Multiscale Spatiotemporal Visualisation

STREP # FP7-248032

White paper on multiscale visualisation (final)

D2.2

Work package 2: Shared vision

Debora Testi (B3C)
Daniele Giunchi (B3C)
Xavier Planer (UPF)
Rubens Cardenes (UPF)
Gordon Clapworthy (BED)
Nigel McFarlane (BED)
Stephen R Aylward (KIT)
Richard Christie (AUK)

20/01/2012

DOCUMENT INFORMATION

IST Project Number	FP7-248032	**Acronym**	MSV
Full title	Multiscale Spatiotemporal visualisation: development of an open-source software library for the interactive visualisation of multiscale biomedical data		
Project URL	http://www.msv-project.eu		
Document URL	http://www.biomedtown.org/biomed_town/MSV/associates/reviewers/		
EU Project officer	Ivo Locatelli		

Deliverable Number	D2.2	**Title**	White paper
Work package Number	2	**Title**	Shared Vision

Date of delivery	**Planned**	31-12-2011	**Actual**	20-01-2012
Status	Version v2		final ☒ Draft ☐	
Nature	Prototype ☐ Report ☒ Demonstrator ☐ Other ☐			
Dissemination Level	Public ☒ Consortium ☐			

Authors (Partner)	Debora Testi (B3C), Daniele Giunchi (B3C), Xavier Planes (UPF), Rubens Cardenes (UPF), Gordon Clapworthy (BED), Nigel McFarlane (BED), Stephen R Aylward (KIT), Richard Christie (AUK)			
Responsible Author	Debora Testi		**Email**	d.testi@scsolutions.it
	Partner	B3C	**Phone**	+39 051 593543

Abstract (for dissemination)	The aim of the MSV project is to develop an open source library for the visualisation and interaction with multiscale biomedical data. The partners revised the current state of the art also in domains outside biomedicine and defined a series of needs and challenges for the MSV project, which are here, summarised. This shared vision document includes the taxonomy of multiscale visualisation derived from the best practice definition, an update in the exemplary problems data collection, and the defined MSV interaction paradigm.	
Keywords	Multiscale visualisation	State of the art review

Version Log			
Issue Date	**Rev No.**	**Author**	**Change**
30/12/2011	V1	Debora Testi	Complete document in all contents
20/01/2012	V2	Debora Testi	Inputs from all partners included and final revision

Project Consortium Information

University of Bedfordshire

LIST OF ABBREVIATIONS

3D	Three-dimensional
AUK	University of Auckland, NZ
B3C	BioComputing Competence Centre - SCS srl, Italy
BED	University of Bedfordshire, UK
CFD	Computational Fluid Dynamics
CT	Common ToolKit Initiative
CT	Computed Tomography
DTI	Diffusion Tensor Imaging
DXA	Dual-emission X-ray Absorptiometry
ECG	ElectroCardioGram
EEG	ElectionEncefaloGram
FPS	Frames per second
GPU	Graphic Processing Units
GUI	Graphic User Interface
HCI	Human-computer interaction
KIT	Kitware Inc, USA
LOD	Level of Detail
LOD	Levels of detail
MR	Magnetic Resonance
MSV	Multiscale Spatiotemporal Visualisation
MSV	Multiscale Visualisation and interaction project
Qt	GUI development package from Nokia
TEM	Transmission electron microscopy
UPF	Universitat Pompeu Fraba, Spain
US	Ultrasound
VPH	Virtual Physiological Human
VTK	Visualisation ToolKit

1 Introduction

1.1 Project domain and aim

The Virtual Physiological Human (VPH)[1] is a methodological and technological framework that, once established, will enable collaborative investigation of the human body as a single complex system. That framework should make it possible to interconnect predictive models that span the full range of relevant spatial and temporal scales, with diverse methods operating at varying levels of detail, into systemic networks that enable the discovery and evaluation of multi-system hypotheses [Fenner, 2008]. The project discussed in this report addresses the specific aim of integrating multiple physiological processes that operate at temporal and spatial scales that may differ by orders of magnitude (Multiscale Spatiotemporal Visualisation, MSV in brief).

As new methods for collecting and modelling multiscale data have begun to emerge from the different VPH-related projects, it has become increasingly evident that there is a shortage of appropriate tools for visualising and exploring data that are defined across a broad range of spatial and/or temporal scales; this is particularly true for data that cannot be entirely represented within the typical human perceptual range.

The number of biomedical problems that will demand multiscale visualisation in the coming years suggests that this area should start to receive urgent attention but, surprisingly, it received almost no mention in the Visualisation Research Challenges [Johnson, 2006] document produced jointly by the USA National Institute for Health and National Science Foundation.

Multiscale visualisation has been investigated in other contexts with, perhaps, the most relevant work being undertaken in the context of geographical data visualisation, as exemplified by well-known solutions such as Google Earth[2]. While these are extremely effective within the context of their specific target problem, not all of the available solutions can be generalised to other domains. In particular, many of the approaches do not translate well into the biomedical research area due to the volumetric, heterogeneous, and time-varying nature of the relevant data.

It is evident that this visualisation challenge is not simply related to the computational efficiency. Researchers should collaboratively work with domain experts on driving tasks to produce tools, which will solve real-world needs and make systems and techniques that better leverage human characteristics.

To address this issue, an international consortium has been established involving leading groups in the development of visualisation software tools[3]. MSV project (FP7-IST-248032) has the aim of designing and providing, over the course of the next year, a first implementation of an open-source library to meet the current and future needs encountered when visualising, simulating and interacting with multiscale biomedical data. The MSV software library will provide a suitable resource for the VPH community and others to be used in this rapidly evolving area.

The aim of this document, *White Paper*, is to provide a review of the state-of-the-art and of the open challenges, which are driving the designing phase of the MSV library. This White Paper also includes a description of the MSV implementation strategy together with a summary of the first demonstrators in course of implementation.

[1] http://en.wikipedia.org/wiki/Virtual_Physiological_Human
[2] http://earth.google.com/
[3] http://www.msv-project.eu/consortium.html

1.2 Some definitions

The literature (books, papers, conference proceedings, etc.) available on visualisation in general and on some of the specific topics we are addressing in this document is extensive and its exhaustive review was out of the scope of this report. We rather focused on those aspects relevant for the MSV project, in particular for this and for the following state-of-the-art sections.

1.2.1 Visualisation

Visualisation is a general term that has specific meanings in different contexts. It can refer to the research discipline, to a specific technique or to a visual result. In particular, in computer graphics, visualisation is any technique for creating images, diagrams, or animations to communicate a message. Visualisation is thus about helping people explore or explain data through software systems that provide a static or interactive visual representation [Johnson, 2006].

Visualisation has been an effective way to communicate both abstract and concrete ideas since a long time ago. However, in its early days the lack of graphics power and computational efficiency often limited its usefulness. We might consider the birth of scientific visualisation, as we currently interpret it, in 1987 with the special issue of Computer Graphics on Visualisation in Scientific Computing [McCormick, 1987]. Since then, there has been an increasing role of visualisation, and, together with computer graphics, it has today an ever-leading role in many scientific applications like science, education, engineering, multimedia, medicine, etc.

Visualisation designers exploit the high bandwidth channel of the human visual perception to allow people to comprehend information orders of magnitude more quickly than they would have by reading raw numbers or text to derive new knowledge from data. Thus, designing effective visualisation is a complex process that requires a sophisticated understanding of human information processing and, in fact, visualisation systems are explicitly designed not to replace the human but to keep him/her in the loop by extending his/her capabilities [van Wijk, 2005].

An example on the role given to the human in the visualisation process is given also by [Ware, 2004] who presents the four stages of the process giving particular emphasis to the role of the human in the loops (Figure 1).

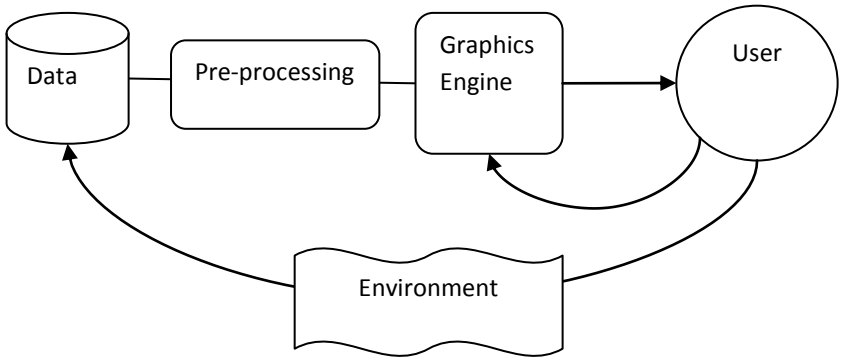


Figure 1 - The visualisation process adapted from [Ware, 2004]. The process is composed by four steps: i) the collection and storage of data itself, ii) the pre-processing designed to transform the data into something the user can understand, iii) the display and the graphics algorithms that produce an image on the screen, and iv) the human perceptual and cognitive system.

Visualisation can be then categorised according to the different types of application; the most relevant for MSV are:

- *Scientific visualisation:* it is the transformation, selection or representation of data generated by simulations, observed from physical processes, or recorded during experiments. These data typically exist in an implicit or explicit geometric framework that gives context to its exploration, analysis, and understanding.

- *Information visualisation*: it concentrates on the use of computer-supported tools to explore, in intuitive ways, large collections of non-numerical/textual meta-data, such as lines of code in software systems, results reported in publications, and citations in bibliographic databases, networks of relations on the internet, and so forth.

As MSV project is dealing with biomedical data on different scales, we are particularly interested in scientific visualisation for which more details are given in the next section. However, information visualisation also deals with multiscale data and some approaches can be of interest, such as methods based on data cubes [Stolte, 2003].

1.2.2 Scientific visualisation

Scientific visualisation is an interdisciplinary branch of science "primarily concerned with the visualisation of 3D phenomena (architectural, meteorological, medical, biological, etc.), where the emphasis is on realistic renderings of volumes, surfaces, illumination sources, and so forth, perhaps with a dynamic (time) component" [Friendly, 2008].

Fields, which are strongly related with scientific visualisation and with MSV, are:

- *Computer simulation* is a program running on a single computer, or on a network of computers, that attempts to simulate a specific system using a numeric model. Computer simulations have become a useful part of mathematical modelling of natural systems in physics, computational biology, and chemistry; of human systems in economics, psychology, and social science; and in analogue and digital systems and manufacturing process in computer science and engineering. The goal of these simulations is to gain insight into the operation of those systems, to observe their behaviour, to diagnose their failures, and to improve their performance.

- *Computational geometry* is a branch of computer science devoted to the study of algorithms which can be stated in terms of operations on geometry. The operations include measures of surface properties, transformations that project data onto geometric forms for visualisation and analysis, and unions and intersections of geometries.

- *Signal processing* is concerned with the segmentation, registration, filtering, and characterisation of data, including images. It encompasses the fields of medical image analysis, statistical pattern recognition, computer vision, and biophysical signal recording and processing.

- *Human-Computer Interaction* (HCI) is the field of study that involves the design of graphical user interfaces (GUIs), including those involving complex visualisations and the interactive display and manipulation of data. Often neglected in early computer applications, particularly research applications; the Nintendo Wii[4] and Apple's iPad[5] are just two recent examples of the power of effective HCI and the priority it is now being given in consumer systems. The surge in the adoption of Nokia's open-source, yet highly polished, Qt GUI development package[6] similarly portrays the emphasis that research application developers are also giving to HCI.

- *Parallel and GPU computing*: Visualisation has always been computationally demanding. As soon as hardware has become capable of delivering the required output within a reasonable

[4] http://wii.com
[5] http://www.apple.com/ipad/
[6] http://qt.nokia.com/

computational time, users have demanded greater realism, improved visual effects or additional features, and the associated computational demands have again exceeded the contemporary hardware capability. Certain graphics techniques, e.g. ray tracing or ray casting, have always been amenable to parallelisation, and there have been several examples of specific hardware being marketed to take advantage of this, though none proved commercially viable. The advent of the multicore computer and, in particular, the modern Graphics Processing Unit (GPU) has brought parallel computing to the desktop and this has greatly increased interest in parallelisation. In fact, the rapid expansion of GPGPU (general purpose GPU) programming, particularly as a result of support from programming languages such as CUDA[7], has meant that many tasks in visualisation are now being addressed using parallel techniques.

Scientific visualisation is usually performed using specialised software. Some of these have been released as Open-source software (i.e. Visualisation ToolKit[8]), while there are also many proprietary software packages of scientific visualisation tools.

