SKETCH SPATIAL QUERY REPRESENTATION MODEL

Lau Bee THENG and Wang Yin CHAI
Faculty of Information Technology
University of Malaysia Sarawak
Malaysia

ABSTRACT

Sketch is the most natural way of expressing a person's spatial query. Sketch-based spatial query requires essential modeling on sketch query to retrieve meaningful spatial configurations from the spatial databases. This research has proposed a spatial query model that retrieves spatial objects from spatial database by using sketch as query. This paper discusses on the conceptual modeling of spatial query representation.

KEY WORDS

Sketch, spatial databases, spatial query and spatial retrieval.

1. INTRODUCTION

In general, sketch consists of sketched objects with ambiguous shapes, scale and direction. The relations between all objects in a sketch are called object relations. Objects' relations are considered as one of the most important source of information for spatial query where object's geometry is considered only for metric refinements as in [1] and [7][8]. In their works, sketch is modeled into an association graph made up of sketched objects' relations alone. The retrieval is based on the matching of object relations where a sketch is translated into a reduced association graph for relative directions, metric refinements and topological relations. The reduced association graph is very complex to compute and requires long processing time even for matching a simple sketch with a database consist of 100 files with less than 10 objects in individual data record. Furthermore, the existing model is lacking of ability to support the real world spatial databases. Consequently an improved model to interpret and formulate spatial query for a sketch is essential.

There are a few areas of improvement for spatial query by sketch for real world spatial database system such as GIS [1]. The areas of improvements are stated below.

1. How to produce a more efficient digital sketch representation for spatial query?

2. How to index spatial objects and relations in a spatial database for spatial query by sketch?

3. How to efficiently query a set of large continuous spatial database by sketch query?

The above three areas of improvement are indeed interconnected. A model that can meaningfully represent a digital sketch with most relevant details using less computations is required, hence the digital sketch representation can be used to perform indexing on the fly for spatial objects and relations in the spatial database, finally the query of large continuous spatial database can be conducted efficiently.

This research focuses only on the above areas to produce a sketch spatial query representation model that is able to extract and represent a sketch effectively and can also help to improve the retrieval from large continuous spatial database. As a sketch provides meaningful information such as relative size, relative distance, shape, relative direction and topology, the proposed model produces a spiral web structure that is able to represent spatial objects and relations in a sketch taking into account all the abovementioned information. By using the proposed model, a sketch is able to query the large continuous database object by assessing the similarity through a spiral web with the spatial objects from database. Hence the proposed model can support the retrieval from a set of large continuous spatial database such as a street map.

2. THE MODEL

This section discusses on the strengths of the proposed model namely the model over other existing model.

First, existing model uses MBR or Tilted MBR to approximate the sketch objects, hence objects A, B, C and D are treated as similar because they have exactly similar bounding rectangles though B, C and D have been rotated from A. The proposed model is able to preserve the originality of the sketch and represent its objects under various rotations differently.

Secondly, the model is sensitive to relative sizes that exist among objects in the sketch. Hence objects in Sketch A are represented with different values then objects in Sketch B. This is crucial as large continuous databases have plenty of possible matches, the sensitivity to scale can filter out the dissimilarities of objects.

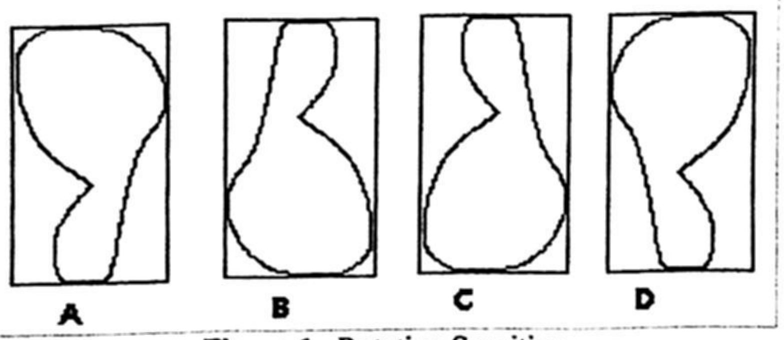

Figure 1. Rotation Sensitive

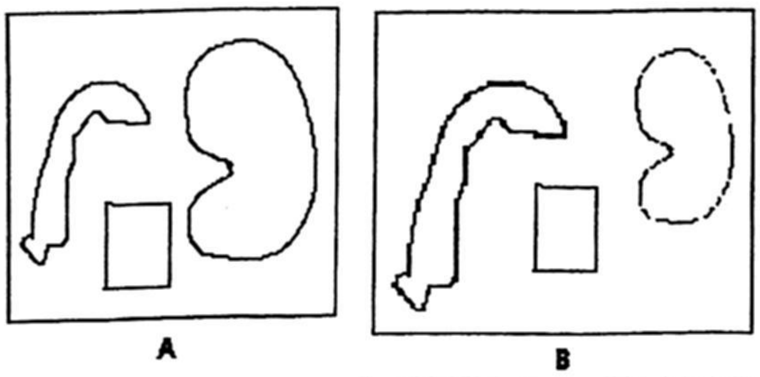

Figure 2. Relative Size Sensitive

Thirdly, the model is sensitive to neighborhood relations that exist between sketched objects within a sketch. It is able to preserve them under any spatial operations like rotation, scaling and movement. For example, C is contained by D, E touches F, G overlaps H, A is covered by B, A disjoin G, D disjoin F etc.

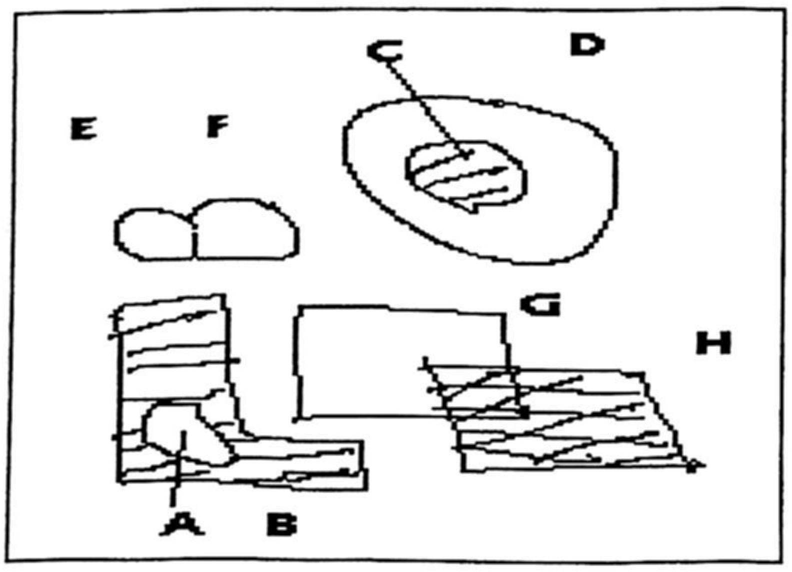

Figure 3. Object Neighborhood Sensitive

Currently, the existing model estimates the reduced object relations by ([n*(n-1)]/2) which still consists of redundancies though it is better than complete association relation (n*n). Fourthly, the model simplifies the number of spatial relations that need to be assessed to further reduce redundancy by using the formula: [2n – 3]. In Figure 5, the object relations computation is [8*(8-1)]/2 = 28 whereby, it is [(2*8)-3]=13 only using the proposed model for the eight objects in the sketch. The algorithm for computing the relation is shown below. Figure 4 shows the number of relations using the above algorithm where there are 5 relations for the 4 objects.

```
FOR i = 1 to TotalObjectCount
  FOR j = 2 to TotalObjectCount
    IF i <> j THEN
      BuildRelation (Object(i), Object(j))
    END IF
  NEXT
NEXT
```

