Extracted Structural Features for Image Comparison

Pornchai Mongkolnam[a], Thanee Dechsakulthorn and Chakarida Nukoolkit[a]
[a]School of Information Technology
[a]King Mongkut's University of Technology Thonburi
Bangkok, Thailand

***Abstract*- We present a method that extracts structural features of images. The method is based on both a region-based analysis and a contour-based analysis. The image is first segmented, based on its pixels' information. Color information of each segmented region is performed by using the hue-saturation-value color space. Area of each region is also extracted by counting the number of bound pixels. Location of each region is computed as a center of the region's convex hull. A contour of the region is approximated by a B-spline approximation to obtain its control polygon and curve in the limit. The region's convex hull is obtained from the control polygon. For multi-scale features, we apply Chaikin's algorithm to the control polygon for finer level of control polygons, which could be used in a coarse to fine comparison. Curvature information of the B-spline curve fitting could also be used in the comparison. Our method could be used in many interesting applications including image retrieval, image classification, image clustering, image manipulation, image understanding, pattern recognition, and machine vision.**

I. Introduction

There is an ever-increasing need for a simple, yet effective and efficient way to analyze, retrieve, cluster, explore, and display digital mages and videos. Similar need for a large document collection is also desired [1, 2]. The most popular image search Web sites such as Yahoo! and Google are the irrefutable evidence. To improve the search results various techniques have been invented in order to incorporate other relevant features such as shape, color, and texture into a mere text-based image search. Keyword-only search has some drawbacks in that keywords are context dependent and do not allow for unanticipated search. Furthermore, language barriers and lacks of uniform textual descriptions make them ineffective.

Several existing work uses B-spline curves to represent profile shapes of 3D objects [3, 4]. A contour of the object is first extracted and then approximated by the B-spline curve, which is in turn used for curve matching in a data retrieval application. The object matching is an integral part for many applications of shape modeling, machine vision, and image processing. The B-spline curves and their curvatures are widely and effectively used for curve representation of the object contours instead of the far higher degree Bezier curves because they possess some very attractive properties such as smoothness, compactness, local shape controllability, and affine transformation invariance. In addition to using B-spline approximation, the Chaikin's algorithm [5] is used to refine the matching curves at many different scales in this work. For images, little work has been done on applying the B-spline concept to structurally represent the images' components. Some work merely allows user to sketch an outline as well as specify color information before submitting the query to the search engine [6]. Therefore, this work has extended the already existing 3D object's curve concepts to 2D images in order to help represent structural shapes or component-level contents within the images. Knowing the shapes, components and their spatial layout relationships would certainly yield a good comparison in image databases, comparatively to understanding a molecular structure of an element. Our work would definitely find a useful place in various important fields such as machine vision, pattern recognition, image retrieval, clustering, exploration, manipulation, understanding and visualization. A good overview of using shapes for content-based image retrieval (CBIR) can be found in [7, 8, 9].

The paper is organized as follows. First, we present overview and related work on B-splines, curve fitting or curve approximation, Chaikin's algorithm, image shape modeling, and content-based image retrieval. Thereafter our work and its results on extracting segmented regions' features are discussed. Last, a conclusion and future work on extending the image shape representation to image retrieval, clustering, exploration, and visualization is given.

II. Related work

This work has been built on some prior work in a 2D and 3D shape matching and curve representation that apply B-spline and its curvature to approximate the image or object's profile. Over the past thirty years work on shape has been an active research area and was mainly driven by object recognition. One of the recent work [4] proposes a novel 2D shape-matching algorithm based on the B-spline curves and its Curvature Scale Space (CSS) image. The CSS image [10] is robust with respect to noise and affine transformation and is chosen as a shape representation for a digital planar curve. The representation is computed by convolving the curve with a Gaussian function at different scaling levels. The CSS is suitable because the B-splines have advantages of being continuous curve representation and affine invariant. The algorithm first smoothens the B-spline curve of an input shape and constructs the CSS image. It then extracts the maxima of CSS image and performs matching.

Due to the B-splines' attractive properties, Reference [3] chooses the B-splines for curve modeling in matching 2D objects such as aircrafts and handwriting over other approaches such as the Fourier descriptors, the polygonal approximation, the medial axis transform, the moments, and

T. Sobh (ed.), Innovations and Advanced Techniques in Computer and Information Sciences and Engineering, 13–18.

the curvature invariant. Their algorithm attempts to match and classify planar curves that are modeled as B-splines, independent of any affine transformations. Two methods are presented. First, the control points of the prototype curves are globally related to the knot points of a given sample curve and then are compared. Second, a sum of the residual error between each prototype curve and the given sample curve is compared.

