

How to Read a Visualization Research Paper: Extracting the Essentials

Robert S. Laramee¹

¹Visual and Interactive Computing Group,
Computer Science Department,
Swansea, Wales, UK
r.s.laramee@swansea.ac.uk

Abstract

PhD students or researchers starting a new research project or initiating work in an unfamiliar research direction often undertake a scientific literature search in order to inform themselves with respect to a chosen topic. This start-up phase involves wading through and reading scores, if not hundreds, of research papers that have already been published in the area of interest. Reading a large quantity of scientific papers and capturing the essential information from them is a very challenging task. Furthermore, this difficulty only increases with the passage of time as the complexity of literature increases as well as the quantity of publications.

This paper aims to instruct a starting PhD candidate or researcher on how to read a visualization research paper. By “read” we mean extracting the essential, most important information from a (previously) published scientific conference or journal paper. During the course of a PhD, the candidate will read many research papers containing a vast amount of information. However, it is not possible to remember all of the details presented, nor is it necessary. Here we identify and describe the essential knowledge that is best extracted and summarized when reading a visualization paper.

1. Introduction and Motivation

When undertaking an unfamiliar research direction, the researcher, possibly a PhD candidate is required undertake a literature search. A literature search is performed to discover what work has already been carried out in a given field (solved problems) and likewise what work has not yet been undertaken, i.e., unsolved problems.

However, in visualization and computer graphics, reading a refereed conference or journal paper can be very challenging due to the high level of specialization, complexity, and detail contained in an individual paper. A scientific paper often presents a detailed mathematical framework accompanied by algorithms and data structures to carry them out and are usually written by (or with the aid of) experts in the field with many years of experience. This complexity does not come as a great surprise since a single paper is often the result of multiple (combined) man-years of work.

Furthermore, a new PhD candidate or researcher starting

in a new direction must wade through scores, if not hundreds, of such research papers. This is a very challenging and even daunting task. Keeping up with the explosion of literature was identified by Jim Blinn as a top future research challenge [Bli99, LK07]. However, gaining a complete understanding of every published research paper including all of its details is neither possible nor necessary.

Here we describe an educational approach to extracting the essential information when reading a visualization (or computer graphics) paper. The aim is to describe how to read a visualization research paper in order perform a literature review, e.g., to write a state-of-the-art (STAR) report. A STAR report, a.k.a. a survey paper, is a very helpful way to get an overview of previously published work in a sub-field of visualization and computer graphics.

Here we build on both our experience of writing surveys [LHD*04, LK07, LHZP07, MLP*09, MLP*10, PVH*02, PL09, PVH*03] and in the classroom. This guide

is given to both undergraduate, masters, and PhD students taking our Data Visualization class as part of a coursework assignment that requires them to summarize a scientific visualization paper, extracting and capturing the most important concepts and information. For many students, this is the first time they have been given this task. These guidelines have demonstrated themselves to be very useful for this task. They can be given out directly in the classroom and discussed. These guidelines have formed the basis for several successful survey papers as part of PhD student degrees [LHD*04, LK07, LHZP07, MLP*09, MLP*10, PL09]

The rest of the manuscript is organized as follows: Section 2 describes related work on this topic. Section 3 discusses the importance of separating a concept from its implementation, including examples. Section 4 then presents the strategy by which the essentials are extracted from a visualization research paper, supported by further examples. A discussion of breadth vs depth with respect to reading a paper is given in Section 5. Section 6 wraps up with a summary and conclusions. The length of this manuscript reflects its goals, i.e., to extract a concise summary of a visualization paper's essential elements.

2. Related Work

There are several online resources available for writing a literature review. Taylor provides guidance on how to write a literature review with a focus on the health sciences [Tay09]. A general how-to guide has been written on writing a literature review [eme09]. Lie has written a literature review guide with a focus on psychology [Lie09]. Several other online guidelines can be found by searching for "literature survey" or "literature review" online. What we present here is not a how-to guide on writing a literature survey however. The viewpoint offered here focuses on a very specific component of the literature review process, namely, given one scientific paper, how to extract the vital, most important information. It is meant to compliment and facilitate the literature review process. It's also duals with previous literature on how to write a visualization paper [Lar09a, Lar09b]. To our knowledge, it's the only one of its kind with a focus on visualization. Although, we believe it can be generalized and adapted to other topics, especially computer graphics.

