PROCEEDINGS

The 1st Postgraduate Colloquium

Faculty of Computer Science And Information Technology

October 6, 2004

Universiti Malaysia Sarawak

UNIVERSITI MALAYSIA SARAWAK
UNIMAS

FOCUS IT
Focus on individual Transformation

Preface

This is the Proceeding of the Inaugural Post-Graduate Colloquium of the Faculty of Computer Science and Information Technology held on the 6th October 2004.

The aim of the colloquium is to showcase outstanding research being carried out at the Faculty while fostering the exchange of ideas and research results amongst researchers across educational institutions in Sarawak.

We are pleased to announce that 22 papers are being presented by the Postgraduate students of the Faculty. I wish to congratulate the postgraduate candidates and their supervisors for their high quality research papers.

It is hoped that this colloquium will provide a platform to bring together local researchers to share their knowledge and experiences in a broad range of topics relating to the core research areas of the faculty.

We would also like to acknowledge and thank the many people who have contributed greatly to the conference. I would like to take this opportunity to thank Professor Khairuddin Ab. Hamid for his support towards the R&D activities in the Faculty and for officiating the event. I also wish to thank the judges of the best paper awards for their invaluable contribution. Last but not least, I wish to record my sincere appreciation of the members of the organizing committee for their hard work and endurance in ensuring the success of this event. The efforts of the Post-graduate candidates who themselves played key roles in the organization of the Colloquium is deeply appreciated.

I welcome all our guests and wish you all an enjoyable colloquium with fruitful deliberations.

Assoc. Prof. Narayanan Kulathuramaiyer

Organizing Committees

Advisor
Assoc. Prof. Narayanan Kulathuramaiyer

Chairman
Dr. Alvin Yeo Wee

Secretary
Ms Zuraini Ramli

Editorial Board
Dr. Rosziati Ibrahim
Dr. Jane Labadin
Mr Lim Tek Yong
Ms Lim Phei Chin

Promotion Committee
Mr Johari Abdullah
Ms Chiu Po Chan

Publicity Committee
Ms Jacey Minoi
Ms Tan Ping Ping

Logistic Committee
Assoc. Prof. Dr. Wang Yin Chai
Mr Sapiee Hj. Jamel
Mr Goh Hock Chui

Paper Award Committee
Mr Sylvester Arnab
Mdm Eaqerzilla Phang

Judges
Prof. Justo Diaz
Dr. Wong Chee Weng
Dr. Tan Chong Eng
Dr. Jane Labadin
Dr. Hong Kian Sam
Mr Lee Nung Kion
Mr Ng Giap Weng

Table of Contents

Page

Knowledge Systems

Page

Information Infrastructure

A Dynamic Approach For Scheme Transformation Of Spatial Data From Relational Database To Object-oriented Database

Siew Teck Kiong
Geographical Information System
siew_tk@yahoo.com

ABSTRACT

Relational data models have proven to have limitations in modeling spatial data in many aspects. In the latest computer technology, object-oriented database technology extends support for the complex data types to deal with the today's complex applications. In view of this, the attempt in this study is to develop a dynamic approach, which involves schema transformation of spatial data from the relational database to the object-oriented database. The methodology consists of reverse engineering process and semantic reconstruction to produce dynamic transformation with minimum loss of semantics. It is implemented using an object-oriented paradigm. Dynamic scheme transformation is achieved through the applications of object-oriented features such as data abstraction, inheritance, polymorphism and behavior modeling.

KEYWORDS

Reverse engineering process, semantic reconstruction, virtual visibility

1 INTRODUCTION

This study is in a period of intensive change and innovation regarding database technology. A different trend can be seen in the history of the last 20 years of software development. Relational database systems (RDBMS) revolutionized database management systems in the 1980s, and object oriented programming languages are revolutionizing software development in the 1990s. This results from the increasing needs to store and manipulate complex data in the databases as the complex data is getting imminent across a wide range of applications such as multimedia applications for the Wed, specialized application domains including medical care, geographical, space, and exploration systems and even financial systems (Frank, 1998). In addition, the classical databases based on the hierarchical, the network, and the relational data model have proven to be insufficient for these complex applications in many ways (Herring, 1992; Egenhofer & Frank, 1992). However, the dominant database in the real world applications is still the relational database due to the proliferation of the relational systems, and their dominance will remain in the near future. In view of the today database market and the needs to store and manipulate the complex data in the databases, there are various efforts in integrating relational and object-oriented systems. However, there are significant problems existing in these approaches in modeling the complex data. In view of this, this paper discusses a dynamic schema transformation for spatial data from relational databases to object-oriented databases.

The rest of this paper is organized as follows: Section 2 gives a brief overview of the database models and related GIS architectures. Section 3 explains the research methodology adopted for the schema transformation. Section 4 describes the overall architecture of the dynamic schema transformation. Section 5 introduces virtual visibility in enhancing object-oriented paradigm in the implementation of the dynamic schema transformation, and finally, conclusion is given in Section 6.

2 DATABASE MODELS AND RELATED GIS-ARCHITECTURES

Choosing a standard data model is critical to making database transformations practical, and represents a wide endorsement of our database transformation approach. In relation to this, this study works with the standard spatial databases. It selects Vector Product Format (VPF) as the source relational database and Object-Oriented Geo-Data Model (OOGDM) as the target object-oriented database as indicated in Figure 1.

