Designing Multi-sensory Displays for Abstract Data.

Keith V. Nesbitt

BMath (Newcastle), MComp (Newcastle),

Diploma in Technical Analysis (SIA)

School of Information Technologies

University of Sydney

NSW 2006 Australia

Submitted in partial fulfilment of the requirements of the degree of Doctor of Philosophy

January 2003

Abstract

The rapid increase in available information has lead to many attempts to automatically locate patterns in large, abstract, multi-attributed information spaces. These techniques are often called *data mining* and have met with varying degrees of success. An alternative approach to automatic pattern detection is to keep the user in the exploration loop by developing displays for *perceptual data mining*. This approach allows a domain expert to search the data for useful relationships and can be effective when automated rules are hard to define. However, designing models of the abstract data and defining appropriate displays are critical tasks in building a useful system.

Designing displays of abstract data is especially difficult when multi-sensory interaction is considered. New technology, such as Virtual Environments, enables such multi-sensory interaction. For example, interfaces can be designed that immerse the user in a 3D space and provide visual, auditory and haptic (tactile) feedback. It has been a goal of Virtual Environments to use multi-sensory interaction in an attempt to increase the human-to-computer bandwidth. This approach may assist the user to understand large information spaces and find patterns in them. However, while the motivation is simple enough, actually designing appropriate mappings between the abstract information and the human sensory channels is quite difficult.

Designing intuitive multi-sensory displays of abstract data is complex and needs to carefully consider human perceptual capabilities, yet we interact with the real world everyday in a multi-sensory way. Metaphors can describe mappings between the natural world and an abstract information space. This thesis develops a division of the multi-sensory design space called the MS-Taxonomy. The MS-Taxonomy provides a concept map of the design space based on temporal, spatial and direct metaphors. The detailed concepts within the taxonomy allow for discussion of low level design issues. Furthermore the concepts abstract to higher levels, allowing general design issues to be compared and discussed across the different senses.

The MS-Taxonomy provides a categorisation of multi-sensory design options. However, to design effective multi-sensory displays requires more than a thorough understanding of design options. It is also useful to have guidelines to follow, and a process to describe the design steps. This thesis uses the structure of the MS-Taxonomy to develop the MS-Guidelines and the MS-Process. The MS-Guidelines capture design recommendations and the problems associated with different design choices. The MS-Process integrates the MS-Guidelines into a methodology for developing and evaluating multi-sensory displays.

A detailed case study is used to validate the MS-Taxonomy, the MS-Guidelines and the MS-Process. The case study explores the design of multi-sensory displays within a domain where users wish to explore abstract data for patterns. This area is called *Technical Analysis* and involves the interpretation of patterns in stock market data. Following the MS-Process and using the MS-Guidelines some new multi-sensory displays are designed for pattern detection in stock market data. The outcome from the case study includes some novel haptic-visual and auditory-visual designs that are prototyped and evaluated.

Acknowledgements

Early motivation for this work was gained while working at BHP Research on a project to investigate the use of Virtual Environments for solving industrial problems. Many people were involved in this early work and I would like to thank Steve Woodyatt, Randall Gallimore, Bernard Orenstein, Harmeet Lamba, Adrian Ugray and Tony Murnain for their support and assistance during that time.

Many of the collaborations that began during this early work have carried on during the course of this thesis. I would like to thank a number of people for their input and for providing access to the Virtual Environments necessary for me to complete my work. In particular, Kevin Smith, Duncan Stevens, Chris Gunn, Stephen Barrass, Mathew Hutchins and others at the Interactive Modelling and Visualisation Systems Group, which is part of the Mathematical and Information Sciences Division, CSIRO, Canberra.

I would like to thank Martin Göbel, Bernd Fröhlich, John Plate and other members of the IMK-VE group for their help and support. This group is located in Sankt Augustin, Germany and formed part of the German National Research Centre for Information Technology (GMD) before becoming the Fraunhofer Institute for Media Communication.

I would also like to express my appreciation for funding support from Rudi Vernik and Tim Patterson at the Information Technology Division of the Defence Science and Technology Organisation (DSTO), Adelaide.

A special note of appreciation is extended to my supervisor, Professor Peter Eades for his invaluable assistance and direction during the course of my work.

Finally, of course this work would not have been possible with out the support of family, friends and the odd angel or two (metaphorical not metaphysical). To those people I would like to offer my sincere thanks for your faith and patience.

Preface

All work presented is the original work of the author unless otherwise indicated. Parts of this thesis have been published in the following works.

Keith Nesbitt (2002) "Experimenting with Haptic Attributes for Display of Abstract Data", Proceedings of Eurohaptics, 2002. Edinburgh, UK, pp 150-156.

Keith Nesbitt and Stephen Barrass (2002) "Evaluation of a Multimodal Sonification and Visualisation of Depth of Market Stock Data", Proceedings of the International Conference of Auditory Display, 2002. Kyoto, Japan, pp 233-238.

Keith Nesbitt and Carsten Friedrich (2002) "Applying Gestalt Principles to Animated Visualizations of Network Data", International Symposium on Web Graphics and Visualisation, IV02-WGV 2002, London, pp 737-743.

Keith Nesbitt (2001) "Modelling the Multi-sensory Design Space", Information Visualisation, 2001. Proceedings of the Australian Symposium on Information Visualisation, Conferences in Research and Practice in Information Technology. Volume 9, Australian Computer Society, pp 27-36.

Keith Nesbitt (2001) "Interacting with Stock Market Data in a Virtual Environment", Data Visualization, 2001. Proceedings of the Joint Eurographics IEEE TCVG Symposium on Visualization, SpringerWein New York, pp273-282

Keith Nesbitt, Randall Gallimore, Bernard Orenstein (2001) "Using Force Feedback for Multi-sensory Display", Proceedings of 2nd Australasian User Interface Conference. AUIC 2001. Australian Computer Science Communications, Vol. 23, No. 5. pp64-68.

Keith Nesbitt (2000) "Designing Multi-sensory Models for Finding Patterns in Stock Market Data" Proceedings of the Third International Conference on Multimodal Interfaces 2000, Springer Lecture Notes in Computer Science 1948, pp24-31.

Keith Nesbitt (2000) "A Classification of Multi-sensory Metaphors for Understanding Abstract Data in a Virtual Environment", Proceedings of the IEEE Conference on Information Visualisation 2000, IEEE Computer Society, Los Alamitos, pp493-498.

Keith Nesbitt, Randall Gallimore, Bernard Orenstein (2000), "Investigating the Application of Virtual Environment Technology for use in the Petroleum Exploration Industry", Proceedings of the 23rd Australasian Computer Science Conference ACSC 2000, Vol 22 (1), pp181-188, Canberra, Australia.

Keith Nesbitt and Bernard Orenstein (1999) "Multisensory Metaphors and Virtual Environments applied to Technical Analysis of Financial Markets", Proceedings of the Conference on Advanced Investment Technology 1999, Gold Coast, Australia, pp 195-205. ISBN: 0733100171.

This work has not been submitted for a higher degree to any other University or institution.

Keith Nesbitt

Contents

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Chapter 9

Chapter 10

Chapter 11

Chapter 12

Chapter 13

Chapter 14

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Chapter 10

Chapter 11

Chapter 12

Chapter 13

Chapter 1

Multi-sensory Display

The Sensorama Simulator patented by Morton Heilig.

"As an environmental simulator, the Sensorama display was one of the first steps toward duplicating the viewer's act of confronting a real scene. The user is totally immersed in an information booth designed to imitate the mode of exploration while the scene is imaged simultaneously through several senses. The next step is to allow the viewer to control his own path through available information to create a highly personalized interaction capability bordering on the threshold of virtual exploration." [Fisher 1982]

Chapter 1

Multi-sensory Display

1.1 The Motivation - Data Mining

A problem facing many areas of industry is the rapid increase in the amount of data and how to deal with it. However, these large amounts of data could also be considered a resource. The data often contains many attributes, far more than humans can readily understand, so there may be an opportunity to better comprehend the data and extract useful patterns. For example, the patterns could take the form of previously unknown relationships between attributes of the data. The term *data mining* has been used to describe the diverse methods used to explore abstract data in the search for valuable and unexpected patterns [Groth 1998].

There are lots of possible applications for data mining and these are from such diverse domains as:

- finding drilling targets in petroleum or minerals exploration data
- uncovering business opportunities in marketing data
- detecting abnormal conditions in medical diagnostic data
- recognising new trading rules in stock market data.

The available tools for *data mining* can be considered in two broad categories, *Automated Intelligent Tools* and *Human Perceptual Tools* (figure 1-1). *Automated Intelligent Tools* implement well-defined strategies for finding rules or patterns in data. These systems take advantage of a computer's capability to perform error-free, repetitive tasks and to process large amounts of data efficiently. *Human Perceptual Tools*, on the other hand, display the data to the user and allow the user to search for patterns. These systems take advantage of the human capability to perform subtle pattern matching tasks.

It can be expected that *Human Perceptual Tools* are particularly useful where:

- unpredictable exceptions may occur in the data
- heuristics are required to filter subtle variations in information
- the target is unknown or cannot be precisely formalised by rules.
- the problem requires knowledge that is hard to formalise, such as past experience or domain proficiency.

During the 1990s, the accent for Human Perceptual Tools was on designing visual displays of data. This approach is called *Visual Data Mining* [Soukup 2002]. However, this thesis explores the design of multi-sensory displays for finding patterns in large abstract data sets. For example, as part of a case study, this thesis prototypes a range of *Human Perceptual Tools* designed for finding relationships in stock market data.

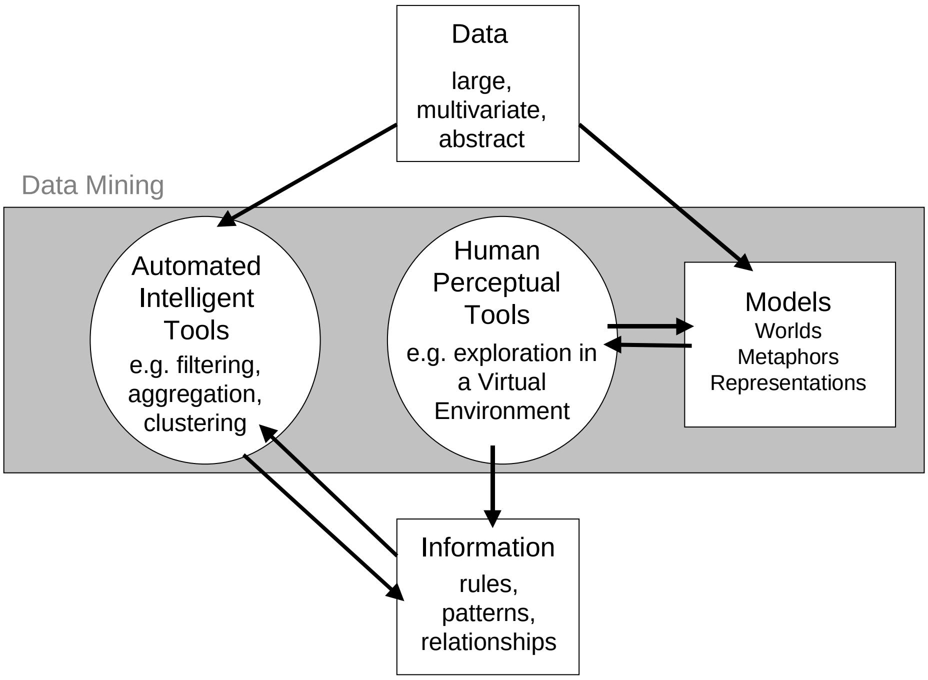


Figure 1-1 A diagram showing the two categories of data mining.

The success of Human Perceptual Tools for data mining relies on effective models of the information being developed. These models should allow for effective exploration of the underlying abstract data. Some issues that need to be addressed are:

- Can new technologies assist in the data mining process?
- How do we build better Human Perceptual Tools for data mining?
- How do we design appropriate multi-sensory models for data mining?
- How do we evaluate the effectiveness of a display for finding patterns in data?

This thesis is concerned with such questions.

To assist with understanding the design of Human Perceptual Tools, this thesis, proposes a new taxonomy called the MS-Taxonomy. The MS-Taxonomy is developed to better categorise the multi-sensory design space. Using this categorization, a new process called the MS-Process is developed. The MS-Process is used to design, prototype and evaluate *Human Perceptual Tools*. The MS-Process incorporates the MS-Guidelines. The MS-Guidelines are a structured set of principles that help to direct design decisions and act as a checklist during evaluation. To illustrate and test these new concepts, the MS-Process is applied in an application domain. In this domain the context becomes one of finding patterns in stock market data.

The fundamental focus of this thesis is the design of multi-sensory displays of abstract data. However, this thesis also indirectly investigates the technology of Virtual Environments[1]. The technology of Virtual Environments is important because it enables Human Perceptual Tools to be built. Indeed the technology itself has motivated the possibility of Human Perceptual Tools [Rheingold 1991]. In the case study described in chapters 11, 12, and 13 this thesis discusses the implementation of a number of applications within different Virtual Environments. For now, the technology of Virtual Environments is an appropriate place to begin.

Section 1.2.1 discusses the origins of Virtual Environments. Section 1.2.2 describes some typical types of Virtual Environments. Section 1.2.3 contrasts the different types of worlds displayed in Virtual Environments. The mechanisms of multi-sensory display in Virtual Environments are discussed in section 1.2.4 and section 1.2.5 concludes with a discussion of how the technology seeks to widen the human-to-computer bandwidth.

1.2 The Technology - Virtual Environments

Computing and network technologies developed rapidly towards the end of the 20th century. However, during that time there were less fundamental changes to the way users interacted with computers. Initially the computer user interface relied on mechanical switches. Later computers were programmed using punched cards. Eventually in the 1970s there was a move to command style interfaces where interaction was primarily with mono-coloured text screens and a keyboard. Throughout the 1980s and 1990s the desktop metaphor became common. Computer users would interact with icons and windows on a high resolution colour display using a keyboard and mouse. At the start of the 21st century there was little change to this user-interface paradigm. The technology of Virtual Environments systems promise to provide the next fundamental change in the way users interact with computers.

Research emphasis is often placed on the systems or hardware used to enable Virtual Environments. It is hardware, such as, high-end graphics computers, 3D display systems or force feedback devices that most readily capture media attention. However, the purpose of this technology is not really to develop new hardware but rather to develop a new style of human-computer interface.

Virtual Environments construct the user interface as a synthetic world. In this computer-synthesized world the user can interact with objects and navigate the environment as if they were in the real world. Virtual Environments promote a natural way of interacting with computers using the human body and all its senses. In Virtual Environments people participate, perform tasks and experience activities within the computer-generated world. The idea is to immerse a person in an environment that allows natural interaction and participation in order to perform tasks.

---

[1] The ideas in this thesis are generated from the cross-section of a number of intersecting but distinct lines of thought. For example, the thesis discusses tools like Virtual Environments, tasks such as data mining and fundamental fields such as perceptual science. Appendix B provides a number of diagrams that can assist navigation through the network of concepts that are foundations for this thesis.

An important aspect of Virtual Environments is that they try to provide interactions that are like those in the real world. This implies interaction between all the sensory modalities. A multi-sensory user interface is a promising enabler of natural interaction. A multi-sensory interface could significantly increase the communication bandwidth between human and computer. The development of Human Perceptual Tools relies on the capability of Virtual Environments to provide the user with such a multi-sensory interface.

1.2.1 The Origins of Virtual Environments

Virtual Environments has been an evolving field in computer science since the first computer-generated Virtual Environment was created by Ivan Sutherland in the 1960’s [Sutherland 1965]. Sutherland went on to demonstrate the first immersive visualisation by creating his prototype head-mounted display in 1968. Sutherland's work in scene generation eventually led to the development of 3D display hardware and the field of computer graphics emerged.

While the science of Virtual Environments was led by Ivan Sutherland, the entertainment industry also recognised the potential of this new concept. In 1960, the concept of a multi-sensory theatre was marketed by Morton Heilig. By 1962 this led to the development of a patented virtual reality arcade ride called the Sensorama Simulator. This arcade ride allowed users to ride a motorbike through New York and experience a 3D view with sound, wind, vibration and even aroma [Hamit 1993].

The potential of Virtual Environments for building aircraft flight simulators drove much of the early work in this field. It cost millions of dollars to build simulator hardware for a single aircraft configuration and the promise of producing the same system in software was enticing. The software could simply be upgraded with the aircraft. This led to a substantial investment by the military in this technology for training. NASA also became interested in simulation for astronaut training and in 1981 they developed their own Virtual Visual Environment Display (VIVED) [Burdea and Coiffet 1994].

The initial focus of the technology was on trying to replicate reality. Being able to mimic the real world was useful for providing experiences, prototyping designs or for training. The concept of a more abstract data world, called *Cyberspace*, was popularised by the science fiction author William Gibson in the 1984 novel, Neuromancer.

> "*A graphic representation of data abstracted from the banks of every computer in the human system. Unthinkable complexity. Lines of light ranged in the nonspace of the mind, clusters and constellations of data. Like city lights, receding*." [Gibson 1984].

However, despite the appearance of a fictional *Cyberspace*, it would be a while before people began to create and explore such abstract computer-generated worlds. Early work concentrated on producing graphics which were as realistic as possible. The focus was very much on generating 3D visual displays. In particular the drive to produce real-time, computer-generated images, for both eyes at something like 60 times a second provided a major research initiative.

Despite this focus on visual display, developing a multi-sensory user interface always remained a goal of Virtual Environments. From 1967, Fred Brooks and colleagues at the University of North Carolina were pioneering the use of force feedback displays for feeling the forces between molecules [Batter and Brooks 1972]. One of the first spatial sound displays was created by Scott Foster and Elizabeth Enzel for NASA in 1988 [Fisher, Wenzel et al. 1988].

By the 1990s Virtual Environments were being publicly demonstrated at SIGGRAPH, where the University of Illinois showed groups of people through their CAVE™ exhibit [Cruz-Neira, Sandin et al. 1993]. It was later in the 1990s that this technology suffered an image problem, due to its strong association with the entertainment industry in the sense of both parlour games (with players wearing large helmet displays) and as portrayed in movies. The technology was the subject of much media hype and therefore was not taken very seriously for business applications.

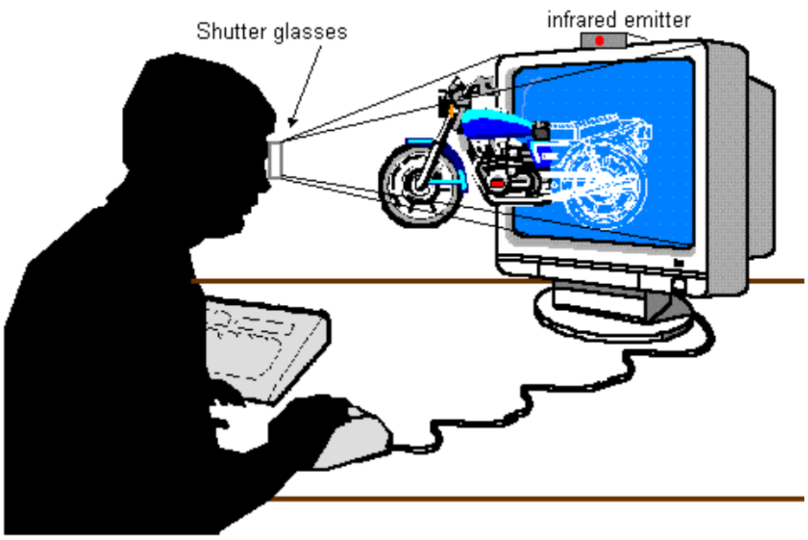
The cost of the technology was a barrier to widespread use, though some industries did make early attempts to take up the technology. In particular, the defence, medical, aerospace and automotive industries experimented with early applications of Virtual Environments. These applications were in areas such as simulation, training and virtual prototyping [Durlach and Mavor 1995].

At the start of the 21st century the viability of the technology increased as computing power increased and costs reduced. This accessibility meant that, a wider range of applications began to be developed. There was particular interest from the petroleum exploration and medical domains. It should be stressed that this technology brings together many disciplines (electronics, software engineering, computer graphics, human-computer interaction, perceptual science, and psychology) with the common objective of significantly improving the human-computer interface.

There have been many different configurations of Virtual Environments developed. These include low-end, single-user platforms such as, a workstation with stereo glasses (figures 1-2), through to more complex head-mounted display systems. Though head-mounted displays were commonly used in the early days, these gradually lost favour in the late 1990s and the focus moved to projection-based displays. Some platforms allow a large group of people to collaborate together in environments such as the CAVE™ [Cruz-Neira, Sandin et al. 1993], the `SGI Reality Center™` [Helsel 2002] or desk-sized displays (figures 1-3, 1-4, 1-5). Another common configuration that found support was the `Responsive Workbench` [Kruger, Bohn et al. 1995] (figure 1-7).

Virtual Environments, like the `Responsive Workbench,` provide the user with the familiar paradigm of manipulating 3D objects above a tabletop. Recently the commercial availability of the `Phantom™` [Salsibury and Srinivasan 1997] device has made it feasible to integrate force feedback into environments such as the `Haptic Workbench` [Stevenson, Smith et al. 1999] (figure 1-7).

A number of different types of Virtual Environments are described in figures 1-2, 1-3, 1-4, 1-5, 1-6, 1-7. These different environments have very different characteristics and these figures include a discussion of the different interaction paradigms provided by each platform.