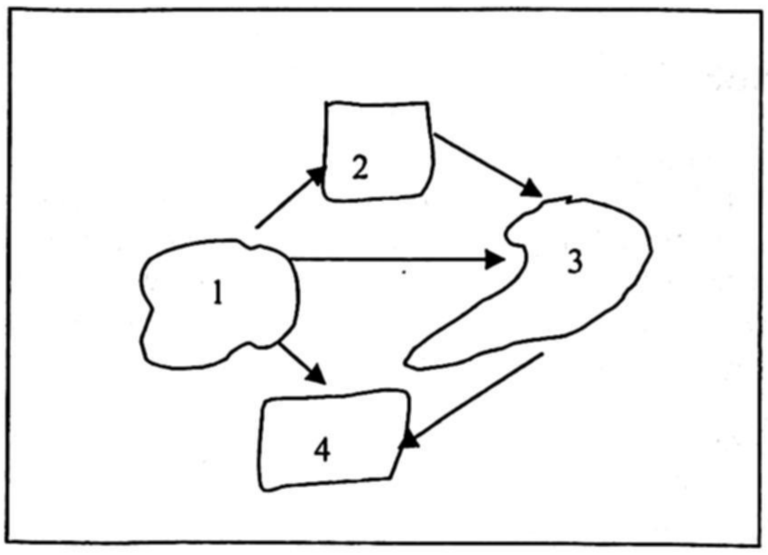

Figure 4. Reduced Relations for 4 Objects

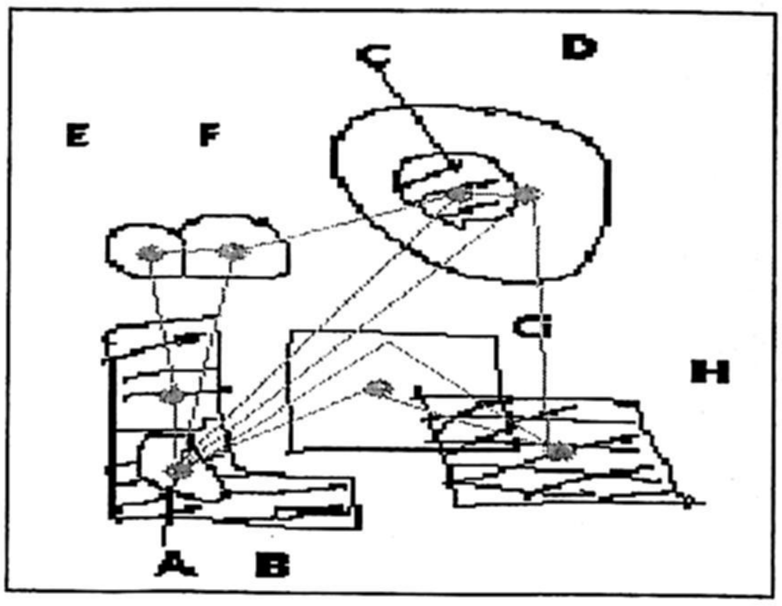

Figure 5. Reduced Object Relations

Basically, the spatial query by sketch model uses approximation to represent the sketch. However the shapes of each object bear a significant meaning when user sketches it. Fifthly, the model represents a sketch with its original object shapes without approximating them using any bounding rectangle or point. The sample in Figure 6 shows a representation of sketch A in Spiral Web B.

Sixthly, the model is inferential so that it can be used to determine the loose or unknown sketch properties from the known representation. For instance, the relative distance among the sketch objects can be viewed and estimated by looking into the ring and zone in the spiral web shown in Figure 6. Furthermore the relative directions among sketch objects can be easily recognized from the location on the spiral web. Other property like relative size can be easily inferred from the objects in the spiral web as well. This inferential attribute has not existed in the existing spatial query by sketch model.

Currently, there is no formal model being developed specifically to meet the requirements of sketch-based spatial query and retrieval processing. Seventhly, the model provides a formal basis of sketch representation for query and retrieval processing.

Lastly, the model provides a sketch representation that each object is quantifiable and distinguishable from one to another. Hence it also provides quantitative assessment to similarity between sketch and objects in the spatial database for spatial objects retrieval.