1.2.2.1 Applications

Scientific visualisation is applied today in many fields of application, which are here briefly listed, while more details on specific multiscale visualisation and interaction techniques can be found in the publicly available D4.1 “Best practices”.

Geography and Meteorology

Cartography and geography have a long history in the use of scientific visualisation since the first world map in the 550 BC [Friendly, 2008]. The amount of information to be represented and their visualisation techniques have been growing steadily. At present, in this field of application, multiscale visualisation is used for rendering terrains, meteorological systems, climate changes, etc. The most famous example of multiscale representation in this field is provided by Google Earth, which is described in the next state of the art section (Figure 2). Strategies for representing nearly every type of multidimensional data in this field have also been identified.

Figure 2 – Example of visualisation of multidimensional and multiscale data in cartography and meteorology: Google earth snapshot.

In the applied sciences, engineering

Engineering is also an influential field of application for scientific visualisation. It typically integrates computer simulation, computational geometry, and parallel/distributed processing to drive its visualisations. Consider, for example, the multiscale visualisation shown in Figure 3. The underlying data was generated by a computational simulation that was conducted on distributed compute system.

[7] http://www.nvidia.com/object/cuda_home_new.html
[8] http://www.vtk.org

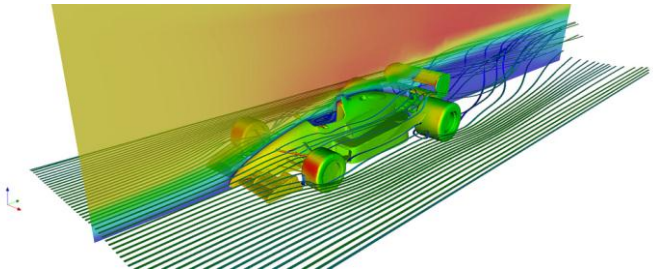


Figure 3 - Streamlines and a flow speed colour map are shown in 3D and projected onto the geometry of a car, as determine by a computation fluid dynamics (CFD) simulation. Author: Renato N. Elias, Associate Researcher at the CFD Group.

Natural sciences

Astrophysics is also challenged by multiscale spatiotemporal data. In Figure 4 the magnetic, particle, and gravitational components of the formation of a star are depicted at one event of the complex, multi-energy process.

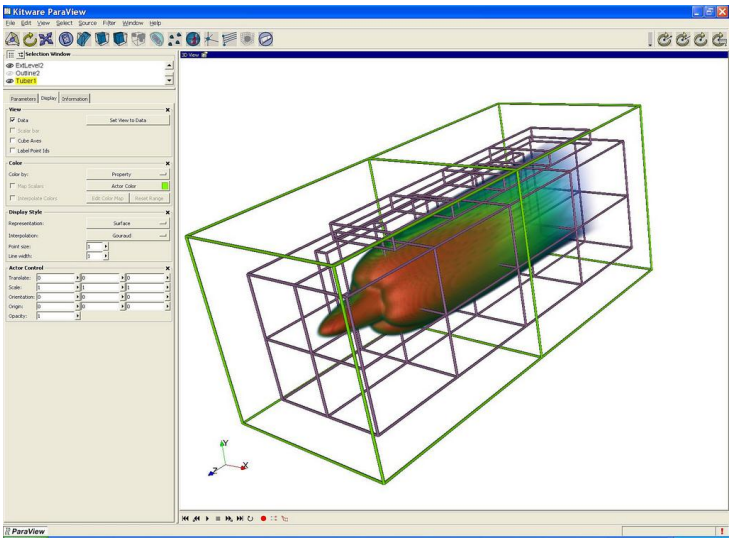


Figure 4 - The magnetic jet from a young stellar object rams into the magnetized molecular cloud from which the star is born by collapse. Data simulation by Rolf Walder and Doris Folini with the A-MAZE code. Adaptive refinement techniques were used to compute the visualisation to different levels of details in different regions of the volume. Author: Data courtesy of the Swiss National Supercomputing Centre

Biomedical sciences use of visualisation techniques has also grown significantly in recent decades. Applications that involving visualisation abound, including diagnostic medical imaging, surgical guidance, cell migration analysis, genomics, proteomics, molecular modelling for drug discovery, and more. Figures 5 illustrates only a small portion of the range of scales and visualisation methods to be considered. More details on this application area will be given in the exemplary problems section.

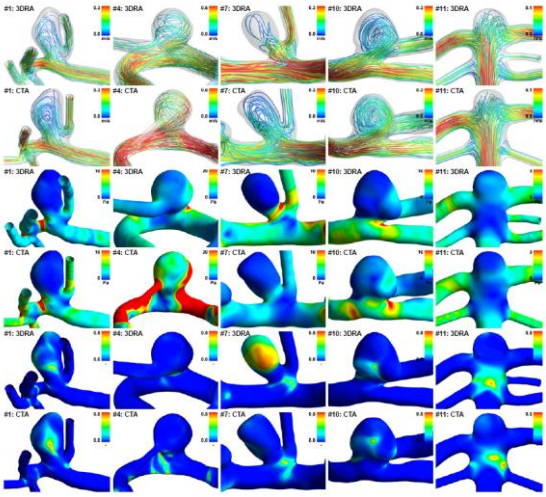


Figure 5a – Examples of scientific visualisation in biomedicine. Visualisation of cerebral aneurysm hemodynamic simulations extracted from models based on different imaging modalities (in this case, CTA and 3DRA). Results correspond to streamlines, wall shear stresses and oscillatory shear index. Courtesy of the @neurIST project.

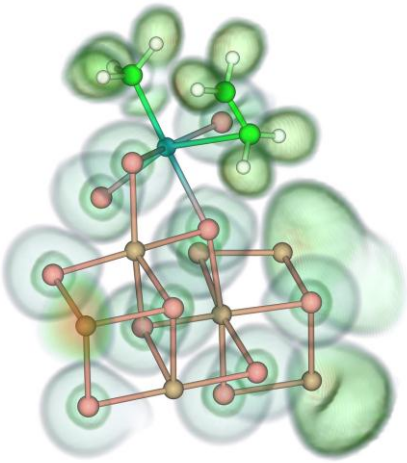


Figure 5b - Volume rendering of the density in an ethylene molecule shown in correspondence with a stick and ball model of the atoms and bonds that form that molecule, as a reference frame. Author: Jean M. Favre (CSCS)

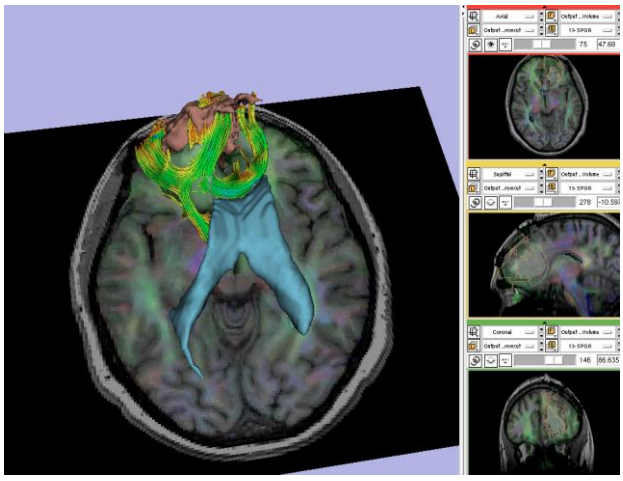


Figure 5c – Visualisation generated using 3D Slicer that includes pre- and post-treatment segmentations of a tumor to illustrate its change, DTI tractography seeded the peritumor boundary, structural and DTI data colored by orientation, and segmentation of the ventricles (included for context). Author: Wendy Plesniak, Brigham and Women's Hospital, Boston, USA.

1.2.3 Interactive visualisation

Another important aspect, we are dealing with in the MSV project, is that visualisation of biomedical data should be interactive, allowing the user to navigate within the data to gain a better understanding of the details and complexity of the information represented in them. Furthermore, in many cases interaction not only enhances the understanding of the data, but it becomes essential when the amount of data to be represented does not fit on the computer screen, or even in the computer's available memory.

Interactive visualisation is a branch of HCI and visualisation in computer science that involves the study of how humans interact with computers to create, understand, and manipulate representation of information, and how that process can be made more efficient.

For a visualisation to be considered interactive it must satisfy two criteria:

- *Human control*: control of some aspects of the visual representation of information, or of the information being represented, must be available to the user, and

- *Response time*: changes made by the user must be incorporated into the visualisation in a timely and effective manner.

1.2.3.1 Human control

Standard types of input that can be used in the human interaction phase are for example: picking some part of an existing visual representation, locating a point of interest, stroking a path, choosing an option from a list, and/or inputting numbers or text. All of these actions require a physical device. Input devices range from the common – keyboards, mice, graphics tablets, trackballs, and touchpads – to the more advanced – wired gloves, 3D mouse, haptic or other Virtual Reality tools. Recently, attention has increased on alternative interaction modalities such as human gestures devices or human tracking systems such as Microsoft's kinect, Nintendo's Wiimote, and multi-touch panels similar to Apple's iPad. These input methods can be used to control both what information is being represented as well as how that the information is being presented. When the information being presented is altered, the visualisation is usually part of a feedback loop. More frequently, the representation of the information is changed rather than the information itself.

All these standard and new ways of interacting with data may be applicable to MSV, or new input methods, specific to MSV, may be needed. For example, MSV-specific control paradigms may be needed to interact with data that are not completely in the screen or that may be hidden by other objects at vastly different scales in time and space.

The development, evaluation, and refinement of alternative input devices, however, is outside of the scope of the MSV project discussed in this White Paper. We have instead chosen to focus our efforts on designing an interaction paradigm that is effective using a standard mouse-monitor environment. In this standard environment, beyond creating an intuitive user experience, we must also create one that responds rapidly to human cues, as discussed next.

1.2.3.2 Rapid response to human input

The term interactive *frame-rate*, measured in frames-per-second (FPS), is often used to measure how much interactive a visualisation is. Frame-rate measures the frequency with which an image (a frame) can be generated by a visualisation system. Essential for good interactivity is to maintain a high and consistent frame rate [Constantinescu, 2000]. A frame-rate between 30 and 15 frames per second (frame/s) is usually considered acceptable, while below 6 frame/s would be considered poor [MCCarthy, 2004; Constantinescu, 2000]. The interactive navigation through large datasets is usually limited by the graphics hardware capabilities and thus the maximum number of frame rate a graphical hardware can achieve depends directly on the complexity of the rendered image [Constantinescu, 2000]. However,

frame-rate alone does not completely characterise interactivity as also latency is highly affecting the user perception of the interaction.

Experiments have shown that a delay greater than 10 seconds leads the user to assume that there has been an error in the processing. Other studies have confirmed the importance of the "immediacy of feedback" in learning. That is, the longer the delay between an action and its consequence, the longer it takes for a user to learn a behaviour. Yet other studies have shown that for many tasks, the tolerable waiting time between when an input is provided and a visual representation is updated is about 2 seconds and that tolerance is reduced to 0.1 second for the user to feel as if the system is interactive [Bouch, 2000; Nah, 2004]. We have therefore chosen 6-10 FPS as the limit that an interactive MSV software application should operate when rendering data based on human input. If the system cannot provide such immediate response, continuous feedback should be provided to the user by reporting processing progress [Myers, 1985].

Several approaches have been explored to provide people with rapid visual feedback based on their input. Some include:

- *Parallel rendering and GPU acceleration*: more than one computer or video card is used simultaneously to render an image. There are different way to technically achieve this parallel rendering like each computer can render a different region of a single frame and send the results over a network for display, or each computer can render an entire frame containing a subset of the information.

- *Progressive rendering*: where a frame-rate is guaranteed by rendering some subset of the information to be presented, like global scene illumination, and providing incremental (progressive) improvements to the rendering once the visualisation is no longer changing.

- *Level-of-detail* (LOD) rendering: where simplified or sub-sampled representations of information are rendered to achieve a desired frame-rate while the user is providing input, and then the full representation is generated once the person has ended manipulating.

- *Frameless rendering*: where the visualisation is no longer presented as a time series of images, but as a single image where different regions are updated over time.

1.2.4 Multiscale data

Besides visualisation and interaction aspects, it is important in the MSV context to define which types of data are going to be taken into consideration. We usually refer to the data we are interested in with the term "multiscale data" and its difference with respect to multidimensional and multivariate data is here described.

In engineering, physics, meteorology and computer science, *multiscale modelling* is the field of solving physical problems that have important features at multiple scales, particularly multiple spatial and(or) temporal scales. Data associated to this type of modelling and including information at different spatial and temporal scales are called *multiscale data*.

