A Data Visualization Course at The University of Paderborn

Gitta Domik*

University of Paderborn, Department of Computer Science, Germany

ABSTRACT

My visualization courses at the University of Paderborn (Department of Computer Science) have been offered for over 20 years – but have been improved and updated many times over the years. Consistently I have used eight themes to teach the "Curriculum of Visualization". I will here shortly describe these themes and also the assignments I use and the method of introducing final projects. Some of my assignments are online for use to other visualization educators. I will also deviate to describe the interdisciplinary version of my vis course. And last but not least I will describe alterations that long-term visualization educators implemented over previous years.

Keywords: data visualization, interactive visualization, interdisciplinary visualization courses, visualization curriculum, assignments for visualization courses.

Index Terms: Human-Centered Computing (Human-Computer Interaction; Visualization)

1 INTRODUCTION

The University of Paderborn has 20.000 students and is divided into five Faculties. My visualization course "Data and Information Visualization" has been offered once a year for the past 20+ years (with few exceptions) as an elective course for our Master program in Computer Science (Faculty of Computer Science, Electrical Engineering and Mathematics). In previous years it was offered under the title "Computer-Generated Visualization" and starting next year it will be titled "Interactive Data Visualization" to emphasize the fact that today we are mostly concerned with interactive visualizations. My course follows a visualization curriculum going back to workshops in the '90s, when the ACM SIGGRAPH Education Committee (or rather its subcommittee on Education for Visualization) took on the responsibility of evaluating existing courses on visualization and making recommendations for educators at universities and colleges for future courses or curricula (e.g. [1]). These recommendations were finalized a few years later, best explained in [2]: We proposed eight themes at the core of visualization-knowledge (Introduction to Visualization; The Data; The User and the Task; Mapping Process; The Representations; Interaction Issues; Concepts of the Visualization Process; Systems and Tools). These eight themes have been updated over the years, most recently in a Eurographics Workshop, on which we reported in [3].

* domik@uni-paderborn.de

These themes are explained in more detail in section 2.1. Successful students receive 4 ECTS, translating to 120 hours of work (including time of attendance in the class room) over a period of 15 weeks. The final grade depends on the results of an oral exam on the theoretical content of the course, and delivery of homework or project results. I have switched over the years between elaborate homework assignments; or fewer and shorter homework assignments plus an extended final project. I will explain examples of assignments and projects both in section 2.2 and 2.3.

Often the course is supplemented by a separate Master seminar on "Advanced Visualization Topics" allowing to go into more depth of the matter. But this paper is describing the course and not the seminar.

The same visualization course has also been offered many times to a wide variety of students from other faculties. In the mix of students from music, economy, media sciences and other departments together with computer science students, I structure the course differently, in order to create interdisciplinary project teams. I have previously reported on such interdisciplinary and collaborative visualization courses and specifically on pedagogical strategies to make it work, e.g. in [4], [5] and [6], but in this paper I will only describe the interdisciplinary version of this course shortly in section 2.4.

2 VISUALIZATION COURSES AT THE UNIVERSITY OF PADERBORN

I will first discuss the eight themes of the course and then discuss assignments and projects.

2.1 The (Current) Eight Themes of the Course

Definitions/History/Examples introduces history and definitions of "computer-generated visualizations", shows the power of visual representations in understanding complex processes and large data sets through well selected examples, and of course explains the content of the course to students. In this theme I also explain three distinct areas of visualization differing in their data domains:

- *scientific or data visualization*, in which data dimensions usually coincide with physical dimensions, such as in medical or remote-sensing scalar or flow data;
- *information visualization*, with typically multi-dimensional data, such as in finance, business intelligence, or large databases;
- *visual analytics*, with massive, multisource, multiscale, heterogeneous, and streaming data.