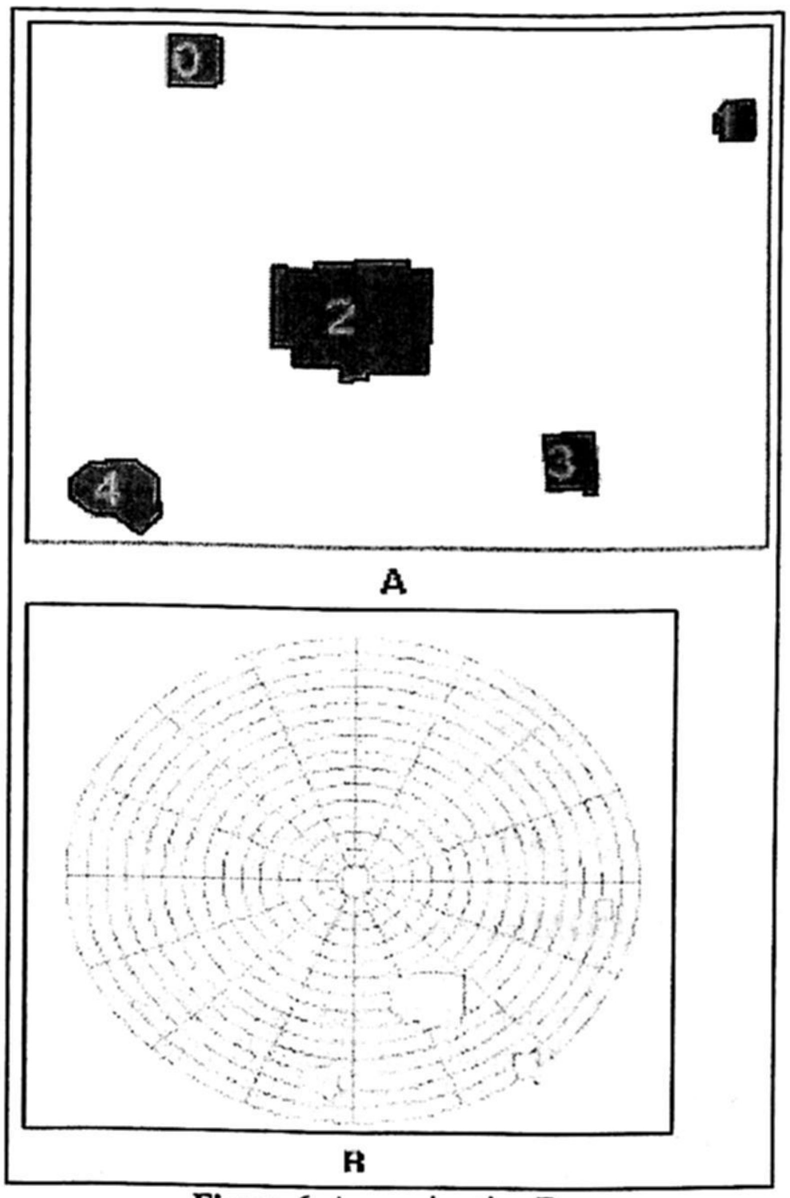

Figure 6. Approximation Free

3. HOW DOES THE MODEL WORK

The graphical illustration on how the model work is shown in Figure 21. Firstly, when a sketch is received, then representation process starts working to create a spiral web for the sketch. During the creation of spiral web, the values of all sketched objects also computed. Then the spiral web is used for retrieval from spatial database that consists of more than a thematic layer. The created spiral web is used as an index on the fly for the spatial database. Then systematic search is applied to retrieve all the selected objects based on the index. After the systematic search with the spiral web index, each set of selected objects is stored as individual spatial scene. Each retrieved spatial scene is assessed for similarity values. For instance in Figure 21, one of the retrieved spatial scene produces the similarity of "1" as compared to each object in the sketch.

4. HOW TO REPRESENT SKETCH

After the brief discussions about the strengths of the model and how it works, this section describes how the query representation in the model works.

A sketch is extracted and represented as a spiral web structure that is made up of rings, zones and cells. The number of cells is a product of numbers of zone and ring They are determined by factors like relative distance of reference object to the furthest object, diameter of the reference object and relative direction in a sketch's orientation. For example, in Figure 9, there are 4 rings and 64 cells. The number of cell is not fixed as it grows or shrinks according to the abovementioned factors. The number of zone represents the relative directions. Figure 6 illustrates a sketch consist of 5 objects is transformed into the designed spiral web. There are 15 rings, 16 zones and 240 cells in the spiral web. All the sketched objects are also created as part of the spiral web. The representation of the sketched objects on the spiral web is determined by which ring and zone it falls on.

There are different computations for determining various positions of sketched objects on the spiral web as discussed in the following sections. The input for query representation is a digital sketch stored as a spatial data file that consists of sketched objects.

A. SPIRAL WEB STRUCTURE

The spiral web is the main component of query representation model. A spiral web is actually a spatial data structure built as a result of processing a sketch in which it contains the data from the sketch and also a set of cells. Spiral web is defined as a collection of zones and rings that is built to relate all sketched objects in a sketch. Each sketch has a spiral web that is tailored made for it. The number of zones and rings depend solely on the number of sketched objects in a sketch. A reference object is chosen to be the center of each spiral web. The reference object is the first sketched object drawn by user. The spiral web is highly sensitive to the reference object.

A spiral web is stored as a spatial object with these attributes: {SpiralWebObjectID, SpiralWebObjectShape}. SpiralWebObjectID is the index of spiral web; SpiralWebObjectShape is the geometry type of a spiral web object in generic format that is polygon or line. The spiral web is in generic format that spatial data file like ESRI Shapefiles can provide. This gives the convenience of applying this model into any generic spatial data management system like GIS. With this simple spiral web structure, it is able to simplify the processing task in spatial query and retrieval by sketch.

B. ZONE COMPUTATION

People manipulate concrete relations rather than continuous quantitative direction i.e. angles to express and reason about directions. Most previous work defines qualitative directions using either object projections or centroids. Each approach has its advantages and shortcomings. In this research, a centroid-based method is applied where the direction between two objects is determined by the angle between their centroids. The set of relative direction relations used is shown in Figure 8. These directions are used to represent the zones in the spiral web as shown in Figure 9. The 16 zones represent the 16 qualitative direction model. There is no fixed numbers of zone for a spiral web but it must follow the direction model such as 4 zones (north, east, south and west), 8 zones (north, northeast, east, south, southeast, northwest, southwest and west), 16 zones (north, northeast, north northeast, east northeast, south, southeast, south southeast, east southeast, east, northwest, north northwest, west northwest, south southwest, southwest, west southwest and west) or more by 4^n. The derivation of zones depends on the number of objects in a sketch, the relative distances of each object pair formulated as shown in the algorithm in Figure 7. The sufficient condition is determined by the number of objects fall in each zone, if there is more than one object in a zone, then more zones need to be drawn. There are 8 zones in Figure 9 and there are 4 zones in Figure 9. The rationale behind these various zones is the less zone, the less details can be provided; the more zones, the higher the processing power is needed. However if there are sixteen zones and only three objects exist in the sketch, then the number of zones may be simplified accordingly such as reduces it to 8 or 4 zones only. For instance, in
Figure 10, there are four sketched objects in the sketch, therefore the spiral web only consists of 8 zones as it is sufficient to describe the sketch where Object B falls in Zone 8, Object C falls in Zone 1 and Object D falls in Zone 2 and 3 and Object E falls in Zone 4.

Algorithm to Determine Number of Zones

```
GET Sketch
FOR each object to total number of objects in sketch
        Check location
        Build 4-Zone
        Estimate Value of object
NEXT object

IF 4-Zone is not sufficient to describe the objects THEN
        FOR each object in sketch
                Check location
                Build 8-Zone
                Estimate Value of object
        NEXT object

        IF 8-Zone is not sufficient for the objects THEN
                FOR each object in sketch
                        Check location
                        Build 16-Zone
                        Estimate Value of object
                NEXT object
        END IF
END IF
```