Smith describes how a manuscript should be read from a referee's point of view [Smi90]. Lee presents another interesting paper on how read a research paper in order to conduct a review [Lee95]. The goal of the current paper is *not* how to read a paper from the point of view of a referee—a different process.

Globus and Raible have written a great how-to-cheat when writing a visualization paper [GR94]. Munzner [Mun08] provides a nice overview of ways to prevent an information visualization paper from being rejected from the review process.

The current manuscript is the fourth installment in what

we call *The PhD in Visualization Starter Kit (PVSK)*. Reading a research paper and extracting the essentials is one of the essential skills we've identified for PhD candidates and researchers in general. The other key skills are (1) how to write a visualization research paper [Lar09a, Lar09b], (2) basic implementation [Lar10a], and (3) how to debug visualization software [Lar10b].

3. Concept versus Implementation

Before going into the details of how to extract the essential information from a research paper it is important to understand the difference between a *concept* and its *implementation*. Looking up the words concept and implementation in the dictionary yields:

Concept: "1. a general notion or idea; conception. 2. an idea of something formed by mentally combining all its characteristics or particulars; a construct. 3. a directly conceived or intuited object of thought." [LLC09]

Implementation: "1. the act of accomplishing some aim or executing some order; 2. the act of implementing (providing a practical means for accomplishing something); carrying into effect " [LLC09]

In other words, a concept is an idea or thought. "Every revolution was first a thought in one man's mind." —Ralph Waldo Emerson.

The same can be said not only about revolutions, but about every man-made object. A concept is abstract. It is not something with mass that can weighed on a scale (if that helps).

An implementation is the actualization, or realization of a concept (or idea). In other words, an implementation is a concept that has been brought into reality. Most implementations are actual objects with mass that can be weighed on a scale. (This is true about implementations in the real world although software does not have mass.) A concept usually starts out as idea in someone's mind. A concept is then often written down on paper perhaps as a hypothesis or a specification. (It's better to write it down than simply communicate it verbally.) Lastly a concept is implemented.

The concept behind a paper can be mapped to the *contribution* of a paper in terms of a literature review.

An Example:

- **Concept:** a writing utensil. A writing utensil is a concept. It is a tool that can be used to communicate with others using symbols that are drawn by a person (or animal).
- **Implementation:** a pencil. A pencil is a writing utensil, or more specifically, it is the implementation of a writing utensil. With a pencil, graphite is used to write down on a piece of paper.

Furthermore, there are *many* different implementations of a

writing utensil. There are also pens, markers, palm pilots, etc. There are essentially an unlimited variety of different implementations of a writing utensil. Think of how many different pens and pencils you have seen in your lifetime. There are many different varieties of pencils alone, e.g., those made out of plastic, wood, or metal and so on.

It is important to note that for a given concept, there may be many different implementations. This is an important distinction when reading or writing a research paper, as well as in most other areas of life that require critical thinking.

With respect to this manuscript, its concept is described in Section 1 whereas its implementation is described in Section 4. Conceptually, the ideas expressed here act as a filter. An entire research paper is the filter input, whereas the essential information is what passes through the filter.

4. Extracting the Essentials of a Visualization Research Paper

When reading a research paper, the following is the essential information that is focused on and written down:

1. *The Concept*: What, conceptually, are the authors trying to achieve? What is the goal of the work? This can also be reformulated as: What is the contribution of the paper? (What's new here?)
2. *The Implementation*: How is the concept realized? How do the authors support their hypothesis? How do they implement the concept?
3. *Related Work*: What previous work does this paper build upon? Almost all research papers build heavily upon the work of one or two previous papers. What are these?
4. *Data Characteristics*: What are the characteristics of the data analyzed and visualized in the paper? What is the spatial dimensionality? (2D, surfaces, or 3D) What is the temporal dimensionality? (static or time-dependent) What is the resolution and size of the data set? Is the dataset multi-resolution or adaptive resolution? Are the data samples given on a structured or unstructured grid? Is it scalar, vector, or tensor data? Is it multi-variate data?