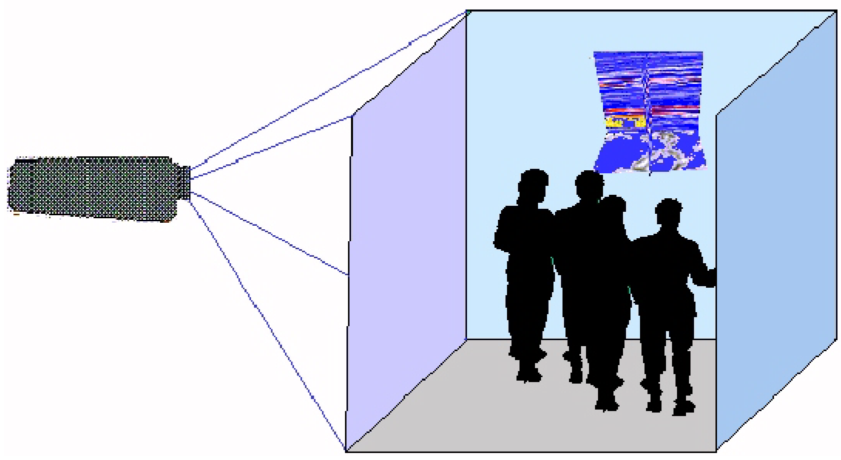
Using active shutter glasses 3D images are generated on a normal workstation. The shutter glasses are controlled by an infrared signal that synchronises the opening of the glasses with the display of left and right eye images on the computer screen. It is possible to use head tracking to support viewpoint changes to the image.

Paradigm: Looking at 3D objects on a screen. Single user. Limited field of view. Low level of immersion.

3D Effect: 3D images on flat screen. Model is restricted by monitor boundaries, so the user cannot look behind the image or from the side.

Control & Manipulation: Direct object manipulation using a mouse or trackball.

Figure 1-2 3D Workstation with shutter glasses.



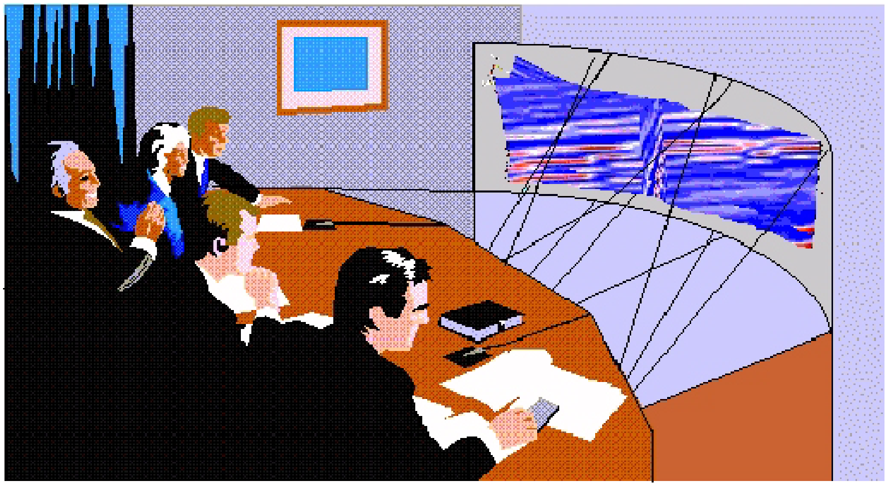
An enclosed space about the size of a small room. The user is surrounded by rear projected displays (left, right, front and floor). By wearing shutter glasses, users perceive the computer images to occupy the same space that surrounds them. A lower cost alternative is the `Wedge` [Gardner and Boswell 2001]. The `Wedge` uses only two walls arranged at an angle to each other.

Paradigm: Immersion of the users within the same space as the object. Supports 4-5 people.

3D Effect: 3D images occupying the same space as the user. Created by active shutter glasses. View is calculated from the perspective of a single head-tracked user.

Control & Manipulation: Menu and object manipulation using a virtual pointer and specialised interface widgets. Typically the virtual pointer is associated with a hand held device that is tracked.

Figure 1-3 The CAVE™ [Cruz-Neira, Sandin et al. 1993].



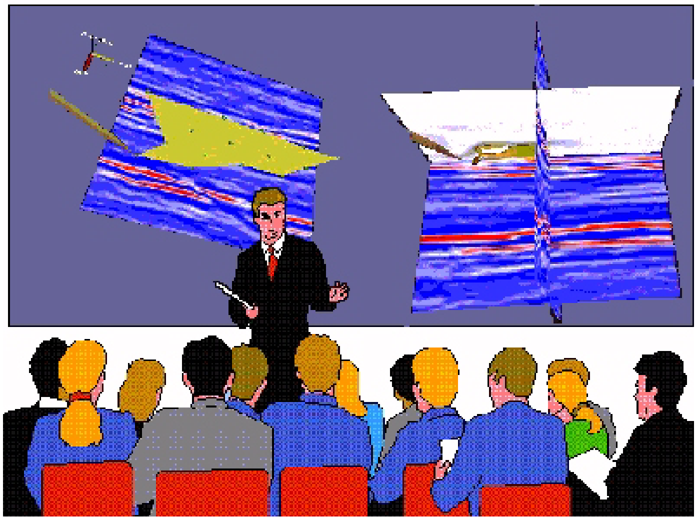
A bench sized workstation for a small group of people to work around. It can support a single application or multiple windows. A small projection screen uses three front projectors to produce a high resolution image with a wide field of view.

Paradigm: A group desk with a wide field of view. Often used as a large screen display, running existing 2D software applications. Supports 2-5 users.

3D Effect: Can support 3D using active shutter glasses or passive polarisation techniques.

Control & Manipulation: Typically, objects are directly manipulated using the keyboard and mouse as this is currently supported by available software.

Figure 1-4 A collaborative desk-size display.



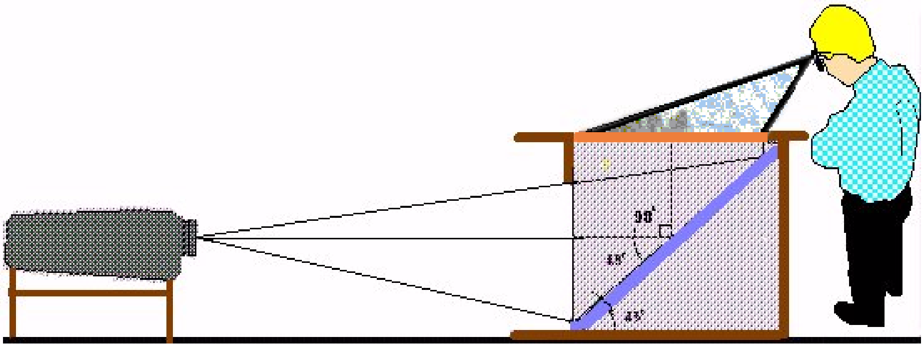
A Reality Centre™ [Helsel 2002] is a room-sized environment with a large projection screen. Multiple projectors are used to produce a bright, high-resolution image. Sometimes these environments use a straight screen, at other times the screen is curved. The aim is to fill as much of the field of view as possible. The latest generation of these environments is the i-CONE™ [Barco 2001]. This is a large curved display design to have improved sound display characteristics.

Paradigm: Large high resolution screen display with a wide field of view. It supports large groups of about 3-20 people. Often used as a large screen display, running existing software applications.

3D Effect: Often not used but it can support 3D images using active or passive stereo.

Control & Manipulation: Typically manipulation is directly of object by keyboard or mouse.

Figure 1-5 The SGI Reality Centre ™ [Helsel 2002].



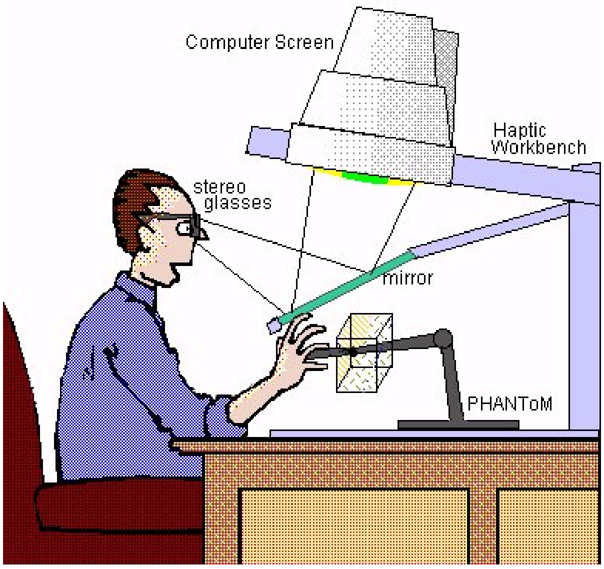
The bench mimics a normal desk but has a projection screen as part of its top surface. This allows a computer display to be projected on this surface. Using shutter glasses it creates the illusion of a 3D Model projected just above the bench top. The original version used just a single display surface [Kruger, Bohn et al. 1995]. Later versions of this display, are L-shaped and project images onto both a top and a rear display surface.

Paradigm: Manipulating 3D objects in a virtual working volume above a tabletop. The normal working paradigm is to be looking down at objects and interacting with them. Supports about 2-3 users.

3D Effect: 3D images are projected in virtual space above the bench top. These 3D images are created using shutter glasses. View is typically calculated from the perspective of a single head-tracked user, though this has been extended to provide head-tracking for two users [Hanrahan and Fröhlich 1997].

Control & Manipulation: Typically manipulation of objects is direct using a virtual pointer or other virtual tools controlled by a hand-held prop.

Figure 1-6 The Responsive Workbench ™ [Kruger, Bohn et al. 1995].



Specially designed environment to allow haptic interaction. A 3D image is created using shutter glasses. The model is reflected in a mirror from a suspended monitor to create the illusion of a virtual workspace. Interaction is by the Phantom™ [Salsibury and Srinivasan 1997]. This device provides force feedback at a point in the 3D image.

Paradigm: Single user environment. Direct manipulation of 3D objects in a virtual workspace coinciding with the user's hands.

3D Effect: Uses shutter glasses with head tracking to maintain perspective correct view.

Control & Manipulation: Left hand manipulation of objects to allow rotation and zooming using a tracked prop. Force feedback to right hand allows objects to be felt. The right hand for is also used for selection and direct manipulation of objects.

Figure 1-7 The Haptic Workbench [Stevenson, Smith et al. 1999].

1.2.2 Virtual Environment Platforms

This thesis does not aim to develop new hardware solutions. However the availability of the hardware enables this work to proceed. An important aspect of Virtual Environment research is the study of different interaction paradigms. However, this thesis does not focus on interaction. Interaction is an important aspect of information displays but it further the possible design space and so is excluded from the scope of this thesis.

As previously mentioned a number of different Virtual Environment systems have been developed. This thesis develops and evaluates multi-sensory prototypes using a wide range of these environments. They include:

- the `CyberStage` [Eckel, Göbel et al. 1997]
- the `Wedge` [Gardner and Boswell 2001]
- the `i-CONE™` [Barco 2001]
- the `L-Shaped Responsive Workbench` [Krüger, Bohn et al. 1995]
- the `Barco Baron™` [Barco 2001]
- the `Haptic Workbench` [Stevenson, Smith et al. 1999]

The `CyberStage` was developed at GMD, the German National Research Center for Information Technology. The `CyberStage` is similar to the `CAVE™` [Cruz-Neira, Sandin et al. 1992] developed at the University of Illinois-Chicago. These environments enclose a space about the size of a small room. Basically the shape of the space is a cube surrounded by 4 projection surfaces [Brown-VanHoozer, Hudson et al. 1995]. Computer generated images are displayed onto the three walls and the floor of the environment. By using shutter glasses the user is immersed in a 3D model. Speakers located in the floor and at the corners of the `CAVE™` provide auditory feedback [Cruz-Neira, Sandin et al. 1993]. A similar configuration provides audio feedback in the `CyberStage` [Eckel 1998].

Figure 1-8 The CyberStage [Eckel, Göbel et al. 1997] at GMD. A group of users are shown interacting in an immersive scene.
Courtesy of: Fraunhofer Institute for Media Communication, Germany.

Many visualisation applications have been developed for the CAVE™ and the CyberStage. These applications include, architectural walkthroughs [Airey, Rohlf et al. 1990], motor car prototyping [Durlach and Mavor 1995], simulation of fluid dynamics [Wesche, Kraemer-Fuhrmann et al. 1997] and scientific visualisations of objects in wind tunnels [Bryson and Levit 1992].

Developed at the Australian National University, the Wedge [Gardner and Boswell 2001] is a more economical version of the CAVE™. It uses a projected display onto 2 walls set at an angle to each other. As is frequently the case, the 3D images are viewed with shutter glasses, and the user's head position is tracked to update the user's viewpoint.

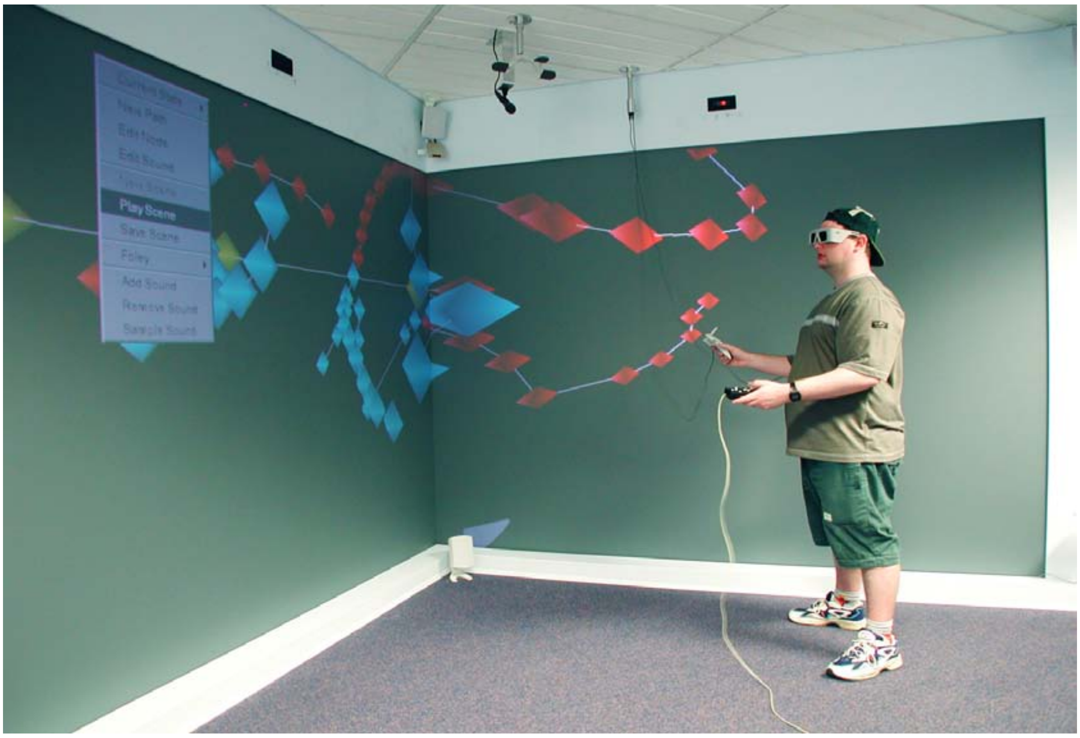


Figure 1-9 The Wedge [Gardner and Boswell 2001] at the Australian National University.
Courtesy of: Australian National University, Canberra.

A number of large screen displays, sometimes called Visualisation Centres have been developed. These include the Reality Center™ [Helsel 2002] developed by Silicon Graphics. These displays are designed to occupy a room and provide a collaborative environment for large groups of people. The screen of the display may be straight, but is often curved. The aim is to encompass as much as possible of the user's field of view. The display is produced using multiple projectors that increase the resolution of the image. Shutter glasses can be used to provide a 3D display although a popular alternative is stereo display based on polarisation. This passive stereo technique has the advantage that users can wear low cost polarised glasses to experience the 3D effects. One drawback of curved screen displays is that a quality auditory display is difficult. This is because sound tends to be focused into a single sweet spot in the floor space. The i-CONE™ [Barco 2001] is one of the most recent evolutions of the large curved screen (figure 1-10). It was developed at GMD and has been designed to prevent the focusing problem with the auditory display.

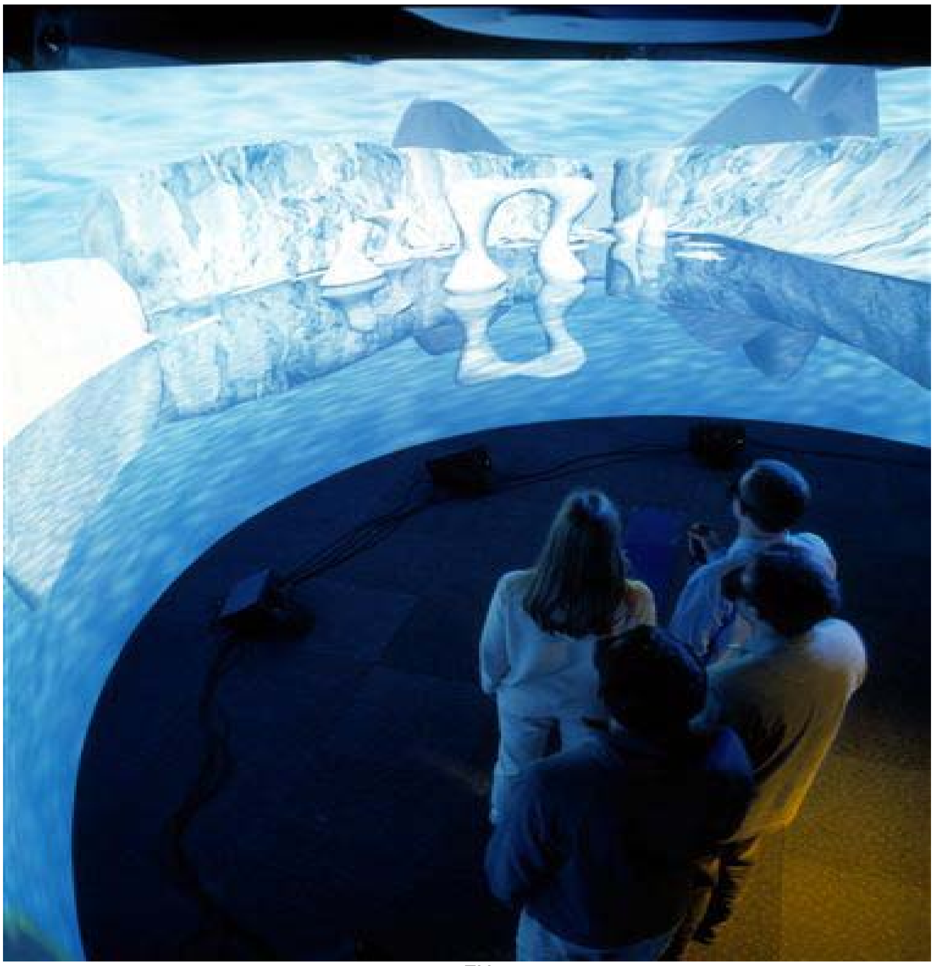


Figure 1-10 The i-CONE^{TM} [Barco 2001] at GMD.
Courtesy of: Fraunhofer Institute for Media Communication, Germany.

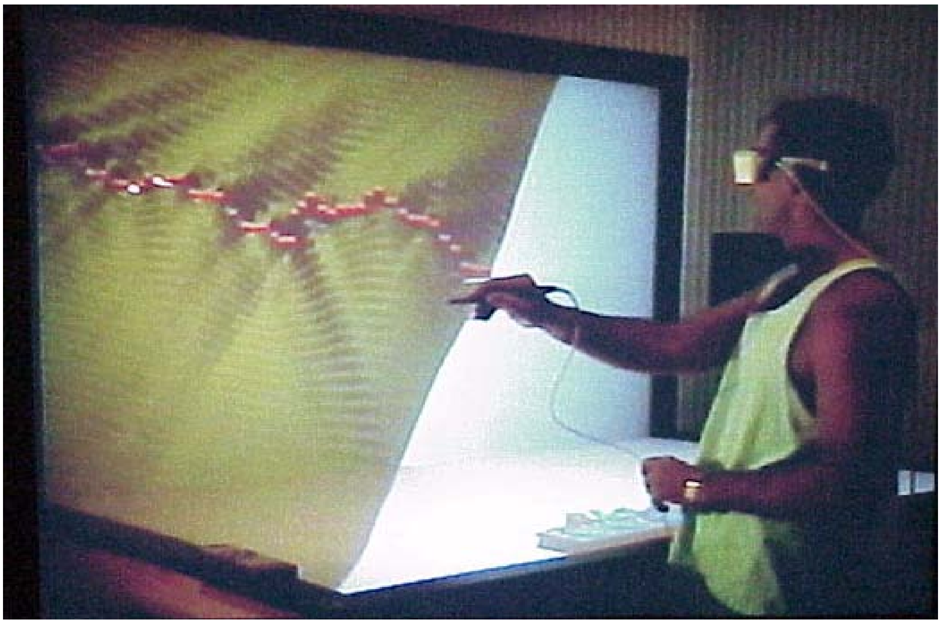


Figure 1-11 The L-Shaped Responsive Workbench [Krüger, Bohn et al. 1995] at GMD.
Courtesy of: Fraunhofer Institute for Media Communication, Germany.

