ISSN 2464–4617 (print) ISSN 2464–4625 (CD-ROM)

CSRN 2603

(Ed.)

- **Vaclav Skala**
  University of West Bohemia, Czech Republic

Computer Science Research Notes

24th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision
WSCG 2016
Plzen, Czech Republic
May 30 – June 3, 2016

Proceedings

WSCG 2016

Posters Proceedings

ISSN 2464–4617 (print) ISSN 2464–4625 (CD-ROM)

ISSN 2464–4617 (print) ISSN 2464–4625 (CD-ROM)

CSRN 2603

(Ed.)

- **Vaclav Skala**
  University of West Bohemia, Czech Republic

Computer Science Research Notes

24th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision
WSCG 2016
Plzen, Czech Republic
May 30 – June 3, 2016

Proceedings

WSCG 2016

Posters Proceedings

Vaclav Skala – Union Agency

ISSN 2464–4617 (print) ISSN 2464–4625 (CD-ROM)

Computer Science Research Notes
CSRN 2603

Editor-in-Chief: Vaclav Skala
c/o University of West Bohemia
Univerzitni 8
CZ 306 14 Plzen
Czech Republic
skala@kiv.zcu.cz *http://www.VaclavSkala.eu*

Managing Editor: Vaclav Skala

Publisher & Author Service Department & Distribution:
Vaclav Skala - UNION Agency
Na Mazinach 9
CZ 322 00 Plzen
Czech Republic
Reg.No. (ICO) 416 82 459

ISSN 2464-4617 (Print) **ISSN 2464-4625** (CD/DVD)
ISBN 2464-4625 (CD ROM**)** **ISBN 978-80-86943-57-2** (CD/-ROM)

WSCG 2016
International Program Committee

Benger,Werner (United States)
Bilbao,Javier,J. (Spain)
Bittner,Jiri (Czech Republic)
Buehler,Katja (Austria)
Chaudhuri,Debasis (India)
Chmielewski,Leszek (Poland)
Coquillart,Sabine (France)
Daniel,Marc (France)
de Geus,Klaus (Brazil)
Dingliana,John (Ireland)
Drechsler,Klaus (Germany)
Durikovic,Roman (Slovakia)
Falcidieno,Bianca (Italy)
Feito,Francisco (Spain)
Ferguson,Stuart (United Kingdom)
Flaquer,Juan (Spain)
Garcia Hernandez,Ruben Jesus (Germany)
Giannini,Franca (Italy)
Gudukbay,Ugur (Turkey)
Guthe,Michael (Germany)
Jeschke,Stefan (Austria)
Juan,M.-Carmen (Spain)
Kim,H. (Korea)
Max,Nelson (United States)
Molla,Ramon (Spain)
Montrucchio,Bartolomeo (Italy)
Muller,Heinrich (Germany)
Murtagh,Fionn (United Kingdom)
Pan,Rongjiang (China)
Pedrini,Helio (Brazil)
Platis,Nikos (Greece)
Renaud,christophe (France)
Richardson,John (United States)
Ritter,Marcel (Austria)
Rojas-Sola,Jose Ignacio (Spain)
Sanna,Andrea (Italy)
Segura,Rafael (Spain)
Skala,Vaclav (Czech Republic)
Slavik,Pavel (Czech Republic)
Sousa,A.Augusto (Portugal)
Szecsi,Laszlo (Hungary)
Teschner,Matthias (Germany)
Tokuta,Alade (United States)
Trapp,Matthias (Germany)
Viola,Ivan (Austria)
Wu,Shin-Ting (Brazil)
Wuensche,Burkhard,C. (New Zealand)
Wuethrich,Charles (Germany)
Zara,Jiri (Czech Republic)
Zemcik,Pavel (Czech Republic)

WSCG 2016
Board of Reviewers

Abad,Francisco (Spain)
Adzhiev,Valery (United Kingdom)
Agathos,Alexander (Romania)
Ahmad,Khurshid (Ireland)
Akleman,Ergun (United States)
Amditis,Angelos (Greece)
Ammi,Mehdi (France)
Ammi,Mehdi (France)
Ariu,Davide (Italy)
Assarsson,Ulf (Sweden)
Aveneau,Lilian (France)
Ayala,Dolors (Spain)
Backfrieder,Werner (Austria)
Barthe,Loic (France)
Battiato,Sebastiano (Italy)
Baum,David (Germany)
Benes,Bedrich (United States)
Benger,Werner (United States)
Benoit,Crespin (France)
Biasotti,Silvia (Italy)
Bilbao,Javier,J. (Spain)
Biri,Venceslas (France)
Birra,Fernando (Portugal)
Bittner,Jiri (Czech Republic)
Bosch,Carles (Spain)
Bouatouch,Kadi (France)
Boukaz,Saida (France)
Bourdin,Jean-Jacques (France)
Bourke,Paul (Australia)
Bouville,Christian (France)
Bruckner,Stefan (Austria)
Bruder,Gerd (Germany)
Brun,Anders (Sweden)
Bruni,Vittoria (Italy)
Brunnett,Guido (Germany)
Buehler,Katja (Austria)
Bulo,Samuel Rota (Italy)
Buriol,Tiago Martinuzzi (Brazil)
Cakmak,Hueseyin (Germany)
Camahort,Emilio (Spain)
Casciola,Giulio (Italy)
Chaine,Raphaelle (France)
Chaudhuri,Debasis (India)
Chen,Falai (China)
Chmielewski,Leszek (Poland)
Choi,Sunghee (Korea)
Chover,Miguel (Spain)
Chrysanthou,Yiorgos (Cyprus)
Chuang,Yung-Yu (Taiwan)
Cline,David (United States)
Coquillart,Sabine (France)
Corcoran,Andrew (Ireland)
Cosker,Darren (United Kingdom)
Daniel,Marc (France)
Daniels,Karen (United States)
de Amicis,raffaele (Italy)
de Geus,Klaus (Brazil)
de Oliveira Neto,Manuel Menezes (Brazil)
De Paolis,Lucio Tommaso (Italy)
Debelov,Victor (Russia)
Dingliana,John (Ireland)
Doellner,Juergen (Germany)
Dokken,Tor (Norway)
Drechsler,Klaus (Germany)
Durikovic,Roman (Slovakia)
Eisemann,Martin (Germany)
Erleben,Kenny (Denmark)
Falcidieno,Bianca (Italy)
Faudot,Dominique (France)
Feito,Francisco (Spain)
Ferguson,Stuart (United Kingdom)
Fiorentino,Michele (Italy)
Flaquer,Juan (Spain)
Fuenfzig,Christoph (Germany)
Gain,James (South Africa)
Galo,Mauricio (Brazil)
Garcia Hernandez,Ruben Jesus (Germany)
Garcia-Alonso,Alejandro (Spain)
Gavrilova,M. (Canada)
Gianelli,Carlota (Germany)
Giannini,Franca (Italy)

Gobron,Stephane (Switzerland)
Goebel,Martin (Germany)
Gonzalez,Pascual (Spain)
Grau,Sergi (Spain)
Gu,Xianfeng (United States)
GuĂŠrin,Eric (France)
Gudukbay,Ugur (Turkey)
Guthe,Michael (Germany)
Habel,Ralf (Switzerland)
Hall,Peter (United Kingdom)
Hansford,Dianne (United States)
Haro,Antonio (United States)
Hast,Anders (Sweden)
Hauser,Helwig (Norway)
Havemann,Sven (Austria)
Havran,Vlastimil (Czech Republic)
Hege,Hans-Christian (Germany)
Hernandez,Benjamin (United States)
Herout,Adam (Czech Republic)
Hicks,Yulia (United Kingdom)
Hildenbrand,Dietmar (Germany)
Hinkenjann,Andre (Germany)
Horain,Patrick (France)
Horain,Patrick (France)
House,Donald (United States)
Ihrke,Ivo (Germany)
Iwasaki,Kei (Japan)
Jeschke,Stefan (Austria)
Jiang,Jianmin (China)
Jones,Mark (United Kingdom)
Juan,M.-Carmen (Spain)
Juettler,Bert (Austria)
Kanai,Takashi (Japan)
Kim,H. (Korea)
Klosowski,James (United States)
Kohout,Josef (Czech Republic)
Kolcun,Alexej (Czech Republic)
Krueger,Jens (Germany)
Kumar,Subodh (India)
Kurillo,Gregorij (United States)
Kurt,Murat (Turkey)
Kyratzi,Sofia (Greece)
Lanquentin,Sandrine (France)
Larboulette,Caroline (France)
Lee,Jong Kwan Jake (United States)
Lengyel,Eric (United States)
Lien,Jyh-Ming (United States)
Lindow,Norbert (Germany)
Liu,SG (China)
Liu,Damon Shing-Min (Taiwan)
Lopes,Adriano (Portugal)
Loscos,Celine (France)
Lucas,Laurent (France)
Lutteroth,Christof (New Zealand)
Maciel,Anderson (Brazil)
Maddock,Steve (United Kingdom)
Magnor,Marcus (Germany)
Manak,Martin (Czech Republic)
Mandl,Thomas (Germany)
Manzke,Michael (Ireland)
Mas,Albert (Spain)
Masia,Belen (Spain)
Masood,Syed Zain (United States)
Matey,Luis (Spain)
Matkovic,Kresimir (Austria)
Max,Nelson (United States)
McDonnell,Rachel (Ireland)
McKisic,Kyle (United States)
Meng,Weiliang (China)
Mestre,Daniel,R. (France)
Metodiev,Nikolay Metodiev (United States)
Meyer,Alexandre (France)
Mokhtari,Marielle (Canada)
Molina Masso,Jose Pascual (Spain)
Molla,Ramon (Spain)
Montrucchio,Bartolomeo (Italy)
Morigi,Serena (Italy)
Muller,Heinrich (Germany)
Murtagh,Fionn (United Kingdom)
Myszkowski,Karol (Germany)
Niemann,Henrich (Germany)
Nishita,Tomoyuki (Japan)
Okabe,Makoto (Japan)
Oliveira, Jr.,Pedro Paulo (Brazil)
Oyarzun Laura,Cristina (Germany)
Pala,Pietro (Italy)
Pan,Rongjiang (China)
Papaioannou,Georgios (Greece)
Paquette,Eric (Canada)
Pasko,Alexander (United Kingdom)
Pasko,Galina (United Kingdom)

Pastor,Luis (Spain)
Patane,Giuseppe (Italy)
Patow,Gustavo (Spain)
Pedrini,Helio (Brazil)
Pereira,Joao Madeiras (Portugal)
Perret,Jerome (France)
Peters,Jorg (United States)
Pettre,Julien (France)
Peytavie,Adrien (France)
Pina,Jose Luis (Spain)
Platis,Nikos (Greece)
Plemenos,Dimitri (France)
Post,Frits,H. (Netherlands)
Poulin,Pierre (Canada)
Praktikakis,Ioannis (Greece)
Puig,Anna (Spain)
Puppo,Enrico (Italy)
Rafferty,Karen (United Kingdom)
Reisner-Kollmann,Irene (Austria)
Renaud,christophe (France)
Renaud,christophe (France)
Reyes-Lecuona,Arcadio (Spain)
Richardson,John (United States)
Ritschel,Tobias (Germany)
Ritter,Marcel (Austria)
Rojas-Sola,Jose Ignacio (Spain)
Rokita,Przemyslaw (Poland)
Runde,Christoph (Germany)
Ruther,Heinz (South Africa)
Sacco,Marco (Italy)
Sadlo,Filip (Germany)
Sakas,Georgios (Germany)
Salvetti,Ovidio (Italy)
Sanna,Andrea (Italy)
Santos,Luis Paulo (Portugal)
Sapidis,Nickolas,S. (Greece)
Savchenko,Vladimir (Japan)
Schultz,Thomas (Germany)
Schumann,Heidrun (Germany)
Segura,Rafael (Spain)
Seipel,Stefan (Sweden)
Sellent,Anita (Switzerland)
Shesh,Amit (United States)
Sik-Lanyi,Cecilia (Hungary)
Slavik,Pavel (Czech Republic)
Sochor,Jiri (Czech Republic)
Solis,Ana Luisa (Mexico)
Sommer,Bj?rn (Germany)
Sourin,Alexei (Singapore)
Sousa,A.Augusto (Portugal)
Sramek,Milos (Austria)
Sreng,Jean (France)
Staadt,Oliver (Germany)
Stricker,Didier (Germany)
Stroud,Ian (Switzerland)
Subsol,Gerard (France)
Suescun,Angel ()
Sunar,Mohd-Shahrizal (Malaysia)
Sundstedt,Veronica (Sweden)
Svoboda,Tomas (Czech Republic)
Szecsi,Laszlo (Hungary)
Tang,Min (China)
Tang,Qian (China)
Taubin,Gabriel (United States)
Tavares,Joao Manuel R.S. (Portugal)
Teschner,Matthias (Germany)
Theussl,Thomas (Saudi Arabia)
Tian,Feng (United Kingdom)
Tobler,Robert (Austria)
Todt,Eduardo (Brazil)
Tokuta,Alade (United States)
Torrens,Francisco (Spain)
Trapp,Matthias (Germany)
Triantafyllidis,Georgios (Greece)
Tytkowski,Krzysztof (Poland)
Umlauf,Georg (Germany)
Vanderhaeghe,David (France)
Vasa,Libor (Czech Republic)
Vazquez,Pere-Pau (Spain)
Vazquez,Pere Pau (United States)
Viola,Ivan (Austria)
Vitulano,Domenico (Italy)
Vosinakis,Spyros (Greece)
Walczak,Krzysztof (Poland)
WAN,Liang (China)
Wang,Charlie,C.L. (Hong Kong SAR)
Weber,Andreas (Germany)
Wenger,Raphael (United States)
Westermann,Ruediger (Germany)
Wu,Enhua (China)
Wu,Shin-Ting (Brazil)
Wuensche,Burkhard,C. (New Zealand)

Wuethrich,Charles (Germany)
Xin,Shi-Qing (Singapore)
Yoshizawa,Shin (Japan)
YU,Qizhi (United Kingdom)
Yue,Yonghao (Japan)
Zachmann,Gabriel (Germany)
Zalik,Borut (Slovenia)
Zara,Jiri (Czech Republic)
Zemcik,Pavel (Czech Republic)
Zhang,Xiaopeng (China)
Zhang,Xinyu (United States)
Zhu,Ying (United States)
Zillich,Michael (Austria)
Zitova,Barbara (Czech Republic)
Zwettler,Gerald (Austria)

WSCG 2016

Poster's Proceedings

Contents

ISSN 2464-4617 (print) WSCG 2016 - 24th Conference on Computer Graphics, Visualization and Computer Vision 2016
ISSN 2464-4625 (CD-ROM)

Benefits of Tablet Interfaces for Immersive Visualization in Information Visualization

Luisa Wurm[1] Rubén García[2] Christoph Anthes[2,1] Dieter Kranzlmüller[1,2] Wolfgang Höhl[1]
1 Ludwig-Maximilians-Universität,
Geschwister-Scholl-Platz 1. D-80539, Munich, Germany.
2 Leibniz Supercomputing Centre, Bavarian Academy of Sciences
Bolzmannstr 1. D-85748, Garching, Germany
l.wurm@campus.lmu.de, garcia@lrz.de, anthes@lrz.de, kranzlmueller@lrz.de, hoehl@cip.ifi.lmu.de

ABSTRACT

This paper describes how tablets can be used to increase the interaction possibilities of virtual reality applications running on CAVE-like[1] environments. A use case is presented using the genome comparison software 3DScover. The tablet interface has been introduced in a user-centric development environment. Multiple user studies have guided the evolution of the interface to ensure optimal ease of use. Our results indicate that the tablet provides useful input and output capabilities which are cumbersome to introduce in typical CAVE-like installations. The results should be extrapolable to other information visualization applications using virtual reality technology.

Keywords

Tablet Interface, Immersive Virtual Reality, Information Visualization, User Study, Genomics

1 INTRODUCTION

Virtual reality (VR) devices submerge the user in a virtual world and allow him/her to explore and manipulate the world in an intuitive manner. In the case of information visualization (infovis), the virtual world is composed of two- or three-dimensional objects corresponding to abstract information. The ease of visualization of 3D plots makes the tool quite useful in the study The interaction devices used in VR installations are varied, but often tracked joysticks or wands are used. In some cases, a PDA or tablet has been used to provide a more complex graphical user interface (GUI), but the use of such devices is relatively rare.

In this paper, we present a use case for a VR application which has been augmented with a tablet interface. The application belongs to the genomic comparison field and makes extensive use of information visualization. We claim that the use of a tablet can be of special value for immersive VR applications using semi-immersive CAVE-like[1] or other projection displays in the information visualization field, as the interaction with the datasets follows more abstract concepts which are not easily mappable to simple hand gestures. In particular, tablets are ideally suited for complex system control and textual input, while their performance in navigation, 3D manipulation and selection tasks is similar to that of standard wand or joystick devices.

GUI-based interaction, while possible within VR, suffers of quality degradation in the presence of text. The availability of a device which can display high-quality text, menus, graphics or animations (including zooms of the virtual world) greatly enhances the interaction possibilities. A high precision touch interface increases accuracy and speed, and is very useful for text input.

To evaluate this claim, we have performed user studies to compare the ease of use, the performance and the satisfaction of the users with the different input devices and interaction methods. While there is some variability in the responses of the different users, on average users showed that the tablet was a valuable resource.

Section 2 describes related work. Section 3 introduces our prototype, and describes the user studies performed to evaluate it. Finally, conclusions are presented.

[1] CAVE™ is a registered trademark of the University of Illinois' Board of Trustee. The term is used here to generically refer to CAVEs™ and CAVE-like displays.

2 RELATED WORK

Personal Digital Assistants (PDAs) and tablets have previously been used to help control VR systems, by displaying 2D interfaces in the PDA and 3D ones in the CAVE [Wat99], or for training in industrial settings [Med13]. The medical field also has some applications [Mor14]. A smartphone has been used for e-learning

Figure 1: CAVE application with tablet interface.

[Bed09]. Finally, a library for creating general GUIs in tablets for VR environments has also been developed [Har02].

There is a high quantity of VR applications in information visualization for genomics data. Examples include a dividing road metaphor to symbolize decision taking and a lab bench metaphor for information display, within gene expression studies [Nis04], and Juicebox [Rao14], a commercial software which displays interaction among loci within the human genome.

3DScover is a suite of graphical applications to perform comparative genomic analysis. The desktop version [Tuk14a] shows a 3D data structure which encompasses all the information available in the most often used two dimensional representations, and optionally exposes the previous representations (dot plot, gradient view, and linear representation) as projections. Multiple operations on the genomic datasets are implemented. The CAVE version [Gar16a] displays the 3D data structure and projections in an inmersive environment, and allows basic interaction using a wand.

3 TABLET INTERFACE FOR THE CAVE VERSION OF 3DSCOVER

The early prototype of 3DScover-CAVE used one wall to provide the interface in form of textual information and clickable buttons to control the application (figure 1, top right). This correspondence between the physical and virtual locations allows clear text to be displayed. However, the screen resolution limits text to simple words. Buttons need to be large to be easily clicked by using the wand. Consequently, only a small subset of the desktop functionality was implemented.

The wand used in our installation has four digital buttons and an analog joystick. Although in principle it is easy to map the different interaction modes to the different options, the users found it too complex to remember which button mapped to which action. The possibility to create virtual buttons on the tablet screen was very welcome by final users.

We used the tablet capabilities to create a new high-precision selection method, although we also provide the wand-like selection method by attaching a tracker to the tablet. We also added advanced functionality which can be easily accessed by using the tablet features. The development of this improved version of the software was guided by a set of user studies to ensure usability. A final study studied the satisfaction of users with the software. Figure 1 shows our final system.

The tablet interface is composed of a series of clickable tabs. The visualization tab helps the users familiarize themselves with the 3D representation and its three 2D projections. The selection tab (figure 3) controls which of the genome is the reference. The other genomes are pairwise compared to it. Subsets of the genomes can be selected here based on their position in the genome, and the application will show the relevant subsets in the rest of the genomes. The edit tab allows the user to perform rigid transformations on the 3D object, allowing them to focus on the relevant parts of the object they want to study. The search tab allows the user to search the genomes by gene name or function. Additional tabs to control connection settings and file loadings also exist.

4 EVALUATION

A description of the user studies performed to validate the interface follows (participants were new to VR).

4.1 Interface Design Questionnaire

The first user study took care of choosing easy-to-understand names and icons for the different possible actions and the navigation across the different modes of the application. The users were shown different possibilities and chose the ones they found best. The most commonly chosen sets of icons and names were used in the following prototypes. In particular, most users were keen on having textual description in addition to icons. The navigation was also tailored to the responses of the users (two thirds of the users preferred tabs versus one third which preferred a slide-in interface).

This user study used a web-based questionnaire. 35 participants took part, 22 (63%) of them male and 13 (37%) female. The average age was 25.

4.2 Mock-up analysis

A mock-up was created which allowed users to test the interaction possibilities The second user study used a tablet with this mock-up application to verify the usability of the interface.

10 users took part in the mock-up analysis, Four (40%) of them male and six (60%) of them female. The average age was 29,3. Since we were targeting an android tablet, we used when possible the native icons for standard actions. A minority of users (who use IOS or windows phone) had issues with the menu icon.

The users provided comments on possible improvements in most of the interaction modes, which were

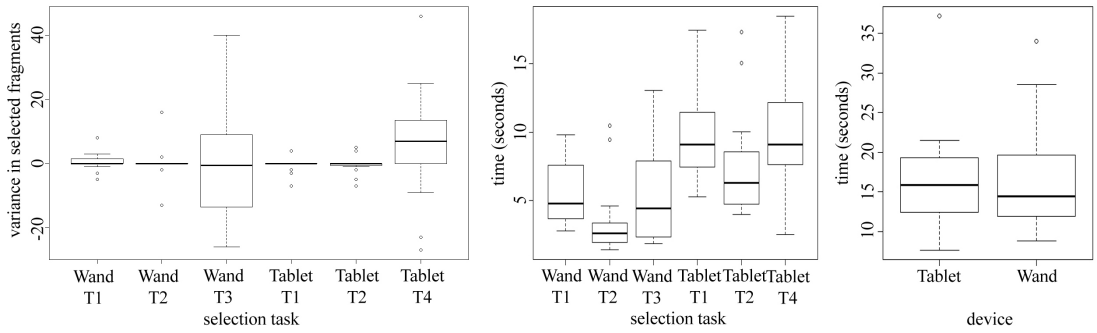


Figure 2: Time (left) and accuracy (center) for tasks 1-3. Time for task 4.

taken into account in the final version of the software. In particular, most of the comments regarded adding additional information to the search, sliders and data input fields. The users were asked to evaluate the interface in a 5 point Likert scale. The mean score was 4.1, validating our design. See figure 3 for the resulting design.

4.3 Final User Study

The final user study aimed to validate whether the use of a tablet would improve accuracy and user satisfaction with the software. 20 participants took part, 17 (85%) of them male and three (15%) female, with an age average of 28. The following hypothesis were tested:

H1 The selection task with the wand is performed faster than with the tablet.

H2 The selection with the tablet is more precise than with the wand.

H3 The user will be satisfied with both devices.

H4 The user will be satisfied with the additional functions provided by the tablet.

4.3.1 Functionalities and tasks

14 basic functions compose the building blocks for all the functionality of the software. Each of the functions corresponds to a primitive interface action. We classified the functions depending on whether the wand could be used to perform the task (testing hypothesis 1-3) or not (testing hypothesis 4). The functions were used to define higher level tasks, which were evaluated empirically to test for accuracy and speed.

T1 First the user has to scale the visualization to the side and select a defined huge area on the front wall. After the selection he has to activate the default view.

T2 Next the user has to move the visualization on the front wall and select a small area on the side wall. The default view has to be activated afterwards.

T3 In the last selection task the user has to deactivate the 3D view and select a square area on the floor. After the selection return to the default view.

T4 In the final task the user has to scale the visualization. The small cube has to be scaled to it's maximum size. Then, the cube has to be grabbed and placed on the top right corner of the front wall.

Figure 3: Wand vs tracked Tablet. For the users view, the pointer is aligned with the left side of the tablet

Figure 2 shows the accuracy and speed for the first three tasks, and the speed for the last task. Mean, standard deviation and outliers are displayed. Interested readers can find detailed statistical information in the additional material. The distribution of responses was examined using the Quantile-Quantile-plot [Wil68], the Shapiro-Wilk-Test [Sha65] and a histogram. All distributions were found to be non-normal and independent, so the Wilcoxon-Mann-Whitney-Test (WMWT) [Man47] was used to measure the significance, with values under 5% considered statistically significant.

The WMWT showed statistical significance for the time difference in tasks 1-3 (with values 6e-4, 6e-6 and 6e-4, respectively), while there was no statistical significance in the difference in accuracy in tasks 1-3 or in the time for task 4 (values were 12%, 39%, 12%, 70%). From this data, we must consider that the new selection method is not useful, and the old one should probably be used. Fortunately, the tablet can use the old method when a tracker is attached to it (figure 3).

4.3.2 Post-questionnaire

This section deals with the subjective evaluation of the differences between wand and tablet by the users. When evaluating the intuitiveness, satisfaction, comprehension and speed of using the tablet interface to do control and selection tasks which could be done previously with the wand, a preference for the tablet appears. The wand is preferred for moving the 3D object and for its smaller weight.

With respect to the evaluation of the tablet interface to do things which could not be done using only the wand, which included load, search, setting the main genome, reordering, visualization selection, etc: except for the

changing of the order of the genomes, all the tasks were found easy to perform.

Users mostly preferred using the tablet, even though they were slower while performing the tasks. For typical information visualization applications such as this one, where we are interested in obtaining insights about the data, user satisfaction is more important than speed, as content users are more likely to spend the needed time within the system to obtain valuable discoveries.

4.3.3 Hypothesis evaluation

In accordance with section 4.3.1, H1 (wand selection faster than new tablet interface) is true. H2 (tablet selection more precise) could not be proven, as the tests show no statistical difference between the two. Since the tablet also allows the use of the wand selection method, by the use of an attached tracking device, it may be possible to obtain high speed selection using the alternative interface. As we saw in the previous section, the questionnaire shows that H3 and H4 are true.

The tablet can thus be used as an alternative to traditional wand interfaces, allowing new tasks to be performed easily and to the user's satisfaction.

5 CONCLUSIONS

We have used an android tablet to enhance the interaction possibilities of an immersive VR installation in the field of information visualization. The screen allowed us to provide an easier to use menu interface, and to increase the online help available during the use of the software. The touch interface in the tablet is higher precision than commonly used wand interfaces. There was no statistically significant difference between the accuracy of tasks when using tablet versus wand, but the tablet allowed symbolic input and enhanced system control.

It is worth noting, though, that the tablet is no panacea: the biggest remark is often that tablets are heavier and more fragile than joysticks or wands. We expect that in the future, these limitations will be reduced. Even now, the use of a mobile phone can alleviate the weight considerations, and phone covers can also be useful in the case of an accident.

6 ACKNOWLEDGMENTS

This work has been supported by the research project MrSymBioMath (European Union program FP7-PEOPLE-2012-IAPP, grant number: 324554).

7 REFERENCES

[Bed09] Bednarz, T. P., Caris, C., and Dranga, O. Human-computer interaction experiments in an immersive virtual reality environment for e-learning applications. 20th Australasian Association for Engineering Education Conference, Pages 834–839, 2009.

[Gar16] García-Hernández, R. J., Anthes, C., Wiedemann, M., and Kranzlmüller, D. Perspectives for using Virtual Reality to extend visual data mining in Information Visualization. IEEE Aerospace Conference, 2016.

[Har02] Hartling, P. L., Bierbaum, A. D., and Cruz-Neira, C. Tweek: Merging 2D and 3D interaction in immersive environments. 6th World Multiconference on Systemics, Cybernetics and Informatics, Pages 57–61, 2002.

[Man47] Mann, H. B., and Whitney, D. R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. Annals of Mathematical Statistics 18 (1), Pages 50–60, 1947.

[Med13] D. Medeiros, D., Carvalho, F., Teixeira, L., Braz, P., Raposo, A., and Santos, I. Proposal and evaluation of a tablet-based tool for 3D virtual environments. SBC Journal on 3D Interactive Systems, volume 4, number 2, Pages 30–41, 2013.

[Mor14] Morehead, M., Jones, Q., Blatt, J., Holcomb, P., Schultz, J., DeFanti, T., Ellisman, M., Doretto, G., and Spirou, G.: Poster: BrainTrek-An immersive environment for investigating neuronal tissue. 3D User Interfaces (3DUI), 2014 IEEE Symposium, Pages 157–158, 2014.

[Nis04] Nishimura, K., Abe, K., Tsutsumi, S., Aburatani, H., Hirota, K., Hirose, M. Virtual environment design guidelines for gene expression analysis: the utility of a lab bench metaphor and a road metaphor. Proceedings of IEEE Virtual Reality, Pages 247–248, 2004.

[Rao14] Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., Lieberman Aiden, E. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell 159, 2014.

[Sha65] Shapiro, S. S., and Wilk, M. B. Analysis of variance test for normality (complete samples). Biometrika 52: 591–611, 1965.

[Tuk14] Tukora, B., Anthes, C., Heinzlreiter, P., and Kranzlmüller, D. Large-scale dynamic visualization of multiple comparative genomic data. IEEE VIS Poster Program, 2014

[Wat99] Watsen, K., Darken, P. R., and Capps M. V. A handheld computer as an interaction device to a virtual environment. Proceedings of the Third International Immersive Projection Technology Workshop, 1999.

[Wil68] Wilk, M.B., and Gnanadesikan, R. Probability plotting methods for the analysis of data. Biometrika (Biometrika Trust) 55 (1), Pages 1–17, 1968.

Pardubice Region Health Data Visualization – Steps of Geoinformatics Project

Ing. Ivana Čermáková
1st author's affiliation
University of Pardubice
Studentská 95
530 09, Pardubice, Czech Republic
st32689@student.upce.cz

Mgr. Pavel Sedlák, Ph.D.
2nd author's affiliation
University of Pardubice
Studentská 95
530 09, Pardubice, Czech Republic
pavel.sedlak@upce.cz

doc. Ing. Jitka Komárková, Ph.D.
3rd author's affiliation
University of Pardubice
Studentská 95
530 09, Pardubice, Czech Republic
jitka.komarkova@upce.cz

ABSTRACT

Increasing importance about health and healthcare leads to utilization of unusual methods to present information to residents. Thematic maps, particularly set of thematic maps, which are readable and understandable is one of possibilities how to inform residents. Creating a set of thematic maps is a whole project of data visualization, because it is necessary to cover all facts and to be able to clearly present facts that are important for readers. The project of data visualization consists of the analytic phase, design phase, testing phase and the final implementation phase. Final set of maps contains 19 thematic maps and using 4 cartographic methods, e.g. directional distribution. The article describes case study: the whole project of creation of the thematic maps of Pardubice Region of the Czech Republic regarding health and healthcare. The article is focused on specific tasks solved within this project.

Keywords

Data visualization, health data, geoinformatics project, project of visualization

1. INTRODUCTION

Health and healthcare belong to the important topics about which public want to be inform. Many institutions are searching for suitable ways how to provide data in a format that will be understandable for citizens. Importance of thematic maps increases in the case of this problematic. One of disadvantages for readers is that maps are usually available for the whole world or country, but they are not available at regional level.

The main aim of the paper is to find a suitable way how to provide information about important topics regarding health and healthcare at regional level, namely on region Pardubice in the Czech Republic. Various cartographic methods are used to present information and create the final set of maps.

The paper describes a case study: creation of a set of thematic maps describing health and healthcare for the Pardubice Region is a complete geoinformatics project. It covers the whole process starting with users and aim definition and available data collection. It ends with published set of thematic maps. The project of creation of the set of maps is divided into 4 phases: analysis, design, testing and implementation. Furthermore, maps can be improved according to obtained results. The final set of thematic maps is created in the last phase (implementation).

2. PHASES OF THE PROJECT

The whole project is divided into 4 phases. Content of particular phases is based on previous geoinformatics projects ([Dob00a] and [Haj00a]), but each phase is modified according to the needs of this project, see Figure 1.

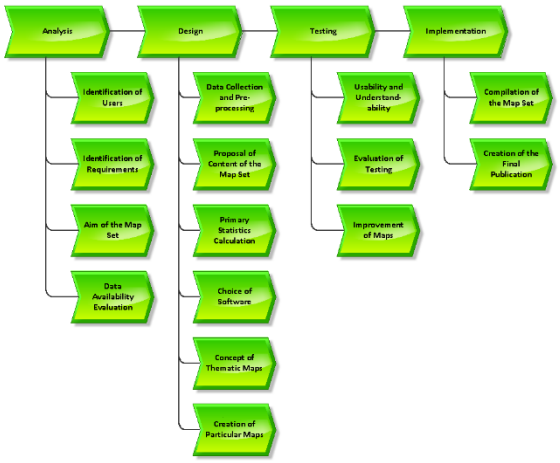


Figure 1: Phases of the Project, source: authors

Phase 1: Analysis

The analytical phase consists of definition of target group of users and their requirements and expectations, definition of aim of the set of maps and evaluation of availability of health data.

The aim of the set of maps is to describe situation of healthcare in the Pardubice Region by thematic maps to citizens. Different people (i.e. readers of the maps) have different level of knowledge so they differently understand these problematics. This fact leads to the necessity of maps readable and understandable by everyone.

Topicality of maps requires adequate and up-to-date datasets. Available and accessible datasets must be identified and evaluated. In the case the following data sources are used: ÚZIS, Regional Public Health Authority of Pardubice Region and NADĚJE. ÚZIS is the key institute providing health data. The institute is an organisational component of the Czech Republic, founded by Ministry of Health. [Uzi00a] Regional Public Health Authority of Pardubice Region is institute which provides health data only for its region of interest. NADĚJE is a non-profit non-governmental organization that provides religious, social, medical, counselling, educational, enlightenment and pedagogical services etc. [Nad00a].

Phase 2: Design

At first, data are collected from the above listed sources. They are evaluated from the point of view of their type, e.g. ordinal, cardinal, etc. This step is necessary for choice of proper statistical, analytical and cartographic methods in next steps.

Next, detailed list of included topics must be proposed according to the aim of the set of maps and according to the available data. In our case, there are chosen 19 topics which are divided into 3 basic parts. First part is focused on general characteristics and it contains 3 topics. The main part is focused on the state of health and healthcare (12 topics desrcribed by 16 maps). The last part of the set of maps is focused on health from other points of view. This part covers 4 particular topics in total: revenues and costs of hospitals, employees in health care and Public Health Control System.

Data pre-processing is the next step. Collected data must be verified and non-complete and non-reliable data must be discarded. Data from primary sources are in different formats so data must be transformed into format suitable for the next steps.

Statistical evaluation is the next step. All the data has one type – ordinal type. It means that features are measures by quantity [Rob00a]. Basic summary statistics, e.g. average, minimum, maximum and other measures, is calculated. These measures are necessary for choosing a suitable method for visualization and dividing data into the intervals.

ArcGIS for Desktop version 10.2.2 is used to create maps because it is one of mostly used software for data analyses and consequent maps making. CorelDRAW (version X5) is used for design of the resulting set of maps.

Concept of thematic maps is following: maps contain topographical background and thematic content. Topographic dataset ArcČR 500 is used as a topographic background. The dataset contains streets, rivers, lakes, forests or administrative boundaries [Top00a].

Design and creation of a template of maps follows. Template must be suitable for the problematic and must be attractive and interesting for map users.

Main advantage of creating thematic maps is that the thematic maps are published as a whole set. A comprehensive set of suitable symbols is proposed for the set. The symbols are unique and increase importance and understandability of maps because they are proposed in an associative way.

For the whole making the set of the maps it is necessary to describe the general characteristics of health and healthcare in the Pardubice Region. It means create maps of population, age index and health centres. Of course, mostly there are thematic maps use traditional cartography methods, for example diagram map and choropleth maps. Of course there are some special maps use more advanced methods, e.g. dots method or spatial statistics (every type of visualization is explained in the following parts).

2.1.1 Directional distribution method

This method is suitable in cases when concentration of features is higher in particular directions [Arc00a]. This method is used for one map only. This map compares births and deaths. Figure 2 shows difference between ellipse of deaths and births. The map clearly shows that both easternmost regions (Moravskoslezský a Jihomoravský) have higher number of deaths than births. It is shown by prolongation of the ellipse to the east.

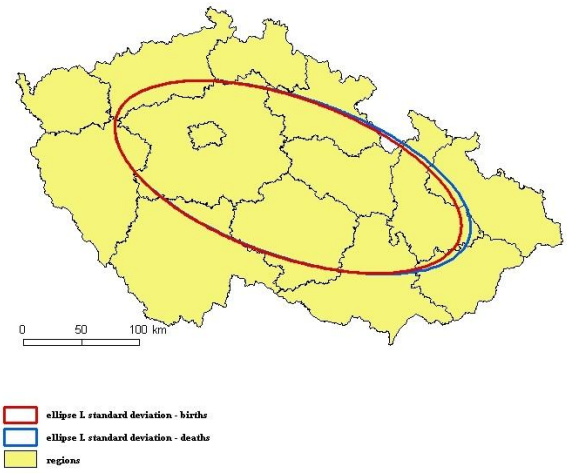


Figure 2: Comparison of deaths and births in 2012, source: [Cer00a]

2.1.2 Dots method

Dots method of visualization allows highlighting number of features better than single features. [Rob00a] For example, Figure 3 shows occurrence of tuberculosis in Pardubice Region by dots method (1 dot =1 occurrence).

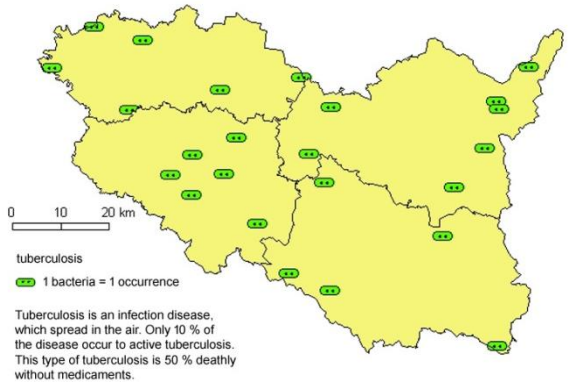


Figure 3: Tuberculosis in the Pardubice Region in 2012, source: [Cer00a]

Dots method is used for map of incidence of chickenpox also. Red dot is used for representation of number of people with chickenpox. Red colour is used because it is visible character of this disease. One red dot represents 90 patients. Icon of green bacteria is used in the case of map of incidence of tuberculosis instead of dots. Green bacteria symbol is a typical symbol for this disease. Dots method is used in the map describing state medical and map of epidemiological investigations.

2.1.3 Diagram map method

This method is suitable for some maps and some topics only. Diagram map method is used for example for thematic map regarding hospitalized persons. In this case, diagram uses enlarging icon of white bed with a man on an orange background and size of the symbol represents number of hospitalized persons. Different sizes represent different numeric intervals. Choice of a suitable number of intervals is important to adequately represents differences between regions on one side and to keep readability on the other side. Statistical analysis helps to set suitable ranges of intervals. Three intervals are chosen for each thematic map describing situation at the national level. Another problem is how data divide into intervals. It is individual for every map. Natural breaks (jenks) are used for most of maps. More about each map and method used for dividing data into intervals is available in [Cer00a].

Diagram maps method is used in the case of 12 maps. One is described above.

2.1.4 Choropleth mapping method

Maps are used for phenomena with spatial variation which can be divided by boundaries to the intervals. [Ter00a] Map of women in working age with hormonal contraceptives is divided to two ranges: < 50 % and > 50 %. Saturated colour (shades of pink in the case of this map) has values with higher intensity. Map regarding hospitalized persons at the regional level used for each particular value one colour (as a whole 4 colours). Map regarding births, described in previous section, uses choropleth mapping method too. Map at the national level uses 3 intervals (yellow, blue, pink) and map at the regional level uses only 2 intervals (blue and pink) – one colour for 2 values. Map is available like Figure 4.

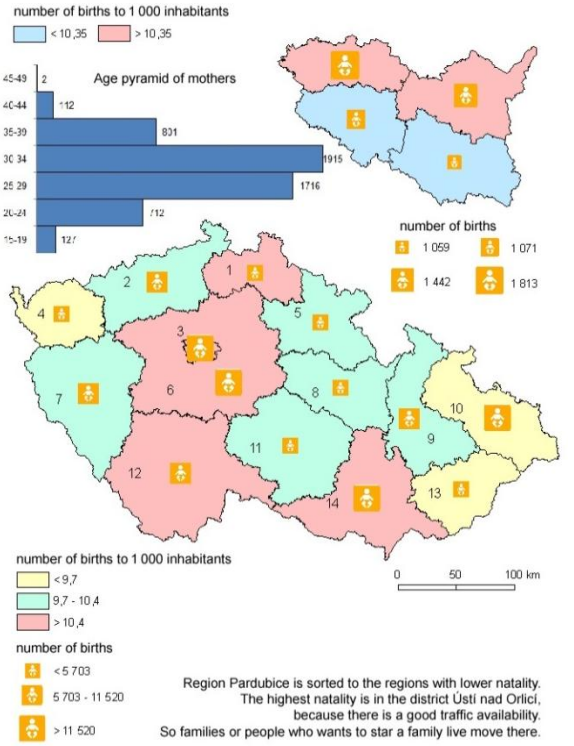


Figure 4: Births in the Czech Republic in 2012, source: [Cer00a]

Map about abortion at the national level uses 3 intervals symbolized by shades of blue. The last map describes employees working in health care. There are used 4 colours in this map – one colour is used for each particular value, i.e. each district. Map is designed in shades of green.

Phase 3: Testing

After creating all maps, it is necessary to find out if the maps are usable and if the readers understand content of the maps.

Several people, representatives of both public and specialists, are chosen for testing thematic maps. The main objective of testing phase is brief user testing of usability and understandability of thematic maps.

Representatives of both groups tested understandability of thematic maps and design of the maps also. All comments and identified problems were evaluated and used to improve maps.

Phase 4: Implementation

Compilation of the final set of thematic maps is the next step. As it was described previously, the set is divided into 3 parts and it contains graph (age pyramid) and 4 tables with important information about monitored area as well. Additional text must be included to describe main aim of the publication, content of the set of maps and target group of readers. The whole set of maps is available in [Cer00a].

3. CONCLUSION

The main aim of the paper is to identify suitable ways how to visualize health and healthcare data at the regional level in a comprehensive set of maps. Proposed methods are applied to the Pardubice Region in a form of a case study.

Creation of the comprehensive set of maps is a whole project. It is necessary to identify target group of users, their requirements, to collect data, to appropriately process and visualize them. The project is divided to the following parts: analytic phase, design phase, testing phase and final implementation phase.

The comprehensive set of thematic maps contains 19 thematic maps, 4 tables and 1 graph. It represents the main result of the case study. Thematic maps are used as a suitable way for visualization of particular topics and their distribution over the area of interest (Pardubice Region). Four cartographic methods for visualization are used to create thematic maps: directional distribution, dots method, diagram map and choropleth mapping. A new set of associative map symbols is proposed for the final set of maps to present health and healthcare data in an understandable way.

The main contribution of the paper is show a possibility how visualize the health data on the regional level, because no many researches are focused on. And it is very important topic in these days.

The output set of maps can be extended in the future by using other methods of visualization, by including new important topics or by combination of both possibilities. Creating other sets of thematic maps on the regional level can compare health and healthcare from various viewpoints also.

4. ACKNOWLEDGMENTS

The paper has been completed with the kind support of SGS project of Faculty of economics and Administration, University of Pardubice.

5. REFERENCES

[Arc00a] ArcREVUE: 2/2014 [online]. Praha: ARCDATA Praha s.r.o, 2014. ISSN 1211-2135. Available: http://www.arcdata.cz/Publikace/casopis-arcrevue/ Access Date: 9 September, 2015

[Cer00a] Čermáková, I. Visualization of selected health data from Pardubice region. Pardubice, 2015. Diploma thesis. Univerzita Pardubice. Available in Czech only, title: Vizualizace vybraných dat ze zdravotnictví Pardubického kraje

[Dob00a] Dobesova, Z. Geographic information systems for botanical garden - Steps of design and realization in Conf. proc. SGEM, pp.377-384. 2012

[Haj00a] Hajkova, L. et. al. GIS and cartography for bioclimatological research: Atlas of phenological conditions of Czechia in Conf.proc. SGEM, pp.1065-1072, 2012

[Nad00a] Naděje. [online]. 2014. Available: http://www.nadeje.cz/vyrocni_zpravy Access Date: 19 September, 2014

[Rob00a] Robinson, Arthur Howard. Elements of cartography. 6th ed. New York: John Wiley & Sons, 1995, 674 s., [8] s. obr. příl. ISBN 0471555797.

[Ter00a] Terry, A. Slocum. Thematic cartography and geovisualization. 3th ed. United States of America: Pearson Prentice Hall, 2010, 561 s., ISBN 0-13-801006-4

[Top00a] Topografická mapa. Atlas map České republiky - GEOČR500 [online]. 1998. Available: http://www.geology.cz/demo/cd_geocr500/stranky/p_topograficka.html Access Date: 30 November, 2014

[Uzi00a] Ústav zdravotnických informací a statistiky ČR [online]. 2014. Available: http://www.uzis.cz/nas Access Date: 19 September, 2014

Multi-modal Controller for Image Manipulation in the Operating Room

Alexandre Sierro Pierre-André Mudry
University of Applied Sciences Western Switzerland
HES-SO Valais
Rte du Rawyl 47
1950 Sion, CH - Switzerland
{pierre-andre.mudry, alexandre.sierro}@hevs.ch

Abstract

In the domain of orthopedics, surgeons often rely on radiology images during operations. In this context, manipulating the images displayed on a computer screen is an issue as their hands have to remain sterile. In this article, we present a multi-modal controller (foot and voice) coupled with an existing state-of-the-art radiology display and analysis software used in operating rooms. The controller itself consists of a battery-operated wireless embedded system integrated into a shoe that provides multiple foot pressure-points as well as an absolute orientation sensors. In addition, a wireless microphone is used to acquire voice commands.
To demonstrate the validity of our approach, we present a randomized user study conducted on ten subjects that had to perform image manipulation tasks using the controller.

Keywords

Foot-based controller, inertial measurement unit, voice control, orthopedics, surgery, radiology imagery.

1 INTRODUCTION

During orthopedic operations, surgeons often rely on existing radiology images (X-ray, MRI, . . .). Displayed on computer screens, those images are often manipulated with a mouse by operating room (OR) assistants as the surgeon's hand are often busy manipulating the patient. Another reason behind those assisted manipulations are sterility issues related to hand-based controllers (keyboards or mice for instance).

In this paper, we present a multi-modal controller based on voice and foot input for radiology image manipulation during surgery. The advantages of this approach are two-fold: first, the advantage of sterility and hand-free operation and, second, the independence in the positioning of the surgeon towards the input device.

1.1 Paper Organization

We proceed as follows: in the next section, a brief overview of existing human-machine interaction methods in the operating room is presented. After that, focus is put on the hardware and software implementation of the controller and the means of interfacing it with a standard, PC-based, radiology image manipulation program. We then present a randomized experimental setup to demonstrate the strengths and weaknesses of the applied approach before concluding.

2 RELATED WORK

High-sterility and non-encumbered interaction are paramount in the OR. For this reason, camera-based approaches tracking surgeon gesture have been successfully applied in the past ([1, 2]). However, one major difficulty with this technique is the proper detection of gestures which still remains a challenge today ([3]). To improve the situation, researchers have demonstrated that integrating the third dimension can be useful (for instance by using a *Kinect* device [4], or *Leap Motion device* [5]). However, using an imaging device requires the surgeon to be positioned at a precise location in the OR.