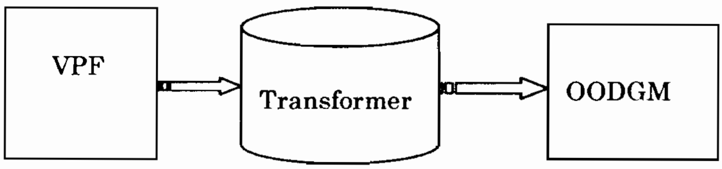


Figure 1. Spatial databases in transformation

Two standard databases are selected due to the various reasons. The Vector Product Format (VPF), developed by the U.S. Defence Mapping Agency (DMA, 1993) has become part of the DIGEST international standard formats for representing geographical data. VPF is a georelational specification, which uses a relational data framework for storing both attribute and geo-spatial information about the geographical features represented. In following the standard rules, it provides a public specification for exchange of geographic data across computer platforms and GIS software products. It occupies a similar role to the Spatial Data Transfer Standard (SDTS) developed jointly by the U.S. Census Bureau and the U.S. Geological survey. Besides, a useful aspect of the VPF specification is that the relational files have their schema description within each file's header. This facilitates dynamic interpretation and processing of feature data as well as a means of coping with some of the differences in structural specifications among the various VPF products.

OODGM (Andreas, 1997) is a new object-oriented data model for GIS. Based on object-oriented design principles, OODGM provides support for data types most commonly found in GIS-applications. Other existing data models for GIS based on object-oriented technology only partially cover the complex structure of GIS data. Comparatively, the type hierarchy offered by OODGM covers the core set of data found in any GIS-application. A set of common operations for GIS-applications is incorporated within the data model. Hence, the spatial data types are organized in a clear hierarchy and support a basic set of functionality covering geometrical, topological, and directional predicates and operations. Besides, OODGM is a standard object-oriented model based on the concepts according to the Object Database Management Group (ODMG).

3 METHODOLOGY

Schema transformation process is the most crucial step in the database transformation. The aim is to produce an equivalent database in the target database as the source database in terms of semantic. This plays an important role especially in preserving the semantics of the relational data when transformed to the object-oriented data structures without the loss of information. Many current approaches use reverse engineering process (Haivant, 2002) to perform schema transformation. Reverse engineering process extracts additional information especially, which is not explicitly available from relational database. The strategies of reverse engineering process undertaken by the current approaches includes analyse of source database, application code, queries and others (Anderson, 1994; Chiang et al., 1997; Petit et al., 1996). The strategies usually involve user interaction and human expert. Besides, current approaches perform reverse engineering of the source database only. This study explores both source and target databases for the semantic information in reverse engineering process. The first step is to analyse the database structures between the relational and object-oriented system. As the two database structures are not compatible, there is no direct representation in between the two databases. So based on the analysis of the databases, it is necessary to make the semantic reconstruction of the object-oriented database without modification to both databases' structures. Only then the mapping rules are derived between the two databases for the second step, data transformation. In addition, the task of schema transformation must be as generic as possible so that it becomes possible to do schema transformations in a system-independent way. The study also attempts to facilitate dynamic interpretation and processing of data to enable an automatic schema transformation process.

4 ARCHITECTURE OF THE DYNAMIC SCHEMA TRANSFORMATION

The architecture of the dynamic schema transformation consists of the three steps as indicated in Figure 2.

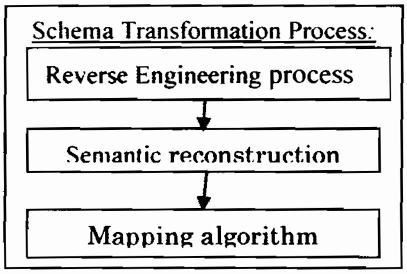


Figure 2. Schema transformation process.

4.1 Reverse Engineering Of Relational And Object-Oriented Databases

Doing a good job at the schema transformation requires a solid understanding of both databases. This is the first step to be taken in the process. The representations of two databases have the similar architecture. Figure 3 indicates the levels of abstraction between VPF and OODGM database architecture. Both the databases have the similar levels of abstraction. However, they are semantically different in the structures. This study is going to analyse these differences semantically in the followings.

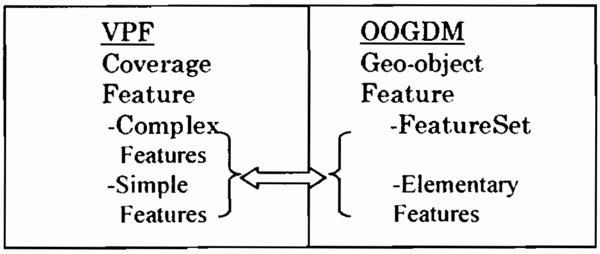


Figure 3. Levels of abstraction between VPF and OODGM

a. VPF employs the topological data structure representation. The nodes are stored as pairs of xy coordinates. Edges are ordered collections of two or more tuples and referenced by their endpoints. Polygons are modelled by referencing the first edge. Topology data are stored together with the geometry in the primitives' tables. For the topology data, only the connectivity information between nodes, edges and faces are provided. These topology data structures represent directed or undirected graphs that do not consider the exact geometry of their vertices.

b. OODGM uses a boundary representation to describe the geometry. The geometry of individual elementary feature is represented directly using xy coordinates. Points are indicated by their xy coordinates. Polylines, and polygons or holes in the regions are modelled directly using a set of xy coordinates. Thus, the geometry of the elementary features can be retrieved directly without computation. This kind of representation is termed object-centred. Besides, there is no support for the topology data.