Example 1: To help illustrate this, here we provide an example of the essential information extracted from a sample research paper:

Title: *Fast and Resolution Independent Line Integral Convolution* by D. Stalling and H.-C. Hege [SH95]:

1. *The Concept*: This paper presents a faster version of the Line Integral Convolution (LIC) algorithm. This work brings the original LIC algorithm by Cabral and Leedom [CL93] towards interactive frame rates.
2. *The Implementation*: They achieve this by reducing the number of redundant streamline computations. They also use an improved streamline integrator with adaptive step-size control.

3. *Related Work*: This work builds upon and improves the original LIC algorithm [CL93].

4. *Data Characteristics*:

- spatial dimensionality: 2D
- temporal dimensionality: steady (static)
- resolution: uniform
- grid structure: uniform resolution, regular
- data type: vector

These essential attributes of a visualization paper are also attributes that can be used to categorize them. This is not a comprehensive list of characteristics that can be used for classification. Some other characteristics that could be used and summarized include:

- *Visualization Techniques*: What basic visualization techniques are used? For example, volume rendering, ray tracing, geometric or texture-based flow visualization, information visualization techniques such as parallel coordinates or treemaps.
- *Application Domain*: What application domain are the visualization techniques being applied to? For example, physics, earth sciences, astronomy, chemistry, biology, etc.

The essential information in a visualization paper can be extracted in a concise way. The above pieces of information can be used as a template to fashion the foundation of a survey paper which gives an overview of research already performed in a given area. Ultimately, a group of papers can be classified in many different ways. A good classification highlights both mature areas of research as well as unsolved problems. Extracting the essential information facilitates the construction of an appropriate classification scheme.

Example 2: As an additional illustration, we can summarize a paper everyone is familiar with, including the additional attributes:

Title: *Marching Cubes: A High Resolution 3D Surface Construction Algorithm* by W.E. Lorensen and H.E. Cline [LC87]:

1. *The Concept*: This paper describes a novel algorithm for the construction of isosurfaces. The isosurfaces produced from the algorithm stem from maintaining the inter-slice connectivity of the original data.
2. *The Implementation*: The basic implementation consists of the following steps: (1) examine a cube, (2) classify each cube vertex as inside or outside the isosurface, (3) build an index into the case table of all possible surface topologies through the cube, (4) get an edge list from the case table, (5) interpolate to find the edge locations, (6) compute gradients, and (7) go to the next next cell.
3. *Related Work*: The marching cubes algorithm builds on and improves the algorithms presented by Chen et al. [CHRU85] and Herman and Udupa [HU83].
4. *Data Characteristics*:

- spatial dimensionality: 3D
 - temporal dimensionality: steady (static)
 - resolution: uniform
 - grid structure: uniform resolution, regular
 - data type: scalar
5. *Visualization Techniques*: volume visualization, isosurface rendering
 6. *Application Domain*: visualization of medical (MR) data and Single-Photon Emission Computed Tomography (SPECT) data

The biggest criticism given to survey papers by reviewers is that the survey paper is simply a list of papers. This is why the third item on the list is emphasized. It is very important to know how a given paper relates to its predecessors and to state this explicitly. This forms a piece of the overall puzzle which can be built piece-by-piece as in Figure 7 of Laramée et al. [LHD*04] or Table 1 of Laramée et al. [LHZP07] or McLoughlin et al. [MLP*10].

We also note that although this process may not seem difficult to the experienced researcher, it is a skill that takes time and practice.

5. Breadth vs Depth

The subject addressed in this paper concerns breadth. In general, the breadth phase of a PhD, involves a literature search spanning many, perhaps even hundreds of research papers. Thus, we would like a way to facilitate and speed up this process. The implementation of this is described in Section 4. The breadth phase is followed by a depth phase. After a specific research project has been identified, a greater understanding of a select few research papers is most likely required. It is during the depth phase of study that more time is spent on individual papers rather than, say, the 1-2 hours time per paper in the breadth phase. During the depth phase, multiple passes through a paper are likely necessary.

Also, a complete understanding of the results presented in a research paper may require knowledge of a previous paper. This previously published literature is identified as one of the essential items to be extracted (in Section 4). Reading and understanding this previous literature may not be necessary during the breadth phase of a project. However, reading and understanding the key previous paper(s) is necessary if a given paper is used in the depth phase of a project, i.e., when a researcher or PhD candidate has decided upon a research prototype they would like to implement.