The Data is the next – and in this course the most extensive - theme. Data is the result of a mapping process from the world (the "reality" or problem domain) to a computer-readable data format. Although the three visualization areas have distinct data domains, main principles of dealing with data are common to all three areas. Distinction is in data pre-processing, e.g. filtering, normalizing, or linguistic analysis, and subsequent visual presentations, e.g. line graphs, line-integral-convolution, or cone trees, all of which strongly depends on data syntax and semantics. As this is a course in Computer Science, many data pre-processing techniques are taught and practiced in assignments. Additionally, students also

learn to categorize data by their characteristics (such as ordinal, nominal, scalar, vectorial, continuous).

The theme ***The User and the Task*** includes human information-processing capabilities and limitations (such as colour blindness or low sensitivity to short wavelength), introduces visual attributes, as well as an understanding of the tasks users bring to visualization processes.

Mapping/Design describes the process of careful mapping of data components to visual attributes and thus to a perceivable representation. This is often called the "design stage", and I explain different strategies of the mapping process, such as Renaissance Teams, bottom-up or top-down mapping.

The theme ***Visualization Techniques*** introduces students to a wealth of different visualization techniques, carefully sorted by their data characteristics. This theme fills the brain of students with images of complex visualizations for later use on their projects or assignments. It also allows me to go into depth for several of the visualization algorithms, e.g. Direct Volume Rendering (DVR), Marching Cubes Algorithm, Streamlines, or LIC (Line Integral Convolution). This depth is only possible if taught in a Computer Science environment with a prerequisite of an introductory Computer Graphics course. For my version of an interdisciplinary visualization course I retreat to a different strategy as explained in section 2.4.

Interaction Techniques have become increasingly important for data and information visualization, and have been a requirement for visual analytics right from the birth of this vis area. Interaction concepts and interaction techniques are therefore taught in theory and practiced in assignments. Different coordinate systems need to be explained, because users need to interact with different spaces (e.g. screen space, data space, secondary data spaces (such as histograms), visualization technique spaces).

Evaluation of the visualization process is taught by giving a broad overview of applicable HCI evaluation methods and details on how to perform empirical user studies to evaluate visualization designs.

Currently available Visualization ***Systems and Tools*** are discussed (this is a theme that needs an update every semester), but it is open to students to choose their favourite systems and tools to work with on assignments or projects.

2.2 Assignments

Because of the limitation of 120 hours of workload for this course, I have switched back and forth between extensive homework assignments or a large final project. In any case, my students' first assignment is project 3 on https://www-old.cs.uni-paderborn.de/fachgebiete/ag-domik/curriculum-for-visualization/sample-projects.html (please be aware that over the summer of 2016 the websites at our University will be updated and that link might not be available continuously; alternatively google for 'uni paderborn domik sample projects') as their first homework. The data in this project is a 81 x 81 vector field to be visualized. It forces students to choose a programming language with a graphics API and forces them into their first practice of reading and displaying data. Also, not having learned yet the methods of the visualization process, the results lend themselves to discussions on good and bad visualizations. A typical student result is presented in Figure 1. The same assignment given after lecturing up to theme 5 (Visualization Techniques) produced Figure 2.

While teaching theme "Data", one assignment is to read in a 512x512 slice of medical CT data; analyze it statistically, produce profile lines through the data set, and filter it with Gaussian and median filters. After the theme "Visualization Techniques", students receive the full CT volume data set (480 slices) and implement a DVR technique (ray-tracing), submitting an animation by rotating around the volume data. This assignment solidifies the knowledge of implementing a DVR algorithm, and at the same time allows students to observe and discuss the 3d effect of rotating objects on a 2d screen.

Another assignment is a solution to online project 1 (see website above, as discussed with first assignment), which involves reading in four astrophysical data sets and visually merge three of these into an RGB or HLS color transformation. Observations of what user tasks can be supported by such a transformation are subsequently discussed in class. An earlier assignment in my class is the implementation of a histogram equalization on the same data sets, though in the original (floating-point sky-flux) data format.

