ヒューマンインタフェース 53−19
（1994. 3. 11）

見本スケッチのダイナミックな構造的組合せによる検索質問をもつ画像データベースシステム

ハッサン・サブリ・シャキール、 長 尾 真
京都大学工学部

見本スケッチのダイナミックな構造的組合せによる検索質問のできる新しい画像データベースシステムを提案する。領域抽出されたスケッチ画像の各領域にそれをおおう包摂長方形と名付けた長方形領域を設定し、これで画像のその領域を代表させる。この包摂長方形はその領域についての種々の画像特徴パラメータをもち、スケッチ画像全体はこの包摂長方形の木構造によって表現される。そしてこの木構造の親ノードと子ノード、子ノード同士の画像特徴関係が付与される。人手で作られたスケッチ画像による画像質問はこのように解析・表現され、画像データベースの各画像の同様な表現との間でファジーマッチングを行なうことによって画像検索を行なう。

Dynamic Token-Based Sketch Query Construction in Image Database Systems

Hussain Sabri Shakir **Makoto Nagao**

Faculty of Engineering, Kyoto University
Yoshida Hon Machi, Sakyo-Ku, Kyoto City, Kyoto 606

This paper presents a new method in sketch query processing of image databases. This method is based on assigning standard tokens called encapsulation rectangles to the regions of the segmented sketch. With the token representation established, processing is focused on the semantic interpretation of the sketch which is drawn from the mapping of encapsulating rectangles. This approach is seen to depend on fuzzy matching measures and weighting functions to control the problems of description approximation and contour correlation differences that accompany the manual construction of the sketch query.

Dynamic Token-Based Sketch Query Construction in Image Database Systems

Hussain Sabri Shakir **Makoto Nagao**

Faculty of Engineering, Kyoto University

Yoshida Hon Machi, Sakyo-Ku, Kyoto City, Kyoto 606

Abstract:

This paper presents a new method in sketch query processing of image databases. This method is based on assigning standard tokens called encapsulation rectangles to the regions of the segmented sketch. With the token representation established, processing is focused on the semantic interpretation of the sketch which is drawn from the mapping of encapsulating rectangles. This approach is seen to depend on fuzzy matching measures and weighting functions to control the problems of description approximation and contour correlation differences that accompany the manual construction of the sketch query.

1. Introduction

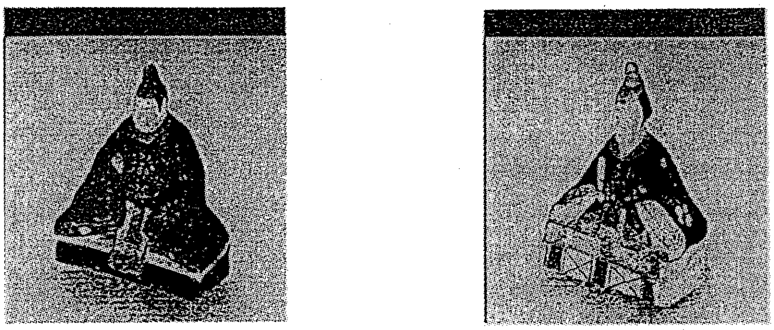
While perfect matching between query and target information is a feature of common sense in conventional database systems, the situation is exactly the opposite when sketch queries in image database (IDB) systems are considered. The reason for this contradiction lays in the conceptual deference between the information types that these two database systems target. In conventional databases architectural instances, redundancy is prohibited in its most specialized form at complete records level. Records are conceptually regarded as semantically simplex entities out of which several levels of complex super entities, such as relations and Cartesian relational product, can be devised. This implies that record components, attribute instances, can be as redundant as the informative combination these attributes' instances allows. On the other hand, extracted shape and texture information that are registered along with their generating natural images during IDB training phase cannot be jointly redundant unless the training image itself is redundant, the case in which the redundant image is omitted.