4.2 Semantic Reconstruction For Establishing Equivalent Schema Semantically

Current approaches such as Jahnke et al. (1997), Daniel et al. (1993) and Behm (2001) involve semantic enrichment that is based on additional information extracted during reverse engineering process. The aim of semantic enrichment is to produce a semantically richer data model to generate target schemas. In contrast to these approaches, the approach persuaded in this study does not generate and modify the target object-oriented schema. This study applies semantic reconstruction rather than semantic enrichment to derive the mapping rules for the transformation. Semantic reconstruction is to resolve any ambiguities between the schemas of VPF and OODGM spatial databases. It performs both static and dynamic transformations between two databases. Based on the semantics of the schema in VPF, this research performs semantic reconstruction according to the schema in OODGM. This is to preserve semantics in both databases without modification during schema transformation. Geometry and topology of simple features are stored together in the primitives' tables in VPF, whereas only geometry of elementary features are represented in the geometric descriptors in OODGM. It is necessary to perform the semantic reconstruction so that the OODGM can accommodate both the geometric data and topology connectivity information being transformed from source database. Based on the above requirements and restriction, this study proposes that the semantic reconstruction is carried out by means of constructing a virtual data structure as indicated in Figure 4. This involves the steps of constructing a virtual data structure on the fly for those primitives in VPF, which do not have the equivalent representation in OODGM. This virtual data structure consists of interfaces to restructure the schema of source database on the fly before being mapped to target database. Then, there is a direct mapping from the virtual data structure to the spatial database in OODGM.

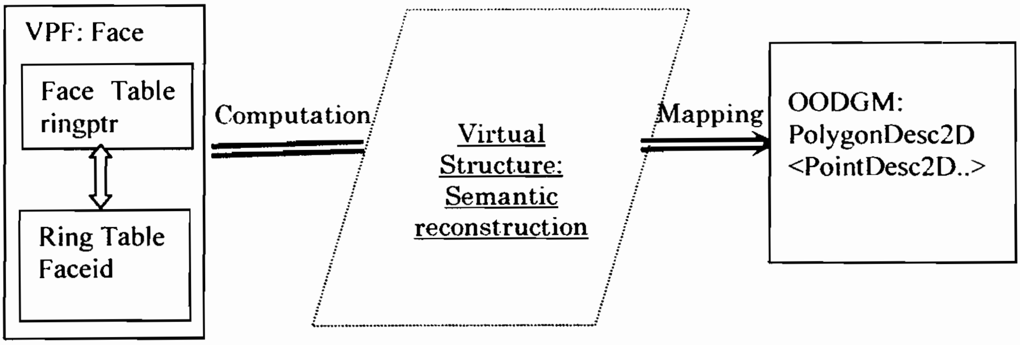


Figure 4. Semantic reconstruction between face and region representation in VPF and OODGM respectively.

4.3 Mapping Rules Between Schemas Of Relational And Object-Oriented Databases

Based on semantic reconstruction, this research derives the following mapping rules for the transformation of the spatial data of the three primitives from the database in VPF to the database in the OODGM. The concept is to map the primitives' tables in VPF to the child classes of elementary feature class in OODGM. Schema transformation involves a schema mapping of a source schema M^s to a target schema M^t by defining a transformation TR such that

$$TR(M^s) = M^t \qquad \text{Equation 1}$$

The transformations are a collection of rules that map from the source database to the target database. Together with semantics reconstruction, this study proceeds to map the schema of primitives in VPF to the respective ones in OODGM. It has to derive the mapping rules for all the primitives. Firstly, the study will determine the data structure that will be used to store a collection of transformed data in the target database such as array or list. It will explore the data structure of primitives' tables in VPF first. Each of the primitives' tables consists of row identities to identify the primary keys of the primitives, primitives' geometry and the topology information. They are stored in an ordered and sequential way according to the primary keys. They are actually the linear lists of data. Thus, equivalently, this study can adopt a linear data structure to represent the primitives' tables in OODGM as portrayed in Figure 4.

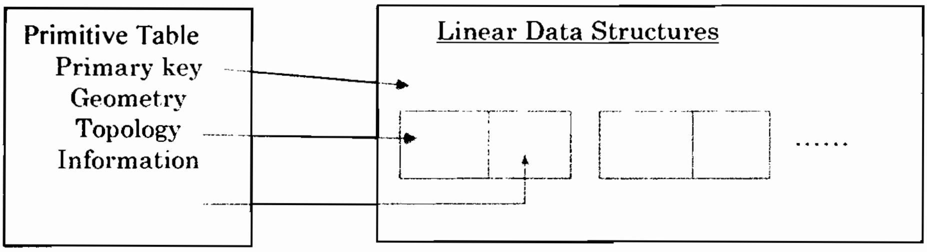


Figure 4. A linear data structure representation of VPF primitives' tables.

5 VIRTUAL VISIBILITY

Dynamic transformation involves interaction between classes. This implies visibility between classes. In object-oriented environment, visibility is implemented using inheritance and polymorphism concepts. Inheritance enables the collaboration between the two classes in such a way that the subclass can inherits the state and behaviour of the superclass. Polymorphism allows more than one implementation for the same event type that can gather data from different information sources. This polymorphism feature allows an object to dynamically change some of its events and to interact with them to perform specific computations.

Inheritance concept is efficient for simple relationships or systematic relationships. These relationships are modelled as simple or multiple inheritances respectively. However, inter-relationship between classes involves complex relationships in this study. It is especially so at level 3 topology. The relationships between classes are inter-related. These complex relationships cannot be modelled as simple or multiple inheritances. If inheritance is persuaded to model the complex relationships, inter-inheritance exists among the classes to perform dynamic modelling of transformation system. This leads to the problem of integrity. The dynamic transformation still looks very relational. It consists of "class-join" which is similar to "table-join" in relational environment as shown in Figure 5.

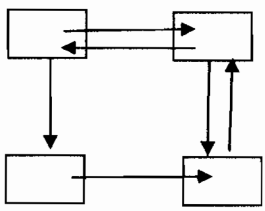


Figure 5. Inter-relationship between objects

This study proposes virtual visibility to model inter-relationship among classes. Since dynamic transformation is implemented using a chain of event-triggering operations, the study proposes to persuade behaviour concept to implement visibility among classes indirectly. This indirect visibility is termed virtual visibility. The approach is to incorporate object identities as the signatures of the methods of the classes. Through the signatures of the behaviour, any class can invoke its methods to interact with another class. In this way, classes can have a direct interaction among themselves virtually as shown in Figure 6.

