The 16-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision

in co-operation with

EUROGRAPHICS

W S C G ' 2008

POSTER Proceedings

University of West Bohemia
Plzen
Czech Republic

February 4 - 7, 2008

Honourary Chair
Wolfgang Strasser, University of Tuebingen, Germany

Co-Chairs
Steve Cunningham, California State University, USA
Vaclav Skala, University of West Bohemia, Czech Republic

Edited by
Steve Cunningham
Vaclav Skala

WSCG'2008 POSTER Proceedings

Editor-in-Chief: Vaclav Skala
University of West Bohemia, Univerzitni 8, Box 314
CZ 306 14 Plzen
Czech Republic
skala@kiv.zcu.cz
Managing Editor: Vaclav Skala

Author Service Department & Distribution:
Vaclav Skala - UNION Agency
Na Mazinách 9
322 00 Plzen
Czech Republic

Printed: University of West Bohemia

Hardcopy: *ISBN 978-80-86943-17-6*

WSCG 2008

International Programme Committee

Baranoski,G. (Canada)
Bartz,D. (Germany)
Benes,B. (United States)
Biri,V. (France)
Bouatouch,K. (France)
Coquillart,S. (France)
Csebfalvi,B. (Hungary)
Cunningham,S. (United States)
Davis,L. (United States)
Debelov,V. (Russia)
Deussen,O. (Germany)
Ferguson,S. (United Kingdom)
Goebel,M. (Germany)
Groeller,E. (Austria)
Chen,M. (United Kingdom)
Chrysanthou,Y. (Cyprus)
Jansen,F. (The Netherlands)
Jorge,J. (Portugal)
Klosowski,J. (United States)
Lee,T.-Y. (Taiwan)
Magnor,M. (Germany)
Myszkowski,K. (Germany)
Pasko,A. (United Kingdom)
Peroche,B. (France)
Puppo,E. (Italy)
Purgathofer,W. (Austria)
Rokita,P. (Poland)
Rosenhahn,B. (Germany)
Rossignac,J. (United States)
Rudomin,I. (Mexico)
Sbert,M. (Spain)
Shamir,A. (Israel)
Schumann,H. (Germany)
Skala,V. (Czech Republic)
Teschner,M. (Germany)
Theoharis,T. (Greece)
Triantafyllidis,G. (Greece)
Veltkamp,R. (Netherlands)
Wu,S. (Brazil)
Zara,J. (Czech Republic)
Zemcik,P. (Czech Republic)

WSCG 2008

Board of Reviewers

Adzhiev,V. (United Kingdom)
Albuquerque,G. (Germany)
Andreadis,I. (Greece)
Andujar,C. (Spain)
Aspragathos,N. (Greece)
Aveneau,L. (France)
Bair,A. (United States)
Bamidis,P. (Greece)
Baranoski,G. (Canada)
Battiato,S. (Italy)
Bellon,O. (Brazil)
Benes,B. (United States)
Bernhard,M. (Austria)
Biri,V. (France)
Bittner,J. (Czech Republic)
Borchani,M. (France)
Bottino,A. (Italy)
Bouatouch,K. (France)
Bouville,C. (France)
Bruni,V. (Italy)
Buehler,K. (Austria)
Carmen Juan-Lizandra,M. (Spain)
Casciola,G. (Italy)
Coleman,S. (United Kingdom)
Comba,J. (Brazil)
Coquillart,S. (France)
Csebfalvi,B. (Hungary)
Daniel,M. (France)
Davis,L. (United States)
de Geus,K. (Brazil)
Debelov,V. (Russia)
Delrieux,C. (Argentina)
Deussen,O. (Germany)
Dingliana,J. (Ireland)
Doleisch,H. (Austria)
Duce,D. (United Kingdom)
Durikovic,R. (Slovakia)
Egges,A. (Netherlands)
Eisemann,M. (Germany)
Erbacher,R. (United States)
Feito,F. (Spain)
Felkel,P. (Czech Republic)
Ferguson,S. (United Kingdom)
Fernandes Marcos,A. (Portugal)
Fernandes,A. (Portugal)
Flaquer,J. (Spain)
Flórez,J. (United States)
Forest,V. (France)
Franke,T. (Germany)
Fuchs,R. (Austria)
Galin,E. (France)
Gallo,G. (Italy)
Galo,M. (Brazil)
Galvez,A. (Spain)
Garcia-Alonso,A. (Spain)
Giannini,F. (Italy)
Gonzalez,P. (Spain)
Groeller,E. (Austria)
Gudukbay,U. (Turkey)
Gumbau,J. (Spain)
Gupta,A. (United States)
Gutierrez,D. (Spain)
Habel,R. (Austria)
Hanak,I. (Czech Republic)
Haro,A. (United States)
Hasler,N. (Germany)
Havemann,S. (Austria)
Havran,V. (Czech Republic)
Hernandez,B. (Mexico)
Horain,P. (France)
House,D. (United States)
Huang,C. (Taiwan)
Chaudhuri,D. (India)
Chen,M. (United Kingdom)
Chen,T.F. (Canada)
Chin,S. (Korea)
Chover,M. (Spain)
Chrysanthou,Y. (Cyprus)
Iglesias,A. (Spain)
Iwanowski,M. (Poland)
Janda,M. (Czech Republic)
Jansen,F. (Netherlands)
Joan-Arinyo,R. (Spain)
Jorge,J. (Portugal)
Kavan,L. (Ireland)
Keller,K. (United States)
Kimmel,B. (Canada)
Kjelldahl,L. (Sweden)
Klosowski,J. (United States)
Knight,M. (United Kingdom)
Kohout,J. (Czech Republic)
Kolcun,A. (Czech Republic)

Krishnaswamy,A. (Canada)
Lee,T. (Taiwan)
Lewis,J. (United States)
Lin,W. (Taiwan)
Lintu,A. (Germany)
Liu,S. (Taiwan)
Maciel,A. (Brazil)
Magnor,M. (Germany)
Maierhofer,S. (Austria)
Mandl,T. (Germany)
Matkovic,K. (Austria)
Mattausch,O. (Austria)
McMenemy,K. (United Kingdom)
Michoud,B. (France)
Mokhtari,M. (Canada)
Mollá Vayá,R. (Spain)
Montrucchio,B. (Italy)
Mould,D. (Canada)
Muigg,P. (Austria)
Muller,H. (Germany)
Myszkowski,K. (Germany)
Nedel,L. (Brazil)
Netto,M. (Brazil)
Odone,F. (Italy)
Ogayar,C. (Spain)
Pan,Z. (China)
Papadakis,P. (Greece)
Pasko,A. (United Kingdom)
Pasko,G. (Japan)
Patera,J. (Czech Republic)
Patil,S. (India)
Paulin,M. (France)
Pedrini,H. (Brazil)
Pelikan,J. (Czech Republic)
Peroche,B. (France)
Pili,P. (Italy)
Platis,N. (Greece)
Plemenos,D. (France)
Podgorelec,D. (Slovenia)
Pratikakis,I. (Greece)
Puig,A. (Spain)
Puppo,E. (Italy)
Purgathofer,W. (Austria)
Ramires Fernandes,A. (Portugal)
Redon,S. (France)
Reichinger,A. (Austria)
Renaud,c. (France)
Ripolles,O. (Spain)
Ritschel,T. (Germany)
Rojas-Sola,J. (Spain)
Rokita,P. (Poland)
Rosenhahn,B. (Germany)
Rossignac,J. (United States)
Rudomin,I. (Mexico)
Sanna,A. (Italy)
Sbert,M. (Spain)
Segundo,M. (0)
Segura,R. (Spain)
Sellent,A. (Germany)
Semwal,S. (United States)
Shamir,A. (Israel)
Schmaltz,C. (Germany)
Schmidt,T. (Germany)
Schneider,B. (United States)
Schumann,H. (Germany)
Singh,M. (United States)
Sips,M. (United States)
Sirakov,N. (United States)
Stanchak,R. (United States)
Stroud,I. (Switzerland)
Suarez Rivero,J. (Spain)
Svoboda,T. (Czech Republic)
Sýkora,D. (Czech Republic)
Teschner,M. (Germany)
Theoharis,T. (Greece)
Theußl,T. (Austria)
Tokuta,A. (United States)
Torrens,F. (Spain)
Triantafyllidis,G. (Greece)
Tytkowski,K. (Poland)
Van Basten,B. (Netherlands)
Van Laerhoven,T. (Belgium)
Vanecek,P. (Czech Republic)
Vasa,L. (Czech Republic)
Veeraraghavan,A. (United States)
Veiga,L. (Portugal)
Vergeest,J. (Netherlands)
Verlinden,J. (Netherlands)
Vitulano,D. (Italy)
Wagner,S. (Germany)
Wan,T. (United Kingdom)
Weiss,G. (Germany)
Wu,S. (Brazil)
Wuethrich,C. (Germany)
You,S. (United States)
Zach,C. (Austria)
Zachmann,G. (Germany)
Zalik,B. (Slovenia)
Zambal,S. (Austria)
Zara,J. (Czech Republic)
Zemcik,P. (Czech Republic)
Zhu,Y. (United States)

Contents

Pajorová,E,. Hluchý,L., Anthes,Ch.: POSTER: 3D Geo-visualization Service for Grid-oriented Applications of Natural Disasters 1

Hebaishy,M.A.: POSTER: Optimized Daugman's Algorithm for Iris Localization 5

Takeuchi,T., Saito,H., Mochimaru,M.: POSTER: 3D-Face model Reconstruction Utilizing Facial Shape Database from Multiple Uncalibrated Cameras 9

Kang,B., Cho,S.: POSTER: Effective Object Sorting Technique for Developing 3D GUI Including Translucent Objects 13

Zeng,Ch., Wang,H.: POSTER: User's Dynamic Interest Description Model Based on Cross-Media 17

Zeng,Ch., Liu,X., Wang,H.: POSTER: Video High-level Semantic Recognizing 21

Thibault,G., Devic,C., Fertil,B., Mari,J.L., Sequeira,J.: POSTER: Shape Indexes: from 2D to 3D. Application to Cells' Nuclei Classification 25

Suzuki,H.: POSTER: Introduction of Animation Assignment in Graphic Science Education Making Use of CG Application of Data Describing Type 29

Oslejsek,R.: POSTER: Virtual Scene as a Software Component 33

García,R., Perales,J.F.: POSTER: Adaptive Templates in Biometrical Authentication 37

POSTER: 3D Geo_visualization Service for Grid oriented Applications of Natural Disasters

[1]Eva Pajorová
[1]Institute of Informatics
Slovak Academy of Sciences
Dúbravská cesta 9
Bratislava 845 07, Slovakia
utrrepaj@savba.sk

[1]Ladislav Hluchy
[1]Institute of Informatics
Slovak Academy of Sciences
Dúbravská cesta 9
Bratislava 845 07, Slovakia
hluchy.ui@savba.sk

[2]Christoph Anthes
[2]Graphics and Parallel Processing
Johannes Kepler University
Linz, Austria
canthes@gup.uni-linz.ac.at

ABSTRACT

Our basic aim is to design a framework that will define a unified way of cooperation between Grid visualization applications and visualization clients. Working with the Grid technology implies increased complexity on one hand, and user demands for high interactivity on the other. The duration of the computations in the Grid environment being exceedingly long, the user naturally wants to see the intermediate results and requires means to modify the running computations if the intermediate results are not satisfactory. Framework is building of lot of tools. This paper describes concrete only one of tool included in framework – 3D Geo_visualization tool for visualization of the running Grid applications to rendering their intermediate or final results in the client application.

Keywords

Virtual Reality, GRID applications, Natural Disasters, 3D_visualization tool

1. INTRODUCTION

Grid applications, by their nature, are very time consuming. Unfortunately, the existing Grid infrastructure supports particularly batch applications; the support for interactive applications is only in its experimental phase. However, the time-consuming nature of Grid applications makes the users to demand the support to be able to check the state of the running application and modify it if the progression is not satisfactory. This encompasses two things; first, to be able to monitor the progress of the computation, by the means of application statistics, and second, to be able to see the actual intermediate results the application has provided so far. Unfortunately, the current situation is far less pleasurable: the user has to wait until the whole application finishes and only then he can run the visualization job, which will present him the data in the form of imagery. Having to wait until the application finishes only to see the computation went wrong can be very irritating.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Copyright UNION Agency – Science Press, Plzen, Czech Republic.

Majority of Grid applications can be extended to contain a subsystem that would, on a demand or in regular time intervals, provide some information about the status of the computation; the phase of the computation or a progress indicator are typical examples of such information. There is a category of tasks in which the time of computation cannot be estimated in advance (some optimalization of tasks, cipher breaking, etc.); for such applications, the progress indicator can be based on statistically obtained data about similar tasks. If the application is running on more cluster nodes (which is nearly almost a rule in the Grid environment) the user might want to observe the computational load and the degree of completeness at each node. In such event it is considerable to create visualization services, which are able to prepare visualization outputs for clients quickly. This paper is describing such visualization services developed and tested in our institute, concrete a 3D visualization tool for GRID oriented Natural Disasters applications. The purpose of 3D Viz. tool is to model and display intermediate or final results of various simulations of natural disasters like fire spread in time, its intensity and erosion or floods in time or landslides as well. The output of the tool is various scenes of terrain by different simulations outputs. Output of the tool can also be the files representing the virtual reality of natural disaster and also files, which are generated as input for VR-Systems. 3D tool was tested with output from applications, which were solved in our institute and

also by data from applications of the MEDIGRID project. Particular demonstrations in this article are from fire in part near Mediterranean Sea at south. Exactly in Marseille industrial port at southeast. The second is from large fire in Krompla region [2], which is part of Slovensky Raj. Flood demonstrations are from Povazie, the region around the river Vah. For applications oriented on Grid computing was tool included to Visualization service (see Fig. 8).

2. RELATED WORK

There are several tools designed for visualizing natural disasters [7]. However, these tools lack common criteria for visualizations, unification of data formats, and a common approach to visualization. The integration of formats and a common solution for visualization of natural disasters, which has been solved in our visual service, allowed to created a visualization service for grid applications. As the applications computing natural disasters work with a large amount of data, grid computing is used frequently. Providing a visualization service could be considered a big contribution. This has been proved by increasing interest of experts working on international projects.

3. ARCHITECTURE OF TOOL

3D visualization tool consists of several modules. The modules are divided into three groups according to what kind of output is generated. The functionality of each group is represent by the appropriate schema. (schema 1, schema 2, schema 3).

- Models of terrain or model of environment.
- Models of simulations.
- Virtual reality models

The first group of modules is used for the creation of 3D terrain or environment models. The following modules has been created: CONVERTOR, TERRAIN, IVFIX and CONVERTTOWRL. Process represent Schema 1 (see Fig. 1).

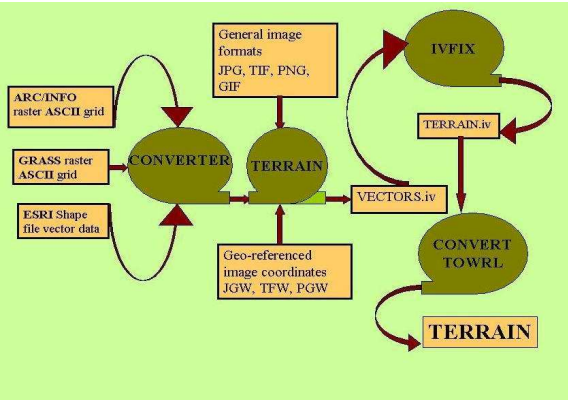


Figure 1.Schema 1.

Most frequently used formats of input data for terrains are: ARC/INFO raster ASCII grid, GRASS raster ASCII grid , ESRI Shape files vector data . Data can be exported to the TERRAIN module. Terrain module computes vectors and creates quad mesh for terrain models. Then modified IVFIX module creates Indexed face sets. Indexed face sets files are converted by module CONVERTTOWRL, to VRML. We created CONVERTTOWRL module according to the requirements of 3D tool. The final model of terrain is a VRML file TERRAIN.wrl. Examples of terrain created by the tool: *Marseille industrial port* (see Fig. 2)

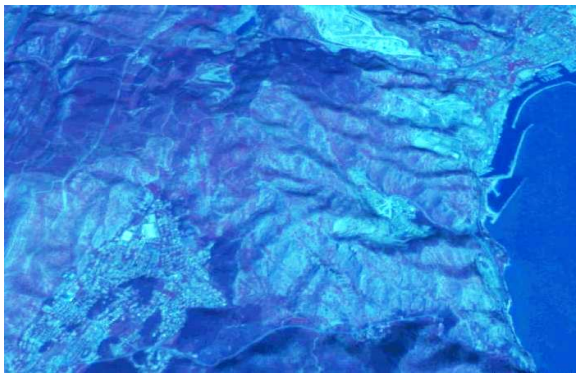


Figure 2. Marsaille industial port_Terrain

Models of simulations

The second group of modules of the 3D tool is designed to generate simulation models. The following modules were created: CONVERTER, CLASSIFIER, COLOR SOTER, TIME SORTER and VIRTUAL SORTER and GENERATORS for static or dynamic outputs. Creation process representet Schema 2. (see Fig. 3).

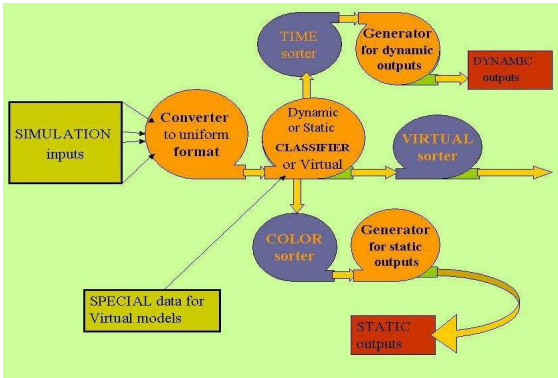


Figure 3. Schema 2.

The Input data for the second part of the 3D tool are output data from simulators in different formats. The module CONVERTER converts the data into a single format. The data input to the CLASSIFIER module. CLASSIFIER sorts the data and exports them to the

 ISBN 978-80-86943-17-6

COLOR sorter and the TIME sorter and some of them, which are used for Virtual models to VIRTUAL sorter. The outputs Data obtained from the sorters are used as input for GENERATORS. Final visualization consists of a terrain model and of some simulation models. For example visualization of *fire spread in time* (see Fig. 4.) consists of TERRAIN.wrl and Firespread.wrl files.

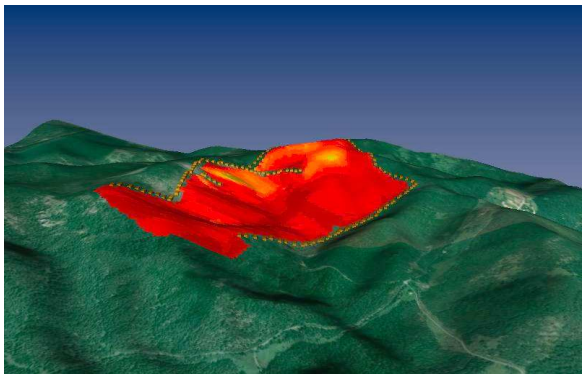


Figure 4. Dynamic model – Fire spread

Virtual reality models

The third group of modules is designed to generate models suitable for virtual reality applications. The following modules were created: DATA SORTER, VIRTUAL TERRAIN SORTER, VIRTUAL SIMULATION SORTER and several GENERATORS. Process of creation Virtual models represent Schema 3. (see Fig. 5).

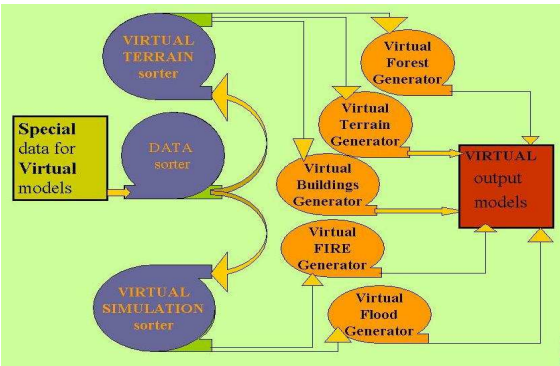


Figure 5. Schema 3.

As input Special data values can be for example values from Grown maps (Grown maps owners are Forestry). These data provide sufficient information for building up a concrete virtual forest composed of quads. For creating concrete forest we used Sibyla system [3]. Then by using 3D tool the same terrain, which has been covered by ortophotomap before is now covered by just prepared concrete virtual forest. Finally red faces representing fire were replaced by virtual fire textures using Virtual fire generator module (see Fig. 6). Since models from the 3D tool are files in VRML format. They can be shown by VRML browsers and viewers as well as by CAVE for (see Fig. 7). In collaboration with Institute of Graphics and Parallel Processing of Johannes Kepler University Linz we are able to view some of the Virtual scenes in a CAVE system.