The conceptual difference above suggests that no two contour information strings belonging to two independent visual entities can perfectly match when the training image set composes solely of natural images. This uniqueness in collective feature prompted several works to address the image retrieval problem from various perspectives. Spatial methods had the early domination in this area due to their straight-forward algorithmic concepts. In Quadtrees [9] spatial kd-Trees [7], and R-Trees [2], training images are either serially partitioned into standard pre-defined shapes and matched against the partitions of the query image, or their pre-extracted semantic partition information is compared with that of the query image. The partition operation is guided by changes of regional intensity with the image region being recursively partitioned if considerable variations are noticed within its boundaries. Queries based on these methods follow the same path in finding their goal through recursive partitioning with the aid of some priori spatial knowledge about the object sought.

Other approaches in IDB information organization is to define an independent a minimal rectangular space that includes a single distinctive image entity. In most of these works this space is found by simply projecting the object itself on the two Cartesian coordinates. Such methods include the cell tree [1], the SPARQ approach [4], and the object-based algorithm developed by Lee and Belford [5]. The performance of both spatial approaches are clearly inversely proportional to the volume of pictorial information involved. Further, it does not provide any information with respect to the shape, texture, and semantic features. The same drawback is shared by the Octree structure [6], which is the 3D version of Quadtrees aiming at remedying another problem with such approaches which is overlapping spaces.

Available contour related query methods generally constitute a specialized extension of existing spatial techniques. The ART MUSEUM system [3] uses smooth sketches as its main query input method. An input sketch is partitioned and a correlation function is used to match the sketch with pre-extracted edge diagrams of registered images serially. Results obtained from the correlation function is directly dependent on perfect matching between the query contour and registered contours and thus is highly sensitive to rotation and while ignoring the semantic contents of the query contour. Consequently, query results shown by the authors included a set of images the are semantically incompatible with each other, all with a considerable rate of compatibility with the input contour. A similar approach is followed in the AI-Mudams system [8].

In this paper we present a new sketch query construction method that takes into consideration the semantic, shape and texture features of the target image. This method is based on the on-line information access facility provided by the nested flat partial power trees architecture [10] and the cross feature relation space that can be accessed through the semantic tree architecture [11].



(a) (b)

Figure 1. An Example on semantic compatibility.

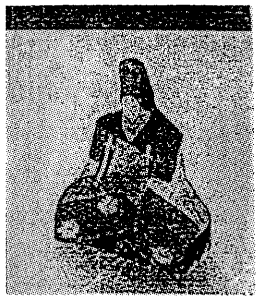


Figure 2. An Example on semantic incompatibility.

2. Inter-feature relationships recognition and utilization

Figure 1 shows an ideal case of conceptual similarity. The image in (a) is incompatible with that in (b) in a pixel-to-pixel manner. Obviously it is desirable, however, to retrieve both of these two images using a single sketch query despite their incompatibility. The Image DataBase Engine (IDBE) should be capable of excluding the image in Figure 2 from its response set although some of its edge candidates resemble those in Figure 1(a) and (b). We can point out the following differences between the three images in the Figures1 and 2:

1) The area ratio between the "FACE" region and the "HEAR" region is considerably higher in the male images, in Figure 1, than the same ratio in the female image of Figure 2,

2) A distinctive rectangular piece of clothes exists starting at the abdominal part and downwards in the images of Figure 1. No comparative edge candidate exists in Figure 2.

With other features closely comparative among the three images, it is obvious that any query method that is restricted to contour matching will produce the image of Figure 2 as a qualified output. Thus additional features has to be considered if satisfactory distinction is to be achieved.

In order to join any number of visual entities, they have to be semantically compatible with each other. In our proposed representation of image information, the semantic tree architecture [11], a visual entity is a multi-level hierarchy of inter dependent sets of features. The region information structure (RIS) as implemented in the project KAMP (Knowledge Assisted Management of Pictures) includes, among others, the following feature attributes:

1- Raw region data,
2- Smallest rectangular shape enclosing the region,
3- Center of gravity of the region,
4- List of shape features such as convex space occupation, circularity, rectangularity, perimeter, area, skeleton, ..., etc.,
5- List of texture features such as mean gray value, local extrema, orientation, texture edges, ..., etc.,
6- List of semantic region labels,
7- List of semantic relation labels that defines the link between this region and other regions within the image.