To partially circumvent this limitation, voice commands can be added to the setup in order to perform some control (see for instance [6, 7]) when not in the field-of-view of the imaging device.

In the last decade, developments in the field of micro electronic mechanical systems (MEMS) enabled the production of cheap and reliable orientation sensors. Of particular interest is the appearance of devices integrating a fusion of accelerometers, gyroscope and geomagnetic sensors which can be used to extract hand or foot move-

ments of a user. This enabled the creation of position-capturing devices which can be used in gaming or control (for instance as described in [8, 9]) and that we will be using in our multi-modal controller to capture foot-orientation information.

3 IMPLEMENTATION

Our multi-modal controller is based on three different sources of information: foot pressure points distribution, foot gestures as well as voice commands. Foot-based information is captured via a dedicated embedded system which has been integrated into the sole of a shoe.

3.1 Architecture

As depicted in Fig. 1, foot sensor information is wirelessly transmitted to a control software which integrates this information thanks to a very efficent open-source voice recognition package called *Sphinx*[1].

Once the proper pointing method (see 3.3) has been selected, the appropriate commands are then generated and sent via telnet to *Weasis DICOM Viewer*[2], an open-source radiology image manipulation program..

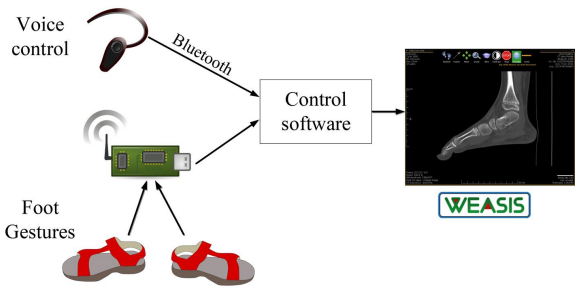


Figure 1: System architecture

3.2 Hardware Implementation of the Foot-based Controller

The foot-based controller embedded system (Fig. 2) contains four main components articulated around an ARM Cortex M0 micro-controller:

- **Pressure sensors** – Foot pressure is measured at three different locations using resistive load cells from *Alpha Electronics*. The resultant resistance is converted into a voltage and then digitized using the micro-controller's analog to digital converter.
- An **inertial measurement unit** (IMU) – The exact model is `BNO055`, which is a module already containing the required sensor fusion algorithms to provide fast and accurate readings of absolute orientations extracted from 9 axes : 3 axes accelerometer, 3 axes gyroscope and 3 axes magnetometer.
- A low-power **wireless communication** chip – Active in the 2.4 GHz range, the `NRF24L01+` chip is connected with an antenna directly printed on the printed-circuit board.
- An autonomous **power supply** which consists in a 3.7 V, 850 mAh Lithium-Polymer battery charged using either with a micro-USB connector or an inductive charger.

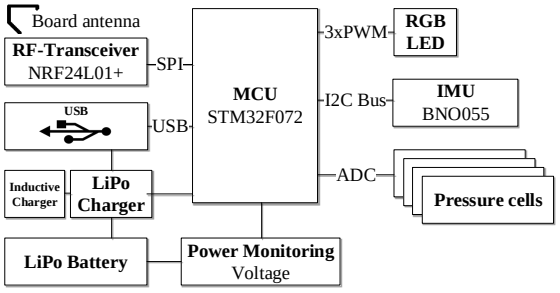


Figure 2: Hardware architecture

The embedded system is integrated into a standard sandal, which can be seen in Fig. 3. The system can be used using one or two shoes, depending on the selected interaction mode, as we will discuss in the next section.

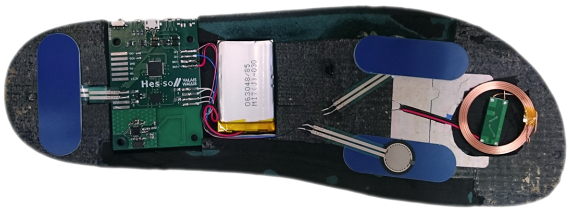


Figure 3: Shoe integration (from left to right) : rear pressure sensor, system board, left and right pressure sensors and wireless charging receiver.

3.3 Software implementation

The visualization software we used is tailored to be used with a standard mouse input. Early tests showed that a direct translation from foot gestures to mouse commands is not feasible. In fact, a clicking gesture with the foot can be very tiresome and therefore a different selection mechanism based on voice commands was chosen.

The valid actions implemented in the context of this project are : *move*, *zoom*, *contrast* and *slice*. For this last point, it is worth noting that radiology data might be three-dimensional and therefore it is possible to navigate into the "depth"of the radiology image by changing the actual slice of the data.

To select between those different actions and interact with the software, three interaction strategies have been implemented:

[1] http://cmusphinx.sourceforge.net/
[2] https://github.com/nroduit/Weasis

1. **Voice** method, which lets the user choose between the different actions using voice commands. In this mode, inclination of the main foot, measured by IMU, only acts on the selected action.

2. **Fusion** method, which combines pressure and inclination of the main foot to act simultaneously on movement and magnifying. In this mode, pressure applied on the tip of the foot will zoom-in and pressure on the back will zoom out.

3. **Two feet** method, which uses the main foot inclination to move the picture and the second foot to control magnifying.

In every strategy, voice commands can be used to cancel the current action or reset the visualization to a known state. As depicted in Fig. 4, a control panel using an icon-based UI appears as an overlay in front of *Weasis*, displays valid commands and provides a feedback of the currently selected mode.

Figure 4: The method selector UI which is displayed atop *Weasis*.

3.3.1 Acquisition process

Raw data output from the sensors is converted to valid user inputs with a relatively simple software on the PC. The conversion algorithms starts by applying specific thresholds and gains to each sensor and then their values are routed to a specific *Weasis* command according to the currently selected pointing method.

In order to improve the user experience of the system and increase its controllability, a profile containing threshold values and main foot selection is generated for each end-user. This profile enables the integration of taste-specific values into the controller and act as calibration for the system.

4 EXPERIMENTAL SETUP

To demonstrate the validity of the approach and to provide a user-based feedback on the multi-modal controller, we implemented an experimental setup reproducing a typical OR scenario.

During the experiment, the three aforementioned pointing methods were evaluated as well as the standard mouse control which serves as a reference.

To test the setup, ten persons were presented with the experimental setup and the detection thresholds of the foot-controller were adjusted to their taste.

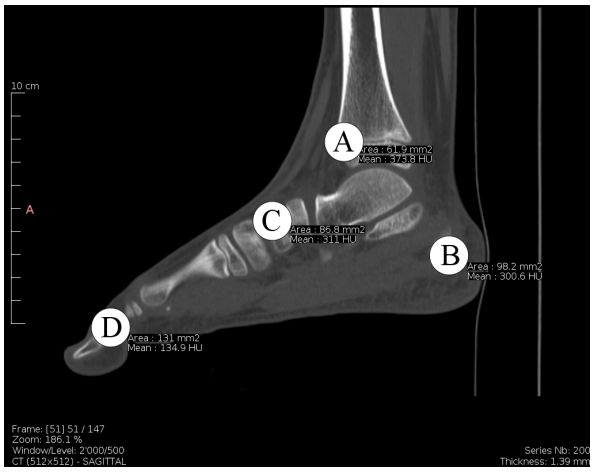


Figure 5: Points of interest that had to be zoomed to.

4.1 User Objectives

For the experiment, the users had to zoom on four points in an X-ray image following a specific order (highlighted as ABCD on Fig. 5).

Measurements are performed twice with different complexities: the first time the user has to zoom-in then zoom-out for every point whereas the second time only the first point has to be zoomed-in before moving over the other points. These complexity levels are labeled respectively *Z+M* and *M*.

In both cases, the time to reach the first point and the subsequent transitions times are measured. Before each measure, users had time for practice. At the end of experiment, users also had to rate their satisfaction level for each of the pointing method in terms of accuracy, speed and usability. The marks given could vary between 1 (not satisfied) and 4 (very satisfied).

4.2 Results

Fig. 6 depicts the pointing duration for the various strategies and task complexities.

When considering interaction speed, a first result that can be extracted is that the reference mouse method is on average 2.5 times faster than any other method. The voice method is the slowest strategy for aggregated zoom and movement ; this can be explained by the fact that changing from one mode to the other requires voice commands. However, as zoom and move commands are clearly separated, this reduces interferences in the movement and allows more accurate movements for the M complexity.

Overall, the fusion method seems to be the most appropriate to achieve a reasonable speed for most users using this setup. Unfortunately, direct speed comparisons with other input techniques are difficult as use-case scenarios differ too much.

From a user evaluation standpoint, Fig. 7 shows how the various methods were evaluated in terms of usability,

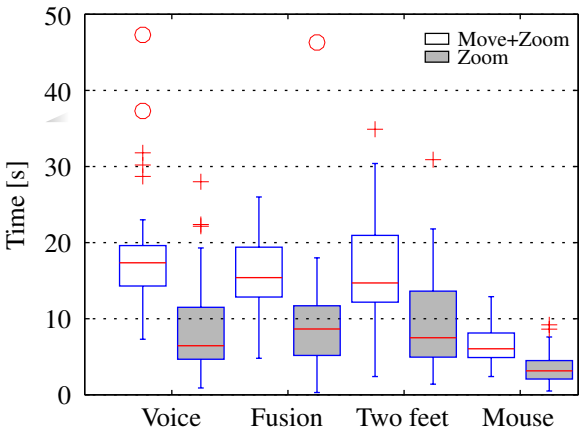


Figure 6: Duration analysis for the different methods and task complexities.

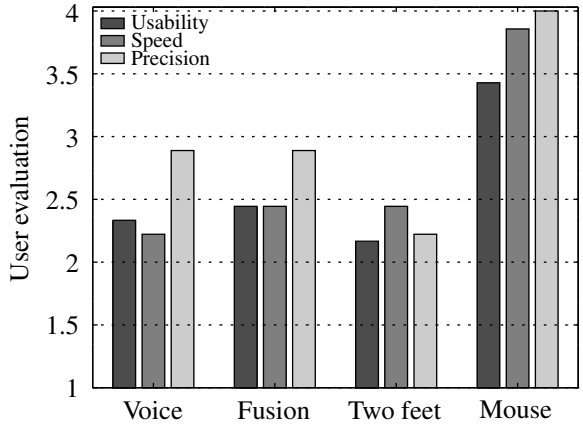


Figure 7: User evaluation (average) of the different methods.

speed and precision. Analyzing this data reveals that test impressions correspond to time measurements, i.e. voice method has a good precision but is slower, two feet is the fastest method but is less precise. For those experiments the reference mouse method is still preferred.

5 CONCLUSION

We showed in this article how a multi-modal controller can be successfully used to provide a robust HMI in the context of an OR. Even if users seem to favor a mouse as input device for image manipulation, we showed that mixing voice commands with foot gestures provides both accuracy and speed whilst preserving sterility and position independence for the surgeon.

Further work will include testing the multi-modal controller in a real OR scenario to adjust the system to real-world constraints and integrate feedback from surgeons.

6 ACKNOWLEDGEMENTS

This research, part of the *Lunamed* research projet, was funded by an UAS Western Switzerand (HES-SO) internal grant.

7 REFERENCES

[1] M. Ma, P. Fallavollita, S. Habert, S. Weidert, and N. Navab, "Device- and system-independent personal touchless user interface for operating rooms," *International Journal of Computer Assisted Radiology and Surgery*, pp. 1–9, 2016.

[2] G. C. S. Ruppert, L. O. Reis, P. H. J. Amorim, T. F. de Moraes, and J. V. L. da Silva, "Touchless gesture user interface for interactive image visualization in urological surgery," *World journal of urology*, vol. 30, no. 5, pp. 687–691, 2012.

[3] T. Kopinski and U. Handmann, "Touchless interaction for future mobile applications," in *International Conference on Computing, Networking and Communications (ICNC)*, pp. 1–6, February 2016.

[4] M. Strickland, J. Tremaine, G. Brigley, and C. Law, "Using a depth-sensing infrared camera system to access and manipulate medical imaging from within the sterile operating field," *Canadian journal of surgery. Journal canadien de chirurgie*, vol. 56, pp. E1–6, June 2013.

[5] A. Zocco, M. D. Zocco, A. Greco, S. Livatino, and L. T. De Paolis in *Proc. of the 2nd Int. Conference on Augmented and Virtual Reality (AVR2015)*, pp. 432–445, Springer, 2015.

[6] A. M. Hötker, M. B. Pitton, P. Mildenberger, and C. Düber, "Speech and motion control for interventional radiology: requirements and feasibility," *International Journal of Computer Assisted Radiology and Surgery*, vol. 8, no. 6, pp. 997–1002, 2013.

[7] Y. Kim, S. Leonard, A. Shademan, A. Krieger, and P. C. W. Kim, "Kinect technology for hand tracking control of surgical robots: technical and surgical skill comparison to current robotic masters," *Surgical Endoscopy*, vol. 28, no. 6, pp. 1993–2000, 2014.

[8] A. Gams and P.-A. Mudry, "Gaming controllers for research robots: controlling a humanoid robot using a WIIMOTE," in *Proc. of the 17th Int. Electrotechnical and Computer Science Conference (ERK08)*, pp. 191–194, 2008.

[9] K.-B. Cho and B.-H. Lee, "Intelligent lead: a novel HRI sensor for guide robots," *Sensors*, vol. 12, no. 6, p. 8301, 2012.

ISSN 2464-4617 (print) WSCG 2016 - 24th Conference on Computer Graphics, Visualization and Computer Vision 2016
ISSN 2464-4625 (CD-ROM)

Min-Max Mipmaps for Efficient 2D Occlusion Culling

Simon Scheckel[1] Andreas Kolb[2]
Computer Graphics Group, University of Siegen, Hölderlinstr. 3, 57074 Siegen, Germany
[1]simon@rollaprint.de; [2]andreas.kolb@uni-siegen.de

ABSTRACT

3D culling techniques are well established to improve rendering performance, but cannot be applied to 2D games in which the scene is composed of partially transparent textures in a known layer arrangement. Commonly, 2D rendering is achieved in a simple back-to-front blending scheme. This paper discusses options to realize 2D occlusion culling techniques using standard OpenGL functionalities, and introduces a novel 2D culling technique based on min-max mipmaps. We evaluate the performance of the different techniques for different scenarios.

Keywords

2D Picture/Image Generation; Textures and Framebuffer Operations; Mipmaps

1 INTRODUCTION & PRIOR WORK

In 2D graphics applications, such as games, the scene consists of a set of *layers* (2D textures with transparency) in a defined "depth" order. The resulting 2D rendering is commonly achieved in applying a simple back-to-front blending scheme.

While the primary target for both, 2D and 3D games is to optimize for high-quality graphics, subject to performance considerations and/or given hardware restriction, the major structural difference lies in the nearly vanishing geometry processing requirements for, and the known "spatial" structure of, 2D applications. Furthermore, 2D games rather target mobile platforms which impose greater hardware restrictions.

3D Culling Techniques: In 3D graphics applications, culling methods are commonly used in order to discard geometry on object level [3]. Culling techniques are designed to deal with the *a-priori unknown depth structure* of the 3D scene to be rendered which consists mainly of *fully opaque objects*.

Early culling techniques usually attempted to exploit specific scene or object structures, such as portal culling for buildings [13], or the aspect graph related to the projetive structure of polygonal models [12]; see Cohen et al. [3] for a in-depth discussion of early culling techniques. Alternative, image based approaches such as a hierarchical z-buffer approach by Green et al. [7] and the hierarchical occlusion maps by Zhang et al. [15] are applicable to generic scenes. Both approaches are related to our work, as they use hierarchical data structures in order to detect objects which are (potentially) hidden by occluders; see Sec. 4.

Hardware occlusion queries offer an efficient utility to test the visibility of bounding geometries against the depth buffer before rendering the contained object. Bittner et al. [1] realize an efficient coherent culling method, which uses an octree and temporal coherence in order to allow for an asynchronous handling of occlusion queries and rendering tasks.

Overdraw Reduction: The performance of 2D applications is steered by the GPU's fill-rate capability and the amount of overdraw present in the current scene. Thus, 2D culling is equivalent to *reduce overdraw operations*. In 3D rendering engines, deferred shading and lighting is used to reduce unnecessary shading operations by applying a fast z-prepass [10]. This technique can be applied to 2D image synthesis as well, however, in the prepass only rendering of opaque areas can be handled, while transparent regions need to be handled differently; see Sec. 2.

An alternative 2D approach is to identify opaque and not fully transparent texture areas and convert them into meshes, thus preventing overdraw for fully transparent regions. Church follows this idea by cutting off transparent regions using "corner cutting" and triangulating the remaining areas [2]. However, this technique cannot handle multiple opaque regions and holes, which frequently occur in 2D games.

Contribution: As the direct application of 3D occlusion culling methods to 2D rendering is of limited use, this paper discusses automated approaches to improve performance for 2D rendering by overdraw reduction. Beside alpha-testing (Sec. 2) and contour tracing with triangulation (Sec. 3), which do not lead to siginifcant performance gains, we introduce a novel 2D occlusion culling technique based on min-max mipmaps. The main advantages of our approach are

- generation of isosceles triangles covering fully opaque or fully transparent regions, resulting in superior geometry and rasterization performance, and
- adaptivity of the triangle refinement level, leading to a performance-optimal ratio between triangle size and triangle count.

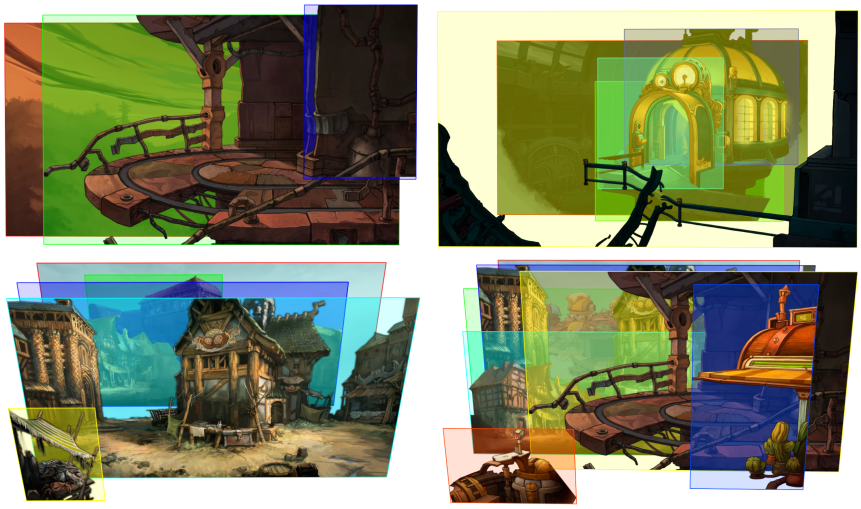


Figure 1: Sample scenes used for evaluation. The scene parameters are depicted in Tab. 1.

2 ALPHA-TEST-BASED CULLING

Alpha testing can naturally be used to identify the foremost, opaque layer at pixel level. Utilizing the alpha test requires a two-pass process. Given a layer sequence `layer[i]` sorted in depth order, in the first pass opaque pixels are drawn in the depth buffer in a front-to-back manner. Due to the front-to-back processing, areas behind near opaque regions are already discarded in the early depth test. In the second back-to-front pass, layer pixels hidden behind opaque pixels are again not drawn. However, in the worst case all layer pixels are drawn twice. Furthermore, transparent regions are not taken care of. An extension of this algorithm involves drawing already in the first pass (later referred as alpha2), but it still has to draw all transparent parts.

The main disadvantage in the alpha test algorithm is that all, i.e. opaque and transparent, regions are rasterized in each pass, even though in the first pass we are dealing with opaque regions only and, in the second pass, only foremost opaque pixel and transparent regions are of interest.

3 CONTOUR TRACING AND TRIANGULATION

A natural approach is to represent fully opaque *inner regions*, as well as the *outer regions* (fully transparent regions) in a layer texture as (triangular) meshes. Hereby, both regions need to fulfill a *consistency criteria*: the inner region may not contain pixels that are not opaque and the outer region must cover all transparent pixels.

We generate these kinds of triangulations on the CPU in a prepocessing step as follows:

Noise Reduction: By applying opening and closing to remove small features without violation of our consistency criteria.

Figure 2: Sample triangulations generated with the Mip8, Mip2 and the contour-based mathods. Images from [4].

Contour Tracing: Tracing the boundary of the opening inner regions using a scanline-based, Moore-Neighbor tracing.

Reduction of Contour Points: In order to avoid a large set of triangles that need to be transfered to the GPU, we further reduce the set of contour points. However, we need to ensure, that this reduction does not violate the consistency criteria. Thus, we only can remove inner, collinear points.

Triangulation: The final triangulation is generated using a sweep-line algorithm and requires a list of closed contours in the image plane as input [6].

4 MAXIMUM MIPMAPS

As contour tracing & triangulation often leads to a large amount of long and thin triangles (see Fig. 2), we propose a novel minimum and maximum mipmaps (*min-max mipmaps*) technique in order to generate nearly ideal triangulations. Min-max mipmaps have already been used in computer graphics, e.g., for soft shadows [9], global illumination [11], collision detection using geometry images [8], and ray casting of terrain data [14].

Our approach computes min-max mipmaps as lower and upper bound for the alpha values in the area represented by each pixel in the mipmap. Thus, a mipmap pixel resembles a fully opaque or fully transparent region if its min and max mipmap values are equal to 1 or

0, respectively. We generate the min-max mipmap using a simple fragment shader, combining 4 pixels into one in each hierarchy generation step.

The triangulation algorithm classifies layer areas as opaque (inner triangulation) or transparent (outer triangulation) directly on the min-max mipmap hierarchy in a top-down manner. Starting from the highest (1x1 px) mipmap level, we recursively check for maximum and minimum pixel values and if the pixel cannot be classified as fully opaque or transparent we traverse to the next finer hierarchy level. If a (refined) mip-map pixel represents a fully opaque or fully transparent rectangular, the related layer region is tessellated using two isosceles triangles. The classification process for both maximum mipmaps and contour tracing is done **only once** for each layer with the outer and inner regions and results in a list of triangles that are used to draw.

The main advantages of the mipmap-based approach is, that we can tune the amount of generated triangles by changing the *maximum level of refinement*. Apparently, a low maximum refinement level yields large "undecided" regions which need to be handled on a per-pixel level in the compositing stage. However, these "undecided" regions most likely contain many fully transparent and fully opaque regions which could be assigned to the inner or outer region, respectively. On the other hand, a high maximum refinement level leads to a large number of small triangles, resulting in a computational effort in the geometry processing and rasterization stages. We discuss the influence of the maximum refinement level in Sec. 5.

5 RESULTS

In this section we compare the 2D occlusion techniques to the standard back-to-front blending for 2D image synthesis. Here 'Alpha2' denotes the changed version of the alpha test with drawing in the first pass and 'Mip*N*' stands for our min-max mipmap-based occlusion culling, where the smallest region generated by the refinement is a NxN pixel region in the final image. For the comparison, we use four scenarios from the 2D games [4, 5]. Tab. 1 depicts the relevant key features of these scenes. The scenes consist of layers with different sizes, which we classified in FullHD, medium (1400x600) and small (600x400). To get genuine results, shader operations consist only of a matrix multiplication for the vertex shader and a texture lookup for the fragment shader. All scenes consist of static layers, that can be preprocessed as is the case for most 2D games. The synthesized images are FullHD throughout.

5.1 Preprocessing

Since we assume static layer geometries, the performance of the preprocessing stages of the individual algorithms is of little practical impact. Still, we present the preprocessing timings in Tab. 2 for FullHD, medium and small texture layers, as a reference for the estimated complexity.

	Layers	Px Cnt	Min Px Cnt
Scene 1	3	6,240k	2,238k (35.8%)
Scene 2	5	8,041k	2,148k (26.7%)
Scene 3	5	8,121k	2,088k (25.7%)
Scene 4	7	11,389k	2,132k (18.7%)

Table 1: Parameters for test scenes: Number of layers, the total count of pixels without culling ('Px Cnt') and the theoretical minimal count of drawn pixels ('Min Px Cnt').

5.2 Performance

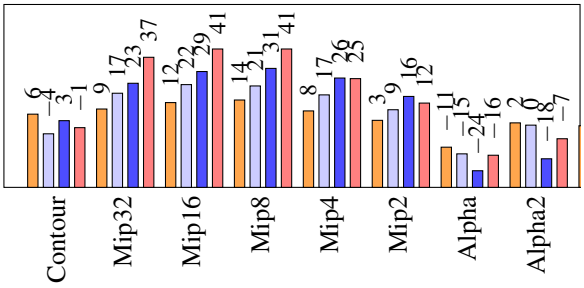


Figure 3: Relative performance gain compared to standard back-to-front blending on a GTX 770 with Nvidia Nsight [%].

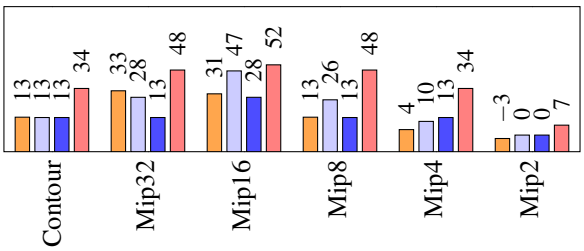


Figure 4: Relative performance gain compared to standard back-to-front blending on an iPad 3 [%]

Figs. 3 and 4 show the relative performance gain of each method compared to the standard back-to-front compositing scheme. Due to the fine granular culling on

Algorithm	1920x1080	1400x600	600x400
Loading layers	5	6	1
Contour tracing	127	126	33
Mip16	38	31	20
Mip8	37	33	22
Mip4	46	37	25
Mip2	49	33	24

Table 2: Preprocessing time [ms] for loading the layers and triangulation (including mipmap generation for Mip$N, N \in \{2,4,8,16\}$).

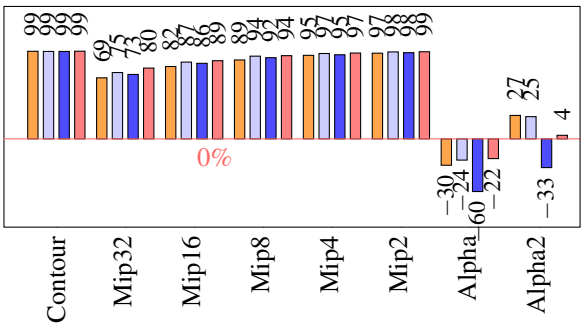


Figure 5: Culling efficiency: The relative amount of culled pixel, where 100% equals to the minimum amount of pixel draw, 0% to the standard back-to-front blending; see Tab. 1.

Method	Scenes 1	2	3	4
Con	7,457	9,895	5,750	16,154
Mip32	3,238	2,764	1,918	5,412
Mip16	7,182	6,724	4,656	12,448
Mip8	15,220	15,080	10,524	28,130
Mip4	27,668	28,814	19,722	53,822
Mip2	51,784	55,760	37,768	100,936

Figure 6: Number of triangle generated by the contour tracing (Con) and the mipmap methods Mip$N$ for $N \in \{2,4,8,16,32\}$ for the scenes (Sc) 1–4.

pixel level and the additional draws required to generate the depth map, the alpha test based approach performs very poorly. The iPad doesn't support Alpha Test, so only triangulation based methods are tested here. Most interestingly, the contour based triangulation approach performs poorly on the PC and somewhat better on the iPad, whereas our mipmap based approach constantly performs best for $N = 8$ and $N = 16$ on the PC and for $N = 16$ and $N = 32$ on the iPad.

On the other hand, we investigate the *culling efficiency* (see Fig. 5), i.e., the relative amount of culled pixels compared to the maximum possible amount of culling (see also Tab. 1). Here, the contour-based approach delivers the most accurate coverage of the regions to be culled. The alpha-blending approaches naturally require many additional draws in order to generate the opaque depth map, thus their efficiency in terms of overdraw reduction is very poor. The mipmap approach's efficiency increases as the level for generating triangles gets smaller, i.e., for decreasing N.

The mipmap approach generates isosceles triangles which can be rasterized more efficiently, whereas the contour based approach generates elongated triangles (see Fig 2). Compared to the PC, the iPad has a bigger performance hit for higher triangle counts, but it shows similar results as the desktop GPU solution 4.

6 CONCLUSION & FUTURE WORK

We present a novel technique to automatically generate efficient, isosceles triangulations for inner and outer regions of partially transparent textures which leads to a significant overdraw reduction in 2D rendering. Our technique can be adopted to the geometry performance capacities of the targeted platform, leading to a performance-optimal ratio between triangle size and triangle count. Future work may include dynamic textures, which requires efficient online triangulation methods.

7 REFERENCES

[1] J. Bittner, M. Wimmer, H. Piringer, and W. Purgathofer. Coherent hierarchical culling: Hardware occlusion queries made useful. *CGF*, 23(3):615–624, 2004.

[2] A. Church. *Depth-Cull Optimization of 2D Scenes for 3D Graphics Hardware*. 2014.

[3] D. Cohen-Or, Y. L. Chrysanthou, C. T. Silva, and F. Durand. A survey of visibility for walkthrough applications. *IEEE TVCG*, 9(3):412–431, 2003.

[4] Daedalic Entertainment. *The Dark Eye: Chains of Satinav (Computer Game)*, 2012.

[5] Daedalic Entertainment. *Deponia (Computer Game)*, 2012.

[6] V. Domiter and B. Žalik. Sweep-line algorithm for constrained delaunay triangulation. *J. Geo. Inf. Sci.*, 22(4):449–462, 2008.

[7] N. Greene, M. Kass, and G. Miller. Hierarchical z-buffer visibility. In *Proc. SIGGRAPH*, pages 231–238, 1993.

[8] A. Greß, M. Guthe, and R. Klein. Gpu-based collision detection for deformable parameterized surfaces. *CGF*, 25(3):497–506, 2006.

[9] G. Guennebaud, L. Barthe, and M. Paulin. Real-time soft shadow mapping by backprojection. In *Rendering Techniques*, pages 227–234, 2006.

[10] T. Harada, J. McKee, and J. C. Yang. Forward+: A step toward film-style shading in real time. *GPU Pro 4: Adv. Rend. Techn.*, 4:115, 2013.

[11] G. Nichols, J. Shopf, and C. Wyman. Hierarchical image-space radiosity for interactive global illumination. *CGF*, 28(4):1141–1149, 2009.

[12] H. Plantinga and C. R. Dyer. Visibility, occlusion, and the aspect graph. *J. Computer Vision*, 5(2):137–160, 1990.

[13] S. J. Teller and C. H. Séquin. Visibility preprocessing for interactive walkthroughs. In *Proc. SIGGRAPH*, volume 25, pages 61–70, 1991.

[14] A. Tevs, I. Ihrke, and H.-P. Seidel. Maximum mipmaps for fast, accurate, and scalable dynamic height field rendering. In *I3D*, pages 183–190, 2008.

[15] H. Zhang, D. Manocha, T. Hudson, and K. E. Hoff III. Visibility culling using hierarchical occlusion maps. In *Proc. SIGGRAPH*, pages 77–88, 1997.

Free form surface modeling and analysis

Karel Dvořák
College of Polytechnics Jihlava
Tolstého 16
586 01, Jihlava, Czech Republic
karel.dvorak@vspj.cz

Michal Bílek
College of Polytechnics Jihlava
Tolstého 16
586 01, Jihlava, Czech Republic
michal.bilek@vspj.cz

Stanislava Dvořáková
College of Polytechnics Jihlava
Tolstého 16
586 01, Jihlava, Czech Republic
stanislava.dvorakova@vspj.cz

ABSTRACT

Creating 2D and 3D models in CAD and following creation of drawings are standard procedures for constructing in the current industrial practice. Most models of industrially produced machine parts can be made by basic processes of modeling using construction or hybrid methods. Listed procedures are based on input of a basic – primitive body and its shaping using predefined design elements, combined with the use of construction methods from the default 2D geometry and using methods of extraction, rotation, or dragging a profile along the curve. Requirements of industrial practice and current production technological possibilities often require the implementation part of a more general form, which cannot be created by mentioned methods. In advanced stages of 3D models it is possible to use the methods of free surface shaping, based on creating of parametric, semi parametric or non-parametric network of 3D curves representing the basis for generating surface that defines the boundary of solid body.

Keywords

Computer Aided Design, Curves, Shape Analysis, Rapid Prototyping, Surface, Virtual Prototype.

1. INTRODUCTION

Computer Aided Design is a standard of the current industry practices in the field of mechanical engineering, electrical engineering, civil engineering, as well as medicine and other relevant fields. 2D and 3D models of components and assemblies are used for the subsequent formation of the manufacturing technical documentation, preparation of production technology, and also for performing simulations and analyzes in these fields. Virtual environments of computer tools enables to perform a wide range of simulations and analysis, without the need for physical prototypes [Sam08]. 3D digital models represent a virtual prototype that contains a large amount of information about the component or assembly for the entire course of the product life cycle. The 3D and 2D drawings are currently still basic information medium for preparation of production and assembly technology [Shi10]. The digital model is also the basis for the generation of data for processing programs that control the machining process for numerically controlled machine tools [Zei10] [Lee99]. Advanced methods of 3D modeling in CAD, based on the free form surfaces modeling generate often components or preparations for the manufacturing of these parts that are difficult to manufacture by "conventional" machining methods [Bro98]. 3D printing methods widen the potential of implementing complex shaped 3D models and simplify their production at reasonable cost. Specific is also the area of industrial design, where combination of procedures of free forming surfaces modeling and visualization methods allows the implementation of proposals combining functional and aesthetic aspect [Sta11].

2. BASIC MODELING METHODS

Body components with unusual shapes can be usually created by the basic methods of design, based on the insertion of the base - primitive body and subsequently predefined standard design elements. It is possible to create a wide range of rather simple components, which are unlikely to stronger later editing and with a relatively small number of objects forming. This is known as a so called Features-method [For07].

If allowed by the used application, Features and design method can be combined using hybrid methods. The result is a 3D surface or solid body, which can be edited by changing the parameters forming elements, features or of sketches [Dvo13]. Sketch is geometry composed of points, lines, circular arcs and certain approach to free shapes is enabled by spline here. An important factor in basic and advanced modeling is associativity and parametrics of objects. Associativity represents the relationship of objects, usually based on a common attribute, which is usually the position, size and orientation of the object, or coherence of parameters directly, or through a mathematical relationship [Cam16]. On the indexed parameters can

be referenced not only within individual objects within the model, but also from other models, which are part of the assembly [Chen04].

Associative parametric models have a transparent history and subsequently created objects linked to the previous cannot be shifted within a list before the one you created earlier. Nonparametric model does not have a detailed history, objects do not have a mutual bond and subsequently created object can be used to edit the previous one. This process is used for example when smoothing freely shaped surfaces when curve, the resulting intersection of a plane and the surface is subsequently used as the parental forming area and when modifying it is adjusted [Ata13].

3. ADVANCED FREE FORM MODELING METHODS

Surface and solid objects created by methods of free shaping are obtained by one or a combination of the following procedures:

- Creating curves by modeling methods.
- Import points from an external file.
- Utilization edges and surfaces of existing surface and solid geometry.
- Defining transitions between multiple body or surface faces.

Primary or basic surface geometry also represent surfaces arising from forming geometry tools for creating surfaces. Geometry arising from previous surface geometry can be considered for secondary surfaces as transitive, related, or editing surfaces, imported in a universal format for third-party for data conversion among different systems.

Primary surfaces are formed of a curved geometry by modeling methods. Areas are considered as primary. Conversely, surfaces that are generated from already existing surfaces, e.g. as transitions or extractions of solids are secondary surfaces. The submitted text is focused on creating and editing primary areas.

Ruled

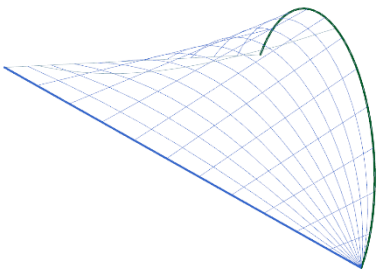


Figure 1. Isolines of ruler's surface.

Forming geometry is represented by any two curves, or a curve and a point. Segments of a surface represent line segments connecting bordering points on forming geometry of curves. This is the simplest type of primary surface. An example of ruler's surfaces - Ruled is shown in Fig. 1. The example shows the area formed by basic units of forming geometry, where curves segments of a surface are formed by a line and a circular arc, not lying in one plane [Bay10]. One of the forming geometry segments can be also a point. Basic procedures of the ruler's surfaces can still be modified by using specifying methods, modifying the segments of used geometry [Kos15].

Through Curve

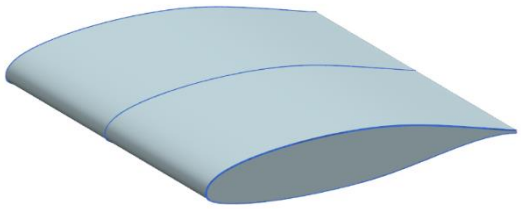


Figure 2. Through Curve on forming curves.

Advanced method of creating curves represents the primary Through Curve Mesh, Fig. 2. It uses a network of curves, which are not interconnected. It consists of slice surface, over which a surface within the modeling tolerance lies.

Editing of a surface is done by editing of each of the segments forming the surface separately. The number of cuts may be arbitrary, min. 2. If the area is related to one of the already developed surface, we can define tangency and curvature. The accuracy and variability of surface is greater than in the case Ruled, but we cannot determine the properties in other directions than along a geometry. E.g. the wing of the plane with a rectangular floor plan is thus created model accurate enough.

Through Curve Mesh

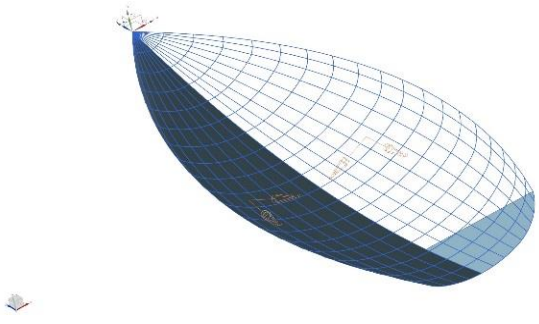


Figure 3. Through Curve Mesh on a net of forming curves.

The method is based on the use of the network of curves, the individual constituting elements are divided into two categories - Primary and Section. Primary Curves can be considered longitudinal and

transverse section curves are used for the ribs. Crosscurves between slices refine the formation of surface and lead to a more exact and more editable result than the previous method Through Curves. Simultaneously, in editing arises more restrictive conditions that may cause difficulty in editing a specific shape, and optionally following selection of Through Curves. Example of surface formed by Curves Through Mesh is in Fig. 3. The number of longitudinal and transverse curves can be arbitrary, min. two in both cases. Just like Through Curve connections to adjacent areas can be defined. Primary surface created by various methods can be combined in one model and determine the accuracy and editability of a geometry as a whole.

Swept

Swept surface is advisable when required pulling cut geometry by one to three curves of the guide with the possibility of specifying the vector of cut surfaces orientation.

Possibility to define the orientation vector surface while creating is obvious in example of the model of a helical coil spring with a square cross-section. Fig. 4 is an example of pulling a square cross-section of the helix using the method Sweep Along Guide without specifying vector of surface creation.

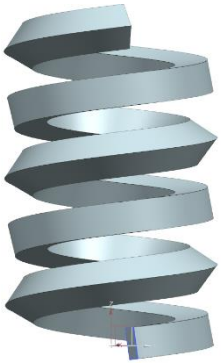


Figure 4. Sweep along guide, without defining section orientation.

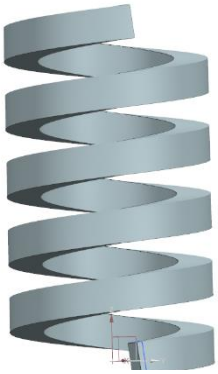


Figure 5. Swept with defining section orientation.

Cut is twisted in proportion to shooting tangent and normal vector to the helix. The following Fig. 5 shows the use of identical forming geometry for formation of the spring with orientation of the vector on the axis of the helix. Using Sweep surfaces is useful when demand constant trailed or variable cut after one to three defining curves.

Effective Methods of free Surfaces editing.

Procedures of the primary surface of the curves are highly efficient and accurate methods of free surfaces, which are editable by changing the geometry forming. The procedures described above can be used in case of direct assignment and the possibility to create, or modify the model fundamentally [Kim16]. In the case of freely deformable surfaces can be used for direct editing groups of tools, allowing a change of control surfaces on the poles of isolines to change the shape. Typical tool is e.g. X-form. Selecting the density of points on field lines, representing surface poles can determine the accuracy of the edits and depth of surface changes. Demonstration of modification and editing of the planar surfaces by the X-form is in Fig. 6.

Managing poles of edited surface can be mounted on points that may be created as the default auxiliary geometry or can be imported as a point cloud from an external file. Such points for editing are often the results of measurements of spatial data, or the results of calculations using external tools.

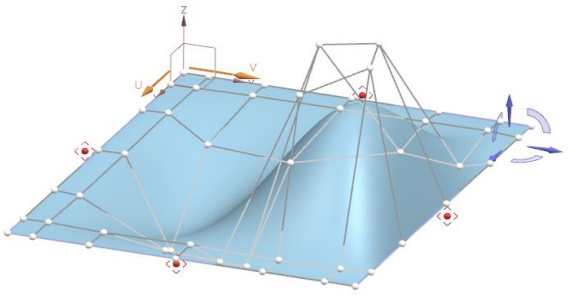


Figure 6. Creating and editing free surface using poles on isolines.

Other editing options are based on the extraction of non-parametric geometry - curves on surface, caused by a section on the plane, or offset curves on a surface. Non-parametric geometry is then modified and used as forming in their place of origin to modify parental surface. Mentioned method is often used especially when smoothing surfaces. Editing surface by methods of free shaping is related to the analysis of forming curves and continuous analysis of the area during editing.

4. SURFACE ANALYSIS

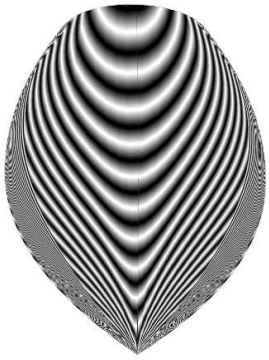


Figure 7. Analysis of the reflectivity of the symmetrical surface.

Another case of surface analyzes is assessment of consistent shaped body while smoothing surface where we perform an analysis of reflectance, involving virtual projection spectrum rays on flat body and the deformation of the image. Example of reflectance analysis is in Fig. 7. Surface analyzes are of great importance especially when constructing cavities molds for casting, or injection molding of complex products.

5. CONCLUSION

Creating a model including free forms offers a wide range of approaches and methods. Specific attributes of complicated freeform surfaces is a difficult identification of negative chamfers, which is a key parameter in the creation of tools, especially molds for plastic injection, which currently represents a significant share of industrial methods. Progressive, currently significantly incoming technology, is the 3D printing - Rapid Prototyping. Free forms, modeled in the virtual environment of a computer application can be easily implemented using a 3D printer at acceptable manufacturing tolerances. Postprocessing of data for 3D printing from digital models is much easier than preparing a program for numerically controlled machine tools, and even complex free form shapes are due to the possibilities of available machine tools difficult to implement. Research and development of methods of free shaping and subsequent analysis of surfaces is based on a systematization of procedures and identifying a key for the best choice when deciding in case of the possible using of multiple approaches.

ACKNOWLEDGMENT

This research was supported by the College of Polytechnics, Jihlava, Czech under Grant no. 1200/04/164 ”Research on advanced 3D Computer Aided Design modeling methods”.

REFERENCES

[Ata13] Atariah, D., S. Ghosh a Günter Rote. On the Parameterization and the Geometry of the Configuration Space of a Single Planar Robot. Journal of WSCG., Vol. 21(1), 11-20, 2013.

[Bay10] Bayar, A. a K. Sami. Parametric Curves Variations with Respect to a Given Direction. Journal of WSCG. Vol. 18(1-3), pp. 33-40, 2010.

[Bro98] Browne, J. a C. McMahon, CAD/CAM: Principles, Practice, and Manufacturing Management. Addison Wesley Publishing Company, 1998.

[Cam16] Camba, J.D., M. Contero a C. Pedro. Parametric CAD modeling: An analysis of strategies for design reusability. Computer-Aided Design, Vol. 74, pp. 18-31, 2016.

[Dvo13] Dvořák, K. Management of parametric CAD model by external tools. Applied Mechanics and Materials. 390, pp. 616-620, 2013.

[For07] Fořt, P. a J. Kletečka. Autodesk Inventor: Functional design in industrial practice. 2. Issue, Brno : Computer Press, 2007.

[Chen04] Chen, G., Y. S. Thimm a S. H. Tang, Unified Feature Modeling Scheme for the Integration of CAD and CAx. Computer-Aided Design & Applications. Vol. 1, 1-4, pp. 595-601, 2004.

[Kim16] Kim, Y.J., G. Elber a M. Kim, Precise contact motion planning for deformable planar curved shapes. Computer-Aided Design, Vol. 70, pp. 126-133, 2016.

[Kos15] Kosinka, J., Malcolm A. Sabin a Neil A. Dodgson. Control vectors for splines. Computer-Aided Design, Vol. 58, pp. 173-178, 2015.

[Lee99] Lee, R., Principles of CAD/CAM/CAE. Michigan University: Pearson Higher Education & Professio-nal Group, 1999.

[Sam08] Samuel, S., B. Stevenson a E. Weeks, Advanced Simulation using Nastran: Project Oriented Learning Manual for People Who design Stuff. San Jose: Design Visionaries, 2008.

[Shi10] Shigley, J. E., Ch. R. Mischke. a R. G. Budynas, Mechanical Engineering Design. 1. Issue, Edi-tor M. Hartl, M. Vlk. Brno: VUTIUM, 2010.

[Sta11] Stavric, M. a O. Marina, Parametric Modeling for Advanced Architecture. International Jour-nal of Applied Mathematics and Informatics. University Press, Vol. 5, Issue. 1,pp. 9-16, 2011.

[Zei10] Zeid, I. a R. Sivasubramanian, CAD/CAM: Theory and Practice. New Delhi: Tata McGraw Hill Education Private Ltd., 2010.

Virtual reality application applied to biomedical models reconstructed from CT scanning

Jordi Torner
jordi.torner@upc.edu

Sergi Gómez
Sergi.Gomez-Gonzalez@upc.edu

Francesc Alpiste
francesc.alpiste@upc.edu

Miguel Brigos
miguel.brigos@upc.edu

Universitat Politècnica de Catalunya, BarcelonaTech
C/Urgell 187
08036, Barcelona, Spain

ABSTRACT

Virtual reality-simulated environments have been used for training for more than 40 years. In recent years, an active development of 3D technologies dealing with medical training, planning and guidance has become an increasingly important area of interest in both research and health-care. Moreover, surgery planning and diagnosis using virtual reality (VR) technology has received much attention. User studies show that the method is both effective and easy to use.

The purpose of this paper is to present a VR application based on three-dimensional (3-D) surface reconstruction CT imaging for surgical planning and simulation. We propose an interactive tool to improve observation before surgery and increase diagnosis efficiency through virtual reality. This tool allows to obtain a more accurate representation comparing with current technology, such as CT or MRI vision. This cost-effective solution can be easily adopted in any hospital for daily use.

Keywords

Virtual reality (VR) · CT scanning · Biomedical models.