A new image database retrieval method based on shape information and deformable template matching process is proposed by using the following two features to represent shape of an image: a histogram of the edge directions and the invariant moments [11]. Euclidean distance between the edge direction histograms is used as a matching score. The shape of the image is also represented in terms of the seven second-order and third order invariant moments. A region-based image retrieval method that employs a new image segmentation technique using circular filters based on Bayes' theorem and image texture distribution is proposed by Reference [12]. After segmentation, extracted features of each region including color, texture, normalized area, shape, and location are recorded and compared against other images.

Reference [13] proposes shape retrieval from image databases, which are composed of boundary contours; the method is based on indexing structural features in terms of convex/concave parts and quantized directional features along the contours. The method exploits the feature transformation rules, which is obtained by an analysis of some particular types of shape deformations, to generate features that could be extracted from deformed patterns. Most work previously mentioned is performed on a single scale shape analysis, which is not able to provide a robust representation. A multi-scale analysis for shapes [14] is used to derive a hierarchical shape representation in that the shape details are progressively screened out whereas the shape characterizing elements are preserved. Using the graph structures representing shape parts at different scales, the coarse-to-fine matching could be performed.

Besides using a curve matching, a shape matching can also be achieved by matching skeletal graphs (medial axis graphs) [15, 16]. The medial axis has been used for matching shapes because outline curves do not meaningfully represent the interior of the shapes. The shock graph, which is the medial axis endowed with geometric and dynamics information, is used because it gives a richer description of shapes. In summary, there are various existing techniques being used on shapes in image processing as shown in Fig. 1. A complete overview could be found in [17].

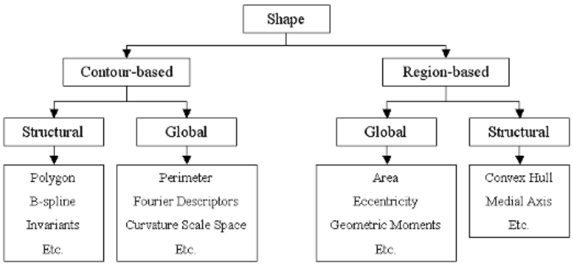

Fig. 1. Classification of shape representation and descriptor techniques.

III. OUR METHOD

Our method begins with image segmentation in order to globally identify structural components of the image by applying the JSEG algorithm [18, 19]. It involves two independent steps: color quantization and spatial segmentation. First, the image pixel colors are quantized to several classes called class maps, which are used to differentiate regions in the image. Second, a region growing method is used to segment the image based on the property of the obtained class maps. A good segmentation method would certainly help obtain good image contours and segments along with their relationships, which significantly impact on quality of subsequent work in image retrieval, clustering, and exploration.

After the segmentation, a boundary of each extracted segment is approximated by a B-spline curve. In general, a B-spline curve [20, 21] is more widely used to represent a complex curve than its far higher degree Bezier curve counterpart because of its local control property and ability to interpolate or approximate a curve with lower degree. The B-spline curve is a generalization of the Bezier curve and has more desired properties than the Bezier curves. The B-spline curve is generated from its control points and contained in the convex hull of its control polyline or polygon. An affine transformation such as rotation, translation, and scaling can be applied to the B-spline control points quite easily instead of to the curve itself. This results in an affine invariance property, where manipulation can be done to the control points instead of to the curve itself. Therefore, speed could be improved when a curve matching is done.

The B-spline curve, *C(u)*, is defined as:

$$C(u) = \sum_{i=0}^{h} N_{i,p}(u) P_i$$

where P_i is a control point, p is a degree, u is parameter, and $N_{i,p}$ is a B-spline basis function and is defined as:

$$N_{i,0}(u) = \begin{cases} 1 & \text{if } u_i \le u < u_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

$$N_{i,p}(u) = \frac{u - u_i}{u_{i+p} - u_i} N_{i,p-1}(u) + \frac{u_{i+p+1} - u}{u_{i+p+1} - u_{i+1}} N_{i+1,p-1}(u)$$

where u_i is known as a knot, a corresponding place where two curve segments join with certain continuity. Fig. 2 illustrates a cubic B-spline curve generated from its eight control points. The control points are shown in dark circles. Lines connecting control points are a polyline, and they capture the overall shape of the curve.