If discrimination between closely resembling images, as those of Figure 1 and 2, is to be achieved, a high degree of perception of the semantic contents of the sketch is required. The query processing issue now becomes that of visual entity integrity as it is recognized by the IDB system through the its trained storage. If this perception is to be achieved in real time, an approximately common ground has to be found using which registered information is accessed. Specifically, specialized processing should utilize the following extractable information:

1) Encapsulation hierarchy,
2) Area ratio among regions,

3) Position relationships among regions,

This is carried out in favor of strict matching, which obviously will not result in any of the three images of Figure 1 and Figure 2 being retrieved with a high certainty rate.

Using the Cartesian product of the instances of the three information types above, sketch retrieval can be enhanced greatly as will be seen in the next section.

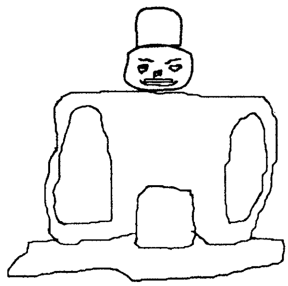


Figure 3. Sample sketch.

3. Token assignment and token hierarchy processing

Sketch query processing operation includes several semantically motivated phases that help narrow the range of compatible images to those with high compatibility.

Image segmentation produces distinct regions which are recursively unified to produce complex visual entities. Thus an input sketch should consist solely of closed regions, as is the case in Figure 3. With this aspect being guaranteed by the sketch drawing utility, the sketch processing algorithm , whose main phases are given in Figure 4, proceeds as follows:

1) All sketch regions are extracted and labeled,
2) The region encapsulation hierarchy is extracted. This is done automatically by a border expansion function that works on the octal neighborhood of the top-left point of each contour which is defined by the characteristic equation:

$$\max(|(x_i-x_j)|,|(y_i-y_j)|)=1 \quad \boxed{1}$$

Where:

(x_i,y_i) = Central coordinates,

(x_j,y_j) = Octal neighborhood member coordinates.

3) Having constructed the encapsulation hierarchy, our next task is to find representative tokens that bring sketches and actual images closer to each other. We propose the Enclosing Rectangle (ER) to be this token. The choice of ER as the standard token offers the following advantages:

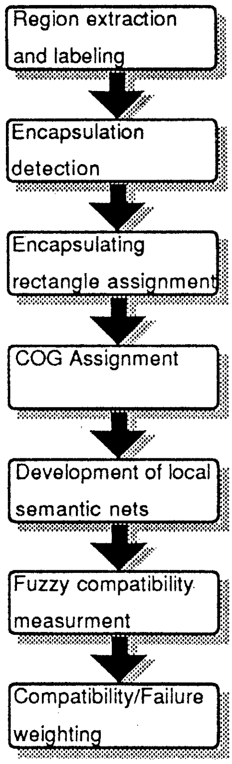


Figure 4. Progress of sketch processing.

1) Spatial relationships are preserved,
2) Area relationships are preserved,
3) Hierarchical composition of the visual entity is preserved,
4) Similarity detection is much easier than when a highly subjective feature, such shape features, are used.

4) With ER specification available for all regions, their center of gravity is found by intersecting the two opposing diagonals of each ER,

5) With centers of gravity available, a semantic net is developed with respect to the daughters of each node in the sketch hierarchy except terminal nodes. Relations among the any two regions covered by a semantic net corresponds to relative size, relative location, and relative distance between of their centers of gravity. Semantic net node attributes include length/height ration of the corresponding ER.

With a semantic model of the sketch completed, domestic search of the IDB begins for semantically compatible hierarchies among registered images. Some important features of the search are:

1) Fuzzy measures are employed for compatibility assessment. Local attribute matching as well as relational compatibility is carried out with a margin of error allowed,

2) It is highly possible that the set of semantic relationships related to a single node of the sketch is found to be equal to an unordered incomplete subset of a registered image node. This possibility is taken into consideration by the search algorithm,
3) Final assessment is achieved be weighting the number of successes against the number of failures hierarchically.