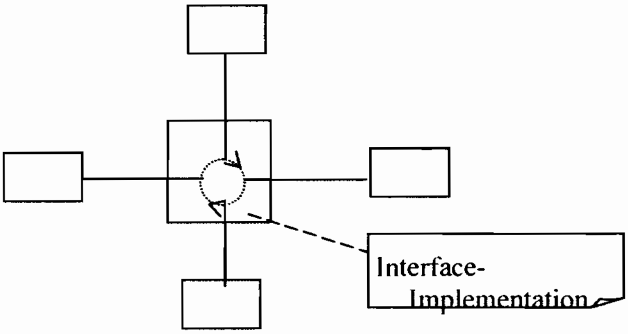


Figure 6. Virtual visibility between classes

6 CONCLUSION

The study produces dynamic schema transformation using object-oriented features.The approach persuaded in the schema transformation does not modify both the source and target databases. Instead, it involves semantic reconstruction which deals with dynamic modelling to derive mapping rules for schema transformation. In this way, organizations can involve in transformation from existing relational database to existing object-oriented database. The study implements virtual visibility concept using the behavioural data of software components. The approach is to registered software components as the signatures of the behaviour. Then, software components have virtual visibility among themselves by declaring their instances through method definitions. This has contributed significantly to the database management systems in view of the existing problems such as problem of integrity.

7 REFERENCES

Andersson. (1994). Extracting an Entity Relationship Schema from a Relational Database through Reverse Engineering. In P.Loucoulos. editor. *13th International Conference on Entity-Relationship Approach*. Volumn 881. 403-419.

Andreas Voigtmann. (1997) An Object-Oriented Database Kernel for Spatio-Temporal Geo-Applications. University Munster.

Behm. A. (2001). Migrating relational databases to object technology. PhD thesis. http://www.ifi.unizh.ch/cgi-bin/db2web/Library/thesis.d2w/report?1key=Behm-2001.

Chiang. R.H.. Barron, T.M. & Storey. V.C.. (1997). A framework for the design and evaluation of reverse engineering methods for relational databases, Data & Knowledge Engineering 12. 107-142. Elsevier Science.

Daniel A. Keim. Hans-Peter Kriegel. Andreas Miethsam. (1993). Integration of Relational Databases in a Multidatabase System based on Schema Enrichment. Institute of Computer Science. University of Munich.

Defense Mapping Agency. (1993). Military standard: Vector Product Format. Draft document no. MIL-STD2407. Dedense Mapping Agency. Fairfax, VA.

Egenhofer. M. J.. & Frank. A.U. (1992). Object-Oriented Modeling for GIS. *URISA Journal*. 4:2. pp. 3-19.

Frank. A. U.(1988). Requirements for a database management system for a GIS. Photogrammetric Engineering and Remote Sensing. 54, 1557.

Haivant, J-L. (2002). Introduction to Database Reverse Engineering. Laboratory of Database Application Engineering. University of Namer.

Herring, J.R., (1992) TIGRIS: A Data Model for an Object-Oriented Geographic Information System. *Computers and Geosciences* 18:4, pp. 443-452.

Jahnke, J.H, Schafer, W. & Zundorf, A.,(1996). A Design Environment for Migrating Relational to Object-Oriented Database Systems. *Proceeding of the International Conference on Softeware Maintenance*

Jahnke, J.H., Schafer, W. & Zundorf, A. (1997). Generic fuzzy reasoning nets as a basis for reverse engineering relational database applications. In *Proc. Of ESEC/FSE 1997*, Zurich, Switzerland. LNCS1301, Sprimger. http://citeseer.nj.nec.com/jahnke97generic.html

Petit, J-M., Toumani F., Boulicaut, J-F. & Kouloumdjian, J. (1996). Towards the reverse engineering of denormalized databases. In *international conference on data engineering*. IEE.

Intra-class Variability Modeling For On-line Signature Using Multi-resolution Approach

Liew Lee Hung[1] and Wang Yin Chai[2]
Geographical Information System
lhliew@fit.unimas.my[1], ycwang@fit.unims.my[2]

ABSTRACT

This paper discusses the findings of intra-class variability modeling for on-line signature. The current function approach of using a single resolution representation for on-line handwritten signature verification is unable to identify the inner variations of the features, which are caused by intra-class and interclass variability contained in the collected short signal. Such an approach might cause the acceptance of forged signatures that have similar patterns as the original and the rejection of genuine signatures that have high intra-class variability. Therefore, on-line signature features represented in single resolution is believed not detail enough to learn one's signature. This paper presents the concept of using multi-resolution approach by applying wavelet decomposition and fractal dimension in the proposed model and its overall performance. The discussion is based on the results of experiments conducted to evaluate the importance of multi-resolution representation of the developed prototype. With the achievement performance of an average improvement of 18% in genuine test verification rate and 7% in forged test verification rate compared to the single resolution approach, it proved that the intra-class variability is importance for on-line signature verification.

KEYWORDS

Wavelet decomposition, fractal dimension, multi-resolution approach, intra-class variability, on-line signature verification.

1 INTRODUCTION

Signature stability is one of the issues in the signature verification task. A great deal of variability can be observed in signatures according to country, age, time, habits, psychological, physical and practical conditions (Plamondon and Lorette, 1989). This has led to the difficulties of determination of the two types of signature variability that have been clearly distinguished: intra-class and interclass variability (Plamondon and Lorette, 1989).

Intra-class variability refers to the differences within genuine signatures produced by the same signer while interclass variability is the differences between genuine signatures, which are produced by two different signers. Therefore, the features selected should have minimum intra-class variability but high interclass variability. This is because minimum intra-class variability shows the stability among the signatures from the same signer while high interclass variability helps in differentiating the signatures from different signers. Hence, signature variability is believed to help in the signature verification task (Liew and Wang, 2002b).