Developed at GMD by Wolgang Krüger, the `Responsive Workbench` mimics a normal desk configuration [Krüger, Bohn et al. 1995]. The `L-Shaped Responsive Workbench` projects the image onto both a top surface and a rear surface (figure 1-

11). Using shutter glasses it creates the illusion of a 3D model projected just above the bench top. Speakers arranged at the top corners of the bench allow auditory display. The `Responsive Workbench` has been configured to support dual users [Hanrahan and Fröhlich 1997] while still providing the natural interaction paradigm where the user manipulates a virtual model on a bench top. Using a combined auditory and visual display it has been applied to such activities as examining well logs for petroleum exploration [Fröhlich, Barrass et al. 1999].

The `Barco Baron` [Barco 2001] is a commercial version of a bench top display (figure 1-12). The display surface that can be tilted at any angle from horizontal to vertical. This platform also uses shutter glasses to create a 3D effect for the user and combines this with head tracking to maintain the user's viewpoint.

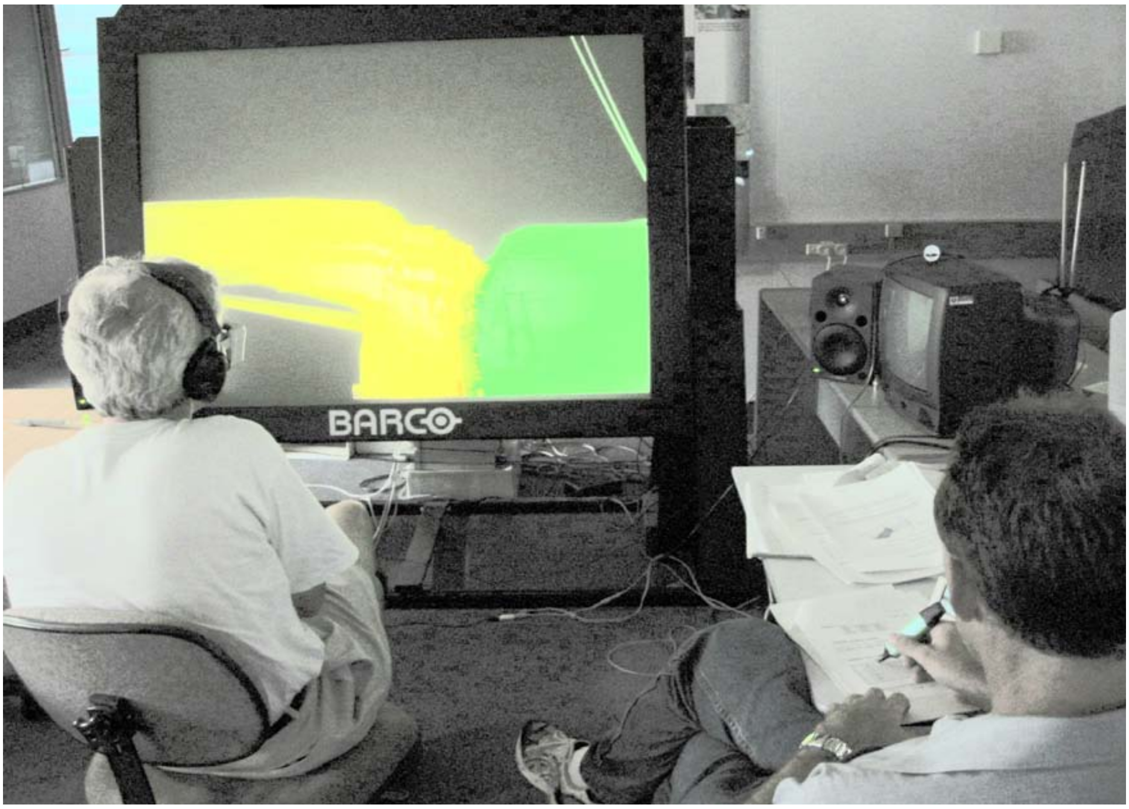


Figure 1-12 The Barco Baron [Barco 2001] at CSIRO in Canberra.
Courtesy of: CSIRO, Mathematical and Information Science, Canberra.

The `Haptic Workbench` was developed at the Department of Mathematical and Information Sciences at CSIRO, in Canberra (figure 1-13). The `Haptic Workbench` is a specially designed single user work environment that allows haptic interaction in a 3D world [Stevenson, Smith et al. 1999]. A 3D image is generated using shutter glasses. The image from a suspended monitor is reflected in a mirror. This creates the illusion of a virtual work space beneath the mirror that is coincident with the user's hands. Primary interaction is by the `Phantom™` [Salsibury and Srinivasan 1997]. The `Phantom™` provides the user with force feedback at a point in space. Information about the orientation and position of the `Phantom™` can be used to help control the display. Sound can also be provided by way of stereo speakers positioned within the bench [Nesbitt, Gallimore et al. 2000]. This environment has been used to develop multi-sensory applications within the petroleum industry to assist the interpretation of seismic exploration data [Nesbitt, Orenstein et al. 1997].

Figure 1-13 The Haptic Workbench [Stevenson, Smith et al. 1999] at CSIRO in Canberra.
Courtesy of: CSIRO, Mathematical and Information Science, Canberra.

1.2.3 Virtual Real Worlds, Virtual Abstract Worlds, Virtual Hybrid Worlds

Section 1.2.2 describes the different types of Virtual Environments used in the case study from this thesis. The accent in this section was on describing the hardware and the interaction styles. This section now focuses on the models displayed within these environments.

Users within a Virtual Environment interact with a computer-generated world or model. There are three main types of worlds the user may work with, *Virtual Real Worlds*, *Virtual Abstract Worlds* and *Virtual Hybrid Worlds*. *Virtual Real Worlds* are models of the real physical world. *Virtual Abstract Worlds* are models of abstract data that has no representation in the real world. *Virtual Hybrid Worlds* overlay abstract data into structures from the real world. Hence these worlds are partly real and partly abstract. The differences between the three types of worlds are described in this section.

Many applications of Virtual Environments focus on mimicking the real world in the display. These displays can be described as *Virtual Real Worlds*. Architectural walkthroughs are one of the earliest applications of this technology [Rheingold 1991]. A building can be rendered in the Virtual Environment for users to examine. Virtual prototyping of many real world goods such as machine parts or automobiles have also been tried in these environments. Caterpillar Inc., the world's largest manufacturer of earth-moving and construction equipment, used Virtual Environments to prototype designs for cabin visibility [Adams 1993]. Major American automobile manufacturers examined using the technology for design and manufacturing [Adams 1993]. Boeing used Virtual Environments for evaluating their aircraft design and, in conjunction with NASA, reviewed critical parts of the design for the space station "Freedom" [Kaufman 1993]. NASA also prototyped the required fixes to the Hubble telescope using a Virtual Environment simulation [Hancock 1993].

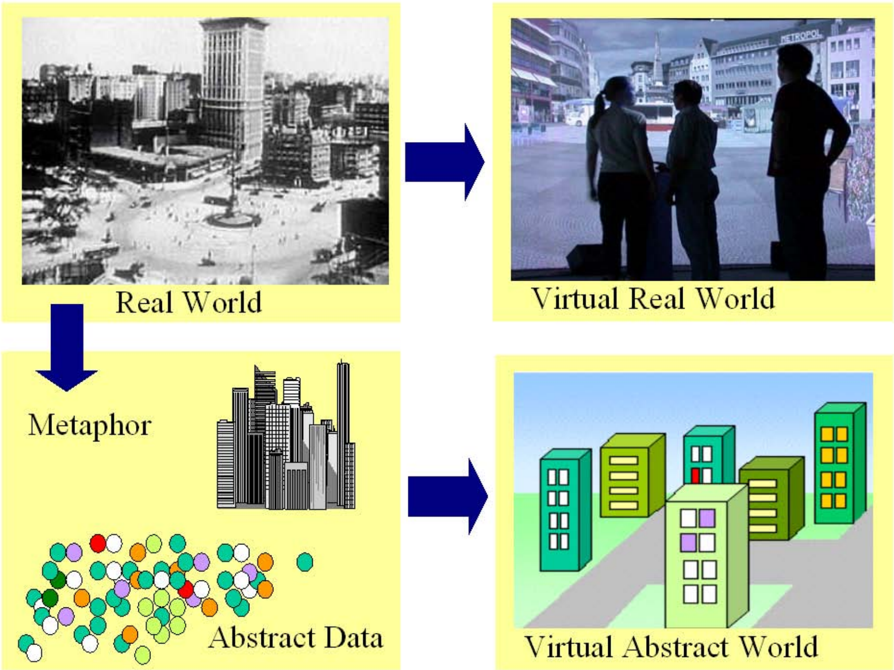


Figure 1-14 Virtual Real Worlds directly model the real world. Virtual Abstract Worlds use a metaphor derived from the real world to build models of abstract data.

Aircraft simulators are a further example of where a Virtual Real World is modelled. In this case the aim is to assist pilots learn to fly. Similarly, many medical applications use this technology to present a virtual body, either for educational purposes or to assist in surgical procedures. The major concern with all these types of applications is to create the most realistic model of the domain that can be rendered in real-time. In practice, limitations in hardware rendering speed often prevent models being created that approach sufficient realism. In fact, there can often be a compromise between producing realism and providing a computationally efficient model that the user can interact with. Despite these technology limits, the goal with Virtual Real Worlds is generally to produce models that are as much like the real world as possible.

Virtual Abstract Worlds on the other hand, focus on producing models of abstract data in a Virtual Environment. Though the idea that Virtual Environments could be used for this role is not new, there have been relatively few applications for this purpose. The cognitive advantages of immersion within Virtual Abstract Worlds have been suggested [Rheingold 1991], but little quantitative supporting evidence is available. Some applications have been developed. For example, Virtual Environments were used for analysing software [Weelner, Mackay et al. 1993]. Another program for exploring and manipulating 3D fractals demonstrates how abstract mathematical entities can be represented as a Virtual Abstract World [DeFanti, Sandin et al. 1993]. The `n-Vision` program allows the user to manipulate abstract 3D models of multivariate relationships

[Beshers and Feiner 1993]. The focus of this thesis is on designing Virtual Abstract Worlds.

A difficult task when designing Virtual Abstract Worlds is to decide on an appropriate model for the abstract data. How should the data be represented in the Virtual Environment? Are there metaphors for displaying abstract information so that the user can better understand the data? This thesis focuses on providing a more structured approach to designing multi-sensory Virtual Abstract Worlds. This structured, engineering-like approach to designing abstract information displays, first develops a structure that categorises the multi-sensory design space. This structure is then used to organise guidelines and to develop a process for designing Virtual Abstract Worlds.

However, the starting point in the search for a better Virtual Abstract World begins with another type of virtual world. Two types of virtual worlds, a Virtual Real World and a Virtual Abstract World have been described. Yet this is a simplification, for there are other types of models that lie somewhere between these two. For example, a number of applications use the framework of the real world and then overlay it with abstract data (figure 1-15). These worlds can be described as Virtual Hybrid Worlds.

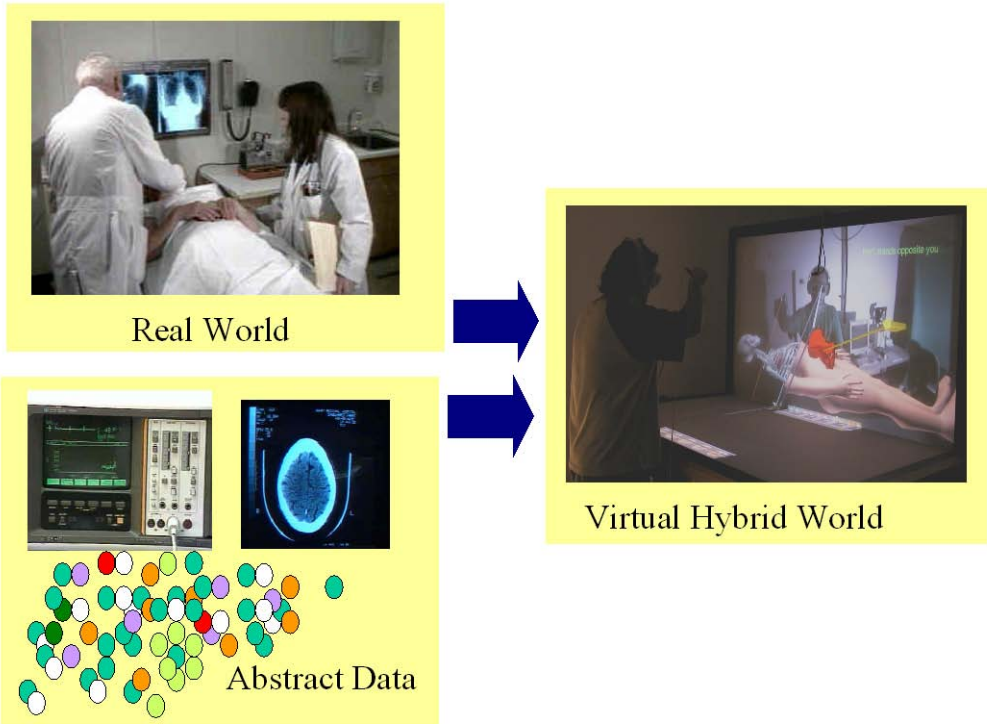


Figure 1-15 Virtual Hybrid Worlds directly overlay abstract data on a model of the real world.

Virtual Hybrid Worlds use a real world structure and overlay this structure with abstract data. For example, a 3D model of the human body can be overlaid with data obtained from MRI or CAT scans [Eben 1995, Satava 1995]. A similar example comes from the petroleum industry where a geological model of the earth was overlaid with abstract data obtained from seismic surveys and well logging [Nesbitt, Gallimore et al. 2000]. Another early example of this approach is the `Virtual Wind Tunnel` created at the NASA Ames Research Center. In this application the user is immersed in a model of a space shuttle that is surrounded by a 3D flow field. The user can investigate the flow by interactively inserting streamlines into the vector field [Bryson and Levit 1992].

What makes Virtual Hybrid Worlds interesting is that they provide users with an easy to understand structural metaphor. For example, physicians understand the structure of the human body and geophysicists understand geological structure. Furthermore it is natural to think of these structures in 3D. They seem to provide intuitive real-world metaphors that are useful for displaying abstract data.

Unfortunately, some domains, such as the stock market, do not have such ready-made metaphors for the display of abstract data. What does a 3D display of abstract data look like? What does the stock market look like in 3D? The focus of this thesis is to find a method for designing better Virtual Abstract Worlds. The design of structural metaphors for abstract data is one area this thesis considers.

The last few sections have introduced the technology of Virtual Environments. Any discussion on this technology can tend to emphasise the 3D visual display of Virtual Abstract Worlds. However, this thesis is concerned with more than visual displays, it is concerned with the display of abstract information to all the senses. The discussion in section 1.2.4 emphasises the multi-sensory aspects of Virtual Environments.

1.2.4 Multi-sensory Feedback

There have been a number of descriptions of Virtual Environments that define their properties [Durlach and Mavor 1995, Stuart 1996]. These properties include immersion in a three-dimensional, synthesized world and natural, intuitive interaction within that world. There is a further key property of Virtual Environments and it is this property that this thesis is most concerned with: that user interaction can be multi-sensory [Durlach and Mavor 1995].

The term *multi-sensory* implies *"more than one sensory modality is used to display the environment"* [Stuart 1996]. Another frequently used term is *multi-modal*. The term *multi-modal* implies that *"interaction can consist of a number of input techniques and provide feedback to several of the user's senses"* [Wickens and Baker 1995]. The emphasis with *multi-sensory* display is on the computer generating output to the user. The emphasis with *multi-modal* interaction is usually on the user generating input to the computer. For example, much work with *multi-modal* applications concentrates on speech and gesture recognition as means for capturing input commands.

There are a number of different types of input and output that can be used in the human-computer interface (table 1-1). The work in this thesis is primarily concerned with the output of information to the user, therefore, the preferred term is *multi-sensory* rather than *multi-modal*.

Visual feedback within Virtual Environments is typically designed to provide the user with a stereoscopic display. This requires displaying a different image to each of the two eyes. Stereoscopy can be achieved by interleaving the display of left and right eye images and using active shutter-glass technology or passive polarisation techniques to ensure each eye sees only the intended image [Burdea and Coiffet 1994]. To create a more immersive display the position and orientation of the user's head is tracked and used to update the image so it is seen from the correct perspective of the user's viewpoint [Burdea and Coiffet 1994].

OUTPUT	MODE	INPUT
Colour Lightning Transparency Shape Text Position	 visual	Eye-gaze direction Head position Head orientation
Force Feedback Vibration Texture Tactile Display Body Motion Temperature Pain	 haptic	Keyboard Mouse (2D, 3D) Props Phantom Body Position Gestures
Speech Pitch Timbre Loudness Rhythm Direction	 auditory	Sound Speech

Table 1-1 Some different modalities of input and output used in Virtual Environments.

Auditory feedback may be provided by the user wearing headphones or by placing multiple speakers within the environment. This feedback can be mono, stereo or designed as a multi-channel, three-dimensional display. One way to provide a three-dimensional sound display is to use a spatially-distributed array of speakers to generate what is called a *sound field simulation* [Evans, Tew et al. 1997]. The alternative approach is called *perceptual synthesis*. The synthesised sound can be displayed on headphones or over a pair of loudspeakers. However, this approach requires an appropriate model of the user's head and ear shape. This model is called the *Head-Related Transfer Function* or HRTF. These complex filters incorporate the human perceptual cues for sound localisation into a source signal [Evans, Tew et al. 1997]. These perceptual cues include interaural intensity differences, interaural time differences and spectral shaping by the pinna [Wenzel 1994]. Using an appropriate HRTF it is also possible to generate a sound that conveys information about the position and distance of the sound source. However, these HRTF models are complex, user-specific and can be unreliable for simulating sounds with a source that is above or below the user [Wenzel, Arruda et al. 1993].

Apart from spatial information, auditory feedback can use sound parameters such as pitch, duration, timbre and loudness to convey information to the user [Stuart 1996]. All of these sound parameters can be controlled in the sound generation process.

Haptic feedback allows a person to interact manually with a Virtual Environment through the sense of touch. Haptic interfaces support the feeling and manipulation of objects and can provide a sense of immersion not possible with visual and auditory feedback alone [Durlach and Mavor 1995]. In the real world, the haptic sense is typically used for exploration and handling tasks. Exploration tasks involve the extraction of object properties such as shape, mass and texture and also provide a sense of contact, position and motion. Handling tasks are dominated by user motor actions such as grasping and object manipulation. For the user, haptic actions require a synergy of sensory exploration and motor manipulation [Srinivasan and Basdogan 1997].

Direct contact and displacement of the skin with an object provides tactile information. Tactile information is what we commonly describe as touch. However, the human haptic system senses both tactile and kinaesthetic information when touching an object [Durlach and Mavor 1995]. Kinaesthetic information provides the sense of position and motion of our limbs and joints. Current tactile displays are inadequate for use in real applications; however, it is possible to integrate force feedback displays into current Virtual Environment systems [Srinivasan and Basdogan 1997]. For example, many platforms use the commercially available Phantom™ force feedback device [Salsibury and Srinivasan 1997]. These displays can mimic a range of haptic sensations that the user senses through a combination of tactile and kinaesthetic receptors.

1.2.5 Increasing the Bandwidth

By enabling multi-sensory feedback, Virtual Environments seek to widen the bandwidth between human and computer. With multi-sensory interfaces the user can potentially perceive and assimilate multi-attributed information more effectively. By mapping different attributes of the data to different senses, such as the visual, auditory and haptic sense, it may be possible to better understand large data sets.

At the start of this chapter, these types of multi-sensory displays were called Human Perceptual Tools when they were designed to find patterns in the data. When designing Human Perceptual Tools the challenge is to choose mappings from the abstract data attributes to the different senses that are intuitive. This thesis develops a design process for building such multi-sensory displays of abstract data. One goal of this design process is to build multi-sensory Virtual Abstract Worlds that increase the human-computer bandwidth.

1.3 The Goal - Multi-sensory Display of Abstract Data

The first section of this chapter describes the motivation for this work, namely, to produce Human Perceptual Tools for finding patterns in abstract data. The next section outlined the technology of Virtual Environments, reviewing their origins, providing some examples of current platforms and focusing on the ability of these environments to enable a multi-sensory display. However, the primary goal of this work is to aid in the design of better Virtual Abstract Worlds for exploring large data sets. This section overviews the various fields that investigate the design of abstract information displays.

The design and display of abstract data is a relatively recent event. For example, it was not until about 1750 that statistical graphics represented data by means of a `time-series chart` and `scatterplot` [Tufte 1983]. The early work by William Playfair in this area helped develop many fundamental visual designs for abstract data.

With the advent of computer technology it became possible to rapidly process larger data sets and produce more complex visual displays of the data.

Despite this relatively recent history, the display of abstract information has given rise to a number of different scientific fields shown in figure 1-16. These include the domains of:

- Information Visualisation (section 1.3.1)
- Scientific Visualisation (section 1.3.2)
- Information Sonification (section 1.3.3)
- Information Haptisation (section 1.3.4)
- Information Perceptualisation (section 1.3.5).

Information Visualisation concerns itself with the visual display of abstract data. *Scientific Visualisation* focuses on the visual display of scientific data. *Information Sonfication* studies the auditory display of abstract data. *Information Haptisation* looks at ways to display abstract data to the sense of touch. *Information Perceptualisation* is used to describe the multi-sensory display of abstract data [Card, Mackinlay et al. 1999].