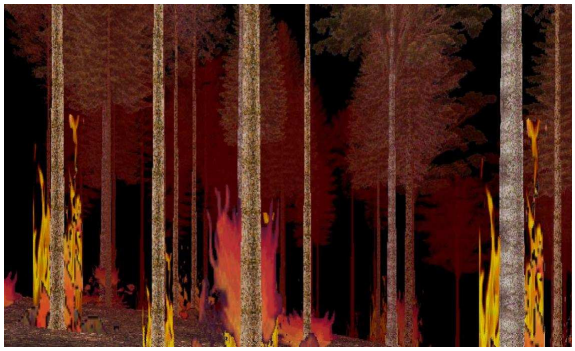


Figure 6. Virtual Fire in Krompla hill

Using the inVRs framework for collaborative immersive visualization

In order to provide a deep insight of the displayed datasets the *inVRs* framework is used for immersive visualization [4]. *inVRs* offers the possibility to render virtual environments in stereoscopic 3D graphics on a variety multi-display output devices like the CAVE [5], curved screens or powerwall installations. The framework consists of independent modules for navigation, interaction and network communication, which are interconnected via its system core. Inside the core databases host objects and user information, like the datasets provided by preprocessing steps. To render the graphical display of *inVRs* applications OpenSG [6] is used as a scene graph. Users are able to navigate trough the scenes depending on their chosen travel methodology and observe the scene as if they are standing inside. Different interaction methodologies can be chosen at the startup of the application. In the case of these visualizations switching of the different simulation steps is performed either automatically by starting an animation sequence or manually by pressing buttons to advance in the simulation. A valuable feature of the framework allows geographically dislocated users to collaboratively view and manipulate the datasets. To implement collaborative visualization the network module of the framework interconnects a set of clients in a peer-to-peer communication topology. The users are visualized by abstract avatars, which indicate the viewing orientation of the remote users. Position tracking is incorporated for pointing at distinct areas of the datasets. The gathered transformation of the users input device is transmitted via the network and visualized as a virtual pointer at the remote participants' site. To increase the immersion in the VE simple shaders have been

 ISBN 978-80-86943-17-6

developed to improve the display quality of the water areas of the datasets. Additionally shaders are used to increase depth perception. The vertex normals of the terrain datasets are scaled in the vertical axis, which leads to a higher contrast if lighting is applied in the scene.

Figure 7. Visualization in CAVE

4. VISUALIZATION SERVICE

The Visualization Service is a common job submission service used to run the modeling by scripts. The service will produces visualization outputs and make them available them for the portal. The results of 3D visualizations require an additional VRML browser plugin installed. It is represent Schema 4. (see Fig. 8). Visual service was used and evaluated during the MEDIGRID project. In a future it can be incorporated in related projects of the domain of disasters management

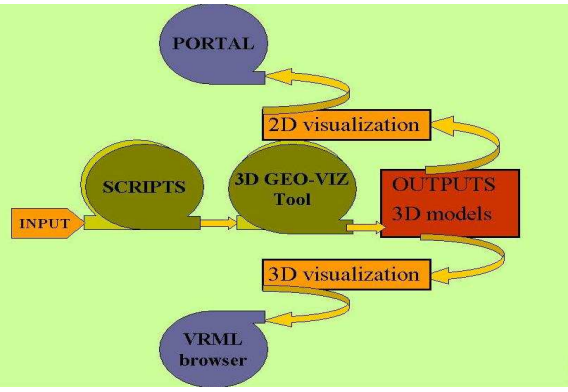


Figure 8. Schema 4.

5. CONCLUSION

Our 3D visualization tool has several approaches. It supports the most frequently used file formats. The 3D tool offers different levels of complexity for the visualization of natural disasters ranging from fast and simplified display up to scenes with a high polygon count. Output files are in format VRML (Virtual Reality Modelling Language), which are displayable by commonly available applications like Internet browsers, Freewrl, Blaxxun contact, Cortona client, Openwrl, VRED as well as in the CAVE. The tool can be used as a virtual service for a huge variety of GRID oriented disaster applications [1].

6. ACKNOWLEDGMENTS

This work was partially supported by projects: 3DSPATIAL No. 53s07 and by the EU project MEDIGRID EU 6FP RTD GOCE-CT-2003-004044. And also is supported by RPEU project GRID_tools No. 0024-06 and VEGA project No. 2/7098/27.

7. REFERENCES

[1] Šimo, B., Ciglan, M., Slížik, P., Mališka, M., Hluchý, L.: *Core services of heterogeneous distributed framework for multi-risk assessment of natural disasters.* International Conference on Computational Science – ICCS 2006, May 28-31, 2006.

[2] Ján Glasa, Eva Pajorová, Ladislav Halada, Peter Weisenpacher Animation of For est Fire Simulation. International Workshop on Environmental applications and Distributed Computing.- EADC October16 – 17, 2006 Bratislava, Slovakia.

[3] Algoritmus and software solution of thinning models for Sibyla grown simulator. Fabrika M., Durovsky J. Journal of forest science 51, 2005 (10) 431 – 445.

[4] Christoph Anthes and Jens Volkert. “inVRs - A Framework for Building Interactive Networked Virtual Reality Systems.” In International Conference (HPCC), pages 894-904, Munich, Germany, September 2006.

[5] Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. Defanti, Robert V. Kenyon, and John C. Hart, “The CAVE: Audio Visual Experience Automatic Virtual Environment,” Communications of the ACM, Vol. 35, No. 6, June 1992, pp. 65-72

[6] Dirk Reiners. „OpenSG: A Scene Graph System for Flexible and Efficient Realtime Rendering for Virtual and Augmented Reality Applications.“ PhD thesis, Technische Universität Darmstadt, Mai 2002.

[7].Hille, M., Goldammer, J.G.: Dispatching and modelling of fires in Central European pine stands: New research and development approaches in Germany. Proc. of workshop, Agronomic Institute of Chania, Crete, Greece (2001), 59-74.

POSTER: Optimized Daugman's Algorithm for Iris Localization

Dr. Mohamed A. Hebaishy
National Authority for Remote Sensing and Space Science
Gozif Titp St., Elnozha Elgididah.
Egypt (11769), Cairo
Ma_hebaishy@yahoo.com

ABSTRACT

Iris localization is considered the most difficult part in iris identification algorithms because it defines the inner and outer boundaries of iris region used for feature analysis. Several researches were made in the subject of iris finding and segmentation. The main objective here is to remove any non-useful information, namely the pupil segment and the part outside the iris (sclera, eyelids, skin). R. Wildes used Hough transforms to detect the iris contour. Daugman proposed an integro-differential operator to find both the pupil and the iris contour. Daugman's method is claimed to be the most efficient one. This paper proposes an implementation for Daugman's algorithm, which was found incompatible with visible light illuminated images. Then this paper proposes new algorithm for solving this problem.

Keywords
Biometrics, Human Identification, Iris localization, Daugman's Algorithm

1. INTRODUCTION

The human iris begins to form in the third month of gestation and the structure is complete by the eight month, even though the color and pigmentation continue to build through the first year of birth. After that, the structure of the iris remains stable throughout a person's life, except for direct physical damage or changes caused by eye surgery. The iris hence parallels the fingerprint in uniqueness but enjoys a further advantage that it is an internal organ and less susceptible to damages over a person's lifetime [1]. It is composed of several layers which gives it its unique appearance. This uniqueness is visually apparent when looking at its rich and small details seen in high resolution camera images under proper focus and illumination. The iris is the ring-shape structure that encircles the pupil, the dark centered portion of the eye, and stretches radially to the sclera, the white portion of the eye see Figure 1, it shares high-contrast boundaries with the pupil but less-contrast boundaries with the sclera.

The iris identification system is to automatically recognize the identity of a person from a new image by comparing it to the human iris patterns annotated with identity in a stored database[2]. A general iris recognition system is composed of four steps. Firstly, an image containing the user's eye is captured by the system. Then, the image is preprocessed to normalize the scale and illumination of the iris and localize the iris in the image. Thirdly, features representing the iris patterns are extracted. Finally, decision is made by

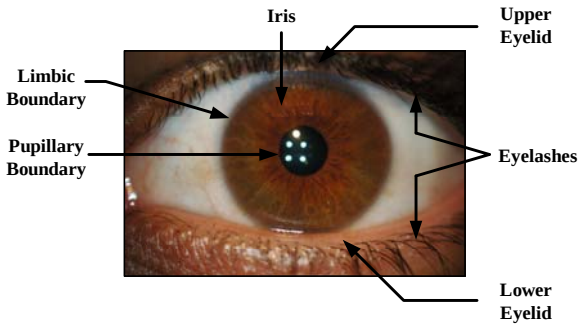


Figure 1. Eye image expected to be used in the iris recognition system

The iris identification system is to automatically recognize the identity of a person from a new image by comparing it to the human iris patterns annotated with identity in a stored database[2]. A general iris recognition system is composed of four steps. Firstly, an image containing the user's eye is captured by the system. Then, the image is preprocessed to normalize the scale and illumination of the iris and localize the iris in the image. Thirdly, features representing the iris

patterns are extracted. Finally, decision is made by means of matching. There are four key parts in the iris recognition system: iris image acquisition, preprocessing, feature extraction, and classifier design[3].

In a world where we will increasingly do business with parties we've never met, and might never meet, authentication will become as integral a part of the transaction as the exchange of goods and tender. The robustness of iris recognition makes it ideal for authenticating parties to commercial transactions, to reduce fraud in applications like check-cashing and ATMs, unauthorized activity in applications like treasury management, and in future, to ensure non-repudiation of sales, or to provide Letter of Credit and other authentication services in an electronic commerce environment[4].

Daugman has shown that iris patterns have about 250 degrees of freedom, i.e. the probability of two eyes having the same iris texture is about 1 in 7 billion. Even the 2 irises of an individual are different thereby suggesting that iris textures are independent of the genetic constitution of an individual. Iris recognition has been successfully deployed in many large scale and small scale applications.

Iris localization is considered the most difficult part in iris identification algorithms because it defines the inner and outer boundaries of iris region used for feature analysis[5]. The main objective here is to remove any non-useful information, namely the pupil segment and the part outside the iris (sclera, eyelids, skin). R. Wildes used Hough transforms to detect the iris contour. Daugman proposed an integro-differential operator to find both the pupil and the iris contour. Daugman's algorithm is claimed to be the most efficient one. After analyzing The Daugman's iris locating and pointing out the some limitations of this algorithm, this paper proposes optimized Daugman's algorithms for iris localization.

2. DAUGMAN'S ALGORITHM:

Daugman's algorithm is based on applying an integro-differential operator to find the iris and pupil contour[3].

The equation is as follows:

$$\max(r, x_o, y_o)\left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r, x_o, y_o} \frac{I(x,y)}{2\pi r} ds \right|$$

Equation 1. Daugman's Integro-Differential Equation

Where: x_0, y_0, r_0: the center and radius of coarse circle (for each of pupil and iris). $G_\sigma(r)$: Gaussian function. Δ_r : the radius range for searching for. $I(x, y)$: the original iris image.

$G_\sigma(r)$ is a smoothing function, the smoothed image is then scanned for a circle that has a maximum gradient change, which indicates an edge. The above algorithm is done twice, first to get the iris contour then to get the pupil contour. It worth mentioning here the problem is that the illumination inside the pupil is a perfect circle with very high intensity level (nearly pure white). Therefore, we have a problem of sticking to the illumination as the max gradient circle. So a minimum pupil radius should be set. Another issue here is in determining the pupil boundary the maximum change should occur at the edge between the very dark pupil and the iris, which is relatively darker than the bright spots of the illumination. Hence, while scanning the image one should take care that a very bright spot value could deceive the operator and can result in a maximum gradient. This simply means failure to localize the pupil. The following experimental results have been getting using UPOL database.

3. OPTIMIZED DAUGMAN'S ALGORITHM:

As a solution to this problem, modification to the integro-differential operator is proposed to ignore all circles if any pixel on this circle has a value higher than a certain threshold. This threshold is determined to be 200 for the grayscale image. This ensures that only the bright spots – values usually higher than 245 – will be cancelled.

Another solution we considered is to treat the illumination by truncating pixels higher than a certain threshold – bright spots – to black. But this method failed in many images, this is because when the spot hits the pupil the illumination spreads on the pupil so as we treat the illumination spots it will leave behind a maximum change edges that can not be determined and the operator will consider it the pupil boundary . The sequence of the Algorithm procedure is cleared in the flowchart shown in Figure 2.

Figure 3 shows the experimental results each daugman's algorithm and optimization Daugman's algorithm.

 ISBN 978-80-86943-17-6

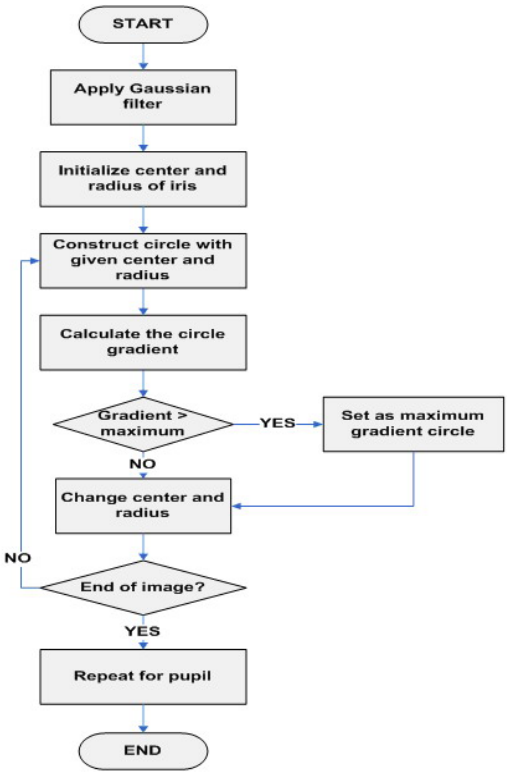


Figure 2 Flowchart of optimized Daugman's localization algorithm operation.

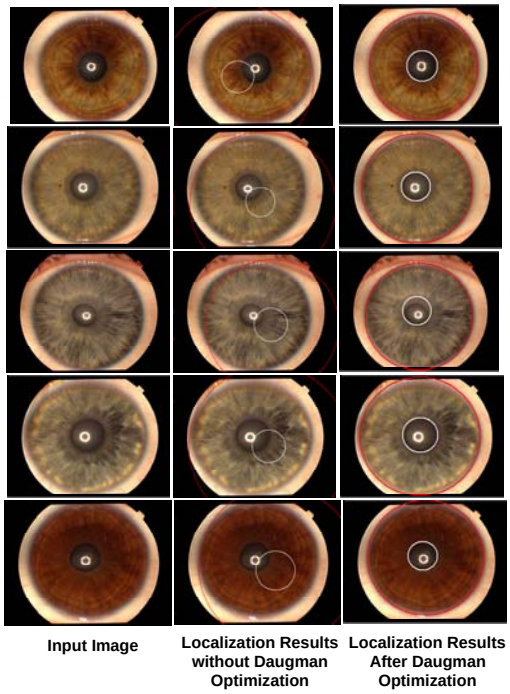


Figure 3 : Localization results

Database	Number of samples	Daugman's Algorithm		Daugman's optimization	
		Fill	Success	Fill	Success
UPOL	384	75%	25%	0%	100%

Table 1. Experimental results

The proposed algorithm is tested by applying it on UPOL database that includes about 384 images for 128 persons the localization successful percentage was 100%.

4. CONCLUSION:

The above studies show that the iris-locating algorithm based on integro-differential operator suffers from bright spots of the illumination inside the pupil , so the optimized Daugman's algorithm overcome this problem an gives a successful results for iris localization process.

5. ACKNOWLEDGMENTS

I would like to express my appreciation to the Egyptian space program at NARSS for their support and encouragements.

6. REFERENCES

[1] J. G. Daugman, "High confidence visual recognition of person by a test of statistical independence," IEEE Trans.PAMI 15, 1148-1161 (1993).

[2] J. G. Daugman, "The importance of being random: statistical principles of iris recognition," Patern. Recognition 36, 279-291 (2003).

[3] Yong Zhu, Tieniu Tan and Yunhong Wang, "Biometric Personal Identification Based on Iris Patterns" National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences, P. R. China, 2003

[4] http://www.unisys.com/news/releases/2001 /jun/ 06218031.asp

[5] J. G. Daugman, "How iris recognition works," IEEE Trans. Circuits and Syst. for Video Tech. 14(1), 21-30 (2004).

POSTER: 3D-Face model Reconstruction Utilizing Facial Shape Database from Multiple Uncalibrated Cameras

Toshio Takeuchi
Keio University
3-14-1 Hiyoshi Kohoku-ku
223-8522, Yokohama, Japan
takeuchi@ozawa.ics.keio.ac.jp

Hideo Saito
Keio University
3-14-1 Hiyoshi Kohoku-ku
223-8522, Yokohama, Japan
saito@ozawa.ics.keio.ac.jp

Masaaki Mochimaru
Digital Human RC, AIST
2-41-6, Aomi, Koto-ku
135-0064, Tokyo, Japan
m-mochimaru@aist.go.jp

ABSTRACT

In recent years, a surveillance camera has come to be attached in various places from a rise of the consciousness to security. Since the surveillance cameras are installed in variety of place, it is possible to take a picture of the same person from multiple uncalibrated cameras though it is asynchronous. In this article, we propose a method for reconstructing a face shape from multiple-view images taken with non - synchronous multiple cameras. In this method, we do not directly reconstruct the shape, but estimate a small number of parameters which represent the face shape. The parameter space is constructed with Principal Component Analysis of database of a large number of anatomical face shapes collected for different people. From the input multiple view images, the region of the face and three feature points on the face are manually extracted. Then the facial pose is estimated by optimizing the evaluation based on the silhouette shape, appearance, and the position of the feature points. According to the facial pose, the parameters representing the facial shape are also estimated by optimizing the same evaluation function. Those optimizing procedures are repeated for obtaining the facial shape for the object face captured with the non-synchronous multiple cameras. The experimental results demonstrate the effectiveness of the proposed method. Since the database used in this paper consists of anatomically aligned shape data, we can obtain anatomical shape of the face, which is suitable to represent the identity of each person.

Keywords

Asynchronous, Multi-view, 3D-Face model, Face shape database, PCA, Anatomical, pattern recognition

1. INTRODUCTION

In recent years, surveillance cameras have been attached in various places from a rise of the consciousness to security. By using surveillance cameras, we might be able to prevent unauthorized people entering into secured area. Surveillance cameras can be used not only as such a criminal deterrence purpose but also for the purpose of recording the action history of the people. For such a purpose, the identification of the person taken in images of the surveillance cameras is one of the important technologies. The surveillance cameras are installed in variety of places, so it is possible to take pictures of the same person from multi view cameras although it is asynchronous. Using the multiple viewpoint images, the reconstruction accuracy improves more than the reconstruction from one image. However, accurate position and poses of the capturing cameras should be estimated, which sometimes makes the computing time become huge. As a method for reconstructing the face shape from the image, there is a method named Active Shape Model that uses the data base with man's head shape information analyzed in PCA and that presumes shape from a peculiar space of the low level. There are researches of the Active Shape Model. Some are using the PCA to 3D shape and texture information and estimate the face shape [Vet99], [Vet02], [Vet03], others fit two dimensional vertex model to the image, and compare that from three dimensional shape model [Yan04], [Yan05]. There is a method using AR-toolkit for estimating face pose and position, and entering the characteristic points manually, then estimating three dimensional face shape [Saito06]. But using AR-toolkit, the scene

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Copyright UNION Agency – Science Press, Plzen, Czech Republic.

used is limited. This research reconstructs three dimensional face model by using that the direction of the face is already-known.

In this paper, we will propose a novel method for recovering face shape from multiple viewpoint images that are taken with asynchronous uncalibrated multiple cameras. Our method estimates both face direction and three dimensional face shape at the same time. We use the anatomy face shape database[Saito06]. Adopting PCA to the database, our method express the face shape by small number of parameters of about 20 dimensions. By estimating the parameter, we can reconstruct face shape. To estimate the face shape, the estimation of face direction is necessary. Our method use random face shape that can be calculated by face database to estimate the face direction and then estimate the face shape. By repeating this face direction estimating and face shape estimating, we reconstruct the three dimensional face shape from asynchronous multi view cameras.

The data base used with this paper is made for an anatomy study, so that we can easily get the anatomy specific vertex after reconstructing the face shape. So, reconstructed three dimensional model can be used not only for the person identification but also for getting face information from pictures took by multi view cameras installed in a free position.

2. PROPOSED METHOD

In our method, we adopt the way to compare the model image described by a few principal ingredients with input image, optimize the face shape and get the answer. Thus, it is a final purpose of this method to estimate the face direction and three dimensional face shape that fit the input image best. To estimate that, we use three evaluated value, facial contour information, specific points information [Sai06] and appearance information. We describe the evaluated value of appearance. We choose the image that is the nearest the front direction among all images given as an input by using the result of face direction estimation. Afterwards, texture mapping is done to three dimensional model by the texture of an image that is the nearest the front at that time. Images in other three directions are synthesized by the three dimensional model which is texture mapped.

We show the outline of the proposed method in Fig.1. First, estimate the rough face direction by using the input image. Next, estimate the three dimensional face shape by using face direction. Then, repeating the face direction estimation and the three dimensional face shape estimation, the face direction and three dimensional face shape when evaluated value converge is assumed to be the face direction and three dimensional face shape of the input image.

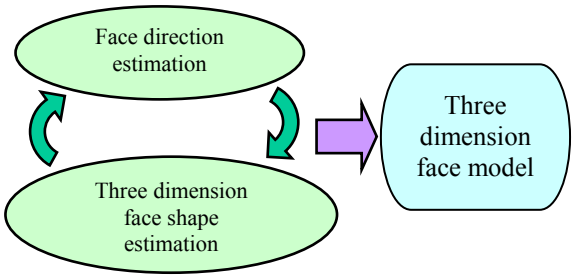


Figure 1. Outline of proposed method

3 . EXPERIMENT AND RESULT

3.1 Error evaluation using plastic model with known shape

To evaluate the accuracy of the three dimensional face model reconstructions, we use a plastic face model with known shape that is the average shape of the face database. We take the picture of the model by non-synchronous multiple cameras, and adopt this method to reconstruct the model.