If the goal of reading a paper is to actually implement the presented approach, then a deeper understanding of the technique is also required. Here, the model and the implementation constituents of the paper are re-read, possibly multiple times. If some aspect of the model or implementation is not clear (this can happen often, since page limits prevent inclusion of some details) then the authors of the paper(s) themselves could be contacted with questions. Our

experience in this regard has generally been quite positive. Authors are often happy to discuss and answer questions regarding their previously published work. They may even be willing to share their implementation, or sub-sets of it. Alternatively, other experts in the field with closely related experience could be contacted.

It is difficult to judge how fruitful a future research direction or project might be based on a single research paper, unless the reader is already an expert in the field or has a good number of years experience. This is where the value of a survey paper lies. A good survey paper identifies mature areas where a lot of work has been done and newer directions where more unsolved problems lie. Luckily, a wealth of surveys have already been written, such as those found in *ACM Computing Surveys* and *Eurographics State-of-the-Art (STAR) Reports*. The *Computer Graphics Forum (CGF)* and *IEEE Transaction on Visualization and Computer Graphics (IEEE TVCG)* journals also publish survey papers. More about the literature search process itself can be found in previous papers [Lar09a, Lar09b].

Also, one way of finding out how easy or difficult it is to implement a technique presented in a paper is to simply ask the authors.

6. Conclusions

We identify and present our view of an essential skill needed by both PhD candidates and researchers in visualization, namely, how to extract the essentials of a research paper. Gathering and summarizing the vital information of scores, even hundreds, of research papers is a very challenging task that requires special methodology. The current manuscript describes an approach to achieving this goal. It serves as a helpful aid to those reading visualization research papers for the first time or to researchers carrying out a literature survey in the field. The methodology described above has demonstrated its utility in both writing survey papers and in the classroom for students of data visualization. It can be brought directly into the classroom for educational use.

7. Acknowledgements

We are indebted to Eugene Zhang of Oregon State University for very valuable discussions and suggestions on this topic. We would like to thank R. Daniel Bergeron from the University of New Hampshire, Durham, NH, for teaching us the importance of a concept versus its implementation. Thanks to Jiang Yang, a former PhD candidate at Swansea University, for providing the impetus behind this paper. We thank Tony McLoughlin, Ed Grundy, and Dan Lipsa of Swansea University for their careful proof-reading and feedback. This research was supported in part by EPSRC Grant EP/F002335/1. Questions, comments, suggestions for improvement, contributions to this paper, or any other feedback is not only welcome but encouraged. Please contact the first author.