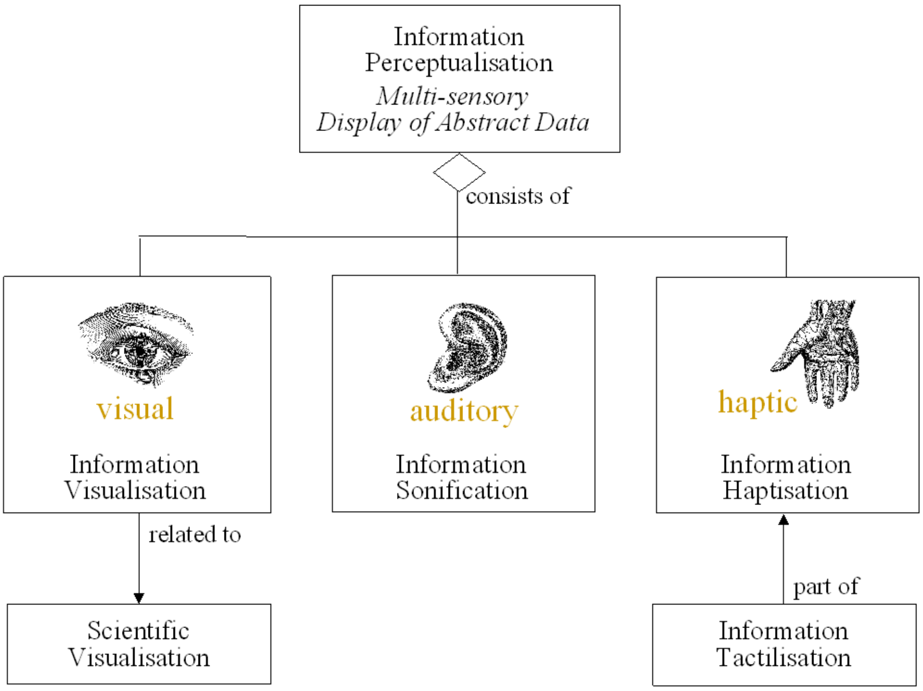


Figure 1-16 The components that make up the field of Information Perceptualisation

As of 2002, all these scientific fields remain fairly immature. Information Visualisation and Scientific Visualisation could be described as emerging fields. Information Sonification is still an undeveloped field of study. Information Haptisation and Information Perceptualisation have been little more than suggested as terms and neither has been investigated in any depth. Although the focus of this work is the multi-sensory display of abstract data, each of the domains shown in Figure 1-17 has the common goal of displaying abstract data and therefore needs to be considered when designing Virtual Abstract Worlds. The next section elaborates a little more on the definition of each of these terms and describes some examples of work from each area.

1.3.1 Information Visualisation

Information Visualisation is the term commonly used to describe interactive computer systems that provide the user with external visual models of abstract data [Card, Mackinlay et al. 1999]. The abstract data is taken from such domains as the stock market, network traffic, software engineering or marketing. These abstract domains often have no real-world model or analog for structuring the information in a way that is intuitive to the user.

Many approaches have been developed in the field of Information Visualisation[2] to allow exploration of large, multivariate data spaces. For example, the `VisDB` system responds to a database query by mapping each database item that satisfies that query to a single pixel [Keim and Kriegel 1994]. The pixels are arranged in space and coloured according to their significance to the query.

Another approach to visualising abstract data is `Parallel Coordinates` [Inselberg 1997]. This technique can be used to explore for relationships between multiple data attributes. The relationships between attributes may be found as 2D visual patterns. For example, one application of this technique allowed the user to look for relationships between 16 process parameters used in the manufacture of VLSI chips.

`Graph Drawings` are a common technique used in many application areas. Large networks of data, such as, electronic mail traffic, have been visualised this way [Eick and Wills 1993]. . For example, the lines drawn between nodes in the drawing can illustrate who in the mail network are communicating. Another visual attribute such as the colour of the link can illustrate the frequency of communication

`Graph Drawings` can be used to produce both 2D and 3D visual models of abstract data. The application called, `SemNet`, implements a 3D display of semantic networks [Fairchild, Poltrock et al. 1988]. Other visualisation approaches have also concentrated on providing 3D visual models. For example, a number of 3D animations of the security market were developed to help evaluate risk [Wright 1995]. In a further example, 3D nested co-ordinate systems were used for exploration of n-dimensional data and applied to analysing financial options [Feiner 1990].

1.3.2 Scientific Visualisation

Scientific Visualisation is a field closely related to Information Visualisation. Unlike Information Visualisation, applications of Scientific Visualisation usually have a physical-world structure or geometry about which they can display the data. These structures are application dependent, but include typical models from the physical sciences such as geological, anatomical and mechanical structures. Overlaid on these structures is the data that may derive from physical simulations, geophysical surveys, weather modelling or collected medical diagnostics [DeFanti, Brown et al. 1989][3].

---

[2] A detailed review of applications within the field of Information Visualisation is provided in chapter 3 and a list of applications and references can be found in Appendix A.

[3] As has been previously noted, the combination of a real world structure overlaid with abstract data provides the user with an easy-to-understand 3D spatial metaphor. This type of model was is also described a Virtual Hybrid World.

While Scientific Visualisations have been tried in Virtual Environments, the focus is not generally on the multi-sensory aspects of these environments [Cruz-Neira, Leigh et al. 1993]. Usually these applications concentrate on the visual mode and in particular use the stereographic capabilities of these environments to present the model in three dimensions. There are some exceptions. Sound, for example, was used to assist car designers to explore the position of air-conditioning ducts [Eckel 1998]. Immersed in a 3D model of the car the designers could visually examine the flow of air using the common technique of flow streams. At the same time, by positioning a tool anywhere in the model the magnitude of air flow could be heard as a wind-like noise. This application is an example of both Information Visualisation and Information Sonification.

1.3.3 Information Sonification

Information Sonification is a newly evolving field that uses sound rather than vision to represent abstract data [Kramer 1994a]. The term, Information Sonification, implies a mapping from the data attributes to the sound parameters. When there is no such mapping the term *Audification* is used. *Audification* describes the direct playing of data as sound [Kramer 1994a]. A good example of Audification is the playing back of seismic events recorded from an earthquake [Hayward 1994].

In some sample applications of Information Sonification, sound has been used to assist in debugging software [Jameson 1994], to display scatter plots [Madhyastha and Reed 1994], to help understand parallel program performance [Jackson and Francioni 1994] and to display computational fluid dynamics data [McCabe and Rangwalla 1994].

1.3.4 Information Haptisation

The term Information Haptisation is used when the sense of touch is used to display abstract data. The word *haptic* derives from the Greek and means to grasp. The sense of touch differs from vision and hearing in that it relies on action from the user to generate the stimuli. This is also described as *bimodal*. For example, a person must tap against a surface to feel its hardness or move their hand across a surface to feel the surface texture.

The term *Information Tactilization* has also been suggested [Card, Mackinlay et al. 1999]. The term, *tactile* refers to the sensation experienced on contact with a surface. However, the word, *haptic*, refers to both the *tactile* and *kinaesthetic* components of touch. The *kinaesthetic* receptors provide a sense of position in space. Since most interactions involving the sense of touch rely on a combination of both tactile and kinaesthetic feedback the term Information Haptisation is preferred in this thesis.

Although Information Haptisation is a very new domain, some interesting applications have been developed. For example, haptics has been used to display soil properties such as density, cohesion and angle of internal friction by allowing the user to move a simulated plough blade through various sandy soils [Green and Salsibury 1998]. Force feedback was used to display a small set of properties such as static friction and surface deviations [Green 1997]. This allowed the user to feel surface textures on simulated surfaces. For example, in this way, different grades of sandpaper can be simulated.

1.3.5 Information Perceptualisation

While Information Perceptualisation is still a fledgling field of research a number of multi-sensory applications have previously been developed. These applications typically enhance existing visual displays by adding sound or haptics, rather than considering the overall design space. This current approach, of enhancing a visual display by the addition of sound or haptics, may not create the best displays. It is the contention of this thesis that considering the entire multi-sensory design space provides a better approach for building good multi-sensory displays.

One example of Information Perceptualisation is seen in the control interface for a scanning probe microscope [Seeger, Chen et al. 1997]. In this application the user can feel the height and friction of the surface. As well as receiving this haptic information, the user also receives visual data from the surface height and colour. In another application, force was used to assist seismic interpreters look for patterns in geophysical data [McLaughlin and Orenstein 1997] (figure 1-17). In this system, force feedback helped the user feel subtle features in the seismic data. These features were often occluded in the visual model.

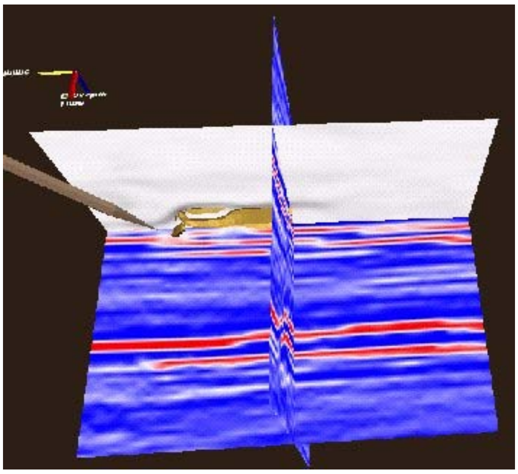


Figure 1-17 A haptic-visual display to assist in the interpretation of seismic data [Nesbitt, Orenstein et al. 1997].
Courtesy of: BHP Research, Newcastle, Australia

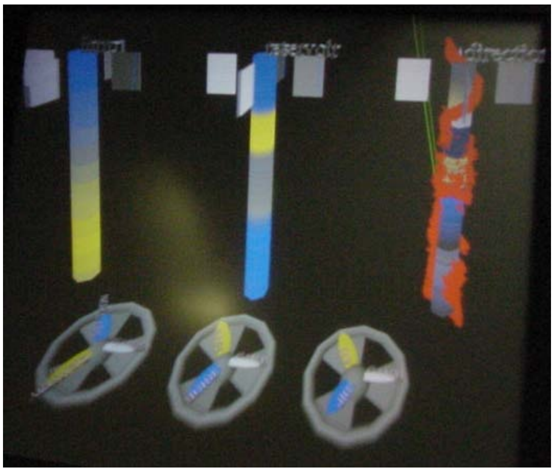


Figure 1-18 An auditory-visual display to assist in the interpretation of well data [Barass and Zehner 2000].
Courtesy of: Fraunhofer Institute for Media Communication, Germany.

Force feedback has also been used as an adjunct to the normal visual user interface. Forces were added to the X-Window System to create a true "look and feel" interface. While the user could see the normal windows they could also feel ridges around icons and menu items. These ridges were designed to provide alignment guides for windows. Haptic button clicks were another feature of this application, so clicking a button on the screen felt like pushing a real world button [Miller 1998].

The previous examples describe where haptic displays were used with visual displays. Sound displays have also been combined with visual displays. For example, auditory signals based on a geiger-counter metaphor were used to display attributes of data collected from a petroleum well [Barass and Zehner 2000] (figure 1-18). The user could probe attributes of the well data with a sound tool. While listening to the well data the user viewed a visual model of the petroleum well. This visual model was colour-coded according to other data attributes. Sound has also been used to display physiological parameters such as respiratory rate, body temperature and heart rate in conjunction with a visual readout of the same data [Fitch and Kramer 1994].

Perhaps one reason why Information Perceptualisation has been slow to develop is that there are numerous problems in designing good displays that use more than one sense. The next section introduces some of the problems with the multi-sensory display of abstract data.

1.4 Some Problems – Designing Multi-sensory Displays

Designing good *visual* displays of abstract data is difficult. Tufte, in his study of statistical graphics [Tufte 1983], suggests some principles to improve the design of 2D statistical visual displays. However, he also comments on the rarity of good design:

> "*On rare occasions graphical architecture combines with the data content to yield uniquely spectacular graphics. Such performances can be described and admired but there are no principles on how to create that one wonderful graphic in a million*" [Tufte 1983].

The immense difficulty of designing visual displays is well known. Tufte gives numerous examples of poor design for statistical graphics. If these were not enough it is also simple to find a number of 2D and 3D information visualisations that suffer from poor design. The evidence for this is fairly easy to gather. (For example, by browsing the daily graphics on the bottom left front page of *USA Today*).

Many visualisations, despite their complexity, ingenuity and aesthetic qualities, fail to support the user in the task they were designed for. For example, in one set of conference proceedings on software visualisations, of the four new ways to visualise software [Churcher, Keown et al. 1999], [Ali and Nishinaka 1999], [Hill, Potter et al. 1999], [Nishikawa 1999] none have been adopted by software engineers in the real world. This is not meant as a criticism of the individual work, but rather as an indication of the difficulty of good design. There are many considerations, *"not only of efficiency, but also of complexity, structure, density, and even beauty"* [Tufte 1983].

Designing good *visual* displays of abstract data is difficult. However, problems have also been encountered when trying to design good auditory displays of abstract data. In the Sonification Report prepared for the National Science Foundation by members of the International Community for Auditory Display in 1997 [Kramer, Walker et al. 1997] the point is made: "*By now it is clear that sonification works and can be very useful. What is not clear is how to go about designing a successful application".*

Designing good *visual* displays of abstract data is difficult. Designing good *auditory* displays of abstract data is also difficult. It is difficult to make such general statements about designing *haptic* displays because so few have been built. However, a particular example does illustrate some problems. A haptic-visual application was developed to help users analyse 3D fluid flow fields and the relationships of those flows to the temperature in a blast furnace [Nesbitt, Gallimore et al. 2001] (figure 1-19). The user was presented with a 3D visual model of temperature in the blast furnace and could feel the fluid flow field by using a force feedback device. The assumption, during design, was that this seemed a very intuitive and natural metaphor. It would be like putting your hand into a flowing stream of water and feeling the direction of flow. However, even for very simple flow fields, the user had trouble discerning the direction of flow. One conclusion from this work was that designing an intuitive interaction metaphor for haptic displays was more complex than originally assumed.

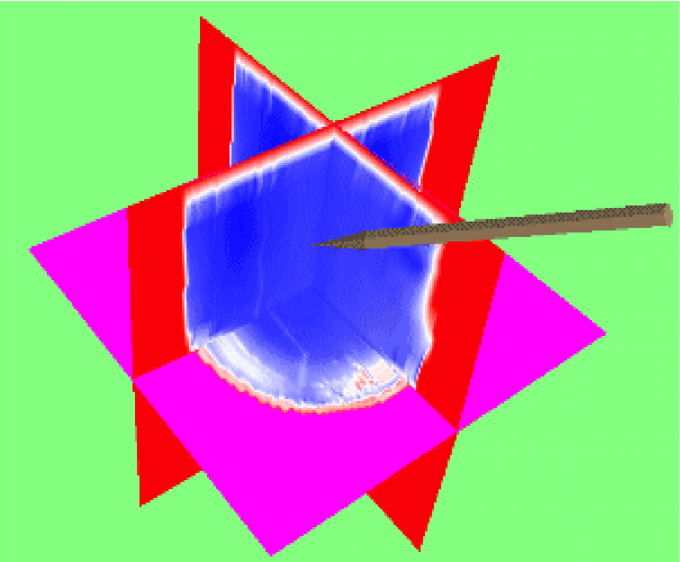


Figure 1-19 A haptic-visual display to analyse fluid flows
[Nesbitt, Orenstein et al. 1997].
Courtesy of: BHP Research, Newcastle, Australia

Designing good *visual* displays of abstract data is difficult. Designing good *auditory* displays of abstract data is difficult. Designing good *haptic* displays of abstract data is difficult. The intuition follows that designing good *multi-sensory* displays is an even more difficult task. This issue is discussed in the next section along with other issues that are the motivating concerns of this thesis:

- How do we design good, multi-sensory displays for abstract data?
- What information do we display to each sense?
- How do we prevent problems with sensory-interaction?
- Can we find useful 3D visual structures for abstract data?

1.4.1 Designing Good Multi-sensory Displays

How do we design good, multi-sensory displays for abstract data? The difficulties of designing displays for a single sense have already been discussed. It is usually fairly simple to gather evidence of this by collecting examples of failed applications. Furthermore, even the few successful applications have encountered numerous problems. These problems are typically solved application by application. Often the design of multi-sensory displays has been something of an ad-hoc process. The design process usually involves the addition of sound or haptics to an existing visual display. For example, a visual display of physiological data was extended to include an auditory display [Fitch and Kramer 1994].

This thesis attempts to develop a more systematic approach to the design of multi-sensory displays. The first step in this thesis is to try and understand the full multi-sensory design space. The design space is large and complex and this thesis tries to provide some structure to the concepts that a designer must consider. The outcome is a new categorisation of the multi-sensory design space called the MS-Taxonomy (chapter 2).

From a human factors perspective, a multi-sensory display must consider the choice of mappings from data attributes to sensory properties. For example, should the visual, auditory or haptic sense be used? Is the visual property of colour or shape more appropriate? Perhaps the auditory property of pitch or timbre would be effective? Is the haptic property of compliance or inertia a good way to display the data? Developing appropriate guidelines about how to use each sense would be useful.

This thesis develops a structured series of guidelines for multi-sensory display called the MS-Guidelines (chapter 6-9). Some new guidelines are developed, but many existing guidelines are also incorporated. These existing guidelines come from a number of fundamental areas, including:

- previous user-interface studies
- human factors research
- perceptual psychology.

Another aspect that needs to be considered when building multi-sensory displays is the limitation of available technology. In particular there are often real-time constraints encountered when using Virtual Environments. For example, to maintain a smooth transition between scenes, a stereo visual display requires images to be refreshed at 60 Hz [Durlach and Mavor 1995]. This places a limit on the complexity of graphics that can be generated. For modelling forces the limitation is even greater as force feedback requires a refresh rate of 1000 Hz [Srinivasan and Basdogan 1997].

One way to ensure that all the relevant engineering factors of design are considered is to follow a process. This thesis also develops the MS-Process for designing multi-sensory displays. The MS-Process considers the limitations of both hardware and software during the implementation phase of the process. In particular the MS-Process considers the following engineering factors:

- the user's requirements (see Section 10.3)
- the nature of the abstract data (see Section 10.4)
- the limitations of development tools (see Section 10.6)
- the specific nature of target environments (see Section 10.6).

1.4.2 Mapping Data to the Different Senses

What information do we display to each sense? Because of the problems with multi-sensory display perhaps the focus should remain only on the visual display of abstract information? After all, some suggest that vision is the dominant sense. While it is true that vision is highly detailed and well suited to comparing objects arranged in space, it is equally true that hearing is effective for monitoring sounds from all directions, even when the source of the sound is not visible. Touch, it has been shown, does equally well as vision at discriminating texture [Morton 1982]. In fact haptic texture cues may be more perceptually prominent than visual texture cues when both sources of information are present [Morton 1982].

"The dominance of vision is wrong" [Welch and Warren 1980]. In fact, the various senses are well suited for different kinds of tasks. However, it is not altogether clear what types of abstract data to display to each sense. To address this issue the multi-sensory display must consider the physiological, perceptual and cognitive capability of each sense.

Understanding the physiology of each sense helps in understanding its performance capabilities and bandwidth. For example, the range of colours that the eye can see or the frequency of sounds that can be heard are limited by the underlying physiology. Perception is dependent on physiology but multiple levels of neural processing also influence it. For example the same wavelength of light can appear to be a different colour depending on the background colour [Itten 1970]. This is a result of the way nerves from the visual receptor cells are organised rather than the actual physiology of the eye's receptors.

The influence of higher neural processes on sensory perception is a general principle and can also be illustrated with the hearing and touch [Sekuler and Blake 1990]. For example, two similar sound frequencies can sound the same and the ability to distinguish them may depend on the musical training of the listener [Kramer, Walker et al. 1997]. When displaying a haptic surface with force feedback, the display can give the impression of objects with a soft surface if the display frequency is low [Srinivasan and Basdogan 1997].

Cognition issues are also important when designing a display to recognise patterns. For example, the haptic sense may not be as useful for remembering complex patterns as the auditory sense. This thesis develops the MS-Guidelines to capture appropriate principles of design for the physiological, perceptual and cognitive capability of each sense. These guidelines are then integrated into the design process and used for both designing and evaluating displays.

If the goal of multi-sensory display is to *widen the human-to-computer bandwidth* then it is important that we strive to display different data attributes to different senses. This type of display has been characterized as a *complementary display* [McGee, Gray et al. 1998, Pao and Lawrence 1998]. *Complementary displays* map different information to each sense. It is expected that task performance with a *complementary display* would be superior to the performance with a single modality display.

McGee, Gray et al. also described two other types of multi-sensory display. They are *conflicting displays* and *redundant displays*[4]. A *conflicting display* maps contradictory information to each sense. With a *conflicting display* performance is worse in the multi-sensory display than with a single modality display. With *redundant displays* the same information is displayed to each sense. In a *redundant display* the performance of the user with the single modality display is the same as with the multi-sensory display. However, users may report a reduction in workload or an increase in confidence with the multi-sensory display.

Using *complementary* display takes advantage of what is known as *Modal Specific Theory* [Friedes 1974]. This psychophysical theory states that each sensory modality

[4] Redundant display has also been described as cooperative rendering [Pao and Lawrence 1998].

has distinct patterns of transduction. So, each sense has unique sensory and perceptual qualities that are adept with certain kinds of complex information. Once again the MS-Guidelines developed in this thesis aim to use the unique capabilities of each sense.

In summary, the solution this thesis provides to the problem of, what data to map to what senses, is a series of guidelines. The MS-Guidelines are both general and specific. The general guidelines assist high level design choices, guiding the overall approach. The more specific guidelines assist in the final choice of sensory mappings.