1. OBJECTIVES

The aim of the study is to describe the method that can build a three-dimensional (3D) high-resolution model using DICOM images from CT/MRI and export them to create a virtual environment that improves medical knowledge. Also establishing an interactive environment between the user and images based on virtual reality. Another objective is to build an easy and effective tool suitable to be disseminated.

2. METHODOLOGY

Functional requirements required in the app are:

1. The application must allow users to move and rotate the reconstructed model facilitating zoom in and out simply pressing buttons.
2. The application should play model rotation animations for easy viewing from different points of view.
3. Layers will have to show different human tissues separately.
4. Some application settings may be controlled by the user, such as speed of movement, brightness and screen contrast.

In order to generate a 3D virtual reality model and animations, two types of software are required; firstly, 3D reconstruction software that translate the DICOM (digital imaging and communications in medicine) files from CT/MRI scans into a 3D CAD files (. STL or .OBJ), and secondly Unity 3D or similar to create a virtual reality environment compatible for Oculus Rift glasses.

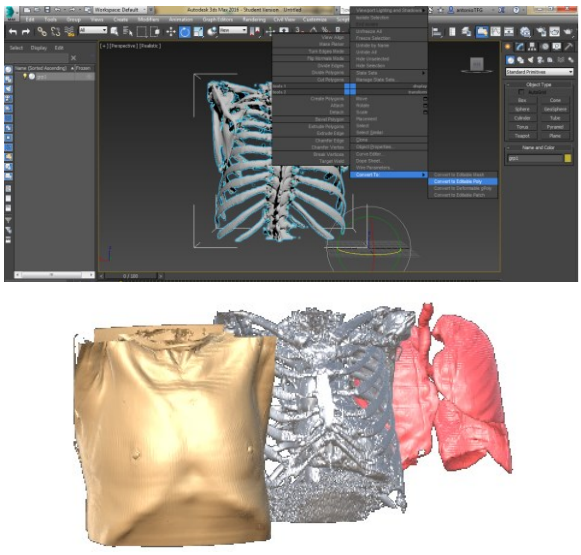


Figure 1. Modelling software 3D Studio Max.

There are many 3D reconstruction software options, free open-source, such as 3D Slicer or InVesalius and 3D professional software like Mimics, 3D Doctor or ScanIP. In all instances the working process is similar. A region of interest (ROI) must be selected and then use threshold and segmentation function to define the tissue to reconstruct. STL (Standard tessellation language) export 3D file is one of the most used. InVesalius was selected as development tool after analyzing other alternatives. It is a free software (GNU GPL 2) than generates easily and quickly 3D medical imaging reconstruction based on a sequence of 2D DICOM files acquired with CT or MRI equipment. The 3D reconstruction models can be save as STL format to its export to other 3D CAD software like Unity 3D or 3D Studio Max (Table 1).

When the desired models meet the required characteristics, they are brought to a development platform (Unity3D), allowing visualization of the object in a more interactive environment where analysis functions allow to provide more information, such as increase areas of interest, rotate the model, select different layers, etc., giving a more realistic and simple vision.

Table 1. Software comparison

Software	Company	Free	OS platform
3D Slicer	Surgical Planning Laboratory	Yes	W,M
Osirix	Pixmeo	Yes	M
MITK	German Cancer Research Centre	Yes	W,M
MIPAV	NIH CIT	Yes	W,M
InVesalius	CTI	Yes	W,M
MeVisLab	MeVis Medical Solutions AG	Yes	W,M
3D Doctor	Able Software	No	W
Amira/Avizo	FEI Visualization Science Group	No	W,M
Mimics	Materialise NV	No	W,M
ScanIP	Simpleware	No	W
VoXim	IVS Technology	No	W
Dolphin Imaging 3D	Dolphin Imaging and Management	No	W

Threshold/segmentation and Export STL included in all software
STL, Standard tessellation language; OS, operating system; W, Windows OS; M, Mac OS.

Surgical preparation with virtual reality may represent a more cost-effective and efficient alternative. Virtual reality training improves operating performance. The use of VR surgical simulation to train skills and reduce error risk in the OR has been applied over the last years. User studies show that the system is both effective and easy to use.

In Figure1, we can see the main menu, with different options to select layers, split layers, and sectioning. There are also available buttons to zoom in / out and rotate, both vertically and horizontally.

Figure 2. Modelling software 3D Studio Max.

Models can be animated and displayed in three different layers.

Three buttons were created to break down (divide the object into parts), create layers (define different tissues) or section (cut model) (Figure 2).

Options rotate and zoom are controlled. Speed and Brightness adjustments have also been taken into account in the application.

The user is able to select options moving the pointer with his own head. Thus, there is no need to use input devices.

3. RESULTS

Surgical preparation or diagnosis with virtual reality may represent a more cost-effective and efficient alternative. Virtual reality training improves operating performance. The use of VR to train skills and reduce error risk in the OR has been applied over the last years. User studies show that the system is both effective and easy to use.

At this point, the app should be tested in the hospitals that are collaborating with our university. Specifically, it will be tested in minimally invasive surgery procedures such as vertebro- and kypho-plasty in a University Hospital. This app is important to reveal the different microstructural features of different vertebra (sacral, lumbar, etc.) in order to provide specific and individual solutions in

accordance to the actual state of the osteoporotic vertebra that is going to be infiltrated by bone cement. This approach will serve to optimize both the instrumental and the surgical procedure.

4. CONCLUSIONS

This process can be applied and should assist future authors to improve similar experiences. This application will be available for different platforms such as Android, iOS and Windows, so it's possible to use it on different devices. The cost-effective solution proposed here can be easily adopted in hospitals.

5. REFERENCES

[AMB+04] Adamovich, S. V., Merians, A. S., Boian, R., Tremaine, M., Burdea, G. S., Recce, M., & Poizner, H. (2004, September). A virtual reality based exercise system for hand rehabilitation post-stroke: transfer to function. In Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE (Vol. 2, pp. 4936-4939). IEEE.

[MWLK00] Mettler Jr, F. A., Wiest, P. W., Locken, J. A., & Kelsey, C. A. (2000). CT scanning: patterns of use and dose. Journal of radiological Protection, 20(4), 353.

[SSR94] Sullivan, P. K., Smith, J. F., & Rozzelle, A. A. (1994). Cranio-orbital reconstruction: safety and image quality of metallic implants on CT and MRI scanning. Plastic and reconstructive surgery, 94(5), 589-596

[Thi98] Thirion, J. P. (1998). Image matching as a diffusion process: an analogy with Maxwell's demons. Medical image analysis, 2(3), 243-260.

A ballistic missile shutdown point estimation method based on double fitting corrections

Jing Hu[1]
Science & Technology on Multi-spectral Information Processing Laboratory,
School of automation,
Huazhong University of Science and Technology
Wuhan(430074), China
[1]hust_h@163.com

Hongtao Gao[2], Xiangyu Kong[3], Fan Dong[4]
School of automation,
Huazhong University of Science and Technology
Wuhan(430074), China
[2]HongtaoGao_1989@outlook.com
[3]734654043@qq.com
[4]M201372322@hust.edu.cn

ABSTRACT

For long-range infrared systems, a new method is proposed in this paper to estimate the shutdown point of ballistic missile. In order to reduce the effect of model error and positioning error of observation point on estimation accuracy, two successive fitting corrections are used in a three-dimensional observation space and a two-dimensional characteristic space respectively. Firstly, the three-dimensional observation data points are fitted by a trajectory plane model, and these points are projected to the trajectory plane for the first correction. Then, a characteristic space l-β is set up to describe these projective points in a two-dimensional polar coordinates and the projective points are fitted and corrected by polynomial curves along two axes respectively. Finally, the motion characters of these projective points are converted to the first three-dimensional observation space and the motion state of shutdown point are estimated. The simulation results show that our method is feasible and valid.

Keywords

ballistic missile, shutdown point, trajectory plane, characteristic space, double corrections.

1. INTRODUCTION

As ballistic missiles have been the main component element in long-distance guidance weapons[1]. Ballistic missile tracking in time is one of the key technologies in the missile defense system. Many related work about ballistic missile tracking have been done. Xu Zhang [1] proposed a method for ballistic missile trajectory prediction based on predictor-corrector method considering the radar's measurement noise. Zhang Feng [2] investigated the method of ballistic missile tracking using dynamic model. Dong Gwan Lee [3] proposed a method to predict ballistic missile trajectory based on the Kepler's law for the missile defense system using cueing information. Robert L. Cooperman [4] developed a tactical ballistic missile tracker with an interacting multiple models framework to solve the problem of tracking a ballistic missile by the varying target dynamics in the boost, exo-atmospheric and endo-atmospheric phases of flight. Bennavoli[5] proposed an approach which combines a nonlinear batch estimator with a recursive multiple model particle filter in order to estimate the launch and impact points of a ballistic target.

In order to predict position and velocity vector of shutdown point, this paper proposes a shutdown

point estimation method based on double fitting corrections.

2. PROBLEM & CLASSIC METHOD

All observation points are from a three-dimensional space which can be represented by the Inner earth space rectangular coordinates $O-XYZ$. The estimated parameters of shutdown point need to be represented in the same space. It is assumed that the launch time of the missile is t_0 , the first time of observation data points is t_m , the shutdown time is t_n ($0 \le m < n$). Target position and velocity vector when the shutdown time can be estimated from three-dimensional observation data points. In the post-boost phase, the ballistic missile targets are mainly acted upon by gravitational force and follow a Keplerian orbit, so fall point depends mainly on parameters of shutdown point. In addition, because of 6 parameters, it is not simple and intuitive to measure the estimation accuracy of shutdown point. So, fall point error(FPE) is used to measure the shutdown point estimation accuracy in this paper.

Classic method estimated the position and velocity vector using polynomial fitting in x , y , z directions respectively. Shutdown point error of classic method relatively large because it is the sum of three directions. If a trajectory plane can be determined , these three-dimensional points can be projected to the trajectory plane and estimation of shutdown point can be processed on a 2-D space. The FPE would become the sum of errors in 2 directions. In theory, the sum of errors in 2 directions is smaller than that in 3 directions.

3. PROPOSED METHOD

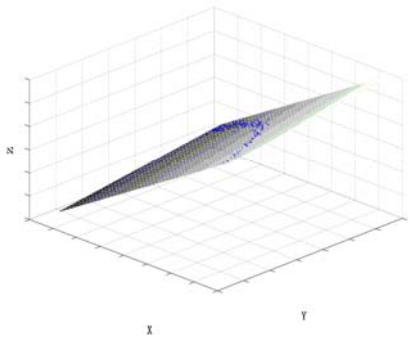


Figure 1 Trajectory plane and projective points

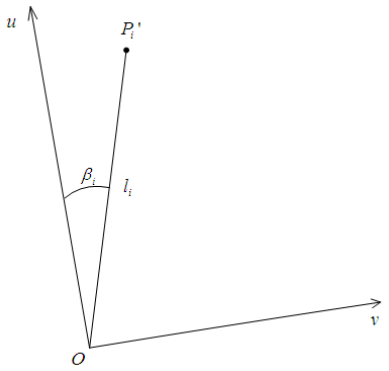


Figure 2 Relationship between $l-\beta$ and $u-v$

The specific steps are as follows:

S1. Determining the trajectory plane:

The trajectory plane: $Ax+By+Cz+D=0$ is obtained by linear regression with some three-dimensional observation points.

S2. Projecting observation points to plane:

According the projecting method of point to plane, we can get the projective point $P_i'(P_{Xi}', P_{Yi}', P_{Zi}')$ of the observation point $P_i(P_{Xi}, P_{Yi}, P_{Zi})$ at the time t_i , $t_k \le t_i \le t_n$. Trajectory plane can be seen in Figure 1.

S3.Calculating l and β :

The first-correction vector at time t_i is defined as

$$\begin{aligned}&\overrightarrow{OP_i'}(OP_{Xi}', OP_{Yi}', OP_{Zi}')\\&= P'-O=(P_{Xi}'-O_X, P_{Yi}'-O_Y, P_{Zi}'-O_Z)\end{aligned} \quad (1)$$

Where (O_X, O_Y, O_Z) is the projective point of the center of earth.

The initial vector is defined as the first-correction vector at t_m : $\overrightarrow{OP_m'}(OP_{Xm}', OP_{Ym}', OP_{Zm}')$, then it is also defined as the polar axis;

The length l_i and the angle β_i of the first-correction vector at time t_i are

$$l_i = \left|\overrightarrow{OP_i'}\right| \tag{2}$$

$$\beta_i = a\cos(\overrightarrow{k_i} \bullet \overrightarrow{k_m}) \tag{3}$$

where $\overrightarrow{k_i}$ is the unit vector of $\overrightarrow{OP_i'}$, and $\overrightarrow{k_m}$ is the unit vector of $\overrightarrow{OP_m'}$.

S4.Transforming characteristic vectors in Coordinates $l-\beta$ to $u-v$:

Build a rectangular coordinates $u-v$, its origin coincide with the origin of coordinates $l-\beta$, and u axis coincide with the polar axis, and v axis is perpendicular to u axis. Transform projective point $P_i'(l_i,\beta_i)$ to $u-v$ coordinates, the result is $P_i'(P_{Ui},P_{Vi})$. The transformation formulas are:

$$\begin{cases} P_{Ui} = l_i \times \cos\beta_i \\ P_{Vi} = l_i \times \sin\beta_i \end{cases} \tag{4}$$

The relationship between $l-\beta$ and $u-v$ can be seen in figure 2.

S5.Second correction in coordinates $u-v$:

$$u(t) = a_1t^3 + b_1t^2 + c_1t + d_1 \tag{5-1}$$

$$v(t) = a_2t^3 + b_2t^2 + c_2t + d_2 \tag{5-2}$$

In coordinates $u-v$, $P_i'(P_{Ui},P_{Vi})$ ($k \le i \le n$) can be fitted by the above curves model (here, we take 3 as the fitting order). $P_i''(P_{Ui}',P_{Vi}')$ is the corrected point of $P_i'(P_{Ui},P_{Vi})$. Then take derivative of $(10-1)$, we obtain:

$$u'(t) = 3a_1t^2 + 2b_1t + c \tag{6}$$

The velocity in u direction can be get as $V_{Ui} = u'(t_i)$. The velocity in v direction, V_{Vi} can also be obtained.

S6.Calculating position and velocity of shutdown point in observation space:

$P_n''(P_{Un}',P_{Vn}')$ and $\overrightarrow{V_n}(V_{Un},V_{Vn})$ are the position and the velocity of shutdown point in $u-v$ coordinates.

Assume that point $P_n''(P_{Xn}'',P_{Yn}'',P_{Zn}'')$ is the estimated position of shutdown point in observation space, let the corresponding second-correction vector

$$\begin{aligned}\overrightarrow{OP_n''} &= P_n'' - O \\ &= (P_{Xn}'' - O_X, P_{Yn}'' - O_Y, P_{Zn}'' - O_Z)\end{aligned}.$$

The position of P_n'' can be obtained by solving:

$$\begin{cases} A \bullet P_{Xn}'' + B \bullet P_{Yn}'' + C \bullet P_{Zn}'' + D = 0 \\ l_n' = \sqrt{P_{Un}'^2 + P_{Vn}'^2} \\ \beta_n' = \arctan(\dfrac{P_{Un}'}{P_{Vn}'}) \end{cases} \tag{7}$$

Where $l_n' = \left|\overrightarrow{OP_n''}\right|$, and $\beta_n' = \arccos(\overrightarrow{k_n'} \bullet \overrightarrow{k_m})$, and $\overrightarrow{k_n'}$ is the unit vector of $\overrightarrow{OP_n''}$.

It is obvious that the following equations can be obtained according to geometric characteristics:

$$\begin{cases} \overrightarrow{k_m}(1)(P_{Xi}'' - O_X) + \overrightarrow{k_m}(2)(P_{Yi}'' - O_Y) + \overrightarrow{k_m}(3)(P_{Zi}'' - O_Z) = \cos(a\tan(P_{Ui}'/P_{Vi}')) * \sqrt{P_{Ui}'^2 + P_{Vi}'^2} \\ A * P_{Xi}'' + B * P_{Yi}'' + C * P_{Zi}'' + D = 0 \\ (P_{Xi}'' - O_X)^2 + (P_{Yi}'' - O_Y)^2 + (P_{Zi}'' - O_Z)^2 = P_{Ui}'^2 + P_{Vi}'^2 \end{cases} \tag{8}$$

Take derivative of Formula (8), we can obtain the following equations:

$$\begin{cases} \overrightarrow{k_m}(1)*V_{Xi}+\overrightarrow{k_m}(2)*V_{Yi}+\overrightarrow{k_m}(3)*V_{Zi}=-\sin(\arctan(P_{Ui}'/P_{Vi}'))*\dfrac{V_{Ui}*P_{Vi}'-V_{Vi}*P_{Ui}'}{P_{Ui}'^2+P_{Vi}'^2} \\ *\sqrt{P_{Ui}'^2+P_{Vi}'^2}+\cos(\arctan(P_{Ui}'/P_{Vi}'))*\dfrac{V_{Ui}*P_{Ui}'+V_{Vi}*P_{Vi}'}{\sqrt{P_{Ui}'^2+P_{Vi}'^2}} \\ A*V_{Xi}+B*V_{Yi}+C*V_{Zi}=0 \\ (P_{Xi}''-O_X)*V_{Xi}+(P_{Yi}''-O_Y)*V_{Yi}+(P_{Zi}''-O_Z)*V_{Zi}=V_{Ui}*P_{Ui}'+V_{Vi}*P_{Vi}' \end{cases} \quad (9)$$

Velocity vector $\overrightarrow{V_n}(V_{Xn},V_{Yn},V_{Zn})$ can be obtained by Formula (9).

4. RESULTS

The performance of the classic method and our proposed method are evaluated by simulation experiments.

3-D observation data points of a ballistic missile trajectory are simulated with positioning error of $\alpha(\alpha \neq 0)$. Take the fitting order to 3, number of fitting points to 150-450, some simulation experiments are carried out. Figure 3 illustrates FPE curves along with the number of fitting points in different methods. The number of points used to determine the trajectory plane in the experiment is 500, and it is the first optimal value obtained by simulation experiments.

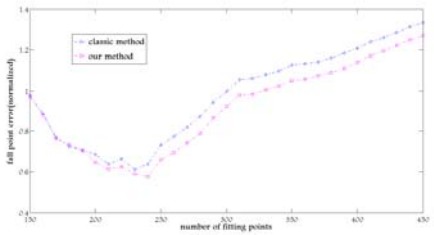


Figure 3 Normalized FPE

It is shown that the predicted FPE by our proposed method is smaller than that by the classic method.

5. CONCLUSIONS

This paper proposed a shutdown point estimation method with higher accuracy by converting comprehensive error of 3 directions to 2directions. The simulation results show that our method is feasible and valid.

6. REFERENCES

[1] Xu Zhang, Hu-Min Lei, Jiong Li and Da-Yuan Zhang. Ballistic Missile Trajectory Prediction and the Solution Algorithms for Impact Point Prediction.Guidance, Navigation and Control Conference(CGNCC), 2014 IEEE Chinese on, pages 879-883, Aug. 2014.

[2] Zhang Feng, Tian Kang-sheng and Xi Mu-lin. The Ballistic Missile Tracking Method Using Dynamic Model. Radar(Radar), 2011 IEEE CIE International Conference on, Vol. 1, pages 789-794, 2014.

[3] Dong Gwan Lee, Kil Seok Cho and Jin Hwa Shin. A Simple Prediction Method of Ballistic Missile Trajectory to Designate Search Direction and its Verification Using a Testbench. Control Conference (ASCC), 2015 10th Asian, pages 1-7, 2015.

[4] Robert L. Cooperman. Tactical Ballistic Missile Tracking using the Interacting Multiple Model Algorithm. Information Fusition, 2002. Proceedings of the Fifth International Conference on, Vol. 2, pages 824-831, 2002.

[5] A. Bennavoli , L.Chisci and others. Tracking of a Ballistic Missile with A-Priori Information. Aerospace and Electronic Systems, IEEE Transactions on, Vol. 43, pages 1000-1016, 2007.

Hardware-Software Embedded Face Recognition System

M.J. Avedillo
IMSE-CNM
(CSIC/Univ. Sevilla)
c/Americo Vespucio
41092-Sevilla, Spain
avedillo@imse-cnm.csic.es

A. Barriga
IMSE-CNM
(CSIC/Univ. Sevilla)
c/Americo Vespucio
41092-Sevilla, Spain
barriga@imse-cnm.csic.es

L. Acasandrei
IMSE-CNM
(CSIC/Univ. Sevilla)
c/Americo Vespucio
41092-Sevilla, Spain
laurentiu@imse-cnm.csic.es

J.M. Calahorro
ETSII
Univ. Sevilla
a/ Reina Mercedes s/n
41012-Sevilla, Spain

ABSTRACT

This paper describes the design and implementation of a hardware-software embedded system for face recognition applications in images and/or videos. The system has hardware components to speed up the face detection and recognition stages. It is a system suitable for applications requiring real-time, due that the response times are deterministic and bounded. The system is based on a previous implementation that had accelerated the image capturing process, and the face detection. This paper will focuses in the face recognition acceleration.

Keywords

Hardware-software codesign, embedded system, face recognition, FPGA implementations, high level synthesis

1. INTRODUCTION

This communication presents the design of an embedded system to accelerate the recognition of faces in images and/or videos. A recognition system consists of four steps: 1. *Face detection* to detect if there is a face in the image (it provides the location and size of the face in the image); 2. *Face alignment* to locate the position of the face and, using geometric transformations, normalizes it with respect to geometric properties, such as size and pose, and photometric such as lighting; 3. *Feature extraction* to provide a feature vector with information to distinguish faces from different individuals according to geometric or photometric variations; 4. *Recognition step* in which the extracted feature vector is compared with the vectors in a database.

The system (Fig. 1) receives data from an image sensor (camera). Each frame is stored in internal memory and is processed by the system. The processing performed by the recognition algorithm requires two components: a software application and the hardware accelerators. The software application runs on a processor, and realizes the initialization and control of the hardware, as well as the recognition algorithms. The hardware accelerators can accelerate those tasks that constitute the "bottleneck" of the recognition algorithm. Thus, the proposed system (Fig. 1) is a hardware-software solution, which includes hardware accelerators to implement the most computationally expensive part of the face recognition algorithms: capturing images from the camera, processing for face detection and recognition algorithm acceleration. This paper focuses on the description of the Face Recognition Acceleration.

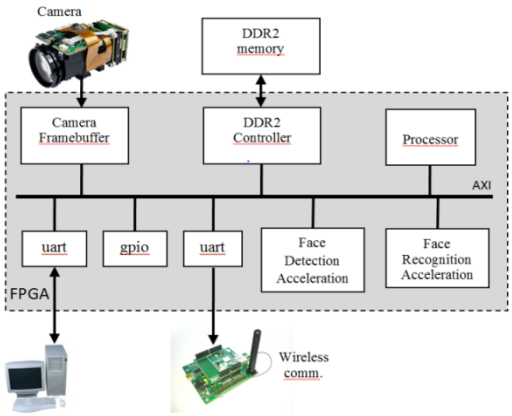


Figure 1. Block diagram of the embedded system

2. HARDWARE ACCELERATOR DESIGN

The face recognition algorithms implemented require matrix product operations which constitute the "bottleneck" of the system, limiting the operation speed. So it was decided to implement this part in hardware. The design of the hardware accelerator has been carried out using the high level synthesis tool from Xilinx Vivado-HLS [Xil14]. Vivado-HLS

generates an RTL description starting from a C/C++ algorithm level description. It realizes the scheduling and the resource allocation in order to map the algorithm to hardware. It also generates a description as IP module so that it can be used as a peripheral of a processor.

This methodology allows the designer to start the design of the system to be implemented in hardware from high-level descriptions. This means that algorithmic descriptions are made in a high-level programming language (in our case C++). Vivado-HLS takes as input the high level description and is able of generating a circuit that implements the desired algorithm. The designer can set restrictions, using directives (pragmas), on latency, throughput or hardware resources. As we will see, this methodology allows for exploration of the design space (due to automation), and design optimization through the application of directives.

Specification

The multiplication of the row vector containing the information of the input image by the principal component matrix (generated in the training stage and stored in the database) is computationally expensive. For this reason, it was decided to design a circuit to perform this operation.

The hardware block multiplies two matrices of fixed point numbers. The first matrix has one row and as many columns as the number of pixels of the image. That is, its size is *1*DIM*1. The dimension of the second matrix, *B*, is *DIM*1* *DIM*2, where *DIM*2is the number of principal components. The elements of the matrices use a 32 bits signed fixed-point representation, with 16 bits for the fractional part. This is given by the type of data used in the software application that it aims to accelerate.

The accelerator works in streaming mode, so it receives and sends data sequentially. There is only one port for receiving the two operand matrices. This is determined by the architecture requirements of the hardware platform, which has been designed to optimize data transfer between the processing system and the accelerator. The operating frequency of the system is 100 MHz.

Description

Figure 2 shows the description in C++ language. The ***accelerator*** function is the top function and it is the function to be synthesized using Vivado-HLS. In this top function, interfaces for the hardware module are configured and the function that encapsulates the functionality of the hardware block (***accelerator_core***) is called.

The pragmas with the directives HLS INTERFACE and HLS RESOURCE are responsible for managing the sending and receiving of data arrays through the AXI buses. The first two associate FIFO communications protocols to the input and output ports to meet the interface requirements of the block. The other three pragmas add the required adapters in order to connect the module to an AXI4Lite bus for control, and to an AXIStream for data transfer. Finally, the function ***accelerator_core*** is called and the multiplication is performed.

In function ***accelerator_core,*** the first pair of nested loops models the storing of the first array in internal memory (array *a*). Next, the system stores the first column of the second array (first principal component) in internal memory (array *b*) and, then, the multiplication *a* by *b* is carried out. That is, it calculates the first element of the resulting array and stores it in internal memory (array *out*). The procedure is repeated for each column of the second matrix. Finally, the loop labeled *converter* describes the sending of the elements of the resulting matrix.

```
void accelerator(AXI_VAL in_stream[num.pix + num.pix
* num.comp], AXI_VAL out_stream[num.comp]) {

#pragma HLS INTERFACE ap_fifo port=in_stream
#pragma HLS INTERFACE ap_fifo port=out_stream
#pragma HLS RESOURCE variable=in_stream core=AXIS
metadata="-bus_bundle INPUT_STREAM"
#pragma HLS RESOURCE variable=out_stream core=AXIS
metadata="-bus_bundle OUTPUT_STREAM"
#pragma HLS RESOURCE variable=return core=AXI4LiteS
metadata="-bus_bundle CONTROL_BUS"

accelerator_core<ap_fixed<32, 16>, num.pix,
num.comp, 4, 5, 5>(in_stream,out_stream);
return;
}

template<typename T, int DIM, int DIM2, int U, int
TI, int TD>
void accelerator_core(AXI_VAL in_stream[DIM + DIM *
DIM2],

AXI_VAL out_stream[DIM2]) {
#pragma HLS INTERFACE ap_fifo port=in_stream
#pragma HLS INTERFACE ap_fifo port=out_stream
T a[UNO][DIM];
T b[DIM][UNO];
T out[UNO][DIM2];
assert(sizeof(T) * 8 == 32);
for (int i = 0; i < UNO; i++)
  readA_int: for (int j = 0; j < DIM; j++) {
                  int k = i * DIM + j;
                  a[i][j] = read_stream<T, U,
                      TI,TD>(in_stream[k]); }

for (int i = 0; i < DIM2; i++) {
  read_B_int: for (int j = 0; j < DIM; j++) {
                   int k = j + DIM * (i + 1);
                   b[j][0] = read_stream<T,U,
                     TI, TD>(in_stream[k]); }
  T sum = 0;
  Multiplier: for (int id = 0; id<DIM; ++id){
                  sum += a[0][id] * b[id][0]; }
  out[0][i] = sum; }

converter:for (int j = 0; j < DIM2; j++) {
out_stream[j] = write_stream<T, U, TI,
            TD>(out[0][j], j == (DIM2 - 1)); }
return; }
```

Figure 2. C++ algorithm description.

AXI_VAL is a structure which represents the sending of data through the AXI bus.

The initial description has been validated in Vivado-HLS environment by running various functional tests. Obtained results have been checked against those of the multiplication in the original software application.

Synthesis

Once it has been verified that the software description of the multiplication algorithm works properly, the next step is to obtain an RTL implementation. This requires synthesizing the C++ description.

Three different solutions have been generated. Each of them with different optimization directives applied. The accelerator is optimized using pipeline. Table I summarizes the obtained results. The results are for an accelerator module with *DIM*1=*320* and *DIM*2=160 on a Zynq-7020device.

	Latency (cycles)	BRAM	DSP	FF	LUT
Sol1	564644	3	4	294	316
Sol2	512966	3	4	306	365
Sol3	104646	3	4	271	305

Table 1. Synthesis results

Solution 1 corresponds to the synthesis of the C++ description without applying any directive. That is, this solution does not include any optimization. In Solution 2, pipeline directives are applied to all the loops, except the one in which the multiplication is performed (loop labeled *Multiplier*). Solution 3 is the most optimized one since it applies the optimizations of previous solution, but also applies pipeline to the multiplication loop. Precisely this is the loop causing the bottleneck in the matrix multiplication algorithm.

It is noted that the three solutions are similar in terms of used resources. No differences are observed, as expected, in the number of memory blocks and DSPs. The differences in terms of latency (in cycles) are significant. Solution 2 slightly improves Solution 1. For Solution 3, which apply pipeline to the calculation loop, latency is significantly lower than in Solutions 1 and 2.

Furthermore, concerning the use of hardware resources, it is observed that Solution 2 requires slightly more LUTs and flip-flops. This is due to the inclusion of pipeline stages which increases the number of required circuit elements. However, it is noted that Solution 3 requires less LUTs and flip-flops. The reason is that, although more pipeline is added, the synthesis tool has achieved a better optimization of the resulting hardware.

Figure 3 shows the latency, measured in clock cycles for each of the three solutions. The impact of the use of directives in the system performance is graphically observed.

Figure 3. Response time of each implementation.

3. FACE RECOGNITION SYSTEM DESIGN

Once a Pcore containing the description of the hardware accelerator for facial recognition has been generated with Vivado-HLS, the next step is the design of a hardware platform incorporating it as peripheral and the development of a software application which makes use of it.

The system has been implemented in the Xilinx ZedBoard development board, which incorporates a Zynq-7000 SoC device. The hardware platform consists of five components: the ARM processor, the AXI buses, a timer, a DMA controller and the IP accelerator module. Figure 4 outlines the architecture of the system.

The AXI buses are responsible of the communication between the various components of the circuit. The ARM processor, in the PS in Figure 4, runs the software part of the face recognition system. It sends the data to the hardware accelerator. In order to configure the processor, it is needed to enable the S_AXI_ACP interface to communicate with the bus for sending and receiving data.

The timer is used to make performance analysis of the proposed system. Specifically, it is used to measure the number of clock cycles consumed by the accelerator module, in order to compare it with the number of cycles without using accelerator.

The DMA is responsible of the communication between the PS and the IP accelerator module. It provides the input matrices to the hardware accelerator and receives the resulting matrix from it. Also it must make this data available to the processor so that it continues its execution. The DMA operates in burst mode. To reduce the memory access time, the maximum number of data to be transmitted in each burst has been set to 256 Bytes.

The starting point for the design of the embedded system is the OpenCV's baseline face detection/recognition application [Ope10]. The OpenCV baseline face recognition application was modified in order to adapt it to a standalone embedded face recognition system. We have considered that the majority of embedded environments are capable of running C applications

with or without operating system (OS) support. This means that the resulting application code has to be compatible to C compilers and, at the same time it should be platform independent.

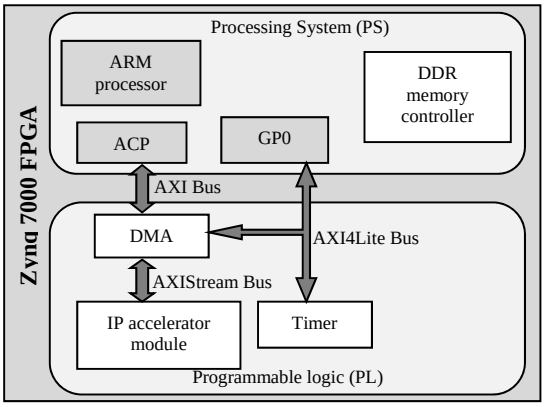


Figure 4. Hardware platform

Another consideration made is the fact that most of the SoC have no floating point support. Because of this, the developed application uses integer operations instead of floating point operations in order to preserve the generality of the application for the embedded system world. Additionally other modifications were made to accelerate the algorithms. Finally, the part of the software application related to the matrices multiplication, needed for projecting the image, has been replaced. This calculation is performed by the IP accelerator module, so it has been necessary to develop different functions to work directly with the system peripherals (DMA and the IP module).

Once the IP module has completed the execution and the DMA has transferred the result of the multiplication, the processor continues running the software application.

4. FACE RECOGNITION SYSTEM EVALUATION

To validate the developed system and to evaluate its performance, different tests have been carried out. Specifically, two types of tests have been made, one with an all-software recognition system, which does not include hardware accelerator, and the other with the system that includes the hardware accelerator. The three implementations of the hardware accelerator described in Section 2 have been tested to experimentally evaluate how optimizations affect the final result.

We have used two image databases containing images of different sizes in order to check the acceleration in relation to the size of the matrices to be multiplied. One database is the AT&T face databases [Att02] (resolution of 92×112 pixels). In AT&T database there are 10 grayscale images for each of the 40 individuals who compose it. The second database is Faces94 [Spa94] (resolution of 180×200 pixels). It contains 3040 color images of faces divided into 20 images for each of the 152 individuals who compose it.

Tests were carried out on four types of systems: the purely software embedded face recognition system and the three hardware platforms with accelerator, as mentioned previously. These three solutions are the hardware acceleration module without optimization (not including pipeline stages), the module that includes pipeline stages in all loops except the multiplication one, and the module including pipeline in all loops.

Figure 5 shows the results obtained for Faces94 database. We can see the impact produced when inserting pipeline on the accelerator, which makes this operation even faster. For the small images (ATT), the multiplication is about 11 times faster, and for the large images (Faces94) it is 17 times faster.

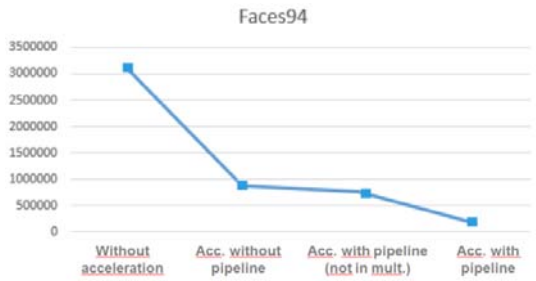


Figure 5. Acceleration measures (in clock cycles)

5. CONCLUSIONS

This communication described an embedded system that accelerates the face recognition process in images or videos. The implemented face identification system is suitable for applications requiring real time because the time response is deterministic.

6. ACKNOWLEDGMENTS

This work was supported in part by Ministerio de Economía y Competitividad under the Project TEC2014-57971-R, co-financed by FEDER.

7. REFERENCES

[Att02] AT&T Laboratories Cambridge, Database of Faces. [On-line] http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.

[Ope10] OpenCV Reference Manual. 2010.

[Spa94] Spacek, L. Collection of Facial Images: Faces94. [On-line] http://cswww.essex.ac.uk/mv/allfaces/faces94.html.

[Xil14] Xilinx. Vivado Design Suite User Guide High-Level Synthesis. UG902 (v2014.1) May 30, 2014.

Data Visualization and Evaluation for Industry 4.0 using an interactive k-Means Algorithm

Dieter Meiller
East Bavarian Technical University Amberg-Weiden
Kaiser-Wilhelm-Ring 23
92224, Amberg, Germany
d.meiller@oth-aw.de

Florian Niewiera
East Bavarian Technical University Amberg-Weiden
Kaiser-Wilhelm-Ring 23
92224, Amberg, Germany
f.niewiera@oth-aw.de

ABSTRACT

The project ISAC@OTH-AW will focus on developing an innovative expert system for data visualization and optimization to produce better manufacturing processes. A mandatory part of the project is the appropriation of industry 4.0 technology benefits, like efficiency, quality and time to market rate, for **SME**s (**S**mall and **M**edium **E**nterprises).

Keywords

Industry 4.0, swarm intelligence, swarm visualization, interactive k-Means, particle swarm optimization, big data, information visualization, visual analytics, machine learning.

1. INTRODUCTION

Modern digital factories are monitored and remote controlled. Large amounts of data are transmitted to interfaces for controlling and monitoring. The next stage is to analyze the data in order to make predictions. The modelling of algorithms for analysis can be supported by interactive visualizations. This topic will be focused in this project.

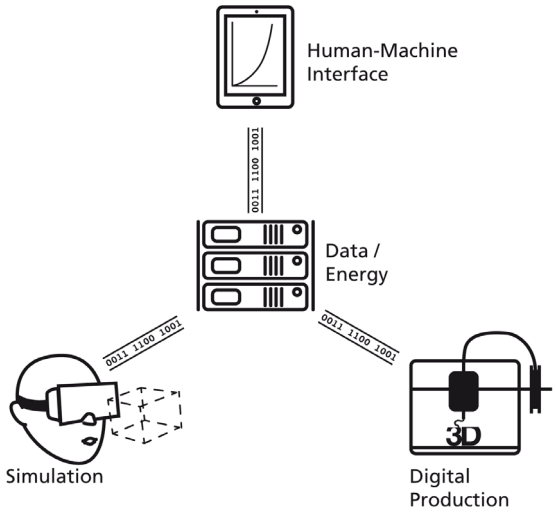


Figure 1. Structure of ISAC@OTH-AW

Four main topics will be addressed within the ISAC@OTH-AW project (Industry Software Application Center at East Bavarian Technical University Amberg-Weiden) :

1 The improvement of production cycles and process workflows, which are based on characteristic production data.

2 Opportunity of facility optimization by virtual plant simulation.

3 Intelligent networked and power supplied controllers and machines.

4 Innovative visualization design for process monitoring and the optimization of data analysis for big data sources.

Four project cooperation partners of two faculties at the Technical University will handle the mentioned topics. The first results are presented in this short paper. They aim to portray the different aspects of human-machine-interface, visualization and the analysis of process data for production optimization in subproject number four.

2. Current requirements in industrial environments by small and medium enterprises (SME)

The desire and need by companies for an optimization, visualization and handling of big data appearances is greater than ever before. Industry 4.0 technologies can eliminate many frequently occurring problems and ease process procedures in a SME. This will be adapted and refined in the

ISAC@OTH-AW project. These problems include many helpdesk center situations for common customer support. The innovative storage and analysis of costumer queries could offer fast and more reliable solutions that are unique to each problem for an individual support. Hence, the reduction of customer support time would raise efficiency and could lead to a higher customer satisfaction with less effort for employees.

Another indispensable operation field for industry 4.0 benefits is the enhancement and optimization of production sequences and machine parameters for example CNC milling, drilling, spray casting, 3d printing. The improvement of these parameters would help accomplish many things including faster process times or a conservative tool usage.

Another possible application is the handling and long term archiving of big data occurrences for selected analysis of relevant data parts in combination with innovative and informative visualization methods. The topic in this periphery is the identification of undetected process correlations, the gain of a clearly represented production overview or an easier detection of essential information.

Furthermore, even data evaluation for logistical problems could benefit from these optimizations to reducing operation time and avoiding an excess use of space and resources.

3. State of the art

As mentioned, the main targets of the analysis component in this project are self-clustering and evaluation algorithms. In particular, big data pools with a variety of different data types will be processed. In the first step, intelligent models and algorithms analog to swarm visualization [Mei15a] and particle Swarm Optimization [Mal15a] [Sha12a] [Kuo11a] will complete the self-organization of these data objects. After this extensive phase of organization, all data objects will be analyzed by using machine learning algorithms [Na15a]. The obtained results of the unsupervised machine learning supported analysis shall deliver new insights.

An advantage of visualizing an amount of m objects in a swarm is, that the provided data could have n features where n can be greater than three. Hence, the dimensionality n is higher than it is suitable for mapping the features to an axis for the drawing. A common way to visualize correlations between data examples is to reduce the dimensionality of the multivariate data using the Principal Component Analysis (PCA) algorithm [Joll02a]. This approach is not suitable when the amount m is big or varying, because the algorithm is expensive to perform. Hence, the data will be visualized as a swarm, where single objects attract each other when they have similar features [Mei15a]. In this approach it is not necessary to reduce the dimensionality because the similarity can be computed from more than three features. The strength of the similarity will increase or decrease the attraction between two single objects and can be observed in a 2D space. This can come to use when performing tasks with industry data. An example of this is to detect classes of support requests in the help desk and to route them to the right administrator. The problem with applying machine-learning techniques such as the k-Means algorithm to identify clusters [Hart79a] is the choosing of the number of clusters K, where K should be smaller than m. This could be solved by visualizing the data and estimate the value of K. Thus, swarm visualization could be an approach to support a subsequent clustering algorithm like k-Means.

4. Procedure model

The computation of the clustering algorithm (e.g. k-Means [Eri11a] [Esm13a]) could be mapped into the swarm visualization as well: Each of the K cluster centroids μ_k ($k \leq K$) could be visualized as an object in the swarm, because it has the same dimensionality as the data objects and will attract similar data. Additionally, clustering algorithms (e.g. k-Means) could be run interactively. An operator could be able to add and remove cluster centroids at runtime, simply by touch:

```
1.  K := 0;
2.  Repeat while clustering
3.    if touch on x(i)
4.      K := K + 1;
5.      add Centroid μK := x(i)
6.    endIf
7.    if touch on μk
8.      K := K - 1
9.      remove Centroid μk
10.   endIf
11.   //Assign Data to centroid
12.   for i := 1 to m
13.     C(i) := index k of
                μk most similar to x(i)
14.   endFor
15.   //Make centroid similar
        to assigned data
16.   for k := 1 to K
17.     n := 0
18.     for i := 1 to m
19.       if C(i) = k
20.         μk:= μk + x(i)
21.         n := n + 1
22.       endif
23.       μk:= μk / n if n > 0
24.     endFor
25.   endFor
26. EndRepeat
```

Algorithm 1. Interactive k-Means algorithm

In Algorithm 1 a new interactive version of the k-Means algorithm is described. $C^{(i)}$ is an index ($i \leq m$) and $x^{(i)}$ is one data sample, a vector of n features ($x^{(i)} \in \mathbb{R}^n$). The algorithm runs continuously.

An endless loop encloses the procedure (line 2-26) to make it run as long as the operator needs to find clusters. At the start, there are no cluster centroids present (number of centroids K is 0 at line 1). The user is able to touch (click) on a representation of a data object $x^{(i)}$ to add a new cluster centroid at this location. The centroids will be removed (lines 8-10) by repeatedly clicking on them. After adding one more centroid, the cluster algorithm will (re-) assign the data points to all centroids available (lines 12–14). In lines 16–25 the centroids will be matched to the data points they correspond with. After that, the centroids will move ahead into swarm visualization to the data points which are assigned to them.

Since we are still in the initial phase of the ISAC@OTH-AW project the algorithms have not been fully developed or chosen yet. This is only an example of how (swarm-) visualizations can support methods of machine learning. We plan to support further techniques of machine learning with swarm visualizations: methods for anomaly detection to determine defects in manufacturing, supervised learning with support vector machines and neuronal networks to predict good machine configurations could benefit from support by interactive visualizations as well.

5. Approach and current status of the implementation

Actual production data from different facility machines like CNC milling machines is collected and accumulated in the first step of the project. This information comprised e.g. timestamps, tool position data, speed-up parameters. All information is saved into a real time compatible NoSQL database like MongoDB. It is an open source, cross platform document oriented database. It is possible to collect different types of data in great quantities because of schema free file storage. Additionally, static information (tool parameter, specimen parameter) belonging to these dynamic machine information parameters will be stored to the data base.

In a second step the first realization and implementation of a purpose build visualization and data analysis framework was initiated. For this, the framework "DataOcean@ISAC" (Figure 2) was created. The application visualizes all process data and parameter as swimming particles in a barrier free environment, such as a fish swarm in the ocean. These individual data objects are self-organized and behave like a fish in a swarm. Algorithms based on Flock Models [Cui00a] [Fow00a], Swarm Intelligence [Cui00c] combined with k-Means [Bin00a] and Particle Swarm Optimization [Cui00b] will be used to organize the data structures by themselves.

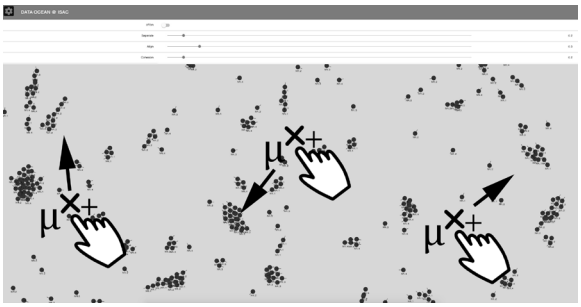


Figure 2. Process of adding centroids to intelligent swarm visualization (DataOcean@ISAC)

The various particles will automatically organize themselves into clusters. These are small groups that are organized by certain features which are unique to all other clusters. This advance grading will separate unnecessary from desired information for the next steps.

Furthermore, a performance optimization of the DataOcean@ISAC visualization algorithm will be strived. The actual algorithm is limited to managing several thousand data objects, depending on the used hardware. The number of presentable particles has to be enhanced in order to use the visualization with big data structures. The use of JavaScript visualization runtime is browser based and adaptive to any display size. Therefore, the enhancement of the performance will guarantee proper usage of inefficient devices like tablets or smartphones in comparison to powerful PCs, touch tables or big touch walls. Figure 3 depicts an early version of the DataOcean@ISAC framework used on an 84" touch wall, which can be used in an authentic facility environment and is able to display big amounts of data.

Figure 3. Human Machine Interface: Using DataOcean@ISAC on an 84" touch wall

6. Conclusion and Outlook

The evaluation of current industry demands [agi05a] [agi05b] and national/international calls and guidelines [bmb15a] shows, that analysis, organization and evaluation of big pools of data objects are desired and needed, to improve the efficiency in several sectors of production and processing. ISAC@OTH-AW addresses these demands and will enhance the efficiency of production in SME. Besides the usage of well-known and proven algorithms, new types and advancements (e.g. interactive k-Means) will be developed.

Furthermore, current approaches of machine learning workflows are not based on interactive visualizations. ISAC@OTH-AW will implement new algorithms which will be based on visual analytics of the given data. This can be a promising approach.

The project started recently at the end of 2015 and will last till the end of 2021. This paper depicts the vision and intent of the project.

7. ACKNOWLEDGMENTS

We thank Hans-Peter Schmidt and Wolfgang Blöchl for their valuable suggestions.

8. REFERENCES

[Agi05a] Studie: Industrie 4.0 im Mittelstand: http://www.agiplan.de/unternehmen/news-und-presseinformationen/detail/studie-industrie-40-im-mittelstand.html

[Agi05b] Industrie 4.0 für den Mittelstand: http://industrie-40-mittelstand.agiplan.de/agiplan-industrie/?page=home

[Bin02a] Bin, W., and Yi Z. (eds.). Evolutionary Computation: CSIM: A document clustering algorithm based on swarm intelligence, in Conf.proc. CEC'02 Honolulu, HI, IEEE, pp.477-482, 2002.