As it can be seen from the above definition multiscale data are different from multivariate (data with several variables for each sampling unit) and multidimensional data (data being represented on more than one dimension like 2D, 3D, 4D, etc.).

Besides the differences illustrated, it should be noticed that in the case of biomedical data visualisation often we can be in presence of set of data which include multiscale, multivariate and multidimensional information at the same time. A proper combination and appropriate representation of all those aspects are important to the users in the data interpretation and understanding, and increase the complexity of the challenge to be solved by MSV.

In addition to this, biomedical data are often multiscale both in the spatial and temporal dimensions.

During data visualisation, the user focuses on a time-space window, which is the one within the human perceptual range, just spanning a part of the whole domain. Thus, even if the whole domain is available in the data (both in space and time) the user does not look at all the data at the same time but only focus on a sub-set of them (Figure 6).

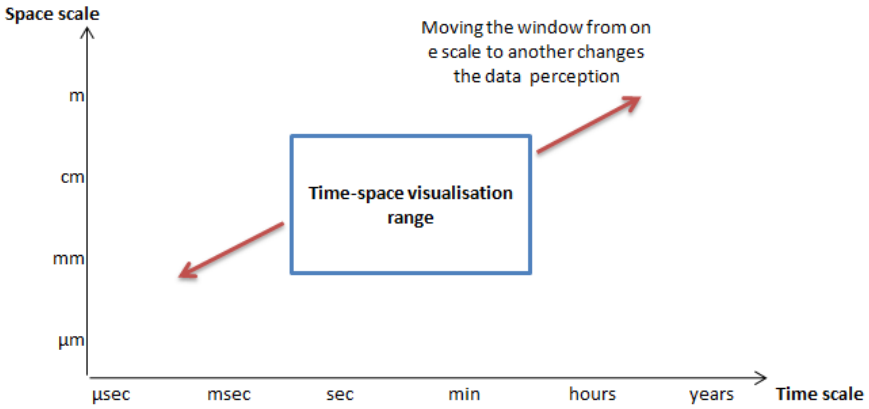


Figure 6 – Perceptual range and multiscale data

By moving this time-space window from one scale to another, the user has a different perception of the information. Another important aspect of multiscale data is that most of the time there are gaps in the scales; thus during the movement of the time-space window continuity is not guaranteed by the data, but in any case a smooth transition should be provided to the user.

Usually the movement of the time-space window will not be un-constrained but limited by the type of biomedical application which is under consideration; for example, when looking and molecule interactions, the user is not interested in the meters/hours scale and the same happens, all the way around, at body level with data like for blood flow, heart studies. Two illustrations of possible ways of constrain user interaction and manage multiscale, heterogeneous data are showed in Figure 7.

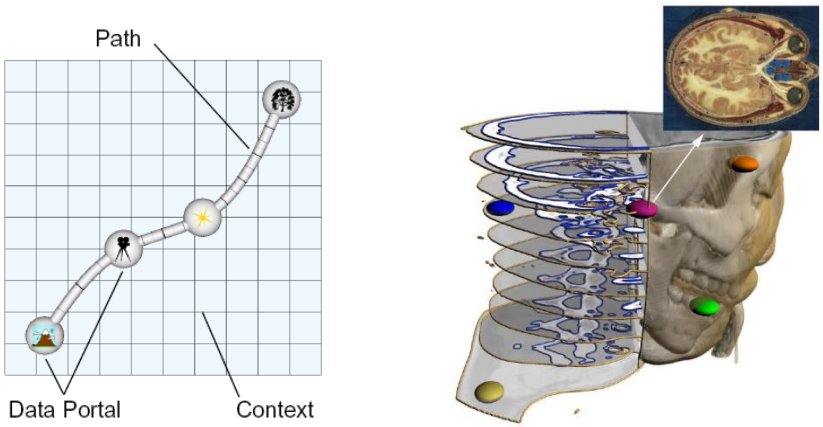


Figure 7 - Left, a path through time and space contains waypoints at which additional details are available. By clicking on a waypoint, an alternative visualisation of the enclosed data is given, perhaps a PDF document is displayed, an interactive volume rendering is shown, or an internet search is issued. Right, instead of waypoints being along a path, the waypoints are embedded in volumetric data. Again, each waypoint (shown as ellipsoids in this visualisation) indicates where alternative data is available. Those data may be the pathology reports from a biopsy sample taken at that location, or biophysical recordings from probes positioned at those location, or links to publications that discuss the anatomic feature in detail.

2 State-of-the-art

2.1 Non biomedical domains

In domains other than the biomedical one, there have been attempts to visualise and interact with multiscale data. Even if the chosen interaction approach highly depends on the type of application, it is interesting to have a look at what has been already proposed to eventually identify possible methods useful when dealing also with biomedical data. Further examples, related to specific multiscale techniques for both visualisation and interaction, can be found in the report on best practice (D4.1[9]).

Astrophysics

Computational astrophysics has rapidly developed in the recent years, enabling, since only recently, multiscale and multi-physics simulations of entire stars, planets, galaxies, and cosmic explosions [Muller, 2004]. Within the past decade, computer facilities and development of computational methods have literally exploded, allowing huge progress in the understanding of star and planet formation, a complex field which combines radioactive- and magneto-hydrodynamics, the physics of shocks, supersonic turbulence and elaborated computational.

Moreover, the latest generations of telescopes gives an unprecedented amount of observational data. However, a proper interpretation of such data always depends on computational models that explore the origin and propagation of the observed photons. On the other hand, astrophysical objects are multiscale and multi-physics data. Flows, magnetic fields, and radiation are in perpetual dynamical interplay. Non-linear couplings among the different ingredients are crucial for the understanding of objects like stars, planets, galaxies, and cosmic explosions like supernovae. Thus, in terms of visualisation astrophysics has very similar problems to those we are trying to address with MSV; the data managed are multiscale in space ranging from 10^{25} to 10^{-15} meters [Hanson, 2000]. It is also often necessary to store and analyse huge datasets, which is also one of MSV challenges.

For all these reasons, it is interesting to understand which approaches have been undertaken in astrophysics. The most interesting findings are summarised here.

The AMUSE project[10] provides a Python based framework for the multiscale simulations of dense stellar system. The framework is meant to be executed only on Unix operative system and presents parallel executing codes to manage large simulations.

Another interesting approach can be found in [Hanson, 2000]. The paper has the aim to describe a method to display data over more orders of magnitude (without significant arithmetic error) by the use of a combination of strategies:

- Definition of an homogeneous power coordinates and interpolation system;
- Systematic treatment of data at unit scale with transformation matrices;
- Systematic blending of icon and very large data group representation with object disappearance criteria and smallest visible object definition;
- Merging environment maps with octrees to make very distant data visible;
- Support to multi-resolution maps;
- Constrained multi-resolution navigation.

[9] https://www.biomedtown.org/biomed_town/MSV/reception/repository/MSV_D41_final.pdf
[10] http://amusecode.org/doc/

Cartography and geography

In geological sciences the possibility to deal with the dimensionality of the data started after 1990, and it is now challenged with an increasing volume and complexity of data to be represented. In terms of visualisation, the need in this field is to allow on-the-fly rendering across a broad range of scales.

Google earth[11] is probably the most famous example. It is a virtual globe, map and geographic information program released with a free licence in 2005 by Google obtained by the superimposition of images from satellites and aerial photography. For large part of the globe only 2D images are available from vertical photography, while for others 3D images of terrains and buildings are also available (created with a digital elevation model) [Yoshimura, 2009]. The resolution is not uniform over the globe and in the most popular area goes down to 15 cm [Yoshimura, 2009]. The interaction is based on a zoom-based and level of detail approach: the more the user zooms at a higher resolution the information is rendered. The user is also allowed to add layers to the images with additional information increasing the multi-variation of the information.

Even if more limited in terms of multiscale range than MSV and not dealing with the temporal scale, the user interaction has been proven to be very effective and user-friendly, and it might be taken into consideration in the navigation also of biomedical datasets.

Another software example is CoViz4D[12], which is more focused on the visualisation on variation over of the data. Also in this case the navigation at different spatial scales is based on zoom-in/zoom-out interaction.

An application example, which focuses more on the time scale aspect, is the climate change visualisation. In this field, there are software solutions, which allow the user to visualise huge amount of data over time[13,14]. However, the spatial multiscale problem is usually less relevant than in other applications, while great attention is posed to the multivariate data representation.

2.2 General tools

A considerable effort has been spent in the last three decades in designing effective visualisation methods for inspecting and analysing the ever-growing data provided by experiments, medical equipment, and simulations. Generally, they have a broad scope of applicability and, as such, may cover a central role also in biomedicine.

The computer graphics community concurrently with the formidable progress in systems architectures has recently proved the feasibility of interactively and accurately rendering static volumetric datasets containing billions of multi-value grid elements on high-end distributed machines [Fogal, 2010; Guitián, 2010] as well as on commodity workstations [Crassin, 2009; Guitián, 2010] by means of sophisticated GPU (graphics processing units) algorithms and level of detail mechanisms. Not only visualisation efficiency is fundamental for augmenting scientific productivity, but also the prerogative to provide feature-driven contextual information is essential when, for example, too many data are concomitantly displayed or some features of interest are hidden by occluding regions. This is more important for gaining insights into large and high-dimensional datasets and has been effectively addressed in [Guitián, 2010; Barakat, 2010].

Several simulations, experiments and imaging techniques provide time-varying data. The consideration of the time dimension poses additional challenges for efficiently exploring large volumes. These have been circumvented in some cases by means of advanced compression/decompression techniques [Nagayasu, 2008; Wang, 2010]. However, two common drawbacks shared by the aforementioned

[11] http://www.google.com/intl/it/earth/index.html
[12] http://www.dgi.com/coviz/cvmain.html
[13] http://vis5d.sourceforge.net/doc/
[14] http://www.eonfusion.com/

visualisation methods are the long pre-processing times, and the unavailability of their implementation in unspecialised solutions, e.g. existing open-source software. Furthermore, little or no attention has been dedicated to address truly heterogeneous and multiscale datasets by current visualisation approaches well-suited for handling large data.

Other interesting approaches can be found in information visualisation methods like the one based on data cubes which are commonly used for abstracting relational databases [Stolte, 2003]. For example, the data cubes approach is based on both data and visual abstraction, and on the navigation trough dimensions by zooming a graph where nodes have different levels of detail.

2.3 Biomedical field

Computer-aided visualisation has shifted the way in which biomedicine and biology are approached. The insights gained through simulation and visualisation tools permit us to reveal fundamental aspects and sometimes lead to the formulation of new hypotheses, which can be tested by means of further experiments. The inherent complexity of biological systems, in turn, justified the development of new simulation, visualisation and imaging techniques.

Several software developments have enabled the effective exploration of large molecular assemblies [O'Donoghue, 2010]. For example, features of interest, such as spatial correlations, can be viewed on 3D structures so as to provide a mean to compactly visualise multiple properties at the same time. Additionally, the ability to superimpose structures may enhance discovery further. Global illumination effects can aid the understanding of the spatial arrangement of complex structures considerably [Gribble, 2008], albeit at greater computational and memory costs.

Visualisation of protein interactions, gene expressions and metabolic profile data has been made possible and effective through techniques based on data clustering, networks, connections between protein complexes and dimensionality reduction. The discovery process may be substantially enriched with editing tools that can synthesise new configurations or curate existing ones [Gehlenborg, 2010].

In the review recently appeared on Nature Methods [Walter, 2010], several other aspects encountered when visualising the spatial scales of the cells up to those of the organisms are pointed out. The integration of data originating from multiple simulation tools means solving the recurring problem of having them available in different file formats, and different reference systems. Moreover, the ability to efficiently and intuitively query, analyse and compare data is paramount and has recently reached a satisfactory status in some contexts; however, the visualisation of complex biological datasets is beyond the capabilities of existing software packages. Long-term patient-specific studies can reveal important time-varying disease patterns; this pushes forward the need to adopt robust registration techniques and methods that can significantly correct scanning-motion issues, nonlinear deformations and geometrical distortions. In addition, the proficiency to relate data and physical/biological phenomena may require employing algorithms developed within the fields of image processing, statistical analysis, mathematical modelling and simulation and machine learning, all possibly approached via user-defined visualisation modules, which is valuable.

The importance of on-the-fly computation to interactively refine and iterate analyses has been already emphasised [Nielsen, 2010] together with the role played by visualisation methods that enable the comparison of data of the same category, and the possibility to hide part of them when too much information is being displayed on the screen.

