsieh and Fan constructed 3D color histograms to prepartition pixels into bins and then merged colors on the color histograms to obtain the final quantization results [12]. Although these methods reduced the computational cost to some extent, they still needed to traverse the pixel set iteratively and possessed high computational complexity.

Segmentation of color space [5,28] is inevitable to color quantization. Through color space segmentation, the image is partitioned into homogenous regions depending on the color properties of pixels and numerous original colors can be concisely represented using only a small set of compressed colors. Existing segmentation methods for color compression can be roughly classified into the approaches as histogram based [6,24], edge based [43,44], region based [20], clustering based [8], and combination of several techniques [15,42]. As a popular segmentation technique, histogram thresholding has the advantages of low computational complexity and no requirement of prior information. However, since the bins in a histogram only count the number of pixels with the same color, the histogram thresholding usually produces over rough segmentation results. To refine the color segmentation based on histogram thresholding, the traditional histogram was improved based on rough sets.

As a soft computing technique, rough set theory was widely used in the area of image analysis to handle the data uncertainty [10,30–32]. Mohabey and Ray proposed the statistics histon as an approximation of traditional histogram through checking the neighborhood color similarity [22]. Compared with histogram, the segmentation based on histon considered the spatial color correlation. However, it is not robust to the imbalanced color distribution. To tackle this shortage, Mushrif and Ray proposed a roughness measure utilizing the boundary between histogram and histon to represent the color homogeneity [24]. For an RGB image F of size $M \times N$, its basic roughness measure is defined as follows:

$$roughness_i(l) = 1 - \frac{|histogram_i(l)|}{|histon_i(l)|}, \quad 0 \le l \le L, \ i \in \{R, G, B\} \tag{1}$$