[Bmb15a] Industrie 4.0 – Innovationen für die Produktion von morgen http://www.bmbf.de/pub/Industrie_4.0.pdf

[Cui06a] Cui, X., and Potok T.E. A Distributed Agent Implementation of Multiple Species Flocking Model for Document Partitioning Clustering. CIA 2006, LNAI 4149, pp.124-137, 2006.

[Cui05b] Cui, X., and Potok T.E. Swarm Intelligence Symposium: Document clustering using particle swarm optimization, in Conf.proc. SIS 2005, IEEE, pp.185-191, 2005.

[Cui09c] Cui, X., and Potok T.E. Swarm intelligence in text document clustering, in Handbook of research on text and web mining technologies, 2009.

[Fow12a] Fowler, R. H., Huerta, R. A. (eds.). Visualization and Clustering of Document Collections using a Flock-based Swarm Intelligence Technique, in Proceedings of the International Conference on Modeling, Simulation and Visualization Methods (MSV). The Steering Committee of the World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2012

[Eri11a] Erisoglu, M. and Calis, N. (eds.). A new algorithm for initial cluster centers in k-Means algorithm, in Pattern Recognition Letters 32, ScienceDirect, pp.1701-1705, 2011.

[Esm13a] Esmin, A.A.A. and Coelho, R.A (eds.). A review on particle swarm optimization algorithm and its variants to clustering high-dimensional data, in Artif. Intell. Rev, 44, Springer Science+Business Media Dordrecht, pp.23-45, 2013.

[Hart79a] Hartigan, J.A. and Wong, M.A. A k-Means Clustering Algorithm, in Journal of the Royal Statistical Society, Series C (Applied Statistics) 28, pp.100–108, 1979.

[Joll02a] Jolliffe, I. Principal component analysis. John Wiley & Sons, Ldt, 2002.

[Kuo11a] Kuo, R. J., and Wang, M. J. (eds.). An application of particle swarm optimization algorithm to clustering analysis, Soft Computing, 15(3), pp.533-542, 2011.

[Mal15a] Maltese, J., and Ombuki-Berman, B. (eds.). Co-operative Vector-Evaluated Particle Swarm Optimization for Multi-Objective Optimization, in Conf.proc. Computational Intelligence 2015, IEEE Symposium Series on IEEE, 2015.

[Mei15a] Meiller, D. Diving into the Data Ocean, International Federation for Information Processing 2015, Interact 2015, Part IV, LNCS 9299, pp.465-468, 2015.

[Na15a] Na, Wenbo, (eds.). Research on well production prediction based on improved extreme learning machine, in International Journal of Modelling, Identification and Control, pp.238-247, 2015.

[Sha12a] Shafeeq, A., and Hareesha, K.S. Dynamic Clustering of Data with Modified k-Means Algorithm, International Conference on Information and Computer Networks (CICIN 2012), IPCSIT vol.27, pp.221-225,2012

ISSN 2464-4617 (print)
ISSN 2464-4625 (CD-ROM)

Personalized Sound Zoning for Communication means - user studies and evaluation

Ferdinand Fuhrmann, Clemens Amon, Christina Leitner, Anna Maly and Franz Graf
JOANNEUM RESEARCH Forschungsgesellschaft
Institute for Information and Communication Technologies
Steyrergasse 17
Austria, 8010, Graz
{forename.surname}@joanneum.at

ABSTRACT

We present a new audio user interface for communication means based on automatic sound zoning (ASZ). Highly directive audio reproduction and capturing devices are combined with the output of a newest-generation depth sensor. We particularly use parametric loudspeaker and microphone arrays for the spatial filtering of sound playback and recording. This enables device-less, mobile and ergonomic communication. We evaluated the system using subjective experiments assessing, on the one hand, the immersion properties of the system and, on the other hand, general usability aspects. We could show that users are highly impressed by the capability of such a system and would greatly benefit from the inherent properties.

Keywords

audio interface, sound zoning, device-less communication

1 INTRODUCTION

Unlike visual and haptic interfaces, the audio interface has not gone through radical changes across the last decades. Starting from the telephone, the communication principles and sensor basics have mostly stayed the same; a loudspeaker – as close as possible to the ear of the user – is used for the playback channel while a microphone – as close as possible to the user's mouth – captures the acoustic speech signal. This configuration leads to several drawbacks, outlined in the following.

In office scenarios, different audio emitting and receiving devices often act together, producing distracting noise. Headsets as a solution suffer from ergonomic deficits, especially during prevalent use and long working hours. Moreover, tethered headsets restrict the user's mobility. Further, audio interfaces are usually distributed; a single interface is used for each application (e.g. radio handset, telephone, and internet communication).

In this work we propose a new audio user interface for communication means, applicable in both home and office environments. The aim is to eliminate ergonomic problems while reducing the noise-level, and guarantee greatest mobility in a multi-purpose interface. The presented system combines motion tracking technology with directional audio reproduction and capturing devices. This enables automatic sound zoning (ASZ) where the information about the position of the user is used to automatically steer the sound zones for audio playback and capturing. Figure 1 shows an illustration of the presented interface. The advantages of this approach are twofold; (i) the user can move freely in the room while communicating device-less, (ii) the cross-talk to other persons in the room is reduced.

Figure 1: Illustration of the presented acoustical interface, generating sound zones.

We specifically use the newest generation of Microsoft's Kinect sensor[1] in combination with a

[1] https://dev.windows.com/en-us/kinect/hardware

Figure 2: Hardware prototype of the presented acoustic interface, consisting of the Kinect 2 depth sensor, a 13 channel microphone array, and two parametric arrays mounted on automatic alignment units.

directional sound reproducing system and a microphone array. Figure 2 shows the resulting hardware prototype. By using such techniques in communication systems or gaming applications, the immersion of the user is supported. In a professional environment such a system can be applied to provide cognitive relief, reduce stress, and increase mobility while reducing the overall noise level, or simply improve ergonomic aspects by freeing the user of a worn headset.

The presented audio interface is part of an interaction concept originally designed for modern control centers [Kai14]. The principles and technology can however be translated in several other domains, where audio interaction is indispensable (e.g. home, office, public facilities, production site, etc.). In this paper we describe the main technical components, both hard- and software, and present a thorough assessment of the usability of the audio interface. We thereby conducted subjective experiments to involve the user in the design and evaluation process of the system. In an accompanying work [Fuh16] we also provide objective evaluation results related to the applied technical components.

2 RELATED WORK

Microphone arrays are widely used for teleconferencing. They are often combined with video systems to improve the accuracy of the beam steering for better sound quality [Hua11] or are used in automatic speech recognition systems to improve recognition rates [Asa04]. Other application fields include control centers such as air traffic control [Gul09]. A different approach for audio capturing are microphone domes that have been investigated for the application in multimodal interfaces for crisis management [Sha03].

Directional sound reproduction systems are realized by either loudspeaker arrays or parametric loudspeakers. They have been applied to create an immersive audio environment for single users in desktop applications [Str03] or for gaming [Gan11]. Besides fixed installations, there exist automatically aligned systems that not only steer towards, but are able to follow a user [Kas14]. For multiple users, the concept of sound zones has been developed; each user is provided with a personal sound zone directed at her location [Bet15]. This can be applied in different environments such as shared offices or exhibitions. For communication applications, directional sound reproduction systems are combined with microphone arrays [Hua11, Gul09] to create a hands-free interface without headset.

3 CONCEPT

The presented audio interface consists of an audio reproduction and capturing system. Both components are characterized by a narrow spatial focus, which is automatically steered towards the user's head. Here, the alignment information is provided by a depth sensor (Kinect 2), enabling head-tracking in a triangular area in front of the sensor with an opening angle of 70° and depth to 3.5 m. More information about the sensor's capabilities can be found in [Amo14].

The reproduction system consists of two parametric loudspeaker arrays, emmitting spatially focused sound. By applying the principle of ultrasonic sound reproduction and the coupling of many parallel transducers, the emission area of these loudspeakers can be limited to a very narrow sound beam [Pom09]. More precisely, an inaudible ultrasonic carrier signal is first used to generate a spatially narrow radiation pattern. Modulation of the audio signal onto this ultrasonic carrier causes the generation of differential frequency components between the two signals. If the carrier's level is high enough – typical parametric arrays use sound pressure levels of 120 dB and more – this differential components become audible while the spatial emission pattern is preserved [Gan12]. The audio signal is hence demodulated along the radiation pattern of the ultrasonic carrier.

In our implementation we use two Acouspade ultrasonic[2] speakers. Special ultrasound preamps are used to generate a carrier frequency of 40 kHz and modulate their amplitude with the desired playback signal. The speakers are placed on the desk of the work station in a typical stereo arrangement and are automatically aligned to the respective ear of the user. For this purpose, the speakers are mounted on engine-driven automatic alignment units. We built prototypes of these alignment systems using two stepper motors – one for the azimuth and one for the elevation angle. The motors are run by two motor drivers and a micro controller, processing the control data for steering.

The audio capturing system consists of a microphone array followed by a beamforming algorithm. The beam-

[2] http://www.ultrasonic-audio.com/products/acouspade.html

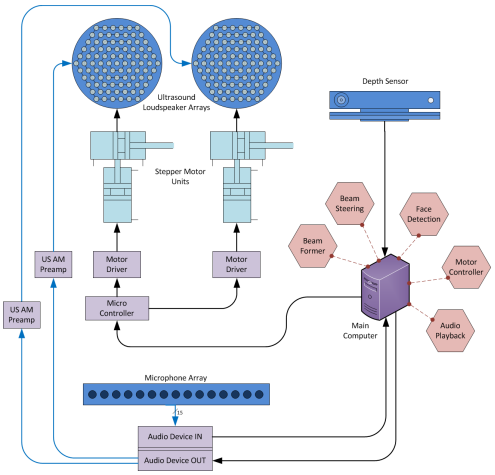


Figure 3: System architecture overview. Hexagonal objects represent software components while rectangular shapes correspond to hardware modules.

former's main directivity is enforced towards the position of the user, enhancing the target speech while suppressing noise and interfering sounds. We constructed a linear array consisting of 13 omnidirectional microphones with an inter-distance of 2 cm. This ensures an effective signal bandwidth up to approximately 8.5 kHz, which is sufficient for the transmission of speech and other acoustical messages (e.g. sound notifications).

Figure 3 illustrates an overview of the system architecture. Here, software components (hexagonal objects) such as head tracking, beamforming, and the motor controller of the loudspeaker alignment system are processed by the main computing unit (PC), which is directly connected via external hardware modules (rectangular objects) with the sensors and actuators.

4 EVALUATION AND RESULTS

4.1 Auditive interference

Here we evaluated the acoustic cross-talk of the parametric arrays in a subjective listening test. The focus of the evaluation was to estimate the cognitive relief with respect to the acoustic information emitted by the proposed system. Overall 24 subjects – 15 male and 9 female – took part in this study, the average age was 32.9.

We considered a two-workplace-scenario where two persons work side-by-side while performing telephone calls or talking via a radio handset. The test subject was seated at one desk, a second person was simulated by an artificial head with torso in front of the second desk. We asked the test subject to read a text (about 500 words) aloud while at the same time an audio signal containing five speech utterances was played back via loudspeakers, which were aligned at the dummy head. After reading the text, we asked the person to write down the utterances she understood. The more utterances the subject could reproduce, the higher the impact of the auditive interference on the cognitive load.

System	*#Utt(mean)*
Presented system	1.792
Reference system	2.792

Table 1: Results of the subjective interference tests. Average number of correctly recognized utterances is shown.

We compared the performance of the subjects within the proposed system against their performance when exposed to a reference system. The reference system was realized by using standard studio loudspeakers together with a band-pass filter imitating the frequency characteristics of a radio handset. Hence, we expected the presented system to exhibit significantly less cognitive load than the reference system.

Table 1 shows the results of the experiment. It can be seen that with the presented system significantly less utterances were recognized than with the reference system (paired t-test, $p < 0.003$). Here, participants could reproduce around 55% more utterances than in the system applying the parametric arrays. This confirms our hypothesis that the presented system supports the cognitive relief of the user, since significant less audio information was transferred to the test subject.

4.2 User acceptance

To evaluate the overall system we performed a user study in a working area where users both are confronted with a lot of acoustical information and work in an office environment. We opted for the traffic monitoring field where users receive audio information via telephone, radio handset, synthetic voice alarms, and other acoustical alerts. Typically, three to five operators work together in specially equipped control rooms, where each operator has a work station desk.

We were able to interview six operators in total (45 mean aged and male), four of them had more than 15 years of work experience. The experimental procedure was arranged into three parts; first, the main functionalities of the presented acoustic interface were demonstrated. We prepared a typical traffic monitoring scenario, where acoustical alerts, telephone calls and radio handset interaction were integrated in a fully automatic sequence of events. The scenario started with an acoustic alarm to grab the attention of the test subject. Then, the operator could accept an incoming call, talk to the calling person, and give a phone call to resolve the situation. In the second part, the operator could explore the acoustic interface in detail. The test person listened to several test signals (radio streams, telephone and radio handset calls), experiencing the immersion properties of the system. Moreover, we asked the operator to move around in the room while listening to the test

signals, judging the quality of the automatic alignment. We the interviewed the operator related to positive and negative aspects of the presented system and the respective components, as well as possible improvements and applications in the traffic monitoring area.

In the interviews all subjects commented positively on the overall impression of the presented system. We were able to transmit the idea of the device-less communication to the operators, who emphasized the integrating and centralizing aspects of the acoustic interface (i.e., all information is transferred over a single user interface). The combination of automatic aligned parametric arrays and the microphone array adds a lot of flexibility and mobility to the daily working routines. All operators were satisfied with the speed of the alignment as well with its mechanical noise properties (very quiet). They were impressed by the narrow sound beam emitted by the parametric arrays; the audio quality was judged between good and excellent.

5 DISCUSSION

The results presented in Section 4 suggest that the acoustic user interface concept – visually guided sound zoning – can be translated into a system implementation. The parametric arrays allow personalized audio playback for multiple users, since the emitted audio signal is only perceived by the target user and all other audible signals in the room originate solely from reflections. This is also shown in the experiments involving the acoustic crosstalk, which can be highly reduced with parametric arrays compared to regular loudspeaker systems. In combination with the microphone array, a device-less communication system is created offering maximum mobility and ergonomy.

The presented system offers a great potential for new applications in both home consumer and office environments. Its mobility property frees the user from any wearable device, while the zoning properties immerse the user. Moreover, for multi-user applications, the overall noise level is reduced which directly influences the cognitive load of the involved people.

6 CONCLUSION

We presented a new audio user interface based on a sound zoning approach. We control highly directional sound reproduction and capturing units by the output of a vision system, tracking the user's position and head movement. This enables a maximum mobility together with a maximal immersion of the user. The applications of such a system range from home usage to typical office environments. In our studies we could show that users are highly impressed by the capability of such a system and would greatly benefit from the inherent properties.

7 ACKNOWLEDGEMENTS

This work was funded by Austrian Ministry for Transport, Innovation and Technology in the fusInC[3] project.

8 REFERENCES

[Amo14] Amon, C., Fuhrmann, F., and Graf, F. Evaluation of the spatial resolution accuracy of the face tracking system for kinect for windows v1 and v2, 6th Congress of the Alps Adria Acoustics Association, 2014.

[Asa04] Asano, F., Yamamoto, K., Hara, I., Ogata, J., Yoshimura, T., Motomura, Y., Ichimura, N. and Asoh, H. Detection and separation of speech event using audio and video information fusion and its application to robust speech interface. EURASIP Journal on Applied Signal Processing, 2004.

[Bet15] Betlehem, T., Zhang, W., Poletti, M. , and Abhayapala, T. Personal Sound Zones: Delivering interface-free audio to multiple listeners, Signal Processing Magazine, IEEE, 32(2), 2015.

[Fuh16] Fuhrmann, F., Amon, C., Leitner, C., Maly, A. and Graf, F. Personalized sound zoning for communication means - design and evaluation of a bidirectional beamforming system, submitted to DAFx, 2016.

[Gan11] Gan, W., Tan, E. and Kuo, S. Audio projection. IEEE Signal Processing Magazine, 28(1), 2011.

[Gan12] Gan, W., Yang, J. and Kamakura, T. A review of parametric acoustic array in air. Applied Acoustics, 73(12), 2012.

[Gul09] Guldenschuh, M. Transaural beamforming. M.Sc. thesis, Graz University of Technology, 2009.

[Hua11] Huang, Y., Chen, J. and Benesty, J. Immersive audio schemes. IEEE Signal Processing Magazine, 28(1), 2011.

[Kai14] Kaiser, R. and Fuhrmann, F. Multimodal interaction for future control centers: interaction concept and implementation, Workshop on Roadmapping the Future of Multimodal Interaction Research, ACM, 2014.

[Kas14] Kashiwase, S. and Kondo, K. Towards a parametric speaker system with human head tracking beam control, 3rd IEEE Global Conference on Consumer Electronics (GCCE), 2014.

[Pom09] Pompei, J. Fundamental Limitations of Loudspeaker Directivity, `https://www.holosonics.com/tech_directivity.html`, 2009.

[Sha03] Sharma, R., Yeasin, M., Krahnstoever, N., Rauschert, I., Cai, G., Brewer, I., MacEachren, A. and Sengupta, K. Speech-gesture driven multimodal interfaces for crisis management, Proc. of the IEEE, 91(9), 2003.

[Str03] Strauss, M., Sontacchi, A., Noisternig, M. and Höldrich, R. A spatial audio interface for desktop applications, 24th International Audio Engineering Society Conference, 2003.

[3] `http://www.joanneum.at/digital/referenzprojekte/fusinc.html`

Image Spatial Verification using Segment Intersection of Interest Points

Marisa Bernabeu, Antonio Pertusa and Antonio-Javier Gallego
Departamento de Lenguajes y Sistemas Informáticos,
Universidad de Alicante, San Vicente del Raspeig (Alicante) - 03690, Spain
{mbernabeu, pertusa, jgallego}@dlsi.ua.es

ABSTRACT

This work presents a new spatial verification technique for image similarity search. The proposed algorithm evaluates the geometry of the detected local keypoints by building segments connecting pairs of points and analyzing their intersections in a 2D plane. We show that these intersections remain constant with respect to different geometric transformations (translation, rotation, similarity and affine). Evaluation has been performed obtaining an initial image similarity ranking with a BOF-based methodology, and then using the proposed method for reranking. The presented algorithm (SIIP) has been compared to the RANSAC spatial verification method, showing that it is more efficient and obtains a higher performance on three different datasets.

1 INTRODUCTION

Image similarity search methods aim to obtain a ranking of the most similar images given a query. In general, the goal is to get fast algorithms that are robust to scale, rotation, noise and illumination. A classical methodology to face this task is to detect *interest point* descriptors such as SIFT [6] or SURF [7] from the query image, and match them to those of the images in the dataset. Such mappings can be computed from correspondences of salient regions between the candidate and query images. In some studies these descriptors are clustered into a *bag of features* (BOF) to increase the efficiency for similarity search. One such approach is TOP-SURF [4] which groups the visual features into a histogram obtained by selecting the highest scored visual words (the top T visual words).

However, BOF models lack spatial information. In order to consider it, spatial verification methods can be added to rerank the BOF results to improve the performance with respect to basic representations. This topic have been studied in [2], both in the transformations and in the estimations required to perform the matching.

Several spatial verification methods can be found in the literature. A well known technique is RANSAC [1], which estimates a transformation between a query and a prototype image based on how well its feature locations are predicted by the estimated transformation. RANSAC-based spatial matching has been used in several works [8] for image retrieval. A weak spatial verification alternative is proposed in [5], a fast spatial reranking method is used in [11] with a vocabulary tree, and in [10] the spatial information is obtained using a BOF.

In this work we propose a simple and efficient spatial verification algorithm to rerank the results returned from a BOF method. This technique is based on the comparison of the intersections between segments built by pairs of interest points (keypoints) from two images. The proposed approach is compared to RANSAC, obtaining better results with a smaller computational cost.

Next section describes the presented methodology. Evaluation results are detailed in section 3, followed by the conclusions and future work in section 4.

2 PROPOSED METHOD

Existing spatial verification methods can improve the performance over a basic bag of features representation, but they tend to be computationally expensive. We propose a new method based on the intersection of segments between interest points or *keypoints* that improves the performance of RANSAC with a smaller computational burden.

In a preprocessing stage, given an image query we get the most similar prototypes from a dataset using TOP-SURF [4]. This algorithm extracts the keypoints from the query image, clusters them into a BOF and compares them to the dataset of prototypes, yielding a ranking of the most similar prototypes. Then, we rerank the top K results from this stage with the proposed spatial verification algorithm.

The proposed spatial verification method consists of extracting the keypoints given a descriptor, match them and evaluate the intersections between segments. Figure 1 shows an example to explain the motivation of the

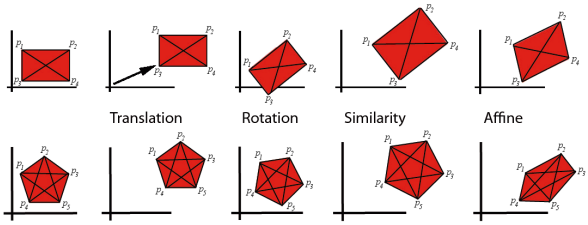


Figure 1: The intersections of segments between interest points are invariant to common transformations.

proposed approach. Given a set P_a of keypoints in an image a, we get all the possible segments and calculate their intersections. As it can be seen, the segment intersections remain invariant to common geometric transformations (scale, translation, rotation, affine, and similarity). In the example of the first row, the segment $\overline{p_1p_4}$ only intersects with the segment $\overline{p_2p_3}$ in all the transformations of the same image.

Using the number of matching intersections as a measure of distance between two images is an efficient method which is robust to 2D transformations since the number of intersection remains constant. The proof of this is trivial since the intersection points belong to the segments, thus the transformation applied to the segments will also be applied to these points.

2.1 Notation and definitions

Consider a given set $P = \{p_1, p_2, \ldots, p_n\}$ of keypoints, where P_a denotes the set of points from the image a, and $\overline{p_1p_2}$ the segment formed by points p_1 and p_2. Let's also denote $\mathscr{P}_2(P)$ as the power set of P with cardinality 2 which contains all the subsets of P with two elements, or in other words, the set of all possible segments between two points in the set P.

The intersections of segments in a set $\mathscr{P}_2(P)$ can be calculated using different methods. In this study, we use an efficient algorithm for line segment intersection from Chazelle et al [12] that has a complexity of $\mathscr{O}(n\log n + k)$, where n is the number of segments in the set, and k the number of intersections.

Given a set of segments $S_a = \{s_1, s_2, \ldots, s_n\}$ belonging to an image a, we define $I = s_1 \cap s_2$ as the intersection function between s_1 and s_2. This function returns whether two segments intersect or not ($\emptyset$). Let also denote I_a as the set containing all the pairs of segments which intersect each other:

$$I_a = \{(s_n, s_m) : s_n, s_m \in S_a, s_n \cap s_m \neq \emptyset, n \neq m\}$$

Spatial verification is performed by comparing the set of intersections. We define the distance $d(a,b)$ between two images a and b as the number of common intersections divided by the maximum number of intersections from both images:

$$d(a,b) = 1 - (|I_a \cap I_b|)/(\max(|I_a|, |I_b|))$$

2.2 Algorithm

Data: Images a, b
Result: Distance $d_{a,b}$
$P_a = \text{SURF}(a)$; $P_b = \text{SURF}(b)$;
$M_{a,b} = \max_N\{(p_a, p_b) : p_a \in P_a, p_b \in P_b,$
$dist(p_a, p_b) < \varepsilon\}$;
for ***each*** *image* i ***in*** a,b **do**
  $P'_i = \{p_i : p_i \in f_i(M_{a,b})\}$;
  $I_i = \{(s_n, s_m) : s_n, s_m \in \mathscr{P}_2(P'_i), s_n \cap s_m \neq \emptyset, n \neq m\}$;
end
$d(a,b) = 1 - (|I_a \cap I_b|)/(\max(|I_a|, |I_b|))$;
return $d(a,b)$

Algorithm 1: Distance between two images a and b.

Algorithm 1 describes the proposed spatial verification method that calculates the distance between two images a and b. We first extract the sets of keypoints P_a and P_b from the input images. Then, these sets are matched to get the subset of related points $M_{a,b}$ that are common to both images. Matching is performed using the Euclidean distance between the feature vectors of each keypoint, defined as $dist(p_a, p_b)$. For efficiency, these sets are ordered by relevance using the inverse of the Euclidean distance between each pair of matched points in order to keep as a maximum only the first (the most correlated) N points to build segments between them.

It is important to note that adding keypoints makes the number of possible intersections to increase. In general, for a set of n segments, there can be up to n^2 intersections in the worst case. This is the reason to keep as a maximum only the N keypoints that are most similar from both images.

From the initial sets of keypoints P_a and P_b, the algorithm selects the subsets of keypoints P'_a and P'_b that are present on the set of corresponding pairs $M_{a,b}$ in order to keep only the matching points for the next stage. Mathematically, we define a bijective function $f_a : M_{a,b} \to P_a$ which given an element of the matching set $M_{a,b}$ returns the corresponding point of the set P_a. Analogously, we define f_b for the set of keypoints P_b.

Then, the segments between all filtered keypoints P'_a and P'_b are independently built for images a and b respectively. Finally we calculate the intersections between these segments in both images, and we define a distance $d(a,b)$ that takes into account the number of common intersections.

For instance, consider the first and last rectangles from the first row in Figure 1, denoted as image a and image b respectively. The interest points p_1, p_2, p_3 and p_4 from image a are matched to the points p_1, p_2, p_3 and p_4 from image b. The segment intersections in a are $\overline{p_1p_3} \cap \overline{p_2p_4}$, that are common to the intersections $\overline{p_1p_3} \cap \overline{p_2p_4}$ in b. Therefore, as there is only one com-

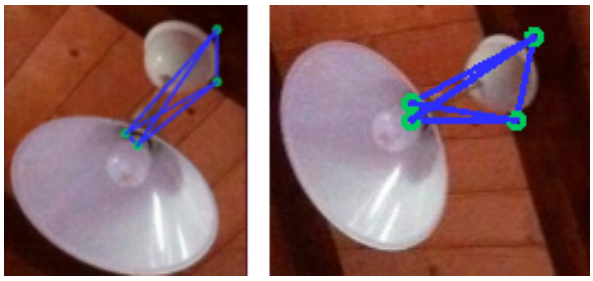


Figure 2: Top interest points and their segments from two images in the MirBot dataset.

mon intersection and each image only contains one intersection, their distance is $d(a,b) = 0$.

Figure 2 shows an example of the segments between matching keypoints of two images. The perspective change is well approximated by a 2D affine transformation and occlusion has little importance.

2.3 Regularization term

The distance $d(a,b)$ was used to perform a fair comparison with RANSAC, which only considers matching keypoints. However, this distance does not take into account the number of unmatched points, which can be a problem in some cases; for example, in Figure 1, if the rectangle is compared to the pentagon, as their geometry is coherent their distance will be 0.

This is the reason to add a regularization term to the original distance in order to consider the proportion between matched and unmatched keypoints. The modified distance $d'(a,b)$ is defined as:

$$d'(a,b) = d(a,b)\frac{|M_{a,b}|}{\max(|P_a|,|P_b|)}$$

3 EVALUATION

This section describes the experimental results. To serve the purpose of comparison, we have evaluated the performance with three datasets commonly used in spatial verification tasks:

Mirbot [3] contains photographs taken by users using smartphones, so they are low-medium quality with variable dimensions (max. 640×640px). The dataset version from October 2014 has been used, with 16327 images distributed in more than 1000 classes.

Oxford 5K [8] contains 5062 high resolution (1024×768px) images divided in 11 classes. Images are assigned to four possible labels: *Good*, *OK*, *Junk* or *Bad*. We have used only the pictures labeled as *Good* or *OK* as positive results, discarding the *Junk* images, and using the *Bad* images as negative results.

Paris [9] contains 6300 high resolution (1024×768px) images collected from Flickr with Paris landmarks. Similarly to Oxford, images are divided in 11 classes. Labels are assigned as in Oxford dataset: *Good, OK, Junk or Bad*. We have also ignored those images labeled as *Junk*.

For the experiments we first use the TOP-SURF method [4] with the default dictionary of 100k words to get a list of the most similar images to a query. Then, only the top K images are taken to be reranked, both using RANSAC and the proposed spatial verification technique for comparison.

In the reranking stage, keypoints could be obtained using any local descriptor. We have chosen SURF [7] to evaluate both RANSAC and the proposed SIIP method.

Performance is assessed using leaving-one-out cross-validation with values of K between 20 and 50. Then the accuracy and the Mean Average Precision at k (MAP@k) are computed. Accuracy (Top-1) measures the ratio between true positives TP and the number of images in the dataset Q:

$$\text{Acc} = \frac{1}{|Q|}\sum_{q\in Q}\text{TP}(q)$$

To calculate the Mean Average Precision at k (MAP@k) we first calculate the Average Precision at k (AP@k) for a query q, and then the MAP is obtained as the mean of the APs for all queries:

$$\text{AP@}k(q) = \frac{1}{N_R}\sum_{n=1}^{N_R} P_k(q)$$

$$\text{MAP@}k(Q) = \frac{1}{|Q|}\sum_{q\in Q}\text{AP@}k(q)$$

where N_R is the minimum between k and the total number of retrieved results, and $P_k(q)$ is the precision at cut-off k in the results list.

3.1 Results

The table 1 shows the results after reranking the $K = 20$ most similar images from the Mirbot dataset with RANSAC and the proposed method (Segment Intersection of Interes Points, SIIP). Different values of N were evaluated to measure the impact on the performance of the maximum number of keypoints, but changing N does not alter significantly the accuracy neither in RANSAC nor with SIIP with the distance function d. A value $N = 24$ has been chosen in the following as it obtains the best MAP@10 and a good accuracy.

The table 2 shows the results obtained in the Mirbot, Oxford 5K and Paris datasets. The accuracy using only TOP-SURF without rerank is shown as the baseline. Then, RANSAC and SIIP (both using SURF features) accuracy and MAP@10 are obtained with $N = 24$. SIIP results are given with the distance function d and also adding the regularization term d'. As it can be seen, SIIP outperforms RANSAC in all the experiments.

		N=8	N=16	N=24	N=32
SIIP (d)	Accuracy	0.3140	0.3148	0.3149	**0.3150**
	MAP@10	0.1782	0.1786	**0.1786**	0.1781
RANSAC	Accuracy	0.3018	0.3017	0.3022	0.3023
	MAP@10	0.1770	0.1774	0.1776	0.1776

Table 1: Results in the MirBot dataset reranking the $K = 20$ first images varying the number of keypoints N.

		Accuracy			MAP@10		
DB	K	RANSAC	SIIP (d)	SIIP (d')	RANSAC	SIIP (d)	SIIP (d')
[3]	20	0.302	0.315	0.305	0.178	**0.179**	0.177
	30	0.304	0.319	0.306	0.177	0.177	0.175
	40	0.305	**0.322**	0.305	0.176	0.176	0.174
	50	0.304	0.319	0.305	0.176	0.175	0.173
[8]	20	0.920	0.923	0.928	0.823	0.826	0.830
	30	0.922	0.925	0.931	0.826	0.829	0.835
	40	0.920	0.925	0.933	0.827	0.831	0.836
	50	0.917	0.925	**0.933**	0.825	0.830	**0.837**
[9]	20	0.796	0.806	0.812	0.621	0.631	0.633
	30	0.800	0.811	0.817	0.629	0.638	0.644
	40	0.803	0.814	0.826	0.631	0.640	0.649
	50	0.803	0.816	**0.828**	0.631	0.643	**0.652**

Table 2: Accuracy and MAP@10 with MirBot [3], Oxford 5K [8] and Paris [9] datasets. The baseline (only TOP-SURF) accuracy without reranking is 0.247 in MirBot, 0.896 in Oxford 5K, and 0.744 in Paris.

In MirBot the function d' did not improve the results of d, but in the other datasets it was consistently better. Contrary to Paris and Oxford, in MirBot each class contains different objects of the same type, instead of the same object from different perspectives. This fact, along with the very large number of classes, explains these accuracy differences.

Besides each image had a very different runtime that depends on the number of keypoints and intersections, SIIP is consistently 3 times faster than RANSAC for all the images, independently of the dataset.

4 CONCLUSIONS

This work presents a new algorithm (SIIP) for image spatial verification based on the comparison of the intersections between segments built by pairs of interest points from two images. It can be used for reranking a subset of images already ranked with a similarity search algorithm such as a BOF.

Evaluation has been performed by obtaining the most K similar images with TOP-SURF, and then reranking the filtered prototypes using SURF interest points. SIIP has been compared to RANSAC, given both superior efficiency (3 times faster) and performance on the three datasets evaluated: Oxford 5K, Paris and MirBot.

As a future work, the proposed methodology could be applied with other descriptors such as SIFT [6], different equations for the distance $d(a,b)$ could be explored, and a comparison with more recent spatial verification methods should be made.

Acknowledgment

This work is supported by the TIMUL project (TIN2013-48152-C2-1-R) and the Instituto Universitario de Investigación en Informática (IUII) from the Universidad de Alicante.

5 REFERENCES

[1] M.A. Fischler and R.C. Bolles. Random sample consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Communications of the ACM, Vol. 24(6):381–395, 1981.

[2] R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, 2nd edition, 2004.

[3] A. Pertusa, A.J. Gallego, and M. Bernabeu. MirBot: A multimodal interactive image retrieval system. Lecture Notes in Computer Science, 7887:197–204, 2013.

[4] B. Thomee, E.M. Bakker, and M.S. Lew. TOP-SURF: a visual words toolkit. In 18th ACM Int. Conf. on Multimedia, pp. 1473–1476, 2010.

[5] H. Jégou, M. Douze, and C. Schmid. Improving bag-of-features for large scale image search. Int. Journal of Computer Vision, 87(3), 2010.

[6] D.G. Lowe. Distinctive image features from scale-invariant keypoints. Int. Journal of Computer Vision, 60(2), 2004.

[7] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-Up Robust Features (SURF). Computer Vision and Image Understanding, 110(3):346–359, 2008.

[8] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Object retrieval with large vocabularies and fast spatial matching. IEEE Conf. on Computer Vision and Pattern Recognition, 2007.

[9] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Lost in Quantization: Improving Particular Object Retrieval in Large Scale Image Databases. IEEE Conf. on Computer Vision and Pattern Recognition, 2008.

[10] S.A. Chatzichristofis, C. Iakovidou, and Y.S. Boutalis. Content Based Image Retrieval Using Visual-Words Distribution Entropy. 5th Int. Conf. MIRAGE, Rocquencourt, pp. 204–215, 2011.

[11] S.S. Tsai, D. Chen, G. Takacs, V. Chandrasekhar, R. Vedantham, R. Grzeszczuk, and B. Girod. Fast geometric re-ranking for image-based retrieval. 17th IEEE Int. Conf. on Image Processing (ICIP), pp. 1029–1032, Sept. 2010.

[12] B. Chazelle and H. Edelsbrunner. An Optimal Algorithm for Intersecting Line Segments in the Plane. Journal of the ACM, 39(1), 1992.

Visualizing astrophysics simulations
New results of the COAST project

Bruno Thooris
On behalf of the COAST team
IRFU, CEA
Université Paris-Saclay,
91191 Gif-sur-Yvette, France
thooris@cea.fr

ABSTRACT
The COAST project at CEA/IRFU at Saclay involves astrophysicists and software engineers developing simulation codes in magneto-hydrodynamics and generic tools for data structuration and visualization. Thanks to the new generation of massively parallel mainframes, computing in astrophysics had made huge progress, generating more complexity and much bigger size of numerical simulations. Two software tools have been developed at IRFU for the visualization of massive amounts of data produced by these simulation codes. SDvision is a code deployed in the framework of IDL Object Graphics, to process all results from different simulation codes developed by astrophysicists. PyMSES is a set of Python modules especially optimized for visualizing the complex data produced by our AMR simulation code RAMSES. These codes are suitable for interactive and immersive navigation for the analysis of 3D results and also for videos and stereoscopic movies productions for people at large. In this paper, we present visualizations of recent numerical simulations in astrophysics, more specifically in the domain of Interstellar Medium and Galaxies Formation.

Keywords
Astrophysics, visualization

1. INTRODUCTION
The COAST (for COmputational ASTrophysics) project [Coa00a] is a program of massively parallel numerical simulations in astrophysics, mixing astrophysicists and software engineers of CEA/IRFU institute. The COAST team is developing 3D magneto-hydrodynamics codes suitable for studying different scales of the Universe. The scientific objective is the understanding of the formation of structures in the Universe, including the study of large-scale cosmological structures and galaxy formation.

Due to the complexity, the geometry or the size of the simulations, the codes are using different numerical techniques, regular Cartesian meshes or structures such as Adaptive Mesh Refinement, spherical coordinates or multi-meshes embedded in the geometry.

The post-treatment software, and in particular the visualizing software tools, must fulfil all these requirements, so two general visualization codes have been developed inside the COAST team: the SDvision code [Pom08b] [SDv00a] , a general tool based on the IDL framework [IDL00a] and PyMSES [Pym00a], a set of Python modules dedicated to the reading and the analysis of our RAMSES AMR code data [Ram00a]. These codes are presented in this paper, with recent examples of use in the domain of interstellar medium and galaxies formation.

2. THE COAST PROJECT
The CEA/IRFU project COAST or Computational Astrophysics [Tho09a] was created in 2005 and includes all the modeling teams of Astrophysics Department; it aims to develop intensive numerical simulations in astrophysics and the generated data post-treatment.
Thanks to the complementarity of its teams, the COAST project will actually complete coverage of astrophysical topics ranging from the formation and evolution of galaxies up to those stars and planets. The COAST project already has to his credit numerous scientific achievements, but also

algorithmic knowledge dissemination. Multiple codes were in fact developed in COAST. This is for example the case of RAMSES code now widely used for the study of various astrophysical themes in the world.
The ASH, HERACLES and FARGO codes that allow us to model stellar streams, respectively radiative flows and the protoplanetary accretion disks are other examples of codes developed by the COAST teams. In addition, major efforts have been made for viewing and disseminating the results to the public. Visualization software was developed and allowed the realization of many films from the simulation results.
Specifically, films combining 3-D simulations ranging from the formation of galaxies at large scales to stellar magnetism and retracing the major stages of structure formation in the universe were produced and widely disseminated.

3. VISUALIZATION OF COAST DATA

The interactive visualization of the complex and massive three-dimensional simulations data is a challenge; the goal is to be able to visualize all data produced by different codes with similarities (written in Fortran90 and parallelized with MPI and OpenMP) and differences (regular Cartesian meshes, spherical coordinates, Adaptive Mesh Refinement structures or multi-meshes embedded in the geometry). The choice was to start from scratch, writing our own visualizing tools [Pom08a], instead of using existing codes like Paraview[Par00a] or Visit [Vis00a]. We start developing the code SDvision in the framework of the IDL Objects Graphics language, about 100000 lines written now, in order to analyze results with the physicists and provide videos in 2D or 3D stereoscopic. Objects including different data can be displayed in the same image, visualizing scalar fields (e.g. density or temperature), vector fields (e.g. velocity or magnetic field) and point clouds (e.g. dark matter). The graphics objects can be coupled with GLSL (OpenGL Shading Language) shaders algorithms to perform scientific computation and visualization by graphics card for time saving. The favored data structure for scalar and vector fields is the uniform grid. More sophisticated data structures such as AMR (Adaptive Mesh Refinement) are handled by projection onto uniform grid. The PyMSES python modules has been developed with the specific objective to visualize the AMR (Adaptive Mesh Refinement) data structures produced by the RAMSES MagnetoHydroDynamics and simulation code, used e.g. for cosmology . It offers two highly efficient CPU-intensive volume rendering algorithms: a ray-casting volume reconstruction based on direct propagation of the rays through the AMR structure, and a splatting algorithm, based on the application of a predefined texture on each cell of the AMR tree. This visualization tool offers both an interactive mode and a batch mode to produce sequence of images for videos. Our visualization tools have been also used for the treatment of simulated data in other domains of physics, such as visualization of turbulences in the ITER plasma [Gra07a] or particles beam movement in the IFMIF accelerator [Tho14a]. Using the very same algorithms and expertise gained in the context of HPC simulations, the SDvision code was also developed in order to visualize observational astrophysical data, specially catalogs of peculiar velocities of galaxies used as input to a Wiener Filter reconstruction algorithm [Cou13a]. This algorithm produces as output the three-dimensional velocity field of gravitational origin in which matter (galaxies and dark matter) is flowing. This study led to the identification of the nature and extent of our Home supercluster Laniakea [Tul14a].

4. THE AMR CODE RAMSES

RAMSES [Tey02a] was developed in Saclay to study large scale structures and galaxies formation. It is now a rather flexible package to be used for general purpose simulations in self-gravitating fluid dynamics. It is written in Fortran 90 with extensive use of the MPI library. It is a free software for non-commercial use only. This code is a grid-based hydro solver with adaptive mesh refinement. As opposed to "patch-based AMR", cells are refined on a cell by cell basis: it is therefore called a "tree-based AMR". A very simple interface can be used to specify runtime parameters. A few routines can be modified to set more complex initial or boundary conditions.

RAMSES calculates the interaction of Dark Matter with baryonic gas, resolving the equations of magneto-hydrodynamics and gravitation. The code is used worldwide by the community of astrophysicists.

5. THE PYMSES CODE

The PyMSES code was written in the last five years to provide a quick and easy way of getting the data out of RAMSES simulation outputs. It is a set of Python modules which helps users analyse and manipulate very large simulations, without worrying more than needed about domain decompositions, memory management, etc… PyMSES is interfaced with a lot of powerful Python libraries (Numpy/Scipy, Matplotlib, PIL, HDF5/PyTables) and can be used with existing users code, as a post-

processing toolbox for data analysis. Optimization comes from a parallelized reading of the AMR data and the import of a sub-set of data corresponding to the restricted domain to visualize.

PyMSES is not an interactive environment but it provides an AMRViewer graphical user interface (GUI) module to look quickly at the AMR data. (Figure 1)

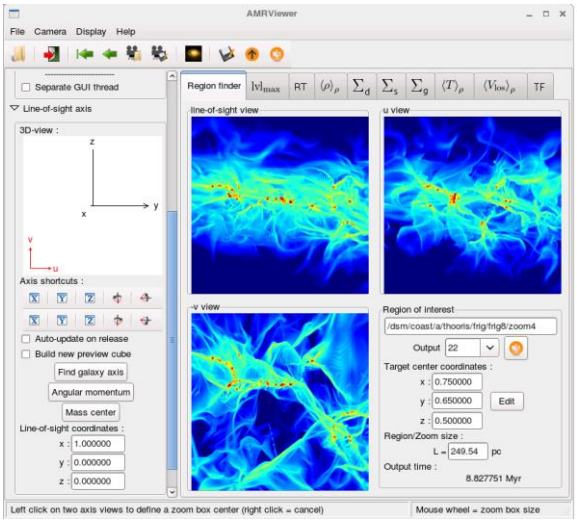


Figure 1. A view of the AMR viewer, graphical module interface of PyMSES, allowing a quick and easy look to the AMR data

6. TWO EXAMPLES OF VISUALIZATION WITH PYMSES

Visualization of Interstellar Medium

Understanding star formation is one of the major challenges of modern astrophysics. Determinant aspects of star formation haven't been yet fully unveiled because star formation is governed by the interstellar cycle, which entails: a huge range of spatial and temporal scales and a broad diversity of coupled physical phenomena such as turbulence, magnetic field, gravity and radiation.

A very recent simulation [Hen12a] has been performed on Curie, the 80000 cores, 1.5 Pflops, Hybrid CPU/GPU PRACE supercomputer at CEA computing center, using 15 million hours calculation. This simulation is including supernovae explosions, which drive the interstellar turbulence and stratification by gravity. In these simulations, dense structures develop and collapse under the influence of large scale turbulence and gravity. (Fig. 2 and 3)

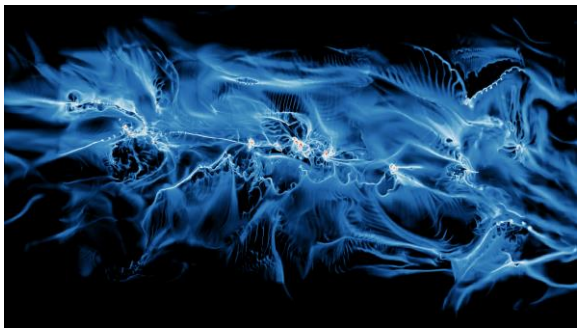


Figure 2. A visualization of a RAMSES simulation of turbulences in the Interstellar Medium; data are coming from the level 12 of the AMR mesh

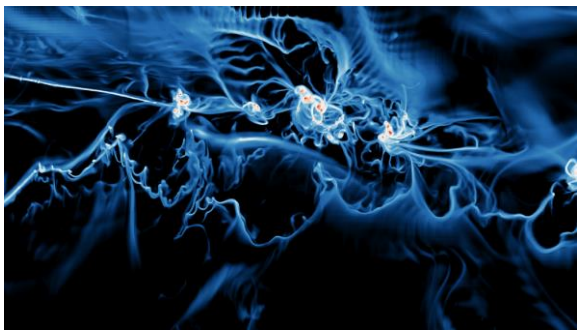


Figure 3. Same simulation as in Fig. 2; zoom using the level 15 of the AMR mesh

Visualization of Galaxies

Another very recent simulation [Roo15a] performed on the Curie supercomputer, realized in 11 million hours of computation and using the RAMSES AMR code is dedicated to star-forming galaxy studies; these are high-resolution simulations[Bou10a] [Cha10a] of star-forming disk galaxies including stellar and AGN (Active Galactic Nucleus) feedback, and it studies particularly the physical origins galactic outflows. The main novelty of this project is to account for all kinds and sources of feedback listed above simultaneously, and follow the galaxy evolution at high resolution over an extended period of time, to really understand the physical origins of galactic outflows and what their main driving mechanisms are. (Fig. 4 and 5)

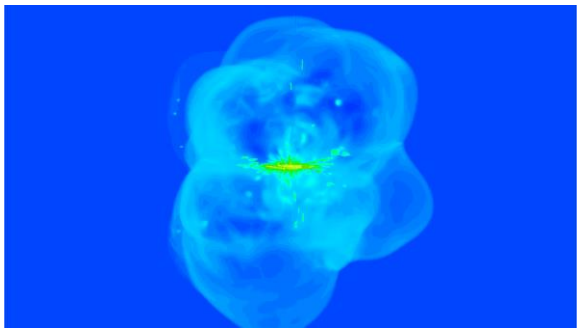


Figure 4. A visualization of a RAMSES simulation of star-forming galaxy; data are coming from the level 11 of the AMR mesh

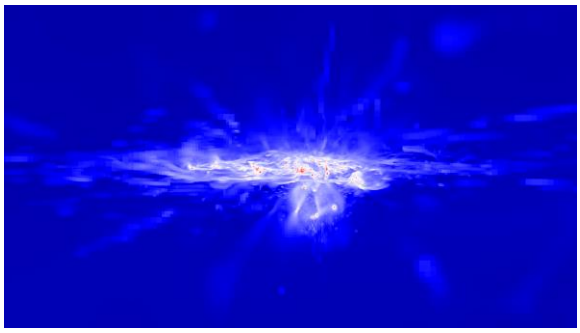


Figure 5. Zoom in a visualization of a RAMSES simulation of star-forming galaxy; data are coming from the level 18 of the AMR mesh

7. CONCLUSION

Thanks to the IDL widget SDvision developed at Saclay, the COAST team could analyze all the data provided by the astrophysicists simulation codes, using volume rendering or isosurfaces techniques. Due to the bigger and bigger amount of data coming from the latest simulations on the PRACE supercomputers in Europe, and the complexity of the data structures in an octree-based AMR code, the development of the PyMSES code was also a very useful work. Even if PyMSES has restricted functionalities (volume rendering only), his powerful capacity of AMR data reading with PYTHON modules provide us a high efficiency to analyze the results and achieve their visualization at the end of calculations. Images, videos and stereoscopic movies are also used to show our laboratory research results to the outside world; videos generated by the PyMSES code have been screened in exhibitions for the general public.

