Digital Signal Processing 79 (2018) 125–135

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Tetrakis square tiling-based triangulated feature descriptor aiding shape retrieval

S. Priyanka, M.S. Sudhakar *

School of Electronics Engineering, VIT, Vellore, Tamil Nadu, India

ARTICLE INFO

Article history:
Available online 3 May 2018

Keywords:
Bulls-eye retrieval
k-nearest neighbor
Law-of-sines
Tetrakis square tiling
Triangulated feature descriptor

ABSTRACT

Ongoing research addressing shape retrieval render highly acute feature descriptors that are computationally intensive. Hence, a simple approach through a novel tessellated version of the Tetrakis square tiling for acute feature characterization aiding shape retrieval is contributed in this paper. The proposed descriptor, named Triangulated Feature Descriptor (TFD), performs feature characterization and abstraction by fusing hybrid geometrical concepts. The mechanism initially tiles the image into square regions that are later tessellated into right-angled triangles. The right-angled triangular neighbors interact locally using the trigonometric identities to produce an angle-based feature map representing an image. The attained feature maps are locally segmented to finally produce a shape histogram. Since, the intended descriptor is modeled based on intrinsic angular interactions between the local neighbors the resulting descriptor is highly invariant to diverse affine transformations. This characteristic is examined by performing rigorous experiments on MPEG-7, TARI-1000, Articulated and PHOS datasets. Competitive analysis reveals a consistent retrieval rate greater than 85% achieved by TFD with increased performance scores.

1. Introduction

With the increased utilization of image data across diversified domains, there has been a drastic surge in archived images. This demands for efficient data management and retrieval systems [1,2]. Also, the ever growing image databases make the computational complexity prohibitively expensive and further demands when extended to real-time faster retrieval. For large image databases, annotation based image retrieval is inefficient as it requires intensive human labor. Thus, Content Based Image Retrieval (CBIR) evolved with the intention of addressing the aforesaid concerns. CBIR performs image abstraction using color, texture and shape characteristics to compose feature vectors that represent the images [3]. This ad-hoc combination is suitable for generic image retrieval systems where the objective is to retrieve images that match the given query. However, in areas that demand human perception for discriminating the structural information of the given image, shape feature stakes claim for subsequent image analysis. This inspired the development of shape-based CBIR systems that had the potential of providing fine distinction among the image classes owing to the higher intra and inter class variations.

Shape remains the key component of human perception that aids in perceiving scenes as being composed of individual objects. From the retrieval point, shape is easy to describe, either by giving example or by sketching. The main constituents of shape retrieval are shape description and representation, shape similarity and shape indexing. Particularly, shape description and representation remain a key issue of concern in shape retrieval [4].

The majority of the accessible literature builds shape descriptors by either extracting shape features from contours or from the whole shape region. Accordingly, they were classified as contour-based and region-based shape descriptors. These descriptors are further categorized into global and local based on their representation as a whole or segment/section. Especially, contour based methods such as Fourier descriptor [5], curvature scale space [6], wavelet descriptor [7,8], shape signature [9], UNL-Fourier transform [10], modified UNL algorithm [11], point distance histogram [12] produced descriptors under different domains for shape characterization. In particular global approaches such as polygonal multi-resolution and elastic matching (PMEM) [13], contour points distribution histogram (CPDH) [14], earth mover's distance (EMD) [15] compact the shape information on a global map into a feature histogram that is intrinsically invariant to translation. The scope of these global schemes is limited by the shape characterization that considers the overall nature of the data. Whilst, neglecting the lo-

* Corresponding author.
E-mail address: sudhakar.ms@vit.ac.in (M.S. Sudhakar).

https://doi.org/10.1016/j.dsp.2018.04.012
1051-2004/

cal information pertaining to edges, corners, points that tend to formulate the shape information.

In contrary, local approaches retain these high frequency features to enhance feature characterization. Local shape descriptors such as the shape context (SC) [16] builds histograms from localized landmark points using the Euclidean metric. To achieve more robustness and to improve the retrieval performance, inner distance shape context (IDSC) [17] was presented. IDSC replaced the Euclidean with the inner distance to build a more accurate descriptor for articulated shapes. Although, effective IDSC suffers majorly in the matching phase that demands higher local matching operations for bounding shapes. This issue was resolved by dynamic programming (DP) [18] that also, escalated the retrieval accuracy. Further retrieval improvement was attained by contour flexibility [19] using the deformable potential of points lying on closed curves. The above methods exploit the two point spatial relationship among the landmark points to construct shape descriptors that demand higher computations.