However, it is hard to separate both the intra-class and interclass variability from the data collected since not every sample from the same signer has the same number of sample points and not everyone sign at the same time sequence (Dimauro et al., 1997). Furthermore, such instability features of an individual signature, which might be caused by the intra-class variability, can be large (Jain et al., 2002).

The current function approach of using a single resolution representation for on-line handwritten signature verification is unable to identify the inner variations of the features, which are caused by intra-class and interclass variability contained in the collected short signal (Liew, 2004). Such an approach only studies the outline of each features used and not consider the inner variation of each features which is not detail enough to study interclass and intra-class variability. It might cause the acceptance of forged signatures that have similar patterns as the original and the rejection of genuine signatures that have high intra-class variability.

This research study proposes an on-line signature intra-class variability modeling using multi-resolution approach, with wavelet as a multi-resolution analysis tool and fractal dimension as a method of signature complexity measurement to overcome the limitation of using a single resolution approach.

2 MULTI-RESOLUTION APPROACH TO ON-LINE SIGNATURE INTRA VARIABILITY

When the function approach is used, each signature feature is characterized in terms of a time function, which is possible to treat as a signal. In signal processing, a signal is frequently contaminated by noise when a signal is received after transmission over some distance (Walker, 1999). A signal can be represented by Equation 1.

(contaminated signal) = (original signal) + (noise) Equation 1

It is believed that signature features also have this same problem but the term *noise* in signal processing would be the variations in signature verification task. Therefore, a signature feature will have the form as Equation 2.

(signature feature) = (original signal) + (variations) Equation 2

Single resolution approach is unable to separate the small changes and the overall signal structure for further analysis either with or without applying any pre-processing techniques. This problem can be overcome when representing the signal in multi-resolution. Multi-resolution wavelet transformation can decompose a signal into different levels of low-pass and high-pass information (Deng et al., 1999). The low-pass (low frequency) information represents the main body of the original data while the high-pass (high frequency) information represents features that contain sharper variations (Deng et al., 1999). These advantages enable one signature signal to be presented into two different forms: low-pass information as global function, which represents the general signal pattern and high-pass information as local function, which represents the detail signal pattern. The detail signal pattern could not be obtained when using the single resolution approach. It is believed that from the detail signal pattern, the analysis of intra-class variability could be performed (Liew and Wang, 2002a).

The signature features of this research study (angle, pressure and speed) are decomposed into multi-resolution in order to study the pattern of sub-signals produced at each level, which are believed to represent intra-class and inter class variability. The wavelet is used as a multi-resolution analysis tool to decompose a signature feature into different levels of high-pass and low-pass information while fractal dimension is used as a pattern measurement of sub-signals to study the complexity of each produced sub-signal. The complexity of different levels of sub-signals represents the characteristic of the signature and it is believed that it contains the individuality of every signer. The fractal dimension indexes are used to form an optimal individual threshold. The test signatures will be verified according to the determined individual threshold. Such complexity index from fractal dimension indirectly simplifies the phase of comparing non-linear functions of different durations.

3 THE INDIVIDUAL CUMULATIVE MODEL THRESHOLD

The individual cumulative model threshold uses the multi-resolution decomposition concept. Referring to the wavelet multi-resolution theory, one signal can be transformed back by adding a certain level of approximation analysis with the same level of detail analysis and all the previous levels of detail analysis.

The standard deviations of the levels considered in one feature are summed to form the feature threshold (T_{f_i}) using Equation 3. Only the fourth level of approximation analysis is used with the first four levels of detail analysis. This is because the sub-signal of the third level of approximation analysis can be formed using the fourth level of approximation analysis and the fourth level of detail analysis. This condition also happens to the sub-signal of the second level of approximation analysis and the sub-signal of the first level of approximation analysis, which can be regenerated using their lower resolution sub-signals. The writer-dependent threshold (T_{w_i}) is the combination of the features thresholds with different weight as in Equation 4.

$$T_{f_i} = \sigma_{D1} + \sigma_{D2} + \sigma_{D3} + \sigma_{D4} + \sigma_{A4} \qquad \text{Equation 3}$$

$$T_{w_i} = \alpha T_{f_{u}} + \beta T_{f_{p}} + \gamma T_{f_{c}}, \ \alpha + \beta + \gamma = 1 \qquad \text{Equation 4}$$

The features thresholds are combined using weighted sum method. The weight of each feature is based on the verification rate of the reference signatures using the feature threshold. Each signer will have his/her own feature's weight (Liew, 2004).

4 INTRA-CLASS VARIABILITY TEST

Intra-class variability occurs within genuine signatures produced by the same signer. Therefore, the genuine test verification of some selected unstable signatures from the database is carried out in order to test the ability of single resolution approach and multi-resolution approach in analyzing such variability.

The signature samples shown in Figure 1 are genuine signatures from the original signer: a) signatures of Signer F, b) signatures of Signer K and c) signatures from Signer L. The difference between the signatures is obvious although they are from the same signer.

The genuine test verification rates of these three signers, shown in Table1 for Signer F, Table 2 for Signer K and Table 3 for Signer L, are used to prove the advantage of multi-resolution approach over single resolution approach in considering the intra-class variability.