Figure 7. An Algorithm to Determine Number of Zones for A Spiral Web

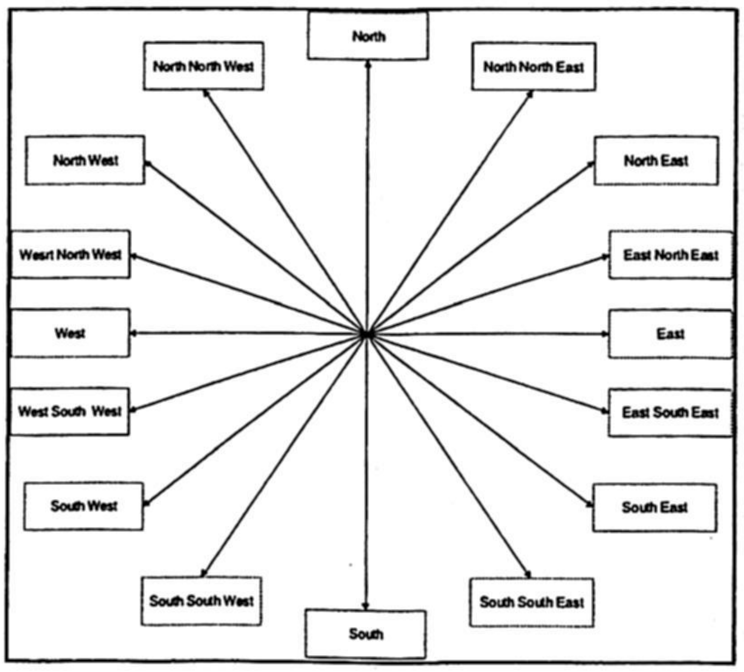

Figure 8. Directional Zones

C. RING COMPUTATION

In this research, a ring is defined as a circular ring derived specifically for the spiral web. Figure 11 shows the ring structure used for spiral web in which the rings are separated to illustrate the ring. In fact, they are joined together to make up the spiral web as shown in Figure 9. The computation of ring depends on the total distance of the reference object from the furthest object in the spatial scene. There is a formula derived specifically to build the spiral web as shown in Equation 1. In Figure 10, there are 5 objects where the furthest object is Object E, therefore the spiral is built with 4 rings based on the diameter of Object A that is the reference object in the sketch. Whereby in Figure 10, there are 5 objects in a sketch, and there are 3 rings created for the spiral web as the furthest object is Object E that falls into the 3rd ring.

$$L = EucDist / Diam,$$

where:
L: Number of ring.
EucDist: Euclidean distance of reference object to the furthest object in the sketch.
Diam: Diameter of the reference object.

Equation 1. Estimate number of rings in a spiral web

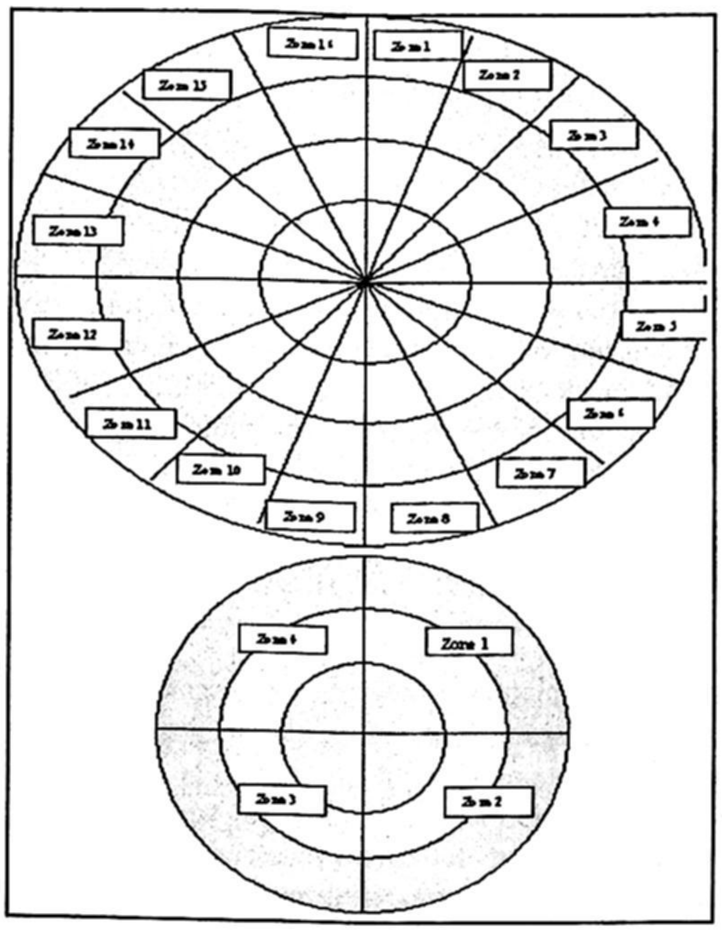

Figure 9. (a) A Sample of Spiral Web Structure Consists of 16 Zones and 4 Rings (b) Spiral Web Structure Consists of 4 Zones and 3 Rings

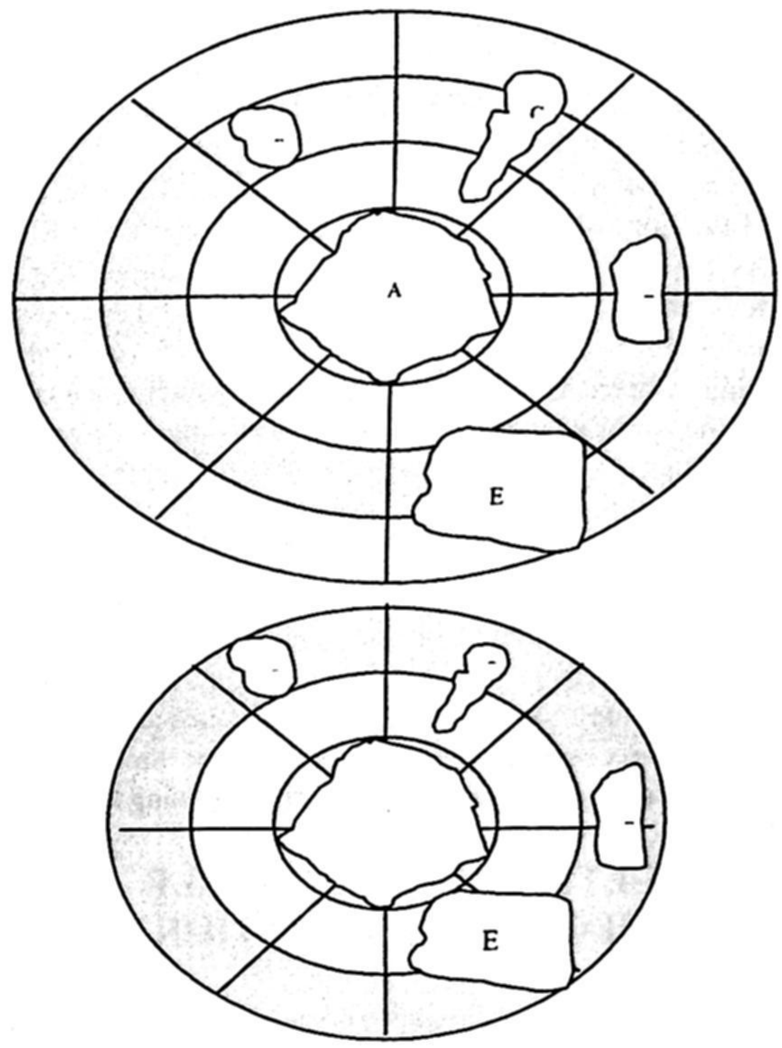

Figure 10. (a) A Sample of Spiral Web Structure Consists of 8 Zones and 3 Rings for A Sketch (b) A Sample of Positioning of Sketch Objects on the Spiral web Structure