Recent shape representation methods [20–22] employ stable three point spatial positional relationship based on the points that fall under the triangular region. The multi-scale triangle area representation (MTAR) [20] constructs feature from the zero crossings of wavelet coefficients confined in a triangular region across different wavelet scales. This method bettered noise immunity, but demanded more computations due to the higher wavelet decompositions. MTAR was spatially extended to triangle area representation (TAR) [21] that derived shapes from fixed common base triangles with varying base lengths. This aided for multi-scale representation, but introduced higher redundancy that made feature extraction difficult and was highly prone to noise. Similarly, the Delaunay triangulation method [22] decomposed the shape contours into small triangles using pre-defined feature triangles, upon which shape trees were constructed. This method was robust to noise but the processing stages involved in shape tree formulation are computationally very expensive. The height function [23] shape description method constructed descriptors based on the distance of sampling points to its tangent line. Feature matching is then attained using dynamic programming which is again an expensive process. Likewise, the hierarchical procrustes matching (HPM) [24] captures shape information among different closed contours that are hierarchically ordered using the shape tree technique [25]. Retrieval accuracy was improved by combining various distance measures in the locally constrained mixed-diffusion [26].

Dual shape descriptors aiding retrieval coined as angular pattern (AP) and binary angular pattern (BAP) [27] captured the angular information among contour points that were invariant to similarity transformation, scaling and rotation. Multi-scale integration of AP and BAP improved the retrieval accuracy with increased computations. The recently introduced hierarchical string cuts (HSC) [28] realized shape descriptors by multi-level segmentation of contours. This arrangement ensured that global information and fine features are packed in the resulting descriptor that were invariant to mirror, scale and rotation. Curvature partition (CP) [29] addressed occlusions and geometric distortions using normalized Euclidean invariant angle length profiles. Formulating CP involves higher space and time complexities owing to the template matching process. To reduce contour deformations in noisy environments, common base triangle area (CBTA) [30] was introduced. CBTA builds features from common base triangles with varying vertices followed by dynamic programming for shape matching. The above methods preserve the exterior shape information while neglecting the interior information that may further escalate the retrieval accuracy. Moreover, the computational load places a higher burden on the feature extraction.

This novel method is proposed based on the identification that interior property of shape is important in feature characterization

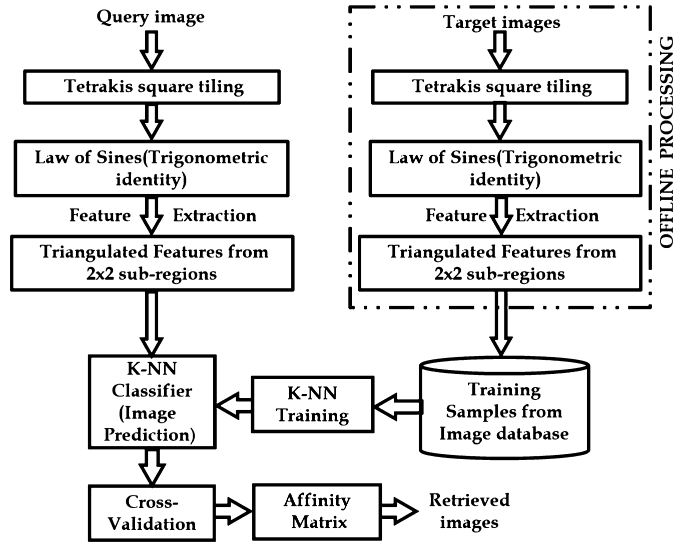

Fig. 1. Outline of the intended shape retrieval scheme.

for efficient retrieval and higher classification accuracy. Hence, the Tetrakis square tiling-based triangulation method extracts the local pixel transitions and then build rich shape descriptors using congruent triangle patterns that makes the method to be naturally invariant to illumination, translation, scaling and rotation.

The remainder of this paper is ordered as follows: Section 1 chronologically lists the available shape descriptors with their pros and cons. Section 2 discusses the formulation of triangulated feature descriptor and its representation. Section 3 covers the mathematical proofs of intrinsic invariance of TFD. In section 4, detailed performance analysis of TFD on diverse datasets is performed. Relative computational complexity analysis in terms of time and space is presented in Section 5. Lastly, the proposed approach is briefly summarized in Section 6.

2. Tetrakis square tiling-based triangulated feature descriptor

Efficiency of retrieval systems heavily relies on feature extraction and characterization for delivering descriptors with accurate discrimination abilities. Shape retrieval schemes producing optimal performance without incurring higher computational cost pose a challenging issue to researchers. Hence, a simple, novel and an acute feature descriptor capturing the coarse, fine features with minimal computational load is the main contribution in the proposed scheme. In this approach every image in the database is considered as a tiled square, from which the triangles are tessellated across each sub-regions to subsequently extract localized features. The main idea behind triangulation is the enhancement of high frequency information that is essential in characterizing shapes.

Prior to retrieval, TFD features are extracted and accumulated into shape histograms. This procedure is repeated for every image in the database offline and the corresponding descriptors are attained. The stored features are spatially classified by the k-Nearest Neighbor algorithm. In the online phase, an input image queries the retrieval framework from which the respective TFD is extracted and compared against the categorized data to identify the relevant matches. Based on the ordered similarity index the images that match with the given query are retrieved and displayed as output at the retrieval window. An overview of the aforesaid process is depicted in Fig. 1.

The concept of tetrakis square tiling is extended to image processing wherein each image is tiled into squares and the neighbors in the local vicinity are connected as four right-angled triangles in