8. REFERENCES

[Bou10a] Bournaud, F., et al. ,Hydrodynamics of high-redshift galaxy collisions: From gas-rich disks to dispersion-dominated mergers and compact spheroids, *eprint arXiv:1006.4782, June 2010*

[Cha10a] Chapon, D., Teyssier, R., Bournaud, F., The Driving Mechanism of Starbusts in Galaxy Mergers. *The Astrophysical Journal Letters*, Vol. 720, Issue 2, pp. L149-L154(2010)

[Coa00a] http://irfu.cea.fr/projets/COAST

[Cou13a] Courtois, H., Pomarède, D., Tully, R.B., Hoffman, Y., Courtois, D., Cosmography of the Local Universe, *The Astronomical Journal, 146 (2013) 69.*

[Gra07a] Grandgirard, V., et al., Global full-f gyrokinetic simulations of plasma turbulence. *Plasma Physics and Controlled Fusion*, 49(12B) :B173, 2007

[Hen12a] Hennebelle, P., Formation of proto-clusters and star formation within clusters: apparent universality of the initial mass function, *Astronomy & Astrophysics, Volume 545, id.A147, 15 pp*

[IDL00a] *IDL* http://www.ittvis.com/idl

[Par00a] *PARAVIEW* http://www.paraview.org/

[Pom08a] Pomarède, D., Fidaali, Y., Teyssier, R., Visualization of Hybrid, N-body and Octree-based Adaptive Mesh Resolution Parallelized Simulations, *Proceedings of 5th International Conference 'Computer Graphics, Imaging and Visualization' (august 25-28 2008, Penang, Malaysia)* ISBN: 978-0-7695-3359-9, p. 295.

[Pom08b] Pomarède, D., Fidaali, Y., Audit, E., Brun, A.S., Masset, F., Teyssier, R., Interactive visualization of astrophysical plasma simulations with SDvision, *Proceedings of the IGPP/DAPNIA International Conference on Numerical Modeling of Space Plasma Flows, ASTRONUM2007, Paris, France, June 11-15, 2008, ASPCS 386 (2008) 327*

[Pym00a] http://irfu.cea.fr/Projets/PYMSES/intro.html

[Ram00a] http://www.itp.uzh.ch/~teyssier/ramses/RAMSES.html

[Roo15a] Roos, O., et al., Thermal and radiative active galactic nucleus feedback have a limited impact on star formation in high-redshift galaxies, *ApJ 800 19, arXiv:1405.7971*

[SDv00a] http://irfu.cea.fr/Projets/COAST/SDvision.html

[Tey02a] Cosmological hydrodynamics with adaptive mesh refinement - A new high resolution code called RAMSES, Teyssier, R., *Astron. Astrophys., 385 (2002) 337-364*

[Tho09a] Thooris, B. et al. , The COAST Project: a Program of Massively Parallel Numerical Astrophysics Simulations, *Proceedings of the HPC-Asia 2009 conference (march 2-5, Kaohsiung, Taiwan)*

[Tho14a]Thooris, B., Nghiem, P.A.P., Pomarède, D., Visualization of Particle Beam Simulations in the IFMIF-EVEDA accelerator, *SC14 supercomputing conference, New-Orleans, 17-20 November 2014*

[Tul14a] Tully, R.B., Courtois, H., Hoffman, Y., Pomarède, D., Laniakea: Our Supercluster of Galaxies, *Nature, 513, 71-73 (4th September 2014)*

[Vis00a] *VISIT* https://wci.llnl.gov/codes/visit

ISSN 2464-4617 (print) WSCG 2016 - 24th Conference on Computer Graphics, Visualization and Computer Vision 2016
ISSN 2464-4625 (CD-ROM)

Reachable Sets for Dubins Car in Control Problems: Physical Visualization

Starodubtsev I.S.
IMM UB RAS[1]
Russia, Ekaterinburg
starodubtsevis@imm.uran.ru

Fedotov A.A.
IMM UB RAS[1]
Russia, Ekaterinburg
andreyfedotov@mail.ru

Averbukh V.L.
IMM UB RAS[1]
Russia, Ekaterinburg
averbukh@imm.uran.ru

Patsko V.S.
IMM UB RAS[1]
Russia, Ekaterinburg
patsko@imm.uran.ru

ABSTRACT

The work deals with an application of the 3D-printing to full-size building the reachable sets in control problems. As an example, a simple car model is considered with nonlinear dynamics, three-dimensional phase vector, and scalar control constrained by modulus (Dubins car). Current state of the sytem includes its position in the plane and the velocity heading. The velocity value is given to be constant. The reachable sets are considered "at the terminal instant" and "till the terminal instant". These sets are nonconvex and their boundaries are not smooth in the whole. Peculiarities of the sets can be better comprehenced being represented in the form of three-dimensional bodies that are built with using the 3D-printing. The sets' boundaries are presented in the VRML and STL formats. Examples of computation and visualization are given including their 3D-printed copies. Problems are discussed that appear under the 3D-constructing of surfaces with complicated forms.

Keywords

control problems, reachable sets, 3D-printing, FDM-printing, physical visualization

1 INTRODUCTION

Let a control system dynamics be described by a stationary vector differential equation $\dot{z} = f(z,u)$ with a geometric constraint $u(t) \in P$, and $z(t;z_0,u(\cdot))$ be a system position at the instant t on a trajectory beginning from the initial point z_0 at the instant $t_0 = 0$. Then *the reachable set at a fixed instant* T is

$$G(T;z_0) = \bigcup_{u(\cdot)} z(T;z_0,u(\cdot)).$$

The join is performed over all *admissible controls* $u(\cdot)$. In many cases, the reachable set is closed in the class of piecewise-continuous controls. So, as admissible controls $u(\cdot)$, we can take piecewise-continuous functions $t \to u(t)$.

The reachable set $G(T;z_0)$ at the fixed instant T comprises of all points, to which the control system can be delivered (by means of admissible controls) at the given instant T from the initial point z_0 given at the instant $t_0 = 0$. If for a stationary system we say about the reachable set $G^*(T;z_0)$ *till the instant* T, then

$$G^*(T;z_0) = \bigcup_{t\in[0,T]} \bigcup_{u(\cdot)} z(t;z_0,u(\cdot)).$$

Here, the join is additionally performed over all instants $t \in [0,T]$.

Appearance of 3D-printers opens opportunity of full-scale visualization of three-dimensional reachable sets. Such three-dimensional bodies performed on the 3D-printer make easy comprehension of geometry of the reachable sets.

An adequate physical visualization gets off the "mysticism" and connected with it "fear" of students and engineers to use results of the mathematical control theory to investigate and solve applied problems.

At the same time, apparency of comprehension can suggest such "reasonable" variants of approximation for three-dimensional objects, for which the most important quality characteristics of the objects are not lost.

2 CONTROL PROBLEM

Let the reachable sets be investigated *at the instant* and *till the instant* for a simple model of car motion (Dubins car [5, 6]). The dynamics is described as follows:

$$\begin{aligned} &\dot{x} = V\cos\varphi, \\ &\dot{y} = V\sin\varphi, \\ &\dot{\varphi} = \frac{k}{V}u, \\ &|u| \le 1, \;\; V = \mathrm{const} > 0, \;\; k = \mathrm{const} > 0, \end{aligned} \tag{1}$$

where x, y are coordinates of the geometric position, φ is the heading of the velocity vector (Fig. 1), V is

[1] Krasovskii Institute of Mathematics and Mechanics, Ekaterinburg, Russia

the velocity value, k is the maximal value of the lateral acceleration. Admissible controls $u(\cdot)$ are piecewise-continuous functions of time that satisfy the constraint $|u(t)| \leq 1$. Values of the heading angle φ are considered in the interval $(-\infty, +\infty)$.

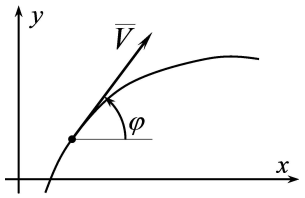


Figure 1: System of coordinates.

The system phase vector (x, y, φ) (1) is denoted as z. To be short, assume $\alpha = k/V$. Peculiarity of system (1) is in the fact that the initial state z_0 affects onto the reachable set with accuracy up to a turn and transfer. So, assume $z_0 = 0$ and write $G(T)$, $G^*(T)$ instead of $G(T; z_0)$, $G^*(T; z_0)$.

In paper [2], the form of the reachable sets $G(T)$ was found in projection onto the plane of geometric coordinates x, y. In [7], the reachable sets $G(T)$ were investigated in the three-dimensional space x, y, φ.

In work [3], the reachable sets $G^*(T)$ were investigated. Also, the sets $G(T)$ and $G^*(T)$ were considered under nonsymmetric constraints onto the control u.

3 PREPARATION OF DATA FOR A 3D-PRINTER

To obtain for system (1) computer images of the reachable sets, the `VRML` (v2.0) format was used.

It was proved in [7], that the boundary of the reachable set $G(T)$ under the symmetric constraint $|u| \leq 1$ comprises of six surfaces given in the three-dimensional space and defined by two parameters. These parameters are the instants of the control switches.

One of these surfaces corresponds to the control values $-1, 0, 1$ that are performed in time namely in the shown order. In this case, the parameter t_1 corresponds to the instant of the control switching from the value -1 to 0, and the parameter $t_2 \in [t_1, T]$ corresponds to the control switching from the value 0 to 1.

Other five surfaces correspond to the control collections $1, 0, -1$; $1, 0, 1$; $-1, 0, -1$; $1, -1, 1$; $-1, 1, -1$.

For each of the surfaces mentioned above, a uniform time-step grid is given for forming collection of the switch instants. Instants 0 and T are formally included into this collection. Note that the third equation of system (3) is stationary and linear w.r.t. the control u. So, values $\varphi(T)$ at an instant T also belong to nodes of the uniform grid on φ. Under this, an interval of possible values of φ is $[0, \alpha T]$ for the control version +1, 0, +1; $[-\alpha T, 0]$ for control -1, 0, -1; and $[-\alpha T, \alpha T]$ for all other four versions of control. This allows one to provide layer-after-layer constructing the considered surfaces with a uniform step in φ.

It is important to note that for correct junction of neighboring surfaces, it becomes possible to choose special time steps of different surfaces in such a way that the set in φ will be the same for all set $G(T)$. To do this in practical constructing of the surfaces, the interval $\pm\alpha T$ of admissible values of φ is divided into even number $2m$ of equal segments to include the zero value. For each node in the interval of φ, possible motions of system (1) are calculated on the interval $[0, T]$ from the point z_0 by application of a concrete version of the control changes.

Under this, the step of over-looking the switching instants for surfaces determined by control collections ± 1, 0, ± 1, and ± 1, 0, ∓ 1 coincides with the step of dividing the interval on φ and is equal to $\alpha T/m$. For two other surfaces, the step of over-looking the switching instants is given twice smaller.

It should be noted some peculiarites of surfaces (composing boundary of the reachable set $G(T)$). As a rule the quantity of points from a layer to layer in φ differs not more than on 1. For the outer layer, consisting of a single point, the difference in number of points with the neighboring layer may be more then 1. This is used for correct junction of layers by means of collections of spatial triangle elements (i.e., by triangulation) during process of forming corresponding three-dimensional objects in formats `VRML` and `STL`.

In the case of nonsymmetric constraint on the control u, the difference is in the fact that the general interval of admissible values of φ is not symmetric w.r.t. zero.

In preparation for 3D-printing the sets $G^*(T)$, the data are firstly formed for building the sets $G(T)$. After that to them, the collections of the space-triangle elements are added. Note that these collections have been built during procession of the layers obtained on the grid of instants from 0 till T under marginal controls with only one switch instant as a parameter [3].

Thus, in all cases, a collection of surfaces (containing boundary of the reachable set) is written into a text-file in the `STL` format. The surfaces' parts that are not on the boundary of the reachable set are placed inside the object to be formed. Triangulation of the surfaces is performed "layer-after-layer" with using the property of continuity of cross-sections on φ for the reachable sets at the instant and on t for the additional surface for the reachable sets till the instant.

4 3D-PRINTING THE REACHABLE SETS

Several versions of reachable sets *at the instant* and *till the instant* have been prepared and printed on the 3D-printer (RepRap Prusa i3 style).

Fig. 2 presents the set $G^*(T)$ till the instant $T = 1.5\pi/\alpha$ (left) and the set $G(T)$ at the instant $T = 1.5\pi/\alpha$ (right). Fig. 3 shows the reachable sets $G(T)$ at the instants $T = 2\pi/\alpha$, $3\pi/\alpha$, $4\pi/\alpha$. Nonsimmetric reachable set $G(T)$ for $T = 5\pi/\alpha$ and $u \in [-1, +0.25]$ is shown in Fig. 4.

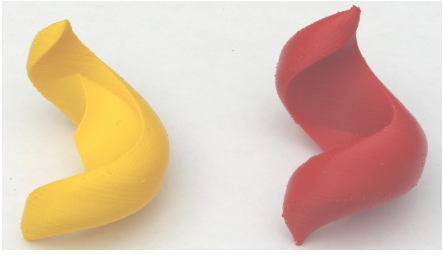


Figure 2: The reachable sets $G^*(T)$ and $G(T)$ for $T = 1.5\pi/\alpha$.

Figure 3: The sets $G(T)$ at $T = 2\pi/\alpha$, $3\pi/\alpha$, $4\pi/\alpha$.

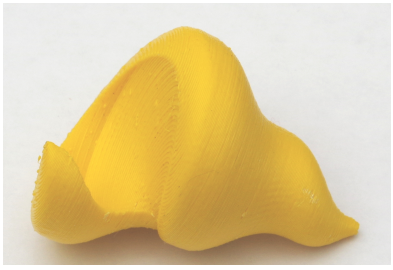


Figure 4: The reachable sets $G(T)$ at $T = 5\pi/\alpha$ in the case $u \in [-1, +0.25]$.

Last years, investigations in physical visualization and, particular, in the 3D-printing, become more and more popular. In work [4], a review is given of application of such technologies in various branches of science and medicine.

Methods of 3D-printing claim more strict demands to a visualized object in comparison with visualization by usual traditional means of computer graphics.

Introduce the following notions:

• *Slicer* is a program for transferring a 3D-model into a controlling code for the printer; this program is necessary since the printer will not be able to load directly the 3D-model; also, there are printers, in which the slicer already is a part of the internal software;

• *Slicing* is a process of transferring the 3D-model into the controlling code.

There exist several types of the 3D-printing that use different technologies and materials. For visualization of the reachable sets, we have used the FDM (Fused Deposition Modeling) printing technology by molten thread (fiber) of the ABS plastic.

The model is cut (sliced) into layers. Each layer are includes the perimeters, infilling, and/or supports. A model can have different level of the infilling, and it can also be without the infilling (a hollow model).

4.1 Peculiarities of the FDM-technology of printing

Below, we describe several facts that might not be classified as "problems". Rather, they can be regarded as peculiarities of the used technology that also should be taking into account for obtaining correct results.

1. Wall's thickness. The walls should be equal to the sprayer's diameter or thicker it. Otherwise, the printer will not be able to correct print them. The wall thickness depends on the sprayer's diameter and number of perimeters to be printed. For example, under three perimeters and sprayer's diameter 0.4 mm, the wall thickness must be 0.4, 0.8, 1.2, 1.6, and 2 mm, but got the larger number the thickness can be any one. So, the wall thickness should be multiple to the sprayer's diameter if it is smaller than $N \cdot d$, where N is the number of perimeters d is the sprayer's diameter.

2. Orientation of the object. The final result of printing can depend on the object orientation. Note the following aspects:

• Since the discrete character of slicing, the horizontal over-layering is imaged on the printed surface; this is the essence of 3D-printing by the FDM technology; to minimize the over-layering, it is possible to change the spatial orientation of the model or to use it to underline the model edges;

• Cylindric edges printed from the side surface can be "lower round" than the similar edges printed in the vertical projection;

• Stiffness/length: thin and direct elements will be more steady if they have been printed rather in the horizontal position than in the vertical one; this is stipulated by the fact that connections between layers are lower strong than inside one layer.

3. Overhanging elements. For each overhanging element (examples are given in Fig. 5), a supporting construction (the support) is necessary; the smaller number of overhanging elements, the smaller number of supports is necessary; so, expenditure of the plastic material and time of printing decrease; moreover, the support deforms the surface touching it.

4. Accuracy. Printing accuracy on axes XY depends on slacks, stiffness of construction, belts, and other details of the printer; for unprofessional printers, the accuracy is about 0.3 mm; accuracy along the axis Z is defined by

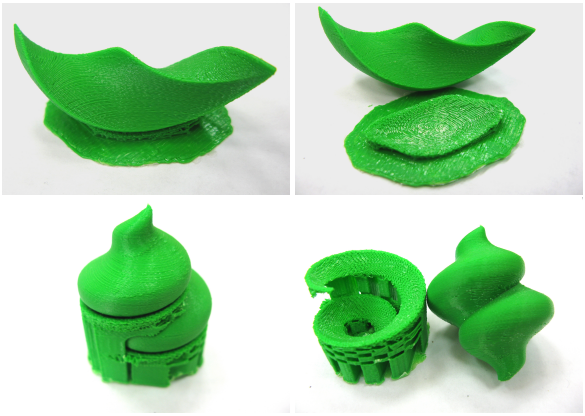


Figure 5: Examples of stable supports during printing for the reachable sets till the instant $T = \pi/\alpha$ (the upper row) and till the instant $T = 4\pi/\alpha$ (the lower row).

the layer height (0.1– 0.4 mm); so, hight of the model will be multiple to the layer height.

There exists and a software problem: not each slicer correctly processes the internal sizes; so, it would be better to increase the holes' diameters on 0.1 – 0.2 mm.

5. Changing the geometry. It is necessary to take into account that after cooling, the material shrinks out; so, the geometry of the printed object changes. On large objects, this can lead to "downturn" of the object's ends and, in some cases, to "detachment" of the printing model from the print table.

Another aspect of the thermo-shrinkage is in appearance of internal stress in the material. If the complete filling of the model was made, then under sufficient square of a cross-section, forces of the internal stress could lead to deforming (up to a break) the printed model. So, a partial filling of model is used in the form of various configurations of the uniform grid. As a result, the following features are provided:

- Compensation of the internal stresses;
- Increasing the model stiffness; the partial filling forms the stiffening ribs that do not allow deforming for the noncooled soft parts of the model under their own weight;
- Decrease of the material expenditure and time of printing.

In printing the reachable sets, we often used configuration of the internal filling in the form of a *square net*. In this case, the minimal value of material is necessary and the printing time is the shortest.

Some of the mentioned peculiarities are taken into account automatically (by the slicer or by the printer software). For example, a formation of the structure of internal filling or calculation of the support elements implemented automatically; but a user check such parameters as "density", "frequency" of the step, and the bevel angle under which it is necessary to built the support.

Other peculiarities have to be processed by-hand, since in the case of ambiguity (a hole in geometry, nonuniformity of the grid, improper normals) the result of procession could be unpredictable.

5 CONCLUSION

The paper describes results of practical calculations of the reachable sets and their reproduction (visualization) by means of 3D-printing for a dynamic object known as "Dubins car".

The sets have nontrivial geometry. Nevertheless, as a result of processing the obtained spatial objects, it was succeeded to achieve not only good quality of the three-dimensional printed images, but, also, to formulate criteria necessary for successful physical visualization.

When preparing this paper, the authors became aware that a close 3D-printing work was done [1] at Mines Paristech / LAAS-CNRS by G. Caner under the supervision by J.-P. Laumond. Namely, reachable sets till the instant were built for the Dubins car, as well as for the Reeds-Shepp car, by means of a 3D-printer. A mathematical model of the Reeds-Shepp car is also widely used in robotics and is described in [5, 6].

6 ACKNOWLEDGMENTS

The work has been partially supported by Program of Presidium RAS "Mathematical problems of modern control theory" and Russian Foundation for Basic Researches under project No.15-01-07909.

7 REFERENCES

[1] Caner G. Modelisation en 3D de la boule de Reeds & Shepp. Rapport de stage d'option, Mines Paristech, LAAS-CNRS, 2015.

[2] Cockayne E. J., Hall G. W. C. Plane motion of a particle subject to curvature constraints. SIAM J. Control. Vol. 13, No. 1, pp. 197–220, 1975.

[3] Fedotov A. A., Patsko V. S., Turova V. L. Reachable sets for simple models of car motion. Ed. by A. V. Topalov. Rijeka: InTech Open Access Publisher, pp. 147–172, 2011.

[4] Horvath J. Mastering 3D Printing (1st edition). Berkely, CA, USA: Apress, 2014.

[5] Laumond J.-P. (ed.) Robot Motion Planning and Control. Lect. Notes in Contr. and Inform. Sci. Vol. 229. Springer, New York, 1998.

[6] Laumond J.-P., Mansard N., Lasserre J.-B. Optimality in robot motion: Optimal versus optimized motion. Communications of the ACM. Vol. 57, No. 9, pp. 82–89, 2014.

[7] Patsko V. S., Pyatko S. G., Fedotov A. A. Three-dimensional reachability set for a nonlinear control system. Journal of Computer and Systems Sciences International. Vol. 42, No. 3, pp. 320–328, 2003.

Validity of Metaphors and Views of Software Visualization for Parallel Computing

Vladimir L. Averbukh[1,2]
[1]IMM UrB RAS,
[2]IMCS UrFU
averbukh@imm.uran.ru

Mikhail O. Bakhterev[1,2]
[1]IMM UrB RAS,
[2]IMCS UrFU
m.bakhterev@imm.uran.ru

Dmitriy V. Manakov
IMM UrB RAS
manakov@imm.uran.ru

ABSTRACT

In this paper approaches to the evaluation of Software Visualization for Parallel Computing are considered on the examples of representation of call graphs and execution traces of parallel programs. The concept of visualization metaphor is described. The visualization metaphors using to depict call graphs and execution traces are surveyed. The validity of visualization techniques is considered on basis of analysis of metaphor properties, Shneiderman's scheme and other approaches to the evaluation of metaphors and views.

Keywords

Software Visualization, Parallel Computing, call graphs, execution traces, visualization metaphors, Shneiderman's scheme

1 INTRODUCTION

Software Visualization (SV) systems were actively developed as late as the 80th and the 90th years of the XX century. Much part of these systems is visual systems for performance tuning and program debugging in the field of parallel computing. But later it can be observed a certain recession in this domain. The reasons of the recession are connected with a number of problems in particular connected with perception, analysis, and interpretation of images depicted huge volumes of data. At the very beginning of Software Visualization evolution, when volumes of data were comparatively inconsiderable, designers used standard "Nodes and Arcs" approaches. However, as early as the 80s more sophisticated views were used. These views were based on one or another metaphors or they were built on the basis of the figurativeness of applications under consideration. In some cases Software Visualization systems were provided with tools of design and drawing so that their users-programmers themselves can develop views for their needs. Ideas of performance tuning are based on the representation of statistics of parallel program executions. For these purposes statistical graphics, first of all, Gantt charts and Kiviat diagrams and their modifications are used. Note that complex views using various modifications of statistical charts to this day are the main most in the "industrial" systems of performance tuning and performance debugging [1], despite the obvious limitations when dealing with real high-performance programs. It appears that in the 90th the designers of visual debugging systems focused on the problems of capture of data in the frameworks (and under restrictions) of the then existing parallel computers. However, other problems exist, for example, the problems associated with visualization – how to choose and how to show entities of parallel programs, as well as to analyze and interpret them. In the case of parallel computation, the very definition of the program entities associated with its "erroneous" states is the tricky problem. The set and the essence of the analyzed software objects strongly depend on used parallel programming paradigm. In the case of performance tuning there is also a lot of problems, because the entities appropriated for analysis are hard to choose. Further in 2000th and 2010th years visualization metaphors were actively used. However use of interesting metaphors was not always clear as users often need the simple picture representation which could be accurately interpreted. In this regard, the important issue is the choice and the evaluation of visualization techniques, the analysis of their applicability to those or other cases. Designers need to evaluate visualization techniques basing both on the qualitative analysis and some formalization. There are various parallel program entities that are subjects of study, analysis and visualization. But in this paper, only the program execution traces and call graphs are considered, although they are not the only software analysis techniques.

2 VISUALIZATION OF EXECUTION TRACES AND CALL GRAPHS

From the beginning of the development of Software Visualization the question of how to graphically represent program entities came up. System designers relied on the standard ("paper") approaches to software visualization (for example control-flow diagrams, etc.) or shifted the task to users-programmers, providing a graphical toolkit. There are serious problems of scal-

ing the views based on one or the other diagram and charts. The next step in Software Visualization may be associated with metaphor using. But using the original metaphor was not always justified because of the need of simple but clearly interpretable presentation methods. "Graph-based" metaphors have significant limitations in this regard also. Execution traces and call graphs are used in one form or another by many debugging systems to describe the dynamics of the programs. Execution traces (also the term "Event Traces" is used) map the dynamics of the certain program executions. Visualization and "replaying" of execution traces are an important element of debugging systems. The visual presentations of the call graph are widely used in the systems parallel program performance tuning systems. In debugging systems realized in the 80s-90s of XX century methods based on charts or diagrams were used. Accordingly, traces are visualized in the form of dynamic "jumping" on the diagrams/charts. There was also the use of "passage" (or "running") on the text of a program highlighting the current position. In the case of high-performance computing, such visualization techniques are hardly suitable. Call Graphs also were represented as very complex and convoluted schemes. Now visualization metaphors were actively used. The metaphor essence consists in interpretation and experience the phenomena of one sort in terms of the phenomena of other sort. Metaphorization is based on interaction structures of source and target domains. During process of metaphorization some objects of target domain are structured on an example of objects of target domain and there is a metaphorical mapping (projection) of one domain onto another. That is the metaphor can be understood as a map from source domain onto target domain, and this map is strongly structured. There are many interesting examples of metaphor using for visualization of execution traces call graphs in parallel programs. For example such metaphors are used as Building metaphor [2], City and Landscape metaphors [3, 4], Dynamic Systems metaphor, [5], Molecule metaphor [2] metaphor. Also less customary metaphors for example Brain Metaphor [6] and Circular Bundle metaphor [7, 8] were suggested.

Views are designed on the basis of the metaphors. A view includes a description of possible visualization objects, their relative positions on the screen, as well as the possible interaction with them. The consideration of specific tasks of debugging and analysis is needed during the phase of the view development. Views are designed on the basis of the metaphors. A view includes a description of possible visualization objects, their relative positions on the screen, as well as the possible interaction with them. The consideration of specific tasks of debugging and analysis is needed during the phase of the view development.

3 PROPERTIES OF VISUALIZATION METAPHORS

The success or failure of debugging and performance tuning systems depends on many factors. Of course, the important factors are the comprehension of correspondence to system specifications and the system reliability. However at the design stage, an important task is the choice of methods of visual representation of objects and entities to be considered during debugging. One approach to the evaluation of visualization involves the examination of properties of visualization metaphors. We analyze the properties to consider the possibility of metaphor using for specific applications of Software Visualization. It is important to understand what objects may be represented with one or another metaphor. We need to analyze the possibility of the visualization metaphors (more precisely – the views based on the visualization metaphors) to represent a large and huge volumes of data and details required to understanding the program's operations. The positive effects of a 3D display and virtual and augmented reality environments are possible in these cases. Therefore it is important to analyze possible applications of metaphors in the frameworks of visualization systems using modern computer graphics environment, in particular the virtual reality environment. For all this, we need to describe how to verify the suitability of metaphor for solving problems under consideration. Note on such metaphor properties as ***"ability to contain any objects inside itself", "restriction of a perception context", "closeness", "inclusion in structure", "presence a structure inside", "naturalness of a metaphor"***. In the cases of popular in Software Visualization systems City Metaphor and similar Landscape Metaphor one may consider the following properties as:

Unlimited context. The user context isn't artificially limited in City Metaphor and Landscape Metaphor. When visualization of large volumes of data is needed, unlimited context allows to have a look-see round the whole picture and to select the key places quickly.

Naturalness. It is known that naturalness of a metaphor reduces efforts on the resultant image interpretation. In the cases of City and Landscape metaphors not only naturalness of spatial orientation, but naturalness of navigation takes place also. In case of a city metaphor the method of navigation is defined by the metaphor itself.

Organization of inner structure. Metaphors suggest the existence of an inner structure. In case of a City metaphor this structure is dictated by the metaphor itself, and it is defined rather rigidly – there are buildings, quarters, streets, districts. In Landscape metaphor a structure choice is nondedicated. In this case one may say about landscape nesting.

Key elements. Metaphors suggest a representation of large volume of information, and in most cases this information is rather homogeneous in visual sense. Users need the key elements to interpret this information. If we want to use a metaphor to reveal specific features and/or exceptions (for example bugs in programs), these elements have to be depicted by easy distinguished image-keys. One may design some key elements in frameworks of City or Landscape metaphors. In these cases some forms of guidance signs or markers may be used as key elements.

Resistance to scaling. These metaphors are stable in the case of increase in information volumes. Moreover, applications of City and Landscape metaphors are reasonable only in the cases of large information volumes.

In the cases of City and Industrial Landscape metaphor *transport corridors* help to design software visualization systems. Transport corridors may be used as means to represent *control flows, data flows*, and other relations between program constructions or parts of program complex.

Note that unlike in the case of Landscape metaphor, the choice of City metaphor strongly limits the set of possible views.

Thus City and Landscape metaphors may form base to represent considerable volumes of the structured information with identifications of specific interest cases that is necessary in the systems for performance tuning and program debugging for parallel computing.

Additionally possibility to fly over a city/landscape creates prerequisites to easy navigation. Flight with changes of height allows to carry out scaling and zooming. Interpretation of the graphical displays based on these metaphors seems to be simple.

4 VIEW EVALUATIONS

A number of papers are devoted to evaluation of views used in Computer Visualization Most of these researches are linked with Information Visualization. The paper [9] contains the outline of visualization analysis based on Visual Information-Seeking Mantra (so called Shneiderman's Mantra). B. Shneiderman presents seven high level user needs that an information visualization application should support [10]. In [11] these needs were refined to evaluate views of three-dimensional information visualization. Let's cite the outline of visualization analysis following [9]: **Overview; Zoom; Filter; Details-on-demand; Relate; History; Extract**.

It is supposed that if the system supports this set of operations, it may be used for Information and Software Visualization. In [7, 8] the criterion based on an expanded Schneiderman's mantra is applied to the analysis of visualization of execution traces constructed on the basis of two synchronized views 1) a circular bundle view for displaying the structural elements and bundling their call relationships, and 2) a massive sequence view that provides an interactive overview. The summary table of how the two synchronized views satisfy each of these seven criteria is provided. In general these views correspond to the chosen criterion.

Let's analyze now from a perspective "Shneiderman's Mantra" the possibilities of visualizations based on City and Landscape metaphors.

Overview task may be realized by the flight over the city or landscape. **Zoom** task is implemented easily by lowering or lifting during the flight. The idea of the realization **Filter** task may be borrowed from cartography (and Information Visualization based on cartography techniques). There is the method of filtration on maps presented geographical data – to eliminate some types of information from the overall picture as for example human settlements may be eliminated from the map presented ground reliefs. The analogy of cartography shows that Landscape metaphor is preferred then City metaphor in the case of **Filter** task. **Details-on-demand** task, as well as **Relate** task may be implemented within the framework of the extended Room-Building-City metaphor by means "passing" down the street and "inputs" inside buildings and rooms. **History** and **Extract** tasks may realized naturally in frameworks of City and Landscape metaphors.

Schneiderman's scheme may be applied to evaluate Software Visualization in cases when corresponding systems are destined to be analyzed large volumes of abstract data similar to Information Visualization systems. In other cases these scheme is not applicable. Schneiderman's criterion is based on check of necessary, but not ampleness conditions of quality of Information Visualization. The use of the Schneiderman's scheme presupposes the existence of large structured data volumes. But in this case the resulted visualization has to be a manageable size. More importantly it is assumed that the user either knows what she/he searches or at least she/he is able to recognize it. In the case of "circular bundle" the new complicated abstract visualization technique (based on the new metaphor) is used. The user should always correlate the images with non-obvious representations of interesting data.

Similarly the visualization techniques based on using different charts and diagrams in many cases generate abstract and nontrivial representations. In the case of high-performance computing the methods of visualization for execution traces may be ineffective, due to the complexity of both the analysis of codes execution, and large data volumes. Such considerations can be applied for many new abstract methods (visualization metaphors) for complex data representation. The views using modification of statistical diagrams and

charts scale insufficiently. They can't map the execution of hundreds and thousands of parallel processes. Also let's state a remark about "natural" metaphors using. Interpretation of graphical displays implemented in the framework of the "natural" metaphors, for example, interesting "animation" Brain metaphor, often is not obvious. The naturalness of imagery in the cases of City and Landscape metaphors can sometimes distract users. Also, there are problems of perception and interpretation of large and huge data volumes. For example, the the flickering and blinking animation displays observation may be unusual and unpleasant. Users of systems based on virtual reality can have discomfort in the form of dizziness and so forth. There is the problem of selecting the objects to be visualized in the debugging process. In the case of parallel computation the definition of program objects associated with its "bug" states is a difficult task. Set and the essence of the analyzed program objects differ markedly in the various paradigms of parallelism. The execution trace is only one of possible entities subjected to analysis and, as a consequence, to visualization. In the case of performance tuning there is also no clarity with selection of entities which can help to improve performance.

5 CONCLUSION

In this paper approaches to the evaluation of Software Visualization for Parallel Computing are considered on the examples of execution traces and call graphs of parallel programs. The validity of visualization techniques was evaluated on basis of Shneiderman's scheme. Additionally cognitive approaches to the analysis of visualization and to the evaluation of the implementation efforts were considered. In the pre-design analysis the consideration of whole range of evaluations is necessary. Contradictions between representation opportunities and visualization perception, analysis and interpretation abilities of the users exist. Interesting metaphors can give pictures difficult to interpret or demand big efforts when developing. Scaling problem remains unsolved for many techniques of Software Visualization for Parallel Computing. This problem is related to fundamental limitations on placing "big pictures" on the screen, and the user perception and interpretation of "big data" generated by debugging and performance tuning systems.

6 REFERENCES

[1] Mohr, B., 2014. Scalable parallel performance measurement and analysis tools - state-of-the-art and future challenges. In Supercomputing frontiers and innovations. Volume 1. Number 2 (2014). Pp. 108-123.

[2] Averbukh, V.L., Baydalin, A.U., Ismagilov D.R.,, Kazantsev, A.U., Timoshpolskiy, S.P., 2004. Utilizing 3D Visualization Methophors. In Proceedings of 14 International Conference "Graphicon". 2004, Moscow, Russia. Pp. 295-298.

[3] Fittkau, F., Waller, J., Wulf, Ch., Hasselbring, W., 2013. Live trace visualization for comprehending large software landscapes: The ExplorViz approach. In Proceedings of the 1st Working Conference on Software Visualization (VISSOFT), IEEE Computer Society, 2013. 4 pp.

[4] Waller, J., Wulf, Ch., Fittkau, F., Dohring, Ph., Hasselbring, W., 2013. SynchroVis: 3D Visualization of Monitoring Traces in the City Metaphor for Analyzing Concurrency. In First IEEE Working Conference on Software Visualization (VISSOFT), 2013. 4 pp.

[5] Osawa, N., 1998. An Enhanced 3-D Animation Tool for Performance Tuning of Parallel Programs Based on Dynamic Models. In SPDP 98 Welches Or. USA. pp.72-80.

[6] Palepu, V.K., Jones, J.A., 2013. Visualizing constituent behaviors within executions. In Proceedings of the 1st Working Conference on Software Visualization (VISSOFT), IEEE Computer Society, 2013. 4 pp.

[7] Cornelissen, B., Holten, D., Zaidman, A., Moonen, L., van Wijk, J.J., van Deursen, A., 2007. Understanding execution traces using massive sequence and circular bundle views. In Proc. of the 15th IEEE Int. Conf. on Program Comprehension. IEEE, 2007, pp. 49-58.

[8] Cornelissen, B., Zaidman, A., Holten, D., Moonen, L., van Deursen, A., van Wijk, J.J., 2008. Execution Trace Analysis through Massive Sequence and Circular

[9] Maletic, J.I., Marcus, A., Collard, M.L., 2002. A task oriented view of software visualization. In International Workshop on Visualizing Software for Understanding and Analysis. 2002. Pp. 32-40.

[10] Shneiderman, B., 1996. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In Proceedings of the IEEE Conference on Visual Languages, September 3-6, 1996. Pp. 336-343.

[11] Wiss U., Carr, D., Jonsson, H., 1998. Evaluating Three-Dimensional Information Visualization Designs. A Case Study of Three Designs. In Proceedings of International Conference on Information Visualisation, London, England, July 29-31. 1998. Pp. 137-144.

ISSN 2464-4617 (print) WSCG 2016 - 24th Conference on Computer Graphics, Visualization and Computer Vision 2016
ISSN 2464-4625 (CD-ROM)

Medical image processing tools for Blender with HPC support

Petr Strakoš*, Milan Jaroš,
Tomáš Karásek, Marta Jarošová,
Daniel Krpelík, Alena Vašatová
IT4Innovations, VŠB-Technical University of Ostrava, Czech Republic
*petr.strakos@vsb.cz

Tereza Burešová, Juraj Timkovič,
Jan Štembírek, Petr Vávra
The University Hospital in Ostrava, Czech Republic

ABSTRACT

We present a medical image processing plug-in in this paper. Our plug-in uses Blender's environment and adds tools for the medical image processing and 3D model reconstructions and measurements. There are several software solutions to provide these tasks, two of which are used for comparison with our tool in this paper. However, there is not a single application that can perform medical image processing and subsequent 3D model reconstruction and post-processing with possible utilization of High Performance Computing (HPC) resources. In our plug-in we aim to use HPC resources to speed up the most computationally extensive tasks. The plug-in is written in Python and C++. The basic functionality and additional tools are implemented in Phyton. Blender has built-in Python interface, thus the development of the tools is quite fast. The C++ language is used for all time critical functions as it has unrestricted potential in parallel programming that can utilize available HPC resources. In the paper we present two examples, where we demonstrate possibilities of our Blender plug-in. We provide 3D reconstruction of human airway and human orbital floor and compare the results with those provided by 3D Slicer and Osirix.

Keywords

Blender, Python, tools, medical imaging, HPC, parallel computing.

1 INTRODUCTION

The fields using powerful computer technologies are getting larger with their availability. Medicine is one of the areas of human activities where such technologies are widely used. Modern sophisticated medical equipment is usually based on a computer technology. As a typical example, Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) scanners can be mentioned. These devices use computer technology to process and provide digital image information of a patient's inner body structure. Such information about individual patients provides a higher level of diagnostic medicine. To move the possibilities even further, the effort is put to the post-processing of acquired data. Image data can be used to provide virtual models of human organs and other tissues. To perform this, besides the powerful hardware, the capable software is necessary.

There are several software solutions specialized in medical image visualization and further processing. Here, we can mention 3D Slicer [4] or Osirix [9] as good examples of long term developed software projects.

3D Slicer is medical image processing software offering tools for the whole image processing cycle, from loading image data stored in the Digital Imaging and Communications in Medicine standard (DICOM), processing the data, to creating 3D reconstructions of selected parts of a human body. Several measurement and result evaluation tools such as distance, area, and volume measurement tools are available.

The other medical image processing software, Osirix, also offers the whole process from loading image data to final 3D reconstructions of selected parts of a human body. Osirix offers several measurement tools for data analysis and result evaluation as well.

Our aim in this paper is to present our own image processing software and introduce its powerful capabilities. Since a lot of image processing methods and techniques is very time consuming, it would be advantageous to have software which can exploit the latest powerful HPC resources. With this in mind, we have started to develop our own medical image processing software implemented as a plug-in for Blender, the 3D

graphics and animation software [1]. Blender is interesting for its Application Programming Interface (API), which allows advanced Python scripting.

Our software is implemented to take advantage of parallel computing and exploit the latest powerful HPC resources. More demanding tasks are implemented in parallel using OpenMP and MPI in C++ language. Our implementation can exploit the computational power of both non-accelerated nodes as well as Intel Xeon Phi accelerated compute nodes of a cluster.

2 MODEL RECONSTRUCTION

To provide 3D model reconstructions from medical image data, the software has to go through several typical steps that can be considered as general steps for any software of such kind. We give the explanation as a description of our plug-in components.

First of all, the data has to be loaded. Concerning medical image data, it is stored in the DICOM format. To load this data into Blender, Grassroots DICOM (GDCM) C++ library is integrated in our plug-in.

After loading the medical image data, its visualization is provided to users in a typical axial, sagittal, and coronal view. It might be beneficial to restrict such image data to only some regions of user interest and thus reduce the computational time of the further applied image processing techniques. In our plug-in, we can restrict the computations only to the volume area of our interest by using the box or sphere cutting tool.

Another step within the 3D model reconstruction process is image filtering. The imaging techniques like CT or MRI produce residual image noise, which can have significant impact on subsequent operations performed with an image. To decrease the adverse effect of noise, we have implemented several filtering techniques such as the simple Gaussian smoothing [10], a more sophisticated anisotropic diffusion filter [8], and a very effective and state-of-the-art Sparse 3D Transform-Domain Collaborative filter (BM3D) [3].

After filtering the images, image segmentation is performed. It serves as image simplification in order to localize the objects and their boundaries. Two methods are currently implemented in our plug-in such as a simple image thresholding [10] and k-means clustering [10]. The parallel implementation of the k-means method by using Intel Xeon Phi was introduced in the paper [11]. Both methods can be used separately or combined together.

If the segmentation of the selected body tissue is finished, a reconstruction technique that provides iso-surfaces as a result can be applied. In our plug-in we use Screened Poisson Surface Reconstruction [6] and Metaballs [2]. Both methods use information about identified boundaries as an input parameter. To create a polygonal mesh from iso-surfaces, the Marching Cubes method is used [7]. This method is time demanding, and so the efficient parallel implementation was necessary. For visualization, Blender Cycles engine is used. To accelerate the computations, the architecture of the Graphic Processing Unit (GPU) (the implementation is already included in Cycles) and Intel's Many Integrated Core (MIC) architecture can be employed. Our implementation for MIC architecture contains all features of the engine with improved performance [5].

The appearance of the plug-in within Blender environment is shown in Figure 1.

3 AUXILIARY TOOLS

When the 3D model is reconstructed, some additional software functionality, which can help in diagnostic and planning of the patient treatment, could be required. In the current version of our plug-in, some painting, erasing, cutting, and measuring tools are implemented. A lot of them are based on simple calculations.

We have implemented a tool to correct the results of the semi-automatic segmentation. It is called 2D PAINTER/ERASER and it can also be used to perform manual segmentation. The tool uses a typical brush and erase buttons, which open the segmentation layer for editing.

A similar tool that allows the user to mark the surface area of a selected 3D object but that serves another purpose than image segmentation is called 3D PAINTER/ERASER. This tool works as an intelligent 3D brush, which can paint on 3D surfaces. The surface area marked by the 3D PAINTER/ERASER is then processed by another tool to obtain the total area size of the marked surface. The algorithm behind the tool sums the size of all coloured triangles. The mesh of triangles creates surface of the reconstructed model.

Since surgeons are interested in using this software for planning surgeries, where, e.g., precise measurement of organ volume or resection of a specific part of a tissue is necessary for successful treatment, we have developed two specific tools for such manipulations. The first one is the CUT VOLUME tool for creating arbitrary cuts of a segmented object volume. It provides an infinite plane, which is used as the cutting tool to restrict spatially segmented data.

The second tool calculates the volume of an object and it is named the VOLUME MEASUREMENT tool. The algorithm sums the volumes of each voxel in the object.

Inasmuch as surgeons and radiologists usually work with 2D images, the 3D models resulting from the reconstructions can sometimes be confusing for them. For this reason and also for checking the correctness of the provided reconstructions, we have developed the

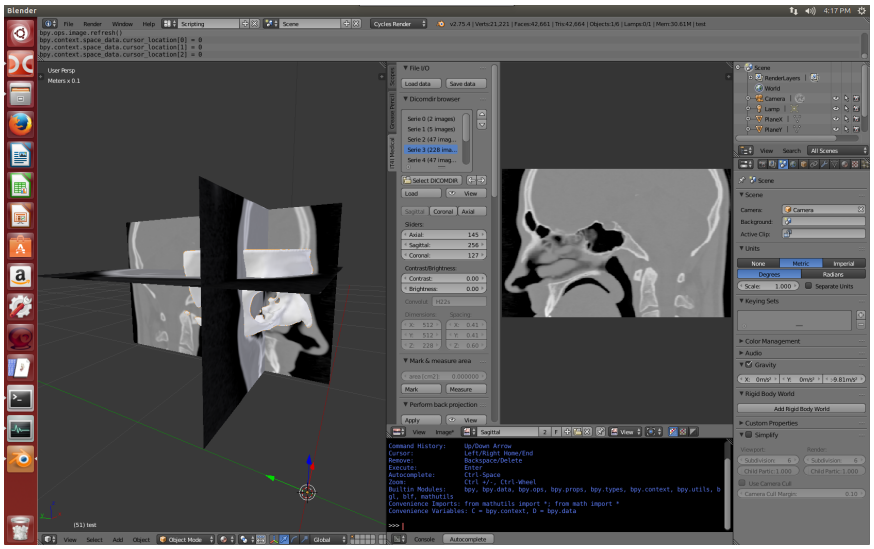


Figure 1: Our plug-in for medical image processing within Blender environment

BACKWARD PROJECTION tool. This tool can show the shape of the reconstructed 3D model as a projection to the original CT data. This allows users to see all important features of the data and at the same time compare accuracy of the provided reconstruction.

4 APPLICATION OF THE TOOLS

To show the capabilities of our plug-in, we present here two medical test cases, in which the different functions of our plug-in are exploited. Specifically, we aimed to test precision of the whole 3D reconstruction concept. As a result, we provided 3D models, the volume of which we then measured along with the size of the selected surface area. Within the tests we compared the results performed by our plug-in with the results obtained by the other two already mentioned image processing software solutions, Osirix and 3D Slicer.

For the first test, we selected reconstruction of the human airway, the volume of which was then evaluated. The process of model reconstruction with volume measurement was performed in each software individually. The final results of the volume measurements were then compared.

In the second test, the model reconstruction of the human orbit was done. In each software we carried out the reconstruction from the input data up to the 3D model. As neither Osirix nor 3D Slicer have a tool for measuring surface area marked on the reconstructed model, we exported the models in STL format to Blender, using our tool to mark the surface of the orbital floor and performing the measurement of the selected area. Again, the final results were then compared.

4.1 Reconstruction of the airway

Precision of 3D model reconstruction in terms of total volume is very important for medical doctors who will use those models for diagnostic purposes or during surgery planning. For example, while diagnosing sleep apnea, the volume of upper respiratory tract is needed. To mention another example, while planning a liver resection, it is important to calculate the exact size of the liver with lesion, which will be removed during surgery. In both cases, precise information based on 3D model reconstruction is a critical issue.

