Shape-Based Image Retrieval For Thematic Database

Kim Meng Liang
Research & Development
APIIT, Malaysia
liangkm@apiit.edu.my

Rajeswari Mandava
School of Computer Science
Universiti Sains Malaysia, Malaysia
mandava@cs.usm.my

Bee Ee Khoo
School of Electrical Engineering
Universiti Sains Malaysia, Malaysia
beekhoo@eng.usm.my

Abstract

Large image databases are used in many multimedia applications such as entertainment, business, art, engineering and science. Searching information in these databases is a crucial problem to be solved for the development of visual information system. Therefore efficient retrieval methods are required for the purpose of finding a desired image from a collection of images. This research is focused on developing a shape-based image retrieval system for use in thematic databases. For this purpose, a novel shape similarity matching methodology is needed in order to detect a small similarity difference between two images. In this paper, a new shape transformation approach, referred to as NURBS-warping approach is implemented. In the proposed approach, Non Uniform Rational B-Spline (NURBS) and Gradient Vector Flow (GVF) are incorporated to ensure accurate and fast retrieval results. NURBS is a compact and accurate shape descriptor, whereas GVF ensure fast and accurate matching results. The effectiveness of this approach is examined by carrying out experiments on a collection of 1100 similar images from a thematic fish database. The overall retrieval results show that the proposed approach is able to derive an accurate similarity measure that resembles human similarity judgement.

Keywords

Image retrieval, similarity matching, thematic database.

INTRODUCTION

With the availability of advance computer vision research in classification techniques, it has become a common practice to group images that are highly similar under the same database known as thematic database. Examples of the thematic databases are fish, industrial tools, aircrafts, human faces, vases and other industrial products. In these types of databases, there exists small difference in the shape of the objects between images. This complicates the tasks of searching and browsing through these databases. Therefore accurate shape similarity matching techniques, which may place heavy demands on the computation resources, including time are needed.

In the past decade, shape similarity matching techniques has evolved and is influenced by two general approaches: feature vector approach and shape transformation approach. Feature vector approach is attractive for use in the large image database retrieval due to its simplicity and short retrieval time. It is most suitable to be adopted in the databases that contain images that have geometric shapes that are not highly similar to each other. Even though the speed of the feature vector approach is high, it does not provide sufficient accuracy in the similarity matching process and ability in resembling human similarity judgement as demanded in thematic databases.

Shape transformation approach stands as one of the most reasonable approach for this purpose. It not only has a high accuracy in similarity matching process but also closely resembles human similarity judgement, making it attractive for use in thematic databases [3]. One of the main drawbacks in most of the methods in this approach is that the retrieval time is longer as compared to the feature vector approach. It is due to the slow convergence of the transformation technique in transforming the shape of the query image towards the shape of the database image, and long computation time of the similarity measure technique in measuring the effort spent in the transformation process.

The transformation technique is evaluated based on three components: representation of the deformable template, transformation process and optimization technique [1]. In representation of the deformable template, an accurate and compact shape representation method is needed in order to ensure that only a small amount of features are adjusted in the transformation process. Various representation methods are introduced for this purpose [6, 7, 12]. Although these methods may provide compact representation, they do not ensure accuracy of the information preservation. An alternative way is to use B-Spline [4,5]. B-Spline is an accurate and compact shape descriptor.

In transformation process, the deformable template is transformed under the constraint of two energy fields: internal and external energies. The internal energy derived from the deformable template and defined in terms of first and second derivatives, which their constraint parameters are responsible in controlling the continuity and smoothness of the deformed template, respectively. There is a problem in finding the suitable constraint parameters for each target object, adaptively [8]. The external energy derived from the edge map of the target image. There are two general problems that are associated with the external energy computation: initialization and ability to move into boundary concavities. Several methods have been proposed to address these problems [6,7]. However, most of the methods still do not solve these two problems, completely. Xu and Prince [14] introduce an important direction for external energy formulation, which is referred as Gradient Vector Flow (GVF). Particular advantages of the GVF are: insensitivity

0-7803-8243-9/04/$17.00 © 2004 IEEE

to initialization and ability to move into boundary concavities.