The reconstruction accuracy can be evaluated by comparing the vertex of reconstructed shape with the correct answer shape, because correct answer shape is already-known. If difference between two become smaller after adopting our method, it can be said that near shape can be made by our method.

We use 4 input images, Fig.**2** shows the input image.

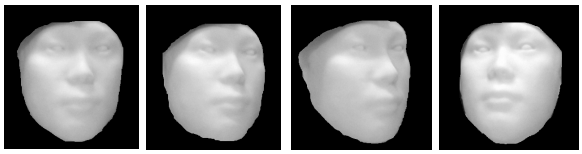
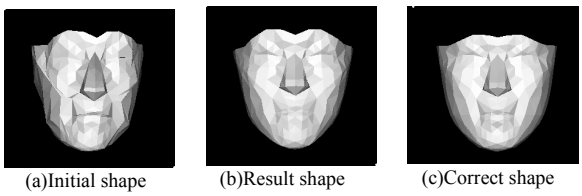


Figure 2. Input images for plastic model with known shape

Fig.**3 (b)** shows the result of recovered shape form the input images shown in Fig.**2**. Fig.**3**(a) indicates the initial shape used for the recovery of the shape in the proposed method. Fig.**3** (c) is the correct shape of the object face. By comparing those images shown in Fig.**3**, it is obvious that the recovered shape shown in Fig.**3** (b) is getting close to the correct shape shown in Fig.**3** (c).



(a)Initial shape (b)Result shape (c)Correct shape

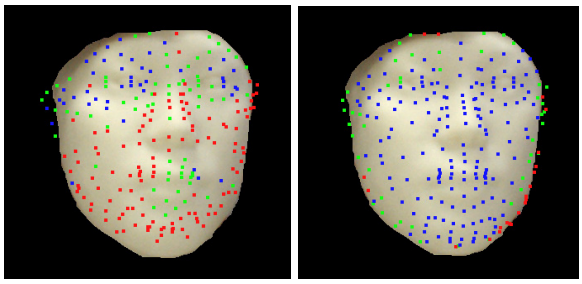
Figure 3. Result image

We also compare the error of the initial shape with the error of the recovered shape. Table.**1** shows the

average error of the shape. Face area means the whole area of face and face front means the area forward from eyes. Table.**1** says that the error of the face front is more than the error of the face area. This is also indicated by Fig.**4**, which shows the distribution of error margin. Red shows the error margin equal or more than 8mm, and green shows the error margin less than 8mm and more than 5mm, blue shows the error margin equal and less than 5mm. Fig.**4**(a) shows initial shape error, and Fig.**4**(b) result shape error. Fig.**4**(a) shows that the top error margin has extended to the entire face, especially the outline of the shapely nose and the cheek. Fig.**4**(b) shows that the error margin decreases overall.

	Initial shape	Recovered shape
Face area	9.83	5.01
Face front	9.26	4.36

Table. 1. Average Error(mm)



(a) Initial shape (b) Recovered shape

Figure 4. Error margin

To examine the average and variance of 3D face model reconstruction error, we used 4 images from 10 input images. There are 210 kinds of combinations that take 4 images from 10 input images. We calculated the error of all 210 combinations and then, calculated the average and variance of 210 combinations. Table.2 shows the result of error. The average error of face area is 4.87mm, and face front is 4.76mm. The data shows face front is better accuracy than face area, but the extension of variance is larger than that of face area. It may be because of the delicacies of face parts. facial contour changes little by direction estimation error, but facial part changes large.

	average	variance
Face area	4.87	0.202
Face front	4.76	0.376

Table. 2. Average and variance

3.2 Performance evaluation using real face

We took a picture of five actual people, and estimate the direction and three dimensional model of the input image. We use the images that are assumed as an asynchronous multiple viewpoints as input images. Fig.**5** shows the example of the input images. Some expressions are different and the direction and the position of the face to the camera have changed into each image as shown here. We clipped the face area, and used it as gray scale.

Figure 5. Example of the input image

Fig.**6** shows input images and result shapes. Steps above are the input images, and below are result shapes. The input image was taken of a picture from various directions, and only an image near the front was taken out here for the comparison. It is understood that the shape that exists in the feature of each face such as nose, eye origin, mouths have been optimized.

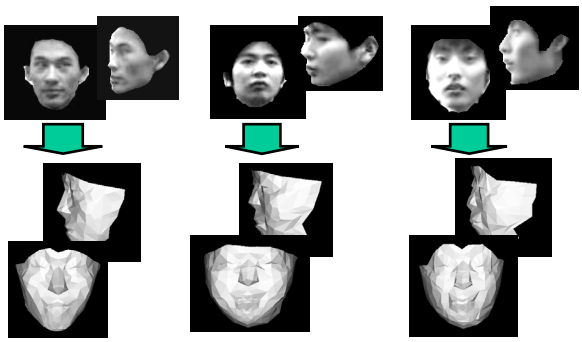


Figure 6. Comparison of image

(a)No.0 (b)No.1 (c)No.2 (d)No.3

Figure 7. Input image

Fig.**7** shows the input image of silhouette experiment. Fig.**8** shows comparison of silhouette of recovered face shape with silhouette of input image. The silhouette of the recovered face shape is closer to the input image silhouette than the initial shape. In Fig.**8**, the green part is a common of silhouette of the input and silhouette made from the model. A blue part is a part only of the silhouette of the input. A yellow part is a part only of the silhouette made from the model.

The protruding area is large in an initial model from the silhouette of the input image. By the direction and three dimensional model of the input image improving since the face direction estimation and the processing of three dimensional model estimation are repeated, it is understood that the area of the silhouette part of the protruding model decreased

from the silhouette of the input image. As a result, it is understood that the direction and three dimensional model of the input image improved by repeating the processing of this method.

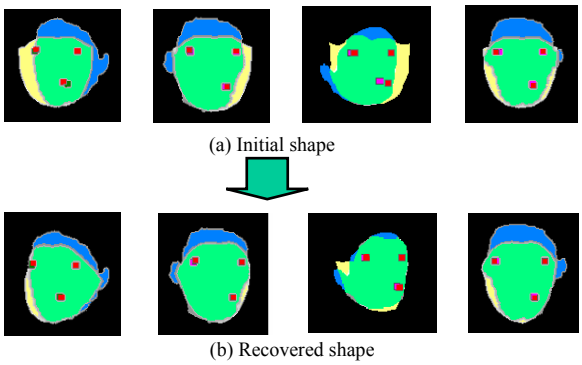


Figure 8. Comparison of the silhouette

Fig.**9** shows comparison of the texture mapping image of recovered shape with the texture mapping image of initial shape. Initial result means the result of the texture mapping on the initial three dimensional model. Reconstructed result means the result of texture mapping on the reconstructed model after repeating the direction estimation and three dimensions model estimation of the face. The texture mapping results for the initial shape are distorted more than the texture mapping results for the reconstructed shapes. This means that the accuracy of the direction estimation by the initial face is bad, the first three dimensional model estimation are thought to be similarly bad accuracy, and the texture mapping result, too. The distortion decreases in reconstructed result, and it is understood that it is near the correct answer image. As The accuracy of face direction estimation and face shape estimation become better by having repeated the direction estimation and the model estimation of three dimensions of the face, the texture mapping result improved as a result, too.

4. CONCLUSION

In this research, we propose the face reconstructing method by using the face image taken from an asynchronous multiple viewpoints. In the image, the face direction and face shape is unknown. And we use the anatomy face shape data base for reconstruction. We examine the process of optimization by this method.

In this method, we don't use that the direction of the face is already-known. We proposed the method for obtaining both by repeating the face direction estimation and the three dimensional face model estimation. In the experiment, we examined the appearance of optimization. It was shown for the distortion of the rotation by texture mapping to decrease, and to be able to make three dimensional models which looked like the object.

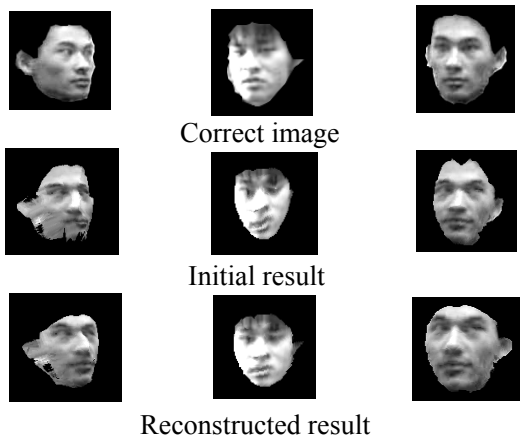


Figure 9. Optimization of the texture mapping

With the advantage that direct three dimensional shape information is obtained after it reconstructs it by using the database that contains a lot of anatomy shape information, it has the possibility to be able to use for the person attestation by the surveillance camera etc. in the future

5. ACKNOWLEDGMENT

This work was supported by Grant-in-Aid for Scientific Research (C)

(17500119).

6. REFERENCES

[Saito06]H. Saito, Y. Ito, M. Mochimaru, Face Shape Reconstruction from Image Sequence Taken with Monocular Camera using Shape Database, Proc. 14th International Conference on Image Analysis and Processing (ICIAP 2007), pp165-170, Sept. 10-14, 2007

[Vet99]T.Vetter, V.Blanz, A morphable model for the synthesis of 3D-faces, ACM SIGGRAPH, 1999

[Vet02]T.Vetter, S.Romdhani, V.Blanz, Face identification by fitting a 3D morphable model using linear shape and texture error functions, European Conf. on Computer Vision, vol.V4, pp.3-19

[Vet03]T.Vetter, V.Blanz, Face recognition based on fitting a 3D morphable model, IEEE Pattern Analysis and Machine Intelligence, vol.25, 2004

[Yan04]S.Yan, L.Zhang, H.Zhang, Y.Hu, Automatic 3D reconstruction for face recognition, IEEE International Conference on Automatic Face and Gesture Recognition, 2004

[Yan05]S.Yan, L.Zhang, H.Zhang, W.Gao, D.Jiang, Y.Hu, Efficient 3D reconstruction for face recognition, IEEE Pattern Recognition, vol38, pp.787-798, 2005

POSTER: Effective Object Sorting Technique for Developing 3D GUI Including Translucent Objects

Byungkwon Kang
Samsung Electronics
416, Maetan-3Dong, Yeongtong-Gu,
443-742, Suwon-city, Gyeonggi-Do, Korea
bk76.kang@samsung.com

Sunghee Cho
Samsung Electronics
416, Maetan-3Dong, Yeongtong-Gu,
443-742, Suwon-city, Gyeonggi-Do, Korea
shane.cho@samsung.com

ABSTRACT

In developing a user interface, it is very important to provide usability, intuitiveness and convenience. To accomplish these factors, 3D user interface can be one good solution. Recently, many kinds of embedded devices have been developed and utilized for various areas of human life; for example, Digital TVs, game consoles, MP3 players, etc. A GUI makes these devices easier to use and more effective. But, their poor hardware performances often prevent developing a high quality visual effect on their GUIs. A visual effect using alpha blended object is one of the typical examples of this kind that limits the development of 3D applications on an embedded hardware. In this paper, we propose an effective object sorting technique that can be used to develop a 3D GUI including many translucent objects.

Keywords

3D graphics, 3D GUI, Real-time rendering, Alpha blending, Object sorting, Translucent object.

1. INTRODUCTION

Alpha blending is a very useful technique that can be utilized in various areas of 3D graphics in order to generate more elegant visual effects. To composite colors of several translucent objects naturally, not only the rule for color composition but also the order of compositing objects must be selected carefully. These days, 3D graphics APIs such as OpenGL and Direct3D support blending of translucent objects' colors by using alpha channels very well. But, depth sorting which determines the order of composition for accurate blending of colors is not supported on current graphics hardware and APIs. Therefore, an effective sorting technique for correct composition of colors in an order of objects' distances is indispensable to high quality 3D graphics.

In this paper, we propose an effective object sorting technique that can be utilized for developing a 3D GUI including many translucent objects. In designing this technique, we defined some restrictions and conditions that are often found in developing a 3D GUI. These factors are very useful to simplify the method of determining spatial relations between 3D objects.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Copyright UNION Agency – Science Press, Plzen, Czech Republic.

This paper is organized as follows. First, in Chapter 2, we introduce previous works about depth sorting methods and in Chapter 3, we summarize the preconditions and restrictions that are often found in developing 3D GUIs including translucent objects. In Chapter 4, we describe our depth sorting technique and summarize the performance and effectiveness in Chapter 5. Finally, we conclude this research in Chapter 6.

2. RELATED WORK

Depth ordering of 3D objects is a classic problem in various areas of computer graphics. There have been many types of research done to sort the primitives in object space [SBGS69, NNS72]. But their performances largely depended on the number of objects to be used. Therefore, there were many attempts to avoid the dependency on the number of objects [SBM94, KLNR97, SL98, WMS98, BOS04].

If the spatial positions of objects or polygons are not changed, a space partitioning data structure can be the most effective solution for ordering 3D objects. A binary space partition (BSP) of a set of objects is a recursive convex subdivision of space. It was designed to compute a visibility order among polygonal primitives used in graphics applications by

Fuchs et al. [HKN80]. The BSP tree has been a general solution and widely used for hidden surface removal, shadow generation, ray tracing, real-time 3D games, and so on. Most of the BSP algorithms work well only in static environments. But, traditional BSP algorithms were not suitable for a dynamic environment where the position of objects can be changed dynamically. Torres [Tor90] and Agarwal et al. [AGMV97] presented dynamic BSP data structures that are able to compute visibility ordering of polygons in a dynamic scene.

3. PRECONDITIONS AND RESTRICTIONS

In this chapter, we defined some preconditions and restrictions that are often found in implementing 3D GUIs. The preconditions and restrictions that will be considered in developing a 3D application focused on this research are as follows.

a. All objects in 3D space have a boundbox which has a shape of rectangular parallelepiped.

b. There is no intersection between each object.

c. It is developed on a dynamic environment that all spatial relations between objects change every time frame.

d. Object sorting is performed per each object using their boundbox.

All of these preconditions and restrictions are based on a premise that our research aims to develop a useful depth sorting method for 3D GUIs. Because our object sorting method is based on these conditions, our method is not suitable for some common 3D applications. But, the conditions described in this section are common features that can be found in most graphics user interfaces and they are very useful to simplify the object sorting process.

4. DESCRIPTION OF THE NEW SORTING TECHNIQUE

In this chapter, we describe a new sorting technique for accurate alpha blending proposed in this paper.

Finding overlapped objects

For the first step, find all objects overlapped with other ones. To make the order of composition of all objects in the scene, the spatial relations between object pairs must be determined. The process of finding overlapped objects in the eye coordination space is very simple. Because all boundboxes are shaped as a rectangular parallelepiped, up to 3 faces of a box are visible from the view point. Using these visible faces, we can construct a quad or hexagon using vertices on visible faces. After then, by applying a 2D overlapping test to these polygons we can easily find overlapping objects.

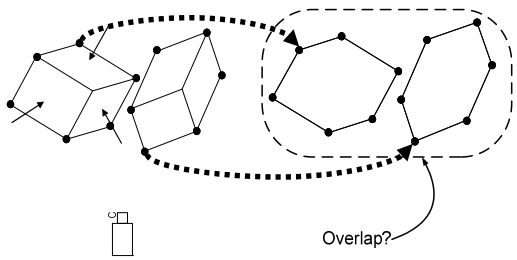


Figure 1. Projecting boundboxes into 2D polygons.

Constructing 2D polygon from boundbox

To construct a 2D polygon, we need to find the sequence of vertices that are constructing the contour of visible region of a boundbox. To achieve this process effectively, we designed an easy way to find the vertex sequence that is constructing the contour of visible faces of a boundbox.

At first, we assigned a fixed order of indexes to all faces and each vertices of a boundbox. As illustrated in Figure 2, all faces and vertices have uniquely and sequentially assigned indexes. And the boundboxes are transformed into the eye coordinates. After this transformation process, we can find visible faces from the camera easily by comparing 4th coefficient of plane equation of each transformed faces.

Third, we must find the sequence of vertices that are constructing the contour of visible faces. Because all faces and vertices have fixed indexes, we can find which vertices are included in the contour vertex list earlier on run-time when the visible faces are found. On run time, we can find the contour vertex list very quickly by making an index that is calculated by finding indexes of visible faces and searching its vertex list from the look up table.

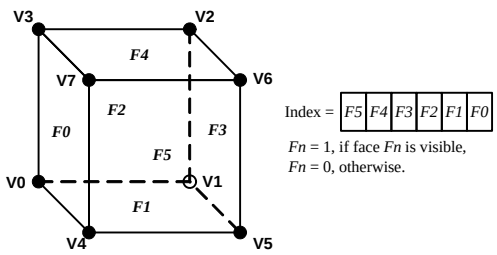


Figure 2. Making an index from visible faces.

After constructing projected polygon, the overlapping test can be performed very fast and effectively because this process is performed in 2D domain. There are many overlapping test techniques for 2D polygons and this test can be done easily [WMS98].

 ISBN 978-80-86943-17-6

Finding front-back object

For each object which is overlapping with other objects, the spatial relation with the overlapped object must be discerned. To make an accurate alpha blended result, we must find which object is occluding the other object carefully and exactly. In this paper, we propose a simple and effective technique to find the spatial relation between overlapped objects based on the preconditions and restrictions described in Chapter 3.

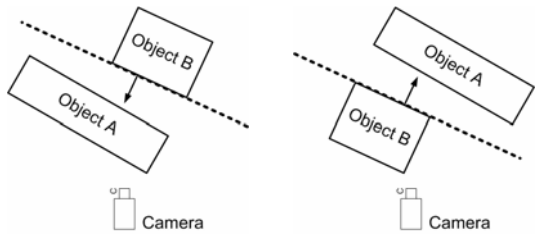
This technique is based on the restriction that all boundboxes must be shaped as a rectangular parallelepiped. From this restriction, we can determine the spatial relation between two boxes by finding the number of planes that all vertices of another object are placed in front of it. There are 3 cases between 2 non-intersecting boundboxes.

a. A places in front of only 1 plane of B.

b. A places in front of 2 or 3 planes of B.

c. There is no plane that A places in front of it.

Let's consider the first case. In this case, as illustrated in Figure 3, we can decide which object is the front object by discerning the spatial position of view point related to the selected plane. As illustrated in Figure 3(a), when the view point is placed in front of the plane, object A is placed in front of object B. Otherwise, when the view point is behind the plane as see in Figure 3(b), object B becomes the front object.



a) front case b) behind case

Figure 3. Spatial relation between selected plane and camera.

In cases except the first case, it is difficult to find front or back objects by using only 1 plane. But it is possible to find nearer object by comparing the nearest distances of vertices because the two objects are explicitly overlapped in the view of camera. After finding front and back objects between overlapped object pair, the front object is inserted to the front list of the back object and the back object is inserted to the back-list of the front object.

Rendering the objects in correct order

After front-back relations of all overlapped object pairs are discerned, the rendering order can be decided by using the result of discrimination. This process is performed based on a simple rule that any object cannot be rendered before its back-list becomes empty.

First of all, find an object having an empty back-list. The fact that an object has an empty back-list means the objects can be rendered immediately because there are no objects covered by the current object. The found object is inserted to the rendering queue and deleted from the back-list of its all front objects. If one of the front objects has an empty back-list, the object becomes a renderable object that can be rendered immediately. And then, the previous process is recursively performed in order to remove the newly selected object from the other object's back-list again. After finish this process for all object in the scene, the remaining objects which are not inserted yet are the non-overlapped objects. So these objects can be rendered without any ordering process.

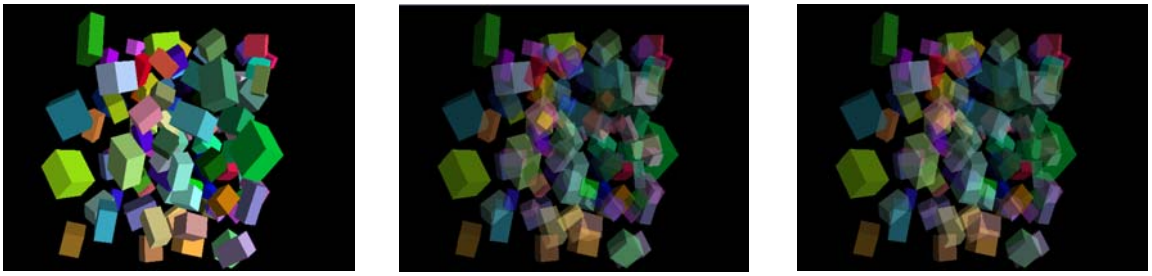
5. EXPERIMENTS AND RESULTS

To verify the performance and effectiveness of out technique, we performed several test on a desktop PC with a 2.13 GHz Intel Core2Duo CPU and 1 GB RAM. To test the performance, we generated random numbers of boxes and panels that are not intersected with any other boxes.