Another aspect of practical interest for characterising and sharing multivariate and multiscale data, common to various research domains, is represented by the skill of annotating data provenance, functionally important features alongside with their integration, and the amalgamation of simulation and clinical information sources. In particular, Gehlenborg et al. [Gehlenborg, 2010] highlights that “truly integrated visualisation of systems biology data across the entire range of possible data types is still very much in its infancy”.

Moreover, the many limitations of modern visualisation tools demand more intuitive software interfaces, new ways to represent spatiotemporal datasets, and the ability to seamlessly navigate across different resolution levels whilst being able to interactively aggregate data at run-time where needed most.

If we look at software tools, we can find a couple of relevant examples here briefly described.

The BrainMaps project[15], which is an interactive multi-resolution brain atlas that is based on over 20 million megapixels of sub-micron resolution, annotated, scanned images of serial sections of both primate and non-primate brains. The interface is a web-based one and it allows zooming in the different scale levels being built on top of one dataset available at very high resolution. Even if very effective, this example does not deal with the time scale problem and the data are only represented with 2D slices, which can be limiting in some biomedical applications.

Another approach with a web-based interface has been recently presented by Google: the Body Browser tool[16]. The tool allows exploring the body organs using a navigation widget and different level of transparencies. The tool is mainly a demonstrator for the WebGL technology and now taken up now from a private company developing 3D anatomical models. Its main limitation in the MSV perspective was that the data were not multiscale (just organs available to be navigated) and not patient-specific (what is visualised is a pre-processed model of a "generic" man and woman). In any case, the ideas and widgets to manage transparency and navigate objects, which are one hidden from another, might be an interesting approach in some use case scenarios.

In terms of software tools, another recent development in the biomedical domain is one of the results of the EU-funded LHDL project [McFarlane, 2008; Viceconti, 2010]. Two prototypes were implemented based on the Multimod Application Framework[17] providing an interactive visualisation environment for biomedical data defined over different spatial or temporal scales.

In summary, the temporal data visualisation and interaction approach, tested with Electromyography scalar data at high frequency, was based on the perceptual limitation of our visual system: there is a physiological limitation to the frequency humans can visually perceive. However, by providing the user the possibility to slow down or to speed up the time in a certain interval of the data, the user was allowed to explore the entire range of frequencies represented in a single signal, or in multiple fused signals. For what concerns the spatial scales, even if more complex, as dealing with 3D dataset (CT and microCT scans), the basic approach was the same. Given a current view, the user could interactively define the level of zoom: when an object is too small or too big to be perceived, it disappeared from the visualisation using a placeholder only for objects that are smaller than the current view. With this approach all data objects in the scene are always visible, independently from their size allowing the user to have an understanding on where the different data are. These two prototypes are very interesting and already provide some understanding on a possible type of interaction, which can be of interest for MSV objectives. The spatial scale example examined many aspects of the click-and-zoom interaction in 3D, such as a minimal set of GUI controls, the speed of zoom, the behaviour of the sub-scale placeholders, coupling between the scale and the GUI, and the importance of knowing which objects were the subject of the camera's current "attention". The main limitations to be overcome with respect to these prototypes are: the missing coupling between spatial and time multiscale navigation; the generalisation of the approach to different types of data (so far only 3D volumes were considered); the visualisation of placeholders embedded inside larger objects; the problems of units and ill-conditioning when combining scales many orders of magnitude apart; and the management of the glyphs in case of a high number of data in each scale should be improved.

[15] http://brainmaps.org/
[16] http://www.zygotebody.com/, recently taken up by the company developing the 3D anatomical models
[17] http://www.openmaf.org/

Figures 8 shows a sequence of example images from the click-and-zoom view. The data consists of a femur volume CT image, visualised in a slice view, with embedded microCT and nanoCT images of the bone structure.

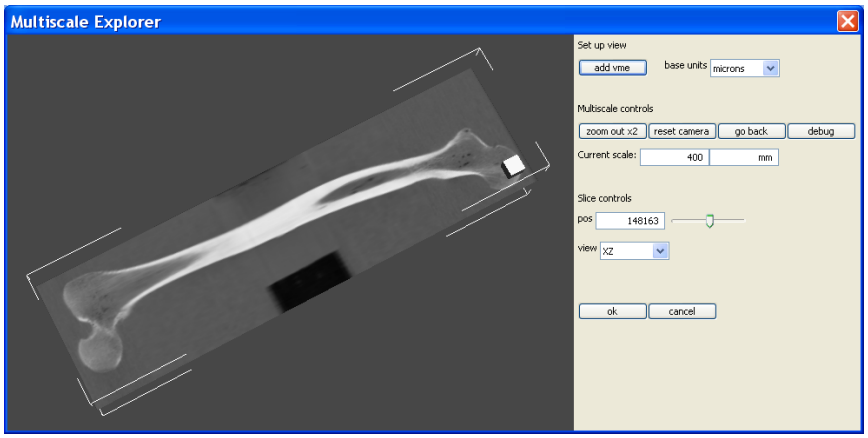


Figure 8a - Slice view of femur (400mm) showing placeholder for embedded microCT

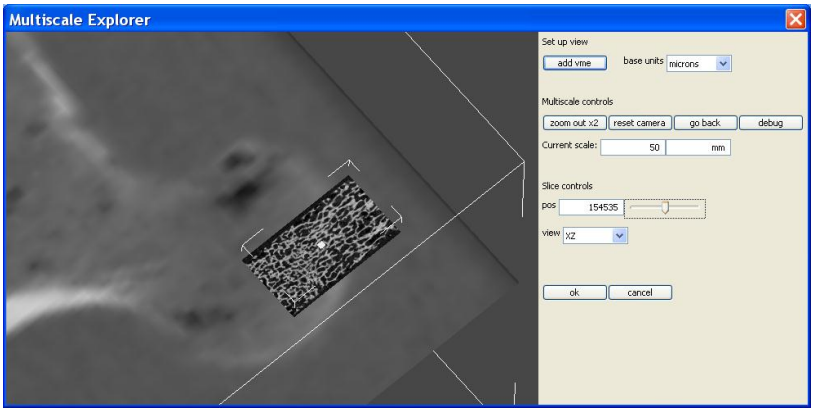


Figure 8b - Slice view of femur head (50mm) showing microCT data of trabecular structure, with further placeholder for embedded nanoCT.

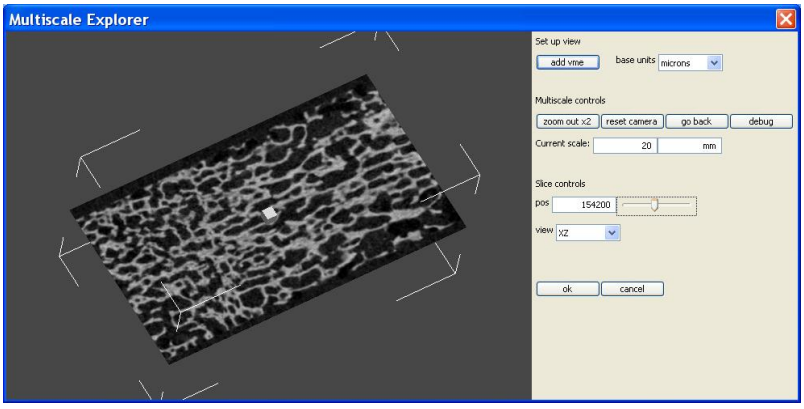


Figure 8c - Slice view of trabecular structure (20mm) with placeholder for nanoCT; the femur is out-of-scale (too large) and is no longer visualised

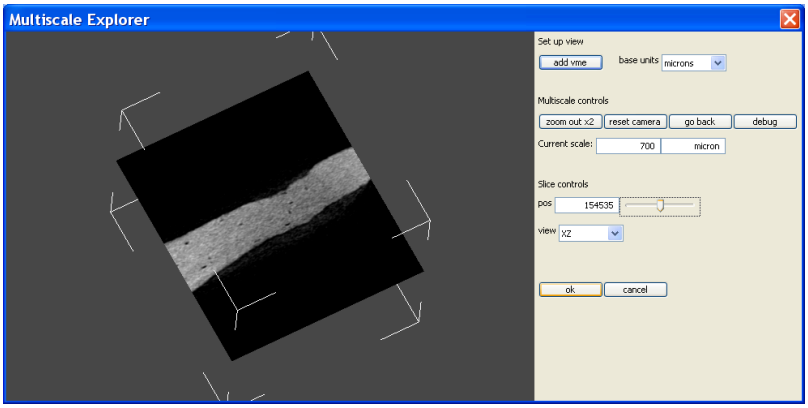


Figure 8d - Result of clicking on nanoCT placeholder: the dataset is expanded to show the single trabecula (700 micron). The larger out-of-scale datasets are not visualised.

3 Exemplary problems (updated section)

In order to motivate better the MSV problem, convincing the reader about the importance of multiscale visualisation in very different biomedical domains, a series of use cases have been collected. Then, among all possible use scenarios, a collection has been set up consisting of data, which are:

- public, that can be used by the MSV partners but also from external users to test the multiscale library implementation, according to their specific data licence (see D3.2 for details);
- presenting one or more multiscale challenges;
- representing a clear clinical value, and potential usefulness for the scientific community in the near future.

All the descriptions, together with the data, file descriptions challenges covered can be found at the MSV public Data wiki (https://www.biomedtown.org/biomed_town/MSV/reception/wikis/Data), while in this report just a brief summary is reported. The data examples have been grouped by biomedical domain, and each example is provided with a description, one or more snapshot, and links to download the data where authorisation have been got.

3.1 Data collection

1. Cardiology

1.1. Fibre multiscale visualisation of the myocardium: it shows the fibre structure of the heart, with its helix shape form in both ventricles. The data is coming from an ex-vivo heart of a pig acquired with a MR scanner using a DTI protocol. In order to check the relation between the data coming from the organ level, and the one coming from the cell level, it is important to be able to see both data in the same reference system, to show that orientation of the macroscopic representation of the fibres corresponds to cellular structure at a particular position.

1.2. Multiscale visualisation of the propagation of the electrical signals in the heart: these data show the electrical activity in the cardiac conduction system and, in particular, in the Purkinje network and the myocardium of a patient specific structure following electrical stimulation at selected points of the above network. The Purkinje network and the myocardium are respectively represented with 1D linear elements and a 3D polyhedron mesh disposed along curved lines. The simulated electrical activity is given in form of a scalar field whose discretization points are the nodes of those computational 1D/3D descriptions. The Purkinje network and the myocardium mesh are associated with the tissue and organ levels, respectively. The electrical activity can be represented in form of a time-independent colour map, wherein the colour represents the arrival time following the electrical stimulus (featuring null time).

1.3. Electrical simulation of the heart: this use case shows an electrical simulation of a heart using several image modalities as input. Input images have been segmented and a surface heart model is used to create a volumetric mesh that will be the input for the simulation engine. A model for human ventricular tissue is used to simulate the electrical propagation in the Midmyocardial Cells.

1.4. Electro-physiological dataset: this use case shows a dataset of electro-physiological CARTO points with some sample ECG signals. This data allows studying electrical behaviour of the heart and can be used to model and plan the success of resynchronization therapy. Electrical data are sampled at 1KHz for a period of 2.5 s, having a total of 2500 samples for each point. The number of channels and their meaning depends on the catheter that was used for the acquisition. In this case, we have the 22 channels corresponding to superficial unipolar and bipolar ECGs (I, II, III, AvR, AvL, AvF, V1, V2, V3, V4, V5 and V6 derivations), six unipolar signals

measured in the tip of the catheter (with six different distances from the catheter tip), and three bipolar signals.

1.5. Heart model: this example shows a left ventricle heart model segmented over the input images with modality US, and MR. This is the first step for patient specific treatment.

1.6. Heart model: this example shows a left ventricle heart model segmented over the input DICOM images with modality MR. This example provides Cardiac MR images and surface model of the beating heart.

1.7. Heart tissue: this example shows a generated tomographic volume of heart tissue from mouse left ventricle acquired using confocal and electron tomography.

1.8. Cardiac coupled electro-mechanics and propagation to torso skin: this model solved coupled electro-mechanics in the left ventricle model over some 300000 cell/grid points embedded in a 128 element mesh over which the geometry is described by a tricubic-Hermite interpolated field. For each element in the mesh, electrical dipoles have been generated to summarize the net effect of the potential flow, and these have been used to simulate the forward problem to model the electrical potential on the torso surface arising from the ventricle model.

2. Cerebral aneurysms

2.1. Cerebral aneurysm flow dynamics: 3DRA image modality is used to obtain a good image quality of the patient aneurysm. Input image before treatment is segmented to obtain a surface mesh of the brain vessels. Aneurysm is located in one of the vessels and some mesh processing filters are used to isolate it. The output surface mesh is used to compute several descriptors. Finally, blood flow is simulated within the segmented vessel using a CFD simulator like ANSYS[18] to analyse the aneurysm haemodynamic.