4. Experimental results

In the project KAMP (Knowledge assisted management fo pictures) we experimented the encapsulation-rectangle-based sketch processing with an image database consisting of 85 images of Japanese articrafts. Out of these images, 13 were resembling the sketch given in Figure 3 in their semantic decomposition. And because the original training of the IDB was conducted with manual assignment of recognition labels and guided region summation, the system *could* discriminate between the images in Figure 1 and Figure 2 with a compatibility margin that excedded 30% in several situations. However, discrimenation between the images of Figure 1, when considered as a sketch query responce, could not exceed 7%. This drawback is largely due to the non-intelligent approach using which the original response set, proposed using encapsulation rectangles, was filtered using specific shape features such as perimeter approximation. This requires further work to improve the systems performance with respect to such situations through a method of automatic knowledge capture and calibration.

4. Conclusions

This paper presented a new sketch processing scheme that is based on semantic compatibility rather than the pixel-by pixel approach. Query results shows that this approach is more flexible than the previous approach. The methodology of sketch processing presented in this paper allows also for on-line information access, fuzzy compatibility utilization, and weighting functions implementation. The introduction of encapsulation rectangles as standard tokens was shown to eliminate the need to directly process the contour element which often constitutes the main problem due to the approximations that appear in the sketch. This standard token was shown to preserve the semantic hierarchy of the sketch and clarifies its resemblance with target images, thus improving the efficiency of the processing scheme.

References

[1] O. Gunther. The Design of the Cell Tree : An Object-Oriented Index Structure for Geometric Databases. in Fifth International Conference on Data Engineering, (pp. 598-605). Los Angeles, USA, IEEE Computer Society Press (1989).

[2] A. Guttman. R-Trees : A Dynamic Index Structure for Spatial Searching. in SIGMOD '84. Annual Meeting (Published in SIGMOD Records Vol. 14, No. 2) (pp.47-57). Boston, USA, ACM Press (1984).

[3] Kato, T., T. Kurita, N. Otsu, and K. Hirata. A Sketch Retrieval Method for Full Color Image Database-Query by Visual Example. in 11th International Conference on Pattern Recognition. Conference A : Computer Vision and Applications (pp. 530-533). The Hague, The Netherlands, IEEE Computer Society Press (1992).

[4] R. L. Kerr and P. C. Y. Sheu. SPARQ : Spatial Relationship Query Based on Spatial Decomposition. Computer and Graphics. **15**, 267-284 (1991).

[5] J.-T. Lee and G. Belford. An Efficient Object-Based Algorithm for Spatial Searching. in Eighth International Conference on Data Engineering (pp. 40-47). Tempe, AZ, USA, IEEE Computer Society Press (1992).

[6] D. J. Meagher. Geometric Modeling Using Octree Encoding. Computer Graphics and Image Processing. **19**, 129-147 (1981).

[7] B. C. Ooi, K. J. McDonell, and R. Staks-Davis. Spatial kd-Tree : An Indexing Mechanism for Spatial Database. in COMPSAC 87. The Eleventh Annual International Computer Software and Applications Conference (pp. 433-438). Tokyo, Japan, IEEE Computer Society Press (1987).

[8] M. Sakauchi and S. Satoh. Data model generation in image database systems. Syst. Comput. Jpn. **23**, 14-26 (1992).

[9] H. Samet. The Quadtree and Related Data Structures. ACM Computing Surveys. **16**, 187-260 (1984).

[10] H. S. Shakir, submitted. An Efficient Dynamic Universal Information Indexing Strategy for Intelligent Image Database Systems. Information Systems.

[11] H. S. Shakir, submitted. Semantic Trees: A Highly Versatile Logical Architecture for Hierarchical Image Information Organization. Information Systems.