To compare volumes of 3D models obtained from 3D Slicer, Osirix, and our Blender plug-in, 3D reconstruction of the airway was performed. This was done on DICOM image series of one patient. The density of the data measured by mutual distance between the slices was 0.75mm.

The numerical results showing the airway volume for each software are written in the Table 1. From the results we can see that our model creation concept offers similar model reconstruction as specified by the amount of the final volume, as those measured by other two software.

Patient	N003
Osirix [ml]	18.07
3D Slicer [ml]	18.33
Our plug-in [ml]	18.22

Table 1: Volume of the reconstructed human airway

4.2 Reconstruction of the orbital floor

Another feature requested from medical doctors is the measurement of an area or surface size. During the treatment of orbital fractures, the information about the size of orbital floor damage is critical for the decision about the type of treatment. Our image processing plug-in can provide the information about the size of the injury in terms of surface area measurement. Based on

this finding, the doctors can make the decision whether to provide a conservative or surgical treatment of the patient.

We performed 3D reconstruction of the orbit. This was done on the DICOM image series of three different patients. The density of the data measured by mutual distance between the slices was 0.6mm in all cases. The comparison of the 3D reconstructions provided by 3D Slicer, Osirix, and our Blender plug-in was made. Individual 3D reconstructions from Osirix and 3D Slicer were, due to the lack of a specific measuring tool, exported as STL models and loaded into Blender. Then the orbital floor area was marked and measured by a trained person. For marking and measurement of such surface area, we used our developed tool as described in Section 3.

The numerical results showing the area of the orbital floor surface for each patient and each software are written in the Table 2. The results show that model precision is similar. Moreover, using our tools, we can provide area measurements on reconstructed surfaces, which is a feature available only in our plug-in.

Patient	N001	N002	N010
Osirix [cm^2]	16.38	10.63	10.15
3D Slicer [cm^2]	16.14	10.37	10.24
Our plug-in [cm^2]	16.28	10.49	10.14

Table 2: Area of the orbital floor for individual reconstruction and patient data

5 CONCLUSION

In the presented paper, we have introduced our Blender plug-in equipped with a number of tools for medical image processing. By using it, users can create 3D reconstructions of selected human body tissues from loaded DICOM data. Several tools in the plug-in can utilize the computational support of HPC and parallelization and thus significant processing speed-up can be achieved.

The created 3D reconstructions are of the similar precision as those created in the other two software (3D Slicer, Osirix) we used as a reference. As a unique thing in our plug-in, we provide users with the surface selection and measurement tool for area measurement on the reconstructed surfaces. This can be highly beneficial when, e.g., evaluating the surface area of the orbital floor for studies focusing on eye injuries.

6 ACKNOWLEDGMENTS

This work was supported by the Ministry of Education, Youth and Sports from the National Programme of Sustainability (NPU II) project "IT4Innovations excellence in science - LQ1602" and from the Large Infrastructures for Research, Experimental Development and Innovations project "IT4Innovations National Supercomputing Center - LM2015070".

7 REFERENCES

[1] blender.org - Home of the Blender project - Free and Open 3D Creation Software. Available from: <https://www.blender.org>. [18 February 2016].

[2] J. Bloomenthal, C. Bajaj, J. Blinn, M. Gascuel, A. Rockwood, B. Wyvill, G. Wyvill, Introduction to Implicit Surfaces, Morgan Kaufman Publishers Inc, 1997.

[3] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image Denoising by Sparse 3D Transform-Domain Collaborative Filtering, In: IEEE Trans. Image Process., vol. 16, no. 8, pp. 2080–2095, (2007).

[4] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S.R. Aylward, J.V. Miller, S. Pieper, R. Kikinis, 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network, Magnetic Resonance Imaging 2012; July PMID: 22770690.

[5] M. Jaros, L. Riha, T. Karasek, P. Strakos, A. Vasatova, M. Jarosova, T. Kozubek, Acceleration of Blender Cycles Path-Tracing Engine using Intel Many Integrated Core Architecture, Proceedings of the 14th International Conference on Computer Information Systems and Industrial Management Applications, Warsaw, Poland, 2015.

[6] M. Kazhdan, H. Hoppe, Screened Poisson surface reconstruction, ACM Trans. Graphics, 32(3), 2013. (Presented at SIGGRAPH 2013).

[7] W. E. Lorensen, H. E. Cline, Marching Cubes: A high resolution 3D surface construction algorithm, Computer Graphics, Vol. 21, Nr. 4, July 1987.

[8] P. Perona, J. Malik, Scale-space and Edge Detection Using Anisotropic Diffusion, In: IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 12, no. 7, pp. 629–639. (1990).

[9] A. Rosset, L. Spadola, O. Ratib, OsiriX: an open-source software for navigating in multidimensional DICOM images, J Digit Imaging, 2004 Sep;17(3):205-16, Epub 2004 Jun 29; PMID: 15534753.

[10] M. Sonka, V. Hlavac, R. Boyle, Image Processing, Analysis and Machine Vision, Thomson, (2006).

[11] P. Strakos, M. Jaros, T. Karasek, L. Riha, M. Jarosova, T. Kozubek, P. Vavra, T. Jonszta, Parallelization of the Image Segmentation Algorithm for Intel Xeon Phi with Application in Medical Imaging, in P. Ivnyi, B.H.V. Topping, (Editors), Proceedings of the Fourth International Conference on Parallel, Distributed, Grid and Cloud Computing for Engineering, Civil-Comp Press, Stirlingshire, UK, Paper 7, 2015.

Implementation of Face Recognition Methods as a First Step for Human Behaviour Analysis in Intelligent Room

Alexander Ronzhin
SPIIRAS
39, 14th line
199178, St. Petersburg, Russia
ronzhinal@iias.spb.su

Irina Vatamaniuk
SPIIRAS
39, 14th line
199178, St. Petersburg, Russia
vatamaniuk@iias.spb.su

Miloš Železný
University of West Bohemia
Univerzitní 8
306 14, Pilseň, Czech Republic
zelezny@kky.zcu.cz

ABSTRACT

This paper discuss an intelligent space concepts, goals and developments. The review presents an analysis of seven intelligent meeting rooms, equipment and developed services. One of the main goals of intelligent space is the development of proactive services. Realization of such types of services requires the analysis of current situation in a room as well as behaviour of the audience. Implementation of methods for facial features analysis are the first step for this goal destination. This paper presents an algorithm for facial image normalization and recognition as well as a developed system for automatic registration and identification of meeting participants in the intelligent room. The developed algorithm is intended for processing of low resolution facial images with approximate size of 30x30 pixels. The experiments shows distribution of participants recognition rate by variety of their face sizes. In addition, implementation of blur estimation method at preliminary stage of facial image processing shows an increase of the whole system speed as well as a decrease in face recognition error rate.

Keywords

Computer vision, face recognition, image normalization, low resolution image, image quality estimation, intelligent room.

1 INTRODUCTION

The development of technologies for the automated processing of audiovisual and other types of data is required at a constant growth of volumes of recorded data. The significant amount of these data is related to photo and video recording. The automatic image processing usually includes preliminary processing (elimination of images of unsatisfactory quality, light normalization, noise removal, etc.), feature extraction, object segmentation and pattern recognition. Often it is necessary to select some images of high quality corresponding to the requirements of a particular technical task from an image database. For an operator such work is not difficult, however, if the amount of data is large, especially at the stage of preliminary processing, it is advisable to automate the process of image quality evaluating and elimination of the frames with no reliable data. In addition, implementation of blur estimation methods at preliminary stage allows increasing accuracy of pattern recognition methods.

This paper is organized as follows. Section 2 presents a concept and goals of an intelligent space. Section 3 discusses facial image normalization methods. Section 4 describes the developed algorithm for facial image normalization and recognition. Section 5 presents the experiments, conditions and results.

2 INTELLIGENT SPACE

The research work of recent years in the development of prototypes of intelligent spaces (in most cases, intelligent meeting rooms) for scientific projects has been analyzed. Let us consider the functionality of these prototypes of intelligent meeting rooms. By functionality we understand their equipment and methods for the implementation of information support services and the automation of the support of activities. The sensory equipment of the intelligent meeting room can be divided into two types: primary equipment for the recording of events in the entire space of the conference room and supporting equipment for a more detailed analysis of specific areas of interest identified by the main equipment.The list of sensory and multimedia equipment of rooms and the methods used for processing audiovisual signals are given in [Ronzhin15a]. Most intelligent meeting rooms are equipped with audio and video recording equipment as well as with advanced I/O multimedia devices. An analysis of the data [Ronzhin15a] shows that the services for recording, broadcasting, and logging of activities, as well as the creation of test audiovisual databases, are implemented in almost all con-

sidered intelligent meeting rooms. However, services that are important for the automation of the course of activities, such as automatic participant recording, current speaker detection, integration of participant devices with the intelligent meeting room's system, and automatic generation of multimedia report, were implemented only in half of these intelligent meeting rooms.

3 METHODS FOR IMAGE NORMALIZATION

The paper [Xie08] presents a framework for solving the illumination problem in face recognition tasks. Features extracted from face images may be divided into two categories small- and large-scale features also may be called as small and large intrinsic details respectively. The first aim in [Xie08] was decomposition of face image into two features images by implementation of logarithm total variation method (LTV) with interpoint second-order cone program (SCCP) algorithm for approximate estimation of LTV model [Chen06]. The second aim was processing of decomposed face image elements for smoothing of small-scale and illumination normalization of large-scale features. For illumination normalization in work [Xie08] methods non-point light quotient image relighting (NPL-QI) and truncating discrete cosine transform coefficients in logarithm domain (LOG-DCT) were separately employed and estimated. All processed features are used for reconstruction of normalized face image. It shows [Xie11] that implementation of NPL-QI allows one to extend the illumination estimation from single light source to any type illumination conditions and simulate face images under arbitrary that conditions. The LOG-DCT is not so likely for image normalization because it loss some low-frequency information during image restoration.

The paper [Liu08] presents a method of illumination transition of an image (ITI) for illumination estimation and image re-rendering. This method is based on the calculation of face similarity factor, which is considered: the more similarity comparison of an image with more weighted pre-referenced personal face image allows generating a person-specific reference face $T_{x,z}$:

$$T_{x,z} = \frac{\Sigma_i k_i A_i z + \varepsilon}{\Sigma_i k_i A_i x + \varepsilon}$$

where $k = (k_1, k_2, \ldots, k_N^T$ is the weight coefficient, A_i is an images array of subject i, x and z are current and expected illumination parameters respectively, ε is a small constant ($\varepsilon = 1$), which is used in [Liu08] in case of being divided by zero. The experimental results [Liu08] show that current method may be used for illumination normalization of facial images as preprocessing in face recognition methods.

In continuation of works [Xie08, Xie11, Liu08], the paper [Xie14] presents the implementation of empirical mode decomposition (EMD) [Huang98] for illumination preprocessing of facial images. This method can adaptively decompose a complex signal into intrinsic mode functions (IMFs), which are relevant to intrinsic physical significances. A face image may be represented as Lambertian reflectance assumption [Xie08]:

$$I(x,y) = R(x,y)L(x,y),$$

where $I(x,y)$ is the observed intensity, $R(x,y)$ is the reflectance (albedo), and $L(x,y)$ is the shading or illumination component. The right side of this formula is multiplicative and the EMD decomposes signal into IMFs in an additive form. So, factorization in the logarithm domain of EMD (LEMD) was used in [Liu08] for conversion of multiplicative model into additive:

$$f = \log I = \log R + \log L.$$

Because the EMD can decompose a signal into a set of IMFs with different frequencies, the R and L components may be estimated as:

$$R = \exp\left(\sum_{k=1}^{K_0} d_k\right), L = \exp\left(\sum_{(} k = K_0)^K d_k + r\right),$$

where d_k are the IMFs with different frequencies k, r is residue, and K_0 is a settable parameter, in [Xie14] $K_0 = 2$.

The complete algorithm for illumination normalization in facial images, presented in [Xie14], is:

1. Computation of the logarithm of image;
2. Perform the 1-D EMD of gathered logarithm (experiment result in [Xie14] shows that 1-D EMD method have better recognition performances);
3. Detection of shadow regions. In this stage each IMFs have been analyzed. There is shadow in a IMFs if $D \cdot e^{m_{IMF}^2} > \theta_1$ is satisfied the binarization operator with threshold $\theta_2 = m_{IMF}/2$ is applied to analyzed IMF or residue, where D and m_{IMF} are variance and mean pixel values recently as well as the threshold $\theta_1 = 0.1$ [Xie14], and then each connected black area is marked as a shadow region;
4. Grayscale adjustment. Gray level Substitution of each detected region by average gray level (with the use of Gaussian weighting) of its neighboring region.
5. Restoration of the image using processed IMFs and residue and conversion of it to original image space.

For experiments only frontal-face images from same databases as in work [Xie11] were carried out.

4 ALGORITHM FOR FACIAL IMAGE NORMALIZATION AND RECOGNITION

The proposed algorithm of participant identification based on group photo processing is a part of an intelligent meeting room software, which is described in detail in [Ronzhin15a]. First of all, what is the group photo? The classic or porters photo is an image with simple background and several peoples, who stay or sit one by one on the foreground and look at the camera point. But the group photo, which was captured during a meeting, has more complexity, including: 1) complex background; 2) occlusions of participants by each other; 3) illumination and shadow variety; 4) blurred participant's faces (by their own movements at the shooting moment); 5) small size of participant's face region on an image.

In this stage of research the algorithm (Fig. 1) will decide first three and fifth problem. The image quality estimation methods would not be implemented in the algorithm, because state of the art methods cannot make a decision about blurriness of an image with a low resolution (approximately 30x30 pixels), details of such problem are presented in [Ronzhin15b]. The high resolution image (1280x1024 pixels), which is a group photo of participants sitting in chairs, is processed for finding their faces by face detection procedure [Viola03]. Then each face region is processed for estimation of facial reference points such as eyes, nose and mouth. This information is used for images dividing into two categories – frontal-face and other (with angle more than 15 degree from camera point) by the head orientation estimation procedure. For frontal-face images the corrected region was extracted as rectangle and for other images as ellipsoid. These procedures are used for decision of complex background and occlusion problems by reduction of background patches in a facial region.

After extraction of the final participant facial region, information about distance to the participant and illumination in the room was analyzed. Such information was gathered from the audiovisual monitoring system of the intelligent meeting room [Ronzhin15a], description of the room can be found in [Yusupov11].

For image decomposition in the proposed algorithm, the LTV model and LEMD method based on the result in [Xie11, Xie14] will be used separately. In addition, after decomposition by LTV model, the threshold-average filtering will be used for smoothing the small-scale features. As presented in [Xie14], the LEMD decompose facial image into several IMFs, and the illumination-insensitive facial representation is stored in first IMFs. So, in this case, for large-scale features normalization will be IMFs from second to residue. In LTV model application for this aim NPL-QI and LOG-DCT algorithms would be separately used for normalization of shadow and illumination as well as for reconstruction of normalized image. The final step of the algorithm is the identification of the participant by recognition of his face, which is based on Local Binary Patterns method (LBP) [Ahonen06].

5 EXPERIMENTS

For the experimental evaluation of the automatic participants registration system during the events in the intelligent meeting room, participants photos were accumulated only at the first algorithm of the system. As a result, the number of accumulated photos was more than 55000 for 36 participants. The training database contains 20 photos for each participant. For estimation of the influence of participants face size change on recognition rate, it was decided to divide them into several groups. Figure (Fig. 2a) shows distribution of participants by variety difference of their face sizes in ranges from 0 to 10, from 10 to 20, and so on. Most of participants have a difference between minimum and maximum face sizes in range from 30 to 40 pixels. (Fig. 2b) shows a distribution of recognition rate for three methods for the groups of participants. From (Fig. 2b) it is obvious that with increasing of the participant's face size variety difference, the recognition accuracy gradually decreases. This is due to the fact that at the normalization of images to a uniform size, a distortion in certain facial features like eyes, nose, mouth may occur.

6 CONCLUSION

Personalized user service and satisfaction of user needs in an unobtrusive and almost invisible form by analyzing their behaviour and determining the current situation is the main idea in the concept of an ambient intelligent space. The awareness of the conference room of the spatial position of participants, their current activities, their role in the activity, and their preferences provides intelligent control over multimedia and other equipment in the room. In order to determine the current behaviour of participants and analyze the situation, modern interconnected methods and a software for audiovisual processing of data that automatically control audio and video recording equipment are required. Considering the experimental conditions (different distances from the camera to the participant, lighting, movement of participants while taking pictures) that influence the quality and quantity of the facial features extracted from the image, which are directly influenced by the accuracy of recognition and occurrence of false positives, we can conclude that the best results are shown by the method LBP 79,5%.

7 ACKNOWLEDGMENTS

This research is financially supported by the Ministry of Education and Science of the Russian Federation, agreement no. 14.616.21.0056 (reference

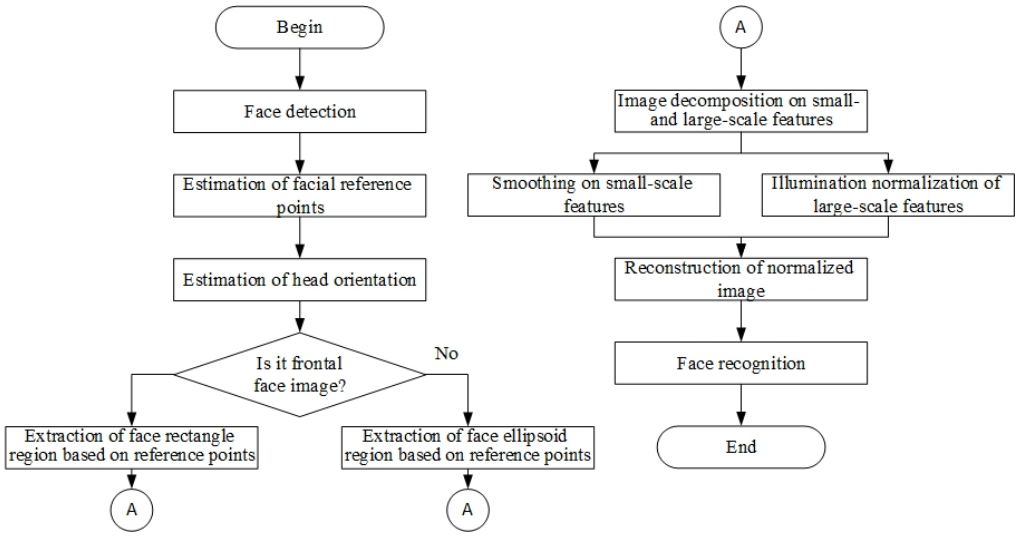


Figure 1: Scheme of the participant identification algorithm

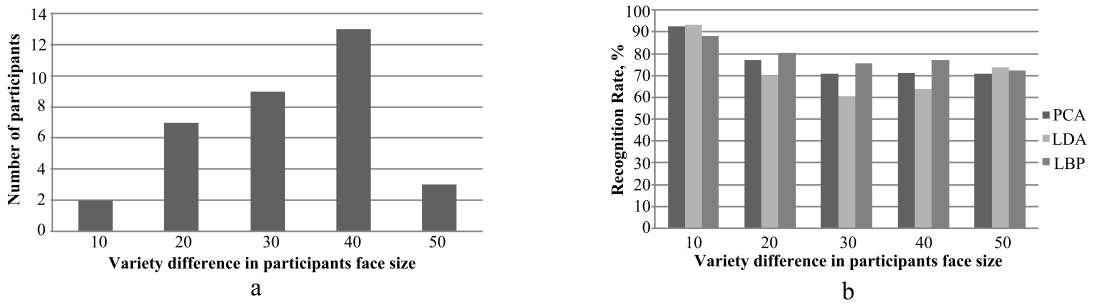


Figure 2: Distribution of participants recognition rate by variety of their face sizes

RFMEFI61615X0056), as well as by the Czech Ministry of Education, Youth and Sports, project no. LO1506.

8 REFERENCES

[Ahonen06] Ahonen, T., Hadid, A., Pietikainen, M. Face description with local binary patterns: Application to face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006. Vol. 28. No. 12. pp. 2037–2041.

[Chen06] Chen, T. et al. Total variation models for variable lighting face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006. Vol. 28. No. 9. pp. 1519–1524.

[Huang98] Huang, N. E. et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. The Royal Society, 1998. Vol. 454. No. 1971. pp. 903–995.

[Liu08] Liu J. et al. Illumination transition image: parameter-based illumination estimation and re-rendering. In Proceedings of the 19th ICPR. 2008. pp. 1–4.

[Xie08] Xie, X. et al. Face illumination normalization on large and small scale features. IEEE Conference on Computer Vision and Pattern Recognition CVPR. 2008. pp. 1–8.

[Ronzhin15a] Ronzhin, A., Karpov, A. A software system for the audiovisual monitoring of an intelligent meeting room in support of scientific and education activities. Pattern Recognition and Image Analysis. 2015. Vol. 25. No. 2. pp. 237–254.

[Ronzhin15b] Ronzhin, An., Vatamanyuk, I., Ronzhin, Al., and Zelezny, M. Mathematical Methods to Estimate Image Blur and Recognize Faces in the System of Automatic Conference Participant Registration. Automation and Remote Control, 2015, Vol. 76, No. 11, pp. 1623–1632.

[Viola03] Viola, P., Jones, M., Snow, D. Detecting pedestrians using patterns of motion and appearance. In Proc. of IEEE ICCV. 2003. pp. 734–741.

[Xie11] Xie, X. et al. Normalization of face illumination based on large-and small-scale features. IEEE Transactions on Image Processing. 2011. Vol. 20. No. 7. pp. 1807–1821.

[Xie14] Xie, X. Illumination preprocessing for face images based on empirical mode decomposition. Signal Processing. 2014. Vol. 103. pp. 250–257.

[Yusupov11] Yusupov, R. M. et al. Models and hardware-software solutions for automatic control of intelligent hall. Automation and Remote Control. 2011. Vol. 72. No. 7. pp. 1389–1397.

Development of the crystal lattice parameter identification method based on the gradient steepest descent method

Alexandr Shirokanev Dmitriy Kirsh Alexandr Kupriyanov
Samara State Aerospace University
443001, Samara, Russia
alexkupr@gmail.com

ABSTRACT

In this paper, we propose a new method to solve a problem of crystal lattice parameter identification. The developed method is based on applying the gradient steepest descent method. The two algorithms of crystal lattice parameter identification on the basis of the developed method was proposed. The comparison of the algorithms with the existing method of lattice identification based on estimation of Bravais unit cell parameters is conducted. The experimental results show a significant increase in the crystal lattice parameter identification accuracy for all seven lattice systems.

Keywords

Crystal lattice, Bravais unit cell, translation vector, parameter identification, gradient steepest descent method.

1. INTRODUCTION

Nowadays, much attention has been concentrated on reconstruction of three-dimensional objects [Fur01a, Kot01a, Kud01a]. In particular in crystallography, reconstruction of a three-dimensional crystal lattice structure is related directly to a parameter identification problem, which is one of the basic problems of X-ray diffraction analysis [Shi01a, Kha01a].

The most well-known crystal lattice model was offered by Auguste Bravais. The Bravais model is based on unit cell representation: the entire lattice can be constructed by translation of a single cell. All unit cells are divided into seven lattice systems according to edge lengths and angle values (Figure 1) [Ham01a].

With evolving technology the parameter identification algorithms, as well as the crystal lattice comparison methods become more relevant [Bra01a, Law01a].

The objective of the crystal lattice parameter identification method is to estimate unit cell parameters. There are several methods that offer a solution to the problem: NIST lattice spacing comparator [Kes01a], parameter identification methods based on estimation of atomic packing factor [Smi01a] and distances between isosurfaces [Pat01a]. However, these methods are not universal and have a number of disadvantages, such as strong dependence between the crystal lattice identification accuracy and the lattice system, high sensitivity to distortions of crystal lattice point coordinates or complexity of the sample preparation.

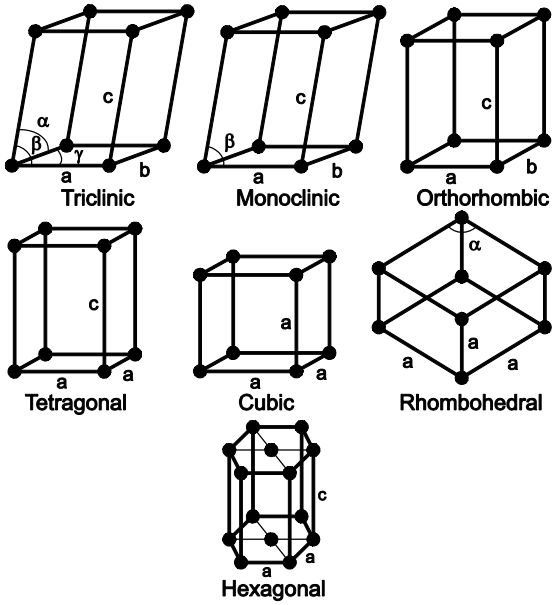


Figure 1. Unit cells of seven lattice systems.

Among the existing universal methods that provides high accuracy of crystal lattice parameter identification, we can distinguish the following one: "The lattice identification method based on estimation of Bravais unit cell parameters" [Bes01a]. In the experimental section, the comparison of the results obtained by this method and the developed one is presented.

2. THE LATTICE IDENTIFICATION METHOD BASED ON ESTIMATION OF BRAVAIS UNIT CELL PARAMETERS

The method is based on calculation of six key parameters of the Bravais unit cell, i.e. three edge lengths and three included angles [Kup01a].

Initial data for the identification method based on estimation of the Bravais unit cell parameters are a finite set of radius-vectors of crystal lattice nodes.

The identification method involves a search of three non-coplanar vectors in the original set of nodes: the first vector has a minimum norm; the second vector does not lie on a straight line with a directing vector equal to the first vector; the third vector does not lie on a plane made by two found vectors.

3. DEVELOPED ALGORITHMS OF CRYSTAL LATTICE PARAMETER IDENTIFICATION BASED ON OPTIMIZATION PROBLEM SOLUTION USING THE GRADIENT STEEPEST DESCENT METHOD

Two algorithms of crystal lattice parameter identification have been developed. In both algorithms we consider the Bravais lattice represented by three translation vectors $\bar{a}_1$, $\bar{a}_2$ and $\bar{a}_3$ [Sha01a]. The set of lattice nodes is expressed as:

$$X = \left\{ \bar{x} = i\bar{a}_1 + j\bar{a}_2 + k\bar{a}_3 \right\} : \; i, j, k \in \mathbb{Z}.$$

In this case, both algorithms shall require initial approximation as an additional input parameter. Particularly, the result vectors of the lattice identification method based on estimation of Bravais unit cell parameters can be used as the initial approximation.

3.1. The Algorithm Based on Joint Optimization of Translation Vectors

The objective function of optimization is as follows:

$$E\left(\bar{a}_1, \bar{a}_2, \bar{a}_3\right) = \sum_{l=1}^{L} \min_{i,j,k} \left\| \bar{x}_l - \left(i\bar{a}_1 + j\bar{a}_2 + k\bar{a}_3 \right) \right\|^2, \qquad (1)$$

where L is the number of nodes in the lattice.

Let us introduce the following notation:

$$\mathrm{A} = \left(\bar{a}_1 \quad \bar{a}_2 \quad \bar{a}_3 \right) \in \mathbb{R}^{3\times 3};$$

$$\bar{n}_l = \left(i_l \quad j_l \quad k_l \right)^T;$$

$$\mathrm{N}_l = \bar{n}_l \bar{n}_l^T;$$

$$\bar{w}_l^s = i_l \bar{a}_1^s + j_l \bar{a}_2^s + k_l \bar{a}_3^s - \bar{x}_l,$$

s is the step number in the descent.

In this case a gradient (2) and a descent factor (3) are as follows:

$$\nabla E\left(\mathrm{A}\right) = 2\left[\mathrm{A} \cdot \sum_{l=1}^{L} \mathrm{N}_l - \sum_{l=1}^{L} \bar{x}_l \bar{n}_l^T \right]; \qquad (2)$$

$$\lambda_s = \frac{\sum_{l=1}^{L} \left(\bar{w}_l^s, \nabla E\left(\mathrm{A}^s\right) \bar{n}_l \right)}{2\sum_{l=1}^{L} \left\| \nabla E\left(\mathrm{A}^s\right) \bar{n}_l \right\|^2}. \qquad (3)$$

Expressions (2) and (3) are recorded in compact form. The desired solution is the matrix (triple translation vectors). The gradient is respectively the matrix, too.

3.2. The Algorithm Based on Independent Optimization of Translation Vectors

Nodes of a crystal lattice are to be set up by translation vectors, however in this case they shall pass through recurrent planes. This property enables to break a problem of translation vector optimization, described above, into three independent problems of optimization of some vectors which describe recurrent planes. The recurrent planes shall be described by a plane normal and a plane period. Thus, the desired vector must have the same direction as the plane normal, and the vector norm should coincide with the plane period.

The objective function shall be as follows:

$$E\left(\bar{d}\right) = \sum_{l=1}^{L} \min_{i} \left[\left(\bar{x}_l, \bar{d} \right) - i \left\| \bar{d} \right\|^2 \right]^2. \qquad (4)$$

For the objective function (4) the gradient shall be determined from the equation (5), and the descent factor – from the equation of the third degree (6).

$$\nabla E = 2\sum_{l=1}^{L} \left(\left(\bar{x}_l, \bar{d} \right) - i_l \left(\bar{d}, \bar{d} \right) \right) \left[\bar{x}_l - 2i_l \bar{d} \right]. \qquad (5)$$

$$y_0 + y_1\lambda + y_2\lambda^2 + y_3\lambda^3 = 0, \qquad (6)$$

where $y_0 = \sum_{l=1}^{L} p_l q_l$; $y_1 = \sum_{l=1}^{L} p_l r_l - \sum_{l=1}^{L} {q_l}^2$;

$y_2 = -2\sum_{l=1}^{L} q_l r_l$; $y_3 = -\sum_{l=1}^{L} {r_l}^2$;

$p_l = \left(\bar{x}_l - i_l \bar{d}^s, \bar{d}^s \right)$;

$q_l = \left(\bar{x}_l - 2i_l \bar{d}^s, \nabla E\left(\bar{d}^s \right) \right)$;

$r_l = i_l \left\| \nabla E\left(\bar{d}^s \right) \right\|^2$.

The iterative process of this algorithm may be controlled. A vector of an individual problem does not correlate with vectors of two remaining problems at all. If the first developed algorithm carries out S iterations for all three vectors (matrix), then in the case of the latest algorithm each desired vector shall

be found only for the required number of iterations. The case when one vector can result in an incorrect solution shall be also excluded due to dependence on other vectors.

The algorithm suggests the connection between a basis of translation vectors and a new basis of "independent" vectors. The transition from the basis of translation vectors to the basis of independent vectors shall be carried out by simultaneous linear algebraic equations (7). The reverse transition is performed using the following expressions (8).

$$\left(\bar{d}_1 \quad \bar{d}_2 \quad \bar{d}_3\right)^T \bar{a}_i = \left\|\bar{d}_i\right\|^2 \bar{e}_i, \quad i = \overline{1,3}, \tag{7}$$

where $\bar{e}_i$ is a unit vector in the i-direction;

$$\bar{d}_i = \frac{\left(\left[\bar{a}_k \times \bar{a}_j\right], \bar{a}_i\right)}{\left\|\left[\bar{a}_k \times \bar{a}_j\right]\right\|^2}, \quad i \neq j \neq k = \overline{1,3}. \tag{8}$$

4. ANALYSIS OF THE DEVELOPED ALGORITHMS OF CRYSTAL LATTICE PARAMETER IDENTIFICATION

The criterion of accuracy of the results obtained by applying the parameter identification algorithm is a solution error. Let us select the function used in the first developed algorithm as the solution error (1).

To calculate this function, the radius-vectors of the original lattice nodes and the parameters obtained by some algorithm shall be required.

For the experiments, crystal lattices have been modeled using the unit cell parameters of the seven existing materials. Each lattice consisted of 125 nodes. Then modeled lattices was distorted through a random offset of each node from its ideal position in a random direction at a distance no more than 0.5 Å.

Parameters of the Bravais unit cell were estimated for each modeled lattice, and the solution error was calculated. Figure 2 gives the research findings as values of the objective function. Both developed algorithms have considerably reduced an error obtained as the result of performing the identification algorithm based on estimation of parameters of the Bravais cell.

The accuracy of the parameter identification algorithm based on independent optimization of translation vectors depends crucially on the selection of the stop criteria. In case of the standard stop criteria, the iteration process continues until the objective function begins to grow or the optimizing vector stops changing. However, the experimental data showed that the second algorithm with the standard stop criteria returns bad results.

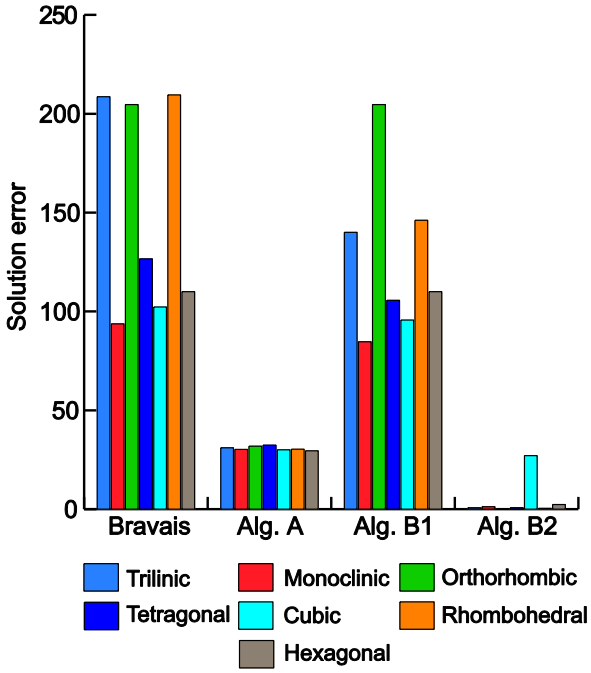


Figure 2. Accuracy of parameter identification: A – algorithm from Section 3.1; B1 – algorithm from Section 3.2 (standard stop criteria); B2 – algorithm from Section 3.2 (weakened stop criteria).

To improve the accuracy, we proposed to weaken the stop criteria by excluding the first condition (non-increasing objective function). The data in Figure 2 indicate that the weakened stop criteria allow to substantially reduce the error.

5. CONCLUSION

In the course of the present work, we have developed two algorithms of parameter identification based on the gradient steepest descent method. In both algorithms, the result vectors of the lattice identification method based on estimation of Bravais unit cell parameters is used as the initial approximation.

The research showed that in most cases the developed algorithms can provide better results than the Bravais unit cell parameter estimation algorithm. Moreover, the second algorithm with the weakened stop criteria demonstrates the minimum errors of parameter identification for all seven lattice systems.

6. ACKNOWLEDGMENTS

This work was partially supported by the Ministry of education and science of the Russian Federation in the framework of the implementation of the Program of increasing the competitiveness of SSAU among the world's leading scientific and educational centers for 2013-2020 years; by the Russian Foundation for Basic Research grants (# 14-01-00369-a, # 14-07-97040-p_povolzh'e_a, # 15-29-03823, # 15-29-07077); by

the ONIT RAS program # 6 “Bioinformatics, modern information technologies and mathematical methods in medicine” 2015.

7. REFERENCES

[Bes01a] Bessmeltsev, V. P., and Bulushev, E. D., Fast image registration algorithm for automated inspection of laser micromachining. Computer Optics, Vol. 38, No. 2, pp. 343-350, 2014.

[Bra01a] Brandon, D., Kaplan U., Microstructure of materials. Methods of research and monitoring. Technosphere Publisher, pp. 39-52, 2004.

[Fur01a] Fursov, V. A., Goshin, Ye. V., Information technology for digital terrain model reconstruction from stereo images. Computer Optics, Vol. 38, No. 2, pp. 335-342, 2014.

[Ham01a] Hammond, C. The Basic of Crystallography and Diffraction. Third Edition. Oxford University Press Inc., pp. 84-95, 2009.

[Kes01a] Kessler, E., Henins, A., Deslattes, R., Nielsen, L., Arif, M., Precision Comparison of the Lattice Parameters of Silicon Monocrystals. Journal of Research of the National Institute of Standards and Technology, Vol. 99, pp. 1–18, 1994.

[Kha01a] Kharitonov, S. I., Volotovskiy, S. G., Khonina, S. N., Kazanskiy, N. L., A differential method for calculating x-ray diffraction by crystals: the scalar theory. Computer Optics, Vol. 39, No. 4, pp. 474-476, 2015.

[Kot01a] Kotov, A. P., Fursov, V. A., Goshin, Ye. V., Technology for fast 3d-scene reconstruction from stereo images. Computer Optics, Vol. 39, No. 4, pp. 600-605, 2015.

[Kud01a] Kudinov, I. A., Pavlov, O. V., Kholopov, I. S., Implementation of an algorithm for determining the spatial coordinates and the angular orientation of an object based on reference marks, using information from a single camera. Computer Optics, Vol. 39, No. 3, pp. 413-415, 2015.

[Kup01a] Kupriyanov, A. V., Kirsh, D. V., Estimation of the Crystal Lattice Similarity Measure by Three-Dimensional Coordinates of Lattice Nodes. Optical Memory & Neural Networks (Information Optics). Vol. 24, No. 2, pp. 145-151, 2015.

[Law01a] Lawrence, C. A., Herbert J. B,. Lattices and Reduced Cells as Points in 6-Space and Selection of Bravais Lattice Type by Projections. Acta Cryst, pp. 1009-1018, 1988.

[Pat01a] Patera, J., Skala, V., Centered cubic lattice method comparison. 17th ALGORITMY Conference on Scientific Computing, pp. 309–319, 2005.

[Sha01a] Shaskolskaya, M. P., Crystallography: Manual for Institutes of Higher Education. Higher School, pp. 10-14, 1984.

[Shi01a] Shirokanev, A. S., Kirsh, D. V., Kupriyanov, A. V., Researching methods of reconstruction of three-dimensional crystal lattice from images of projections. CEUR Workshop Proceedings, Vol. 1490, pp. 290-297, 2015.

[Smi01a] Smith, W., Foundations of Materials Science and Engineering. McGraw-Hill, pp. 67-107, 2004.

Applications of Immersive 3D Visualization Lab

Jacek Lebiedź
Gdańsk University of Technology, Faculty of ETI
Dept. of Intelligent Interactive Systems
ul. G. Narutowicza 11/12
80-233 Gdańsk, Poland
jacekl@eti.pg.gda.pl

Jerzy Redlarski
Gdańsk University of Technology, Faculty of ETI
Dept. of Intelligent Interactive Systems
ul. G. Narutowicza 11/12
80-233 Gdańsk, Poland
jerredla@pg.gda.pl

ABSTRACT

The paper presents Immersive 3D Visualization Lab (at the Faculty of Electronics, Telecommunications and Informatics at Gdańsk University of Technology in Poland) and its applications prepared after its launch in December 2014. The main device of the lab is a virtual reality cubic cave with six screen-walls, where a person can navigate in virtual space using different controllers. The most sophisticated controller is a freely rotating transparent sphere, supported on rollers and equipped with a motion tracking system (something like an omnidirectional hamster wheel). A person's walking motion causes the sphere to revolve and triggers changes in the computer generated 3D images on VR cave screens surrounding the sphere, thus creating an illusion of motion. The projection system is supplemented with a spatial sound generation system.

Keywords

Virtual reality, VR systems and toolkits, VR caves, walk simulation, immersion.

1. INTRODUCTION

Similar to many universities in the world, Gdańsk University of Technology in Poland decided to build their own virtual reality laboratory. The main idea was to allow users to feel immersed and to navigate freely in a computer-generated virtual world. We wanted to ensure the highest possible degree of immersion (unrestricted freedom of movement and stereoscopic 3D projection) while reducing the amount of equipment worn by users (e.g. virtual reality helmets) to provide them with maximum comfort and impression of natural activity.

Virtual reality is often described as I^3 or $3 \times I$: interaction, immersion, and imagination [1]. Interaction relates to the input devices, immersion – to the output devices. However imagination does not concern the devices, it is associated with the user. He or she should not be distracted by the physical reality.

In the case of virtual reality caves [3, 4, 8, 15], interaction is achieved with navigation controllers (from wands to locomotion simulators), immersion – via surround audio-visual projection.

Virtual reality caves are intended to generate scenes and situations that look very realistic and to allow easy navigation through them. Moreover the sophisticated locomotion interface allows us to move through the virtual world in a natural way. Thus, VR caves may be used in applications such as [11, 12]:

- military/rescue training and operations,
- industrial inspection training,
- scientific and architectural visualization,
- virtual tourism, exhibitions and museums,
- analysis of human behavior,
- phobia treatment,
- entertainment (e.g. computer games),
- telepresence.

Projection in virtual reality caves takes place onto flat screens arranged in the form of a cuboid. The majority of VR caves have only four screens: three vertical walls and a horizontal floor. This allows a user to easily access the interior of a VR cave. For full immersion though, a VR cave should have six rectangular screen-walls (one as a gate). To provide users with the ability to walk on their own feet through the virtual scene, without changing physical

location, a number of devices have been proposed [2, 5, 9]. One such device is a rotary sphere with the user in the middle [6, 7, 14, 17]. Its transparent surface allows us to use it inside the VR cave and see a virtual scene (projected onto the screens) without need to wear an uncomfortable VR helmet.

2. IMMERSIVE 3D VISUALIZATION

The Immerse 3D Visualization Lab contains a virtual reality cave in the form of a closed cube with edges of about 3.4 meters each. It consists of six square screens: four vertical walls, a horizontal floor and a ceiling. All screens are two centimeters thick acrylic plates with special coating for highest brightness uniformity. The floor is strengthened with an eight centimeters thick glass plate with a carrying capacity equal to 700 kg (7 persons). To allow users entrance to the VR cave, one of the screen-walls is an automatic sliding gate (Fig. 1).

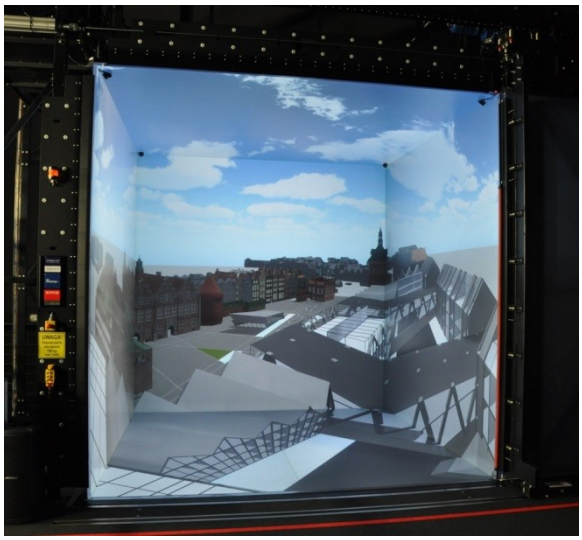


Figure 1. VR cave in I3DVL (with open gate).

The 360 degree view is achieved by the stereoscopic rear projection onto all the six flat screen-faces of the cube [12, 13]. It requires projection from six different directions. In order to cover each whole square screen two Barco Galaxy NW7 projectors with the use of the edge blending technique are applied per screen.

Twelve three-chip DLP digital projectors, with WUXGA resolution (1920 × 1200 pixels) and 7000 ANSI lumens luminous flux each, produce six images with the final resolution equal to 1920 × 1920 pixels and the total luminous flux above 11000 ANSI lumens each (taking into account the losses caused by the edge blending). The square pixel side length equal to 1.77 mm implies the 0.06 degree (3.6 minutes of arc) size of pixel seen form the center of the VR cave [10, 12, 13].

The Barco Galaxy projectors can use two alternative methods of stereoscopy: separation in time with active shutter glasses (nVidia 3DVision Pro) and spectrum channels separation using passive glasses with selective interference filters (active Infitec+). Both methods need a frequency of 120 Hz (i.e. 60 Hz per eye).

The dozen Barco Galaxy projectors are driven by 14 Dell Precision T3600 computers with nVidia Quadro K5000 GPU each. Twelve of these computers generate images directly for the projectors, the remaining two are in charge of synchronization, tracking done by ART IR-optical tracking system and the generation of sound for Bose and Apart speakers. The computers are connected by two networks: 40 Gb/s InfiniBand and 1 Gb/s Ethernet.

The computers of the virtual reality cave can communicate directly, through the fast optical fiber InfiniBand connection, with the high performance Academic Computer Center in Gdańsk (CI TASK) [18]. It allows for real time visualization of very complex calculations (e.g. physical simulations or artificial intelligence).

A user in the middle of the virtual reality cave may navigate using different controllers. The simplest one is a handheld wand (flystick), where the user steers movement with a joystick and buttons. More sophisticated controllers allow the user to navigate by his own steps.

The Immerse 3D Visualization Lab contains the spherical walk simulator Virtusphere [17] that has the form of a transparent (openwork) sphere rotating on rollers. It can be inserted into the center of the virtual reality cave (Fig. 2) and is entered afterwards through a manhole. The user sees images that are projected onto the screens surrounding the rotary transparent sphere. The diameter of the sphere is equal to 3.05 m and provides a sufficiently large radius of curvature, therefore the abnormality of movement will be imperceptible or at acceptable levels [10, 12]. Moreover, the user's head is almost in the geometrical center of the sphere.

Figure 2. Spherical walk simulator in the middle of VR cave in I3DVL (with open gate).

Instead of the rotary sphere other walk simulators (but lighter than 600 kg) can be used. This will allow other models of locomotion interfaces to be tested, such as improved rotary spheres or sliding surfaces with the mechanism of placing feet centrally (e.g. Virtuix Omni [16]). It is worth adding that all these walk simulators (including spherical) can work without the VR cave, just with a virtual reality helmet.

3. APPLICATIONS

The unique combination of hardware in Immersive 3D Visualization Lab makes it possible to execute different simulations. The first year of the lab's operation yielded numerous VR applications: some of them are results of applied research, others have been prepared during student projects.

Military and rescue training and operations

It is much cheaper and safer to simulate battlefields and rescue operations (such as fires, floods, earthquakes) with virtual reality, as opposed to real training grounds. Infantrymen, firemen and other specialists can train the procedures using virtual reality. I3DVL has not yet had the opportunity to cooperate with uniformed services, though we do have a number of indoor and outdoor locations modeled. Some of them include scripted events, objects that can be interacted with, as well as AI controlled agents. One example is a student's application where one can attempt to catch horses on an enclosed meadow (Fig. 3).

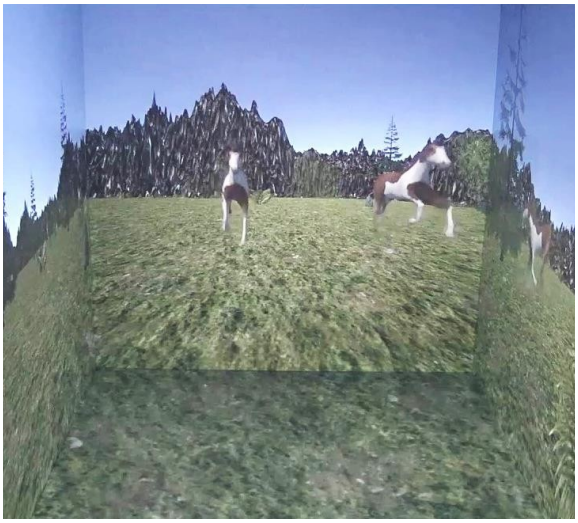


Figure 3. Virtual horse catching simulation.