3. Musculoskeletal modelling

3.1. From body to microCT data of the human bones: this set of data was collected as part of the LHDL EC-funded project and represents a data collection from the body level down to the microCT level. The data collected are of different types: images, surfaces, measurements, etc.

3.2. Lumbar spine surgery: this example shows a MR image of the patient's spine with a preliminary segmentation of the spinal discs.

3.3. Osteoporosis: this example shows a DXA image of a patient's femur head. Dual-emission X-ray absorptiometry (DXA) is a means of measuring bone mineral density.

4. NeuroImaging

4.1. DTI with follow up: two DTI image data sets have been scanned for a healthy control subject at two different time instants using two different scanners and acquisition protocols, and with a separation period of 4 years. In this example the fibre's structure can be extracted and visualised for the two different time instants to look for significant differences or structural changes.

4.2. SISCOM analysis: about 10% of patients with epilepsy have very frequent seizures that are not controlled by medication and adversely affect their quality of life. These patients should be evaluated for epilepsy surgery. In doing so, it is important that we can clearly spot seizure location in the brain. Subtraction Ictal SPECT Co-registered to MR (SISCOM) is an imaging technology developed to pinpoint epilepsy uptake. In SISCOM imaging inter-ictal brain perfusion SPECT images are subtracted from ictal SPECT and the final image is co-registered and superimposed with an MR image.

4.3. EEG-fMRI: localization of the epileptogenic zone is pivotal in the evaluation and treatment of patients with intractable partial epilepsy. The epileptic zone will be removed to improve

[18] http://www.ansys.com

patient's life. Combining EEG simultaneously acquired with functional magnetic resonance imaging (fMRI) yields regions of activation that are presumably the source of spiking activity. These regions are highly linked with epileptic foci and epileptogenic lesions in a significant number of patients.

4.4. Purkinje neuron: these data sets consist of a Purkinje neuron from rat cerebellum injected with Lucifer Yellow and imaged using confocal microscopy; a Purkinje dendrite which shows a tomographic reconstruction of stained Purkinje cell dendrite from rat cerebellum; and a Purkinje Neuron Actin, which is the subcellular distribution of F-actin (protein) in cerebellar Purkinje cell spines.

4.5. Cortical neurons from Alzheimer disease patients: this example shows a reconstruction of a cortical neuron from biopsy material obtained from Alzheimer disease patient. The image shows the major subcellular structures in the cell body of a cortical pyramidal neuron: Golgi apparatus, lipofuscin, nucleus, nucleolus, paired helical filament, and plasma membrane.

5. Oncology

5.1. MR data, histopathology and gene data of a cerebral tumour: this example shows cerebral tumour at different space scales levels: MR data of the head of the patient and microscopy image of the tumour tissue, the microscopy image used to perform several genetic analyses, and some clinical data.

5.2. Breast Radiotherapy DICOM Data: Computed Tomography scan of a patient's breast with the radiotherapy rays of the treatment.

5.3. MR guided prostate interventions: prostate cancer has the second-highest mortality rate of all cancers in American men. This dataset contains two MR images (T1 and T2) and a level map (Label) of 4 prostate regions: the prostate, the tumour, and structures to be avoided (such as the neurovascular bundle) that help to locate these regions during the intervention. The dataset also contains the segmented regions as a surface mesh. To use this pre-operative data, the intra-operative MR image needs to be registered in real time.

5.4. Mammography: the Digital Database for Screening Mammography (DDSM) is a resource for use by the mammographic image analysis research community. Each study includes two images of each breast, along with some associated patient information (age at time of study, ACR breast density rating, subtlety rating for abnormalities, ACR keyword description of abnormalities) and image information (scanner, spatial resolution, etc.). Screening mammography typically involves taking two views of the breast, from above (cranial-caudal view, CC) and from an oblique or angled view (mediolateral-oblique, MLO).

5.5. Lung cancer: this example shows a CT image of a lung with the cancer nodules annotations.

6. Virtual Colonoscopy

6.1. 3D high resolution CT image of abdomen: CT scan image at resolution of 512x512x889 of the patient abdomen allows performing virtual colonoscopy, after colon segmentation. Basic segmentation of colon has been performed using connected threshold algorithm and marching cubes.

7. Human Anatomy

7.1. BodyParts3D: this is a dictionary-type database for anatomy in which shapes and positions of body parts are represented by 3D human models. The original dataset has 865 files in OBJ format. In order to manage the data easily, these files have been grouped depending on the organ system. The number of files has been reduced to 17 VTK files with a size of 217 MB.

8. Mouse Atlas

8.1. µMRI Atlas of Mouse Development: this data set is a 3D digital atlas of normal mouse development constructed from magnetic resonance image data. There are six atlases Theiler Stages (ts) 13, 21,23, 24, 25 and 26 and MR data for an un-labelled ts19 embryo.

8.2. EMAP, the e-Mouse Atlas and EMAGE, e-Mouse Atlas of Gene Expression: this is a 3D anatomical atlas of mouse embryo development including detailed histology. EMA includes the EMAP ontology of anatomical structure. EMAGE is a database of in situ gene expression data in the mouse embryo and an accompanying suite of tools to search and analyse the data. mRNA in situ hybridisation, protein immunohistochemistry and transgenic reporter data is included.

9. Zebrafish embryo

9.1. Digital Fish Project: the aim of the data is to use in toto imaging of developing transgenic zebrafish embryos on a genomic scale to acquire digital, quantitative, cell-based, molecular data suitable for modelling the biological circuits that turn an egg into an embryo. In toto imaging uses confocal/2-photon microscopy to capture the entire volume of organs and eventually whole embryos at cellular resolution every few minutes in living specimens throughout their development. The embryos are labelled such that nuclei are one colour and cell membranes another colour to allow cells to be segmented and tracked as they move and divide. The use of a transgenic marker in a third colour allows a variety of molecular data to be marked. In toto imaging generates 4-D image sets (xyzt), which can contain 100,000 to 1,000,000 images per experiment.

10. Genetics

10.1. Cardiac Disorders: this example tries to summarize how different types of data ranging from medical imaging to genetics data (including computational models and simulations) can be integrated into a single application to study a relevant disease. Long QT Syndrome (LQTS) is an inherited cardiac disease that affects ionic conduction in the Purkinje fibres. In this example, cardiac images from a sample case were processed to obtain a patient-specific 3D model. A file describing micro array data (Cancer Program Publication) was edited manually to include dummy data showing an over-expression of LQTS susceptibility genes. Then, a simulation of electrical propagation wave was done based on the suspected changes in electrical activity produced by the LQTS susceptibility genes.

3.2 Challenges (updated section)

When a multiscale visualisation problem is faced, there are a series of problems to be addressed. These issues have been defined based also on the analysis of the above-presented use cases and ordered into categories classifying them depending on what they entail and require. It seems natural to think on a first high-level categorization to separate the kind of problem at hand, as visualisation, interaction and data management. Still, all these categories are not exclusive; thus we decided not to group the identified challenges according to any of the categories above.

1. *Visualisation across scales* in time and/or space (different orders of magnitude): going from the molecule level (nano-meters and nano-seconds) to the systemic level or even entire populations (meters and years). This is one of the most challenging problems to address, since VPH problems deal with modelling of different levels, and the interaction between them. In this sense here we can divide this problem into two:
    - Spatial: when the data covers multiple spatial levels: molecules, genes, cells, tissues, organs, systems, populations.
    - Temporal: when the data covers multiple temporal levels: from nano-seconds to years.
    - Of course both types of problems are coupled, being molecules and cell processes orders of magnitude faster than organs or systems processes and being both types of multi-scaling present in many biomedical applications.

2. *Integration of data in the same reference system*: different datasets have to be placed at a correct spatial localization using a common reference system, even if they pertain to very different scales. The proper alignment in a common reference system or the possibility for the user to identify correspondences among the data is mandatory for an adequate interpretation of several datasets either if they are being visualised simultaneously or not. This problem arises from the visualisation of data from different acquisition systems, different acquisition procedures, data obtained from simulations, etc.

3. *Very large (Gb) amounts of data in each scale*: at the same scale, medical images, biomedical signals, computational simulations of biological processes, etc. require a high amount of space either for storage, and for loading and visualisation. For this reason, smart solutions for efficient rendering of large datasets are required [Crassin, 2009; Agrawal, 2010]. Due to the increasing performance of acquisition machines, simulation servers, computer software, etc., the amount of pre and post processed data available for clinical studies are also increasing. Therefore, there is a need for fast filtering, storage and visualisation of the data, to show what is relevant to the clinicians to take diagnosis and treatment decisions, and also to researchers to implement and validate their models.

4. *Gaps between different scales* (not all levels are available): the different spatial and temporal resolution of the acquisition devices, makes the representation of biomedical data a multiple data problem, where only sparse scales are modelled. With the current technology, gaps between different levels are always present. With these gaps, the amount of information described by one level in a given space scale, will cover a small amount of space in the level below it, thus, only a very small proportion of the data at a given level can be actually available at the lower scales. It is evident that this will be one of the major issues when dealing with multiscale problems. For this reason, visualising in a continuous way across different scales either is not possible or it has to be simulated.

5. *Multivariate data*: even within the same dimensionality, the data can be of different types and nature, multiple medical image modalities (CT, MR, US), different sizes, spatial resolutions, origins, and also computational models; this entails concurrently approaching several visualisation capabilities, which need to be merged into the same virtual environment [Van Sint Jan, 2006].

6. *Heterogeneous dimensionality of data*: there is also a need for the integration of data with different dimensions: 0D, such as tags or scalar values, 1D, such as physiological signals, 2D, such as images, 3D such as volumetric data, 3D+t such as time-varying ones [Wang, 2010]. The fusion of information with different dimensionalities represents another challenging issue when dealing with different biomedical data from different sources. There exists several solutions to these kind of problems, such as for instance, the use of volume and surface rendering techniques, or the use of colour maps to represent scalar values, etc. However, the integration of these techniques along different scales requires a common strategy for a seamless visualisation of multiple data.

7. *High dimensionality datasets*: in addition to the integration of different data dimensionality, it is more and more common to have high dimensionality datasets, such as time varying vector fields (6D+t), or tensor fields [Cardenes, 2010] (9D, 18D, 29D, etc.). These entail adopting special visualisation paradigms due to the large amount of memory and processing requirements in the CPUs and GPUs, and the difficulty in pursuing correct and meaningful visualisation. The visualisation of these kind of datasets, have been explored in different fields, especially in Diffusion Tensor Imaging (DTI), and Diffusion Spectrum Imaging (DSI) where tensor fields are rendered to explore the fibre structures of the brain. The integration of these solutions together with the rest of multiscale visualisation paradigms will certainly provide the most advanced visualisation solutions for biomedical data.

8. *Interactive visualisation*:

    - Access at interactive frame rates: the ability to explore such complex datasets at interactive rates is essential for effectively capturing and understanding the phenomena of interest [Lum, 2006], but its realisation represents a major challenge underpinning the need to improve or extend existing rendering methods.
    - Scene exploration: the interaction with multiscale data allows the user to actually browse them, being able to change the point of view of the observer, change between scales, zooming in and out within the same scale, etc.
    - Selection of data: finally the user needs to be able to select a portion of the data, in a given scale for many possible reasons: visualise it better (zooming in or out), centre the view on it, hide or show it, save it, remove it, crop it, perform measures, perform post processing, etc. This can be implemented in different ways, but must be compatible with the multiscale nature of the data, for example using a command panel, where the user can select data from a list, or directly from the viewers, picking a zone, selecting a ROI, using a region based criteria, etc.

The data presented before represent quite a wide range of biomedical domains and, all together, they can be used to test all the identified challenges. This is the summary table mapping the challenges to the datasets available:

	Examples																													
Challenges	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	2.1	3.1	3.2	3.3	4.1	4.2	4.3	4.4	4.5	5.1	5.2	5.3	5.4	5.5	6.1	7.1	8.1	8.2	9.1	10.1	11.1	11.2	11.3
Ch1: Different spatial scales	x	x	x	x			x	x	x	x					x			x								x			x	x	x
Ch2: Registration issues	x		x		x				x	x	x	x	x	x	x	x	x	x		x	x								x	x	x
Ch3: Very large data	x		x				x	x	x	x					x	x	x	x					x						x	x	x
Ch4: Gaps between scales	x						x			x								x							x						x
Ch5: Heterogeneous data types	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x		x	x	x	x	x
Ch6: Heterogenous dimensionality	x							x	x									x													
Ch7: High dimensionality	x							x	x				x															x			
Ch8: Interactive visualisation									x	x										x			x	x	x				x	x	x
Ch9: Time varying issues		x	x	x				x	x																						x

4 MSV visualisation problem formalisation

In this section, we try to provide a formal representation of the visualisation process and in particular of its multiscale aspects.