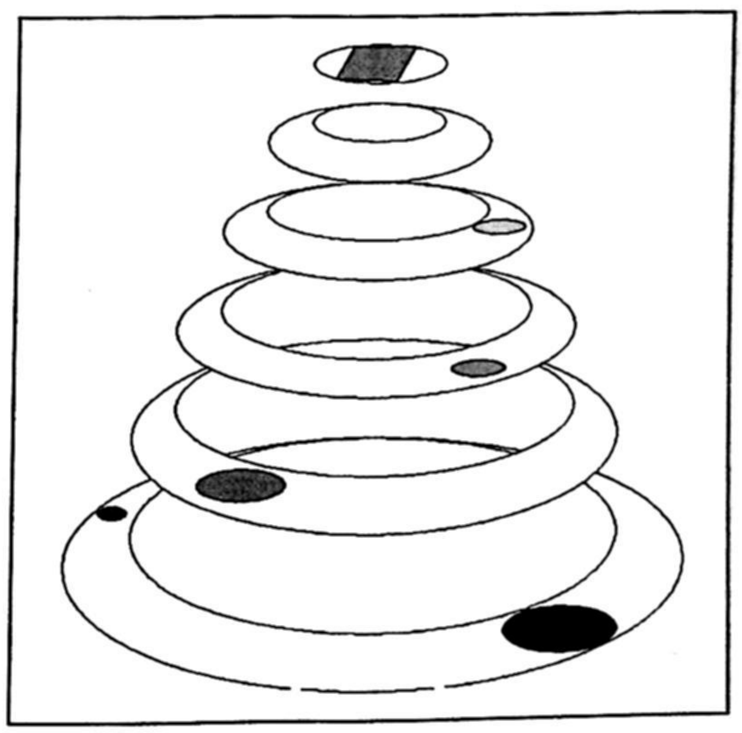

Figure 11. A Sample of Ring Structure Consists of 6 Rings for A Sketch with 6 Objects

D. SKETCHED OBJECT POSITIONING

Once a sketched object is being placed on the spiral web, it has a possibility to fall on single cell or multi cells. For cases of multi cells, there are Multi Zones Single Ring, Multi Zones Multi Rings, and Single Zone Multi Rings. As for single cell, there is Single Zones Single Ring. As for various positioning, the estimation of sketch object value also varies. In Figure 9, there are 16 zones and 4 rings. Object A is the reference object which is the center of the spiral web, the diameter of each ring is determined by Object A's diameter. Object B is a sample of Single Zones Single Ring object. Whereby, Object C is Single Zone Multi Rings object. Object D is Multi Zones Single Ring object. Lastly, Object E is Multi Zones Multi Rings. Due to this different sketch object positioning on the spiral web, there are 4 different computations used to estimate the object values of the sketched object on the spiral web. The detailed computation of the objects in these 4 categories is discussed in the following sections.

▪ MULTI ZONES SINGLE RING OBJECT REPRESENTATION

The multi zones single ring is a category given to describe object that falls into more than one zone but only one ring in the spiral web. Object D shows the multi zones single ring sketched object representation where Object D is in the zone 4 and 5 but in ring 3 of the spiral web. Figure 9 shows zone categorization and Figure 11 shows ring categorization. Figure 10 shows different positioning of sketched object on the spiral web. By using the estimation in this model, the object value of Object D that has 0.4 of its area falls in Zone 4 and 0.6 area falls in Zone 5 and all area fall in Ring 3.

▪ MULTI ZONES MULTI RINGS OBJECT REPRESENTATION

The multi zones multi rings is a category given to describe object that falls into more than one zone but only one ring in the spiral web. Object E shows the multi zones multi ring sketched object representation where Object E is in the zone 7, 8 and 9 but in ring 2, 3 and 4 of the spiral web. Refer to Figure 9 for the zone and ring categorization is shown in Figure 11.

Figure 10 shows different positioning of sketched object on the spiral web. Object E has 0.3 of area falls in Zone 10, 0.5 of area falls in Zone 9 and 0.2 of area falls in Zone 8. 30% of area of Object E falls in Ring 2, 45% area falls in Ring 3 and 0.25 of area falls in Ring 4.

▪ SINGLE ZONE SINGLE RING OBJECT REPRESENTATION

The single zone single ring is a category given to describe object that falls single zone and single ring in the spiral web. Object B shows the single zone single ring sketched object representation where Object B is in zone 15 and ring 3 of the spiral web. Refer to Figure 9 for the zone and ring categorization is shown in Figure 11.

Figure 10 shows different positioning of sketched object on the spiral web. Object B has 100% of area falls in Zone 15, 100% area falls in Ring 3 alone.

▪ SINGLE ZONE MULTI RINGS OBJECT REPRESENTATION

The single zone multi rings is a category given to describe object that falls single zone and multi rings in the spiral web. Object C shows the single zone single ring sketched object representation where Object C is in zone 15 and in ring 2, 3 and 4 of the spiral web. Refer to Figure 9 for the zone and ring categorization and refer to Figure 10 for the different positioning of sketched object on the spiral web. Object C shows the single zones multi rings sketched object representation.

5. HOW TO FORMULATE SPATIAL QUERY

After the completion of the query representation stage, the input to the spatial query formulation stage is a spatial data file containing a spiral web. The spatial query formulation is formed based on spiral web created in the Section 4. In short, for each processed sketch, there are more than a spatial query is formulated. The number of spatial query formulated is determined by the number of zones exists in the spiral web created for a sketch. If there are 4, then four spatial queries shall be formulated. The algorithm to formulate the spatial query is extracted and shown below.

Spatial Query Formulation

```
LET PI = 3.142
LET a = Zone
FOR angle = PI /Zone to (2*PI) STEP PI / a
        FORMULATE Query
        STORE Query
NEXT
```

Figure 12. An Algorithm for Spatial Query Formulation

For instance, for the sketch shown in Figure 6, the spatial query formulation produces a list of spatial queries in spatial data file format as shown from Figure 13 to Figure 20 Query. Each spiral web represents a spatial query for the sketch. Query 2 is formulated when the original sketched objects are rotated for 45 degrees its original position in the spiral web.

As a result, Query 3 is a resulted query by moving the original sketched objects by 90 degrees from the original position. Consequently, Query 4, Query 5, Query 6, Query 7, and Query 8 are formulated by an increment of 45 degrees from the previous query. The retrieval of spatial objects for a sketch is done based on the queries made in the form of spiral web.