References

- [Bli99] BLINN J.: SIGGRAPH 1998 Keynote Address. *Computer Graphics* 33, 1 (February 1999), 43–47. Transcribed by J. M. Fijii.
- [CHRU85] CHEN L., HERMAN G., REYNOLDS R., UDUPA J.: Surface Shading in the Ciberille Environment. *IEEE Computer Graphics and Applications (IEEE CG&A)* 5, 12 (December 1985), 33–43.
- [CL93] CABRAL B., LEEDOM L. C.: Imaging Vector Fields Using Line Integral Convolution. In *Proceedings of ACM SIGGRAPH 1993* (1993), Annual Conference Series, pp. 263–272.
- [eme09] *How to... write a literature review*. Technical report, Emerald Group Publishing Limited, 2009. <http://info.emeraldinsight.com/authors/guides/literature.htm> [Accessed December 2009].
- [GR94] GLOBUS A., RAIBLE E.: Fourteen Ways to Say Nothing with Scientific Visualization. In *IEEE Computer* (July 1994), pp. 86–88.
- [HU83] HERMAN G., UDUPA J.: Display of 3D Digital Images: Computational Foundations and Medical Applications. *IEEE Computer Graphics and Applications* 3, 5 (August 1983), 39–46.
- [Lar09a] LARAMÉE R. S.: How to Write a Visualization Research Paper: A Starting Point. *Computer Graphics Forum (CGF)* (2009). (forthcoming, available online).
- [Lar09b] LARAMÉE R. S.: How to Write a Visualization Research Paper: The Art and Mechanics. In *Eurographics Education Papers 2009* (2009), Eurographics, pp. 59–66.
- [Lar10a] LARAMÉE R.: Bob's Concise Coding Conventions (C³). *Advances in Computer Science and Engineering (ACSE)* 4, 1 (2010), 23–26. (available online).
- [Lar10b] LARAMÉE R.: Debugging Visualization: Guidelines for Eliminating Bugs in Visualization Software. *IEEE Computer Graphics and Applications (IEEE CG&A)* 30 (2010). (forthcoming, available online).
- [LC87] LORENSEN W. E., CLINE H. E.: Marching Cubes: a High Resolution 3D Surface Construction Algorithm. In *Computer Graphics (Proceedings of ACM SIGGRAPH 87, Anaheim, CA)* (July 27–31 1987), ACM, pp. 163–170.
- [Lee95] LEE A. S.: Reviewing a Manuscript for Publication. *Journal of Operations Management* 13, 1 (July 1995), 87–92.
- [LHD*04] LARAMÉE R. S., HAUSER H., DOLEISCH H., POST F. H., VROLIJK B., WEISKOPF D.: The State of the Art in Flow Visualization: Dense and Texture-Based Techniques. *Computer Graphics Forum* 23, 2 (June 2004), 203–221.
- [LHZP07] LARAMÉE R., HAUSER H., ZHAO L., POST F. H.: Topology-Based Flow Visualization: The State of the Art. In *Topology-Based Methods in Visualization (Proceedings of Topo-in-Vis 2005)* (2007), Mathematics and Visualization, Springer, pp. 1–19.
- [Lie09] LIE K.: *Writing the Literature Review*. Technical report, Queen Margaret University, Edinburgh, 2009. http://www.qmu.ac.uk/psych/Rtrek/study_notes/web/sn9.htm [Accessed December 2009].
- [LK07] LARAMÉE R. S., KOSARA R.: *Human-Centered Visualization Environments*. Springer Verlag, 2007, ch. Future Challenges and Unsolved Problems, pp. 231–254. Springer Lecture Notes in Computer Science (LNCS) 4417.
- [LLC09] LLC D.: *Dictionary.com*, 2009. <http://www.dictionary.com> [Accessed December 2009].
- [MLP*09] MCLOUGHLIN T., LARAMÉE R. S., PEIKERT R., POST F. H., CHEN M.: Over Two Decades of Integration-Based, Geometric Flow Visualization. In *Eurographics 2009, State of the Art Reports* (March/April 2009), pp. 73–92.
- [MLP*10] MCLOUGHLIN T., LARAMÉE R. S., PEIKERT R., POST F. H., CHEN M.: Over Two Decades of Integration-Based, Geometric Flow Visualization. *Computer Graphics Forum (CGF)* (2010). forthcoming.
- [Mun08] MUNZNER T.: *Information Visualization: Human-Centered Issues and Perspectives, Springer LNCS Volume 4950*. Springer, 2008, ch. Process and Pitfalls in Writing Information Visualization Research Papers, pp. 134–153.
- [PL09] PENG Z., LARAMÉE R. S.: Higher Dimensional Vector Field Visualization: A Survey. In *Theory and Practice of Computer Graphics (TPCG '09)* (June 2009), pp. 149–163.
- [PVH*02] POST F. H., VROLIJK B., HAUSER H., LARAMÉE R. S., DOLEISCH H.: Feature Extraction and Visualization of Flow Fields. In *Eurographics 2002 State-of-the-Art Reports* (2–6 September 2002), pp. 69–100.
- [PVH*03] POST F. H., VROLIJK B., HAUSER H., LARAMÉE R. S., DOLEISCH H.: The State of the Art in Flow Visualization: Feature Extraction and Tracking. *Computer Graphics Forum* 22, 4 (Dec. 2003), 775–792.
- [SH95] STALLING D., HEGE H.: Fast and Resolution Independent Line Integral Convolution. In *Proceedings of ACM SIGGRAPH 95* (1995), Annual Conference Series, ACM SIGGRAPH, ACM Press / ACM SIGGRAPH, pp. 249–256.
- [Smi90] SMITH A. J.: The Task of the Referee. *Computer* 23, 4 (April 1990), 65–71.
- [Tay09] TAYLOR D.: *The Literature Review: A Few Tips On Conducting It*. Technical report, University of Toronto, Health Sciences Writing Centre, 2009. <http://www.writing.utoronto.ca/advice/specific-types-of-writing/> [Accessed December 2009].