The transformation process is optimized by minimizing the computed energy function. Various optimization techniques have been suggested such as variational framework [7], dynamic programming [2] and greedy optimization [13]. Although all the above techniques may cause the transformation process result to be trapped in the local minima, these techniques are widely used because of their simplicity and low computation time.

An equally important factor that influences the computed retrieval results from the shape transformation approach is the similarity measure technique. Sclaroff and Pentland [12] have used modal vibrations to identify the amount of deformation for both rigid and non-rigid needed to align the query image to the database image. The mode amplitude of the modal vibrations indicates the amount of energy required to align two images and thus provide a similarity measure. Bimbo and Pala [4] address the elastic template matching using user sketch. In this, internal energies as well as degree of matching achieved are used to evaluate the similarity ranking of the database images. The similarity measure is computed through back-propagation neural network training.

Although Bimbo and Pala's work utilizes B-Spline as a shape descriptor for the query image, this method still suffers from several drawbacks. First, the edges of the database image are spatially smoothed in order to increase the capture range. However this blurring effect may results the query image being unable to accurately transform to the database image. Second, there is a problem in finding an appropriate selection of the constraint parameters of the internal energy. Third, the B-Spline control points are implicitly adjusted, as the new positions of the control points are determined by recalculating the B-Spline formulation at each iteration of the optimization. This consumes extra time in the transformation process. Fourth, extra time is also needed in the back-propagation neural network training.

This paper aims to extend the work by Bimbo and Pala [4], with an intention to improve the effectiveness and efficiency of the shape transformation approach based on the transformation and similarity measure techniques. In NURBS-warping approach, NURBS is chosen to describe the query image. NURBS not only inherits all the advantages of the B-Spline but also has a better accuracy and compactness. GVF is selected as the external energy to provide a fast and accurate convergence results of the transformation. A new similarity measure metric based on the local properties of the NURBS control points is used.

This paper is organized as follows. In the next section, the overview of our proposed approach that is adopted in the shape-based image retrieval system is described. This section also introduces the method to determine the similarity parameters, degree of matching and similarity measure. The following section presents the results of the experiments conducted in order to evaluate the effectiveness of the proposed approach. This section also compares the retrieval results with the results given by the Curvature Scale Space (CSS) approach [10]. The subsequent section is devoted to the overall discussion based on the advantages of NURBS and GVF towards the proposed approach. The final section concludes the approach presented in this paper and discusses the future directions of our research.

THE PROPOSED APPROACH

The proposed shape-based image retrieval system based on the NURBS-warping approach is carried out in two phases: offline process and online process. These are illustrated in Figure1 and Figure 2, respectively.

In Figure 1, the GVF field is generated from the edge map for all the database images and all these are kept in the GVF repository. The images selected for this work are solely silhouette image, therefore their edge map only consists of the boundary points of the objects. This process is done offline. The details of the GVF generation are given in [14].

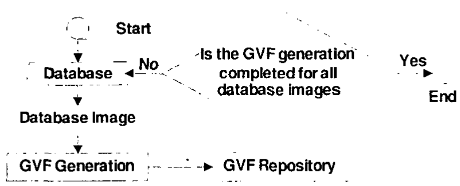

Figure 1. The overview of the offline process, showing the GVF generation for all the database images.

Figure 5.2 illustrates the online process of the proposed system, starting from the input of a query image that is requested by the user until the display of a set of relevant images from the database. The online process consists of the steps described in the following.

Step 1: NURBS parameters derivation

The boundary of the query image is described by the NURBS parameters: control points and their corresponding weights. The details about the derivation of the NURBS parameters are given in [9].

Step 2: NURBS parameters normalization

The query image is normalized with a transformation matrix that is determined based on the translation, scaling, rotation and reflection of the respective database image. The process of normalization is required to ensure that the similarity measure is solely determined from the transformation process without being affected by the orientation differences between the query and database images. If this process of normalization changes the orientation of the query image, the derived NURBS parameters in Step 1 do not corresponded to the new normalized query image, anymore. Owing to the characteristic of affine invariance of NURBS representation, NURBS parameters for the normalized image are automatically calculated by applying the same normalization to the initial NURBS parameters [11].