Table 1. The results of genuine test verification between single resolution approach and multi-resolution approach for signatures of Signer F

Approach	Angle	Pressure	Speed	Combine
Single resolution	45%	45%	85%	40%
Multi-resolution	80%	80%	80%	90%

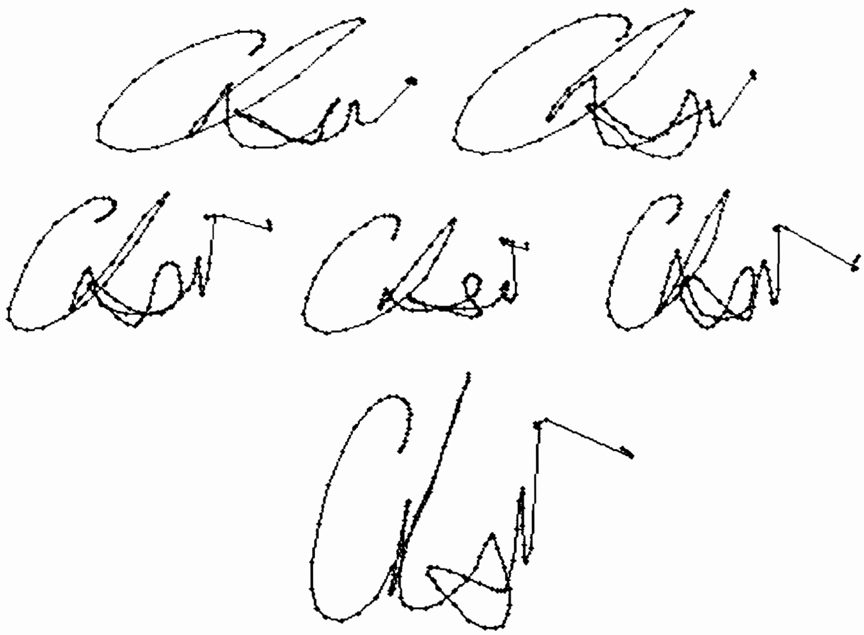
Table 2. The results of genuine test verification between single resolution approach and multi-resolution approach for signatures of Signer K

Approach	Angle	Pressure	Speed	Combine
Single resolution	40%	40%	55%	35%
Multi-resolution	90%	60%	80%	65%

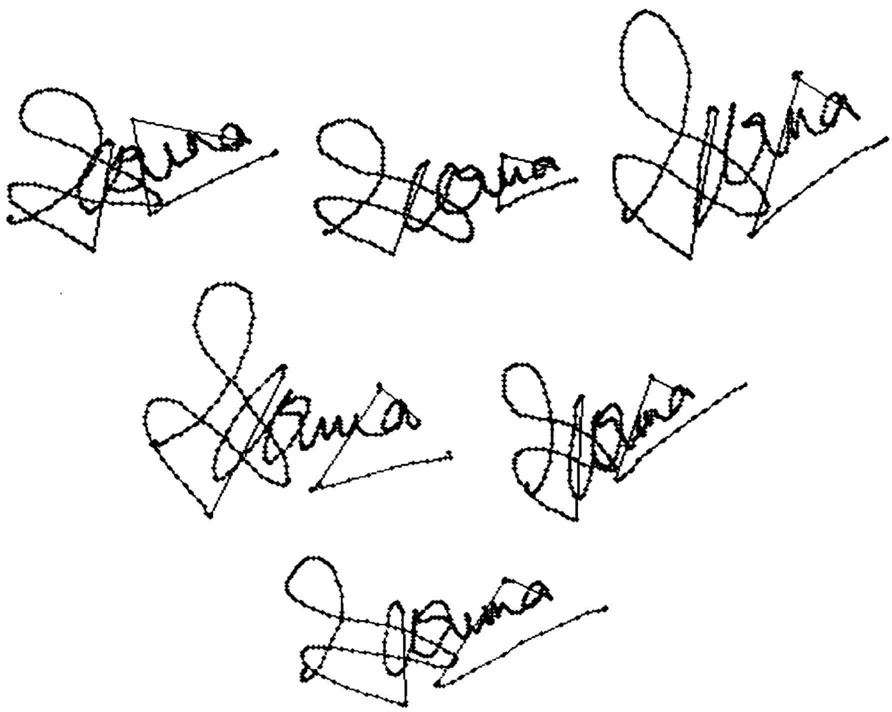
Table 3. The results of genuine test verification between single resolution approach and multi-resolution approach for signatures of Signer L

Approach	Angle	Pressure	Speed	Combine
Single resolution	45%	50%	65%	40%
Multi-resolution	85%	70%	80%	90%

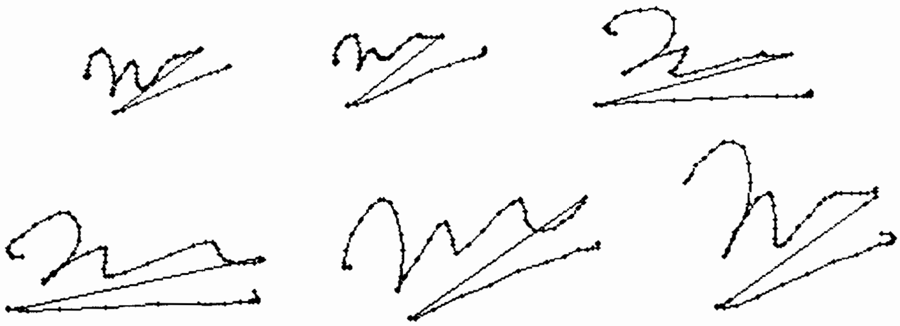
Based on the verification rates obtained, the genuine test verification rates for the above unstable signatures are very low when using the single resolution approach. However, the results were improved when using the multi-resolution approach, which considers the intra-class variability with the detail analysis performed. It is noticed that the multi-resolution approach improves the individual result of Signer F from 40% to 90%, Signer K from 35% to 65% and Signer L from 40% to 90%.



a) Six signature samples from Signer F



b) Six signature samples from Signer K



c) Six signature samples from Signer L

Figure 1. Examples of unstable signatures from the collected database

5 RESULTS AND DISCUSSION

Overall, this research used a total of 1500 genuine signatures (500 signatures as reference signatures, 1000 genuine test signatures) and 1000 forged signatures (500 random forged test signatures and 500 simple forged test signatures) as summarized in Table 4.