Table 1 shows the statistics of results of implemented applications using the proposed sorting technique. The test was done using several numbers of randomly-generated parallelopipedons and parallel panels. In Figure 4 and 5, we illustrated the results of rendered image under standard OpenGL environment. Figure 4 shows the comparison of not blended, incorrectly blended and correctly blended images. Figure 5 precisely shows the difference between correctly and incorrectly rendered results. As shown in Figure 5(b), the overlapped objects can be correctly rendered by using sorting technique proposed in this paper. In Table 1, we summarized the performances for 100 or 300 numbers of objects and we illustrated the performance for all number we considered in figure 6.

6. CONCLUSIONS

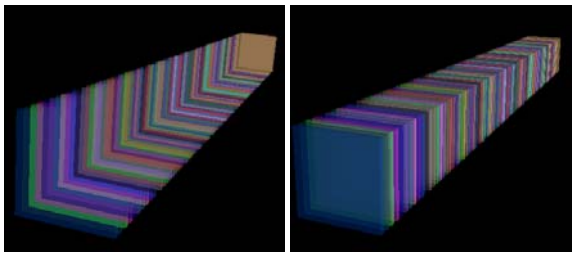
In this paper, we proposed a new depth sorting technique for accurate and fast alpha blending for developing 3D GUIs. In this technique, we designed several effective techniques to construct a projected 2D polygon from a boundbox and to determine the spatial relation between overlapped 3D objects. The proposed sorting technique is very easy to implement and it is possible to quickly arrange the depth order of each object.



a) without alpha blending b) blended without depth sorting c) correct result with depth sorting

Figure 4. Comparison of the correct and incorrect alpha blended results using 100 random boxes.

Although this technique is able to calculate the depth orders very quickly, the restrictions make it difficult to apply our technique for some application areas. In spite of the restrictions, this technique can be applied to various 3D applications in a limited hardware environment if there is not enough hardware performance.



a) incorrectly blended result b) correctly blended result

Figure 5. Comparison of the correct and incorrect alpha blended result using 100 parallel panels.

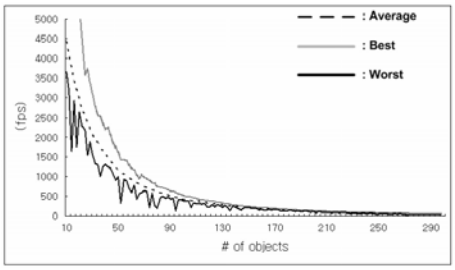


Figure 6. Rendering performance of proposed sorting technique for parallel panels.

	Random 100 boxes	Parallel 100 panels	Parallel 300 panels
Mean	293.5 fps	396.3 fps	55.3 fps
Best	1651.2 fps	476.3 fps	62.8 fps
Worst	95.9 fps	209.4 fps	49.6 fps

Table 1. Rendering statistics.

7. REFERENCES

[AGMV97] Agarwal P. K., Guibas L. J., Murai T. M. and Vitter J. S., Cylindrical static and kinetic binary space partitions. In *ACM Symposium on Computational Geometry*, pp. 39-48, 1997.

[BOS04]. Berg M., Overmars M. and Schwarzkopf O., Computing and verifying depth orders. In *SICOMP*, pp. 437-446, 2004.

[FDFH96] Foley J. D., Dam A., Feiner S. K. and Hughes J. H., Computer Graphics : Principles and Practice, 2nd edition, Addison-Wesley, 1996.

[HKN80] Huchs H., Kedem Z. and Naylor B., On visible surface generation by a priori tree structures. In *Proc. of ACM SIGGRAPH*, Vol. 14, No. 3, pp. 124-133, 1980.

[KLNR97] Karasick M. S., Lieber D., Nackman L. R. and Rajan V. T., Visualization of three-dimensional delaunay meshes. *Algorithmica*, Vol. 19, No. 1-2, pp. 114-128, 1997.

[NNS72] Newell M. E., Newell R.G. and Sancha T. L. A solution to the hidden surface problem. In *Proc. of ACM Nat. Mtg.*, 1972.

[SBGS69] Schumacker R., Brand B., Gilliland M. and Sharp W., Study for applying computer-generated images to visual generation. AFHRL-TR-69-74, US Air Force Human Resources Lab, Tech. Rep., 1969.

[SBM94] Stein C., Becker B. and Max N., Sorting and hardware assisted rendering for volume visualization. In *Symposium on Volume Visualization*, pp. 83-90, 1994.

[SL98] Snyder J. and Lengyel J., Visibility sorting and compositing without splitting for image layer decompositions. In *Proc. of ACM SIGGRAPH*, pp. 219-230, 1998.

[Tor90] Torres E., Optimization of the binary space partition algorithm (BSP) for the visualization of dynamic scenes. In *Proc. of Eurographics '90*, pp. 507-518, 1990.

[WMS98] Williams P. L., Max N. and Stein C., A high accuracy volume renderer for unstructured data. IEEE Trans. on Visualization and Computer Graphics, pp. 1-18, 1998

An Unified User Dynamic Interest Description Model

Cheng Zeng, Hui Wang, Haiping Cao, Ying Zhang
State Key Lab of Software Engineering, Wuhan University, Wuhan, China, 430072
zengc@whu.edu.cn

ABSTRACT

The amount of information available world-wide and its network-based linkage will continue its rapid growth in the foreseeable future. It is urgent that we need a new retrieval mode to span the gap between information explosion and user requirement. This paper reports an user interest mining method based on cross-media and a unified model which describes user's dynamic interest changing through time to overcome the contradiction between short-time interest and long-time interest description in traditional methods. We also present a developing personalized retrieval system FizZ which mines user's interest by installing IE plug-in.

Keywords

UDM, MIS, Cross-media, interest community

1. INTRODUCTION

Current web search engines are very successful and are increasingly big business. They sharply shorten the time that people obtain information and bring lots of enterprises huge profits. But we have to recognize that these existing technologies could not satisfy users' particular information requirement. As evidenced in a recent study, "Contrary to many enterprises' expectations, search technology hasn't settled into a general satisfying phase. The biggest increases in the relevance of web search results to users will come neither by fine-tuning existing search algorithms nor from defining completely new algorithms based on web content and structure alone, the biggest increase in user' satisfaction will come when the results returned are tailored to the individual as well as to the question asked". This information deluge calls for new semantic extraction and the most innovative retrieval and exploration techniques.

The first problem of improving user satisfaction for retrieval results is to mine and understand user interest, and describe it with a reasonable mode. This paper presents a new method which mines user's interest and preference by analyzing various types of media documents that user browsed based on cross-media technology. Besides, we add time axis into user interest model to distinguish short-time interest from long-time interest and discovery the periodic interest change.

2. RELATIVE RESEARCHES

The preparation for constructing user interest description model(UM) is to collect user interest which includes explicit and implicit collecting. The former has been broad applied to lots of retrieval systems. The latter denotes that the system mines user implicit interest by analyzing Web log[Jou0a], such as LOGSOM[Jou0b], and tracking user's manipulation. Shahabi[Jou0c] sends agent to collect user interest in remote computers. It overcomes some problems such as error translating of Web cache and IP address etc., but it is possible to be limited if the client has security mechanisms. In addition, lots of systems [Jou0d] combine explicit and implicit methods in practice.

Current UM mainly includes vector space model, classes or layers model, catalogue structure model and so on. In [Con0a], Rachid adds user profile and weighted index item expression based on vector space model to realize personalized retrieval. Lin[Jou0e] presents a new mechanism describing and updating user interest model to realize personalized information service.

In addition, many researchers describe user interest with a tree structure based on the semantic relationship among interest lemmas. In [Con0b], ontology is utilized to improve the efficiency of personalized information retrieval and a framework is put forward which includes relationship measuring, user interest expressing and automatically updating. Yang[Jou0f] provides a complex preference operation for personalized retrieval and recommending in digital library.

3. DYNAMIC INTEREST DESCRIPTION

These methods introduced above emphasize particularly on text content mining or user action analyzing, but most of information in our daily life exist in multimedia form and the content that user is interested in begins to extend to more wonderful media types such as video, image, audio, etc.. The traditional interest mining method based on text will be not able to exactly and adequately collect user preference. Most of UM are inclined to describe long-time

This work is supported by the Doctor Subject Fund of Education Ministry No.20070486064 and the Hubei Province Natural Science Foundation of China No.2007ABA038 and National Basic Research Program of China(973 Program) No.2007CB310800 and the High School 111 Project of China No. B07037

interest, but ignores the instance that user interest fluctuates in short time, e.g. a user is used to browse investigative Web sites at work, but watch online movie during leisure time. The traditional UM is unfit for describing this kind of dynamic user interest. So this paper put forwards the user dynamic interest description model(UDM) based on multi-granularity interest space to resolve the problem.

Multi-granularity Interest Space

The multi-granularity interest space(MIS) is based on an interest concept hiberarchy(ICH) which is similar to ontology tree and is the tree catalog of all user interest points. Fig.1 shows the ICH and a 5+1 dimensions MIS based on a subset of ICH in terms of some application requirement, where d_i, t respectively represents the codes of interest concept and time dimension.

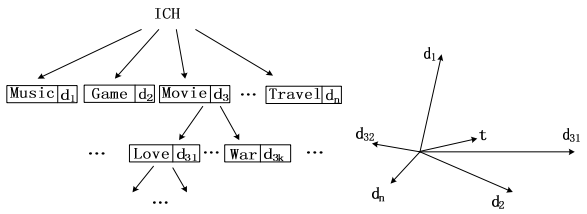


Figure 1. ICH and MIS

The goal of constructing MIS is to compute the similarity between different users' interest or between user interest and retrieval result. In actual application, we could casually choose concept nodes in appropriate ICH layer to construct a certain MIS. So the similarity computation not only considers the time consuming, but also is more desirable in real user interest.

The time dimension is added into MIS so that each user interest is represented by a dynamic twisted n- dimensions cube which corresponds to a interest sub-space. We look the cube as the time-space representation of UDM.

Mining User Interest Based on Cross-media

This paper mines user interest with the strategy combining explicit and implicit methods. The difference with traditional method is that cross-media technology is applied in user interest mining which discoveries cross-media relationship in different aspects. So different types of media with similar semantic could express each other.

For explicit user interest mining, we use the manner that user appoints local folders as analyzed objects other than inputting keywords in the course of registration. Users usually arrange the stored manner and types of local files according with their interest so that we could mine them in terms of some certain rules, e.g. the deeper path the file is stored, the lower interest user will show.

Then each document will be analyzed to mine the latent user interest. For visual media, semantic templates [Con0c] are utilized to extract semantic concepts. So text analysis could be used to cluster semantic concepts and discovery interest points. Audio is analyzed directly based on header information. By analyzing Web pages, online movie that user has browsed, user interest could be implicitly obtained. The text circling round them, or existing hyperlink will be utilized to indirectly mine the semantic information in video or image. It is similar that analyzing relative annotation in Web pages or header information in local cache will gain the semantic content of online music.

UDM

3.3.1 User background information

User background information contains some user private information that is possibly obtained such as sex, IP, blood type, age and so on. The information couldn't be directly utilized to describe user interest, but they have certain regular relationship each other. For example, woman is always interested in cosmetic and those whose ages are between 25 and 35 seem more interested in it.

3.3.2 Interest sub-space(ISS)

ISS is the kernel of UDM which is the projection of user interest in multi-granularities interest space. In traditional UM based on vector space, user interest is represented as multi-dimensions vector. It is convenient to compute user interest similarity with vectors angle. But it ignores user interest changing in short time and will spend much time if interest dimension is very high.

By interest community discovering, we could construct the most suitable multi-granularities interest sub-space for each user. An ICH sub-tree can be chosen to construct collective interest community sub-space. For each user interest point, we record the curve function constructed by the interest rank changing through time. By union process, we could calculate the fluctuating curve function $f_{d_i}^{k}(T, D_i)$ of the interest point for the user k in general one day. So user's ISS could be represented by a function set:

$$ISS(k) = \{f_{d_1}^{k}(T, D_1), f_{d_2}^{k}(T, D_2), \dots, f_{d_i}^{k}(T, D_i), \dots\} \quad (1)$$

Where there are n user interest points, namely the number of leaf nodes in ICH sub-tree. ISS corresponds to an irregular cube (UIC) in n+1 dimensions interest space. All cubes and their relationship will construct a huge user interest network which provides some latent knowledge.

3.3.3 Interest community information share

There is inevitable interest relationship among different users which is inclined to develop in the community manner. Those UIC corresponding to different users who have similar interest generally gather together. By clustering in different MIS, it is able to discovery different degree of interest communities.

Interest point set of each user's corresponds to an ICH

sub-tree. A unified deputy sub-tree is constructed for these users by similarity computation of tree structure, which is the base of the public interest space. Each user interest point corresponds to a function $f_{d_i}^{k}(T, D_i)$. The curve function of their interest father point based on ICH could be calculated by average curve energy function:

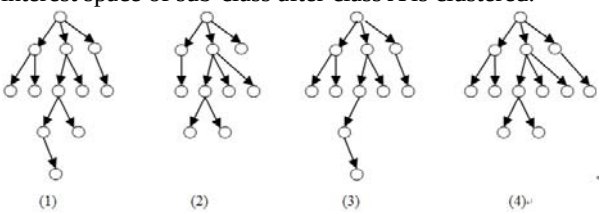
$$f_{d_f}^{k}(T, D_f) = \sqrt{\sum_{u=i}^{j} E(f_{d_u}^{k})/n^2} \quad (i < j) \qquad (2)$$

where $E(f_{d_u}^{k}) = \int f_{d_u}^{k} dt$

d_f denotes the father point of set $d_i,\ldots,d_j$. The public interest space ensures that all UIC could be clustered in the same measure system to realize interest community discovery.

The clustering algorithm chosen for discovering community is DBSCAN where the representation of UIC plays an important role in the determination of different clusters. Usually, centroid or other key point is used to replace UIC. It is possible that the high layers of different users' ICH sub-tree have enormous difference, although they have very similar interest and the space dimension corresponding to sub-tree is high. Using UIC centroid directly is not efficient in this status. A set of core points P_{cor} is used as the representatives. Let $CorC_i^x \subseteq C_i^x$ (cluster C_i^x in a set of cluster C^x at interest subspace x) be a set of the core points belonging to this cluster. The definition of P_{cor} is as follows: $P_{cor}C_i^x \subseteq CorC_i^x$ is a set of core object iff ($\forall p_k, p_l \in P_{corC_i^x}: p_k \neq p_l \Longrightarrow p_k \notin N_{\epsilon x}(p_l)$) and ($\forall c \in CorC_i^x, \exists p \in P_{corC_i^x}: c \in N_{\epsilon x}(p)$ with $N_{\epsilon x}(d_i)$ of a point d_i is defined as $\forall d \in N_{\epsilon x}(d_i): \|di - d\| \leq \epsilon_x$.

For the incipient clustering result, we mine the public portion of interest sub-tree corresponding to all UIC in class A, which will be used to construct the interest space for next time clustering, shown in Fig.2. The integrating course accords with the rule that the nodes in the same layer use union operation but those in different layers use intersection operation. The rule covers as more user interest as possible, but ignores each user's personalized interest which will be advised in the course of constructing interest space of sub-class after class A is clustered.



(1)-(3) 3 different ICH sub-tree;(4) The common interest tree

Figure 2. The conformity of interest sub-trees in class A

The clustering in the interest spaces of different granularities realizes different levels of descriptions for the same user. So the scale of interest clustering could be adjusted by selecting the size of similar interest community in terms of requirement. If the granularity is coarse, lots of users would have similar interest. On the contrary, if the granularity is fine, there would produce interest divaricating among those users so that we will obtain more interest communities. The shared interest in a community is represented by a set of information pair:

$$CIS(k) = \{(\text{Subtree-ID, C-centroid})\} \qquad (3)$$

Where Subtree-ID denotes the ID number of ICH sub-tree corresponding to the granularities space and C-centroid denotes the interest community centroid.

Computing Interest Relationship

Interest relationship includes that among different users' interest and that between user interest and retrieved results. The former is calculated by distance between different ISS centroids. However, the complexity for calculating the centroid of multi- dimensions geometric object does not satisfy for the demand of dynamic relationship among large numbers of online users. This paper uses the method of regular exampling that transforms ISS to a set having some exampled points, where each point denotes the interest status for a user at a moment. So the points set corresponding to a moment could describe user interest in short time. The sparseness or denseness of exampled points reflects the extent that user has the tendency for such interest region in long time. The user interest description is integrated into a unified mode. The centroid of exampled points set will substitute for the user interest, where the centroid coordinate is the mathematical expectation of n exampled points in every dimension x_i:

$$\overline{x_i} = \sum_{j=1}^{n} x_j / n \qquad (4)$$

The Euclidean distance between centroids is calculated to estimate the difference between users.

The retrieved results, such as Web page, image and so on, exist as objective entities which generally contain fixed information. So they are also able to be represented as points in interest space and the relationship between user interest and retrieved result is calculated by space distance.

4. EXPERIMENTS AND ANALYSIS

We construct 148 UDM by analyzing browsed Web pages and submitted interested folder of each user's based on cross-media technology. We track users' online action for two days. Because the first aspect needs a rather big data size to validate the algorithm efficiency, the data of 148 users' are distinctly insufficient. So we simulate the data of 3000 users' based on existing data and the sampling frequency is every 2 minutes. At last there are about 600 interest describing points for each user. Each point is represented by a tuple {ID, Username, Time, Space, Cor}

where Space and Cor respectively denote current interest space and point coordinate. The algorithm complexity is n*log(n).

We random select 8 users of different communities in low granularity interest space and track their clustering results in high granularity interest space which is repeated several times. The result proves the upward restriction of user interest relationship in different granularity spaces. By counting and analyzing the clustering results, we obtain some funny relationship information, shown as follows:

Specialty 1	*Sum1*	*Specialty 2*	*Sum2*	*COM(%)*
Medicine	19	Biology	15	75.6
Computer	25	Commerce	21	69.2
Law	18	History	14	42.9
Art	8	Computer	25	33.9
Biology	15	Law	18	16.7

Table 1. Comparability between different specialties

Age 1	*Sum1*	*Age 2*	*Sum2*	*COM(%)*
30-35	24	40-45	16	57.3
25-30	19	30-35	24	39.6
25-30	19	20-25	26	35.3
Under 10	2	Above 60	3	28.7
15-20	6	40-45	16	13.9

Table 2. Comparability between different ages

Type 1	*Sum1*	*Type 2*	*Sum2*	*COM(%)*
O	28	B	43	41.2
AB	29	A	48	32.1
B	43	A	48	15.3
O	28	AB	29	10.2

Table 3. Comparability between different blood types

The result shows that specialty background has more effect for forming an interest community, but blood type plays a lower effort. In table 1, users of medicine specialty have 75.6% interest similarity with those of biology specialty. In table 2, users whose ages are 30~35 have more interest similarity with those between 40 and 45. Besides, users under 10 have not lower interest gap with those above 60 than our imagination. In table 3, users of blood type O have the most interest comparability with those of blood type B, and the most interest gap with those of blood type AB. Although the size of experiment data is not big, the analyzing results approximate the factual circs to a certain extent.

The recommended information based on interest communities is also displayed in result page of FizZ(shown in Fig.3). After users install the FizZ IE plug-in and startup it, the Web pages they browsed will be tracked and analyzed to mine user interest. Besides, user could appoint local folder with different media documents to analyze user interests. If user agrees to share his own interest, these private information will be able be utilized to discovery interest communities. By user feedback, we could evaluate whether retrieval results accord with the user interest.

Fig. 3 Retrieval result page of FizZ

5. CONCLUSIONS

The key problems of realizing personalized service are how to obtain and describe user interest. In this paper we present an user interest mining method based on cross-media. Besides, we put forward a unified UDM which is suitable to describe user's short-time and long-time interest, and discovery share information in interest communities. The experiment evaluates the feasibility and efficiency of our method. Especially, the developing prototype system FizZ achieves a fairly high user satisfaction.

6. REFERENCES

[Jou0a] Guo Yan, et al. Analyzing Scale of Web Logs and Mining Users' Interests**,** Journal of Computers. Vol.28, No.9, 2005

[Jou0b] Smith.KA, Ng.A. Web Page Clustering Using A Self-organizing Map of User Navigation Patterns. Decision Support Systems, 2003 ,35 :245-256

[Jou0c] Shahabi C, et al. An Adaptive Recommendation System without Explicit Acquisition of User Relevance Feedback. Distributed and Parallel Database, 2003, 14 :173-192

[Jou0d] Smyth.B, Bradley.K, Online Recruitment Services. 2002, 45(5) :39-40

[Con0a] Rachid Arezki, et al. Information Retrieval Model Based on User Profile. AIMSA 2004, p.490-499

[Jou0e] LIN Hong Fei, YANG Yuan Sheng. The Representation and Update Mechanism For User Profile. Computer Research and Development, Vol.139, No.17, p.843-847, 2002

[Con0b]Pablo Castells1, et al. Self-tuning Personalized Information Retrieval in an Ontology-Based Framework. OTM Workshops 2005, p. 977-986, 2005.

[Jou0f] YANG Yan, LI Jian-Zhong, GAO Hong. Ontology-Based Preference Model in Digital Library**.** Journal of Software, Vol.16, No.12, pp. 2080-2088, 2002

[Con0c] Cheng ZENG, et al. Cross-media Database Retrieval System Based on TOTEM. The 7th Web Information Systems International Conference. 2006, p182-193

POSTER: Video High-level Semantic Recognizing

Cheng Zeng, Xiaozhu Liu, Jun Li, Feng Dan
State Key Lab of Software Engineering, Wuhan University, 430072, China
zengc@whu.edu.cn

ABSTRACT

This paper presents an event semantic recognizing method based on Markov chain by stages, which takes object semantic as the bridge and is valid for recognizing complex event semantic. Semantic concept mapping mechanism based on semantic template is used to realize the automatic recognition of video semantic. In the experiment contrasting with IMAT system, our method shows more extensive recognition range and higher accuracy. Experimental results are encouraging, and indicate that the performance of the proposed approach is effective.