5. CONCLUSION

Though sketch as a query to spatial database management systems like GIS suffers plenty of deficiency such as in deficiency in relative size, shape, relative direction and relative distance similarity assessment, we are not supposed to ignore all these sketch behavior in processing spatial query. In fact, neighborhood relations and objects' geometry are equally important in sketching spatial query.

To model a good sketch representation model, this research has proposed a spiral web to represent a sketch's configuration setting. This paper details the conceptual modeling of the proposed model. It explains the model that can utilize the unique behavior of sketch in formulating a good spatial query representation model. It explores a new way of modeling sketch as spatial query in a real world spatial database system. With the spiral web representation, it shall be able to enhance the spatial retrieval by sketch. Currently, testing and evaluation of the model is underway. Hence the results discussions shall be published in the next paper.

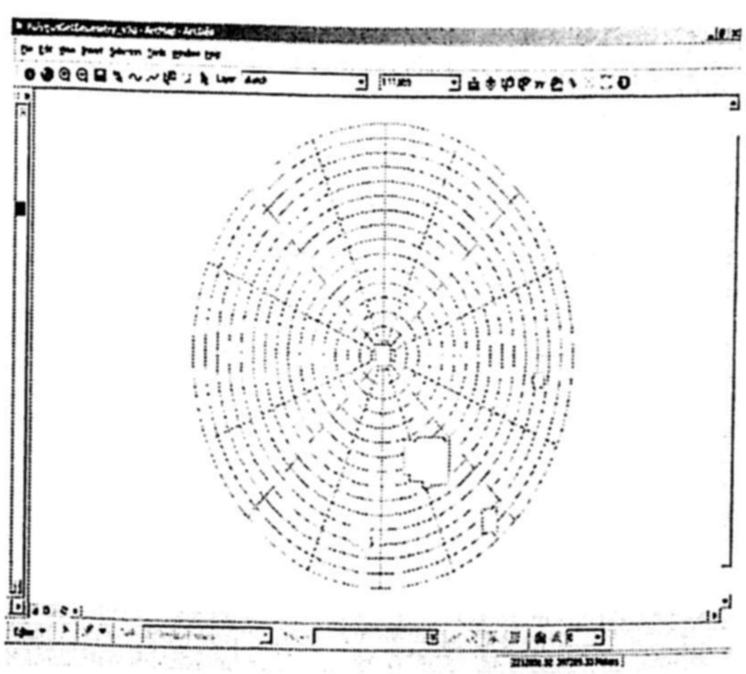

Figure 13. Query 1

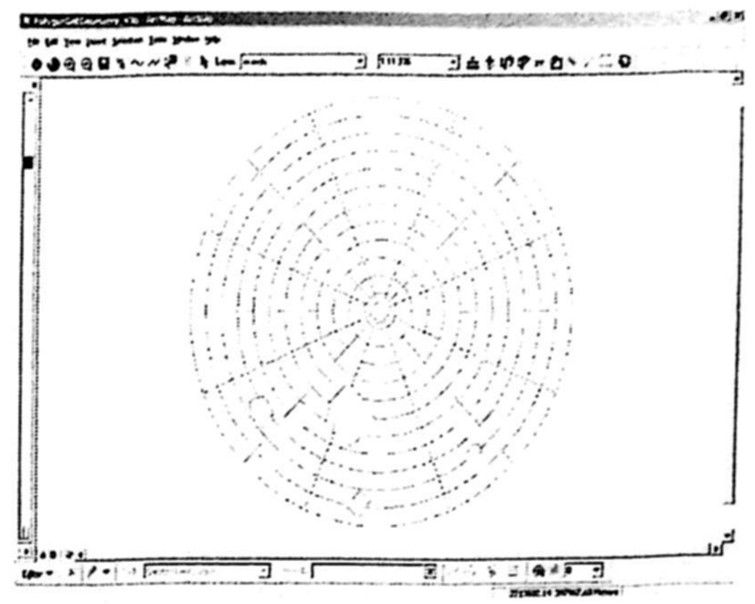

Figure 14. Query 2

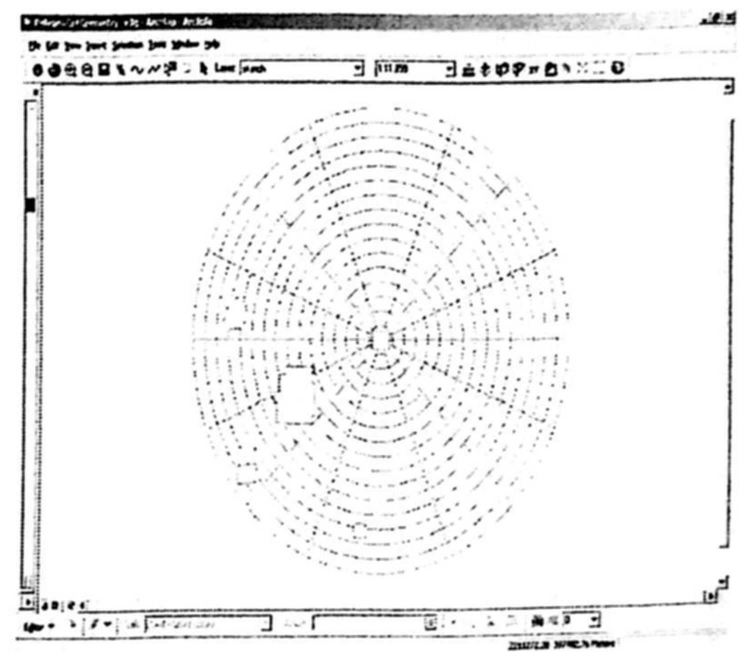

Figure 15. Query 3

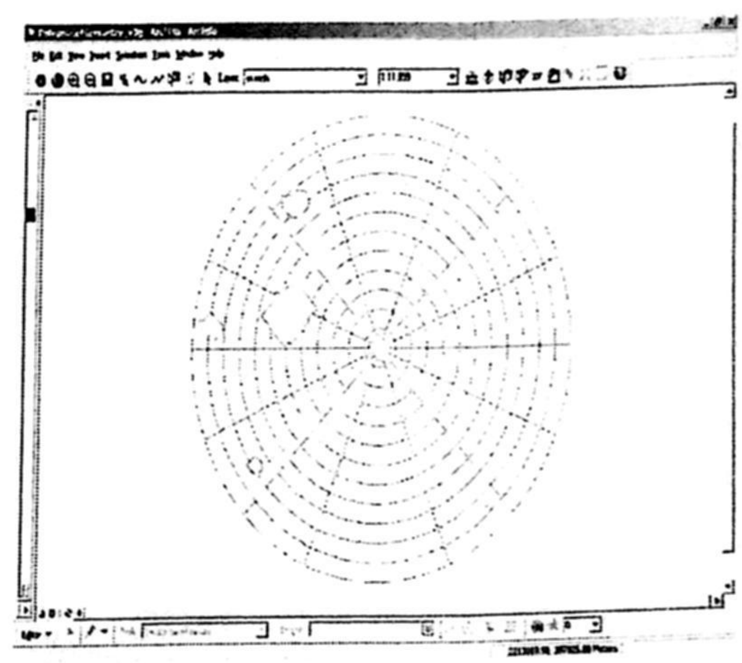

Figure 16. Query 4

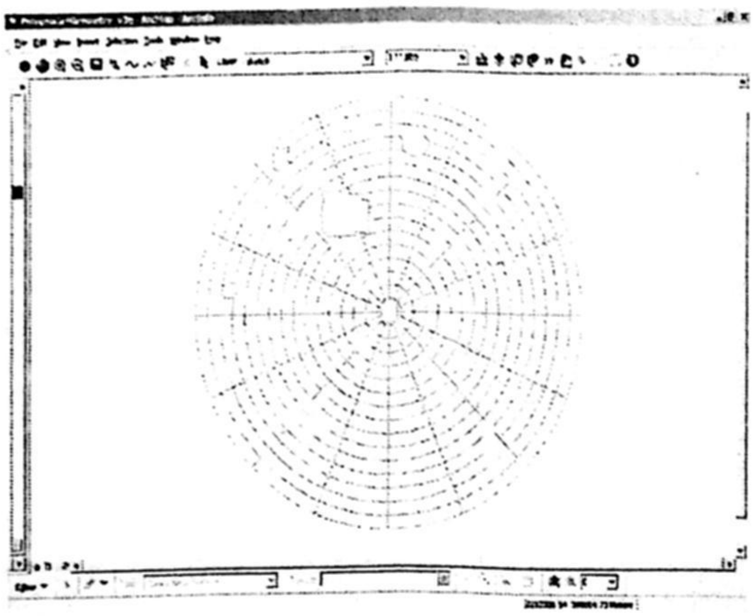

Figure 17. Query 5

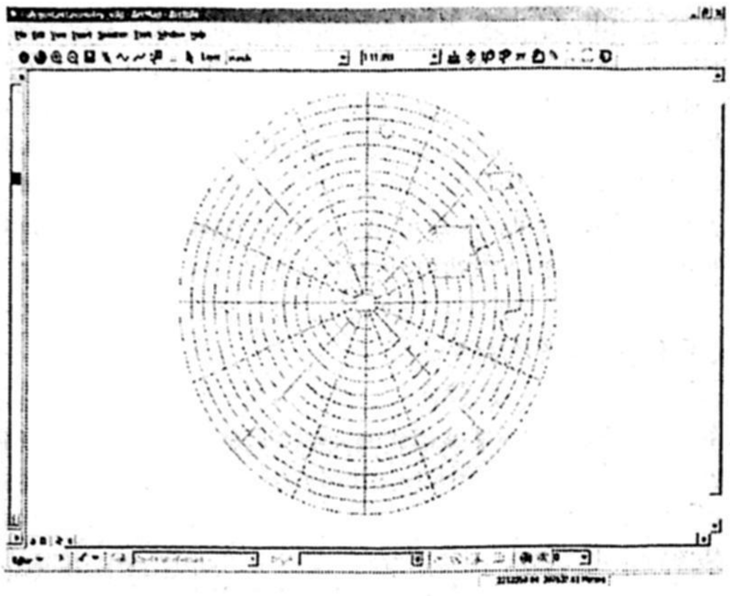

Figure 19. Query 7