Table 4. Database for the proposed signature verification

Signature Usage	Number of signers	Number of signatures	Total number of signatures
Reference	50	10	500
Genuine test	50	20	1000
Forged test	50	20	1000

The same database has been used to carry out the signature verification using single resolution approach and multi-resolution approach. The features used for single resolution approach are the same three features of angle, pressure and speed for the developed prototype. The single resolution approach also uses fractal dimension to measure the feature patterns as done in multi-resolution approach.

The results obtained for different features of the two tests are as shown in Table 5. The result of the single resolution approach is compared with the result of the multi-resolution approach to evaluate the proposed model.

Table 5. The verification result obtained using single resolution approach and multi-resolution approach

	Angle		Pressure		Speed	
Verification Result	Genuine Test (%)	Forged Test (%)	Genuine Test (%)	Forged Test (%)	Genuine Test (%)	Forged Test (%)
Single resolution approach	58	58	55	63	59	63
Multi-resolution approach	75	52	70	74	75	48

Referring to the result of the proposed intra-class variability model of multi-resolution approach and the result of the single resolution approach, it is noticed that the proposed method only has lower verification rates in the forged tests for the signature feature of angle and speed. But their verification rates in the genuine tests are much higher than the verification rates from the single resolution approach. Signature verification usually not only depends on one signature feature; the more features being used the more moderate the result obtained (Rigoll and Kosmala, 1998). Therefore, these three signature features are combined to have a total verification rate.

The comparison of the overall verification result between the multi-resolution approach and single resolution approach is illustrated in Figure 2. In genuine test, the multi-resolution approach has a result of 78% while the single resolution has a result of 60%. In the forged test, the multi-resolution approach obtained a 70% verification rate while the single resolution obtained a 63% verification rate. Therefore, the total verification rate of the proposed model is better than the total verification rate of the single resolution approach in both genuine and forged tests. It showed that multi-resolution approach using wavelet analysis and fractal dimension in on-line handwritten signature verification, which emphasizes the variability within signatures' features and the signal pattern itself, could perform better compared to single resolution approach.

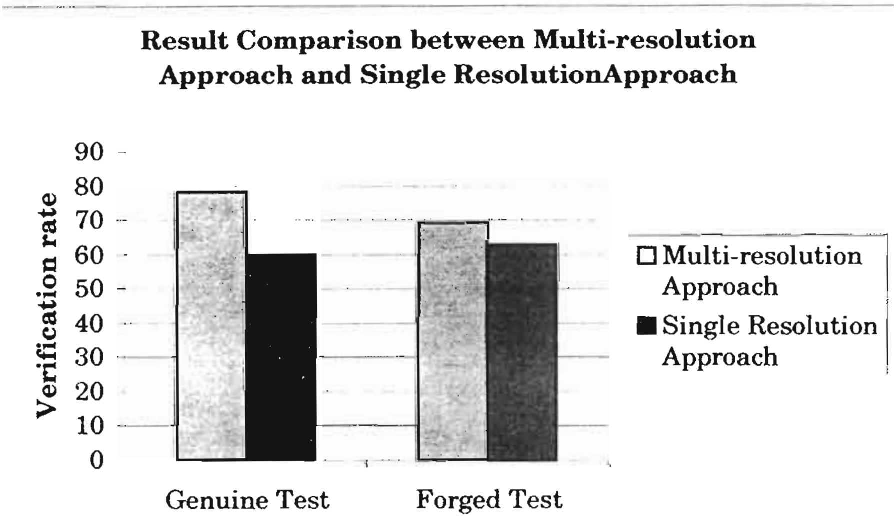


Figure 2. Comparison of the overall verification results of multi-resolution approach and single resolution approach

6 CONCLUSION

As a conclusion, this research has proven that the multi-resolution approach with concentration on the variation analysis provides better analysis of the signature features than single resolution approach and it is possible to use fractal dimension in measuring the complexity pattern for on-line handwritten signature verification. Therefore, it is valuable to include wavelet multi-resolution analysis and fractal dimension in interpreting intra-class variability. Wavelet multi-resolution analysis is suitable to investigate signature variations and fractal dimension is a potential pattern measurement for on-line handwritten signature verification. Thus, the integration of wavelet analysis inspired from the signal processing field and fractal dimension from pattern recognition into multi-resolution representation has been the core of the proposed the model of on-line signature intra-class variability for this research.

7 REFERENCES

Deng. P.S., Liao. H.Y.M., Ho. C.W. & Tyan. H.R. (1999) Wavelet-based off-line signature verification. Computer Vision and Image Understanding, 76(3): 173-190.

Dimauro. G., Impedovo. S. & Pirlo. G. (1997) Algorithms for automatic Signature verification. In Handbook of Character Recognition and Document Image Analysis (Bunke. H. & Wang, P. S. P.: eds.). Singapore: World Scientific Publishing Company, pp.605-621.

Jain. A.K., Griess. F.D. & Conell. S.D. (2002) On-line signature verification. Pattern Recognition. 35 (12): 2963-2972.

Liew. L.H. & Wang. Y.C . (2002a) Extracting signature feature using Wavelet and fractal in on-line signature verification. Proceedings of the Postgraduate Research Colloquium in Science and Technology 2002. pp. 83-85. February 2002.

Liew. L.H. & Wang. Y.C. (2002b) Signature feature extraction using Wavelet and fractal in on-line signature verification. Proceedings of the CoGRAMM' 02, National Conference on Computer Graphics and Multimedia 2002. pp. 348-353. October 2002.

Liew. L.H. (2004) A multi-resolution approach to on-line signature intra-class variability modeling. Faculty of Computer Science & Information Technology. Universiti Malaysia Sarawak. Malaysia. Unpublished Master Thesis.

