Optik 145 (2017) 346–364

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Improved shape matching and retrieval using robust histograms of spatially distributed points and angular radial transform

Pooja Sharma

Department of Computer Science and Applications, DAV University, Jalandhar, 144012, Punjab, India

ARTICLE INFO

Article history:
Received 9 June 2016
Received in revised form 26 April 2017
Accepted 30 April 2017

Keywords:
Histograms of spatially distributed points (HSDP)
Angular radial transform (ART)
Shape matching
Precision and recall

ABSTRACT

In this paper, the problem of shape based image retrieval is addressed by proposing a hybrid shape descriptor. The proposed descriptor conforms to human visual perception along with its low computational complexity. Since global features are related to the holistic characteristics of images, whereas local features describe the finer details within objects of images, in the proposed hybrid descriptor both global and local features of images are used to describe the entire aspects of image shape. For global features extraction, we use angular radial transform, which is also adopted by MPEG-7 as a region based shape descriptor. On the other hand, for local feature extraction, a novel local descriptor is proposed, which is referred to as histograms of spatially distributed points (HSDP). It is based on two components: radial distance and differential coefficient, which are used to build 2D histograms. Global and local features are combined using effective distance measures viz. Min-Max and Bray-Curtis. Their superiority is validated by experimental results. Apart from that, an extensive range of image databases is employed to assess the performance of the proposed hybrid descriptor. These databases represent several characteristics of shape such as partial occlusion, distortion, subject change, gray scale objects, rotated and noise affected objects, unstructured images, trademarks, blurred images, Corel images, etc. The results of wide range of experiments reveal that the fusion of ART and HSDP significantly improves the image retrieval accuracy and provides a robust and invariant solution for effective shape matching.

1. Introduction

Effective image retrieval from large databases is a challenging task, which has not been completely solved yet. Relevance based image retrieval is applied to many applications such as medical imaging, trademark matching, digital library, computer aided design, military services, crime prevention, architectural and engineering design, geographical information, etc. Traditional image retrieval systems [1,2] were based on featuring the original data such as file name, note title, keywords, and indexing icon. However, textual annotations of images using keywords require intensive manual labor. When applied to large scale databases; these textual features become troublesome and time consuming. Besides; this method inadequately describes the image content; which does not meet the human visual perception exactly. In order to address these drawbacks; content based image retrieval (CBIR) systems have been proposed [3–9]. In CBIR; an image is generally represented by a set

E-mail addresses: sharma_pooja@live.com, pooja10013@davuniversity.org

http://dx.doi.org/10.1016/j.ijleo.2017.04.102
0030-4026/

of features; where the feature vector is a point in a multidimensional feature space. Each feature tries to capture only one property of image; such as color; texture; or shape. Among other low level properties color and texture; shape is the most significant one because human recognize objects solely from their shapes. The discrepancy between actual information and its representation using computed features is known as semantic gap. Various shape based representation and description techniques have been proposed to bridge this semantic gap. These techniques can be classified into two categories: contour based and region based.