Keywords

Markov chain, STTS, Semantic template, Object semantic

1. INTRODUCTION

At present, the retrieval technology based on keywords has been adopted widely by many commercial search engines. Many retrieval systems of images or video based on content have also come into being in some research institutions. However, these technologies depend too much on user's participating, or retrieval results betray the user requirement. It is urgent to strengthen the research on media semantic information mining.

Video semantic mining is one of the crucial problems of intending cross-media search engine. The current researches about it focus on mining object semantic information, static scene. But they are restricted within some certain application domains, and don't adapt to deal with complex semantic mining such as dynamic structure, long term, multi-objects, multi-relations etc.. This paper presents a novel method which integrates object semantic with relationship information among objects to realize the transition from low-level feature to high-level semantic. Semantic template, which stores the mapping relationship between physical features and semantic, is constructed to automatically recognize video semantic.

2. RELATIVE RESEARCH

Virage[And0a] is a video automatic annotating system which provides an open framework to expand other video or audio analysis tools. But it doesn't refer to the gap problem between low-level features and high-level semantic. V. Mezaris[Jou0a] corresponds low-level features description in MPEG-7 to relevant middle-level description, and constructs a simple word database called object ontology.

Benitez[Con0a] presents an intelligent information system framework MediaNet, which integrates low-level features and media knowledge concepts into a single system. ASSAVID project[Jou0b] supported by Europe IST fund develops a video retrieval system. It realizes the automatic classification of sport scenes by researching special characters of different sports. Each class corresponds to a sport name that is convenient for user to retrieve based on keywords.

Navid[Jou0c] presented a method which recognizes some special regions, such as sky, grass and so on by low dimensions color feature based on support vector machine and wavelet texture analysis. Multi-label[Jou0d] is utilized to recognize more objects and classifies visual media containing multi-objects. IBM develops a MPEG-7 video annotation system [Jou0e], which realizes video retrieval based on object or event concepts by manual annotating beforehand. In addition, the system provides a training and learning mechanism based on HMM for automatic annotation.

Jurgen et al.[Jou0f] realize the automatic marking of some objects or wonderful scenes information such as goal, football, athlete and so on in football game. Visual and audio features are utilized at the same time to recognize video event semantic in [Jouog].

3. OBJECT SEMANTIC TEMPLATE

Semantic Template Training System

Semantic template is looked as the bridge between feature layer and semantic layer. Each template maps to several semantic concepts. This paper constructs different semantic template respectively corresponding to object and event semantic. Object

This work is supported by National Basic Research Program of China(973 Program) No.2007CB310800 and the Doctor Subject Fund of Education Ministry No.20070486064 and the Hubei Province Natural Science Foundation of China No. 2007ABA038 and the High School 111 Project of China No. B07037

 ISBN 978-80-86943-17-6

semantic template (OST) contains not only static feature relationship obtained by statistic learning, but also best suitable segment granularity for each object region, object sub-region features, space relationship description and so on. In addition, best suitable feature will also be stored to realize retrieving with multi-granularities. Event semantic template(EST) is based on OST. It stores time-space topology relation among regions, state transferring through time, objects appearing list and so on.

The key problem of constructing OST is how to correspond object region in video with object semantic concept (OSC) in real-world. We train visual media set by Semantic Template Training system(STTS).

Pre-processing and Concept Matching

Video pre-processing includes features extracting, video segmenting, video frame segmenting based on object and others. We take the begin-frame of video segment as original information. STTS will be utilized to obtain the mapping relationship, between object region information and OSC.

After ascertaining the object regions, we will dynamically adjust the segmenting granularity of this region and verify its tracking effectiveness in follow-up frames for each semantic object.

Constructing OST

This paper improves parameters accuracy in semantic template by iterative processing based on statistical probability until all parameters level off. The aim is that each object region segmented could be automatically mapped to the best suitable semantic concept. Moreover, we take into account the relationship among different OSC which will improve the effect of mapping and enhance the validity of video semantic expression.

We define $f_i = [f_i^1, f_i^2, \cdots, f_i^X]$ (i∈[1,N]) to denote video feature vector, where N is the number of video segments in training set and X is feature dimension. c^j (j ∈[1,Q])denotes an OSC, Q is the amount of OSC. OST is represented as o ∈O= $\{o_1, o_2, \ldots o_K\}$, where K is the amount of OST waiting for being constructed. Each feature f is able to belong to several OST, but each OST corresponds to a single concept. The essence of overcoming semantic gap is to compute P(c|f), where OST plays the bridge role.

(f_i, c^j) constructed by feature vector and semantic concept is supposed to be independent each other that could be represented as follows:

$$P(f_i, c^j \mid o_k) = P_\lambda(f_i \mid o_k) P_\nu(c^j \mid o_k) \quad (1)$$

Feature vector and concept are looked as random distributing in video frame. The probability, that a OST is chosen, is represented as $P(o_k)$. $P_\lambda(f_i \mid o_k)$ and $P_\nu(c^j \mid o_k)$ respectively denote conditional probability of feature f_i and concept c^j. If the effect of OST o_k is able to be ignored, we will compute the probability of (f_i, c^j):

$$P(f_i, c^j) = P(c^j) \sum_{k=1}^{K} P_\lambda(f_i \mid o_k) P(o_k \mid c^j) \quad (2)$$

$P(o_k \mid c^j)$ could be transformed by Bayesian formula. Suppose that feature vectors in whole media document accord with K Gaussian mixture distribution. So each OST o_k will correspond to a Gaussian distribution.

$$p_\lambda(f_i \mid o_k) = \frac{1}{(2\pi)^{X/2} \left|\sum_k\right|^{1/2}} e^{-\frac{1}{2}(f_i - \mu_k)^T \sum_k^{-1} (f_i - \mu_k)} \quad (3)$$

where μ_k, $\sum_k$ respectively denote the mean and covariance matrix of f_i in the object region corresponding to o_k. In term of maximum likelihood principle, function L(θ) and ln (L(θ)) will tend to maximum with the same independent variable. Therefore, this paper takes $p_\nu(c^j \mid o_k)$ as independent variable to form the probability formula for annotating the whole training set.

$$\ln \prod_{j=1}^{Q} \prod_{i=1}^{N} [p_\nu(c^j \mid O)^{\delta_i} \prod_{t,s=1,t\neq s}^{\delta_i} P(c^t \mid c^s)]$$

$$= \sum_{j=1}^{Q} \sum_{i=1}^{N} [\delta_i \ln(\sum_{k=1}^{K} P(o_k) P_\nu(c^j \mid o_k))$$

$$+ \ln \sum_{t,s=1,t\neq s}^{\delta_i} P(c^t \mid c^s)] \quad (4)$$

$P(c^t \mid c^s)$ denotes relationship between OSC. By similar principle, $P_\nu(f_i \mid o_k)$ could be calculated.

The above formulas gain the probability that feature f_i and concept c^j are chosen, while o_k is taken as prior probability. However, it is necessary to gain the posterior probability of o_k for transiting low-level feature to object semantic. We deduce the following formulas based on Bayesian rule:

$$P(o_k \mid f_i, c^j) = \frac{P(o_k) P_\lambda(f_i \mid o_k) P_\nu(c^j \mid o_k)}{\sum_{a=1}^{K} P(o_a) P_\lambda(f_i \mid o_a) P_\nu(c^j \mid o_a)} \quad (5)$$

$P(o_k \mid f_i)$ is calculated by similar way. The expectation of full likelihood estimate of $\ln P(\lambda, \upsilon, O)$ is calculated based on $P(O \mid \lambda, \upsilon)$:

$$\sum_{(i,j)=1}^{K} \sum_{i=1}^{N} \sum_{j=1}^{Q} s(c_i^j) \ln[P(o_{i,j}) P_\lambda(f_i \mid o_{i,j}) P_\nu(w^j \mid o_{i,j})] P(O \mid \lambda, \nu)$$

$$P(O \mid \lambda, \upsilon) = \prod_{i=1}^{N} \prod_{j=1}^{Q} P(o_{i,j} \mid f_i, c^j) \quad (6)$$

where $s(c_i^j)$ denotes the probability that the concept corresponding to feature vector f_i emerges in media, $o_{i,j}$ denotes the OST corresponding to(f_i, c^j).

With the similar method, we gain the expectation of likelihood estimate $\ln P(\lambda, O)$ based on $P(O \mid \lambda)$.

The calculated $P(o_k \mid f_i)$ and formula (3) will provide the initial distribution of features in OST. And $P(o_k \mid f_i, c^j)$ provides the probability that OST is chosen.

$$P(o_k) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{Q} s(c_i^j) P(o_k \mid f_i, c^j)}{\sum_{i=1}^{N} \sum_{j=1}^{Q} s(c_i^j)} \quad (7)$$

Feedback Study of OST

After an OST has been constructed, it must be validated by retrieving in larger media set to update itself. The formula is as follows:

$$P(f_i \mid c^j) = \int P_v(f_i \mid o) P(o \mid c^j) do$$

$$= \mathrm{E}_o[P_v(c^j \mid o) \frac{P_\lambda(f_i \mid o)}{P(c^j)}]$$

$$\approx \frac{\sum_{k=1}^{K} P_v(c^j \mid o_k) P_\lambda(f_i \mid o_k)}{\sum_{t=1}^{K} P_v(c^j \mid o_t)} \quad (8)$$

In terms of formula (8), system will return M documents of maximal $P(f_i \mid c^j)$. Then user marks those satisfying results. We design a decision function ε(ξ), $\varepsilon(\xi) = \varepsilon_i$ which denotes that ξ belonging to the template o_k is accepted.

The key that judges the suitability of the constructed OST for object semantic recognition is to make the risk minimal. Otherwise, it is necessary to update OST by increasing training examples.

4. EVENT SEMANTIC TEMPLATE

This paper takes the event semantic as the result of semantic objects evolving with some certain rules, structure, or motion relation. Event semantic is inclined to describe certain interested object regions, but ignore other regions, namely observing different regions in different degrees.

For expediently constructing EST, we divide event semantic into unitary and multiunit event semantic. The first is possible to refer to one or two interested objects, but exist only one mainly observed semantic object. Another is combined with lots of unitary events with certain relationships. The interested object of each unitary event contained in multiunit event is likely to be different.

1. Recognizing unitary event semantic

Unitary event semantic is looked as the minimal unit describing event semantic. If a video segment contains one interested object, the object will directly be taken as observed object. On the contrary, if it contains several interested objects with different motion state, it is necessary to classify them.

Unitary event is actually used to describe semantic objects (classes), theirs own motion states and possibly existing action relation between two objects (classes). We are inclined to recognize abstract event semantic without domain restriction. As a result, we define 5 and 6 motion state types for single object and two objects, respectively.

Thereby, the recognition of unitary event semantic is transformed into the problems of recognizing object semantic and classifying motion states of interested semantic objects.

2. Recognizing of multiunit event semantic

Multiunit event semantic is looked as the combination of some unitary event semantic with certain temporal sequence relation. If each unitary event semantic is taken as a kind of state, multiunit event semantic will evolve into the result that a series of states transfer in turn.

This paper utilizes the theory of Markov chain to construct EST. Suppose the probability that unitary event E^t happens in time t is $\pi_t(E^t)$ and similarly E^{t+1} corresponding to $\pi_{t+1}(E^{t+1})$. If $\pi_t(E^t)$ is known, we could calculate the sum of these products which is equal to $\pi_{t+1}(E^{t+1})$:

$$\pi_{t+1}(E^{t+1}) = \sum_{E^t} \pi_t(E^t) p(E^t \rightarrow E^{t+1}) \quad (9)$$

If π_t and π_{t+1} are equal, it means the stable distribution of Markov chain. In other word, multiunit event semantic has been combined with stablest unitary event set with certain temporal sequence relation. We utilize the Sheskin algorithm to calculate the stable distribution probability by decreasing dimension and analogizing. At last, we could gain all stable distribution vectors, and then ascertain each parameter in EST.

5. EXPERIMENTS

We gain 15862 video segments in 866 video documents and train semantic templates based on different initial training sets and feedback times. The initial training numbers are 57, 130, 286 and 457. In the experiment, we define 59 OSC, 21 event semantic concepts. Each template will be used to automatically recognize and annotate the whole video database. We update these templates based on multiple times of feedback study where increase additional 20 examples at a time.

The experimental result shown in Fig.1 proves the average recognizable rate of object semantic more rely on the initial example number than feedback times. If the initial example number is much little, it

will require more times of feedbacks, and it stays at a lower recognizable rate even after many times of feedbacks.

Fig.2 reflects the change of high-level semantic recognizable rate when that of object semantic gradually increases. We found event semantic has fairly high requirement for object semantic recognition accuracy. Fig.3 displays the average semantic recognizable ratio for the whole video database. It shows that STT is better than IBM' IMAT in the average recognizing ratio for most of semantic concepts. Especially for multi-objects and multi-relations event semantic, STT improves recognition accuracy and expands the recognition range.

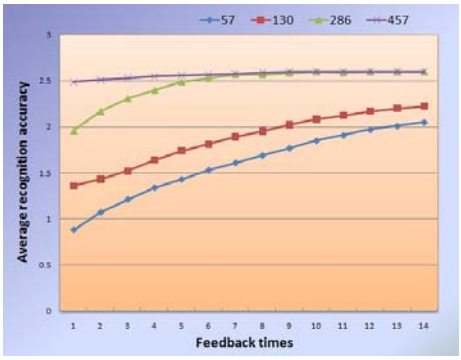


Figure.1 Object semantic recognizable rate

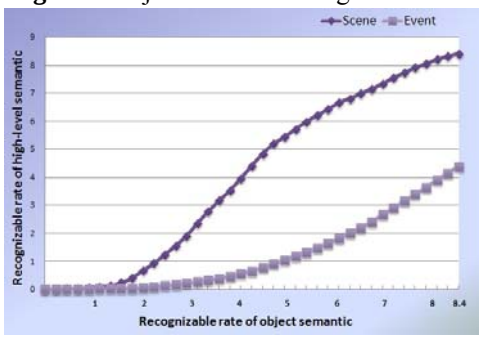


Figure.2 The relation of recognizable rate

6. CONCLUSION

This paper presents an object semantic mining method based on maximum likelihood estimate. It constructs OST by interactive study which stores the mapping relation between low-level features and semantic concept. Object semantic is taken as a bridge for realizing the transiting from low- feature to event semantic by stages. During contrasting our system with IMAT, our method is better in the range and accuracy of recognizing semantic concepts. It promotes the progress of video retrieval based on semantic.

7. REFERENCES

[And0a] Virage Inc. http://www.virage.com.

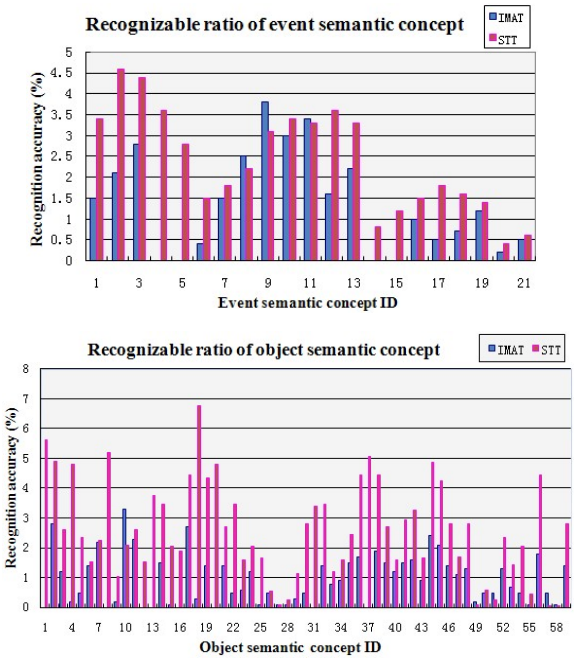


Figure.3 Average recognizable ratio of semantic concepts

[Jou0a] V. Mezaris et al.. Real-time Compressed-domain Spatiotemporal Segmentation and Ontologies for Video Indexing and Retrieval. IEEE Transactions on Circuits and Systems for Video Technology, 14(5):606– 621, May 2004.

[Con0a] Benitez A.B., Smith J.R., Chang S.F., "MediaNet: A Multimedia Information Network for Knowledge Representation", Proceedings of the SPIE 2000 Conference on Internet Multimedia Management Systems, Vol. 4210, 2000.

[Jou0b] K. Messer, et al. A Unified Approach to the Generation of Semantic Cues for Sports Video Annotation. Signal Processing 85 (2005) 357–383

[Jou0c] Navid Serranoa, Andreas E. Savakis, Jiebo Luo. Improved Scene Classification Using Effcient Low-level Features and Semantic cues. Pattern Recognition, 37 (2004) , p1773–1784

[Jou0d] Matthew R. et al. Learning Multi-label Scene Classiffcation. Pattern Recognition 37 (2004):1757 -1771

[Jou0e]Arnon Amira, Sankar Basub, Giridharan Iyengar. A Multi-modal System for the Retrieval of Semantic Video Events. Computer Vision and Image Understanding, V96, 2004, p216-236

[Jou0f] Jurgen Assfalg, Marco Bertini, Carlo Colombo, Alberto Del Bimbo, Walter Nunziati. Semantic Annotation of Soccer Videos: Automatic Highlights Identification. Computer Vision and Image Understanding, v92, v2003, p285–305

[Jou0g] Goh, et al. Audio-Visual Event Detection based on Mining of Semantic Audio-Visual Labels. http://www. merl.com/papers/docs/TR2004-008.pdf, 2004

Poster: Classification of cell nuclei using shape and texture indexes

Guillaume Thibault[1], Caroline Devic[2], Jean-François Horn[2],
Bernard Fertil[1], Jean Sequeira[1] and Jean-Luc Mari[1]
[1] LSIS Laboratory, Aix-Marseille II University, France
[2] Inserm U. 678, Paris VI University, France
Thibault@univmed.fr

ABSTRACT

In this paper, we present a study on the characterization and the classification of binary digital objects. This study is performed using a set of values obtained by the computation of "shape and texture indexes". To get the shape indexes, we extract a set of data called "measures" from 2D shapes, like for example surface and perimeter. These indexes are then used as parameters of a function returning a real value that gives information about geometrical and morphological features of the shape to analyze. A model characterizing the shape (and the texture) of objects is subsequently built. An application to the classification of cell nuclei (in order to diagnose patients affected by the Progeria syndrome) is proposed.

Keywords: Pattern recognition, shape and textures indexes, Haralick's features, cell nuclei classification.

1 INTRODUCTION

Pattern recognition is a major part of artificial intelligence that aims to automate the identification of typical situations. It is a major objective for many applications: handwritten character recognition (optical character recognition, etc.), video surveillance (facial recognition), etc.

At the heart of the pattern recognition issue, there is a first and unavoidable step: object characterization. It is often helpful to distinguish 2 classes of characteristics: the shape (polar signature [THK06], projection histograms [SR04], multi-scale curve smoothing for generalised pattern recognition (MSGPR) [KR06], etc.) and the texture [Har79]. Characterization of shape with shape indexes is more and more popular [IP97, TLG+03] espacially for learning-based classification [SEB+03]. Their flexibility, their simplicity of implementation and ease of use with a classifier make this approach an appropriate choice for many problems.

The aim of this paper is to create a model to classify blood cell nuclei in patients affected by Progeria syndrome. This rare syndrome is a laminopathy [GBC+03] that causes patients to age prematurely. To visualize nuclei, images are obtained using a fluorescent microscope that detects FITC tag (Fluoresceine IsoThioCyanate) showing the shape and the lamin A/C protein distribution.

The first part of this paper defines the concepts of measures, shape indexes, examine their main properties and their usage for pattern recognition. Then the model used to solve the classification task is presented, studied and validated.

A set of over three thousand cell nuclei (figure 1) from patients with Progeria syndrome has been gathered. These nuclei were manually classified as *healthy* or *pathological*. The shape criterion was the most important diagnostic clue for 89% of the nuclei, but complementary information was obtained by a textural analysis relative to the homogeneity of the nucleus.

2 SHAPE INDEXES AND MEASURES

Shape indexes were presented for the first time in 1976 in the book by Santalo [San76] related to mathematical properties of convex shapes. The definition and properties of shape indexes can be found in [CC85, Fil95].

Shape indexes definition: *We call* shape indexes *parameters, coefficients or a combination of coefficients capable of providing numerical information about the shape of objects. A shape index must be dimensionless and invariant by translation, rotation and homothety.*

The majority of shape indexes is derived from an equality or an inequality observed on the shape being analysed. In [San76] the authors have established a set of inequalities about convex shapes in a continuous space: $P^2 - 4\pi A \geqslant \pi^2(\rho_e - \rho_i)^2$, with A the surface, P the perimeter, ρ_i (respectively ρ_e) the radius of the biggest (respectively smallest) inscribed (respectively circumscribed) sphere. All these inequalities use different shape parameters called "*measures*". The calculation of these measures is an unavoidable step for getting shape indexes (indexes are based on at least one

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WSCG 2008 conference proceedings, ISBN 80-903100-7-9
WSCG'2008, February 4 – February 7, 2008
Plzen, Czech Republic.
Copyright UNION Agency – Science Press

measure). Measures can occasionally possess dimension: two dimensions for the surface, no dimension for the number of holes. Measures without dimension are considered as shape indexes.
The ability to build indexes from inequalities inherent to the shape under investigation is a major advantage: every kind of shape can be characterized and classification tasks are facilitated.