If there is no final project, six assignments will be given and graded during the course of the semester and influence the final grade otherwise determined by a final oral exam.

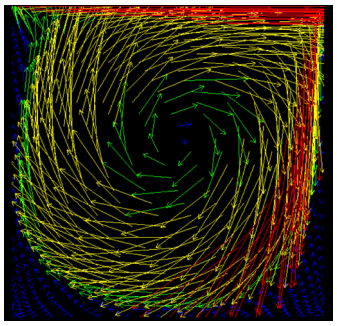

Figure 1: Exemplary result of first assignment.

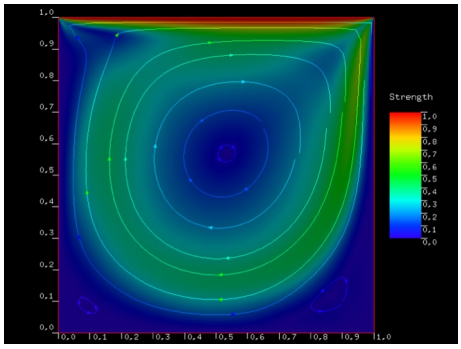

Figure 2: Result of same assignment as in Figure 1, after teaching flow visualization.

I also use In-Class exercises (up to 10 Minutes per exercise) extensively in my class. Examples of those include:

- A hand-drawn solution to project 2 on same website as indicated for assignment 1
- Discuss homework results along "visual attributes" as taught in theme User and Task
- Draw parallel coordinates for small data sets
- Histogram equalization and median filtering of 5x5 integer data set
- Find approximate contour lines using Marching Square algorithm on a 5x5 elevation field.

2.3 Final Projects

In case of an extended final project I stop assignments before mid-term and have students find appropriate data for a final project on their own (during some semesters we have also used data from one of the IEEE Vis challenges). Final projects are group work. Parallel to working on the project, assignments still need to be submitted every (second) week, such as: describe characteristics of data; describe tasks suitable for goal of user; propose (and then implement) interactive visualization techniques. This allows me to watch over the different steps of the visualization process. After developing the project, it is being presented to the class (and sometimes outside visitors).

2.4 The Interdisciplinary Visualization course

During the past 20+years I sometimes offered an interdisciplinary version of the vis course, allowing Master students from different degree programs into the course. While I still taught all 8 themes (section 2.1) in class, I took out the depth I usually teach to CS students (e.g. Volume or Flow rendering algorithms), but taught a parallel rendering course for CS students in which this depth was offered. The interdisciplinary course confronted me with a new challenge: in a single semester course students need to find project partners out of a group of students they have never met before, establish a "common ground" with their new partners and create an expressive and effective solution to a visualization problem. I enforced strictly interdisciplinary teams, so that engineering and non-engineering students were necessary to form a project group. Several years I worked on the collaboration challenge, until I was quite satisfied with the "matching" strategy:

- Students got the opportunity during the first week of class to find out about other students, by having each student fill out a form on „private" interests, including hobbies, memberships in clubs, town of high school graduation, degree program, etc. Also included was a suggested data set that student wanted the group to work on for the final project. This information was distributed to all students.

- For each group, a candidate with a suitable data set was selected by myself, and then had to fill his/her group with students of as many disciplines as possible;

- By learning about the 8 themes (of which the concepts were new to all students), students learned a "common language" suitable for all disciplines that they could use to define the goals of their project, discuss characteristics of data, aims and tasks of users, and quality criteria for good visual representations;

- weekly meetings of project teams were enforced.

One such course that I reported on in [4] had 48 participants (29 computer science students, 14 business information system students, 5 students from other programs). A survey revealed that two thirds of the students had never experienced interdisciplinary team work before, though 90% wished to work in an interdisciplinary team later. Learning to collaborate with other disciplines was thus an additional skill taught in this course.