1.4.3 Preventing Problems with Sensory Interaction

How do we prevent problems with sensory-interaction? Designing *complementary displays* seems a simple enough goal, yet often the senses can interact. For example, using sound in conjunction with haptics can alter the perceived stiffness of a surface [DiFranco, Beauregard et al. 1997]. So when a *hard* sound is played on contact, the surface is reported as being harder than when a *soft* sound is played. This is despite the fact that, in each case, the same haptic model is used to represent the surface contact. Likewise, changing the visual representation of the object can alter the perceived haptic stiffness of a spring. Thick visual representations of a spring feeling stiffer than thinner ones, despite the same force being required to compress the spring [Srinivasan and Basdogan 1997].

This mismatch between two senses is called *sensory discrepancy* [Welch and Warren 1980]. There is a strong tendency for our perceptions to produce an experience that is consistent across our senses. When conflicting information about an event is received, there is a tendency to perceive the situation as a single consistent event rather than two separate events. If a conflict occurs between two or more sensory modalities one or both modalities tend to bias each other. For example, when a stationary hand is viewed through a 14 degree displacing prism, it immediately feels as if it is located very near its seen position. Here the visually displaced view overrides haptic information about the actual physical location of the hand [Welch and Warren 1980].

Preventing multi-sensory interactions is a very complex problem, as it is not even fully understood how the senses interact to perceive information from the real world. What is clear, however, is that we interact with the real world in a multi-sensory way all the time. The approach of this thesis is to attempt to model the abstract data in a way that it resembles our real world interactions. The MS-Taxonomy is based on a categorisation of high-level information metaphors. These metaphors aim to take advantage of the way we perceive the real world. The structure of the MS-Taxonomy is used to organise the MS-Guidelines. These guidelines are integrated into the MS-Process and suggest ways to avoid the problem of *sensory discrepancy.* However, the problem of *sensory discrepancy* is so complex it is not surprising that unforeseen problems can still occur. Hence the MS-Process integrates a number of different types of evaluation. These evaluations can be used to test for unexpected sensory interactions in the display.

1.4.4 Finding Useful 3D Visual Structures

Can we find useful 3D visual structures for abstract data? As we have discussed, designing good multi-sensory displays of abstract data is difficult. However, even before considering a multi-sensory display the issue of using 3D visual models arises. It could be simpler to use only a 2D visual display. This is an important issue in this work as a key facet of Virtual Environments is that they provide the user with a 3D model of the data.

Therefore a key part of the approach described in this thesis is to try and find a useful 3D structure for abstract data. After all, multi-sensory interaction in a 3D world is the way we operate in our real world. The intuition is that 3D visual skills should be useful in an abstract world. Yet 3D visual displays have so far only met with limited success. Perhaps the most appropriate 3D structures for abstract data are yet to be developed.

One domain where a natural 3D metaphor has been successful is the use of for Virtual Environments for oil exploration [Nesbitt, Gallimore et al. 2000]. This domain provides a natural spatial model, namely geological structure for organising the 3D data space (figure 1-20). This 3D spatial model proved to provide a well understood model for experts in the domain. Geophysicists immediately understood the 3D framework provided by rock strata and faults. For this work it was not a question of comparing the effectiveness of a 2D display against the effectiveness of a 3D display. In this application the feedback from domain experts was unequivocal. For understanding the overall relationship of the geological data and communicating concepts to colleagues, the 3D models were clearly superior to a 2D model.

In Section 1.2.3 this type of application was called a Hybrid Virtual World because it mixes a real world structure with abstract data. In this case, the 3D geological structure forms the basis of displaying the seismic data. Using the MS-Process, designing the 3D visual model is a key step. Indeed the first step of the MS-Process seeks to find an appropriate 3D spatial structure for the multi-sensory display. This approach is demonstrated in the case study of the MS-Process (chapter 11-13) where the initial focus is to develop appropriate 3D visual structures of stock market data.

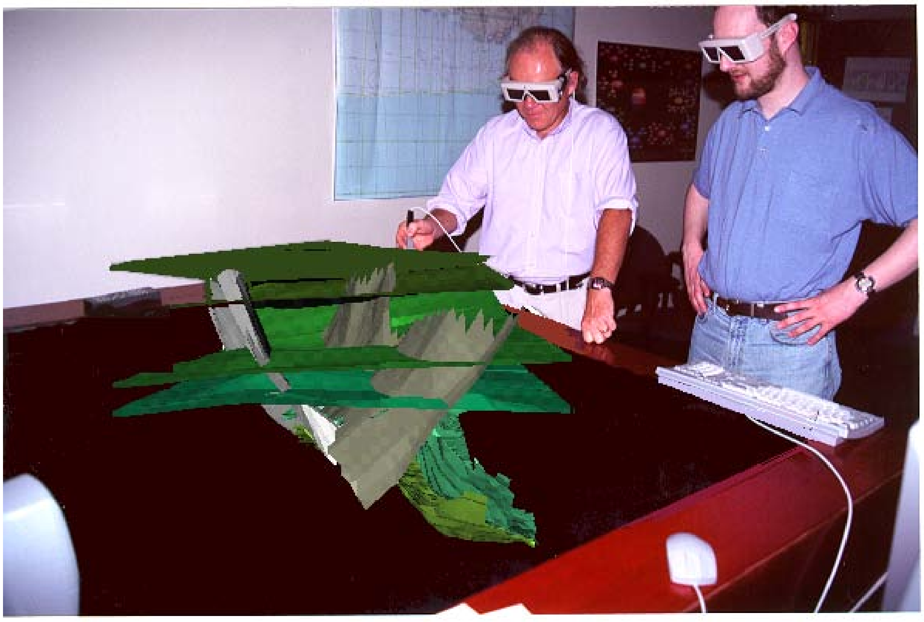


Figure 1-20 Users collaborate on a 3D geological model during a trial of Virtual Environment technology at BHP Petroleum [Harding, C., B. Loftin, et al. 2000].
Courtesy of : BHP Research, Newcastle, Australia

1.5 A Solution - Assisting Design of Multi-Sensory Displays

In summary the vision of this thesis is to use Virtual Environments, particularly the 3D and multi-sensory nature of these environments to *widen the human-to-computer bandwidth* and provide *Human Perceptual Tools* for *data mining* (figure 1-21). These *Human Perceptual Tools* require the design of effective *Virtual Abstract Worlds*.

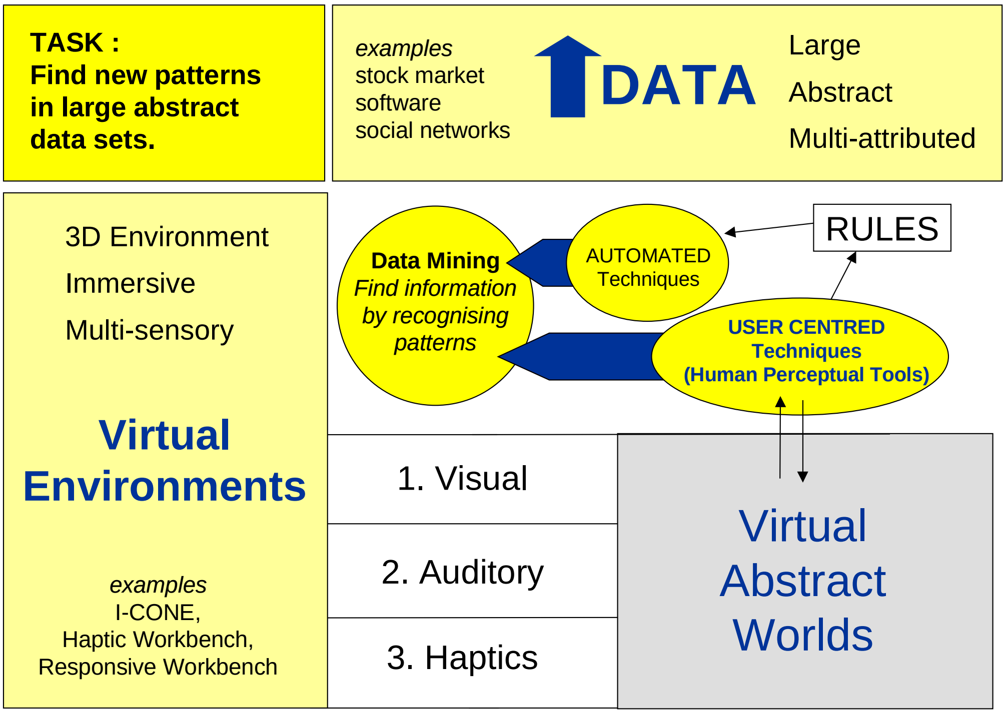


Figure 1-21 The multi-sensory approach to finding new patterns in large abstract data sets.

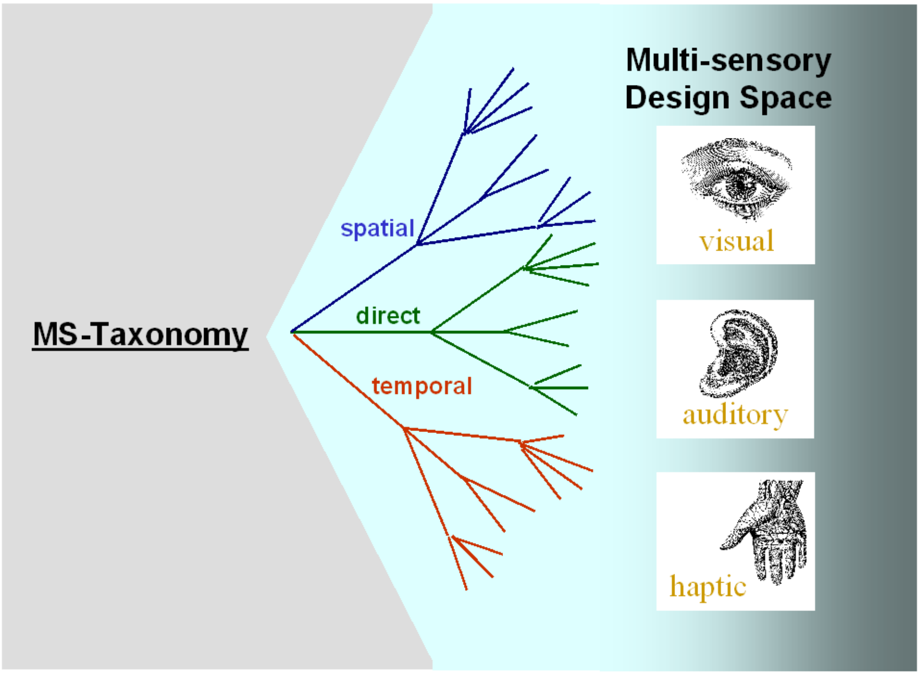


Figure 1-22 The MS-Taxonomy is derived by considering spatial, direct and temporal metaphors for the multi-sensory design space.

The solution adopted by this thesis is to first try and understand the design space of multi-sensory displays. For this purpose a new high-level categorisation of the multi-sensory design space called the MS-Taxonomy is developed (figure 1-22) (chapter 2). The MS-Taxonomy is then used to consider previous work in the domain of multi-sensory display (chapter 3-5). This serves to both review previous literature and explain the concepts that make up the MS-Taxonomy.

This review gives rise to the MS-Guidelines which are a series of design guidelines for multi-sensory display (chapter 6-9). The MS-Guidelines are organised by reusing the structure of the MS-Taxonomy. Finally a new engineering methodology called the MS-Process is developed (chapter 10). The MS-Process uses the MS-Guidelines and also reuses the structure provided by the MS-Taxonomy (figure 1-23).

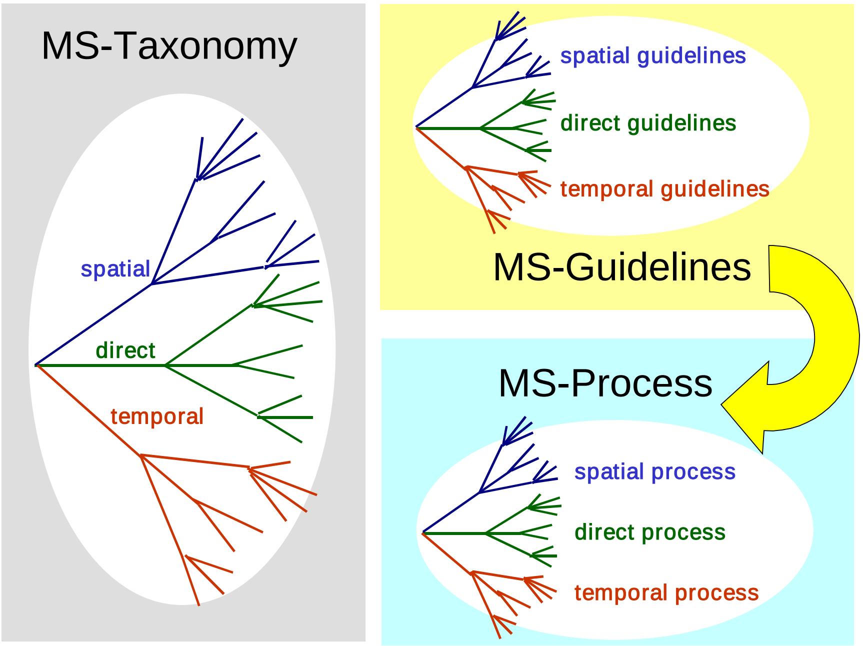


Figure 1-23 The structure of the MS-Taxonomy is used to organise both the MS-Guidelines and the MS-Process. The MS-Guidelines feed into the MS-Process.

1.6 Methodology

The goal of this thesis is to help design better Human Perceptual Tools for finding patterns in abstract data. A more targeted list of questions that the thesis attempts to answer is provided in table 1-2. However a real difficulty such high level goals is how to evaluate the outcomes. The major outcomes from this work are:

- the MS-Taxonomy
- the MS-Guidelines
- the MS-Process

This thesis would first like to evaluate the issue: does the MS-Taxonomy provide a useful, structured description of the multi-sensory design space? The thesis investigates this question by first using the concepts of the MS-Taxonomy to review existing literature (chapter 3-5). This is an unusual approach to performing a literature review. However it is intended as much more than a literature review. It is also intended to illustrate and validate the concepts of the MS-Taxonomy. If the concepts do not make sense in terms of existing applications of information display then it would be hard to argue the validity of the MS-Taxonomy.

#	**Question**
1	Can new technologies assist in the data mining process?
2	How do we build better Human Perceptual Tools for data mining?
3	How do we design appropriate multi-sensory models for data mining?
4	How do we evaluate the effectiveness of a display for finding patterns in data?
5	Can multi-sensory feedback widen the bandwidth between human and computer?
6	How do we design intuitive mappings from the data to the senses?
7	How do we design good, multi-sensory displays for abstract data?
8	What information do we display to each sense?
9	How do we prevent problems with sensory-interaction?
10	Can we find useful 3D visual structures for abstract data?

Table 14-2 A summary of the questions that this thesis tries to address.

A further concern with developing the MS-Taxonomy is whether or not it adequately covers the full multi-sensory design space. To specifically investigate this question an alternative taxonomy of the multi-sensory design space was constructed called the XCM-Taxonomy. The MS-Taxonomy and the XCM-Taxonomy are compared and shown to be equivalent. This provides confidence that both taxonomies cover the full multi-sensory design space. This work is published [Nesbitt 2001b] and a copy of the paper is included in Appendix C.

Apart from detailed concepts, the general structure of the MS-Taxonomy needs to be validated. To evaluate this, the MS-Taxonomy is used to structure the literature review. In particular the different layers of abstraction are highlighted. In fact during the review a number of new displays are suggested by transferring abstractions across the senses.

Using it to develop the MS-Guidelines and the MS-Process further validates the structure of the MS-Taxonomy. The issue at this stage is to also validate the guidelines and the process. Do the MS-Guidelines and the MS-Process assist in the design of a useful multi-sensory display for finding patterns? The approach taken by this thesis is to apply the MS-Process and MS-Guidelines within a case study from a real world domain (chapter 11-13). The domain in this case is the *Technical Analysis* of stock market data. The MS-Process is followed in some detail and many guidelines are used. The case study thus allows the process and guidelines to be assessed in the course of developing a real application.

Within the context of the case study the question arises: do the MS-Guidelines and the MS-Process assist in the design of a useful multi-sensory display for finding patterns in stock market data? By following the MS-Process the case study produces a number of designs. These designs are prototyped and evaluated. This evaluation is both heuristic and experimental[5]. The final success of the overall process can be judged in part by the results of these evaluations.

The approach of a single case study has the disadvantage that the MS-Process and MS-Guidelines are not tested on a wide cross section of abstract data types. For, of course, many other types of application data suggest themselves. The MS-Process could, for

[5] In fact the evaluation of designs is an important step within the MS-Process.

example, be applied to statistical census data, software metrics or social networks data. Clearly more long-term empirical evidence is required to measure the usefulness of the conceptual tools developed within this thesis. Unfortunately, the scope of this thesis topic is too broad to cover more than a single application domain with a suitable level of granularity. In particular the MS-Process requires a proper task analysis of the user's requirements to drive design decisions. It also requires a number of iterations through the steps of design. Furthermore, developing and evaluating any application to a suitable level of complexity within a Virtual Environment is a time consuming task.

However, there is also an advantage of choosing a single application domain. That is, it allows the full multi-sensory design space to be investigated. For example, during the case study a range of both conceptual and detailed designs are explored across the complete multi-sensory design space.

1.7 Contributions

A new categorisation of the multi-sensory design space, called the MS-Taxonomy, is developed. This taxonomy provides the first detailed concept map of the full multi-sensory design space for abstract information display. This model allows the abstract concepts that make up information displays to be both compared and transferred between the different senses.

Using the MS-Taxonomy a set of structured guidelines for multi-sensory design are developed. The MS-Guidelines contains many new guidelines but also incorporate previously proposed guidelines for visual, auditory and haptic display.

To integrate the MS-Guidelines into multi-sensory design, a new process, called the MS-Process is developed. The intent of the MS-Process is to assist in designing multi-sensory displays of abstract data for mining within Virtual Environments.

The application of the MS-Process is validated in the domain of stock market trading. This case study results in a number of new, multi-sensory models being developed for this stock market trading. Included in these models are the first haptic displays of stock market data and a innovative and highly successful auditory-visual display.

Chapter 2

Metaphors and Senses

More than one way of looking at things (1999)

*"I was always puzzled by the fact that people have a great deal of trouble and pain when and if they are forced or feel forced to change a belief or circumstance which they hold dear. I found what I believe is the answer when I read that a Canadian neurosurgeon discovered some truths about the human mind which revealed the intensity of this problem. He conducted some experiments which proved that when a person is forced to change a basic belief or viewpoint, the brain undergoes a series of nervous sensations equivalent to the most agonizing torture." **[Sidney Madwed]***

Chapter 2

Metaphors and Senses

2.1 Introduction

The field of *Information Visualisation* has progressed largely by invention as applications are developed for specific domains. The field is disjointed as it brings together not only disparate application domains but all so many research areas. These research fields include *computer graphics, information theory, perceptual and cognitive science*, s*oftware engineering, human computer interaction and user interface design*. Sometimes it seems that the design of an information display may be better described as an 'art' and not as a 'science'.

It would be desirable for non-specialists to adopt an engineering approach to the design of information displays. An engineering approach would provide a process that could be followed and would integrate design guidelines. This approach could reduce the need to understand complex perceptual theories beforehand. It could also reduce the need to run time-consuming human factors experiments after the display is built. This is particularly critical when attempting to build a multi-sensory display where sensory interactions and conflicts complicate perceptual design issues.

Categorising the multi-sensory design space is an important first step to assist in the development of general principles of design. This is necessary, as any design should consider the full range of possibilities offered by the design space. A typical division of the multi-sensory design space bases categories around the different senses. However, this type of division makes it hard to compare or transfer display concepts between the senses. However, this thesis uses a more novel division of the space by considering both the senses and the different types of metaphors used in information displays.

Section 2.2 introduces the idea of metaphors (section 2.2.1) and the motivation for using metaphors to divide the design space (section 2.2.2). The MS-Taxonomy, a new categorisation of the multi-sensory design space based on a high-level classification of information metaphors is introduced in section 2.2.3. The MS-Taxonomy uses both metaphors and senses to classify the multi-sensory design space. The three senses of vision, hearing and touch are described in section 2.3.

The MS-Taxonomy can also be thought of as a multi-level network of abstractions that typically describe a software system. In this sense the MS-Taxonomy forms a hierarchy or framework that describes the multi-sensory design space. The Universal Modelling Language (UML) notation is a commonly used notation for describing software systems [Booch, Rumbaugh et al. 1999] and is used throughout this thesis to help in the description of the design space.

2.2 Metaphors

2.2.1 What are Metaphors?

A *metaphor* is defined as *"a figure of speech in which a word or phrase is applied to an object or action that it does not literally denote in order to imply a resemblance"* [Wilkes and Krebs 1990]. In the context of data, a metaphor can be considered to be a model or mapping between the abstract application domain and the real world. The intent is to allow the abstract domain to be better understood by referencing a familiar model from the real world (figure 2-1).