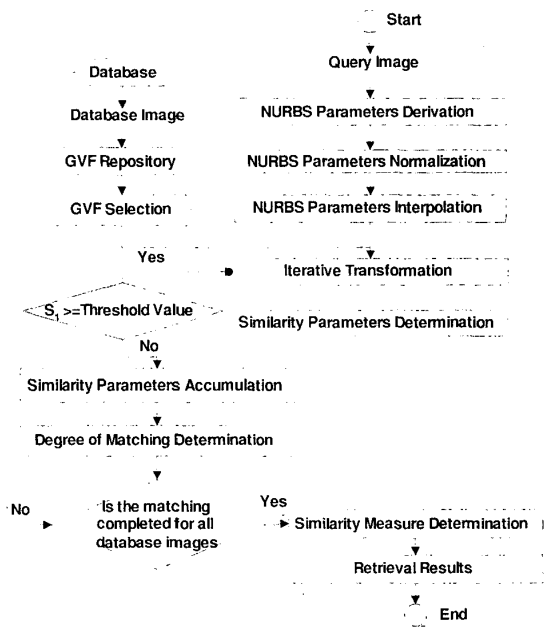

Figure 2. The overview of the online process, showing the adoption of the NURBS-warping approach in the similarity matching process.

Step 3: NURBS parameters interpolation

This is performed using knot insertion in order to relocate the positions of the control points closer to the boundary of the query image. The control points are initially not located on the boundary itself [11]. In this case, the effect of adjusting the control points may not have direct correspondence with the transformation of the boundary of the query image. Thus, the relocation process is required, so that the task of adjusting the control points is similar to the process of adjusting the boundary points of the query image itself, to match with the database image. Knot insertion adjusts the original control points and introduces additional control points without changing the representative shape of the query image through interpolation. This is performed via a single processing step.

Step 4: Gradient Vector Flow (GVF) selection

The GVF field corresponding to the respective database image is selected from the GVF repository, to be used in the iterative transformation.

Step 5: Iterative transformation

Prior to the transformation, the interpolated NURBS parameters are superimposed on the selected GVF field. Even though both the control points and their corresponding weights may be adjusted to transform the shape of the object, this system chooses to vary only the positions of the control points while fixing the weights to ensure simple and fast transformation process. Thus the interpolated control points are adjusted under the attraction of GVF field vf, through an optimization technique. In the NURBS-warping approach, the internal energy is eliminated because the NURBS representation has the inherent capability in controlling the continuity and smoothness of the deformed query image [11]. Thus the adjustment of the positions of the control points is achieved through a simple gradient descent technique as follows.

$$B_t = B_{t-1} + vf(B_{t-1}) \quad (1)$$

where B_t is the positions of the control points at t^{th} iteration, B_{t-1} is the positions of the control points at $t-1^{th}$ iteration.

In GVF field, the attraction energy is stronger at the location that is further away from the boundary of the database image. The shape of the query image is iteratively transformed until the total change in the positions of the control points (S_1) at the current iteration is below a certain threshold value. This happens when all the transformed control points converge and may lock themselves on the boundary of the database image or may be close to it. In this work, the convergence of the query image towards all the database images is guaranteed due to the ability of the NURBS parameters in controlling the continuity and smoothness of the deformed query image, and also the strength of the GVF. It may be noted here that S_1 is one of the similarity parameters, computed at each iteration of the transformation.

Step 6: Similarity parameters determination

Three similarity parameters are determined to reflect the effort spent at each iteration until the completion of the iterative transformation as described in Step 5. The parameters are listed as follows.

- Total change in the positions of the control points (S_1).
- Total voting distance score (S_2).
- Total change in the curvatures at the control points (S_3).

Step 7: Similarity parameters accumulation

Upon completion of the iterative transformation, all the computed similarity parameters at each iteration are accumulated. The accumulated similarity parameters reflect the total effort spent throughout the transformation and they are listed as follows.

- Cumulative total change in the positions of the control points (CS_1).
- Cumulative total voting distance score (CS_2).
- Cumulative total change in the curvature at the control points (CS_3).

Step 8: Degree of matching determination

Upon completion of the transformation, a final parameter, the degree of matching is computed. This parameter measures the effectiveness of the transformation.

The Steps (2-8) are repeated for the remaining database images. Thus for each database image, the accumulated similarity parameters and degree of matching are computed.