4.1 Visualisation process

In the literature it is possible to find different attempts to formalise visualisation [van Wiji, 2005; Onodera, 1990; Hutchins, 1999]. However, each of them focused on a specific aspect of visualisation, like the process, its cost model etc. An example of visualisation process representation is given in Figure 10.

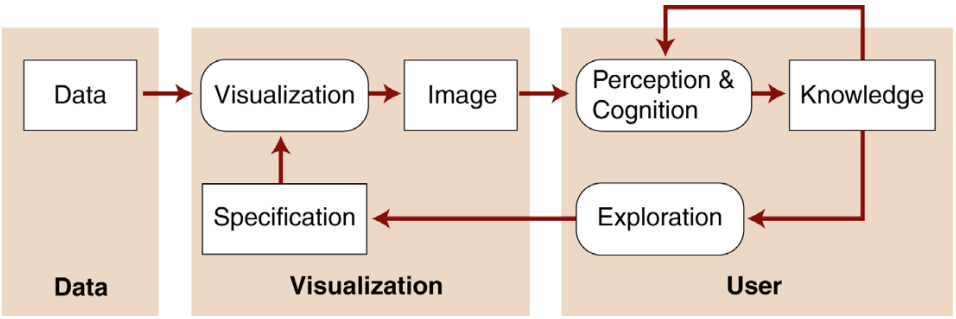


Figure 10 - From [Johnson, 2006]. The visualisation discovery process. Data encompasses the range from a single bit, to time-varying 3D tensor fields, to multi-modal data sources requiring alignment, registration, and fusion, and to non-spatial information sources integrating broad areas of human knowledge. The visualisation specification includes the hardware, the algorithms, and the specific parameters. Users adjust the specification of the visualisation, requiring the introduction of interactive controls. The resulting image will often be an image in the usual sense, but it can also be an animation, or auditory or haptic feedback.

However to the authors' knowledge, nothing is available with respect to the multiscale aspects of visualisation. In this section, we tried to provide a formalisation of the visualisation aspects which relate with multiscale data representation. A glossary of different terms used in the visual process can be found in Annex 1.

4.1.1 Definition of visualisation

First of all we have to define in formal terms the objects, which are visualised. These can be generally represented by a dataset, which can be multidimensional (i.e. a 3D volume), a series of attributes or variables (i.e. attenuation of radiation through a tissue), and a mapping of those variables on the elements of the dataset (i.e. the attenuation value in each cell of the 3D volume); which results in:

- a multidimensional dataset **D** = $\{e_1,\ldots, e_n\}$ contains n samples points or data elements e_i;
- **D** represents the data attributes A = $\{A_1,\ldots, A_m\}$, m => 1;
- The data elements encode values for each attribute, that is, $e_i = \{a_{i,1},\ldots, a_{i,m}\}$, $a_{i,j} \in A_j$;

As described in section 1.2.1, visualisation converts the raw data into images that are presented to the viewer, and it can be represented by a mapping M(V,ϕ) between the dataset attributes above and a visual feature (i.e. colour mapping) thus

- $V = \{V_1 ,\ldots, V_m\}$ identifies a visual feature V_j used to display data attribute A_j.
- $\phi_j : A_j \rightarrow V_j$ maps the domain of A_j to the range of displayable values in V_j.

Based on these definitions, visualisation is the selection of appropriate mapping M(V,ϕ) together the viewer's interpretation of the images produced by M. An effective visualisation chooses among all possible M functions the best one to support the exploration and analysis tasks the viewer wants to perform in the specific field of application. The whole chain of the visualisation formalisation can be represented as in Figure 11.

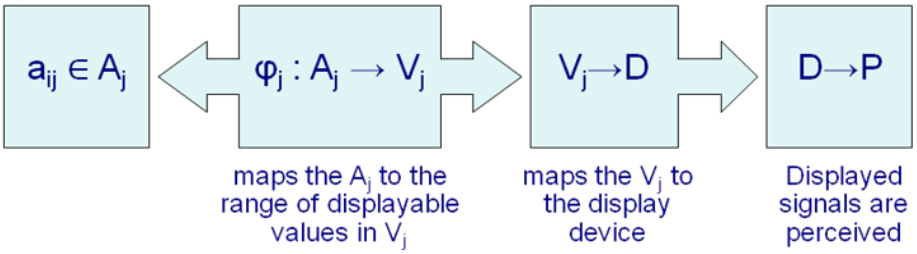


Figure 11 - Visualisation formalisation

Visualisation is thus dependent on a number of aspects: the physical characteristics of the display device (e.g. resolution in terms of the total number of pixels, and the physical size of the display), on the viewer's visual acuity, on the visualisation technique (the methods used to map a data elements values to a visual representation), and on the properties of the data (e.g. its dimensionality, number of elements), and the specific task to be performed by the viewer.

If we consider more in details, the case of a multiscale and/or multidimensional dataset:

- The elements of the dataset D are sampled over space and time, where the attributes A_j they encode are space-time related.
- In order to be representative the visual feature V_j used to represent the attribute A_j is also space-time related.
- The mapping M(V,ϕ) usually preserves the space-time ratios (i.e. zoom, or isotropic scaling)

In choosing the appropriate visualisation approach (the definition of the mapping M) an important aspect is objects distinguishability or resolution acuity, which represents the condition in which two visual features are seen as distinct by the user. As also from the definitions in Annex 1, this characteristics depends on the resolution of the display device and on the visual angle, and in terms of the above V_i can be schematised as:

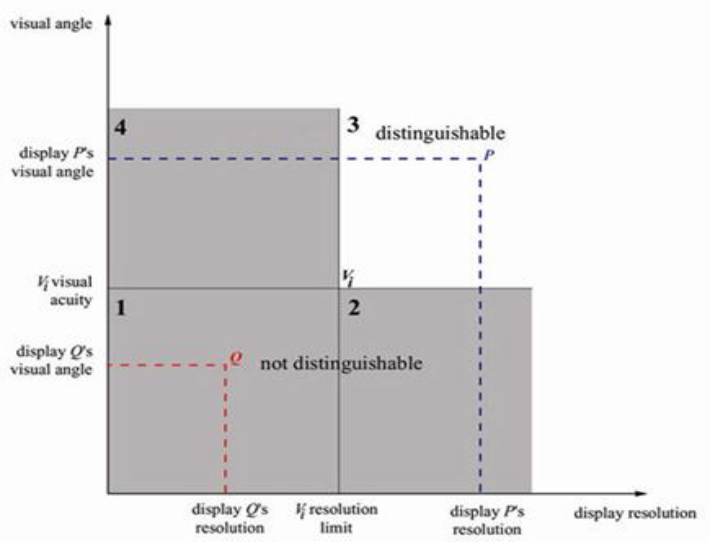
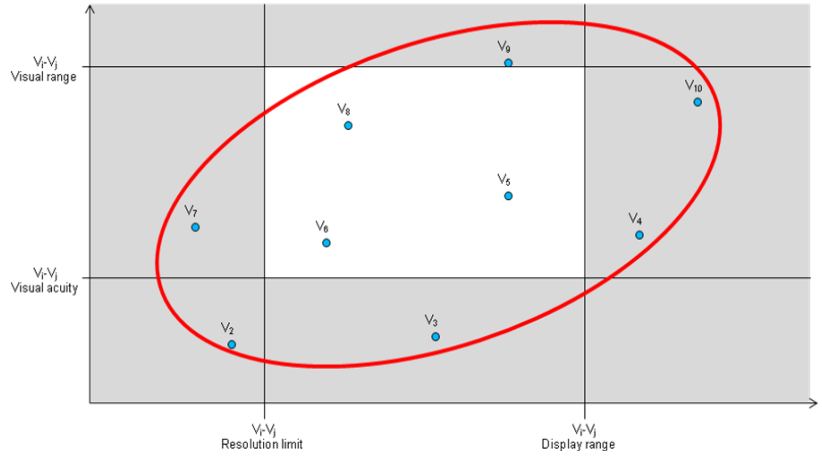


Figure 12 - Distinguishability graph

In reality the distinguishability graph has also upper boundaries for the visual angle (maximum view angle) the for the display resolution (maximum display size). Usually this is not relevant, because we can choose the mapping function M(V,ϕ) to scale the size of visual feature V_j (zoom) used to represent the attribute A_j so that it fits the upper and lower boundaries (i.e. between the pixel and the whole screen).

But what happens if we have to display *simultaneously* multiple V_j that map very different space-time regions? No single scaling can fit them all in the screen!



4.1.2 Degree of multiscaling

In the above conditions, it would be important to have a formal indicator of the multiscaling degree. Thus, we propose to define a degree to indicate severity of the multiscaling (S_d and T_d):

- In space: $S_d = L/a_s$ where:
  - S_d spatial dynamics
  - L largest diagonal of bounding box
  - $a_s \in A_j$ is the smallest size of the smallest attribute of the dataset we need to visualise
- Similarly for time $T_d = T/a_t$

Based on the above indicators we will have for example that the following degrees of multiscaling:

- Whole VPH: from body 10^1m to atoms 10^{-10}m ➔ Sd = 10^{11}
- Bone example: CT 10^{-1}m ; nanoCT 10^{-6}m ➔ Sd = 10^5

But which is the limit an MSV application can tackle? Defining a threshold value to define MSV is very difficult as representational and perceptual limits tend to overlap. Roughly speaking we can say that it is hard to perceive visual features for:

- $S_d > 100$
- $T_d > 100$.

Based on this threshold we will have that a single multiscale dataset contains data attributes to be visualised defined over space-time with a degree of multiscaling > 100, while a multiscale scene contains multiple datasets, each with a degree of multiscaling < 100, but that when displayed together exceed this threshold.

4.2 Formal taxonomy for multiscale visualisation (new section)

Taxonomy is the science of identifying and naming species, and arranging them into a classification scheme or hierarchical structure. The aim of defining a taxonomy for multiscale visualisation to help developers to grasp key techniques of visualisation and domain users to understand how some visualisation methods work and choose the most appropriate one for the specific visualisation purpose.

A number of visualisation taxonomies has been proposed in the literature mostly related to the Information Visualisation fields [Daassi et al, 2006; Chi, 2000; Shneiderman, 1996; Ward, 2002; Lengler et al, 2007]; however, most of them focus on one particular aspect of visualisation (or data to be visualised), and they are not defined to be of general purpose.

With this section we aim to report a taxonomy as much general as possible, but focused on the multiscale visualisation aspects with the objective to provide users of the MSV library with an overview of the most relevant techniques related to multiscale visualisation and interaction. As reported also in other MSV documents (like D3.2 and D4.1) the data type and structure to be visualised is an aspect as important as the techniques when designing an MSV environment. Thus, data classification has been included in the MSV taxonomy, which has not been restricted to the biomedical field. The multiscale techniques have been summarised from D4.1 on the best practices.

- Type of data
  - Field
    - Structured
      - Volume
      - Image
    - Unstructured
      - Point & point cloud
      - Line & polyline
      - Polygon & surface
      - Finite element & Mesh
  - Time-varying
    - Time stamp series
    - Periodicity
  - Metadata
    - Text
    - Dictionary/Ontology
    - Links
  - Dimensionality
    - Scalar
    - Vector
    - Tensor
    - Multidimensional
- Nature of multiscale
  - Dimension
    - Temporal range
    - Spatial range
  - Continuity
    - With gaps
  - Data types
    - Heterogeneous
    - Homogeneous
  - Data size
  - Periodicity in time
  - Scene structure
    - Number of target objects

        - Need for automatic detection of salient scales
        - Occlusions
        - Fibrous structures
        - Ill-conditioning
- Style of interaction
    - Interaction modes
        - Scene-in-hand
        - Fly-through
    - Spatial Interaction techniques
        - Spatial zoom
        - Pan & rotate
        - Click & zoom
        - In-view tokens and widgets
        - External dialog buttons
        - External maps, overlays and panels
    - Temporal interaction techniques
        - Temporal zoom
- Style of visualisation
    - Out-of-core serialisation
        - Chunking
        - Pre-fetching
    - Indication of sub-scale data
        - Placeholder tokens
        - Annotating call-outs
        - Hyperlinks
        - Volume of interest
        - External maps, panels or overlays
    - Magnification of sub-scale data
        - Zoom
        - Hyperlink
        - Split screen
        - Lensing
        - Magnifying call-out
    - Level of Detail
        - Bundling
        - Image resampling
        - Cartographic
        - Data cubes
        - Environment textures
        - Decimation and subdivision
    - Global context
        - External map or overlay
        - Wayfinding widgets
        - Tree
    - Handling of ill-conditioning
        - Scene hierarchy with relative positioning
        - Power-scaled coordinates
        - Mobile origin
        - Double precision
    - Miscellaneous techniques and features
        - Scale saliency
        - Cube maps
        - Temporal gating
        - Dialog buttons

5 Multiscale Visualisation and interaction (new section)

5.1 Problem definition

MSV essential problem is the visualisation of data and information that can be on a wide range of scales both in space and time (from nano scale to the body level scale). The typical data will be a 3D+t dataset of which multiple instances at different scales will have to be displayed together. However, we should also account for the presence of data of other dimensions (0D, 1D or 2D) as described in Section 3. As for the visualisation device, the MSV target will be usually a standard monitor with a resolution of about 1000x1000 and a mouse as input device. The use of advanced hardware (like VR technologies or devices with high degrees of freedom) is not considered in this phase as mandatory for the designing and development of the MSV library. However, the integration of "advanced" devices will not be prevented by the final implementation for testing or use in special application fields.