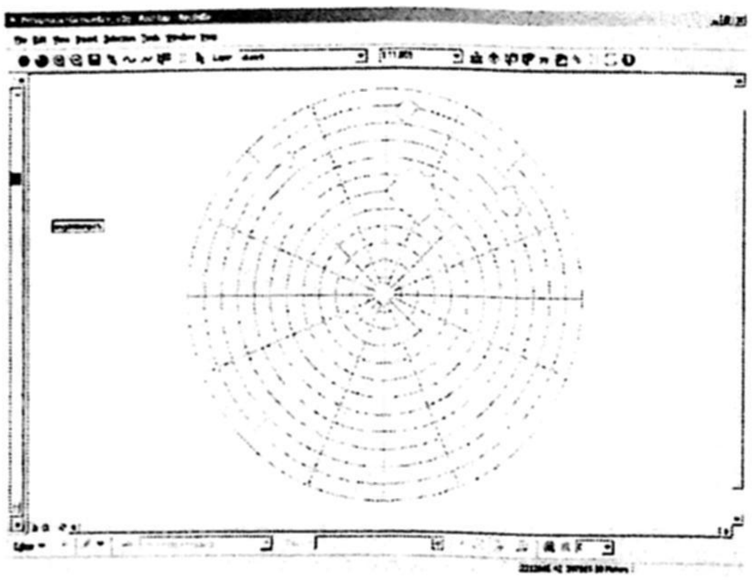

Figure 18. Query 6

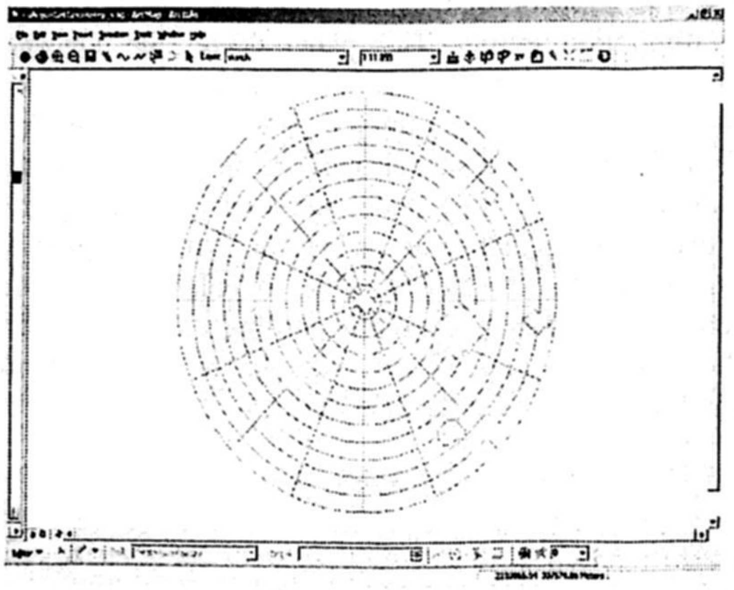

Figure 20. Query 8

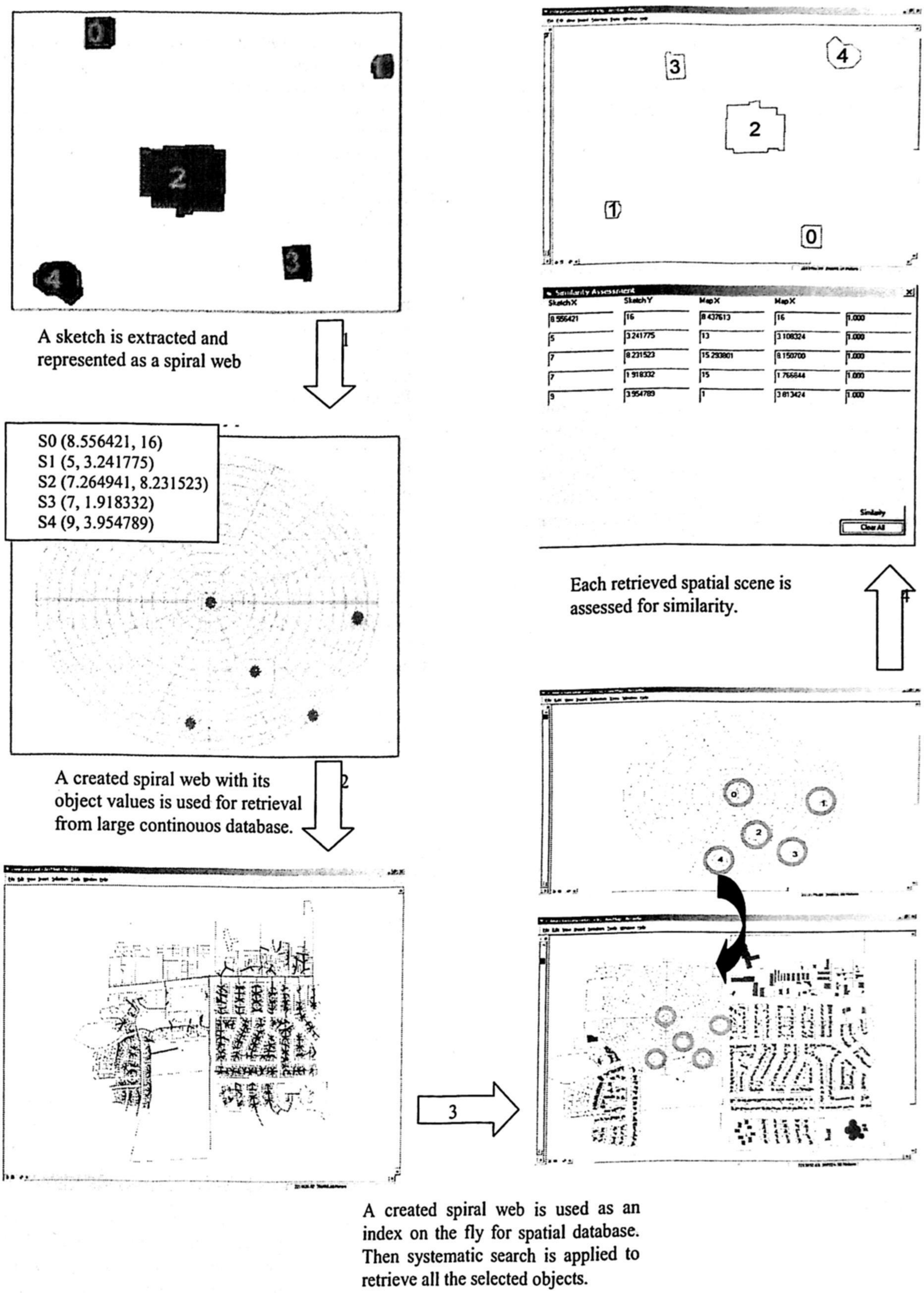

Figure 21. How Does the Model Work

ACKNOWLEDGEMENT

A bundle of thanks to my advisor, Assoc. Prof Dr. Wang Yin Chai for his passion and patience throughout the research.

REFERENCES

[1] Blaser, A. (2000). Sketching Spatial Query. Dissertation of PhD. University of Maine.

[2] Blaser, A. and Egenhofer, M. (2000). Visual Tool for Querying Geographic Databases *AVI2000 – Advanced Visual Databases*, Salerno, Italy in V. Di Gesù, S. Levialdi, and L. Tarantini (eds.), pp. 211-216, May 2000.

[3] Blaser, A., Sester, M., and Egenhofer, M. (2000). Visualization in an Early Stage of the Problem Solving Processing GIS *Computers and Geosciences*, 26 (1): 57-66.

[4] Egenhofer, M. (1991)). Reasoning about Binary Topological Relations. In Gunther, O. an Schek, H.J. (eds). Advances in Spatial Databases, Springer Verlag LNCS.

[5] Egenhofer, M. (1994). Pre-Processing Queries with Spatial Constraints. *Photogrammetric Engineering & Remote Sensing* 60 (6): 783-790.

[6] Egenhofer, M. (1994). Spatial SQL: A Query and Presentation Language. *IEEE Transactions on Knowledge and Data Engineering* 6 (1): 86-95.

[7] Egenhofer, M. (1996). Spatial-Query-By-Sketch. *VL'96: IEEE Symposium on Visual Languages*, Boulder, CO in M. Burnett and W. Citrin (eds.), pp. 60-67, September 1996.