Step 9: Similarity measure determination

A combination of the three accumulated similarity parameters is used to derive the overall similarity measure. This similarity measure is meaningful and accurate only if the query image effectively transforms towards the database image. Thus similarity measure is only computed for the selected database images if their related degree of matching is equal to or more than a desired value.

Step 10: Retrieval results

The database images are ranked based on the similarity measure, with highest rank allocated to the image with the highest similarity measure. In this work, a number of database image that are most similar to the query image are displayed to the user in the order of descending similarity. The user may select this number.

Similarity parameters and degree of matching computation

In this section, it may be noted that the steps to compute the similarity parameters: S_1, S_2 and S_3 are only shown for a single iteration. Similar steps are applied for the following iteration until the completion of the transformation. These similarity parameters are computed as follows.

S_1 is computed by cumulating the sum of distances between the positions of the control points of two subsequent iterations by using Euclidean distance function. S_2 is computed by cumulating all the voting score for each control point. For each control point, a voting score that is 1 or 0 is given if the Euclidean distance between the positions of the control points of two subsequent iterations is greater or smaller than a prespecified threshold value, respectively. For each query image, a new threshold value is computed prior to the iterative transformation in Step 5. In order to compute this value, the query image is reconstructed from the NURBS parameters and this reconstructed image is transformed to match with the original image. The largest distance between the positions of the control points of any two subsequent iterations throughout the transformation is computed as the threshold value. The computed threshold value is used to computes S_2 for all the database images.

S_3 is computed by cumulating the sum of differences between the curvature values at the control points of two subsequent iterations. The curvature value at the control point is computed by using the curvature approximation function, introduced by Williams and Shah [13].

Upon completion of the iterative transformation, the degree of matching is computed. The reconstructed boundary generated from the final transformed control points, is superimposed on the boundary of the database image. The overlapping points between these two boundaries are detected. Two types of matching features: F_1 and F_2 are calculated. Mathematically, the two matching features are derived as follows.

$$F_1 = \frac{\text{number of overlapping points between two boundaries}}{\text{number of boundary points of the database image}} * 100$$

$$F_2 = \frac{\text{number of overlapping points between two boundaries}}{\text{number of boundary points of the deformed query image}} * 100 \quad (2)$$

In this work, the degree of matching is determined by computing the average value of F_1 and F_2.

$$\text{degree of matching} = 0.5 * (F_1 + F_2) \quad (3)$$

Similarity measure computation

The final measure of similarity referred to as Similarity Measure *SM* is only computed for those database images with degree of matching higher than a threshold value. In this work, this value is chosen as 70. However this value may be higher if the user demands retrieval results with higher precision. For each selected database image, *SM* is computed by the weighted additions of the normalized accumulated similarity parameters $\overline{CS_1}$, $\overline{CS_2}$ and $\overline{CS_3}$ as follows.

$$SM = w_1 * \overline{CS_1} + w_2 * \overline{CS_2} + w_3 * \overline{CS_3} \quad (4)$$

where

w_1 is the similarity weight for $\overline{CS_1}$.

w_2 is the similarity weight for $\overline{CS_2}$.

w_3 is the similarity weight for $\overline{CS_3}$.

It may be noted that w_1, w_2 and w_3 are set in the range of $[0,1]$. The choice of the similarity weights depends on the importance of each type of similarity terms in determining the SM. Thus, the choice of the suitable similarity weights depends on the purpose of the retrieval system itself. In our work, all three types of accumulated similarity parameters are chosen to play an equal role in the *SM* determination. We choose $w_1 = w_2 = w_3 = 1$.

EXPERIMENTS

An experiment is carried out to validate the effectiveness of the NURBS-warping approach in the proposed shape-based image retrieval system. The validation is performed by comparing the retrieval results from this system with the Curvature Scale Space (CSS) [10] results. CSS is a multiscale approach that compares two sets of maxima curvature zero crossing and the sum of pairwise distances between corresponding pairs of maxima is defined as similarity measure. The following sections describe the methodology for the experiment, present the experimental results and discuss the results in validating the objective of the experiment.