Thus in summary, different constraints will have to be taken into account in the design and development of the MSV library:

- information will be
  - on very different spatial and temporal scales, going from the molecule (nano-metres and nano-seconds) up to the body level (metres and years) (Figure 13 as example);
  - in different forms (medical images, computer models, signals, etc.);
  - of heterogeneous dimensionality (1D, 2D, 3D, 3D+t);
- visualisation should be interactive even if very large (GB) volumes of data are available at each scale;
- depending on the problem at hand, there may be gaps in the scales (not all levels are available) but visualisation should, nevertheless, be continuous across the scales;
- data in different scales usually have different systems of reference: proper definition of the relative position or correspondence in a common reference system is essential for an adequate interpretation of this information.

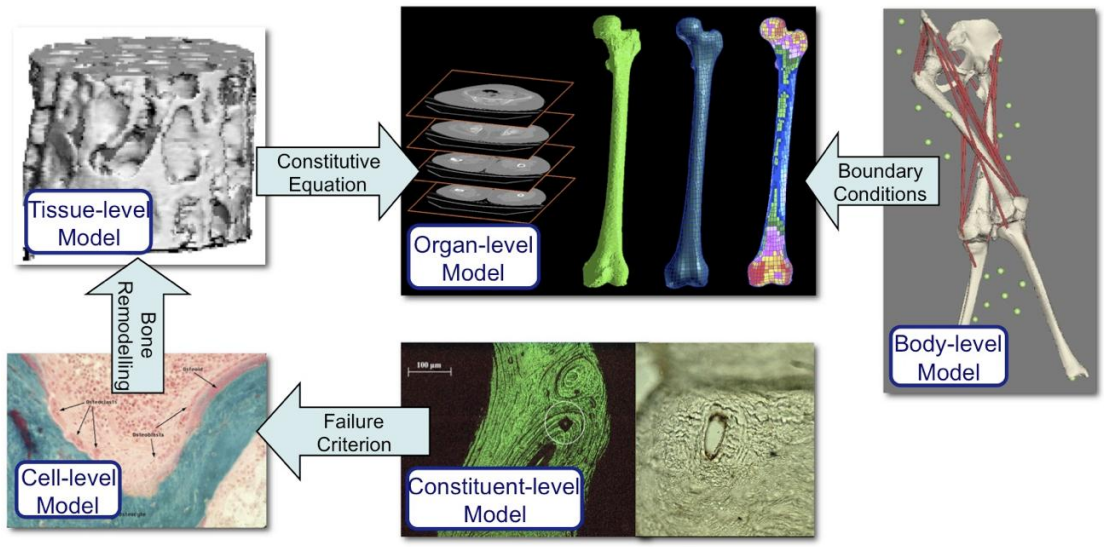


Figure 13 - Example of multiscale data integration and visualisation needs from the VPHOP project for the prediction of fracture in osteoporotic patients.

It is evident from the table at the end of Section 3, that not all the challenges are equally represented by the exemplary data collections. Thus, a priority has been given to the challenges according to the number of available examples: addressing with MSV a more frequent challenge would give higher benefits to the community.

These are the challenges listed in priority order:

- Heterogeneous data types
- Registration issues
- Different spatial scales
- Very large data
- Interactive visualisation
- Gaps between scales
- Time varying issues
- High dimensionality
- Heterogeneous dimensionality

For what concerns heterogeneous data types, all the end-users frameworks, which will be integrating MSVTK in the course of the project (i.e. MAF, GIMIAS, and VTK itself), already provide support to heterogeneous data types. For this reason, no specific implementation is foreseen on this challenge within MSVTK, but the demonstrators will take into consideration the linkage between data of heterogeneous types.

As already mentioned in previous documents, the registration problem is very dependent on the type of data to be registered and to the purpose of the registration itself. As MSVTK is meant to be a general-purpose library, the implementation of specific registration methods is out of the scope of the project. Manual registration techniques will be provided in the demonstrators to deal with use cases data which are not already registered in space.

Following these considerations, the first priorities to be addressed by MSVTK will be the spatial multiscale issue and the large data object management.

5.2 Proposed interaction and visualisation paradigms

This section provides a description of the strategies and paradigms to be adopted by MSV. MSV will implement as open-source a library, called MSVTK, which will address the interactive visualisation of multiscale data.

The MSVTK will implement the click and zoom interaction paradigm: visual cues are provided for the positions of lower scale data with respect to the whole data, which leads to an intuitive interface for data navigation. MSVTK will use placeholders not only for the representation of lower-scale data, but also for hyperlinks to provide extra information such as documentation, etc. At the same time, it will also investigate the possibility of conveying meaningful information about the represented data through the icon/placeholder shapes and colours, with the aim to optimize information transfer and user experience.

The MSVTK library is being implemented as an extension of VTK (Visualization ToolKit). The implementation has been designed to be as general as possible, and the library will provide a series of software elements, like widgets, that might be used later on in other software development projects to add support to the multiscale visualisation.

In particular, MSVTK relies and extends *vtkButtons*, which will be used to provide the visual cue to the user on the data available also at lower scales, and for the data navigation. A “button” is a geometry with a rectangular region that can be textured. The button is divided into two regions: the texture region and the shoulder region. The points in both regions are assigned texture coordinates. The texture region has texture coordinates consistent with the image to be placed on it. All points in the shoulder regions are assigned a texture coordinate specified by the user. In this way the shoulder region can be coloured by the texture.

As previously mentioned, at all scales there may be large datasets (size not possible to be dealt with on the available hardware), running into many GBs, so a proper management of these data types during the interaction has to be included. The MSV project will rely on previous work carried out to manage out-of-core data using a multi-resolution

and bricking approach[19]. The possibility to apply the same technique to different data types is being investigated, together with its effective integration with the zoom-based navigation approach.

Also extension to basic VTK functionalities to better deal with the time-varying data and/or information is being added.

5.3 Implementation strategy

MSVTK will be released at the end of 2012 under an open-source model (Apache 2.0 licence), relaying on state-of-the-art collaborative development tools.

The core implementation language is C++. C++ was chosen for its flexibility, performance, and familiarity to team members. The toolkit uses the full spectrum of C++ features including const and volatile correctness, namespaces, partial template specialization, operator overloading, traits, and iterators.

Because of the agile nature of MSVTK development, the MSVTK team uses the latest development versions of VTK, CTK, and CMake along with the latest stable version of Qt.

Compile-time binding using methods of generic programming and template instantiation is the preferred implementation style. This approach has demonstrated its ability to create efficient, flexible code. Use of the STL (Standard Template Library) is encouraged. In contrast with many libraries, STL containers are the acceptable for passing collections of data between public and private member functions. The STL is typically used by a class, rather than as serving as a base class for derivation of classes.

Most applications as well as the core of the toolkit are designed to compile on a set of target operating system/compiler combinations. These combinations are:

- Windows XP/Vista/7 with Microsoft Visual C++ (Released within the past 5 years)
- Linux with GCC
- Mac OSX 10.6 (Snow Leopard), OSX 10.7 (Lion)
- Other *nix systems with GCC

Some applications and modules make use of specific GPU configurations to optimize certain calculations. These vary according to system, but there shall always be an un-optimized implementation that will comply with the above portability standards.

5.4 Early demonstrators

The MSVTK components are being used in the development of prototypes, which allow checking, on the collected exemplary problems, the efficacy of the proposed approach. Moreover, in order to verify the implementation generality, different prototypes applications are being developed integrating the MSVTK library tools in other frameworks like the Multimod Application Framework (MAF), GIMIAS, and VTK itself.

This is a first list of demonstrators the consortium is working on:

1. vtkButtons draft implementation: Aim is to visualise multiple datasets inside a MAF-based application which uses vtkButtons as a in interactive tool placed in a mafView. The prototype will first re-implement the LHDL prototype trying to overcome its limitations, and then will be verified to other domains data. The user will be allowed to navigate 3D images datasets at different spatial scales (CT scan at organ level, microCT scan, and nanoscan).
2. Large dataset support management: Original BED code is being generalised and integrated in MAF for the management of out-of-core volumes.

[19] A. Agrawal, J. Kohout, G. Clapworthy, N. McFarlane, F. Dong, M. Viceconti, F. Taddei and D. Testi (2010) Enabling the interactive display of large medical volume datasets by multiresolution bricking. *Journal of Supercomputing* 51(1):3-19

3. Electro-physiological dataset: An application designed using Qt, CTK and VTK will be dealing with a cardiological data example with the heterogeneous data type, sparse data, and time scale issues. Sparse points on the 3D heart surface are registered to ECG data, which are varying over a certain time frame. The 3D points are displayed in a VTK render window. The ECG(s) in a separate VTK chart view(s). Time is represented differently depending on the data.
4. Fiber multiscale visualization of the myocardium: This example focuses on the challenge concerning integration of information in different spatial scales with a big gap between the two spatial scales.
5. Human anatomy interaction: The resolution and the large number of components with a lot of points and cells makes the standard rendering and the interaction very slow. The main MSVTK idea is to dynamically load, use, and manage different resolutions when interacting and rendering the scene and components. In this last case, the issue is not related to the nature of the data, but from the hardware limits. It is challenging to load all the data at once due to memory limits, and have an interactive frame rate (limited by the CPU). MSVTK will allow loading a low resolution of the entire anatomy by default, and then have every component (vtkButton) dynamic and clickable to increase/decrease its resolution.

The technical aspects of the demonstrators will be revised and detailed during their implementation and made publicly available via wiki pages as soon as they will be completed.

6 Related efforts and potential users

Potential users for the open-source library, the MSV project is going to develop, are obviously (apart the project partner which will exploit the library in their internal software development) the different software initiatives, which are ongoing at international level. The most important ones are here summarised:

- CTK: Common ToolKit international initiative[20], whose goals are to provide a unified set of basic programming constructs that are useful for medical imaging applications development, to facilitate the exchange and combination of code and data, to document, integrate, and adapt successful solutions, and to avoid the duplication of code and data; some of the MSV partners are also participating to the CTK initiative and this will facilitate the communication between the two groups, the identification of synergies with other visualisation expert groups, and cross-fertilisation of ideas;

- SOFA: SOFA project[21] an Open Source framework primarily targeted at real-time simulation, with an emphasis on medical simulation promoted by INRIA (France). This framework is more focused on the simulation aspects and thus quite complementary to MSV; in this context it may take advantage of the MSV library to deal with the visualisation aspects;

- XIP: the eXtensible Imaging Platform[22] is another open source initiative mainly promoted in the US; being an environment for the rapid development of medical imaging applications from an extensible set of modular elements, it might integrate the MSV library as well.

From the biomedical research perspective potential users of the library will be:

- VPH toolkit initiative: The VPH ToolKit[23] is a technical and methodological framework to support and enable VPH research - the collaborative investigation of the human body as a single complex system. It is supported by the VPH Network of Excellence[24] and it might be a perfect mean of diffusion of the MSV library within the VPH community.

- VPH related research efforts: as stated before in the document different VPH-related projects have multiscale data visualisation issues in their researches. As already done, direct contacts will be established in the different VPH dissemination events in order to let them aware of the project results. The projects already contacted will be invited to evaluate the MSV prototypes for preliminary feedbacks.