2.1 Shape indexes and classification

The objective of this study is to differentiate *healthy* from *pathological* cell nuclei.
Classification methods are divided into two important families: supervised and unsupervised methods. Supervised methods are more powerful but require expert knowledge to learn from. In this study we benefit from the biologists' and geneticists' knowledge who have specified classes (healthy and pathological) and subclasses (*ellipsoidal* and *puffy* shapes, homogeneous and non-homogeneous textures), which has allowed us to use supervised methods.
The aim of classification methods is to build a classification model based on the data under investigation. Although being applied to a specific problem, the model must remain general within the framework of data. For this reason, the data are split into two sets: a learning set and a validation set. The classifier must have comparable performances for both the learning and the validation sets.
It is necessary to construct a characteristic vector for each data prior the classification phase. The vector must be relevant to the problem in order to allow accurate classification and prediction. The major risk when providing too many characteristics to the classifier is overfitting. The greater the vector's dimension is, the greater the flexibility of the model and the better the classification are, but the greater the likelihood that the model's performance will be poor for a data set not used during the validation is. It is therefore necessary to validate each model with respect to overfitting. In this study, the characteristic vector is composed of shape indexes and Haralick's features [Har79] for the texture.
Classification is achieved by the logistic regression [DS89]. It is a linear model particularly well adapted to classification problems with two classes. Logistic regression performs a statistical analysis on the learning set and uses a logical distribution function to predict a membership probability: $P = P(Y/x) = \frac{e^{f(x)}}{1+e^{f(x)}}$ with $x = (x_1, ...x_n)$ the characteristic vector of the initial data, $f(x) = \sum_i \alpha_i x_i$ and $P(Y/x)$ the conditional probability P of the variable x to belong to the class Y.

3 CHARACTERIZATION OF CELL NUCLEI WITH SHAPE INDEXES

Healthy nuclei have ellipsoidal shape while pathological nuclei are puffy so that concave border areas are visible. For this reason, a set of fourteen shape indexes was obtained from the scientific literature. It also appeared interesting to create three additional indexes specifically designed for this study.

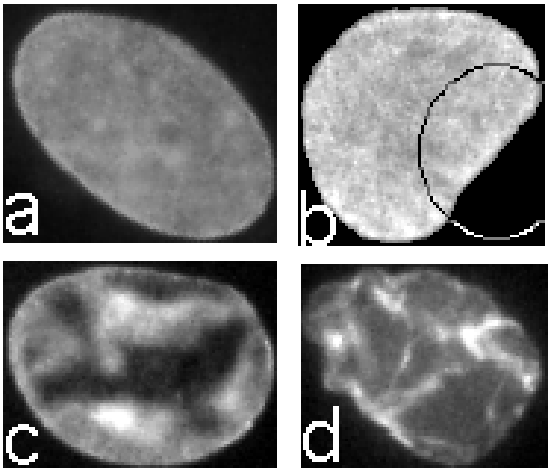


Figure 1: Four cell nuclei: *a* healthy, *b* puffy, *c* non homogeneous texture, *d* puffy and non homogeneous texture.

3.1 Three new shape indexes

Cell nuclei have a near elliptic shape when they are healthy. It is consequently judicious to build indexes characterizing the elliptic nature of the cells. The area of an ellipse A equals πab, with a the semi major axis and b the semi minor axis. Ellipses have some interesting properties that can be "measured": $R_{max} = a$ and $R_{min} = b$ or $R_{max} = \frac{1}{2}L_{AP1}$ and $R_{min} = \frac{1}{2}L_{AP2}$, with R_{max} (respectively R_{min}) the greatest (respectively smallest) radius and L_{AP1} (respectively L_{AP2}) the length of the principal (respectively secondary) axis. Two shape indexes can be derived based on these equalities:

$$\psi_{1\,ellipse} = \frac{\pi R_{min} R_{max}}{A}\ ,\ \psi_{2\,ellipse} = \frac{\pi}{4}\frac{L_{AP1}L_{AP2}}{A}$$

Denominators and numerators are equal in the case of an ellipsis and the index values are 1.
Pathological nuclei are currently not convex and consequently have concave border areas. To quantify these concave areas, it is possible to calculate the number of connected components N_{Cce} remaining when the shape is substracted from its convex hull. In the following, those connected components will be called "gap components". In this work, a normalized version of N_{Cce}, $\psi_{N_{Cce}}$, is used by the classifier:

$$N_{Cce} = card(ConvexHull(F)\backslash F)\ ,\ \psi_{N_{Cce}} = \frac{1}{1+N_{Cce}}$$

 ISBN 978-80-86943-17-6

$\psi_{N_{Cce}}$ is equal to 1 if no gap component is found and tends towards 0 as the number of gap components in the shape increases.
In practice, the components the surface of which equals one pixel are due to resolution errors cannot be considered as gap components. In fact, even small gap components (i.e a few pixels) may not be significant at least with respect to the classification of nuclei. The size and number of gap components N_{Cce} must be taken into account when diagnosing nuclei elements. A systematic analysis of the percentage of correct classifications versus the minimum size and the number was conducted. The highest classification rate (90%) is obtained by considering that nuclei with one 32-pixel at least gap component or two 12-pixel at least gap components have abnormal shapes.

3.2 Construction and validation of the model for nuclei characterization

The model of classification that is used in this study build a linear combination of the indexes in order to predict the class. The efficiency of classification (on the validation test) relies on the selection of the best subset(s) of indexes.
Best all subsets research on the seventeen indexes is performed to find the best combination of indexes. For the validation, the *K-Fold* protocol is used (with $K = 10$). A subset composed of eleven indexes (see appendix A) yields the best classification rate. The clearly bimodal distribution of the classification probabilities of the nuclei associated with the high classification efficiency demonstrate the relevancy of the selected indexes (figure 2).

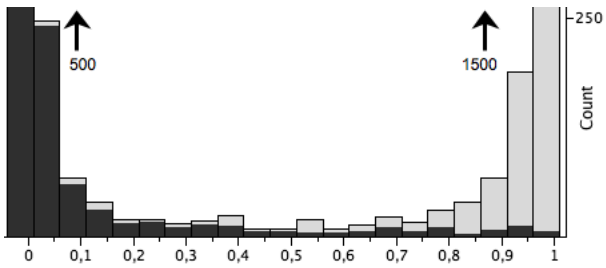


Figure 2: Distribution of the classification probabilities as given by the model to the nuclei from the validation set. The closer to one the probability, the more convex the nucleus. Dark areas and light grey areas are for puffy shapes and ellipsoidal shapes respectively.

4 TEXTURE CHARACTERIZATION

Although very good results have been obtained by modelizing shape, this only resulted in 89% of the global classification of nuclei into "healthy" or "pathological" groups. In order to improve this performance, it is necessary specifically analyse the homogeneity of the texture of the nuclei.
The lamin A/C distribution is homogeneous for healthy nuclei. However, the experts only consider nuclei as having non homogeneous texture when it is not "highly homogeneous" (figure 1).
In order to characterize texture, a co-occurrence matrix (32 grey levels) is built, out of fifteen Haralick's features are calculated [Har79]. One of these features is the homogeneity. Homogeneity is higher when the same pair of pixels is frequently found, as it is the case when there is an uniform area or a spatial periodicity.
Contrasting with the analysis of the shape, the analysis of the texture provides totally unballanced classes ("*homogeneous*" and "*non homogeneous*"), with approximately twenty times more nuclei in the homogeneous class. To efficiently build the model, the number of elements in each class must be roughly comparable. For this reason, the learning phase was carried out with all the nuclei from the non-homogeneous class (116 items) and an equal number of nuclei chosen by selecting prototypes in the homogeneous class with the *K-means* procedure ($K = 116$). The validation is realized according to the "Leave One Out"-protocol. Best all subsets research on the fifteen indexes is performed to find the best combination of indexes. Best subset is made of height indexes (listed in appendix B) which performs 90% of good classification of texture. In addition, the distribution of the probabilities is less constrasted than the distribution of the probability for the shape (figure 3). Several causes may be invoked to explain these differences: the main reason is the far lower number of nuclei belonging to the "non homogeneous" class which reduces the learning possibilities. The second reason is the noise introduced by using the tags, which reduces the reliability of the prediction.

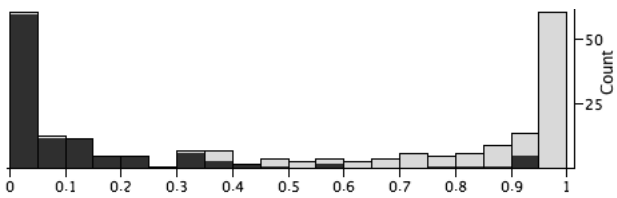


Figure 3: Distribution of the classification probabilities as given by the model to the nuclei from the validation set. The closer to one the probability, the more homogeneous the nucleus. Dark areas and light grey areas are for homogeneous and non homogeneous textures respectively

5 FULL MODEL: DIAGNOSIS OF NUCLEI

Two classification models have been built, characterizing the two main diagnostic parameters: the shape of the nucleus and its texture. These two models must be combined in order to establish the final model, capable of predicting the pathological aspect of nuclei. This

 ISBN 978-80-86943-17-6

model takes advantage of the 11 shape indexes and the 8 texture model features. The classification success rate hence reaches 90% on the learning set and 89.5% on the validation set.

6 CONCLUSIONS AND PERSPECTIVES

In this study, we have presented a method for the classification of cell nuclei of patients affected by the Progeria syndrome. The first step was based on the study of the shape of cell nuclei using shape indexes. Appropriate indexes were specifically built. These indexes have subsequently shown the ability to correctly classify shape of nuclei with a success rate above 95%. With reliability and validation in mind, it is not planned to try and improve this result for two reasons: the first being that it would be necessary to introduce additional characteristics that would jeopardize learning (overfitting). The few tenths of a percent gained in correct classification would be lost in validation. The second reason is that the efficiency of this approach already matches the reproducibility rate of experts.

Next a model based on Haralick's features was built in order to handle the problem of texture characterization. This model has provided a satisfactory handling of the texture characterization (90% of good classification) and has allowed improving the final model. The shape model alone obtained a ratio of 88.9% correct classifications of the nuclei (healthy/pathological). The addition of the texture model allows an improvement in the diagnosis of less than 1%.

In light of these results, it seems necessary to improve the texture model. For this reason it is planned to extend the notion of shape indexes and to use it to analyze textures. A possible solution would be to consider a texture like an elevation map ; with each pixel no longer representing a greyscale but rather a altitude.

A SHAPES INDEXES

With F a form, A the surface, P the perimeter, B the barycenter, R_{max} (respectively R_{min}) the greatest (respectively smallest) radius, ρ_i (respectively ρ_e) the radius of the biggest (respectively smallest) inscribed (respectively circumscribed) sphere, L_{AP1} (respectively L_{AP2}) the length of the principal (respectively secondary) axis, D the diameter, E the thickness, N_{Cce} number of gap components.

$\text{Extension}_{\text{Diameter}} = \frac{E}{D}$, $\text{Extension}_{\text{Radius}} = \frac{\rho_i}{\rho_e}$

$\text{Circularity} = \frac{R_{min}}{R_{max}}$, $\text{Deficit} = 1 - \pi \frac{(\rho_e - \rho_i)^2}{P^2}$

$\text{Convexity}_{\text{Perimeter}} = \frac{P(ConvexHull(F))}{P(F)}$

$\text{Convexity}_{\text{Surface}} = \frac{A(F)}{A(ConvexHull(F))}$

$\text{Symmetry}_{\text{Besicovitch}} = \sup_{x \in F} \frac{A(F \cap Symmetric(F,x))}{A(F)}$

$\psi_{1\ \text{Ellipsis}} = \frac{\pi R_{min} R_{max}}{A}$, $\psi_{2\ \text{Ellipsis}} = \frac{\pi}{4} \frac{L_{AP1} L_{AP2}}{A}$

$\psi_{\text{N}_{\text{Cce}}} = \frac{1}{1+N_{Cce}} \in]0,1]$, $\psi_{2\ \text{Parallelogram}} = \frac{A}{E \times D}$

B HARALICK FEATURES

With $p(x,y)$ the element (x,y) of the grey levels co-occurrence matrix, N the number of pixels of the texture to analyze, N_g the greyscale.

The standard deviation $\sigma = \sum_x \sum_y (p(x,y) - m)^2$

The correlation $\sum_x \sum_y (x-m)(y-m)p(x,y)/\sigma^2$

The average of the sums

The entropy of the sums

The entropy $\sum_x \sum_y p(x,y) log(p(x,y))$

The standard deviation of the differences

The homogeneity $\sum_x \sum_y \frac{1}{1+|x-y|} p(x,y)$

The dissimilarity $\sum_x \sum_y |x-y| p(x,y)$

REFERENCES

[CC85] Michel Coster and Jean-Louis Chermant. *Précis d'analyse d'images*. Editions du CNRS, 1985.

[DS89] Hosmer D.W. and Lemeshow S. *Applied Logistic Regression*. John Wiley & Sons, Toronto, 1989.

[Fil95] Isabelle Fillere. *Outils mathématiques pour la reconnaissance de formes*. PhD thesis, Université de St Etienne, Septembre 1995.

[GBC+03] Annachiara De Sandre Giovannoli, Rafaelle Bernard, Pierre Cau, Claire Navarro, Jeanne Amiel, Irene Boccaccio, Stanislas Lyonnet, Colin L. Stewart, Arnold Munnich, Martine Le Merrer, and Nicolas Levy. Lamin a truncation in progeria. *Science*, 300(5628):2055, 2003.

[Har79] R. M. Haralick. Statistical and structural approaches to texture. In *Proceedings of the IEEE*, volume 67, pages 786–804, 1979.

[IP97] Jukka Iivarinen and Markus Peura. Efficiency of simple shape descriptors. In *International Workshop on Visual Form*, pages 28–30, May 1997.

[KR06] Kidiyo Kpalma and Joseph Ronsin. Multiscale contour description for pattern recognition. In *IEEE Transactions on Pattern Recognition Letters*, volume 27, pages 1545–1559. Elsevier, October 2006.

[San76] L.A Santalo. *Integral Geometry and Geometric Probability*. Addison Wesley, 1976.

[SEB+03] A. Sboner, Claudio Eccher, E. Blanzieri, P. Bauer, Mario Cristofolini, G. Zumiani, and Stefano Forti. A multiple classifier system for early melanoma diagnosis. *Artificial Intelligence in Medicine*, 27(1):29–44, 2003.

[SR04] Hasan Soltanzadeh and Mohammad Rahmati. Recognition of persian handwritten digits using image profiles of multiple orientations. In *IEEE Transactions on Pattern Recognition Letters*, volume 25, pages 1569–1576. Elsevier, 2004.

[THK06] Abdelmalek Toumi, Brigitte Hoeltzener, and Ali Khenchaf. Classification des images ISAR pour la reconnaissance des cibles. In *XIIIème Rencontres de la Société Francophone de Classification (SFC)*, 2006.

[TLG+03] V. M. Tuset, I. J. Lozano, J. A. González, J. F. Pertusa, and M. M. García-Día. Shape indixes to identify regional differences in otolith morphology of comber. In *Journal of Applied Ichthyology*, volume 19, pages 88–93, April 2003.

POSTER: Introduction of Animation Assignment in Graphic Science Education making Use of CG Application of Data Describing Type

Hirotaka SUZUKI
Osaka City University
3-3-138, Sugimoto, Sumiyoshi-ku
558-8585, Osaka-city, Japan
suzuki@arch.eng.osaka-cu.ac.jp

ABSTRACT

In Osaka City University, Graphic Science Education has been located not as an entrance of drawing and designing exercise but as design language education. Design language is constructed by knowledge and technique which facilitates intelligible communication making good use of figures and drawings. From academic year 2006, number of lecture units in half year increased 1 unit in Osaka City university. Making use of the unit, we secured 3 units for animation exercise and introduced geometry CG art for further understanding about parametric control method in POV-Ray. Due to these change, we decided to regard animation work as compulsory assignment. Comparing to GUI interface CG modeler, it is difficult to construct animation with parametric modeler. As users have to consider relationship of parameters and produced image manually with parametric CG modeler in return for vast possibility of animation idea. Students submitted a lot of type of animation works. And almost all students evaluated CG animation as "interesting" exercise comparing to other topics. In this paper, the content of the subject and submitted animation works is introduced. And the comparison analysis between result of class evaluation for animation and that for other topics is discussed.

Keywords

computer graphics, animation, graphic science education

1. INTRODUCTION

In Osaka City University, Graphic Science education has been located not as an entrance of drawing and designing exercise but as design language education. Design language is constructed by knowledge and technique which facilitates intelligible communication making good use of figures and drawings (See [Suz03]). Based on the concept, Graphic Science has been provided as a subject of liberal arts. Half of the graphic science education is allocated for hand writing education (learning of projection method) and the rest is allocated for computer graphics (CG).

CG application software POV-Ray (Persistence of Vision Ray Tracer) has been selected as CG education tool due to following reasons; 1)Freeware, 2)Free platform (hardware, software), 3)High quality, 4)Data description style(readable text format) (See [Suz03] as well). In this subject, students have to submit following 4 assignments; 1)Scarecrow (for mastering combination of primitive figures), 2)Stairs (for mastering repetition procedure), 3)Lighting Equipment (for mastering concept of Constructive Solid Geometry and knowledge about behavior of light, from academic year 2004, See [Suz06]), 4)Final assignment (applying all knowledge learnt in the subject).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Copyright UNION Agency – Science Press, Plzen, Czech Republic.

As student's interest for CG animation is quite high, we had tried to introduce animation into the subject. However, as there were no computer terminal room for the CG exercise, we started teaching animation topic by only lecture. In the lecture, we selected good works from submitted student's works (still works, not animation) and generate CG animation with them. We introduce required tools for making animation and explain how to construct moving animation.

In academic year 2003, the room for CG graphic science education with 80 notebook PCs was established and we started 2 units (2 weeks) CG

exercise in the subject making use of format exchange tool for CG images and animation constructing tool. However 2 units was insufficient for students and we could not set animation works as compulsory assignment. Though some of students could submit the works and we could not reflect evaluation of the works to grade of students.

From academic year 2006, number of lecture units in half year increased 1 unit in Osaka City university. Making use of the unit, we secured 3 units for animation exercise and introduced geometry CG art for further understanding about parametric control method in POV-Ray. Due to these change, we decided to regard animation work as compulsory assignment. Comparing to GUI interface CG modeler, it is difficult to construct animation with parametric modeler like POV-Ray. As users can easily create animation by key frame setting with GUI interface CG modeler, users have to consider relationship of parameters and produced image manually with parametric CG modeler in return for vast possibility of animation idea. Students submitted a lot of type of animation works. And almost all students evaluated CG animation as "interesting" exercise comparing to other topics.

In this paper, the content of the subject and submitted animation works is introduced. And the comparison analysis between result of class evaluation for animation and that for other topics is discussed.

2. CURRICULUM CONTENTS

2.1 Enrolled students

The number of students enrolled in this subject is 174 in total. Because the enrolled students are from a wide variety of departments (Architecture, Civil Eng., Environmental Urban Eng., Applied Physics, Information Eng. and Environmental Design), there was a concern over a difference in the level of information literacy that each student had acquired by then, which might be an obstacle for the smooth implementation of a class exercise. As it turned out, however, a disparity in information literacy has not emerged as a major problem. As a matter of fact, the quality of assignments submitted by students does not always have a direct link with the departments they belong to.

2.2 Class Contents

The lecture given prior to exercise is designed to enlighten students of the basic points concerning animation: 1) an animation is made by combining numerous still images with motion, 2) specific application software programs are used (In this class, Animation GIF Maker and Irfanview are used) in the process of converting format of still images and that of combining still images into unified animation file, 3) while it is possible to produce still images with motion by gradually changing parameters of scene files, POV-Ray uses the parameter called "clock," which sequentially changes from 0 to 1 at the time of rendering, and 4) by using "clock," it will become possible to add such motions as "moving objects little by little (combining with the parallel movement parameter)," "changing brightness little by little (combining with the light-source strength parameter)," and "conducting a walk-through by changing viewpoints (combining with the coordinate axis of the viewpoint)."