Methodology

In this experiment, a collection of one thousand and one hundred highly similar images from a thematic fish database is used. Fifty images are chosen from this database, for use as the query image in this experiment. For each query image, the retrieval results are obtained using NURBS-warping approach. These results are then compared with the results obtained from the CSS approach.

Results and discussion

For discussion purposes, two sets of retrieval results are shown. The retrieval results for these two query images using the NURBS-Warping method are displayed in Figures 3(a) and 4(a). Figures 3(b) and 4(b) illustrate the retrieval results, using the CSS method. The image shown on top left corner of each figure is the query image and its *SM* is equals to 1. For the CSS approach, no similarity measures are recorded because their similarity process is different and cannot be directly compared with the computed similarity measure in this system. In each figure, the first 15 retrieved images, with decreasing order of similarity measure, are shown.

From these results, it is seen that both the methods are equally capable of retrieving relevant images from the database at the coarse level. The retrieval results in Figure 3(a) shows that all the first 15 retrieved database images are very similar to the query image. In this figure, the database images at the top 10 ranking are very similar to the query image. This may be seen from the retrieved results that having a straight long tail and a flat body. However the results illustrated in Figure 3(b) do not comply with the human judgement as those in Figure 3(a). This is shown by the images that have rank of 3, 6, 10 and 14. Those images have a curved tail.

Thus this maybe shown that the NURBS-warping approach outperforms CSS approach if the similarity between the database image and the query image are compared in detail. Comparison at fine level is an important property for two highly similar images, especially if the shape of the query image is complex. This is illustrated in through the results in Figure 4(a) and Figure 4(b). In Figure 4(a), the top five retrieved images have a very rough top fins that are similar to the query image, whereas the top five retrieval results in Figure 4(b) do not obviously have such feature. Therefore it may be concluded that the retrieval system proposed in this paper is able to perform coarse as well as fine similarity matching, simultaneously.

OVERALL DISCUSSION

Based on the experiment results, the proposed NURBS-warping approach as a new shape transformation approach is successfully measures the similarity between two highly similar images. The incorporation of NURBS and GVF into the proposed approach plays a vital role in making the retrieval system to be effective and efficient.

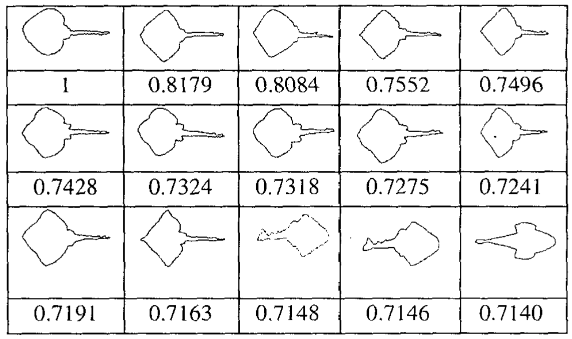

Figure 3(a): The search results of the first query image using NURBS-Warping approach.

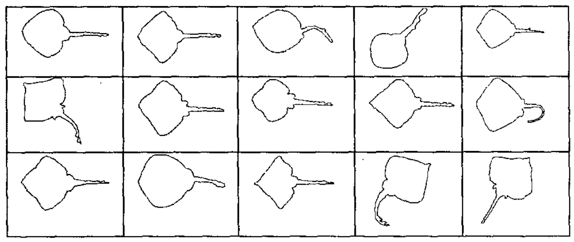

Figure 3(b): The search results of the first query image using CSS approach.

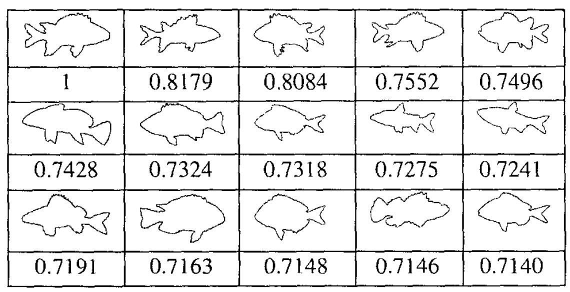

Figure 4(a): The search results of the second query image using NURBS-Warping approach.