[8] Egenhofer, M.(1997).Query Processing in Spatial-Query-By-Sketch. *Journal of Visual Languages and Computing* 8 (4): 403-424.

[9] Papadias, D. (2000). Hill Climbing Algorithms for Content Based Retrieval. Proceedings of the ACM Conference on Information Retrieval (SIGIR), Athens, July 24-28, pp. 240-247.

[10] Papadias, D. and Mamoulis, N.(1999). Improving Search Using Indexing: A Study with Temporal CSPs. Proceedings of International Joint Conference of Artificial Intelligence, Stockholm, Sweden, 436-441.

[11] Papadias, D. Kalnis, P, Mamoulis, N. (1999). Hierarchical Constraint Satisfaction in Spatial Databases. Proceedings of AAAI Conference, Orlando, FL, 142-147, AAAI/MIT Press.

[12] Papadias, D. Mamoulis, N., Theodoridis (2001). Contraint-Based Procesing of Multiway Spatial Joins. Algorithmica, Vol. 30(2), 188-215.

[13] Papadias, D., Mantzourogiannis, M., Kalnis, P., Mamoulis, N., Ahmad, I. (1999). Content Based Retrieval using Heuristic Search. Proceedings of the ACM Conference on Information Retrieval (SIGIR), Berkeley, CA, 168-175.

[14] Papadias, D., Momoulis N., Delis, V. (1998). Algorithms for Querying by Spatial Structure. Proceedings of Very Large Databases Conference (VLDB). New York.

[15] Papadias,D., Karacapilidis, N. and Arkoumanis, D. (1999c). Processing Fuzzy Spatial Queries: A Configuration Similarity Approach. International Journal of Geographic Information Science Vol. 13(2), 93-128.