Industrial inspection training

Similarly, virtual inspection of ships or buildings is safer and less expensive than on physical training grounds. With the cooperation of one leading international certification body and classification society (DNV GL), I3DVL began work on the adaptation of its ship inspection simulator to the VR cave. At the time of this writing, it is possible to view a 3D boat model and streamlines of airflow simulated beforehand around sailboats (Fig. 4). Work on a larger vessel is under way.

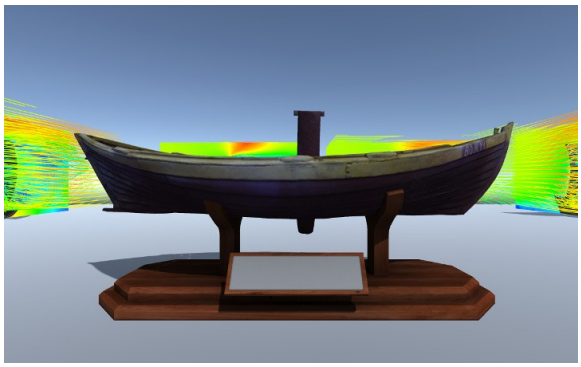


Figure 4. Ship sightseeing simulator.

Scientific visualization

Virtual reality provides very useful tools for different scientific applications. For example, chemists from our university requested an application that allows for visualization and manipulation of chemical structures (e.g. proteins, enzymes). At the moment, an application is being developed for the presentation of 3D structures described in the Protein Data Bank (PDB) format (Fig. 5).

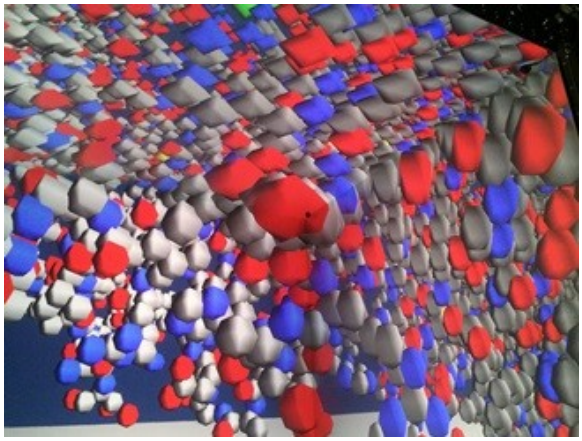


Figure 5. Protein Data Bank structure presenter.

Architectural visualization

The Coal Market Square is a historical part of Gdańsk. The area is being redesigned and a new gallery is planned. Architects have come up with many designs, three of which have been modeled. I3DVL allows the governing decision makers to see the submissions for themselves and walk through virtual galleries and the Coal Market Square itself (Fig. 6). It is possible to switch between day and night, move freely or choose a point of view such as from the roof of one of the buildings, or start a virtual tour along a premade route.

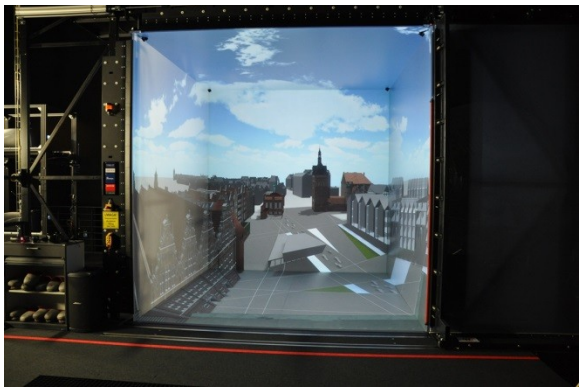


Figure 6. The virtual Coal Market Square.

Virtual tourism, exhibitions & museums

A masterpiece created by eighteenth century craftsmen in Gdańsk, the Amber Room, was considered one of the most valuable works of art ever created in Europe. Until it's mysterious disappearance during World War II, it's history took place in Berlin, Tsarskoye Selo (for over two centuries) and finally Königsberg. It took 24 years between 1979 and 2003, to reconstruct the room in Tsarskoye Selo, although there was a major division between the craftsmen over the question whether the new room should be artificially aged, with the proponents ultimately winning. In I3DVL we can admire the Amber Room as an example of immersive virtual tourism (Fig. 7). This also shows how the laboratory could be used by museums to plan and test expositions before building them, or by artists who might want to see how their sculpture or other form of art might look in a specific room, or even how changing the different parameters (such as material or age) would affect the final look.

Figure 7. The virtual Amber Room.

Analysis of human behavior

Observing how people react in an immersive virtual environment can provide a lot of useful information. One could for example present planned products to customers and assess their attractiveness from their reactions. A group of psychology students from University of Social Sciences and Humanities developed scenarios for a game that should help in the treatment of acrophobia (fear of heights). Two applications were developed by students, one where a person can ascend high skyscrapers to measure at what heights he begins to feel uncomfortable. The other included coins placed on rooftops, ledges and glass bridges and measured fear of height by the number of coins a user was able to collect (Fig. 8). This could provide psychologists with valuable insight. While final versions of this software prevented people from actually falling off the edges, the ones that allowed it showed varying levels of reaction from people, with some people largely unaffected, while others experienced strong emotions.

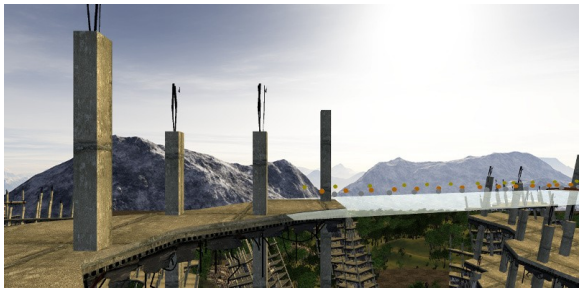


Figure 8. Acrophobia treatment by gamification.

Phobia treatment

A number of anxiety disorders and phobias can be treated through controlled exposure to the real phobic stimuli (flooding technique) or to its image in the patient's brain (implosive therapy). In I3DVL, virtual stimuli can be used instead or real or imagined ones. For example, arachnophobia can be treated by controlled exposure to a virtual spider in different forms and situations (Fig. 9), peristerophobia – to a virtual pigeon, glassophobia – to a virtual crowd, acrophobia – to a virtual height. Wandering on virtual ledges and floors of a skyscraper allows a patient to overcome their fear of heights.

Figure 9. Virtual implosive therapy.

Entertainment – computer games

I3DVL is a very attractive platform for the development of games focused on movement through different terrains. We present a game made by students, taking place in an old hospital, where a player escapes from zombies (Fig. 10). The main focus of this game was to test different navigation schemes. For example, the forward direction could be determined based on the direction a player is looking, or the direction he is pointing the controller. Another option was to use the player's position in the VR cave (i.e. how much off-center they are and in what direction).

Figure 10. Escape from zombies.

Telepresence

I3DVL can display not only a virtual, computer-generated world, but also photos and video recordings (Fig. 11). This can also be done in real time, allowing users to feel as if present in a distant, possibly dangerous, place. Telepresence could be used to control robots, giving the operator 360 degree view (assuming the robot has appropriately placed cameras). This would be immensely useful for operations in dangerous environments, but also any other situation where the user cannot be physically present at the chosen location.

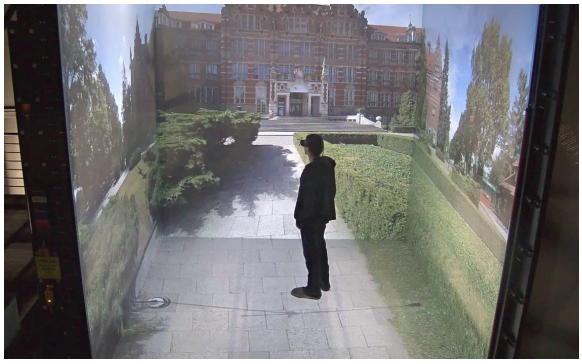


Figure 11. Real scene projected onto cave screens.

4. DIRECTION OF DEVELOPMENT

Currently, the laboratory is expanding in a number of ways, such as installation of new hardware and development of new software. For example, enabling more users to interact at the same time with the virtual world will require both hardware and software solutions.

Hardware development

New equipment will allow developers (e.g. students) to test their projects on a smaller VR cave with only three walls and a floor, so that the main VR cave is only used when projects are nearing completion. Initial programming work can be implemented on single-computer stations, or a 4-computer mini VR cave built on the basis of four monitors. All the computers controlling these devices have identical hardware and software configuration. The clusters supporting the VR caves have the same network architecture (InfiniBand and Ethernet). This solution allows developers to use scalability rules during development of virtual reality applications.

Interaction development

The VR caves will be equipped with more wands and markers, which will allow extensive interaction and accurate tracking of groups of users and their gestures. Further improvements are planned to make movement in the spherical walk simulator more natural. We are also looking for ways to improve the collisions with virtual obstacles – it is easy to detect the collisions thanks to the user's position given by tracking devices, but responding to collisions is problematic. If a user tries to walk into a virtual wall (by physically stepping forward), we can't stop him – we can either allow him to walk through the wall, or we can push the wall away, but both solutions break immersion. The virtual amber room worked around the problem by resizing the room to the exact size of the CAVE, thus placing virtual walls where the physical screens are located. The problem of reduced immersion can also be alleviated by making colliding objects appear fluid when approached closely. In addition to the above solutions, visual and audio feedback can be given to users when they collide with objects. Haptic feedback would increase immersion, but we are yet to acquire any haptic devices for our lab. Similarly, interaction with the virtual world may seem less or more natural depending on the controllers used and how the application reacts to the inputs, so there's plenty of work that can still be done to find solutions that work best for the given scenarios.

Software development

Development of software applications for the VR caves requires appropriate development frameworks. Dedicated tools are very expensive (e.g. VBS), so we are negotiating conditions for obtaining them for educational and research purposes. On the other hand, standard environments for game development (e.g. Unity) tend to be useful, especially since their cost is relatively low, but they require specialized

libraries for synchronization of image generation on each VR cave screen. Despite the fact that we already have a simple library for Unity projects, we are working on a new, improved library and framework that will be more robust and extensible. Additionally, we want to expand it to support connections with other VR devices (caves, helmets, vehicle and flight simulators etc.) to manage distributed VR simulations. The framework should also support usage of the high performance cluster (CI TASK) to execute complex calculations (nontrivial physics, advanced artificial intelligence, prediction of user behavior etc.) Once our library is ready for use we will also be able to make meaningful performance measurments to see the full capabilities of our CAVE's hardware (the temporary Unity library we're using experiences major slowdowns with stereoscopy active).

5. CONCLUSIONS

Immersive 3D Visualization Lab is a complete virtual reality environment. The greatest advantage of this laboratory is the possibility to walk through a computer-generated virtual world without limitations. This creates endless capabilities to implement scenarios for a wide range of practical applications. The first year of I3DVL's operation proves that virtual reality is a powerful tool for different human activities.

6. ACKNOWLEDGEMENTS

We would like to thank our students for developing very interesting simulations in I3DVL.

7. REFERENCES

[1] G. Burdea, P. Coiffet: *Virtual Reality Technology*, 2nd Ed., Wiley, New York 2003.

[2] H. P. Crowell III, J. A. Faughn, P. K. Tran, P. W. Wiley: Improvements in the Omni-Directional Treadmill: *Summary report and Recommendations for Future Development*. Army Research Laboratory, ARL-TR-3958, 2006.

[3] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti: The CAVE: A Virtual Reality Theater. *HPCCV Publications*, 2, 1992.

[4] T. A. De Fanti et al.: The future of the CAVE. *Central European Journal of Engineering*. November 2010.

[5] A. De Luca, R. Mattone, P.R. Giordano: The Motion Control Problem for the CyberCarpet, *Proc. of the 2006 IEEE International Conf. on Robotics and Automation*, Orlando, 2006.

[6] K. J. Fernandes, V. Raja, J. Eyre: Cybersphere: The Fully Immersive Spherical Projection System. *Communications of the ACM*, Vol. 46, No. 9ve, pp. 141-146, September 2003.

[7] J. Gantenberg, K. Schill, C. Zetzsche: Exploring Virtual Worlds in a Computerised Hamster Wheel. *German Research* 1/2012.

[8] Iowa State University News Service: *The most realistic virtual reality room in the world*. 2006.

[9] S. "Cb" Kuntz: *A VR Geek Blog*. http://cb.nowan.net/blog/state-of-vr/state-of-vr-devices/.

[10] J. Lebiedź, J. Łubiński, A. Mazikowski: An Immersive 3D Visualization Laboratory Concept. *Proc. of the 2nd Int. Conf. on Information Technology – Information Technologies*, Vol. 18, pp. 117-120, 2010.

[11] J. Lebiedź, A. Mazikowski: Launch of the Immersive 3D Visualization Laboratory. *Szybkobieżne Pojazdy Gąsienicowe*, vol. 34, nr 1, 2014, pp. 49-56.

[12] J. Lebiedź, A. Mazikowski: Innovative Solutions for Immersive 3D Visualization Laboratory. *22nd International Conference on Computer Graphics, Visualization and Computer Vision WSCG 2014 – Communication Papers Proceedings* (ed. Vaclav Skala), Plzeň 2014, p. 315-319, 2014.

[13] A. Mazikowski, J. Lebiedź: Image projection in Immersive 3D Visualization Laboratory. *18th International Conference in Knowledge Based and Intelligent Information and Engineering Systems KES 2014, Gdynia 2014, Procedia Computer Science* 35, 2014, p. 842-850.

[14] E. Medina, R. Fruland, S. Weghorst: Virtusphere: Walking in a Human Size VR Hamster Ball, *Proc. of the Human Factors and Ergonomics Society Annual Meeting* 52, No. 27, pp. 2102-2106, 2008.

[15] A. Mitchell: Lost in l-Space. *INAVATE IML*, 10 2012, pp. 3÷4. http://www.inAVateonthenet.net .

[16] Virtuix Inc.: *Virtuix Omni. Enter the future of gaming*. http://www.virtuix.com/ .

[17] VirtuSphere Inc.: Virtusphere. *The Virtual World Immence. Let's Immerse Together*. http://www.virtusphere.com/index.html .

[18] Ł. Wiszniewski and T. Ziółkowski: Real-Time Connection between Immerse 3D Visualization Laboratory and KASKADA platform. *TASK Quarterly* 19, No 4, 2015.

Estimating landmarks on 2D images of beetle mandibles

Lê Vãnh L.
ITDLU
Dalat University
Vietnam
linhlv@dlu.edu.vn

Beurton-Aimar M.
LaBRI- CNRS 5800
Bordeaux University
33400 Talence - F.
beurton@labri.fr

Salmon J.-P.
LaBRI- CNRS 5800
Bordeaux University
33400 Talence - F.
jean-pierre.salmon@labri.fr

Marie A.
IGEPP
INRA 1349
35653 Le Rheu - F.

Parisey N.
IGEPP
INRA 1349
35653 Le Rheu - F.
alexia.marie/nparisey @rennes.inra.fr

ABSTRACT

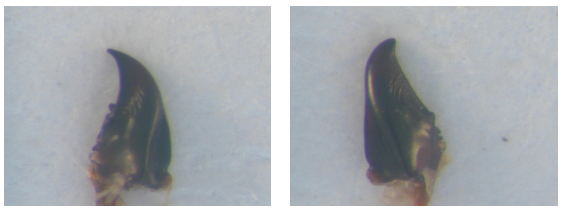
Studying links between phenotype/genotype and agricultural practices is one of the main topics in agronomy research. Phenotypes can be characterized by informations like age, sex of animals/plants and more and more often with the help of image analysis of their morphology. From now, getting good quality of images for numerous individuals is easy but that leads to design automatic procedures to replace manual exploration of such amount of images. Several bottlenecks have been identified to analyze automatically images. One of them is segmentation of selected area and/or shapes, and another well-known one is setting automatically morphometric landmarks. Landmarks are points on the object which can be used to identify or to classify the objects.
It exists a lot of methods to experiment landmarks setting, depending on the image contents. This work has been initiated by using the article of Palaniswamy et al. *"Automatic identification of landmarks in digital images"*[6]. They proposed a method based on calculus of a probabilistic Hough transform coupling to a template matching algorithm. They applied their method to the Drosophilia wings. In our study, we have gotten a set of 291 beetles . For each one 2D images of 5 different parts of their anatomy have been taken: mandibles left and right, head, pronotum and elytra. The first part of the project was to test how the Palaniswamy's method could be used to analyze them. We have implemented all the required algorithms to compute positions of mandibles landmarks and compared the obtained results to landmarks which have been manually set by biologists. We will see that even positions automatically obtained are not fully precised, if we used centroid size to characterize mandibles, the size computed from automatic landmarks is closed to this one computed from the manual ones. Future works will focus on definition of a semi-landmarks procedure which would add some features as the measure of the curve between two landmarks.

Keywords

Landmarks identification, probabilistic Hough transform, morphometry of beetle mandible.

1 INTRODUCTION

Morphology analysis is a way to characterize biological shape variations. In the aim to study potential links between these variations and agricultural ecosystems, a set of 291 beetles has been collected. Informations as sex, place where they were found and agricultural practices in this field were set. To grow richer phenotype data, morphometric operations could be done. To do that, a set of landmarks has been defined. Morphometric landmarks are points that can be defined in all specimens and located precisely [5]. Landmarks are widely used in many biological studying and analysis of geometric characteristics are currently included into classification procedures.



(a) Right mandible (b) Left mandible

Figure 1: The mandibles of beetle

In this paper, we focus on a method which addresses automatic identification of landmarks in digital images. Palaniswamy et al. [6] have proposed a method to set landmarks on images of Drosophila wings. We have investigated how this method can be implemented to work on images of beetle mandibles (figure 1). The method contains four stages: a features extraction of mandible structure (segmentation stage), a recording of the features using pairwise geometric histogram (PGH),

an estimation of the landmarks positions using Probabilistic Hough Transform (PHT) and finally a refinement of the estimated landmarks by cross-correlation.

2 METHODS

For each mandible image, a set of 18 landmarks have been manually set by biologists corresponding to morphological points of interest (see figure 2). It will constitute our ground truth. The automatic procedure to estimate these positions extracts features by analyzing the image histogram firstly. The obtained parameters are then used to approximate edges of the mandible by line segments. These edges are presented to PGH using geometric relationships between them. The shape correspondence is determined by comparing the PGHs of model and scene data. A PHT is then used to identify hypothetical location of model landmarks on scene image. Finally, the hypothetical landmarks are performed by template matching. We now describe in details all these steps.

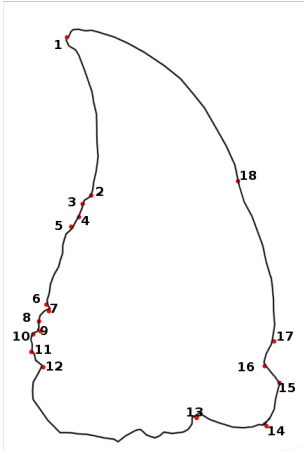


Figure 2: Manual landmarks of the right mandible

Segmentation step

Usual way to obtain automatically threshold value for segmentation is to take a look to the image's histogram (figure 5). In our case, per image we have only one object, the mandible, into a pretty uniform background, consequently the histogram exhibits only two picks. In this case, the retained threshold value is the average value (blue point in fig.3) between two mean values (red and green points in fig.3) of these two pick regions. The first region, begins from the beginning to the median of histogram and the second region is the rest.

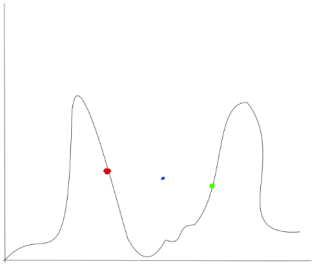


Figure 3: Image Histogram

The Canny algorithm [4] is one of the relevant algorithms to detect segmentation edges. The result is a list of points for each retrieved edge. To compute the PGH another kind of geometric form, lines, is needed. Extraction of approximated lines from the list of points can be achieved by using the recursive algorithm [3] as below:

- Create a line connected by two edge endpoints
- For each point in the edge :
  - Calculate perpendicular distance to the line
  - Keep the point at max distance, i.e. max point
  - Divide edge at max point into two parts:
  - Repeat with the two new parts of the edge.

The algorithm stops when the edge cannot be broken more. Concretely, we stop the algorithm when the maximum perpendicular distance of max point is less than 3 pixels, i.e. enough small to create an approximated line.

Comparison between model and scene

To determine the correspondence between the model and the scene image, we compute the PGH[2] using the approximated lines of each image. The steps to construct the PGH as follows:

- Create a PGH matrix,
- Choose a reference line,
- For each other lines in the shape,
  - Calculating the perpendicular distance from two endpoints to the reference line,
  - Computing the angle between the considered line and the reference line,
  - Recording the perpendicular distance and angle into the matrix.
- Repeat step 2 (choose reference line) to all the lines in the shape considered as reference lines,

The algorithm stop when all lines of shape have been considered as reference line.
To be able to compare model and scene, a similarity metric is needed. The Bhattacharya[6] similarity metric is used to compare the distribution (PGH) for the model and the scene data. It computes the degree of match between them as a dot product correlation of the PGHs (equation 1).

$$d_{Bhatt}(H_iH_j) = \sum_{\theta}^{\pi} \sum_{d}^{d_{max}} \sqrt{H_i(\theta,d)H_j(\theta,d)} \qquad (1)$$

Where $H_i(\theta,d)$ is an entry at row θ (i.e. angle) and column d (i.e. perpendicular distance) in the PGH of the image i.

Selection of matching points

The Probabilistic Hough Transform (PHT) is then used to determined the presence and location of the model in the scene image, as well as to determine the hypothesis of the model landmarks in the scene image[1]. Applying PHT includes two steps: first, we find the pair of scene lines that similar with a pair of model lines (named training process); second, we estimate the model landmarks in the scene image.

Training process includes the duration to construct the reference table for model image and process to find the similar pair of lines between model and scene image. The steps as follows:

- Create the reference table,
  - Choose an arbitrary point in the model,
  - Create a table to record the information,
  - For each pair of model lines, calculate the perpendicular distance and angle from each line to the point and save into the table.
- Create an accumulator (a two dimension matrix (angle and perpendicular distance)),
- For each pair of scene lines, find the pair of model lines within correspondence in position, orientation and scale. Select the respective value in reference table,
- Increase the value in accumulator at respective position and keep the cell that have the maximum value.

The pair of scene lines having the best value is chosen. The estimated landmarks in the scene obtained by calculating the relatedness between the model's reference point and the model's landmarks are recorded. Besides, we also record the difference angle between model image and the scene image. Fig. 4 shows an example of result, the red points are estimated landmarks on the scene mandible (right one) from a model mandible (left one) landmarks.

(a) The model image (b) The scene image

Figure 4: The estimated landmarks by PHT

Template matching

The template matching is process to verify the landmarks estimation provided in the PHG stage. Cross-correlation method is hired for this work. By sliding the template on image by each pixel, cross-correlation will detect the best similarity between model and scene image. The progress of template matching as follows:

- Rotate the scene image (the angle has indicated by PHT),
- Create a bounding box around a model manual landmark (in model image),
- Create a bounding box around a estimated landmark (in scene image),
- Apply cross-correlation between the two bounding boxes.

The template matching finishes when all estimated landmarks are refined. Fig 5 shows a complete result on one scene mandible with the segmentation (red lines), manual landmarks (yellow points) and estimated landmarks (green points).

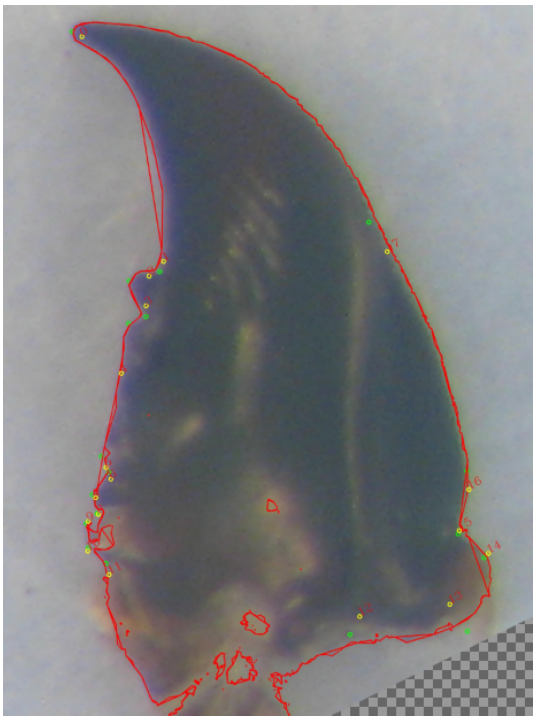


Figure 5: Automated landmarks in scene image after refining

3 EXPERIMENTS AND RESULTS

All the algorithms have been implemented in a framework MaeLab in C++ language[1]. The set of beetles images have been analyzed, right mandibles have been first studied. After verification of the image correctness, it remains 288 usable images. From the 3 images removed, 2 do not contain mandible and in the last one, the mandible is broken in 2 parts. All valid images have been segmented and the 18 landmarks have been set for each. Biologists have chosen to use in a first attempt the centroid size to measure the mandible. This size is obtained by determination of the centroid of the mandible and by sum of all square distances between each landmark and the centroid (see [5] for details).

[1] MaeLab is a free software, it can be directly obtained by request to the authors.

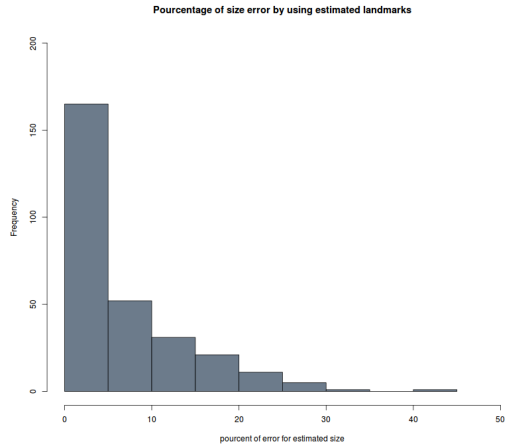


Figure 6: Percentage of error in computing centroid size from estimated landmarks

In that way, we have compared the size computed from manual landmarks and this one from estimated landmarks. The percentage of errors has been evaluated as below:

$$PercentOfErr = \frac{100 \star |(OriginalSize - EstimatedSize)|}{OriginalSize}$$

We can observe in fig. 6 that for more than 150 images, the error is less than 5%. Only 2 mandibles could be considered as wrongly measured with the estimated landmarks and exhibit more than 30% of errors. Finally 90% of images have less than 10% of error in their size computing and for which we can consider estimated landmarks as good enough to replace manual landmarks.

Perspectives and future works

Of course, centroid size is not the only feature we want to consider. It is also possible to compare image per image the exact position of manual and estimated landmarks, for example if we want to work with semi-landmarks by adding of curve measure between 2 landmarks.

In our case, the landmark couples 1 and 2 or 1 and 17 (figure 2) are good candidates to play this role. Figure **??** shows for one mandible the results which have been obtained for each landmark. What one can note is that for some of them, an offset appears. For example

4 CONCLUSION

Morphometric analysis is a powerful tool in biology in order to characterize species. Unfortunately, setting landmarks to run such analysis is time consuming and difficult to replicate through different experiments. In this project we have begun to design set of procedures to segment 288 beetle mandibles and to identify automatically landmarks which have been described by biologists. Each mandible is segmented by computing a approximated lines set. Using the Probabilistic Hough Transform method, these lines are used to align all mandibles scenes with one mandible model. The first results shows that in order to compute the mandible centroid size, the estimated landmarks are accurately enough. A framework in C++ language has been developed to facilitate the using to biologists. From now, a next stage of this studying is to add features as measure of curves, in that way the landmark positions have to be set more precisely. To solve this problem, algorithms based on design of shape skeleton will be tried.

5 REFERENCES

[1] Ashbrook A., Thacker N. A, Rockett P., and CI Brown. Robust recognition of scaled shapes using pairwise geometric histograms. In *BMVC*, volume 95, pages 503–512, 1995.

[2] Evans A., Thacker N. A, and Mayhew J. EW. The use of geometric histograms for model-based object recognition. In *BMVC*, volume 93, pages 429–438. Citeseer, 1993.

[3] Thacker N. A, Riocreux PA, and Yates RB. Assessing the completeness properties of pairwise geometric histograms. *Image and Vision Computing*, 13(5):423–429, 1995.

[4] Canny J. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):679–698, 1986.

[5] Webster. M. and Sheets. H.D. *Quantitative Methods in PaleoBiology*, chapter A Practical Introduction to Landmark-based Geometric Morphometrics, pages 163–188. J.Alroy and G. Hunt, 2010.

[6] Palaniswamy S., Thacker N.l A, and Klingenberg C. P. Automatic identification of landmarks in digital images. *IET Computer Vision*, 4(4):247–260, 2010.

ISSN 2464-4617 (print) WSCG 2016 - 24th Conference on Computer Graphics, Visualization and Computer Vision 2016
ISSN 2464-4625 (CD-ROM)

Resilience of Luminance based Liveness Tests under Attacks with Processed Imposter Images

Luma Omar
Durham University
School of Eng. & Comp. Sciences
South Road, Durham, DH1 3LE, UK
luma.omar@durham.ac.uk

Ioannis Ivrissimtzis
Durham University
School of Eng. & Comp. Sciences
South Road, Durham, DH1 3LE, UK
ioannis.ivrissimtzis@durham.ac.uk

ABSTRACT

Liveness tests are techniques employed by face recognition authentication systems, aiming at verifying that a live face rather than a photo is standing in front of the system camera. In this paper, we study the resilience of a standard liveness test under imposter photo attacks, under the additional assumption that the photos used in the attack may have been processed by common image processing operations such as sharpening, smoothing and corruption with salt and pepper noise. The results verify and quantify the claim that this type of liveness tests rely on the imposter photo images being less sharp than live face images.

Keywords

Liveness tests; face recognition; luminance models; difference of Gaussians; logistic regression

1 INTRODUCTION

Compared to the other main biometric authentication methods which are based on fingerprints or high resolution iris images, face recognition has the unique characteristic that it is based on data that can easily be found in the public domain. For example, in many cases can be very easy to obtain a photo of someone's face, either doing a quick online search or by logging into a social network. As a result, face recognition based authentication is particularly vulnerable to imposter attacks, when, for example, an attacker holds someone's photo in front of the camera trying to gain access through the face recognition system.

To counter such concerns, *liveness tests* are binary classification algorithms aiming at determining whether the recognized face is a live face, or a photo or video played in front of the system's camera. Developing accurate liveness tests is a challenging task and they often require use of specialized hardware, such as infrared cameras. In [10] several commercial user authentication systems that do not use additional hardware to support liveness tests were evaluated and they were found vulnerable to even very crude imposter image attacks.

The research on liveness tests that do not rely on specialized hardware is motivated by the desire to have secure face recognition based authentication on machines of everyday use, such low end laptops and smartphones. The current state of the art, such as the Tan et al. paper [16], is based on machine learning algorithms trained to distinguish between images of live faces taken by the face recognition system and images of photos fed to the system by the imposters. In particular, it has been established that the different reflectance properties of these two categories of can lead to the development of an effective liveness test.

In this paper we study the effect of image processing operations applied on the imposter images on the performance of a variant of the Tan et al. test based on differences of Gaussians and sparse logistic regression. As a possible explanation of why their algorithm is effective, Tan et al. observe that images of face photos fed into the system by imposters tend to be smoother, as they lack detail. The main contribution of our paper is a verification and quantification of this claim by processing the imposter images and measuring the effect on the performance of their algorithm. As expected, the sharpening of the imposter images reduced the accuracy of the liveness test, while the smoothing of the imposter images increased it.

The main limitation of our paper is that we process the imposter images of an existing database (NUAA), which are images of a photo of the subject, rather than processing the photos of the subject and then taking photos of them. While this approach allows for a better quantitative understanding of the basic principle underlying the Tan et al. algorithm, we note that we have not yet measured the effect of a direct attack consisting of processing the photo of the subject and feeding it into the system.

2 RELATED WORK

As biometrics based security applications become increasingly popular, the study of their vulnerabilities and the development of countermeasures is becoming a research topic of current interest. Attacks on face recognition systems fall into two main categories; direct and

indirect attacks. Direct attacks rely on the use of stolen biometric data of some form; digital images displayed on a screen, printed photos, or gummy fingerprints. Pan et al. [11] classified direct attacks into three categories: a photograph of the real user is used; a video; or a 3D model. One particular strength of direct attacks is that they do not require knowledge of the face recognition system they attack.

Indirect attacks use algorithms to construct an input face that will gain entry to the biometric security system. They presuppose some knowledge of the attacked system and a certain level of information leakage from it. Martinez et al. [9] uses similarities scores assumed to be outputted by the face recognition system and the hill climbing technique to construct an image giving access to the system. Galbally et al. [3] tested the vulnerability of a Principal Component Analysis (PCA) based face recognition system against a Bayesian hill climbing attack algorithm and reported a 85% success rate for such attacks. In [5], two face recognition systems using a Gaussian Mixture Model (GMM) and PCA, respectively, were tested against a Hill Climbing indirect attack and were both found vulnerable, with the GMM system being nevertheless the more robust.

Liveness tests are binary classification algorithms developed as countermeasures to the imposter attacks to distinguish between live faces and imposter images, video or 3D models. Robust liveness tests that do not require the use of any specialized hardware have been developed based on the observation that the high frequency components of the imposter images are weaker than those of the live faces.

The algorithm in Tan et al. [16] we study in this paper is based on this observation and will be discussed in more detail in Section 3. Peixoto et al. [13] further improved the Tan et al. algorithm by addressing limitations related to bad illumination conditions. Another anti-spoofing approach proposed by Komulainen et al. [7] uses support vector machine classification of histograms of gradient descriptors. Maatta et al. [8] uses Local Binary Patterns to analyze the local texture of the face and the resulting single feature histograms are classified with Support Vector Machines. Galbally et al. [4] use image quality metrics and classify the lower quality images as imposter.

Other approaches to rely on biometric motion analysis, focusing on different types of motion such as: head tilting [2], mouth movement [6] and eye-blinking [12]. Foreground and background motion correlation is used in [1]. Finally, the accuracy rates of liveness tests can be boosted with the use of specialized hardware. Socolinsky et al. [14] analyzed face thermograms acquired by a thermal imaging camera, while Steiner et al. [15] recently proposed a liveness test based on the analysis of the spectral signatures in the infrared.

3 IMPLEMENTATION

In [16], Tan et al. proposed a series of liveness tests where information sensitive to the reflectance properties of the scene is extracted from the image and it is used to train a binary classifier so that it can distinguish between images of live faces and images of photos of faces. The variant we implemented extracts a *difference of Gaussians* from the image and uses it to train a sparse logistic regression classifier. The implementation was done in Matlab and the *SLEP* package was used for the sparse logistic regression.

Following the recommendation in [16], we smooth the image using Gaussian filters with $\sigma_1 = 0.5$ and $\sigma_2 = 1.0$ and compute the difference of the two smoothed images. Regarding the machine learning part of the algorithm, following the notation and parameter choices in [16], we use the class labels $\{-1, 1\}$, where -1 corresponds to client images and 1 to imposter images and the conditional probability of the the imposter class $y = 1$ is given by

$$\text{Prob}(y|x) = \frac{1}{1+\exp(-y(w^T x + b)} \tag{1}$$

where x is the sample image, and w and b are the weight vector and the intercept. To avoid overfitting, the values of w and b are computed through the minimization of the cost function

$$\min_{w,b} \text{loss}(w,b) + \lambda \|w\|_1 \tag{2}$$

where λ is a user defined constant favoring sparse weight vectors and loss given by

$$\text{loss}(w,b) = \frac{1}{m}\sum_{i=1}^{m} \log(1+\exp(-y_i(w^T x_i + b))) \tag{3}$$

where m is the size of the training set of images x_i with associated labels y_i.

The choice of λ has a significant effect in the performance of the algorithm and depends on the size of the training set. In our implementation, using a training set of 1000 images we found experimentally that $\lambda = 0.25$ gives good results. Figure 2 (top) shows the ROC curves of the liveness test for several values of λ.

3.1 Experimental design

We used the NUAA Photograph Imposter Database, which contains grayscale images of 15 different subjects in various poses under different illumination conditions. The images are organized into the two categories: the *client images* which are images of live faces, and the *imposter images* which are images of photos of the subjects. The size of all images is 64×64 pixels.

Our training dataset consisted of 1000 client and imposter images. The test set consisted of several subsets,

Figure 1: Test images. From left to right: (i) client, (ii) imposter, (iii) sharpened imposter, (iv) sharpened and blurred imposter, (v) imposter with salt and pepper noise added to it.

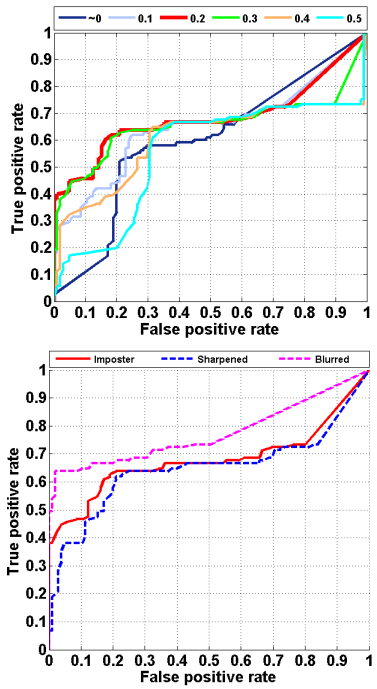


Figure 2: **Top:** ROC curves for several values of λ. **Bottom:** ROC curves for the unaltered imposter images in test set (ii); the sharpened images in (iii); and the sharpened blurred images in (vi) with $\sigma = 1.25$.

each one containing 105 images, i.e. seven from each subject. More specifically, we used test sets of:

(i) client images.

(ii) imposter images without any alteration.

(iii) imposter images sharpened by subtracting from them the response of the Laplacian filter.

(iv) imposter images sharpened with the *imsharpen* Matlab function with parameter values 0.5, 1.0, 1.5 and 2.0.

(v) the sharpened imposter images in (iii) are blurred with a Gaussian filter with $\sigma = 0.1$, 0.5, 1.25 and 2.0.

(vi) imposter images with 0.01, 0.1, 0.5 and 0.9 amount of added salt and pepper noise.

Figure 1 shows typical examples of test set images.

4 RESULTS

Figure 2 (bottom) shows the ROC curves of the liveness test when the imposter test images are either sharpened by subtracting from them the response of the Laplacian filter, or first sharpened and then blurred by a Gaussian filter. The performance decreased when the imposter images were sharpened even with the very basic sharpening algorithm we used. The performance increased when the imposter images were first sharpened and then smoothed, further demonstrating the the sharpness of the image is a key factor in distinguishing between client and imposter images.

Next we want to establish that the differences in the performance of the liveness test are commensurable with the amount of sharpening and blurring applied on the imposter images. Figure 3 (left) shows the ROC curves when the imposter images are sharpened using the *imsharpen* Matlab function which subtracts from the images a blurred version of it. The strength of the *imsharpen* command is controlled by a user defined parameter and we used values of 0.5, 1.0, 1.5 and 2.0. We notice that larger amounts of sharpening on the imposter images result into larger decreases in the performance of the liveness test. Similarly, in Figure 3 (middle) we show the ROC curves when the Laplacian filter sharpened imposter images are blurred with Gaussian filters with $\sigma = 0.1$, 0.5, 1.25 and 2.0, respectively. We notice that larger amounts of smoothing result to larger increases in the performance of the liveness test.

Finally, we experimented with the addition of various amounts of salt and pepper noise on the imposter images. This test is relevant in our context since in [10] it was shown that commercial face recognition systems can cope with large amounts of salt and pepper noise and as a consequence they are also vulnerable to imposter image attacks even when imposter images contain large amounts of salt and pepper noise. Figure 3 (right) shows the results when salt and pepper noise with probability 0.01, 0.1, 0.5 and 0.9 was added. We notice that the addition of noise increases the performance of the liveness test and the performance gain is commensurable with the amount of added noise.

5 CONCLUSIONS

In a real life situation, we should expect that an attacker will process an imposter photo before using it, increasing their chances of successfully evading a liveness test. Motivated by that observation, we evaluated the resilience of a standard luminance based liveness

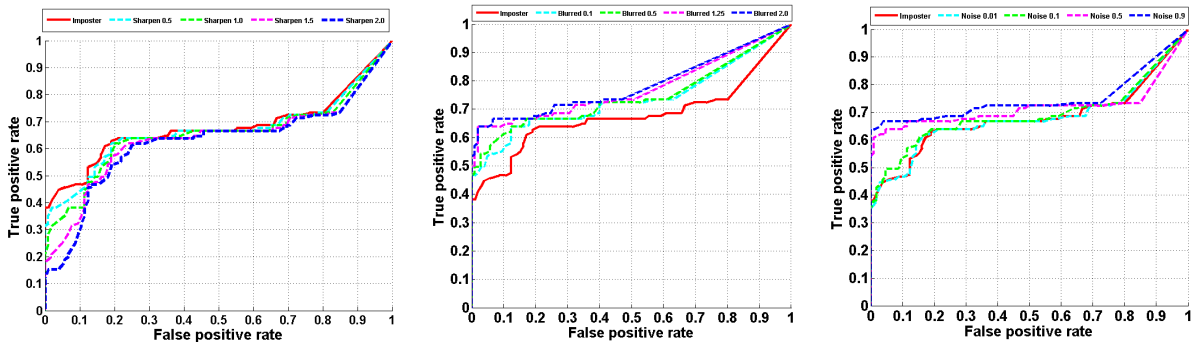


Figure 3: From left to right: ROC curves for the liveness test with different amounts of sharpening, blurring and salt and pepper noise applied on the imposter images.

test in conjunction with certain image processing operations of the imposter database. Our results verified and quantified the assumption that luminance based tests rely on the different amount of sharpness between images of live faces and imposter images. In particular, the sharpening of the imposter images decreased the accuracy of the liveness test while sharpening followed by smoothing increased accuracy rates.

While we expect that the sharpening of an imposter image before using it for an attack will result into a sharpening of the image acquired by the face recognition system, this is an assumption that still has to be verified. Thus, in the future we plan to simulate and evaluate imposter image attacks that use processed images.

6 REFERENCES

[1] A. Anjos, M.M. Chakka, and S. Marcel. Motion-based counter-measures to photo attacks in face recognition. *IET Biometrics*, 3(3):147–158, 2014.

[2] W. Bao, H. Li, N. Li, and W. Jiang. A liveness detection method for face recognition based on optical flow field. In *IEEE IASP*, pages 233–236, 2009.

[3] J. Galbally, J. Fierrez, J. Ortega-Garcia, C. McCool, and S. Marcel. Hill-climbing attack to an eigenface-based face verification system. In *BIdS*, pages 1–6, 2009.

[4] J. Galbally, S. Marcel, and J. Fierrez. Image quality assessment for fake biometric detection: Application to iris, fingerprint, and face recognition. *IEEE Trans. on Image Processing*, 23(2):710–724, 2014.

[5] J. Galbally, C. McCool, J. Fierrez, S. Marcel, and J. Ortega-Garcia. On the vulnerability of face verification systems to hill-climbing attacks. *Pattern Recognition*, 43(3):1027–1038, 2010.

[6] K. Kollreider, H. Fronthaler, M.I. Faraj, and J. Bigun. Real-time face detection and motion analysis with application in "liveness" assessment. *IEEE Trans. on Information Forensics and Security*, 2(3):548–558, 2007.

[7] J. Komulainen, A. Hadid, and M. Pietikainen. Context based face anti-spoofing. In *IEEE BTAS*, pages 1–8, 2013.

[8] J. Maatta, A. Hadid, and M. Pietikainen. Face spoofing detection from single images using micro-texture analysis. In *IEEE IJCB*, pages 1–7, 2011.

[9] M. Martinez-Diaz, J. Fierrez, J. Galbally, and J. Ortega-Garcia. An evaluation of indirect attacks and countermeasures in fingerprint verification systems. *Pattern Recognition Letters*, 32(12):1643–1651, 2011.

[10] L. Omar and I. Ivrissimtzis. Evaluating the resilience of face recognition systems against malicious attacks. In *BMVW*, 2015.

[11] G. Pan, L. Sun, and Z. Wu. *Liveness detection for face recognition*. INTECH Open Access Publisher, 2008.

[12] G. Pan, L. Sun, Z. Wu, and S. Lao. Eyeblink-based anti-spoofing in face recognition from a generic webcamera. In *IEEE ICCV*, pages 1–8, 2007.

[13] B. Peixoto, C. Michelassi, and A. Rocha. Face liveness detection under bad illumination conditions. In *IEEE ICIP*, pages 3557–3560, 2011.

[14] D.A. Socolinsky, A. Selinger, and J.D. Neuheisel. Face recognition with visible and thermal infrared imagery. *Computer Vision and Image Understanding*, 91(1):72–114, 2003.

[15] H. Steiner, S. Sporrer, A. Kolb, and N. Jung. Design of an active multispectral swir camera system for skin detection and face verification. *Journal of Sensors*, 2016, 2015.

[16] X. Tan, Y. Li, J. Liu, and L. Jiang. Face liveness detection from a single image with sparse low rank bilinear discriminative model. In *ECCV*, pages 504–517, 2010.

ISSN 2464-4617 (print) WSCG 2016 - 24th Conference on Computer Graphics, Visualization and Computer Vision 2016
ISSN 2464-4625 (CD-ROM)

2D-Clothing Shape Design Using Spline Curve

Hoang Tra Linh
National Taiwan University of Science and Technology, Taiwan
tralinhoang@gmail.com

Chuan-Kai Yang
National Taiwan University of Science and Technology, Taiwan
ckyang@cs.ntust.edu.tw

ABSTRACT

Fashion design expresses modernness, reflects changes in a society, economy, politics and culture. As a result, fashion also changes very fast and distinctively, and for that reason the improvement and creativity are indispensable. Nowadays, there are numerous fashion design systems/tools. However, they are more about making the clothes software, therefore the provided shape is limited and the way of rendering is stiff since it does not take the shape of the body into account. Tools like Telesia Creator Fashion Design CAD software provides more freedom in the way of representing the clothing shape. However, it adopts a quite unintuitive way to change the shape of a curve. Consequently, we propose the idea of a 2D clothes design system which allows a designer not only to manually create clothes but also to semi-automatically extract a clothes Spline shape from the input image. In other words, our system is able to provide the reference shape for a designer, as well as allows a designer to create his/her work by integrating some provided shape together or even directly modifying the reference shape to create a new one.

Keywords: 2D Clothes, Spline, Interpolation Curve.

1 INTRODUCTION

In the world of fashion, ideas are foundations of every creative task. Coming up with a new idea is always encouraging and necessary especially when the fashion trend changes dramatically. Moreover, since computer has become more and more popular, clothing design software has been used not only for professional purpose, but also for personal interests. It has inspired us to propose a 2D clothing shape design system which can be served as a supporting tool to make the clothing design process easier for both professional and amateur designers. In the fashion field, clothing design includes many steps that range from marketing research, creativity to sketching and fabric selection; and fashion designers work in different ways. Some sketch their ideas on a piece paper, while others drape fabric on a dress form. The proposed system is designed to be suitable for the first type of designers. It is different from the software that gives the finished garment or so-called a pattern marker. It helps a user easily create a new clothing shape from sketch or by modifying an existing design.