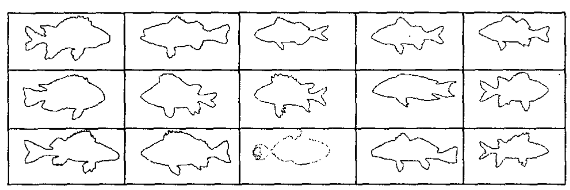

Figure 4(b): The search results of the second query image using CSS approach.

In the proposed approach, the shape of the query image is represented by the NURBS parameters. With the high compactness of NURBS representation, the number of control points that are adjusted during the transformation process is less than 25% of the total boundary points of the query image. The high accuracy of the NURBS parameters eliminates error in similarity measure determination caused by the inaccuracy of the shape descriptors.

In the transformation process, the internal energy is eliminated. This elimination avoids the problems in approximating the first and second derivative terms of the internal energy, involving high computation time. In addition, the use of NURBS eliminates the difficulty in determining constraint parameters that are related to the internal energy. GVF is a selected as a good choice of external energy in increasing the capture range of its energy fields and guides the shape of the query image towards the desired boundary of the database image. Because of the elimination of the internal energy and the use of GVF, the proposed NURBS-warping approach reduces the computation of the transformation process. In addition, the use of three accumulated similarity parameters in measuring the effort being spent in the transformation process further assist in reducing time consumed by shape similarity measuring technique. Therefore, intuitively, the proposed approach is more accurate and faster than the previous proposed shape transformation approaches.

CONCLUSION

This paper has proposed and validated an image database retrieval system using NURBS-warping approach specially suitable for thematic databases. Although the computation involved in the proposed approach is higher as compared to the feature vector approach, it has intuitively improved the speed of the retrieval in the previous shape transformation approaches. Thus the NURBS-warping approach has successfully enhanced the abilities of the shape transformation approach in terms of accuracy and the computation required. In addition, this current work is able to derive a similarity measure in detecting small difference between two highly similar images. The overall analysis of this current work not only shows the potential of this approach in the shape-based image retrieval system for thematic databases but also foresees the future use of this approach in medical image analysis.

REFERENCES

[1] Abrantes, A. J. & Marques, J. S., A class of constrained clustering algorithms for object boundary detection. IEEE Transactions on Image Processing, 5(11): 1507-1521 (1996).

[2] Amini, A. A., Weymouth, T. E., Jain, R. C., Using dynamic programming for solving variational problems in vision. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(9): 855-867 (1990).

[3] Bimbo, A. D., *Visual Information Retrieval*, Morgan Kaufmann, San Francisco, USA (1999).

[4] Bimbo, A. D. & Pala, P., Visual image retrieval by elastic matching of user sketches. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(2): 121-132 (1997).

[5] Brigger, P., Hoeg, J., Unser, M., B-Spline snakes: A flexible tool for parametric contour detection. IEEE Transactions on Image Processing, 9(9): 1484-1496 (2000).

[6] Cohen, L. D. & Cohen, I., Finite-element methods for active contour models and balloons for 2-D and 3-D images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(11): 1131-1147 (1993).

[7] Kass, M., Snake: active contour models. International Journal of Computer Vision, 3(6): 321-331 (1988).

[8] Larsen, O., Radeva, P., Marti, E., Guidelines for choosing optimal parameters of elasticity for snakes. International Conference on Computer Analysis and Image Processing, Prague, Czech Republic, pp. 106-113 (1995).

[9] Liang, K. M., Mandava, R., Khoo, B. E., Free form shape representation using NURBS modeling. WSCG, Plenz, Czech Republic, pp. 141-146 (2002).

[10] Mokhtarian, F. & Abbasi, S., Shape similarity retrieval under affine transform. Pattern Recognition, 35(1): 31-41 (2002).

[11] Piegl, L. & Tiller, W., *The NURBS Book*, Springer-Verlag, New York (1997).

[12] Sclaroff, S. & Pentland, A., Modal matching for correspondence and recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(6): 545-561 (1995).

[13] William, D. J. & Shah, M., A fast algorithm for active contours and curvature estimation. Computer Vision, Graphics and Image Processing: Image Understanding, 55(1): 14-26 (1992).

[14] Xu, C. & Prince, J. L., Snakes, shapes, and gradient vector flow. IEEE Transactions on Image Processing, 7(3): 359-369 (1998).