An exemplary final project had one student bring data from his 80+ runs (location-dependent and location-independent data from his smart phone, including speed, length, altitude, weather, music…) and the group worked on visualizations supporting the identification of parameters describing successful training sessions. Figure 4 shows a visualization of the overview of the records of the runs and Figure 5 shows a particular run in more detail.

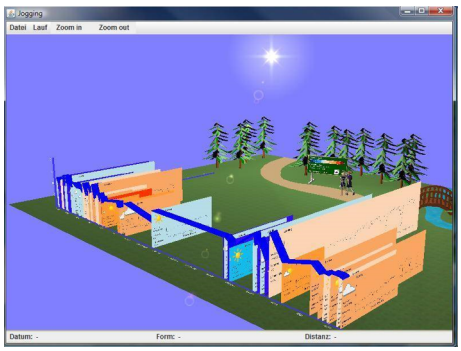

Figure 4: Overview of data records from 80+runs.

Figure 5: Location-dependent and location-independent data of one run on a map.

This project group – having found a topic they felt all passionate about - became very closely knit and ended up working *and* running together.

3 How Visualization courses have changed over the years

Because of profound changes, such as ([3])

- extensive use of GPUs
- improved algorithms for flow or volume visualizations
- emphasis on highly interactive visual interfaces
- the advent and increasing importance of information visualization and visual analytics
- an increase in nontechnical students in our courses
- greater need for professional use of visualization in the workplace
- evaluation frameworks for effective visualization

changes had to be made to visualization courses. In [3], a group of long-term visualization educators discussed such changes. Still, the main themes as laid out in the '90s did not change too much, interaction techniques as well as evaluation techniques became more important, communication and collaboration in visualization became more prominent. In my case, I traded the theme "Concepts" (now a part of the introduction) with "Evaluation", the algorithms have multiplied, and my students use updated tools and techniques.

4 Resources used

I have developed my own slide sets over the years, but I point students towards a textbook [7] that is very suitable to accompany

my lectures. I also include snippets of appropriate MOOCs for a break from my slides and I am making use of my on-line Curriculum at https://www-old.cs.uni-paderborn.de/fachgebiete/ag-domik/curriculum-for-visualization/curriculum-for-visualization.html.

5 Acknowledgement

Over the years I have had many collaborators to discuss visualization curricula and I'd like to thank all of them for providing me with always new ideas of what to improve in my courses. Most of them are named as authors or collaborators in my publications. Key publications appear below.

References

[1] G. Domik. An Agenda for Education in Scientific Visualization. Visualization ´92 Workshop Report, Computer Graphics (ACM SIGGRAPH), 27:1, p.6, January 1993.

[2] G. Domik. Do We Need Formal Education in Visualization? Visualization Viewpoint. IEEE Computer Graphics and Applications, Vol. 20, No. 4, 2000.

[3] G.S. Owen, G. Domik, D.S. Ebert, J. Kohlhammer, H. Rushmeier, B. Sousa Santos, D. Weiskopf. How Visualization Courses Have Changed over the Past 10 Years, IEEE Computer Graphics and Applications, vol. 33, no.4, pp. 14-19, July-August 2013

[4] G. Domik. Who is On My Team: Building Strong Teams in Interdisciplinary Visualization Courses, Education Program, Siggraph Asia, Yokohama, Japan, December, 2009.

[5] G. Domik and G. Fischer. Coping with Complex Real-World Problems: Strategies for Developing the Competency of Transdisciplinary Collaboration, IFIP World Computer Congress (WCC2010) - Key Competencies in the Knowledge Society (KCKS 2010), 20-23 September 2010, Brisbane, Australia.

[6] G. Domik, G. Fischer. Transdisciplinary Collaboration and Lifelong Learning: Fostering and Supporting New Learning Opportunities, Lecture Notes in Computer Science, Springer, ISSN 0302-9743 and ISBN 978-3-642-19390-3, pp. 129-144. 2011.

[7] M. Ward, G. Grinstein, and Daniel Keim. Interactive Data Visualization. A K Peters, 2010.