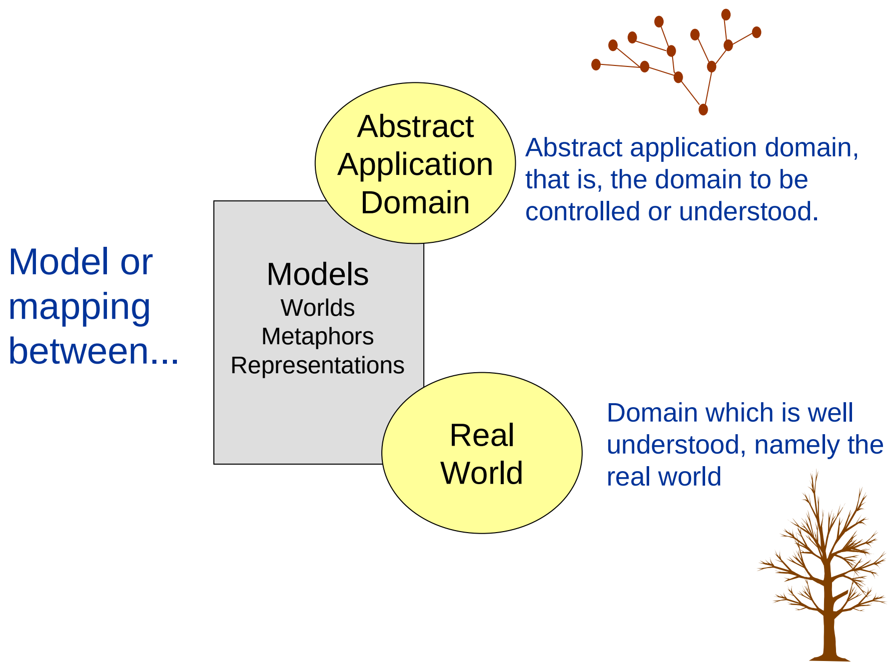


Figure 2-1 A metaphor uses a real world concept to model abstract data.

The role of metaphors in the user interface is defined by Hutchins to occur at three different levels [Hutchins 1989]:

1. <u>Activity metaphors</u>. These metaphors support the user's high level goals. In the case of this thesis the user's high level task is to find new patterns in the data. The usual metaphor for this explorative activity is *mining* and hence this activity is often called data mining.

2. Mode of interaction metaphors. These metaphors refer to the relationship between the user and the computer. The interaction metaphor in this thesis is one of direct object manipulation as supported by all Virtual Environments. There are subtle differences in the interaction paradigm between these environments (see figures 1-2, 1-3, 1-4, 1-5, 1-6). However, they all support a level of natural, real world interaction where objects are directly manipulated. This thesis does not consider interaction metaphors.

3. Task domain metaphors. These metaphors provide a structure for understanding a particular task. They include models, representations, or mappings that describe how the raw data attributes are displayed as sensory artefacts in the multi-sensory display. It is a categorisation of these task domain metaphors that this thesis uses to develop the MS-Taxonomy.

2.2.2 Why use Metaphors in the Design Process?

Why use metaphors at all to categorise the design space? Recall, from section 1.4 that two of the questions that this thesis addresses are:

- How do we design good, multi-sensory displays for abstract data?
- How do we prevent problems with sensory-interaction?

It is these two problems that provide the main motivations for using metaphors.

Firstly, how do we design good, multi-sensory displays? Metaphors can provide cognitive models to help users to browse unfamiliar information spaces. In an exploration mode the user may be seeking to build up a conceptual model of a new information space. This is an iterative process of interaction and exploration that allows the user to form a mental model (figure 2-2). Metaphors provide a useful starting point for the formation of this mental model. It is presumed that *good* metaphors give users a *good* mental model and so enable *good* displays of abstract data.

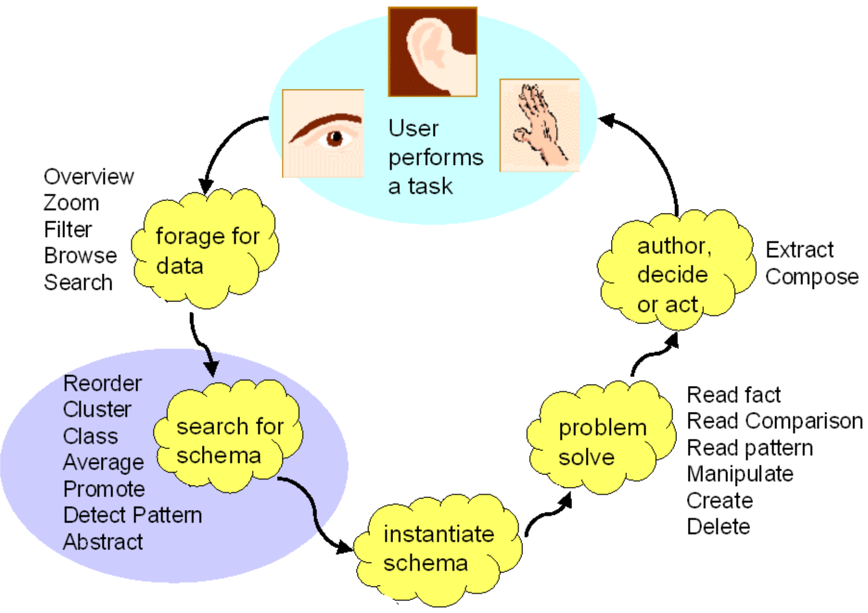


Figure 2-2 A metaphor provides a schema for understanding, ordering and clustering data as shown in this adaptation of the steps of "Knowledge Crystallization" from Card-Mackinlay-Shneiderman [Card, Mackinlay et al. 1999]

Secondly, how do we prevent problems with multi-sensory interaction? The solution to this question is motivated by common sense. Everyday we interact with objects in the real world to do tasks. We interact with all our senses. Our senses have evolved to do this and have been trained to cooperate this way. Our perceptual skills have developed to cope with the demands of this world. Metaphors can provide a mapping from an abstract information space to real world spaces thus making it possible to use our existing perceptual skills in a sensible way and without conflict. Hence we can use our senses to explore the abstract world in the same way we explore the real world. This has been called an ecological approach to design and this approach has, for example, been described for designing sound displays [Gaver 1993].

2.2.3 Deriving the Metaphor-Sensory Taxonomy

The idea of using metaphors to represent abstract information is not new. These metaphors have developed from the 1800s and Tufte describes the evolution of some 2D displays for statistical graphics [Tufte 1983]. By the end of the 1990s, *task domain metaphors* were frequently being used to develop domain-specific displays of information [Spence 2001].

However, the metaphor classes developed by this thesis are more abstract and high-level than this previous work. The first novel idea of this thesis is to further abstract these existing information metaphors. This meta-abstraction, results in three general classes of metaphors called spatial metaphors, direct metaphors and temporal metaphors (figure 2-3). These three general classes of metaphors are applicable to every sense and will be described in more detail during later chapters (3,4,5).

To integrate the human senses, this thesis considers the visual, auditory and haptic artefacts of multi-sensory display[1]. A multi-sensory display consists of a mapping between the artefacts of the visual, auditory and haptic displays and the attributes of the abstract data (table 2-1).

	SENSORY DISPLAY MODES		
	 visual	 auditory	haptic 
DISPLAY ARTEFACTS System Displays OUTPUT to the user as:	Colour, Lighting Transparency Shape (2D, 3D) Patterns, Texture Structure (2D, 3D) Movement Spatial relations Text, Icons	Speech Pitch Timbre Rhythm "Earcons" Melody (music) Direction (2D,3D) Movement	Force Feedback Vibration Texture Tactile Display Body motion Temperature Pain Chemical

Table 2-1 Display artefacts used to output information to the different senses.

[1] Note that while this thesis focuses on only three senses, this taxonomy is thought to be more generic. It could, for example, be simply extended to include other senses such as the olfactory sense. Olfactory displays are currently being developed but the technology is not readily available for use in real applications [Davide, Holmberg et al, 2001].

The name MS-Taxonomy is used because the categorisation is based on types of metaphors (M) and types of sensory (S) displays. Indeed, the MS-Taxonomy is derived by combining the three general types of information abstraction (metaphor) with the different senses (sensory) used for information display.

The sensory classes (visual, auditory and haptic) are combined with the three general metaphor classes (spatial, direct and temporal) as shown in table 2-2. This produces the nine main classes of the MS-Taxonomy (figure 2-3):

- Spatial visual metaphors.
- Spatial auditory metaphors.
- Spatial haptic metaphors.
- Direct visual metaphors.
- Direct auditory metaphors.
- Direct haptic metaphors.
- Temporal visual metaphors.
- Temporal auditory metaphors.
- Temporal haptic metaphors.

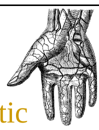
		SENSORY DISPLAY MODES		
		 visual	 auditory	 haptic
METAPHOR CLASSES	**SPATIAL METAPHORS (chapter 3)**	Spatial Visual Metaphors (section 3.2)	Spatial Auditory Metaphors (section 3.3)	Spatial Haptic Metaphors (section 3.4)
	DIRECT METAPHORS (chapter 4)	Direct Visual Metaphors (section 4.2)	Direct Auditory Metaphors (section 4.3)	Direct Haptic Metaphors (section 4.4)
	TEMPORAL METAPHORS (chapter 5)	Temporal Visual Metaphors (section 5.2)	Temporal Auditory Metaphors (section 5.3)	Temporal Haptic Metaphors (section 5.4)

Table 2-2 The nine basic classes of the Metaphor Sensory (MS) Taxonomy.

2.2.3.1 What are Spatial Metaphors?

Spatial metaphors relate to the scale of objects in space, the location of objects in space and the structure of objects in space. Spatial metaphors concern the way pictures, sounds and forces are organised in space and can be described for the visual, auditory and haptic senses.

There are three classes of spatial metaphors:

- *Spatial visual metaphors* concern the way pictures are organised in space.
- *Spatial auditory metaphors* concern the way sounds are organised in space.
- *Spatial haptic metaphors* concern the way forces are organised in space.

Spatial metaphors are discussed in detail in chapter 3. Spatial metaphors involve the perception of a quality (space) that is not associated with any particular sense. Although three different classes of spatial metaphors (visual, auditory and haptic) are described, the concepts that define a spatial metaphor are general and therefore independent of the senses.

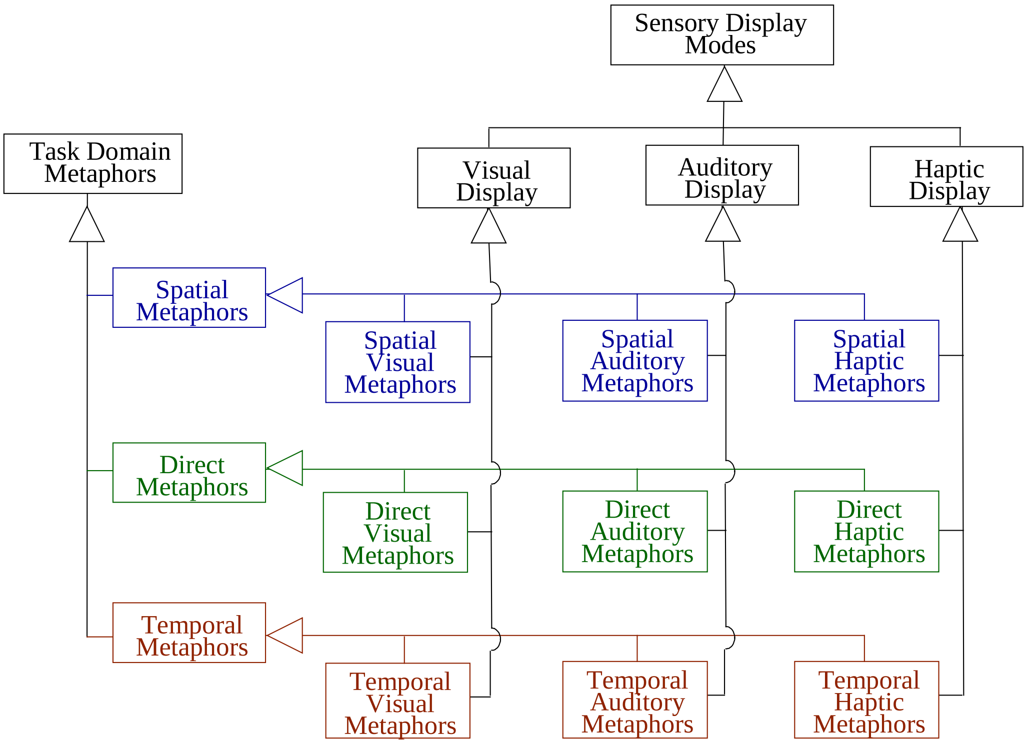


Figure 2-3 A UML diagram showing the MS-Taxonomy. The nine metaphor-sensory classes inherit from three general metaphor classes and the three sensory classes.

2.2.3.2 What are Direct Metaphors?

Direct metaphors are concerned with direct mappings between sensory properties and some abstract information. For example, sensory artefacts such as a specific colour, the volume of sound or the hardness of a surface may be used to represent a particular data attribute. Once again, a class of direct metaphors can be defined for the visual, auditory and haptic senses.

There are three classes of direct metaphors:

- *Direct visual metaphors* concern the perceived properties of pictures.
- *Direct auditory metaphors* concern the perceived properties of sounds.
- *Direct haptic metaphors* concern the perceived properties of touch.

Direct metaphors are discussed in detail in chapter 4. Unlike spatial metaphors, direct metaphors are highly specific for each modality. Each sense is described by it's own capability to perceive sensory properties. For example, the eye perceives colour, the ear perceives pitch and the haptic sense can perceive the hardness of an object. These different sensory capabilities are described as direct properties. The concept of a direct property is general for all senses. Thus, for example, it is possible to compare a direct property of one sense with another.

2.2.3.2 What are Temporal Metaphors?

Temporal metaphors are concerned with how we perceive changes to pictures, sounds and forces over time. The emphasis is on displaying information by using the fluctuations that occur over time. Temporal metaphors can be considered for the visual, auditory and haptic senses.

There are three classes of temporal metaphors:

- *Temporal visual metaphors* concern the way pictures change with time.
- *Temporal auditory metaphors* concern the way sounds change with time.
- *Temporal haptic metaphors* concern the way haptic stimuli change with time.

Temporal metaphors are discussed in detail in chapter 5. Temporal metaphors are like spatial metaphors in that they involve the perception of a quality (time) that is not associated with any particular sense. Though the three different classes of temporal metaphors (visual, auditory and haptic) are described, the concepts that define a temporal metaphor are general and therefore independent of the senses.

Temporal metaphors are closely related to both spatial and direct metaphors. For example it is changes that occur to a particular spatial metaphor or direct metaphor that may display the information.

2.3 Senses

The MS-Taxonomy is derived from considering metaphors and the human senses of vision, hearing and haptics. The basic function and capability of each of these senses needs to be understood when designing information displays. Section 2.3.1 describes the visual sense, section 2.3.2 describes the auditory sense and section 2.3.3 describes the haptic sense. Section 2.3.4 concludes with a discussion on the preferred modality of each of these senses.

2.3.1 Vision

The visual system perceives the world by processing electromagnetic energy that enters the eyes. The majority of this energy is reflected from objects in the world. This reflected energy is perceived by the visual sense. The wavelengths of electromagnetic energy that the eyes can process are called visible light and they range in frequency from 400 to 700 nanometres [Sekuler and Blake 1990] (figure 2-4).

Light that enters the eye is focused as an image on the retina of the eye (figure 2-5). The retina is a thin network containing nerve cells and light receptors. There are two types of light receptors called *rods* and *cones*. These receptors absorb light causing a chemical reaction that activates nerve cells. Rods and cones respond differently to light and also adapt differently as light conditions change. Only cone receptors are associated with colour vision. Rods and cones are not evenly distributed in the retina. Rods greatly outnumber cones in the peripheral retina and there are no rods at all in the fovea.

The fovea is a small area on the retina composed entirely of cones. The fovea is located directly on the line of sight so anytime we look directly at an object the centre of its image falls on the fovea. The fovea is moved frequently and shifts towards areas of visual attention. Foveal vision has much higher acuity than the rest of the retina. This is partly because cone receptors perceive higher detail than rods. However, there is also what is called a magnification factor related to the number of receptors and the degree of neural processing. The fovea occupies 0.01% of the area of the retina yet it occupies about 8% of the area in the brain responsible for visual processing [Goldstein 1989, p89].

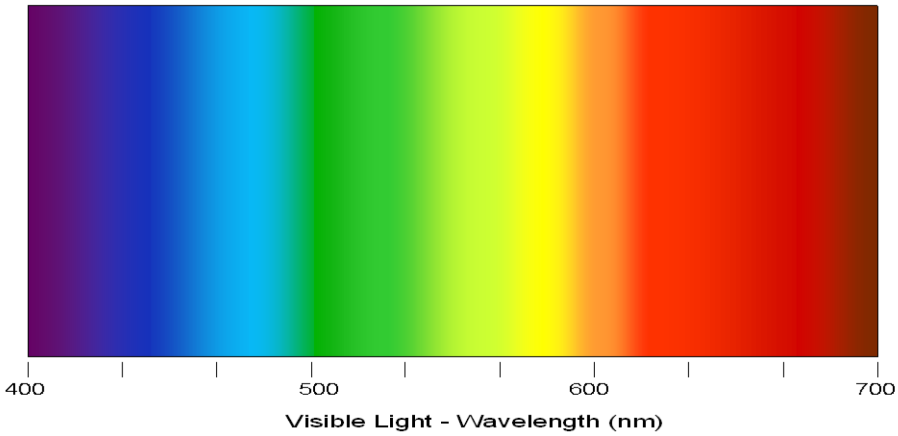


Figure 2-4 The spectrum of visible light.

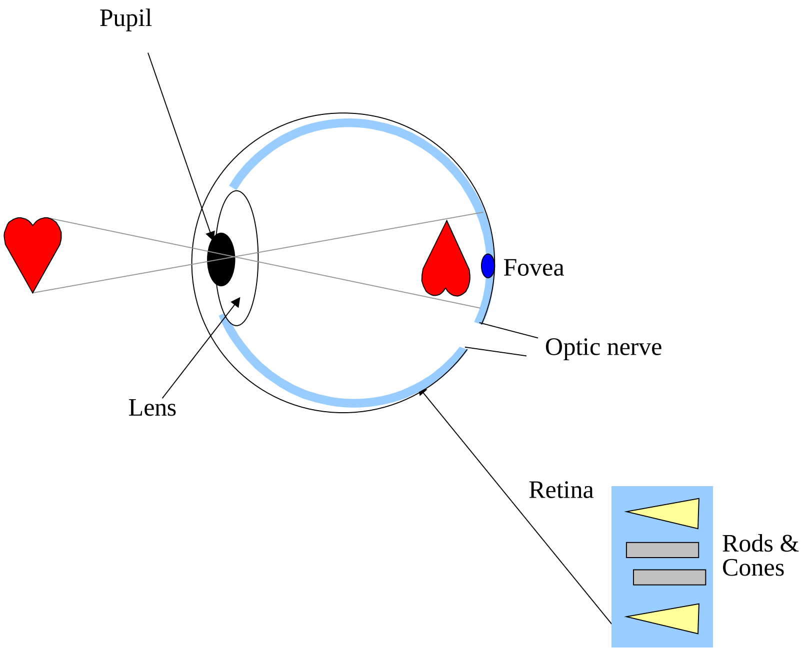


Figure 2-5 The anatomy of the eye.

Peripheral vision is less detailed than foveal vision and the number of rod receptors outnumber the number of cones by about 20 to 1 [Goldstein 1989, p36]. While details are better seen with cones, rod vision is more sensitive due to a property called spatial summation that allows rod vision to operate under lower levels of light than cone vision [Goldstein 1989, p55]. The peripheral region of the retina is good at detecting movement and changes in the visual environment [Card, Mackinlay et al. 1999].

The visual sense is characterised by a wide field of view. The centre of this field contains highly detailed colour information while the periphery is much less detailed but provides a spatial context. The field of view is 200º horizontally and 160º vertically [Card, Mackinlay et al. 1999] (figure 2-6).

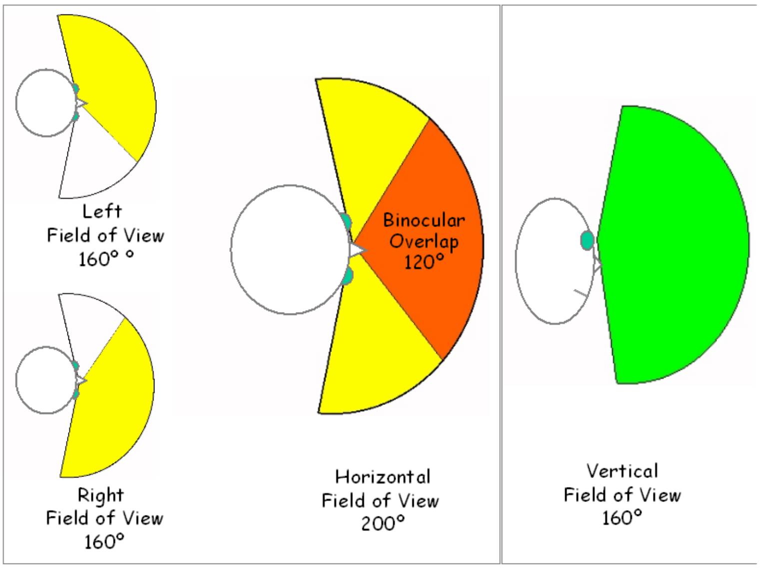


Figure 2-6 The field of view for vision.

It is estimated that we can perceive about 2 million colours [Goldstein 1989, p134]. Colour can be described as chromatic or achromatic. White, black and shades or grey are achromatic colours and are perceived through the rod receptors in the retina. Chromatic colours, such as blue, red, green and yellow are associated with the cone receptors in the retina. The *Trichromatic Theory* of colour perception is based on light stimulating 3 different cone receptors [Goldstein 1989, p139]. The degree of stimulation and the pattern of activity in the three receptors results in the perception of colour. The *Opponent Process Theory* of colour perception is based on the perceptual pairing of red-green and yellow-blue [Goldstein 1989, p139]. Both these theories apply at different levels of visual processing and help explain the perception of colour.