- NAMIC: the National Alliance for Medical Image Computing[25] is a multi-institutional, interdisciplinary community of researchers, who share the recognition that modern healthcare demands improved technologies to ease suffering and prolong productive life. Members of NAMIC already showed interest in the MSV objectives in the early stage of the project. Further contacts will be established in the future.

[20] http://www.commontk.org
[21] http://www.sofa-framework.org/home
[22] http://www.openxip.org/
[23] http://toolkit.vph-noe.eu/
[24] http://www.vph-noe.eu/
[25] http://www.na-mic.org/

7 References

1. Agrawal A, Kohout J, Clapworthy GJ, McFarlane NJB, Dong , Viceconti M, Taddei F, and Testi D, “Enabling the Interactive Display of Large Medical Volume Datasets by Multiresolution Bricking”, Journal of Supercomputing, 51(1):3-19, 2010.
2. Barakat S, and Tricoche X, “An image-based approach to interactive crease extraction and rendering”, International Conference on Computer Science, ICCS, Amsterdam, NL, 31 May - 2 June 2010, 1709-1718, 2010.
3. Bouch A, Kuchinsky A, and Bhatti N, “Quality is in the eye of the beholder: Meeting users’ requirements for internet quality of service”, CHI Letters, 2(1):297-304, 2000.
4. Cardenes R, Muoz-Moreno E, Tristan-Vega A, and Martin-Fernandez M, “Saturn: A software application of tensor utilities of research in neuroimaging”, Computer Methods and Programs in Biomedicine, 97(3):264–279, 2010.
5. Chi EH, “A taxonomy of visualization techniques using the data state reference model”, InfoVis, 69-75, 2000.
6. Constantinescu Z, “Levels of Detail: An Overview”, The Journal of LANA, No. 5, 2000.
7. Crassin C, Neyret F, Lefebvre S, and Eisemann E, “Gigavoxels: Ray-guided streaming for eficient and detailed voxel rendering", in ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games, I3D 2009, Boston, MA, USA, 27 Feb - 1 Mar 2009, 15-22, 2009.
8. Daassi C, Nigay L, and Fauvet M-C, “A taxonomy of temporal data visualization techniques”, Revue Information-Interaction-Intelligence, Revue en Sciences du Traitement de l'Information (A journal in the Sciences of Information Engineering), 5(2), 41-63, 2006.
9. Fenner JW, Brook B, Clapworthy GJ, Coveney PV, Feipel V, Gregersen H et al., "The EuroPhysiome, STEP and a roadmap for the Virtual Physiological Human" Phil Trans Roy Soc A, 366(1878):2979–99, 2008.
10. Fogal T, Childs H, Shankar S, Krüger J, Bergeron RD, and Hatcher P, “Large Data Visualisation on Distributed Memory Multi-GPU Clusters”, Proc. High Performance Graphics, HPG 2010, Saarbrucken, Germany, 25-27 June 2010, 57-66, 2010.
11. Friendly M, "Milestones in the history of thematic cartography, statistical graphics, and data visualization", In Seeing Science: Today. http://www.math.yorku.ca/SCS/Gallery/milestone/ or http://datavis.ca/milestones/, 2008.
12. Gehlenborg N, O'Donoghue SI, Baliga NS, Goesmann A, Hibbs MA, Kitano H, Kohlbacher O, Neuweger H, Schneider R, Tenenbaum D, and Gavin A-C, “Visualization of omics data for systems biology”, Nature Methods, 7(3):S56-S68, 2010.
13. Guitián JAI, Gobbetti E, and Marton F, “View-dependent exploration of massive volumetric models on large-scale light field displays”, The Visual Computer, 26:1037-1047, 2010.
14. Hanson AJ, Fu C- W and Wernert EA, “Very Large Scale Visualization Methods for Astrophysical Data”, Proceedings of the Joint Eurographics and IEEE TVCG Symposium on Visualization, 115-124, 2000.
15. Hutchins MA, “Modelling Visualisation Using Formal Algebra”, A thesis submitted for the degree of Doctor of Philosophy of The Australian National University, http://www.ict.csiro.au/staff/matthew.hutchins/thesis/thesis.pdf, 1999.
16. Johnson CR, Moorhead R, Munzner T, Pfister H, Rheingans P and Yoo TS, “NIH-NSF Visualisation Research Challenges Report”, IEEE Press, 2006.
17. Lengler R, and Eppler M, “Towards a Periodic Table of Visualization Methods for Management”, IASTED Proceedings of the Conference on Graphics and Visualization in Engineering (GVE 2007), 2007.
18. Lum E, Shearer J, and Ma K-L, “Interactive multi-scale exploration for volume classification”, in Pacific Graphics, Conference, also as a special issue of Visual Computer, pp. 622–630, 2006.
19. McCarthy JD, Sasse MA, Miras D, “Sharp or smooth?: comparing the effects of quantization vs. frame rate for streamed video”, Proceeding CHI '04 Proceedings of the SIGCHI conference on Human factors in computing systems ACM New York (USA), 2004.
20. McCormick BH, DeFanti TA and Brown MD (eds.), “Visualization in Scientific Computing”, ACM Press, 1987.
21. McFarlane NJB, Clapworthy GJ, Agrawal A, Viceconti M, Taddei F, Schileo E and Baruffaldi F, “3D Multiscale visualisation for medical datasets”, Proc. 5th Int. Conf. on Biomedical Visualisation, MediVis08, 47-52, London, UK, 8-11 July 2008.

22. Myers BA, “The importance of percent-done progress indicators for computer-human interfaces”, Proc. ACM CHI'85 Conf. (San Francisco, CA, 14-18 April), 11-17, 1985.

23. Müller E, “Simulating astrophysical phenomena: challenges and achievements”, Computer Physics Communications, 169:353-361, 2004.

24. Nagayasu D, Ino F, and Hagihara K, “A decompression pipeline for accelerating out-of-core volume rendering of time-varying data”, Computers and graphics, 32(3):350–362, 2008.

25. Nah FFH, “A study on tolerable waiting time: How long are web users willing to wait? Behaviour & Information Technology”, 23(3):153-163, 2004.

26. Nielsen CB, Cantor M, Dubchak I, Gordon D, and Wang T, “Visualizing genomes: techniques and challenges”, Nature Methods, 7(3):S5-S15, 2010.

27. O'Donoghue SI, Goodsell DS, Frangakis AS, Jossinet F, Laskowski RA, Nilges M, Saibil HR, Schafferhans A, Wade RC, Westhof E, and Olson AJ, “Visualization of macromolecular structures”, Nature Methods, 7(3):S42-S55, 2010.

28. Onodera T and Kawai S, "A Formal Model of Visualization in Computer Graphics Systems", Lecture Notes in Computer Science, Volume 421, Springer-Verlag, 1990.

29. Qin C, Zhou C, and Pei T, “Taxonomy of Visualization Techniques and Systems – Concerns between Users and Developers are. Different”, The State Key Lab of Resources and Environmental Information System, Institute of Geographic Science and Resources Research, Chinese Academy of Sciences, Beijing, China, http://www0.hku.hk/dupad/asiagis/fall03/Full_Paper/Qin_Chengzhi.pdf

30. Shneiderman B, “The eyes have it: a task by data type taxonomy for information visualizations”, Proceedings of Visual Languages, 336–343, 1996.

31. Stolte C, Tang D, and Hanrahan P, “Multiscale Visualization Using Data Cubes”, IEEE Transactions on Visualization and Computer Graphics, 9(2):176 – 187, 2003.

32. Van Sint Jan S, Demondion X, Clapworthy G, Louryan S, Rooze M, Cottenand A, and Viceconti M, “Multimodal visualization interface for data management, self-learning and data presentation”, Surg Radiol Anat, 28:518–524, 2006.

33. van Wijk JJ, "The Value of Visualization," 16th IEEE Visualization 2005, pp.11, 2005.

34. Viceconti M, Clapworthy GJ, Testi D, Taddei F and McFarlane NMB, “Multimodal fusion of biomedical data at different temporal and dimensional scales”, Computer Methods and Programs in Biomedicine, 2010, in press.

35. M.C. Villa-Uriol, I. Larrabide, J.M. Pozo, M. Kim, O. Camara, M. De Craene, C. Zhang, A.J. Geers, H. Morales, H. Bogunovic, R. Cárdenes, A.F. Frangi, "Toward integrated management of cerebral aneurysms", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences; 368(1921): 2961-2982 , 2010.

36. Wang C, Yu H, and Ma KL, “Application-driven compression for visualizing large-scale time-varying data", IEEE Computer Graphics and Applications, 30(1):59-69, 2010.

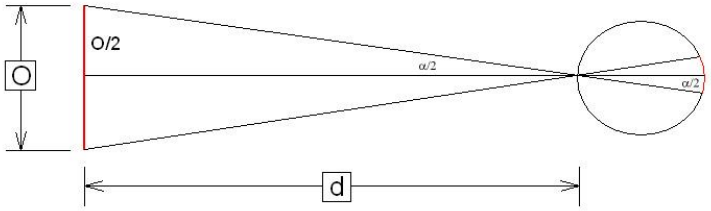
37. Ware C, “Information visualization: perception for design”, (2nd Edition) eds. Morgan Kaufman, 2004.

38. Walter T, Shattuck DW, Baldock R, Bastin ME, Carpenter AE, Duce S, Ellenberg J, Fraser A, Hamilton N, Pieper S, Ragan MA, Schneider JE, Tomancak P, and Hériché JK, “Visualization of image data from cells to organisms“, Nature Methods, 7(3):S26-S41, 2010.

39. Ward MO, “A taxonomy of glyph placement strategies for multidimensional data visualization”, Information Visualization, 1(3), 194-210, 2002.

8 ANNEX 1: Visual process glossary

- **Visual Angle**: Visual angle is the angle subtended by an object on the eye of an observer. Visual angles are generally defined in degrees, minutes, and seconds of arc (a minute is 1/60 degree and a second is 1/60 minute).



If *O* is the size of an object and *d* the distance of the object from the eye, the visual angle α can be calculated as:

Tan(α /2)=(O/2)/d and therefore, α =2 ArcTan [(O/2)/d]

- **Visual Acuity:** is acuteness or clearness of vision, which is dependent on the sharpness of the retinal focus within the eye and the sensitivity of the interpretative faculty of the brain. Visual acuities are measurements of our ability to see details and are important because they define absolute limits on the information densities that can be perceived. Normally visual acuity refers to the ability to resolve two separated points or lines, but there are other measures of the ability of the visual system to discern spatial differences. There are different types of acuities and relative measurements, here are just some examples:

    - *Resolution Acuity:* This is sometimes called "minimum separable" acuity. It refers to the ability to detect a separation, or gap, between objects. A gap is made progressively smaller until the two bars cannot be distinguished from a single bar. This is different from grating acuity where the bars of the grating are progressively narrowed until the pattern cannot be distinguished from a uniform gray field. Resolution acuity thresholds are typically around 0.5 minutes of arc.

    - *Stereoscopic acuity*: is the ability to detect tiny differences in depth with the two eyes. For more complex targets, stereo-acuity is similar to normal monocular visual acuity, or around 0.6-1.0 arc minutes, but for much simpler targets, such as vertical rods, may be as low as only 2 arc seconds.

    - *Recognition Acuity:* Probably the most familiar kind of acuity. The task requires the viewer to name the target stimuli. Many eyecharts use either the Snellen or Sloan lettering system to measure this form of acuity. The rows of letters are made progressively smaller until the subject cannot reliably recognize them. Recognition acuity is typically measured in optotype units, e.g., 20/100, 20/40, 20/20. The optotype line which health, young emmetropic eyes can read at a distance of 20 feet sets the standard.

    - *Detection Acuity:* This is sometimes called "minimum visible" acuity. The subject's task is to detect a either a light or dark target against a background of opposite luminance polarity. As target size gets very small we can no longer resolve differences in size, so the minimum visible acuity reduces to a test of sensitivity (since small targets reflect fewer photons to our eye).

    - *Localization Acuity:* Task is to detect a misalignment of two (or more) objects. One type of localization acuity is called Vernier acuity. This type of acuity is often referred to a hyperacuity, since humans are extraordinarily sensitive to misalignment. The displacement is made progressively smaller until the subject can no longer reliably distinguish the direction. Vernier acuity thresholds are typically around 5-10 seconds of arc.

- *Dynamic Acuity:* All of the above types of acuity are static. Each can be put in motion to measure dynamic visual acuity.

- **Display Resolution:** The resolution of an element on a particular display device (i.e. the number of distinct pixels in each dimension that can be displayed). It can be an ambiguous term especially as the displayed resolution is controlled by all different factors in cathode ray tube (CRT) and flat panel or projection displays using fixed picture-element (pixel) arrays.