Contour based methods include: Fourier descriptor (FD) [10], curvature scale space (CSS) [11], contour point distribution histograms (CPDH) [8], elastic matching [12], contour flexibility [13], local binary patterns (LBP) [14,15], local ternary patterns (LTP) [16], Weber's local descriptor [17], etc. Among them FD, CSS, CPDH are discrete descriptors while LBP, LTP, and WLD are dense descriptors. The discrete descriptors perform calculation on the interest points or contour points of an image, while dense descriptors extract local features pixel by pixel on the input image [22,23]. The contour based descriptors provide local characteristics of the image shape. Therefore, they are also termed as local descriptors. As per Zhang and Lu [10], FD is found to perform better than CSS because CSS is not suitable for efficient indexing due to expensive matching and variable feature dimensions. Contour flexibility represents the deformable potential at each point along a contour. CPDH is based on the distribution of points on the object contour under polar co-ordinates. CPDH fits for shapes with closed contours and does not deal with images having multiple connected regions, such as trademark images. LBP has gained increasing attention due to its simplicity and excellent performance in various texture and face image analysis tasks [14]. Many variants of LBP have been proposed recently and have achieved considerable success in various tasks. Tan and Triggs [16] have changed the thresholding of LBP and propose LTP [16], which performs better than LBP in various conditions among them robustness to noise is predominant. WLD [17] is based on Weber's law, which states that the change of a stimulus (such as sound, lighting) that will be just noticeable is a constant ratio of the original stimulus. WLD performs better on texture images, other recent descriptors include robust histogram based descriptor [18], bioinformatics based approach [19], image to class similarity [20], adaptive local binary patterns [21]. On the other hand, region based descriptors include: moment invariants (MI) [24], angular radial transform (ART) [25], grid descriptor [26], generic Fourier descriptor [27], Zernike moment descriptor [28], etc. These descriptors consider the complete image content for its computation, thus, they are termed as global descriptors. In GFD, the unstable polar mapping makes it unsuitable for noise and other variant transformations. The grid descriptor is also prone to noise and other transformations due to its major axis normalization. In addition, the region based descriptors are computationally complex as compared to contour based descriptors. Other types of the region based descriptors include distribution based descriptors in which histograms are used to represent various characteristics of the shape. The features acquired using these descriptors belong to an interest point (keypoint) or an interest region. The distribution based methods include scale invariant feature transform [29], histograms of oriented gradients [30], speeded up robust features [31], gradient location and orientation histogram [28], etc. However, some limitations exist with these techniques also. Distribution based descriptors are computation intensive and produce very high dimensionality of features [32].

Recently, it has been observed that hybrid approaches provide superior results rather than using region (global) and contour (local) based descriptors autonomously [33–36]. In hybrid concept, complimentary features from different modalities are combined to obtain higher recognition rates. In [33], Jain and Vailaya propose a hybrid approach in which local features of images are obtained by using histograms of edge directions. On the other hand, global features are extracted using Hu's seven moment invariants. Apart from that, a weight based strategy is followed to combine both acquired local and global features. Although features acquired using edge directions provide good performance, merely seven moment invariants are not adequate enough to provide complete global essence of an image. Wei et al. [34], propose another hybrid approach in which global features are obtained using first four low order Zernike moments (ZMs) and for local features contour curvature and histograms of centroid distances (HCD) are used. Nevertheless, contour curvature requires second order derivative, which is sensitive to image noise. In addition, these local features do not provide relationship among adjacent contour points. Besides, merely four low order ZMs do not provide adequacy in image description and representation [3]. Inspired by the hybrid approach proposed by [34], Qi et al. [35] propose another hybrid approach. The global features are extracted using spatial distribution of feature points. Local features are obtained by using relationship among two adjacent boundary points and the centroid (RAPC) and HCD as proposed by Wei et al. [34]. However, it has been observed that the histograms of normalized radii do not provide satisfactory results. Because in certain situations, the radius of the circumscribed circle passing through two adjacent boundary points and the centroid becomes quite large, which results in incorrect formation of histogram, which considerably affects the retrieval accuracy [3]. Another recent hybrid approach proposed by Wang et al. [36], considers both shape and texture features of images for image retrieval. In this approach, the shape features are extracted using exponent moment descriptors and texture features are extracted using localized angular phase histogram.

Motivated by hybrid approaches, in this paper, we propose another hybrid approach, which makes use of both global and local features. In complex applications like image retrieval, it is observed that one kind of feature set is not rich enough to capture the entire object information within the image. Thus, finding and combining the complementary feature sets have become an active research subject in recent years. Specifically, global features are related to the holistic characteristics of images, whereas local features describe the finer details within object images. Hence, it seems logical to combine both of these feature sets since the information conveyed by them belongs to different attributes of the object in an image. Therefore, for describing the complete essence of an image both global and local features of images are necessary. In our proposed solution, for global feature extraction, we use ART which provides superior retrieval accuracy with extra advantage of being quite efficient in its computation. The radial kernel function of ART is simple to compute because it comprises of