Table 1. Weekly subjects of Graphical Science

	Content	Assignment
1st	Introduction	Impression for the submitted works by past students, and Comments and Suggestions for the Lecture
2nd	The Basis of POV-Ray	
3rd	The Description Method of Primitive Figures	Giving Small Assignment 1 (Scarecrow)
4th	Rotation, Enlargement & Reduction, Translation and Repeat	Assignment 1 Due, Giving Small Assignment 2 (Stairs)
5th	Conditional Branch, Use of Defined Colors, Block Pattern	Assignment 2 (Hand writing plan) Due
6th	CSG Model	Assignment 2 (Scene File) Due, Giving Small Assignment 3 (Lighting Equipment)
7th	Midterm Examination	Assignment 3 (Hand Writing Plan) Due
8th	Behavior of Light Flow(Light Source, Reflection, Transmission)	
9th	Giving Final Assignment and Use of Texture, Group, Prism, Solid of Revolution	Giving Final Assignment
10th	Other Useful Knowledge	Assignment 3 (Scene File)
11st	CG Animation 1	
12nd	CG Animation 2	
13rd	Geometry Art and CG Animation 3	Assignment CG Animation
14th	GUI-type Modeler Exercise	

After that, before going into the instruction using scene files, students are taught how to put together images of the staplers photographed with a digital camera to make an animation in which the staplers move from one place to another or their jaws open and close. This session is designed to make students realize that these tools (the animation producing tool and the image-file format-conversion tool) are not only applicable to still images produced by POV-Ray but also has many other potential uses. The session also aims to give students tips on making animations with a limited number of still images by, for example, reusing or reversing them.

After taking this step, some of the still images submitted by students were selected to demonstrate how to make animations that express rotation, movement and the change of brightness. Then simple scene files (with the "clock" already combined with

their rotation angles to rotate the objects) were distributed to the students for hands-on exercise.

These scene files are used for rendering. After rendering several still images all at once, Irfanview is used to convert the images from BMP format files to GIF format files. After that, Animation GIF Maker is used to combine the files into Animation GIF format. When doing so, several parameters are set to shift the "clock" from 0 to 1 to indicate how resultant animations are affected by how parameters are set (quick and rough ↔ slow and smooth) and what will happen if parameters are set inappropriately (if a range from 0 to 1 is divided into four, the rectangular parallelepiped in the center will not rotate.)

2.3 Lecture on "Geometric Art"

The first 30 minutes of the 13th class is spent for a lecture titled "Geometric Art" designed to raise students' awareness about the possibilities of using POV-Ray in a more parametric way. Specifically, instructions are given about how use POV-Ray to express mathematically defined figures such as a 2D curve, a 3D curve and a 3D curved surface, using, as a sample for each figure, a hypotrochoid, a double helix and a hyperbolic paraboloid, respectively (See Fig. 1). Due to time constraints, we only distributed scene files to prepare each figure and showed students how changes of various parameters result in different figures. Students were allowed to spend the rest of the exercise time for finishing their animation assignments. And the students who had finished their animation assignments can spend the rest time for Geometric Art exercise.

3. SUBMITTED ANIMATION WORKS

Fig. 2 shows examples of the animation works submitted by students. The top-left work is a mechanical animation based on a student's Final Assignment, which he had constructed in a parametric way with an eye to using it in his animation assignment as well. The top-right work, incorporating what is taught during the Geometric Art lecture, comprises several hypotrochoids turning around different axes. The bottom-left work is an animation showing the process of a Shogi game. The

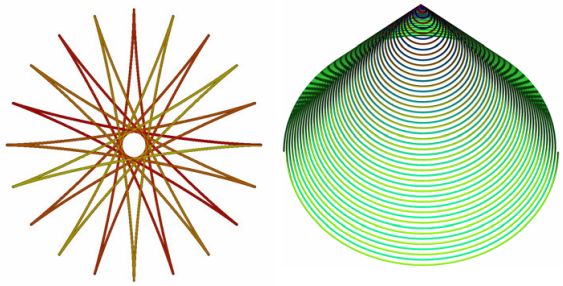


Figure 1. Example figures of geometric art works subjects (See [Suz07])

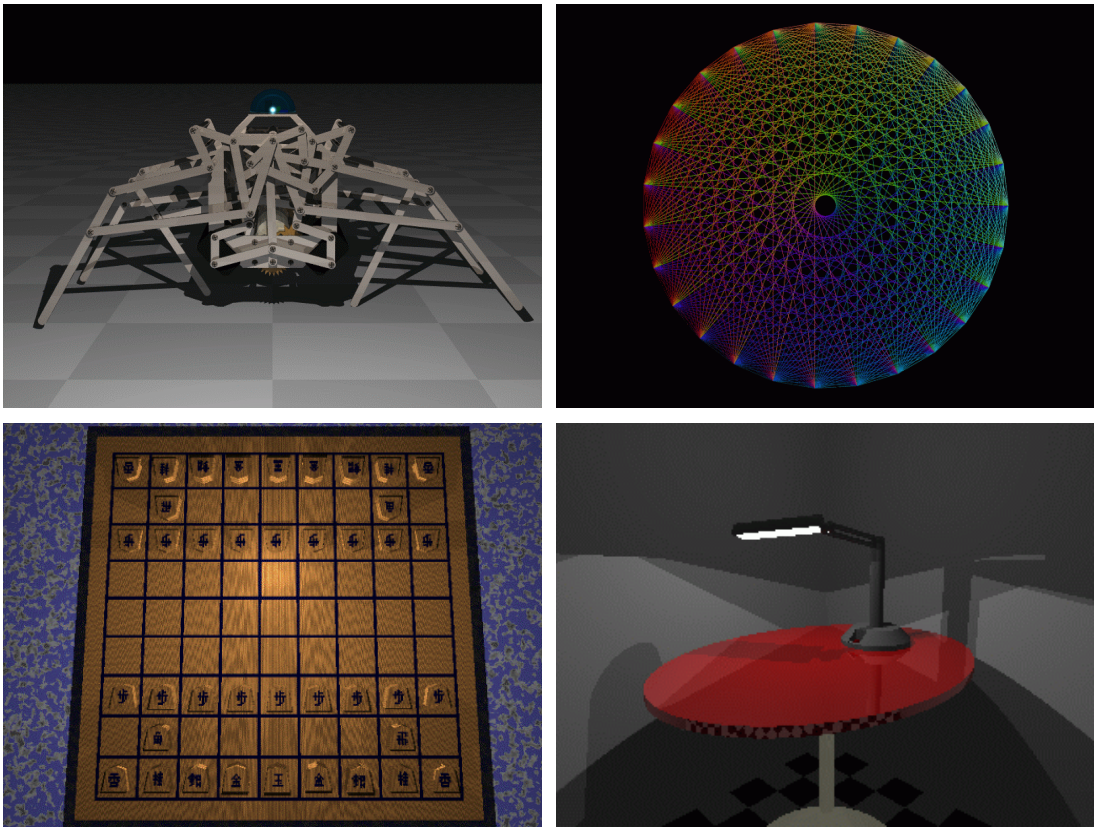


Figure 2. Examples of submitted works

 ISBN 978-80-86943-17-6

Rotation		Movement		Generation and Deletion		Change of Property		Transfor -mation		Viewpoint Movement		Brightness Change		Change of Light Source		Rotation of Light Source	
S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C
67	37	20	19	1	12	7	4	2	5	1	2	16	1	2	1	3	1

- "S" means "simple" and "C" means "complicated".
- "Simple" animation works almost produced by single scene file combined with single clock parameter.

Table 2. Breakdown of submitted works

author of this work have prepared one scene file for each move, a production method different from other works using "clock". The bottom-right work realized complicated movement by using several scene files prepared by making different "clock" combinations and linking the last frame of one scene file with the first frame of following scene file. The work is simple but able to bring about a story-like impression.

Table 2 shows the breakdown of submitted animation works by content. In this table, an animation that falls into several different categories in terms of contents is counted as different works. That is why the total number of the works exceeds 151, the actual number of works submitted. Of the 151 submitted assignments, 31 works fell into more than 1 category and 9 works into more than 2 categories.

All submitted works are introduced in following web page as well.

http://graphics.arch.eng.osaka-cu.ac.jp/zukeikagaku/pov2006/animation/index.html

4. CLASS EVALUATION RESULTS

During the term-end exams, a class evaluation survey was conducted in which students are asked to answer such questions as "Was the class interesting?," "Was the class easy to understand?," "Was the class useful,?" "Was the class recommendable to other students?" Students were asked to answer each question in the scale of 5 from "Yes, very much" to "No, not at all."Fig. 3 contains the graph of the results of students' evaluation with respect to "Animation" which, from this academic year, requires students to submit assignments, and "Geometric Art," a newly introduced lecture, as well as "CSG," which has been taught before and beyond 2005 as the most essential contents of this class. Fig. 6 indicates the averages of the total scores determined on the basis of the class survey results (From "Yes, very much" = 5 points, to "No, not at all" = 1 point).

As the figure shows, many students feel that "Animation" is "interesting." As for other aspects, "Animation" earns about the same scores as "CSG." As for "Geometric Art," average scores are lower than the other contents of the class in all aspects but students' evaluation is particularly low with regards to the "easy to understand" aspect. Students in some

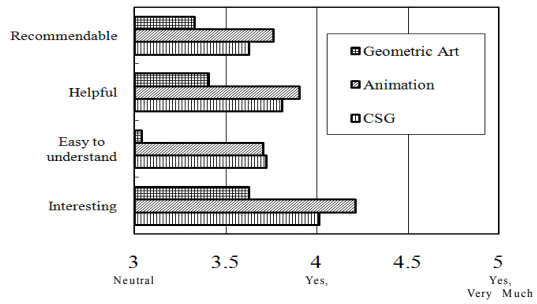


Figure 3. Class evaluation results

departments, however, give comments that they wish to hear more detailed explanation as to what Geometric Art is all about. The lack of thorough explanation, therefore, may be one of the reasons for the low evaluation of the Geometric Art session.

5. CONCLUSION

In this report, we have outlined the schedule, contents, submitted assignments, and class evaluation results with respect to an animation exercise incorporated into Graphic Science II, one of the basic subjects common to all departments of Osaka City University. Since we have not analyzed the results of the recent class survey by students' department, sex and the level of their submitted assignments, we should work out more accurate results by accumulating data based on various analyses and make use of such analysis results to improve the quality of our graphic science education in the years ahead.

6. REFERENCES

[Suz03]Hirotaka SUZUKI and Nobuhiro MIKI, A Graphic Science Education as a Training of Communication, Journal for Geometry and Graphics, Heldermann Verlag, Vol.7, No.2, pp.253-262, 2003.

[Suz06]Hirotaka SUZUKI, A Study on Impact of Introduction of Lighting Equipment Design Assignment into Graphic Science Education., Proc. of The 12th Intl. Conference on Geometry and Graphics, #E17, 2006.

[Suz07]Hirotaka SUZUKI, How to construct Geometric Art Works by CG (in Japanese), Distributed material in the lecture Graphic Science Education II, Osaka City Univ., 2007.

POSTER: Virtual Scene as a Software Component

Radek Ošlejšek
Masaryk University
Botanicka 68a
602 00, Brno, Czech Republic
oslejsek@fi.muni.cz

ABSTRACT

Graphics systems use many advanced techniques that enable to model and visualize a virtual scene with varying level of realism. Unfortunately, rendering algorithms significantly differ in the way how they process a virtual scene. Concrete implementations therefore usually lead to monolithic solutions. In this paper we propose the concept of a component-based scene graph, i.e. an independent scene graph, which can be used by many rendering strategies simultaneously and, moreover, which can be easily replaced with another implementation.

Keywords: Computer graphics, scene graph, software component.

1 INTRODUCTION

Nowadays, computer graphics offers a huge collection of rendering algorithms. They differ in the speed, in the quality of produced images and, unfortunately, also in the way how they handle and process a virtual scene. This different nature of rendering algorithms poses great difficulty in developing a unified rendering architecture, i.e. the architecture, which is able to handle a wide variety of rendering techniques via just a single unified interface. In spite of the difficulties, there exist several experimental architectures, e.g. those in [Fel95, SS95, DH02, OS03]. These architectures attempt to integrate more illumination strategies into a single unified system.

All rendering algorithms share one common concept called *scene graph* – a tree-based container of virtual object. Well-designed scene graph is the basic building block of any generic rendering architecture. Precisely proposed scene graphs can be found in [Opea, Opeb, Rei02, OS05]. Many of them are based on the VISITOR design pattern [GHJV95, Bus96], which enables to manage the scene traversal and inspection comfortably.

Visitors represent an individual operations over the scene, e.g. ray-intersection detection, shading operation, etc., with high level of encapsulation. Once the visitor (operation) is applied to the root node of a scene graph then the operation is automatically applied to all subnodes and thus to the whole scene. On the other hand, this high level of encapsulation means that the visitors are very tightly interconnected with the scene graph. They have to know the implementation details of the scene graph structure as well as the implementation of individual graphical objects. Moreover, the concept of visitors makes it difficult to perform changes in existing scene graph, e.g. extending the scene graph with a new type of graphical objects, features, etc.

A rendering system, even a generic, using a visitor-based scene graph therefore has to adapt itself to concrete scene graph implementation and the scene graph becomes the inseparable part of the rendering system. Moreover, the scene graph usually has to run on the same computer as the rendering system just due to the high encapsulation and interconnection.

Our goal is to make a scene graph as an independent software component [CD04]. Behaviour of a software component is precisely defined by the interface. On the other hand, internal implementation of the behaviour is not limited. Having a scene graph as a software component could bring many advantages. It could be easy to employ many heterogeneous rendering algorithms on the same scene just because they can share the common interface. Internal implementation of the scene could be easily changed without the impact to the existing algorithms. A component-based scene could run on dedicated computer and could be shared by more applications, e.g. in collaborative environments of virtual reality.

Precise definition of a unified interface is the key for successful proposal of component-based scene graph. Simultaneously, it is very serious challenge and difficult task. Unfortunately, the high level of encapsulation of visitors disqualify them from their direct usage because components have to be more autonomous and independent of their implementation.

2 SCENE GRAPH COMPONENT

Any operation over the scene has to perform two basic tasks. It has to traverse the scene graph and inspect indi-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright UNION Agency - Science Press, Plzen, Czech Republic.

vidual nodes. The visitor-based solution do both these tasks in a single step, as discussed above. To find interfaces suitable for software component we have to resign to the automatic application of scene operations inside the entire scene graph container. Instead, we have to separate the traversal and nodes inspection tasks and thus allow the invoker to gradually traverse the scene graph tree node by node and to inspect nodes individually.

Separation of the traversal from nodes inspection slightly breaks the strict encapsulation but enables us to manage scene operations outside of the scene. It allows to build different rendering strategies over the single unified scene graph, because the strategy can fully control behaviour of its operations.

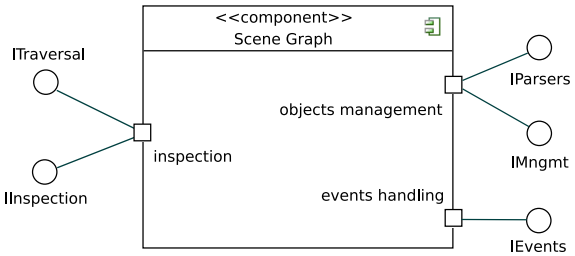


Figure 1: Scene graph component

Fig. 1 shows the basic component diagram of the scene graph. The *ITraversal* and *IInspection* interfaces are the key interfaces used for the scene traversal and nodes inspection respectively and they are discussed in the rest of this paper. However, a practically usable scene has to provide a lot of another useful operations, e.g. those related to the scene graph management, events management, etc. They are only suggested in the diagram for completeness.

3 SCENE GRAPH TRAVERSAL

Tree structure of scene graphs enables us do define a simple but unified interface for the traversal. A client traversing the scene has available a pointer to the scene graph tree as well as a private stacks and pipes enabling to store and restore the pointer. The client can instantiate various stacks and pipes, use them to store and restore position in the tree and thus implement various traversal strategies, e.g. depth-first or breadth-first search. Because every client has its own pointer and the set of stacks and pipes, there can be many clients traversing the scene simultaneously.

4 INSPECTION OF NODES

Scene graph nodes present a wide range of graphical information, e.g. description of shapes, material, transformations in space, etc. To handle all these miscellaneous properties in a uniform way it is necessary to classify them in smaller groups and to define specific interfaces for these groups.

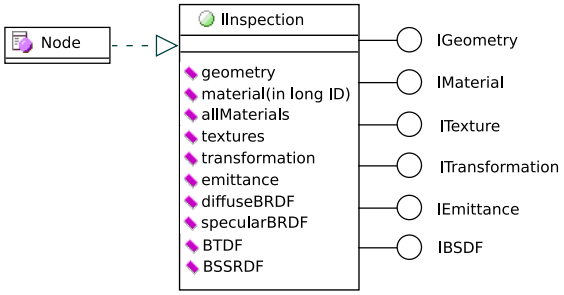


Figure 2: Inspection interface

Actually we have defined 6 groups of properties: (1) *geometry* defines shape of a virtual object in various ways, e.g. as a polygonal mesh, analytic surface, etc. (2) *material* determines coefficients for energy reflection and refraction, e.g. RGB color, transparency, etc. (3) *emittance* is necessary to model sources of energy. (4) *BSDFs, Bidirectional Scattering Distribution Functions*, determine what portion of incoming energy is reflected back to the scene with respect to incoming and outgoing directions and what portion is transmitted through translucent objects. (5) *transformations* are represented by 4x4 transformation matrices and allows manipulation with groups of objects in the space. Uniform management of matrices is clear and well-known and thus uninteresting. We therefore omit the 3D transformations from detail discussion. (6) *textures* are a common way to define color patterns on a surface. This property is also omitted from detail discussion.

Our complete inspection interface consists of two levels, as shown in Fig. 2. The *IInspection* presents the coarse-grained interface related to the actual node of a scene graph. Any scene graph node has to implement this interface. The *IInspection* itself is very simple. It just contains operations related to above discussed categories, each operation returning a relevant fine-grained interface of the category. An invoker selects required inspection category first, calls appropriate operation and then exploits returned fine-grained interface for final inspection. If some property is not present in the node, e.g. the node has no texture defined, then the fine-grained interface retrieval fails and the invoker continues with another inspections. In what follows we discuss individual fine-grained interfaces.

4.1 Geometry

Geometry represents shape of a surface. Computer graphics uses various kinds of geometry description, e.g. analytical surfaces, triangle meshes, etc. Our aim is to not restrict possible implementations of geometry. The unified interface in Fig. 3 therefore consists of only a general operations, which allow to “touch” the surface in a sense and to retrieve the necessary information about the shape. The set of operations cover a ray-intersection inspection, random sampling, basic

volume and spatial information retrieval as well as collision detection.

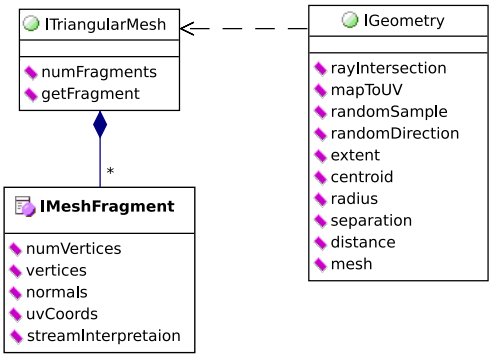


Figure 3: Geometry interface

Many existing graphical algorithms are based on polygonal surfaces. Any kind of geometry should be therefore transformable to this approximate description. The *mesh()* operation instantiates a triangular mesh, represented by the *ITringularMesh* interface in Fig. 3. Unified inspection of this mesh is ensured by using the mechanism known from the OpenGL [WNDS99], for instance. Externally, the mesh is composed of a stream of vertices, normals and mapping coordinates. Streams are interpreted as a *triangle strip*, *triangle fun*, etc. Thus, the concrete interpretation enables to reconstruct original triangles from the streams.

4.2 Material

Material characteristics are in the computer graphics expressed as n-tuples of real numbers, e.g. RGB colors, transparency, roughness, etc. But because different clients and algorithms can use different names even for the same material properties, then our unified model is based on the name-service concept similar to the Domain Name Service (DNS) of the Internet. Client can register various materials under various names and aliases at runtime and thus share these names with another clients. Concrete instances of materials inside a scene graph are then easily identifiable by their common natural names in the unified way.

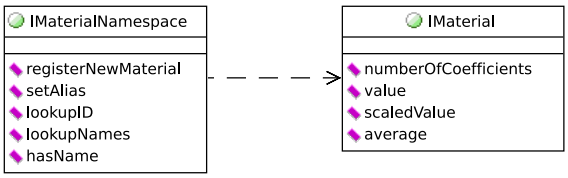


Figure 4: Material interface

IMaterial interface in Fig. 4 represents an n-tuple of concrete material coefficients, while the *IMaterial-Namespace* interface provides the names registration.

4.3 Emittance of Energy

Light sources pose very important but very complicated part of virtual scene. A light source can by understood as an object emitting energy. Description of the emittance can vary, but we can find several common principles that enable us to define emission in a uniform way.

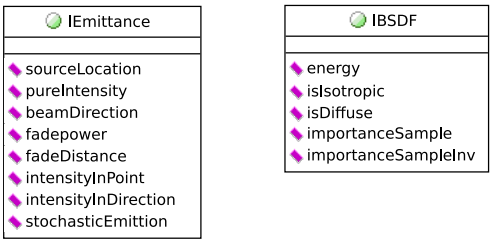


Figure 5: Emittance (left) and BSDF (right)

The basic characteristics of any emmitance consists of space location, beam direction, initial intensity (color) and attenuation. Unified *IEmittance* interface in Fig. 5 therefore contains appropriate inspection methods. Attenuation is controlled by two factors. Fade distance is used to specify the distance at which the full light intensity arrives, while fade power determines the falloff rate beyond the fade distance.

intensityInPoint() and *intensityInDirection()* methods computes concrete amount of energy in a space point. *stochasticEmission()* operation casts a ray stochastically with respect to the properties of the light source.