The perception of visual information is determined not only by the stimulus detected by rods and cones, but also the neural filtering that takes place in later nerve structures. Many areas of the brain responsible for the neural processing of visual information are spatially organised. There are a number of direct mappings between the retinal space and the organisation of nerve cells in the brain. These are called *retinotopic* maps and directly support the perception of spatial information [Goldstein 1989, p75].

The slight difference between the images received by the two eyes is called binocular disparity. This produces the stereoscopic effect and this is one of a number of cues used to perceive depth (table 2-3). Other depth cues are generated from the eye muscles as the eyes converge and focus (accommodate) to look at an object. Movement of an observer produces the depth cues of motion parallax, deletion and accretion. Motion parallax is the effect that distant objects appear to move more slowly than close objects. Sideways movement may cause a closer object to cover (delete) or uncover (accrete) an object behind it. A number of other depth cues are sometimes called the painter's cues. These cues include occlusion, relative height, relative size, familiar size, atmospheric perspective, linear perspective and texture gradients (figure 2-7)

Depth Cue	0-2 meters	2-30 meters	> 30 meters
Accommodation and Convergence[2]	X		
Occlusion	X	X	X
Relative Size	X	X	X
Relative Height		X	X
Atmospheric Perspective			X
Motion Parallax	X	X	
Binocular Disparity[3]	X	X	

Table 2-3 Distances at which the different depth cues are effective.
Adapted from Goldstein [Goldstein 1989, p 231].

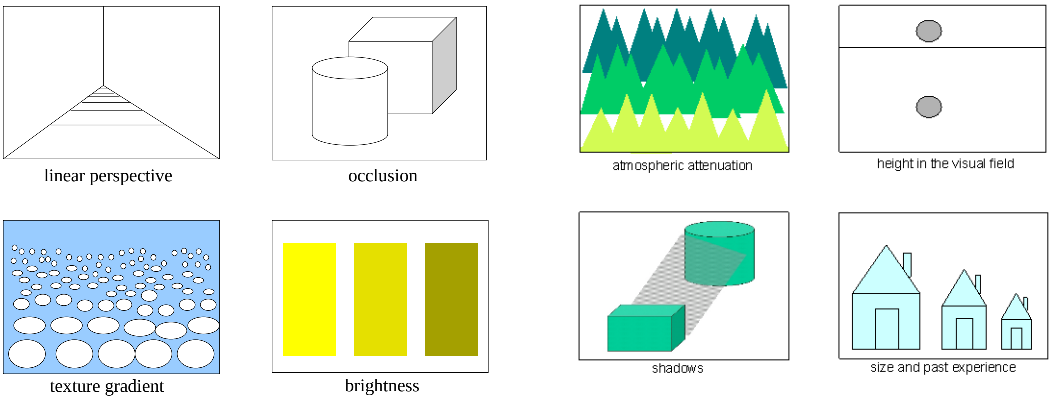


Figure 2-7 The painter's algorithms are techniques used by artists to create a perception of depth.

2.3.2 Hearing

In the real world hearing provides a signalling function [Goldstein 1989, p310]. It notifies a person of events that cannot be seen. Sounds are caused by rapid pressure changes in the air or other mediums. The vibrations of an object usually generate these air pressure changes which are described as a sound wave [Sekuler and Blake 1990, p292]. The perceived properties of sounds are closely related to the frequency and amplitude of sound waves (figure 2-8).

Pure tones have a sound wave described by a sine-wave. The frequency of a pure tone is indicated by how many times the sine-wave repeats itself. The perceived pitch of a pure tone is directly related to the frequency of the wave. Humans can hear frequencies between about 20 and 20,000 Hz [Goldstein 1989, p75]. However, normal sounds are described by much more complex waves than a pure tone. These stimuli usually

[2] In many Virtual Environments the images are projected or displayed on a screen. The image may be perceived at a different depth to the display screen. In these cases the user may experience problems with accommodation-convergence conflict. The eyes converge on the position of the image but must focus (accommodate) on the display surface. Normally the eyes converge and focus at the same point in space.

[3] Virtual Environments often provide a mechanism for providing the user with stereoscopic views of the scene. This enhances the perception of depth at distances of 0-30 meters. Virtual Environments may also provide a mechanism for tracking the user's head movement and updated the view based on head position. This provides useful motion related depth cues.

contain waves of many frequencies and these may all influence the perceived pitch [Kramer 1994a].

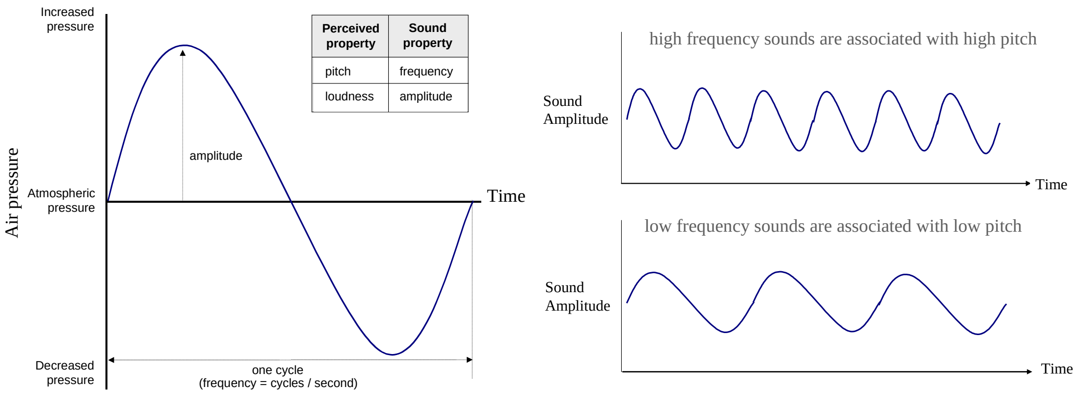


Figure 2-8 The perceived pitch and loudness of a sound are related to the frequency and amplitude of the sound wave [Goldstein 1989, p313].

The perceived loudness of the sound wave is directly related to the amplitude of the wave [Goldstein 1989, p317]. Amplitude is usually described in a logarithmic scale of pressure levels called decibels (dB) [Sekuler and Blake 1990, p296]. A barely audible sound has a pressure level of 0dB while a normal speaking voice is about 60dB. A jet engine produces about 140dB and this corresponds to a level at which pain is felt [Sekuler and Blake 1990, p297].

Apart from loudness and pitch the third type of perceptual experience related to sound is called timbre. Loudness is the perceived magnitude of auditory sensation. Pitch relates to how high or low a tone sounds. Timbre is the quality that describes why sounds with the same pitch and loudness are perceived to be different. For example, timbre distinguishes two musical instruments, such as a flute and an oboe, playing the same note at the same loudness.

The ear is the organ responsible for sensing sounds and is composed of three main parts; the outer ear, the middle ear and the inner ear. The outer ear is composed of the pinna and the auditory canal (figure 2-9). Although the pinna assists in the location of sounds, the main function of the outer ear is to amplify sounds with a frequency between 2,000Hz and 5,000Hz [Goldstein 1989, p319]. As a consequence of the amplifying effect of the auditory canal even a very quiet sound with a frequency of about 3000Hz can be heard [Sekuler and Blake 1990, p300]

The middle ear begins at the eardrum [Sekuler and Blake 1990, p301]. Sound waves cause the eardrum to vibrate and these vibrations are transmitted along a small group of bones called the *ossicles*. The ossicles consist of the *malleus*, the *incus* and the *stapes* (figure 2-10). The *stapes* rests against the *oval window*. The oval window is a membrane-covered hole that separates the middle ear from the inner ear. The ossicles act together to concentrate the changes in air pressure exerted on the ear drum onto the much smaller surface of the oval window. This is necessary because the inner ear is full of a fluid that is much denser than air and fluctuations in pressure are not easily transmitted between air and fluids.

The inner ear contains two main components; the *semi-circular canals* and the *cochlea* (figure 2-11). The semi-circular canals help to maintain posture and balance while the cochlea is the part of the inner ear associated with hearing [Sekuler and Blake 1990, p305]. Vibrations transmitted from the eardrum through the ossicles and onto the oval window create pressure waves in the fluid-filled cochlea. These pressure changes cause movement in the small receptor cells in the cochlea. Stimulation of these receptor cells, called hair cells, results in the perception of sound.

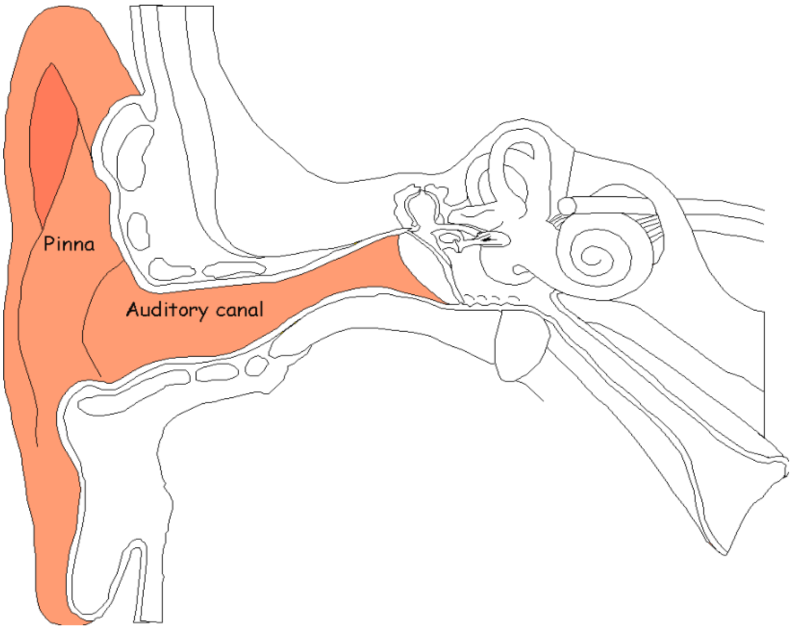


Figure 2-9 The outer ear.

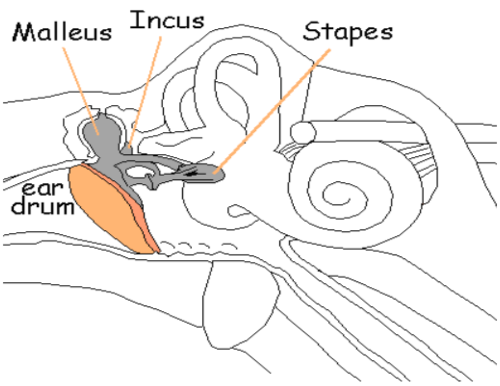


Figure 2-10 The middle ear.

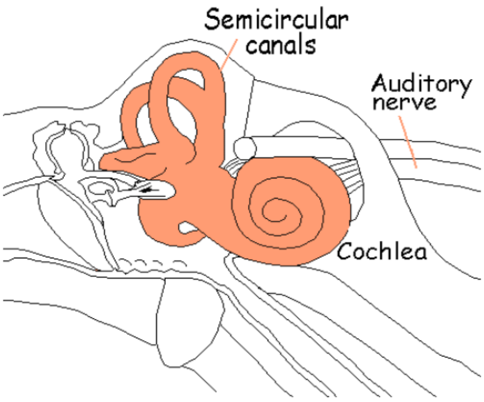


Figure 2-11 The inner ear.

Sound waves of different frequencies stimulate hair cells at different locations along the cochlea. There is in fact a very ordered relationship between the frequency of sound waves and the location of hair cells that are stimulated. This is known as a *tonotopic* organisation [Sekuler and Blake 1990, p314] and contrasts with the spatially organised maps that occur in visual pathways. The tonotopic organisation of receptor cells in the cochlea is also preserved in the organisation of higher brain structures [Goldstein 1989, p98].

There are three auditory cues for localising sound in space:

- *interaural time difference*
- *interaural intensity difference*
- *pinnae cues.*

The difference between the arrival times of a sound at each ear is called the *interaural time difference* (figure 2-11) [Sekuler and Blake 1990, p324]. Sounds arriving at each ear may also differ in loudness. For example, a sound to the right side of the head may be clearly heard in the right ear while on the left ear it will be softer. This is because the head itself blocks some of the sound, but also because the sound loses energy as it travels further. This localisation cue is called the *interaural intensity difference* (figure 2-11) [Sekuler and Blake 1990, p325].

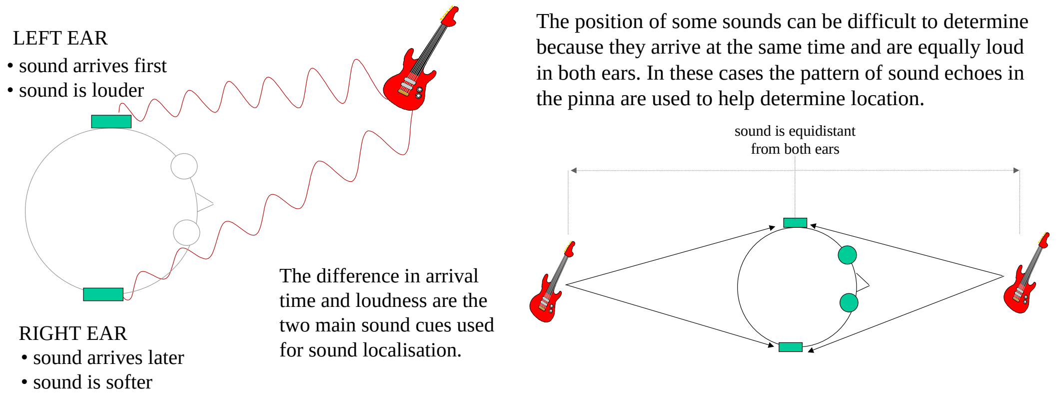


Figure 2-12 The cues used for locating sounds.

The third localisation cue is from the distinctive pattern of echoes that a sound causes in each ear. These patterns result from the distinctive shape of the outer ear and are called *pinnae cues* [Goldstein 1989, p367]. These cues help resolve discrepancies about sound location that can occur when sounds are equidistant from both ears (figure 2-12)[4].

The auditory sense has limited spatial precision compared to the visual sense. Within the arc covered by central vision we can resolve very fine detail. Kramer notes that the typical resolution of fovea is about 0.016° of the visual arc and can even be as detailed as 0.005° for some tasks [Kramer 1994a]. However, the capacity of the auditory system to locate a position in space is much coarser, about 1° of resolution in front of the listener and 5-10° degrees to the periphery [Kramer 1994a].

2.3.3 Haptics

The word *haptics* comes from the Greek and means to grasp. Real world tasks performed by the hands usually involve exploring and manipulating objects. Haptics differs from other sensory modalities in that it is intrinsically requires both input and output [Salsibury and Srinivasan 1997].

The haptic system usually initiates action to stimulate sensory perception of the world. For example, to feel the hardness of a surface a person might tap on the surface. In this example, the motor system must initiate the sensory reception. That is, output is required to generate sensory input. Therefore haptic display devices are not passive displays, as is usually the case with auditory and visual displays but instead requires an element of interaction from the user.

---

[4] However, another way a person can resolve these ambiguous sound locations is by moving the head.

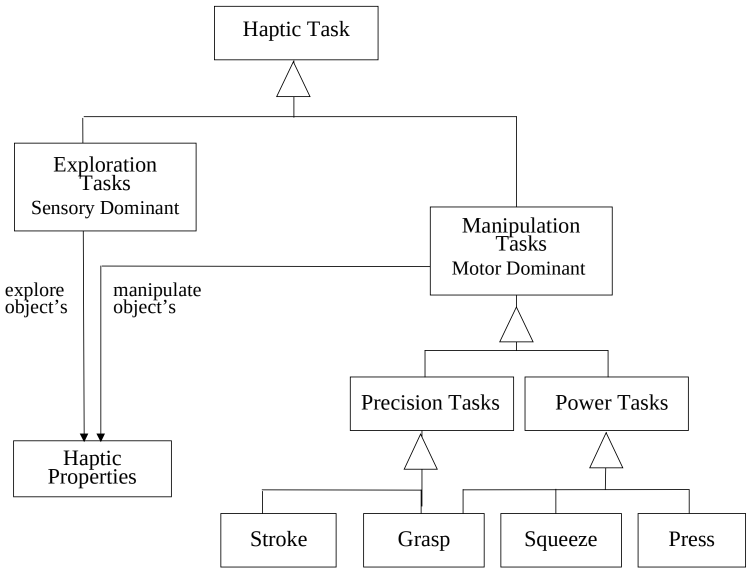


Figure 2-13 Haptic tasks can be divided into exploration and manipulation tasks. Both types of tasks are used together to explore the haptic properties of an object.

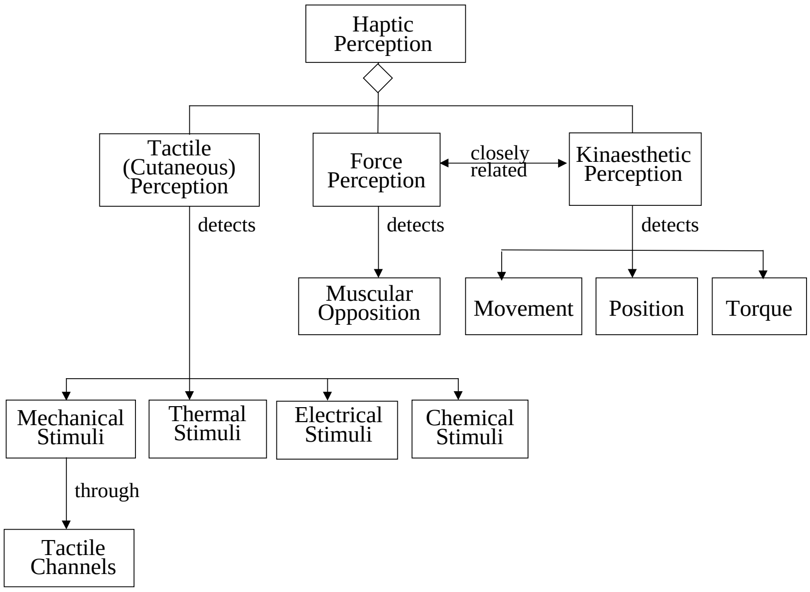


Figure 2-14 Haptic perception requires the integration of information from tactile, force and kinaesthetic receptors.

Haptic tasks can be divided into two main groups, exploration tasks which are primarily sensory oriented (output) and manipulation tasks which are primarily motor driven (input) [Durlach and Mavor 1995]. To investigate an object's properties typically requires a combining of exploration and manipulation skills (figure 2-13). For example, picking up an object and feeling its shape, texture and compliance may involve stroking, squeezing, pinching and pressing the object.

Haptic perception requires the complex integration of stimuli sensed through *tactile*, *kinaesthetic* and *force* receptors [Stuart 1996] (figure 2-14). Hence haptics itself involves multi-sensory perception. For example, picking up an object and feeling its properties requires integrating information from *tactile*, *force* and *kinaesthetic* senses (figure 2-15).

The *tactile* sense is associated with receptors in the skin and refers to the sense of contact with surfaces. This is what people often refer to as touch. Tactile receptors can discriminate very fine surface properties such as small shapes and fine textures. Tactile information is quite complex and encoded in patterns of spatial and temporal distribution activation across the skin [Salsibury and Srinivasan 1997]. The tactile sense detects mechanical properties through four main channels. These four different channels are provided through different anatomical structures [Goldstein 1989] and combine to produce the different sensations of touch (figure 2-16).

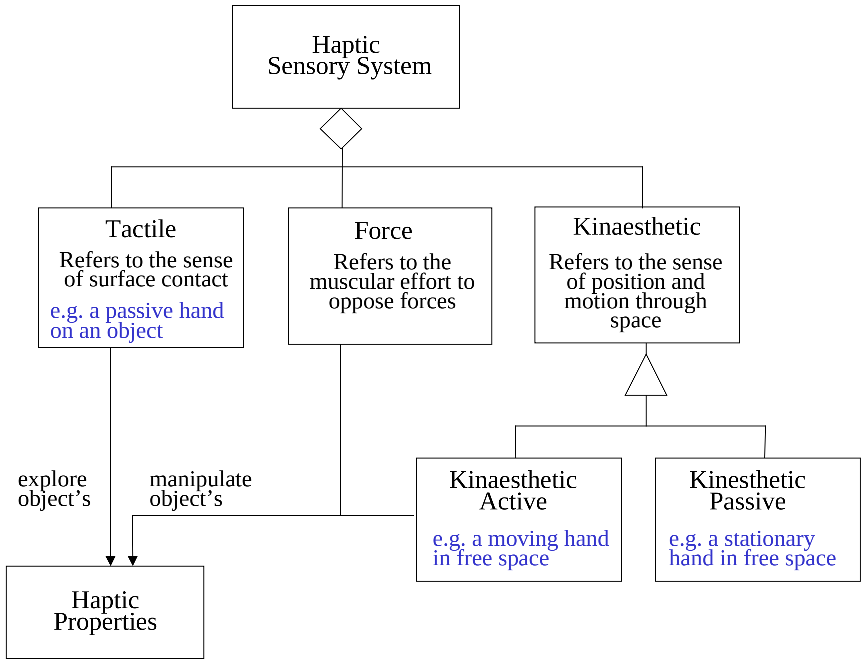


Figure 2-15 To explore haptic properties requires the combined use of tactile, kinaesthetic (active) and force perception.