Nowadays, there is plenty of different fashion design software, such as: Virtual Fashion, Dress Assistant, Poser and Kaledo by Lectra, etc. For those 2D fashion design software, most of them use lines or simple arcs to render clothes which might limit the shape creation. Besides, common software like Photoshop or other drawing tools might be more flexible by allowing a user to create random shapes using an approximate spline. However, using an approximate curve means the target curve doesn't go through control points (points used to define the shape of the curve) and therefore, the created shape will be modified by approximately controlling those points not necessarily on the shape. Some people may find it difficult to manipulate such a shape and therefore need time to get used to this process. Understanding the situation, our system provides a better way for designing a clothing shape by using the interpolating *Natural Cubic Spline Curve*, or *NCSC* for short. This kind of curves also allows us to create complex clothing shapes; moreover, it allows a user to design the shape in an intuitive way. It's because the interpolating NCSC has its curves go through control points. In this way, when a user wants to modify the shape of clothes, he/she only needs to directly manipulate the control points near that desired region.

Furthermore, our system can be used in the case users want to create a clothing shape which is similar to the shape within some kind of fashion image. Our system allows a semi-automatic creation of the clothing Spline shape which is specified in the input image provided by a user. This way, the designing task becomes faster and doable for people who do not draw well. On the other hand, this system is more about giving an idea for creating a clothes rather than making a clothes. And since the trend is not what it used to be, many people now follow their own fashion rules, inspired by what they see on the fashion-animated streets, or the Internet. Hence, together with the blooming of Internet, this system is suitable for anyone with interest in fashion and

WSCG 2016 conference proceedings,
WSCG'2016, May 30 – June 3, 2016
Plzen, Czech Republic.

is able to access the Internet to get his/her own source of clothes. This system will allow a user to extract the shape that he/she wants from an image. As a result, the system helps catch the clothing shape, as well as the current trend, and then allows a user to modify shapes according to his/her taste.

2 RELATED WORK

Several prior works have used B-spline as a curve representation due to its compactness, continuity, and local shape control. Cohen et al. [2] uses *B-splines* curves for independently matching 2D objects of any affine transformations. Similarity, Mongkolnam et al. [5] uses B-spline curves to approximate the boundary of each segmented shape with finer details. Our work is mostly similar to Tseng et al.'s work [8] in a way of sampling a number of control points from the edge elements on the shape. Our system also uses B-spline curves which are more suitable to represent complex shapes in comparison with a higher degree Bezier curve due to its local control property. However, we, instead of using an approximate curve, we use an interpolation curve NCSC as an image contour representation. We think NCSC has an advantage of providing an intuitive way for design since NCSC allows a user to directly manipulate the shape by controlling the points on the curve.

For clothing shape design, over the last few decades, several image segmentation approaches have been proposed, for example, the work by Wu et al. [9], Friedland et al. [3], Mortensen et al. [6], Chuang et al. [1], and Rother et al. [7].

3 PROPOSED SYSTEM

The purpose of this system is to let a user create a new shape by modifying an existing shape. Therefore the outline of a shape is a primary concern. In order to have this outline, we need an image including the target shape which we wish to extract. Then the system will do all the works to give the user a modifiable version of his/her desired shape in the input image. The flow chart of the whole system is shown in Figure 1. We firstly choose the Natural Cubic Spline Curve (NCSC) to be our basic tool of rendering shapes which can be manipulated in the system. Then for clothes design modification of a shape-boundary retrieval approach we do the following stages: 1) Shape extraction, 2) Boundary extraction, 3) Spline-boundary creation. Our shape extraction uses SIOX, which needs some simple interaction from a user, to extract the desired object. After that, the boundary extraction can be divided into two steps: the first step uses the thresholding method to create a binary image; and the second step uses an image morphology method to extract the boundary of that binary shape; the third step is to recreate the boundary of the extracted shape under the form of a spline curve which makes use of the data information including: Harris corner detection and boundary chasing algorithm. Finally, our system allows a user to show the shape with different textures or pattern provided in the library in order to give the clothing shape result a fruitful look.

4 EXPERIMENTS AND RESULTS

4.1 Experiments

Our implemented system runs on an Intel Core i7-2600 CPU at 3.4 GHz with a 32-bit Operating System. The algorithm is written in Java, using NetBeans IDE 8.0.1.

Though most clothing worn for everyday wear falls within a narrow range of conventional styles, unusual garments are usually sought for special occasions such as evening wear or party dresses. We observe that a neckline or a collar and sleeves with cuffs of clothes are the most distinguishable parts that make them different from one other. They are more complex and sophisticated. Other parts like a body of clothes or pants are easier to manipulate. The body part usually covers most of the upper human between neck and waistline. The variety in shape of this area is not much. For example, the length of this area can be either mid-torso or mid-thigh. Or, it can be loose or tight around the bust or waist. Pants are also not too various in style and therefore, are simple to modify. Hence, in our library, we provide templates of different type of neck/collar, several types of sleeves and hemlines. Figure 2 shows an example of template library provided by our system. These templates were used to create all the results shown in Figure 3.

Our system is designed for designers who would like to sketch their ideas on a piece paper or electronic device to create a new clothes shape from sketch or modify an existing design. In order to understand more about the contribution of our system, we want to make a comparison between our system and others that are available in the market. Nowadays, in the market there are various kinds of design software. In general, we divide them into two main categories of design software: one is an image processing tool bundling with multiple design functions, for instance Photoshop; the other type serves for specific professional design purposes, for example: CAD for precise and technical drawing; logo and poster design likes CorelDraw; or artistic fashion illustration likes KaledoV4R1 of Lectra.

For the first type of software described above, we choose Adobe Photoshop CC as the candidate. Since Photoshop published by Adobe Systems is a very popular bitmap graphics editor that not only allows users to create and edit images, but also features even more creative possibilities with new tools for design. Moreover, at the time of writing this thesis, Adobe Photoshop CC is the latest version in the Photoshop series. Meanwhile, for the professional design software that serves

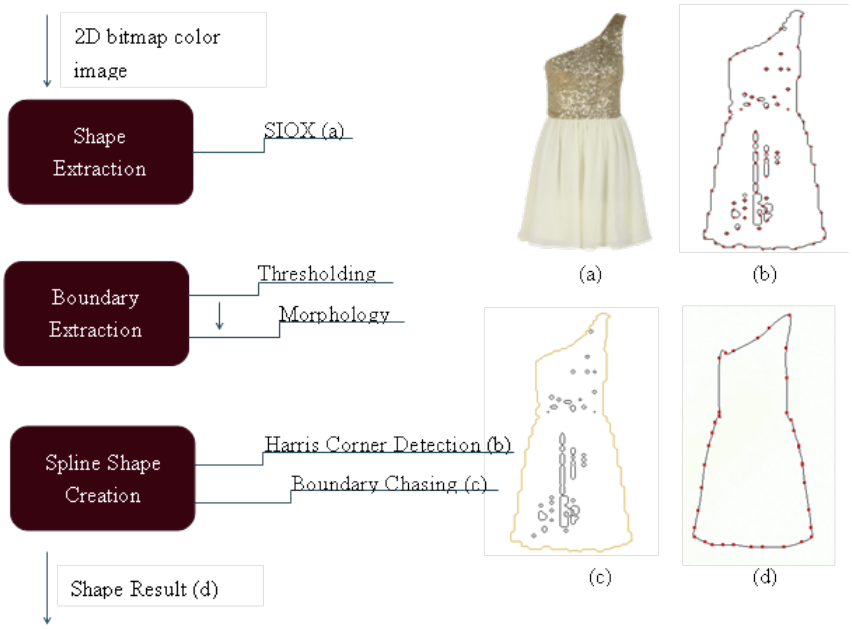


Figure 1: Flow chart and results in each step (a) SIOX extracted result; (b) Harris corner points on the extracted shape; (c) Yellow shape boundary being chased; (d) Final result represented in NCSC.

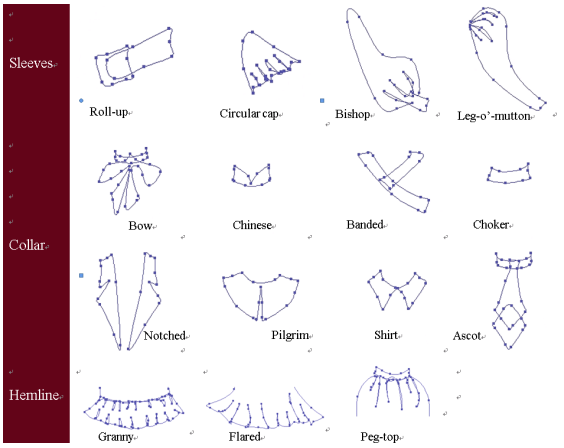


Figure 2: Template library.

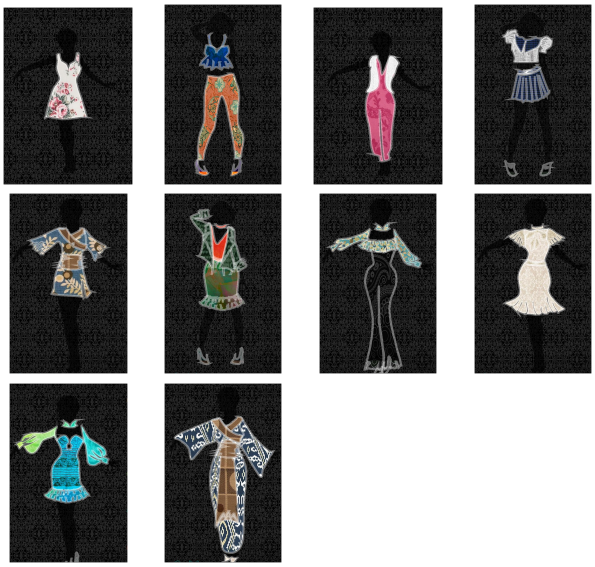


Figure 3: Results.

for specific purpose we choose KaledoV4R1 of Lectra [4] since it is more relevant to our system when it comes to clothing design. Lectra is a well-known company that provides technology solutions & associated services for industries using fabrics, leather, technical textiles & composite materials. Kaledo, professional software for clothing design, is one of the most famous creations of Lectra. It has all the necessary elements and features that are ideal for creating designer clothing meant for the catwalk, end even for a clothing business. KaledoV4R1 is the latest 2D release of Lectra on June 24, 2014.

Table 1 shows the comparison features among three different systems: Adobe Photoshop CC, Kaledo by Lectra and our system.

Currently our system only works well with the image that satisfies the three requirements listed as follows. First, the clothes from the input image should not be covered by the body part or hair (for woman garments). Second, the extracted clothes should not appear more than one time in the input image. Third, the target shape has to be fitted well inside the image, which means the shape cannot be cut by the image margin.

Finally, some results generated by our system are shown in Figure 3.

5 CONCLUSION AND FUTURE WORK

Fashion design systems have been used more and more by not only professional but also amateur designers.

System / Feature	Adobe Photoshop	Kaledo	Our System
Drawing Tools:			
Curve Tools	approximating curve	approximating curve	interpolating curve
Smart Fill Tool	✓	✓	✓
Design Assets:			
Customizable Clothing Template	✗	✓	✓
Semi-automatic Spline Shape Creation	✗	✗	✓
Design Rendering:			
Real Fabric Rendering	✓	✓	✓
Textile Design	✓	✓	✗

Table 1: System Comparison Matrix.

Considering that, we propose a new and easier way of doing clothing design which allows for rapidly acquiring a clothing shape and rendering it in a way that makes the designing task more intuitive. The whole procedure consists of three steps: 1) obtain the silhouette of an object chosen by a user using SIOX; 2) apply Harris corner detection to extract a feature points from the silhouette; 3) sort the feature points and present the shape clothes using NCSC. Compared with other systems, ours provides the functionalities of semi-automatic creating a clothes Spline shape from an input image and a reference which helps to save users a lot of time. Besides, it is different from the previous works which use the approximation curve to do the shape creation. Here we change the way of curve manipulation from approximation to direct control. As a result, it is intuitive and more convenient for a user. Moreover, our system can be used as a useful tool to complement the existing design tool, such as Photoshop or Kaledo of Lectra. The ability of using an arbitrarily provided image enables the user to get more interesting results.

From our experiments, we observed some limitations of our system, and the improvement over which leads to directions for our future works. Firstly, the inaccurate shape extraction (parts missing or incorrectly classified foreground) when it comes to extract the clothes from real-life photos or if background and foreground share many identical shades of similar colors. It is because SIOX is color dependent. Using a better method for shape extraction may significantly improve the results. Secondly, texture removal is also an important step which needs to be considered before doing the shape extraction. It helps to improve the result of shape boundary creation. Next, our current system only provides the 2D texture which makes the garment look flat and therefore deteriorates the liveliness of the design. For that reason, instead of using 2D B-Spline, we could use Non Uniform Rational B-Spline(NURB) surface which allows a user to modify the surface of the clothes to create folds of the garment, thus being able to apply and render textures in a 3D form. Lastly, the speed of the system decreases while working with more clothing shapes, and this is also a concern and needs to be improved in the future.

REFERENCES

[1] Yung-Yu Chuang, B. Curless, D.H. Salesin, and R. Szeliski. A bayesian approach to digital matting. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 2, pages II–264–II–271 vol.2, 2001.

[2] F. S. Cohen, Zhaohui Huang, and Zhengwei Yang. Invariant matching and identification of curves using b-splines curve representation. *Trans. Img. Proc.*, 4(1):1–10, Jan 1995.

[3] K. Jantz G. Friedland and R. Rojas. Siox: Simple interactive object extraction in still images. In *Proceedings of the 7th IEEE International Conference on Multimedia*, pages 1–7, 2005.

[4] Lectra. Kaledo. `http://www.lectra.com/en/fashion-apparel/design-software-kaledo`, 2014.

[5] Pornchai Mongkolnam, Thanee Dechsakulthorn, and Chakarida Nukoolkit. Image shape representation using curve fitting. In *Proceedings of the 6th WSEAS International Conference on Signal, Speech and Image Processing*, SSIP'06, pages 1–6, Stevens Point, Wisconsin, USA, 2006. World Scientific and Engineering Academy and Society (WSEAS).

[6] Eric N. Mortensen and William A. Barrett. Intelligent scissors for image composition. In *Proceedings of the 22Nd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '95, pages 191–198, New York, NY, USA, 1995. ACM.

[7] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. "grabcut": Interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.*, 23(3):309–314, Aug 2004.

[8] Chin-Hsien Tseng, Shao-Shin Hung, Jyh-Jong Tsay, and Derchian Tsaih. An efficient garment visual search based on shape context. *W. Trans. on Comp.*, 8(7):1195–1204, Jul 2009.

[9] Xiao Wu, Bo Zhao, Ling-Ling Liang, and Qiang Peng. *Advances in Multimedia Modeling: 19th International Conference, MMM 2013, Huangshan, China, January 7-9, 2013, Proceedings, Part II*, chapter Clothing Extraction by Coarse Region Localization and Fine Foreground/Background Estimation, pages 316–326. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

An Application of Triangle Mesh Models in Detecting Patterns of Vegetation

Mihai-Sorin Stupariu
University of Bucharest
Faculty of Mathematics
and Computer Science
Academiei Str., 14
010014 Bucharest,
Romania
stupariu@fmi.unibuc.ro

ABSTRACT

This contribution presents the results of a work in progress, attempting to use discrete curvature for triangle meshes in order to automatically identify specific structures in remote sensing data. Specifically, the focus was on determining isolated trees, on the basis of data acquired through airborne laser scanning. Five methods for discrete curvatures were tested and compared for a triangle mesh derived from a high density point cloud. The best performance was obtained for the mean curvature computed by using the shape operator.

Keywords

discrete curvatures; shape detection; LiDAR point cloud; Hough transform

1 INTRODUCTION

One of the applications of pattern recognition techniques is detecting specific vegetation structures or measuring its characteristics. The development of novel remote sensing technologies triggered the appearance of various approaches attempting to automatically extract information related to vegetation features, such as estimating tree height or crown diameter, determining the location of isolated trees, etc. The techniques used cover a wide range of mathematical tools, such as detection of local maxima [PWN03], second-order image derivatives [BWLM03] or multi-resolution techniques [FSH*06]. A comprehensive overview of recent developments and several alternative methods can be found in [KHY*12]. Generally, the techniques are based on the use of the so-called canopy height model in the form of a regular grid.

Since the canopy grids are derived from point clouds through interpolation techniques, some information might get lost during the averaging process. In particular, fine structures might remain undetected. An alternative to the regular grids could be the use of triangle meshes, which can be derived directly from the point cloud by using appropriate triangulation techniques. The presence of vegetation structures can be related to the lack of flatness of the canopy model, which is quantified, in turn, by discrete curvatures computed for the triangle mesh.

The aim of the study was to examine to which extent discrete curvatures for triangle meshes indeed capture remote sensing data information related to presence of isolated trees. This was achieved by performing numerical experiments and by considering several state of the art methods for computing Gaussian and mean curvatures.

2 APPROACH DESCRIPTION

We used a high density LiDAR point cloud ($\sim$20 points/m^2) acquired through customized airborne laser scanning in a mountain region characterized both by variable topography and presence of various vegetation structures. The high quality of the laser scanning together with a pre-processing step mitigated the noise influence in the remote sensing data. The point cloud contained approximately 400K points and covered a surface of 20480m^2. Then, in the next step, a 2.5 Delaunay triangulation was generated by using a standard approach: the points were projected on the horizontal plane, a Delaunay triangulation was generated and then the topology was lifted to the original set of points. The (x,y) duplicates (occurring due to the multiple returns in regions with vegetation) were removed. For the triangle mesh created in this

way, we computed the values of discrete curvatures as provided by five methods widely used and which rely on various geometric elements of the triangle mesh (Figure 1). Specifically, we considered a method based on Euler's theorem (ET), see [WB01]; two methods based on the Gauss-Bonnet scheme (GB1, GB2), see [DHKL01, MDSB03]; a method relying on the shape operator (SO), see [HP04] and the tensor approach (TA), see [Tau95]. A list of the methods as well as the geometric elements required by each method for defining the Gaussian and the mean curvature can be found in Table 1.

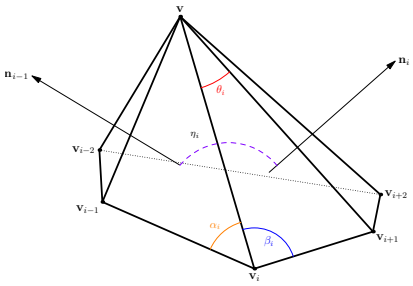


Figure 1: Geometric quantities and the one-ring neighbourhood $\mathcal{N}_{\mathbf{v}}$ of a vertex $\mathbf{v}$. One denotes by θ_i the angle between the adjacent edges $\vec{\mathbf{e}}_i = \overrightarrow{\mathbf{v}\mathbf{v}_i}$ and $\vec{\mathbf{e}}_{i+1} = \overrightarrow{\mathbf{v}\mathbf{v}_{i+1}}$. The angle between the vectors $\vec{\mathbf{n}}_{i-1}$ and $\vec{\mathbf{n}}_i$, normal to the adjacent faces $\Delta\mathbf{v}\mathbf{v}_{i-1}\mathbf{v}_i$ and $\Delta\mathbf{v}\mathbf{v}_i\mathbf{v}_{i+1}$, is denoted by η_i. The angles $\widehat{\mathbf{v}\mathbf{v}_i\mathbf{v}_{i-1}}$ and $\widehat{\mathbf{v}\mathbf{v}_i\mathbf{v}_{i+1}}$ are denoted by α_i and β_i, respectively.

Subsequently, we constructed a regular grid having cell size 1m; such that each vertex belonged to a unique cell. In this way, for each discrete curvature one got a 'grid' of averaged curvatures, having cell size 1m, obtained as follows. For each cell one firstly considered the vertices lying in that cell and then one computed the average value of the curvatures corresponding to these vertices. Finally, we used the Hough transform (e.g. [IK88]) for detecting circles. The main reason for proceeding in this way was that horizontal projections of tree crowns usually yield circular shapes. This step was achieved by using an appropriate function of the Matlab software. The sensitivity factor was set to be equal to 0.85, which is the default value.

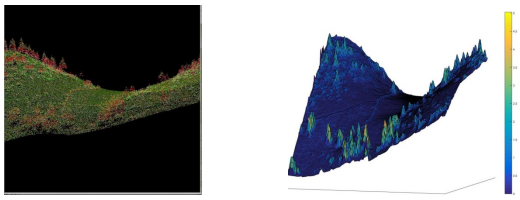
3 RESULTS

The points of the point cloud were classified in a preprocessing step. Due to the high point density, vegetation structures are already visible in the cloud (Figure 2a and Figure 3a), facilitating an empirical comparison. The regular grids derived from discrete curvatures capture for certain methods the occurrence of isolated trees (Figure 2b). While for the Gaussian curvature none of the methods does provide consistent and conclusive results, the situation changes for the mean curvature. The

Method	*Gaussian curvature*	*Mean curvature*
ET	Weighted normals Edges Edge angles θ	Weighted normals Edges Edge angles θ
GB1	Edge angles θ Face areas	Face / normal angles η Edge lengths Face areas
GB2	Edge angles θ Mixed areas	Cotangent of angles α, β Edges Mixed areas
SO	Normals Edges Face / normal angles η	Normals Edges Face / normal angles η
TA	Weighted normals Edges Face areas	Weighted normals Edges Face areas

Table 1: Methods for computing discrete curvatures considered in the study. The geometric elements are denoted as in Figure 1.

structures that can be visually detected in the LiDAR point cloud are recovered very good when one uses the shape operator method SO (Figure 3e). Other methods (such as the Gauss-Bonnet schemes GB1, GB2) detect false positives (Figure 3c,d), while the method ET based on Euler's theorem and the tensor approach TA detect fewer trees (Figure 3b,f). It is worth to notice that by using the shape operator one can detect not only circular structures such as trees, but also linear structures such as fences (Figure 3e). Moreover, the variation of the shape operator reflects the terrain variability (almost null in flat regions), higher values were recorded exactly where some variability occurs.



a. **b.**

Figure 2: **a.** The LiDAR point cloud. **b.** 3D Representation of the grid (the height of each cell is the average of the height of the vertices lying in the cell; the colour is the averaged mean curvature, computed by using the shape operator).

It is a natural question for which reasons the methods provide different outcomes. A first hypothesis could be that the methods providing the best results are those that explicitly use the lengths of the edges of the triangle mesh. This could provide a certain 'scale' sensitivity of these methods, enabling the detection of fine scale structures. Another hypothesis refers to the existence of a local autocorrelation for the methods providing better results. This makes it possible the detection of spatial structures and decreases the noise effects.

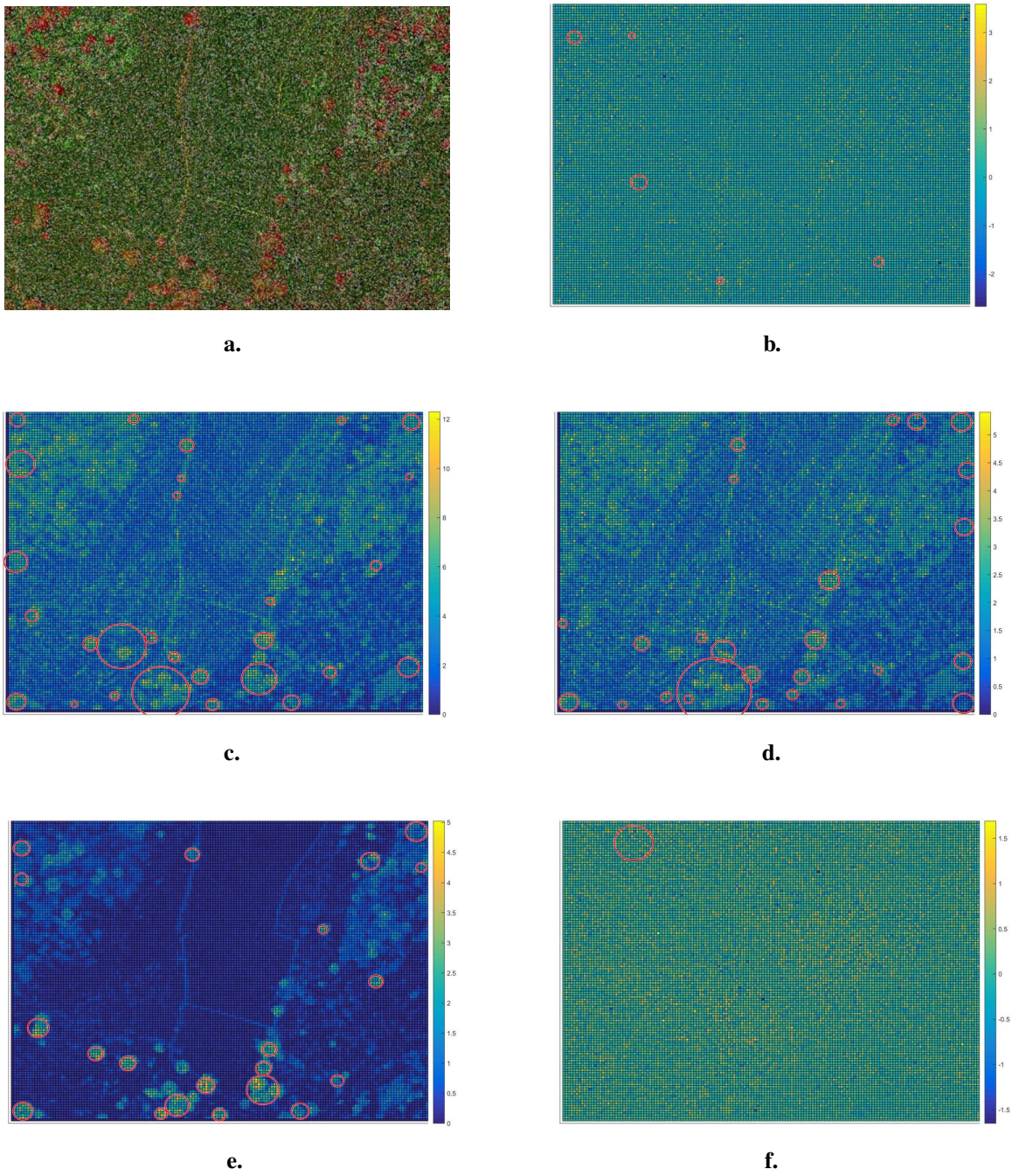


Figure 3: **a.** 2D view of the LiDAR point cloud. Data was pre-processed and colours represent height above ground. Trees are coloured in red. **b.**–**f.** Grids of averaged mean curvature: method based on Euler's theorem ET (**b**); methods based on the Gauss-Bonnet scheme GB1, GB2 (**c, d**) shape operator SO (**e**), tensor approach TA (**f**). The colours are determined by the values of the averaged mean curvature: for each cell one considered the vertices lying in that cell and computed the arithmetic mean of the curvatures corresponding to these vertices. The circular shapes were detected by using the Hough transform, as implemented in the Matlab software.

4 DISCUSSION

The method discussed in the paper has some strengths compared to other methods. For instance, standard state of the art techniques, such as the one based on local maxima in [PWN03] require a preliminary field survey, enabling an appropriate calibration and developing suitable regression models, while the method presented in the paper is completely independent on any *a priori* knowledge. Another advantage is the independence on tree species, while other approaches are species sensitive. For instance, the method in [FSH*06] is based on the Mexican Hat wavelet and is appropriate for coniferous trees. Finally, since one estimates the lack of flatness at local level, the method can be applied both to the canopy height model and to the digital surface model: the former includes only vegetation information and assumes the terrain to be flat while the latter includes both the vegetation and terrain variability.

There are still some shortcomings that can be mitigated and some issues that need to be further investigated. A major assumption of the method is that the trees to be detected are perfectly circular, which is not always the case, and that they are completely isolated. In particular, a major challenge is to detect groups of trees or to identify single trees in a forest. Both the lack of roundness and the neighborhood influence could be addressed by adapting the shape detection algorithm. The method works well for very high resolution data and it is a natural question to what extent it could be useful for less accurate input. The time complexity is a weakness of the method, since generating a triangulation for a high resolution point cloud requires a higher computational effort. However, efficient algorithms, combined with the increase of the computational power of modern hardware could decrease the running time even for larger data sets.

5 FUTURE WORK

Instead of conclusion, we briefly discuss some tracks that can be further investigated. From a rather theoretical perspective, it is a need to understand the hypotheses related to the role played by the formal definitions and whether a certain autocorrelation exists for some methods. Another important issue, regarding both the theoretical and practical perspective, is to understand how the outcomes change when the sensitivity parameter is changed in the circle detection step. Thus, 'fine tuning' could be crucial in improving the performances of the method. Last but not least, comparative analyses with other methods relying on different techniques, combined with a field survey and a cross-check of the results would be desirable. One could understand whether the method presented in this study could be used as a stand-alone technique or could be useful in hybrid approaches, by bringing added-value to other well-established methods.

6 ACKNOWLEDGMENTS

This work was partially supported by the Swiss Enlargement Contribution in the framework of the Romanian-Swiss Research Programme, project WindLand, project code: IZERZO 142168/1 and 22 RO-CH/RSRP.

7 REFERENCES

[BWLM03] Brandtberg, T., Warner, T.A., Landenberger, R.E., McGraw, J.W. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. Remote Sensing of Environment 85 (2003), 290–303.

[DHKL01] Dyn, N., Hormann, K., Kim, S., Levin, D. Optimizing 3d triangulations using discrete curvature analysis, in Mathematical Methods in CAGD: Oslo 2000, Lyche T., Schumaker L., (eds.). Vanderbilt University Press, Nashville, TN, pp. 135–146, 2001.

[FSH*06] Falkowski, M.J., Smith, A.M.S., Hudak, A.T., et al. Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of lidar data. Canadian Journal of Remote Sensing 32 (2006), 153–161.

[HP04] Hildebrandt, K., Polthier, K. Anisotropic filtering of non-linear surface features. Computer Graphics Forum 23 (2004), 391–400.

[IK88] Illingworth, J., Kittler, J. A survey of the Hough transform. Computer Vision, Graphics, and Image Processing 44 (1988), 87–116.

[KHY*12] Kaartinen, H., Hyyppä, J., Yu, X., et al. An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. Remote Sensing 4 (2012), 950–974.

[MDSB03] Meyer, M., Desbrun, M., Schröder, P., Barr, A. Discrete differential-geometry operators for triangulated 2-manifolds, in Visualization and Mathematics III, Hege H., Polthier K., (eds.). Springer-Verlag, Heidelberg, pp. 35–57, 2003.

[PWN03] Popescu, S.C., Wynne, R.H., Nelson, R.F. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Canadian Journal of Remote Sensing 29 (2003), 564–577.

[Tau95] Taubin, G. Estimating the tensor of curvature of a surface from a polyhedral approximation, in The Fifth International Conference on Computer Vision, pp. 902–907, 1995.

[WB01] Watanabe, K., Belyaev, A. Detection of salient curvature features on polygonal surfaces. Computer Graphics Forum 20 (2001), 385–392.

Distributed video coding for periodic video sequences

Moussaab Laraba
SP_Lab, Department of Electronics, Frères Mentouri University, Constantine, Algeria
moussaab_laraba@hotmail.com

Said Benierbah
SP_Lab, Department of Electronics, Frères Mentouri University, Constantine, Algeria
bnyrbhsaid@yahoo.fr

Mohammed Khamadja
SP_Lab and Elect. Eng. Dept., Larbi Ben M'Hidi University, Oum El Bouaghi, Algeria
m_khamadja@yahoo.fr

ABSTRACT

Distributed Video Coding (DVC) is a very active research field that aims to provide simple encoders, needed by many low resources applications. Unfortunately, all the proposed implementations of this type of coding claim that there is no way to design such a system as described by the distributed Source Coding (DSC) theory. With unconvincing results, such a system can only be combined with conventional coding. In this paper, we will show that there is at least one situation where DVC can be applied directly and efficiently. Thus, we adapt and apply the DVC concept to periodic video sequences (PVS). For these sequences, we will propose a new technique to create the side information (SI), where intra coded frames and motion estimation are no longer needed, which makes this technique very simple and yet very effective. The experimental results show a very good performance, and in some cases, we can even outperform H.264 Inter coding.

Keywords

DVC, periodic video sequences, Side Information, WZ coding.

1. INTRODUCTION

Periodic video sequences (PVS) are a very interesting field to study, because this phenomenon surrounds a lot of imaging and video applications. As interesting applications of PVS, we can list: the artificial satellite motion and panning surveillance cameras, where the scenes are monitored in a periodic way. When this movement is regular, it can be perfectly predictable. This allows both the encoder and decoder to know exactly the correlation between the frames. This situation makes them a good application candidate for the distributed video coding (DVC). This paradigm is based on two major theorems of source coding: Slepian-Wolf (SW) [1], for lossless distributed source coding and Wyner-Ziv (WZ) [2], for lossy coding with side information (SI). DVC is a new coding paradigm that shifts the compression complexity task from the encoder toward the decoder while maintaining the compression efficiency, where it exploits the correlation between the main information and the SI at the decoder, the DVC claims that the result achieved is similar as when the correlation is exploited at the encoder. In this case the SI is the most important information in the whole DVC paradigm because it plays a crucial role on the overall achieved RD performance. It seems that there are no situations where WZ coding can be applied directly. It can be only combined with classical video coding techniques. As a result, in all the proposed DVC coding models [4], the video sequence is divided into two parts: key frames and WZ frames. The channel codes based DVC techniques exploits the SI at the decoder to correct the main information. In state of the art, all proposed SI generation techniques are based either on interpolation or extrapolation, using Motion Estimation (ME) and Motion Compensation (MC). The intra coded key frames [4], key blocks [5] or a hash [6] are selected as spatial or temporal neighbors of the WZ information, to ensure that a correlation exists (but not completely known) between the two parts, to extract a good SI. This problem doesn't exist for PVS, because when video sequence is divided into a group of periods, we know that all frames located at the same temporal position of each period are extremely correlated. Thus, if we pick just one period to be SI, we will have a very high degree of correlation between SI and the main information. Therefore, the DVC coding concept is well-suited for PVS.

In this paper, we propose to apply WZ coding to the PVS. In this case, WZ coding can be applied directly without dividing the video sequences into group of pictures; also ME and MC can be avoided. It results

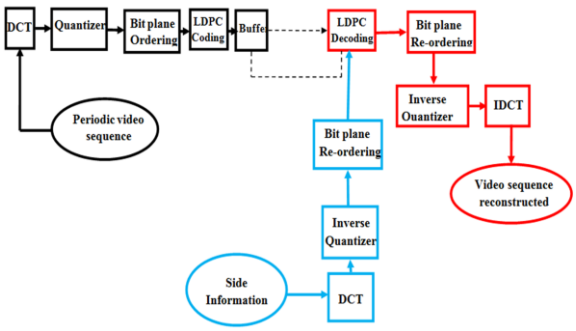


Fig 1. The proposed DVC architecture for PVS

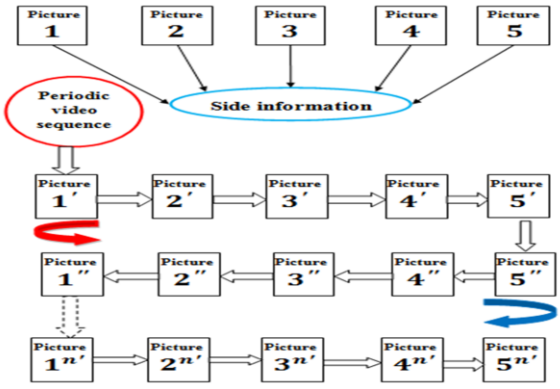


Fig 2. DVC coding for a periodic panning camera

on a very simple technique compared to others. Differently from what is presented in the state of the art on SI generation [3], our proposed technique is based on the following four points, which make its originality and its interest:

1- It is possible without use of ME or MC.

2- No image classification, i.e. there is no intra images; we work only with WZ images

3-The SI can be generated before the starting of coding and decoding procedure.

4- SI can be stored directly in the decoder without being coded and decoded with a traditional coding, to maintain the highest quality and degree of correlation. The rest of this paper is organized as follows: section 2 presents a short state of the art on WZ coding. In section 3, we explain how SI is generated in the case of PVS and how the correlation is exploited. In section 4, we discuss the experimental results. Finally conclusion is presented in section 5.

2. WINER-ZIV CODING

The Slepian-Wolf and Wyner-Ziv theorems state that the rate achieved by encoding two statistically dependent sources X and Y, is almost the same, either when the correlation is exploited by the encoders or by the decoder only. Practical DVC systems are based on WZ theorem for lossy coding with exploiting the SI. The source X and the side information Y are considered to be two independent sources, separated by a virtual dependence channel, where Y is considered as a "noisy" version of X. Channel coding is used to correct the errors of Y to reconstruct X. The capacity approaching Turbo codes or Low Density Parity Check (LDPC) codes are the most dynamic and efficient codes used nowadays, because they only transmit a minimum number of parity or syndrome bits [7].

In a WZ coding system, the main information to code is X, while Y is produced by the decoder based on local information that is supposed to be already highly correlated to X. Where X is only a part of the sequence and the other part is Intra coded [4] [5]. Y is generated by motion compensation based on prediction from the already decoded frames. The obtained compression ratio depends on the amount of used channel bits. If Y is very similar to X, only a small number of channel coding bits are needed and the bit rate is reduced. This makes creating the best possible SI the most important way to improve the coding efficiency of the WZ coding [3].

3. THE PROPOSED DVC SYSTEM

In this section, we describe the proposed DVC system designed specifically for PVS, along with the technique used for SI generation. In these applications, the correlation between the frames is well known by the decoder. For example, if a camera monitors a scene, by turning from left to right and from right to left or around, the result is still a periodic video sequence, composed of a group of periods, where each period is composed of a defined number of frames, as a result frames located at the same temporal position will be extremely correlated. In this case, if a frame is WZ coded, the corresponding frame at the same position at any other period will be highly correlated, and make a good SI candidate. We exploited this idea to design an efficient DVC system for PVS. In order to do this, we picked only one period to be SI, and use it to decode all frames which belongs to the other periods.

Fig.1 illustrates the DVC architecture used in this work. The coding and decoding procedures are exactly the same as in [8] and the only changes are:
- SI generation and Estimation of the noise correlation parameters

3.1 SI generation

Fig.2 explains how the SI is generated and how the video is decoded and reconstructed using this SI. In this example we assume that the video sequence is captured by a panning surveillance camera. Each period is composed of 5 frames. We extract only one period from the whole sequence, store its frames

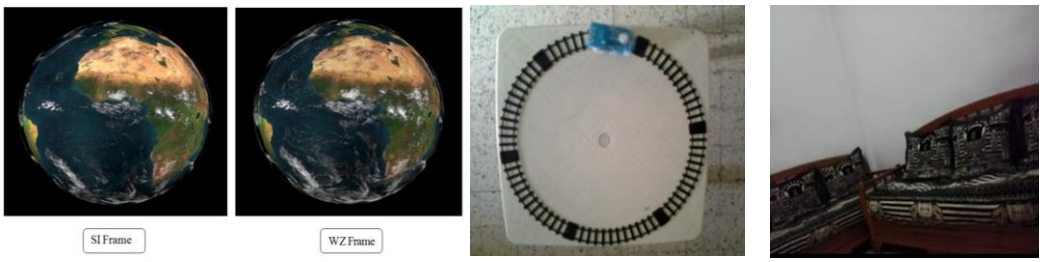


Fig 3. (a) Spinning earth sequence, (b) Train sequence and (c) Video surveillance sequence.

independently at the decoder and consider it as our SI. To ensure the best decoding quality we may also use the last decoded period as SI. To reconstruct the sequence we proceed as follow: Each frame of each period is reconstructed independently, using its corresponding frame in the SI i.e. the frame located at the same temporal position. For example, if we want to reconstruct the frame number 1'of the first period, we perform a correction procedure on the frame number 1 of the SI (frame 1 is located at the same temporal position as frame 1') and so on, for all frames of the being encoded sequence. If the movement of the camera is perfectly or sufficiently synchronized, this technique provides a very high correlation between the SI and the main information, and thus provides a very high quality of SI.

3.2 Estimation of the noise correlation parameter:

In DVC, decoding efficiency of WZ frames critically depends on the ability of modeling the statistical dependency between the main information and the SI at the decoder. In state of the art, a Laplacian distribution is used to model the correlation noise, the distribution parameter is estimated online as in [8]. In our system, we use a Laplacian distribution too, but the distribution parameter is estimated offline, because in our case the correlation between all the frames of the different periods is almost the same. This allows us to access them and to choose the best distribution parameter. Here, the distribution parameter is estimated by two different ways:

128	64	32	16
64	32	16	8
32	16	8	4
16	8	4	0

6

64	32	16	8
32	16	8	4
16	8	4	4
8	4	4	0

5

64	16	8	8
16	8	8	4
8	8	4	4
8	4	4	0

4

32	16	8	4
16	8	4	0
8	4	0	0
4	0	0	0

3

32	8	4	0
8	4	0	0
4	0	0	0
0	0	0	0

2

16	8	0	0
8	0	0	0
0	0	0	0
0	0	0	0

1

Fig 4. quantization matrices

Test1: we test a number of different values and then we pick the one which gives the best RD performance.

Test2: the distribution parameter is estimated offline between frames of two periods and we use it to decode the frames of another period. For example, in fig.2, we estimate the parameter from frames 1 and 1' and we use it between images 1 and 1''. Each band has its own distribution parameter.

4. EXPERIMENTAL RESULTS

Because of the non-availability of standard periodic video sequences, we were forced to design our own periodic sequences. We designed three different test sequences, illustrated in figure 3. In order to evaluate the RD performance of the proposed technique, we compare two different types of DVC with H.264 inter and intra coding: The first is DVC1, where the Laplacian distribution parameter is estimated as in test1 and the second is DVC2, where the distribution parameter is estimated as in test2. To obtain the six RD points we used the six quantization matrices illustrated in figure 4. The first matrix corresponds to the lowest RD point and the sixth matrix corresponds to the highest RD point.

4.1 Spinning earth sequence

The Spinning earth is a QCIF synthetic periodic sequence with 30 fps with 350 frames designed with 3DS max, characterized by a slow motion. This sequence represents two spheres: the first is the earth and the second is the moving clouds located above and around the first one, as shown in fig.3a. By adding clouds, we give to this sequence some sort of reality. The pictures contained in SI period and those in any WZ period are slightly different i.e. WZ images are cloudier and darker than the others.

Fig. 5 illustrates the RD performance of our codec for the earth sequence. We can see that the two DVC models outperform the H.264 intra coding for low, medium and high rates. We can also see that DVC1 outperforms DVC. The RD performance of DVC1 for the same PSNR is better than DVC2 by approximately 50Kbps, for low medium and high rates. DVC1 outperforms DVC2 because its distribution parameter provides better probabilities

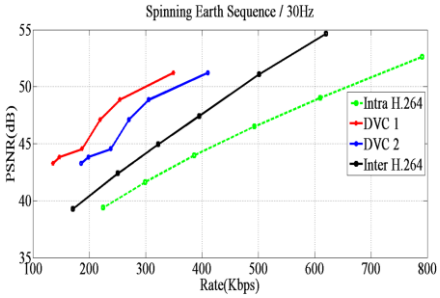


Fig 5. RD performance of Spinning earth sequence.

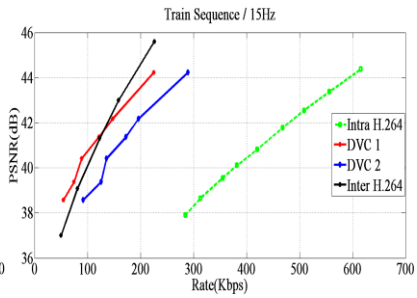


Fig 7. RD performance of Video surveillance sequence.

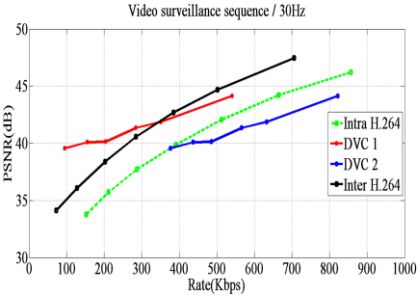


Fig 6. RD performance of Train sequence.

and thus fewer correction bits used. Most important DVC1 outperforms, even, H.264 inter coding due to the high quality of SI. Although, the pictures are not exactly the same due to the clouds movement, but it seems that DVC prediction using the pictures from different periods, is more efficient than the MC of the complex clouds movement using by inter coding.

4.2 Train sequence

The train sequence is a QCIF test periodic sequence with 15 fps with 290 frames, filmed by a camera above the railway and used to follow a moving train, as illustrated in fig 3b.

Fig. 6 illustrates the RD performance of our codec. We can see that the two DVC models outperformers intra H.264 coding for low, medium and high rates. The results show also that the RD performance of DVC1 is better than of DVC2, for the same PSNR, by around 40-80 Kbps. DVC1 outperforms DVC2 because its distribution parameter provides better RD performance.

Here also, DVC1 outperforms H.264 inter coding for low rates. But for higher rates inter is better, because in this test sequence the camera is static and only the train moves. This makes current and precedent images highly correlated, which is easy to code for inter coding.

4.3 Video surveillance sequence

Figure 3c shows this sequence. It is a natural QCIF test periodic sequence with 30 fps with 390 frames, filmed by a surveillance camera. It contains more texture, which makes it more difficult to study. The camera pans from right to left performing a periodic motion. Fig 7 illustrates the RD performance of our codec. We can see that only DVC1 outperforms intra H.264 coding for low, medium and high rates, because the distribution parameter of DVC1 provides better RD performance than DVC2.Fig.7 shows also that DVC1 outperforms H.264 inter coding for low rates, but for higher rates inter is better.

5. CONCLUSION

The main contribution of this paper is to target a new application on DVC field of research, by proposing a new technique of SI generation. It is designed especially for PVS. The major advantages of this technique are the absence of ME or MC, and also no need for intra frames, since we work only with WZ frames.

This makes this technique very simple but it remains very efficient. The results show that our DVC outperforms the H.264 intra coding for all test sequences, and it is even better than H.264 inter coding in some cases.

6. REFERENCES

[1] D. Slepian and J. Wolf, “Noiseless Coding of Correlated Information Sources” IEEE Transactions on Information Theory vol.19, July 1973.

[2] A. Wyner and J. Ziv., “The rate-distortion function for source coding with side information at the decoder,” IEEE Transactions on Information Theory, 22(1):1–10, Jan1976.

[3] C. Brites, J. Ascenso, F. Pereira, “Side information creation for efficient Wyner–Ziv video coding: Classifying and reviewing” signal processing: image communication Vol. 28 pp. 689–726, 2013.

[4] B. Girod, A. M. Aaron, S. Rane, and D. Rebollo-Monedero, “Distributed video coding,” Proceedings of the IEEE, vol. 93, no. 1, pp. 71-83, Jan. 2005.

[5] S. Benierbah and M. Khamadja “Hybrid Wyner-Ziv and intra video coding with partial matching motion estimation at the decoder,” in Proc. of IEEE Int. Conf. on Image Processing ICIP09, pp. 2925-2928. Cairo, Egypt, November 2009.

[6] R. Puri and K. Ramchandran, “PRISM: a “reversed” multimedia coding paradigm,” in Proc. of IEEE Int. Conf. on Image Processing ICIP03, pp. 617–620, Barcelona, Spain, September 2003.

[7] D. Varodayan, A. Aaron and B. Girod, “Rate-adaptive codes for distributed source, coding,” EURASIP Signal Processing Journal, Special on Distributed Source Coding, pp. 3123 - 3130, vol. 86, nº 11, Nov. 2006.

[8] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, M. Ouaret, "The DISCOVER codec: Architecture, Techniques and Evaluation," Picture Coding Symposium 2007, Lisbon, Portugal.