4.4 BSDFs

BSDF, Bidirectional Scattering Distribution Function, determines what portion of incoming energy is reflected back to the scene from a reflective surface or what portion is transmitted through a translucent material. There exist three variants of BSDF: BRDF, BSDF and BSSRDF. They differ only in details and thus the BSDF can be taken as their unified interface.

energy() method in Fig. 5 computes the portion of reflected or transmitted energy. *isIsotropic()* function distinguishes isotropic and anisotropic surfaces.

Some BRDFs do not depend on the outgoing direction but distributes the energy omnidirectionally. e.g. Lambertian function. They are called to be *perfectly diffuse*. Many real algorithms exploit this feature to accelerate energy distribution process. The *isDiffuse()* method informs an invoker about this property.

The rest of the *IBSDF* methods is used by stochastic algorithms of energy distribution.

Single virtual object can have assigned all three types of distribution functions. The basic coarse-grained inspection interface *IInspection* therefore contains inspection methods for these three variants of distribution functions. But all these variants share one fine-grained inspection interface *IBSDF*.

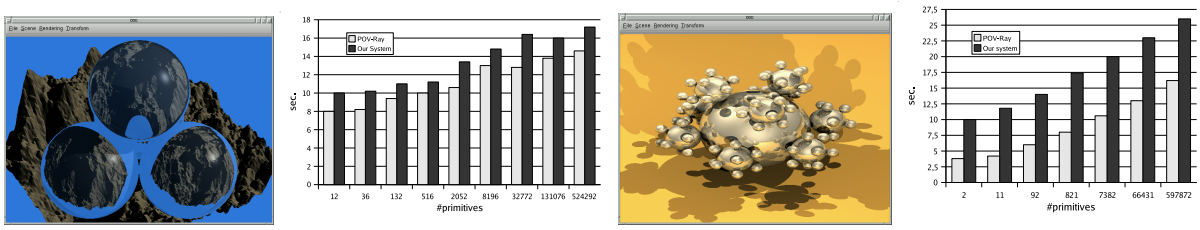


Figure 6: Fractal scenes used for efficiency tests: *Mountains* (left) and *Sphereflake* (right)

5 EXPERIMENTAL RENDERING ARCHITECTURE

To confirm the results of analysis we designed two experimental libraries.

The first library implements scene graph using inspection interface as discussed in this paper. Actually, the library does not represent real software component running under some kind of component system, e.g. CORBA, but it is a standalone C++ library implementing discussed interfaces. This library was developed in order to ensure that proposed concept is practical and functional. Translation of this standalone library into the real CORBA component is in progress.

The second library implements various rendering strategies, e.g. few variants of local illumination, Whitted ray tracing, Monte Carlo ray tracing and photon mapping. This library is used to check that proposed inspection interface of a scene graph is sufficiently general.

The most important result of this project is the existence of the unified component-based scene graph. But practical usage depends mainly on the rendering speed. Although our scene graph is not yet implemented as a real CORBA software component, we performed several preliminary efficiency tests.

Implementation of the local illumination is based on the OpenGL and thus we did not measure any significant decrease of performance in comparison with native OpenGL applications.

Ray tracing algorithm was compared with the POV-Ray system [Tea91]. Fig. 6 shows an overview of tested scenes and the rendering times. Tests were performed on Pentium 4 3.0GHz, 1GB RAM, image resolution 800x600. Results of the tests show a less efficiency of our system. The reason is that the unified inspection interface does not allow direct access into the scene graph and then forbids implementation of various acceleration tricks. Memory requirements are very similar for both the systems.

6 CONCLUSION AND FUTURE WORK

We discussed a unified interface of the scene traversal and inspection. This interface does not restrict implementation of the scene graph, just prescribe necessary inspection operations. On the other hand, proposed inspection operations are sufficiently general for wide range of rendering strategies, from real-time local illumination to photorealistic image synthesis. The interfaces therefore enable to develop a scene as an independent software component, which is very useful mainly in distributed and collaborative environments.

Inspection and traversal interfaces present only a fragment of all required functionality. Another unified scene graph interfaces, e.g. scene creation and maintenance, event handling, etc., have to be proposed.

7 ACKNOWLEDGMENTS

This work has been supported by the Czech Science Foundation under Contract No. GA 201/06/P247 and the Ministry of Education, Youth and Sports of the Czech Republic under the research program LC-06008.

REFERENCES

[Bus96] F. Buschmann. *Pattern-Oriented Software Architecture: A System of Patterns*. John Wiley & Sons, 1996.

[CD04] John Cheesman and John Daniels. *UML Components*. Addison-Wesly, 2004.

[DH02] J. Döllner and K. Hinrichs. A generic rendering system. *IEEE Trans. Visualization & CG*, 8(2):99–118, 2002.

[Fel95] D. W. Fellner. Mrt - an extensible platform for 3d image synthesis. Computer Graphics Lab., Dept. of Computer Science, University of Bonn, Germany, December 1995.

[GHJV95] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns Elements of Reusable Object-Oriented Software*. Addison-Wesley, 1995.

[Opea] Open scene graph. http://openscenegraph.sourceforge.net/.

[Opeb] Opensg. http://www.opensg.org/.

[OS03] Radek Ošlejšek and Jiří Sochor. Generic graphics architecture. In *Theory and Practice of Computer Graphics*, pages 105–112. IEEE Computer Society, June 2003.

[OS05] Radek Ošlejšek and Jiří Sochor. A flexible, low-level scene graph traversal with explorers. In *Spring Conference on Computer Graphics*, pages 194–201. Bratislava : Comenius University, May 2005.

[Rei02] Dirk Reiners. A flexible and extensible traversal framework for scenegraph systems. In *OpenSG Symposium*, 2002.

[SS95] P. Slusallek and H.-P. Seidel. Vision - an architecture for global illumination calculations. *IEEE Trans. Visualization & Computer Graphics*, 1(1), 1995.

[Tea91] POV Team. Persistency of vision ray tracer (pov-ray), version 1.0. Technical report, 1991.

[WNDS99] M. Woo, J. Neider, T. Davis, and D. Shreiner. *OpenGL Programming Guide*. Addison-Wesley, 1999.

POSTER: ADAPTATIVE TEMPLATES IN BIOMETRIC AUTHENTICATION

Ricardo García Noval
University of the Balearic Islands
Son Lledó Building,
Valldemossa Road, km 7.5.
07122, Palma, Balearic Islands, Spain
ricardo.garcia@uib.es

Francisco Perales López
Computer Graphics, Vision, and Artificial Intelligence Group,
University of the Balearic Islands
Anselm Turmeda Building,
Valldemossa Road, km 7.5
07122, Palma, Balearic Islands, Spain
paco.perales@uib.es

ABSTRACT

Biometric authentication systems are usually based on features extraction. Features are a collection of measurable details, obtained from the biometric trait that defines the identity of a certain person. This collection of data is known as template, and it's stored in the database. The acquired biometrics quality must be controlled in order to model the identity of the individual in a unique and distinct way. The creation and update of templates is a critical task for the correct use of a biometric application. In this paper we propose the implementation of a model that, using biometric-independent tools, intends to update, select and improve the templates stored in the database, in what we have called "adaptive biometric templates". It has been tested with a fingerprint biometric database of 60 users. We have obtained an average improvement over traditional templates of 26% for FMR and of 53% for FNMR, we consider these results very successful.

Keywords

Biometric authentication, Template selection, Multi-modal interfaces.

1. INTRODUCTION

A generic biometric system can be defined with a very simple working paradigm that is used in most applications and commercial solutions.

First, when a security system is being installed, the biometric traits of all users with access to the resource must be acquired, thus creating a database that models the identity of the individuals by means of templates.

This step or working mode is known as enrollment. Each time a new genuine user wants to access the resource we must enroll him, acquiring his biometrics. Currently, user's identities stored as templates in the database don't change after enrollment and remain invariable in the database.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Copyright UNION Agency – Science Press, Plzen, Czech Republic.

Subsequently, users who want to enter the system must show their biometrics for comparison with the stored templates and verify that his identity is found in the database and, thus, grant him access. At this moment biometric security system works in "authentication" mode.

With this working scheme the enrollment process is essential since it's the only moment when templates stored in the database are modified. These templates are the main link between the designed user's model and his real identity, and they remain static as long as we don't acquire another set of biometric traits, process that can be annoying to the user.

But achieving biometric templates that represent the user's identity in an accurate way can be a difficult task as a consequence of several factors: it is not easy to measure the quality of a biometric trait (the only "objective" values are FMR and FNMR [1]), the user's biometrics are not in good condition at the time of acquisition (e.g., dry fingers in a fingerprint system or irritaded eyes in iris detection [2]) and, though it's not a desirable feature, some biometrics, like voice, could change with time.

Moreover, if the number of database users is high, a "manual" control over biometric templates quality that rejects incorrect traits and acquires a new (correct) set of traits could not be possible.

Two problems surge from the limitation explained above: first, the need of updating biometric templates in order to accommodate them to the trait's real evolution in the individuals and second the proper (correct) selection of templates in order to turn down deficiencies or errors in acquisition, therefore reducing error rates associated with authentication.

In the next sections, we propose a new adaptative biometric template system. The proposed system improves the update template process increasing inter-class differences and reducing intra-class differences, using the standard authentication procedure to attain more precise ROC curves. Also our system is designed in an open way, so that future new templates from other biometrics features can also be included and therefore offer a multibiometric approach.

2. ADAPTATIVE BIOMETRIC TEMPLATES

A lot of schemes that bring successful solutions to these problems have been implemented. X. Jiang and W. Ser [3] propose a recursive technique for improving biometric templates that compute average values of minutiae included in each instance of a fingerprint template. Other known methods are biometric independent, like the ones proposed by Jain, Ross and Uludag, [4] that use binary trees between the different instances that form a template (dendograms, DEND method) or average distances of similarity between these instances (MDIST). Scheidat et. al. [5], on the other hand, focus the update problem as if it was a "cache" pages' issue. They propose the use of classics algorithms (FIFO, LRU, clock) for replacing the biometric traits that became obsolete. The paradigm that we will show next (implemented in the structure of a multimodal biometric library [9]) doesn't intend to replace none of the mentioned above techniques, whose efficiency and performance has been proved. The main idea is to provide an automatic tool for supporting adaptative biometric templates that, using the information obtained from the access of the different users, could make the templates stored in the database more different between them and more similar to the real trait of the individual.

The working scheme until now was:

1. *Acquire user's biometrics and store its features in a biometric template in the database.*
2. *When a users tries to access:*
   a. *Verify that the biometric given is similar to the one stored in the template.*

We propose the following:

1. *Acquire user's biometrics and store its features in a biometric template in the database*
2. *When a users tries to access:*
   a. *Verify that the biometric given is similar to the one stored in the template*
   b. *Store the biometric trait used in the access.*
3. *Periodically and for each user:*
   a. *Evaluate the quality of the biometric traits used in the access.*
   b. *If the quality of this traits is better then include them in the template, else reject them.*

In order to implement this system we need to use a second biometric database, parallel to the main database. This second database stores the different "attempts to access" that occur when the security system works in authentication mode, for its later evaluation. The information stored is:

- Date and time of the access.
- Name of the user whose identity was claimed in the access.
- Set of biometric features given in the access.

These entries are stored in different lists. First, for each user we store a list of all the successful accesses that he has made, in order to examine directly the evolution of the biometric trait along the different genuine entries. Second, we store a list of users than haven't achieved the access to the system, representing a database of potential impostors. We want the stored users to be as less similar as we can to these impostors.

Second chance verification

To support adaptative biometric templates we need a method that distinguishes user's access that has produced a false non match and the access of an impostor that tries to pass off as a genuine user. If we keep the two lists described above and include genuine user's features to the stored template we can reject features of potential impostors.

That's why we have to implement what we call second chance verification algorithm. The goal of this algorithm is to give, in the moment of verification, a classification mechanism that could speed up the template selection process and the computing rate of the false non match. This algorithm uses two authentication mechanisms *A* and *B*. Mechanism *A*, which we take as predominant, is the biometric trait that we will make all the improvements over, and *B* would only be used in case of error or rejection in verification using *A*. Mechanism *B* could be a password, an ID card or even another biometric trait (so we would have a multimodal biometric system [6]). The algorithm is so simple:

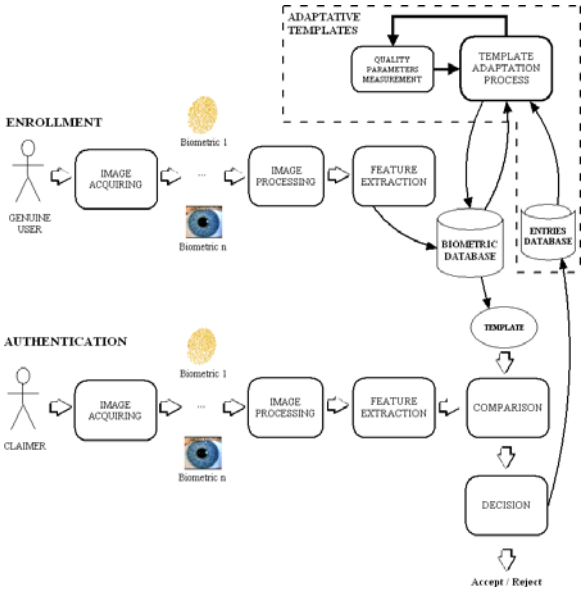


Figure 1. Working scheme of a biometric system with adaptative templates

1. *We use biometric **A** for verifying the identity of the user:*
   a. In case of accept by A =>
      *i. The access is **granted** to the user and the verification is over.*
      ii. The access data is stored in the list of genuine access of the user.
   b. In case of reject using A => step 2.
2. *We use mechanism **B** for verifying the identity of the user:*
   a. In case of accept by B =>
      *i. The access is **granted** to the user and this entry counts as a **false non match** for biometric A.*
      ii. The access data is stored in the list of genuine access of the user.
   b. In case of reject by B =>
      *i. The access is **denied** to the user.*
      ii. The access data is stored in the list of access of impostor users.

If we take a close look to the algorithm we can see the need of a second mechanism that could tell us the difference between an impostor who tries to introduce his traits in the stored templates (in an illegal way) and a genuine user that has suffered a false non match but that can be authenticated using a second mechanism. If a multimodal biometric system is used, the choice of main and secondary biometric must be rational. Logic tells us to use as mechanism *B* a more robust biometric trait, with less loss of quality in the templates but slower than biometric *A* in verification or identification process. *A* doesn't need to have a outstanding initial performance, because adaptative templates would improve error rates. In this way we could avoid the two biometric disadvantages and make the most of them.

Quality parameters measurement

The adaptative templates scheme needs a module that measures parameters and gives an idea of the global quality of the biometric features, in order to select those accesses with a quality potential better than the current stored template. The implemented system uses the following parameters for each template and user's access:

Similarity with the other users *(SO – Similarity Others)*: is the average of the similarity scores obtained in the comparison between the user and the rest of the users stored in the database and the users found in the impostor access list. The smaller the value, the more accurate is the biometric trait. A small value indicates a great distinction with the rest of the users stored in the database and potential impostors. This small similarity score doesn't overcome the threshold. Minimization of this value reduces the false match rate (FMR).

Let s(x,y) be the matching similarity score between users x and y, given N genuine users, $G_1,G_2,\ldots,G_n$, and M impostors $I_1,I_2,\ldots,I_m$, we define $SO(G_i)$ as:

$$SO(G_i) = \frac{\sum_{j=1, j\neq i}^{j=N} s(G_i, G_j) + \sum_{k=1}^{k=M} s(G_i, I_j)}{N+M}$$

Similarity with himself *(SS – Self Similarity)*: is the average of the similarity scores obtained in the comparison between the user and the different access he has made. The greater the value, the more accurate is the biometric trait. A large value indicates similarity between the different stored versions of the user's biometric trait. This large similarity score overcomes the threshold. Maximization of this value reduces the false non match rate (FNMR).

Let s(x,y) be the matching similarity score between users x and y, given N genuine users, $G_1,G_2,\ldots,G_n$, and P(i) successful and genuine access of user i, $A(i)_1, A(i)_2,\ldots, A(i)_{P(i)}$, we define $SS(G_i)$ as:

$$SS(G_i) = \frac{\sum_{j=1}^{j=P(i)} s(G_i, A(i)_j)}{P(i)}$$

Template adaptation process

This process improves stored template's quality by checking the stored biometric features for each user's access.A singularity that we must not forget is that a biometric template could use several instances or a set of features of the biometric traits (such as the 3 biometric user's fingerprints in out test) and generalizes them into a single template, or simply use a single instance and template. This fact, as we will see below, bears upon the way the biometric template is updated in the database.The process work with this algorithm:

 ISBN 978-80-86943-17-6

For each genuine user G_i, i=1,…,N in the database:

1. *Assign quality parameters of templates stored in biometric database: best_so=SO(G_i) and best_ss=SS(G_i)*
2. *For each genuine access $A(i)_j$, j=1,…,P(i) of the ith user*
   a. *If template only uses one instance of the trait*
      i. *Compute quality parameters of the biometric features used in the access: SO($A(i)_j$), SS($A(i)_j$)*
      ii. *If SO($A(i)_j$)<best_SO and SS($A(i)_j$)<best_SS*
         1. *Replace template stored in database with features in access $A(i)_j$.*
         2. *Assign best_SO= SO($A(i)_j$), best_SS= SS($A(i)_j$)*
   b. *If template uses several instances of the trait*
      i. *Assign replaced_feature=0*
      ii. *For each instance or set of features F_k, i=1,…,L of the stored template:*
         1. *Obtain generalized template T_{ijk} replacing F_k instance with features used in access $A(i)_j$*
         2. *Compute quality parameters of resulting template SO(T_{ijk}) and SS(T_{ijk})*
         3. *If SO(T_{ijk})<best_SO and SS(T_{ijk})<best_SS*
            a. *Assign replaced_feature=k*
            b. *Assign best_SO= SO(T_{ijk}), best_SS= SO(T_{ijk})*
      iii. *If replaced_feature>0*
         1. *Replace template stored in database with template $T_{ijreplaced_feature}$*

In short, it's basically a maximum algorithm whose goal is to store in the database the combination of features (template) that has given the better value of the quality parameters (low SO and high SS). The computational complexity of this algorithm is polynomial, though that's not a critical factor.

3. VALIDATION & RESULTS

In order to test the adaptative biometric templates system we have emulated a scenario similar to the one found in a small university or research center lab or in an office with confidential information. The system has been tested during a two month period using a database of 60 users, 15 of them with periodical access. The total number of accesses has been of 100, and since we use 3 fingerprints in enrolment for each user, the total number of samples is higher than 250. The alternative method of authentication used (*B*) has been password, due to its simplicity in implementation and testing.

4. CONCLUSIONS

In this paper we propose a solution to a common problem in most biometric systems, the update and selection of biometric templates in a database. The solution developed here offers a new paradigm of biometric authentication that intends to achieve two goals at a time: the evolution of stored templates with the real trait of the individual and the selection of those features that are characteristic of the individual (reducing intra-class differences) and that also differentiates him from other individuals (increasing interclass differences).

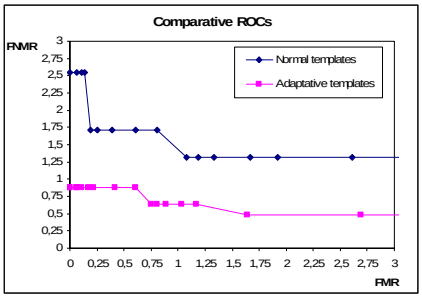


Figure 2. Comparative ROC on performance between normal and adaptative emplates

The system proposed here has been validated with real users in a university environment, obtaining successful and promising results. Furthermore, the results obtained encourages us to study this method in depth, combining it with others such as, multimodal biometric algorithms based on user-specific parameters, in order to make greater improvements, and test its performance in bigger scenarios with a higher level of access. This work has been supported by the Spanish Dirección General de Investigación del Ministerio de Educación, Ciencia y Tecnología through the TIN2004-07926 and TIN2007-67993 projects.

5. BIBLIOGRAPHY AND REFERENCES

[1] A. K. Jain; R. Bolle; S. Pankanti. *"Biometrics: Personal Identification in Networked Society"*, Kluwer Academic Publishers. ISBN 0-7923-8345-1. United States of America, 1999.

[2] R. Bolle; J.H. Cornell; S. Pankanti; N. K. Ratha; A. W. Senior. *"Guide to Biometrics"*, Springer-Verlag. ISBN 0-387-40089-3. United States of America, 2004.

[3] X. Jiang; W. Ser. *"Online Fingerprint Template Improvement"*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24. United States of America, 2002.

[4] A.K. Jain; U. Uludag; A. Ross. *"Biometric Template Selection: A Case Study in Fingerprints"*, in Proc. of 4th Int'l Conference on Audio- and Video-Based Person Authentication (AVBPA). Guildford (UK), 2003.

[5] T. Scheidat; A. Makrushin; C. Vielhauer. *"Automatic Template Update Strategies for Biometrics"*, Otto-von-Guericke University of Magdeburg. Germany, 2007.

[6] A.K. Jain; A. Ross. *"Multimodal Biometrics: An Overview"*, in Proc. of 12th European Signal Processing Conference (EUSIPCO). Viena (Austria), 2004.

[7] A.K. Jain; A. Ross. *"Learning user-specific parameters in a multibiometric system"*, in Proc. International Conference on Image Processing (ICIP). Rochester, New York (USA), 2002.