Plamondon, R. & Lorette, G. (1989) Automatic signature verification and writer identification – The state of the art. Pattern Recognition (1989). 22: 107-131.

Rigoll. G. & Kosmala. A. (1998) A systematic comparison between on-line and off-line methods for signature verification with hidden Markov models. In Proceedings of Fourteenth International Conference on Pattern Recognition (Jain. A.K., Venkatesh, S. & Lovell, B.C.: eds.). vol. 2, pp. 1755-1757.

Walker. J.S. (1999) A primer on wavelets and their scientific applications, Boca Raton: CRC Press LLC. pp. 1-92.

An Enhannced Structural Spatial Query Retrieval Model

Lau Bee Theng[1] and Wang Yin Chai[2]
Geographical Information System
laubt@fit.unimas.my[1], ycwang@fit.unimas.my[2]

ABSTRACT

Structural spatial query retrieval is an active research area in spatial databases. The assessment of structural spatial query similarity is called structural similarity, also known as configuration similarity. This research developed an enhanced Structural Spatial Query Retrieval Model for spatial databases with the developed Spiral Web Model. The model was proven to be more effective than existing models in three main areas that are single similarity measure, improved object association, and object approximation free.

KEYWORDS

Structural spatial query, structural spatial similarity, and spatial retrieval

1 BACKGROUND

Similarity queries constitute an active area in spatial query processing. Currently there is an increasing interest for structural spatial similarity developing in the context of digital libraries, spatial databases and Geographic Information System. Structural spatial queries constitute a special form of content-based retrieval where the user specifies a set of spatial query variables and requests for all configurations of actual objects that match these variables. The problem of querying by structure has been around since the early stage stages of computer vision. There are many researches provide solutions to the structural similarity.

In the earlier stage, structural similarity based on qualitative spatial relations was introduced that entails a mechanism for representing and reasoning on spatial relations, but it also introduces uncertainties due to certain spatial queries can be rather fuzzy (Papadias and Sellis, 1994). The qualitative structural similarity measures included topological relations, cardinal directions and approximate distances (Sharma, Flewelling and Egenhofer, 1994). Then, structural similarity is quantified by using 2D space relation measures for topology, direction and distance with the aim to better represent the spatial knowledge quantitatively (Papadias and Delis, 1997). After that, fuzzy structural similarity was introduced to resolve the uncertainties introduced in the previous methods but the similarity measurements used remain the same (Papadias, Karacapilidis and Arkoumanis, 1999). Blaser (2000) implemented another structural query retrieval model for spatial databases. His structural spatial query is formulated from freehand sketches. His method shows the usefulness of structural query for object retrieval in spatial databases. His model uses the common structural similarity measurements that are topological, directional and metrical relations. He also added a non-spatial measure that is object geometry. He added an extra geometry measurement for shape similarity. In Cohn and Hazarika (2001), they model the qualitative spatial similarity measures with direction and distance, they added the non-spatial measures i.e. orientation and shape.

To deal with the structural similarity retrieval with time and processing effort constraints, there are many refinement models appear with spatial indexing (Papadias, Mamoulis and Delis, 1998; Papadias, Sellis, Theodoridis and Egenhofer, 1996) and heuristics i.e. simulated annealing (Papadias, Mantzourogiannis, Kalnis, Mamoulis, and Ahmad, 1999), constraints satisfaction (Papadias, Kalnis and Mamoulis, 1999), genetic algorithms (Papadias et. al., 1999), hill climbing algorithms (Papadias, 2000; Papadias et. al., 1999). In short, these refinements do not focus on improving the performance of structural similarity assessment but they are more concern on how to speed up and reduce the processing efforts in the retrieval. In SQbS (Blaser, 2000), no search heuristics and indexing are used. The performance of the pure structural similarity with multiple measures is not camouflaged with any external circumstances i.e. members of artificial intelligence. Hence the limitations of generic structural similarity assessment are uncovered.

The aforesaid models have a few similar characteristics. First, their structural similarity assessments are based on topological, directional and distance (or metrical) relations. Second, they neglect the importance of object geometry in the spatial query. These models approximate the objects in spatial query into Minimum Bounding Rectangles (MBR) where the topological relation between MBR does not necessarily coincide with the topological relation between the objects (Papadias, Sellis, Theodoridis and Egenhofer, 1996). Third, the structural similarity measurements utilize the association relation computation using $[(n*(n-1))*r]$ (Papadias, Karacapilidis and Arkoumanis, 1999) and $[((n*(n-1))/2)*r]$ (Blaser, 2000) where n represents the number of objects in a spatial query and r represents the number of similarity measurements.

1.1 Research Problem

The existing structural spatial query models use multiple spatial similarity measures: topology, direction and distance (or metrics) for structural spatial query retrieval in spatial databases. With the multiple measures (e.g. $r = 3$), time and processing efforts spent on the similarity assessment increase as compared to single measure (e.g. $r = 1$). Generally, the multiple spatial similarity measures are given weightage to give an overall similarity value for a structural query. However the similarity value indicated by individual measure is not sufficient to describe how similar a pair of objects as an overall when the value is analyzed separately. A proper integration of multiple measures is required.

The conventional model uses object association relation computation by $[(n*(n-1))*r]$. The computation increases by $((n*2)*r)$ for every additional object exists in a query. Consequently, this increases the complexity of structural similarity assessment, as the query has more objects, the relation computation increases by two times the number of added objects.

On the other hand, the objects in a structural spatial query are approximated into bounding rectangles that cause the query loosing its original geometrical settings. During the filtering step of spatial query retrieval, the bounding rectangles that are unlikely to satisfy the query are eliminated and a set of potential candidates are selected. Since bounding rectangles are only the approximations, they might cause the right objects being eliminated at the early stage of spatial query retrieval.