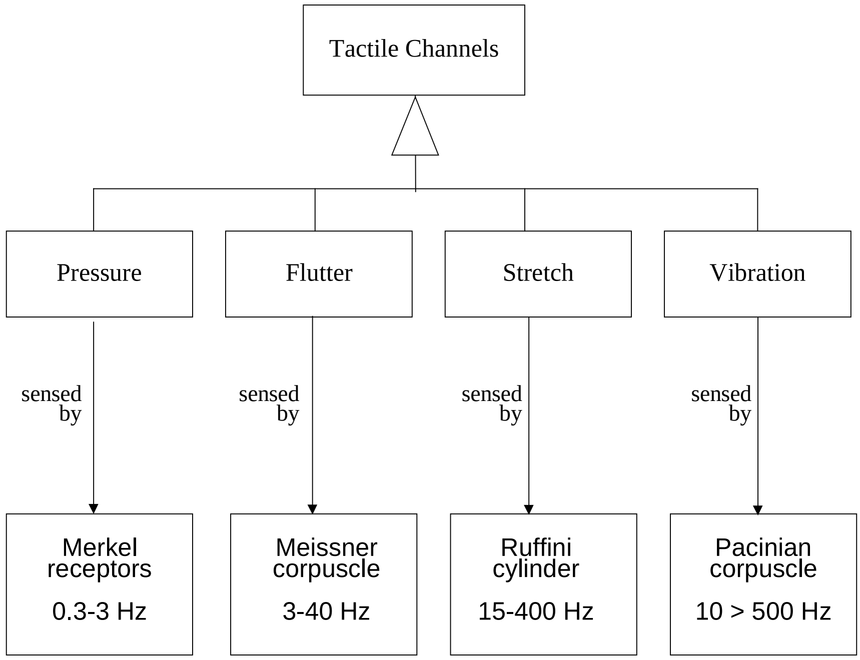


Figure 2-16 Mechanical tactile properties are sensed through the combination of four main

channels [Goldstein 1989]

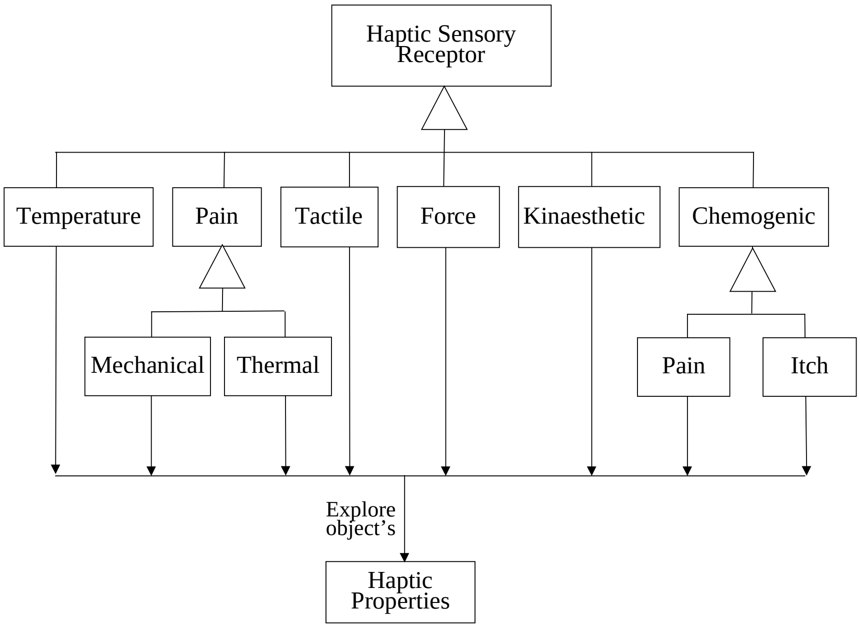


Figure 2-17 The different types of haptic sensory receptors.

The *kinaesthetic* sense is associated with receptors in muscles, tendons and joints and is concerned with detecting the position, movement and torque of limbs. *Kinaesthetic* receptors can only sense coarse properties of objects that require hand or arm motion. These include properties such as, large shape, coarse texture or surface compliance.

The *force* sense has muscle receptors that respond to muscular effort in opposing external forces. Although the *force* and *kinaesthetic* senses are separate systems they work in close cooperation and are often treated as a single system.

There is no simple way to measure the accuracy of the haptic sense. It has been shown that joints close to body can distinguish finer rotations than those more distant. For shoulder rotations the JND (Just Noticeable Difference) is about 0.8°, 2° for the wrist and elbow and 2.5° for finger joints [Durlach and Mavor 1995]. In tasks requiring locating targets with a pointing finger, the accuracy of the gesture depends on the speed, direction and magnitude of movement as well as the shape of the target [Durlach and Mavor 1995]. The ability to judge length using the finger spanning method varies depending on the length of the span. The JND for objects of 10mm is 1mm, while for longer objects of the order of 80mm the JND increases to 2-4mm [Durlach and Mavor 1995].

Haptics is quite a complex sense and apart from *force*, *kinaesthetic* and *mechanical tactile* receptors it also includes receptors for detecting temperature, pain and various chemical stimuli (figure 2-17). Haptic perception is more like vision in terms of high-level neural organisation. For example, in the brain, nerves that process haptic inputs are characterised by a spatial organisation based on the position of receptors on the body (*somatotopic map*) [Goldstein 1989, p98][5].

[5] The neurones of the visual cortex are also arranged around a spatial mapping derived from the position of receptors in the retina (*retinatopic map*) [Goldstein 1989, p98].

The most commonly used haptic display is currently the Phantom™ [Salsibury and Srinivasan 1997]. The Phantom™ displays forces to a finger point and can mimic a combination of tactile, kinaesthetic and force stimuli. A shortcoming of this device is that it only provides force feedback at a single point[6]. In the real world our haptic sense integrates information over multiple spatial locations simultaneously. Because the Phantom™ provides forces only at a single spatial location, global spatial information about space needs to be integrated over time. The Phantom™ display has been compared to *"feeling the world with a stick"* [Srinivasan and Basdogan 1997].

Haptic displays are rapidly being developed, however as of 2002 it is still not possible to find commercial displays that use all the haptic channels. Some more unusual haptic displays have been developed for research purposes. For example, the `ThermoStylus` can display thermal information to users on the end of force feedback device [Ottensmeyer and Salsibury 1997]. Note that while some attempts have been made to develop simple *tactile* displays [Shimago 1992] as yet none have become commercially available. Furthermore no commercial devices are available that combine both *tactile* and *kinaesthetic* force feedback [Oakley, McGee et al. 2000].

2.3.4 Sense of Time and Space

The *Modal Specific Theory* [Friedes 1974] is a *psychophysical* theory. *Psychophysics* is a field of psychology that studies the relationships between physical stimuli and sensory responses. The *Modal Specific Theory* maintains that each sense has a unique method of transferring information. Thus each sense is adept with certain kinds of complex information.

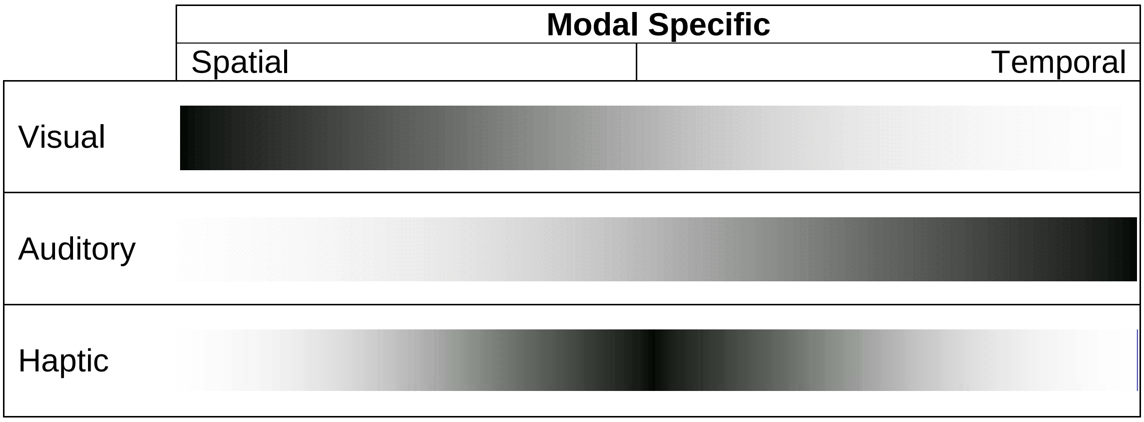


Table 2-4 A comparison of the modal specific nature of the different senses.

Vision is described as a *spatial* sense and so is adept at interpreting *spatial* relationships. Hearing is a *temporal* sense and so is adept at interpreting *temporal* relationships. Haptics is adept at sensing movement. Understanding movement requires the interpretation of both *spatial* and *temporal* relationships.

Table 2-4 compares the modal specific nature of each sense. The immediate intuition from this comparison is that vision should be used for spatial metaphors and hearing

---

6 With these limitations in mind this thesis will develop some haptic-visual displays (chapter 12) that make use of the Phantom™. As more sophisticated haptic display devices become available it is expected that the usefulness of haptics for displaying abstract information will improve.

should be used for temporal metaphors. Haptic display might be effective for displaying both spatial metaphors and temporal metaphors.

2.4 Conclusion

It could be suggested that the abstraction of the multi-sensory design space into nine metaphor classes is an over simplification. However, that is precisely one of the strong points of this meta-abstraction. At the high-level it provides a very simple description of the design space. Even with this quite abstracted view of the multi-sensory design space it is possible to develop a useful design process and to incorporate some sensible high-level design guidelines.

Abstraction is a technique used in software engineering to help mask complexity [Rumbaugh, Blaha et al. 1991]. An advantage of abstraction is that it hides lower level details by creating general classes. This then allows more general classes to be compared. For example, the concept of *Holden*, *Ford* and *Volkswagen* can all be abstracted to the more general concept of *car*. The concepts of *sailing boat*, *powerboat* and *rowboat* are all types of an abstract concept called *boat*. It may be appropriate to compare the features of a *Ford* with a *Holden*. However, it is also possible to compare concepts at a higher level. For example, the more general concept of *cars* can be compared with the general concept of *boats*.

The abstractions that make up the MS-Taxonomy also mask complexity and can be used to compare general design concepts at different levels. For example, at the highest level spatial metaphors can be compared with direct metaphors. At a lower level, direct visual metaphors can be compared with direct auditory metaphors. The thesis aims to develop these higher-level concepts of the MS-Taxonomy. However, the MS-Taxonomy also provides a framework that encapsulates the lower-level complexity of the multi-sensory design space. This requires a broad and detailed look at the lower level concepts[7].

Chapter 3, chapter 4 and chapter 5 discuss the nine metaphor classes that make up the MS-Taxonomy. These chapters spend some time discussing and exploring the lower-level concepts of the multi-sensory design space. These more detailed concepts are illustrated using previous work from visual, auditory and haptic display. These chapters serve to both explain the concepts that make up the taxonomy and also to review previously developed displays of abstract data. Of course, this previous work has never before been considered in terms of the MS-Taxonomy.

The goal of developing the MS-Taxonomy is to provide a new, structured way to look at the multi-sensory design space. The hope is that this categorisation will allow for better comparison of display concepts across the different senses. This may lead to new insights about how best to use the different senses in abstract displays. These insights may hopefully be captured as principles or guidelines for future designers. This structured look at the multi-sensory design space may also highlight opportunities for applying design principles across the different senses. For example, previous findings from visual display may be generalised to auditory and haptic display.

[7] While the high-level abstractions of the MS-Taxonomy provide a useful conceptual tool for design, there is bound to be some general debate over the lower level concepts in the taxonomy. This is a fairly common problem with software abstractions. For example, is a hovercraft a *car* or a *boat*?

One problem with developing a new categorisation such as the MS-Taxonomy is to verify that it adequately covers the multi-sensory design space. Since the design of multi-sensory, abstract information displays is a new field no other such taxonomies have been developed. However, Card, Mackinlay et al. proposed a categorisation of the visual design space for display of abstract data [Card, Mackinlay et al. 1999]. This visual taxonomy is based on an alternative set of criteria to the MS-Taxonomy. As part of this thesis, this visual taxonomy was extended to include both auditory and haptic display and so create an alternative categorisation of the multi-sensory design space [Nesbitt 2001b]. In this same publication, the MS-Taxonomy and the extended visual taxonomy were compared and both taxonomies were shown to cover the same design space. There are some distinct differences in structure of the two taxonomies, but many of the concepts from each taxonomy overlap. More detail on the extended Card, Mackinlay et al. taxonomy is provided in Appendix C.

Chapter 3

Spatial Metaphors

City Escape (1992)

"Equations are just the boring part of mathematics. I attempt to see things in terms of geometry." [Stephen Hawking]

Chapter 3

MS-Taxonomy: Spatial Metaphors

3.1 Introduction

In the real world a great deal of useful information is dependent on the perception of space. For example, driving a car requires an understanding of the relative location of other vehicles. Parking the car requires a comparison of the size of the car with the size of the parking space. Navigating the car requires an understanding of the interconnections and layout of roadways. Real world information is often interpreted in terms of spatial concepts like position, size and structure. Abstract information can also be interpreted in terms of these spatial concepts.

Spatial metaphors relate to the scale of objects in space, the position of objects in space and the structure of objects in space. This chapter describes in more detail the three classes of spatial metaphors (table 3-1):

- *Spatial visual metaphors* concern the way pictures are organised in space.
- *Spatial auditory metaphors* concern the way sounds are organised in space.
- *Spatial haptic metaphors* concern the way forces are organised in space.

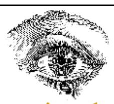
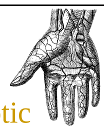
		SENSORY DISPLAY MODES		
		 visual	 auditory	 haptic
METAPHOR CLASSES	**SPATIAL METAPHORS (chapter 3)**	**Spatial Visual Metaphors (section 3.2)**	**Spatial Auditory Metaphors (section 3.3)**	**Spatial Haptic Metaphors (section 3.4)**
	DIRECT METAPHORS	Direct Visual Metaphors	Direct Auditory Metaphors	Direct Haptic Metaphors
	TEMPORAL METAPHORS	Temporal Visual Metaphors	Temporal Auditory Metaphors	Temporal Haptic Metaphors

Table 3-1 The focus of chapter 3 is on the three classes of spatial metaphors.

The MS-Taxonomy distinguishes between spatial visual, spatial auditory and spatial haptic metaphors. However, the general concepts that describe spatial metaphors are independent of each sense. It is simply the different ability of each sense to perceive space that needs to be considered for each of these metaphor classes. Because the concepts abstract across the senses it is possible for spatial metaphors to be directly compared between senses. For example, the ability of the visual sense to judge the position of objects in space can be compared with the ability of hearing to locate a sound in space. This sensory independence also enables concepts to be reused[1] between senses. For example, a spatial visual metaphor, such as a `lineplot`, can be directly transferred to a spatial haptic metaphor to create a `haptic lineplot`.

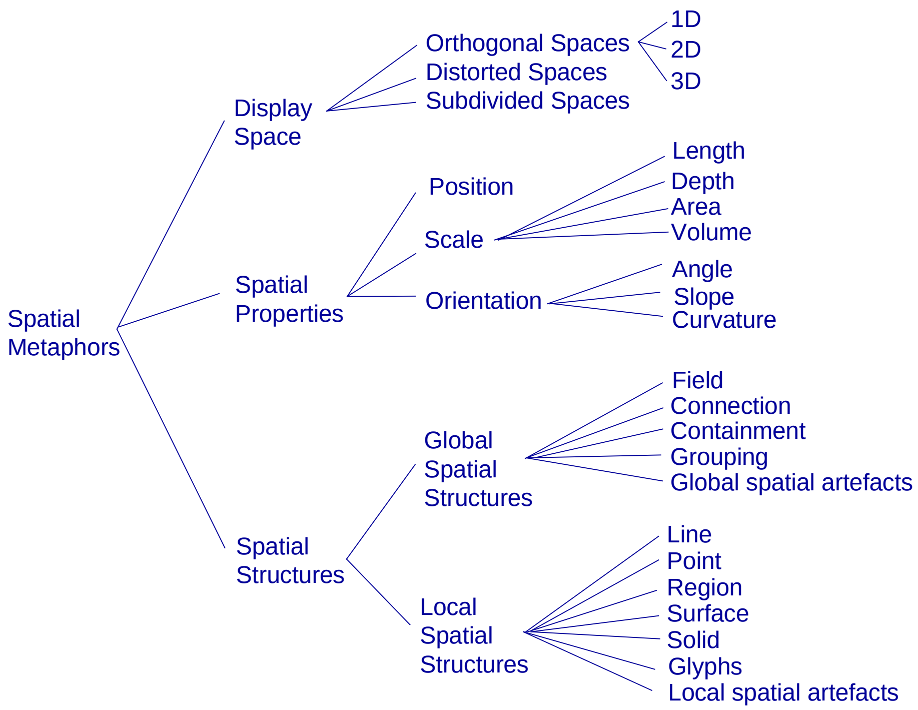


Figure 3-1 The general hierarchy of concepts that describe spatial metaphors.

The concepts that describe spatial metaphors[2] are shown in figure 3.1. To understand spatial metaphors requires an overview of the high-level concepts and a detailed understanding of low-level concepts. This chapter first introduces the high-level concepts that describe spatial metaphors (section 3.1.1). These descriptions may initially seem vague. However, to clarify the concepts, a more detailed discussion is included for each of the three senses, visual (section 3.2), auditory (section 3.3) and haptic (section 3.4). This detailed discussion includes many concrete examples from existing information displays[3] and serves a number of purposes. It provides a review of existing literature, it helps to explain the abstract concepts that describe spatial

[1] Reuse is often a goal of abstraction in software engineering [Pfleeger 1998]. This has also been called *sharing* [Rumbaugh, Blaha et al. 1991]. Sharing spatial metaphors across the senses is one potential advantage offered by the MS-Taxonomy.

[2] The classification used in the MS-Taxonomy for spatial metaphors is less intuitive than the structure for direct metaphors (chapter 4). For direct metaphors it seems simple to discuss the interpretation of direct visual, auditory or haptic properties. On the other hand, spatial metaphors deal with concepts that are more difficult to understand and explain. This is perhaps not surprising as spatial metaphors are concerned with how to design and interpret the more abstract notion of space.

[3] A summary of the applications discussed in this chapter is available in Appendix A.

metaphors and it demonstrates how these generic concepts apply across the different senses.

3.1.1 Spatial Metaphors

The design space for spatial metaphors can be described using the following general concepts (figure 3-2):

- the display space
- spatial structure
- spatial properties.

The display space is the region where the data is presented. All spatial metaphors have as their basis an underlying display space that is used to arrange the display elements. For example, the `scatterplot` defines a 2D orthogonal display space by mapping data attributes to the x and y axis. Points are then interpreted in terms of this display space. In the real world, space is perceived as constant, however in an abstract world the properties that define the space can also be designed. For example, one axis of the `scatterplot` could be defined as a logarithmic space. This has the effect of changing the way the position of points is interpreted.

There are a number of strategies for designing the display space when presenting information and these include using orthogonal spaces (1D, 2D, 3D), distorted spaces and subdivided spaces (figure 3-3). To help describe the display space the terminology of *type*, *metric* and *continuity* are also used. These terms are discussed in section 3.1.2.

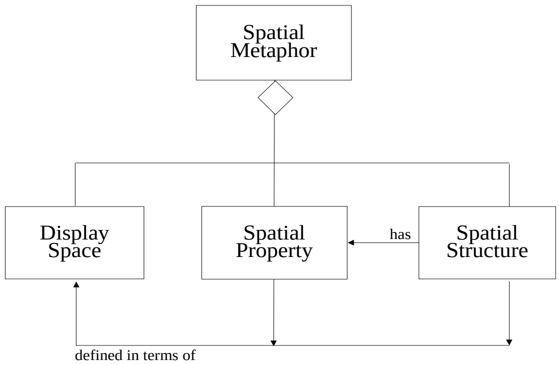


Figure 3-2 A UML diagram showing the high-level components of spatial metaphors.

The entities that occupy the display space are called spatial structures. For example in the `scatterplot`, the points are spatial structures. Spatial structures also describe the arrangement of entities within the display space. For example, a group of points in the `scatterplot` can be considered a more global spatial structure. The MS-Taxonomy distinguishes two levels of organisation for presenting information and these are global spatial structures and local spatial structures (figure 3-4).

Spatial structures may have spatial properties. Spatial properties describe qualities that are defined in terms of the display space. For example, in the `scatterplot` the position of points is used to convey information. This information is interpreted in terms of the abstract space defined by the x and y axis. The spatial properties used for presenting information include position in space, scale in space and orientation in space (figure 3-5).

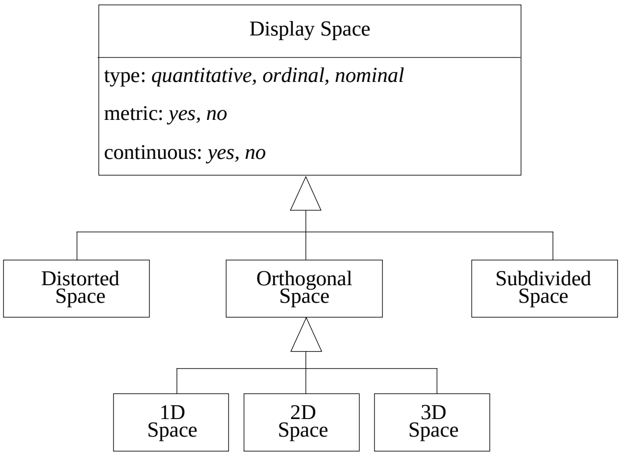


Figure 3-3 A UML diagram showing the types of display space.

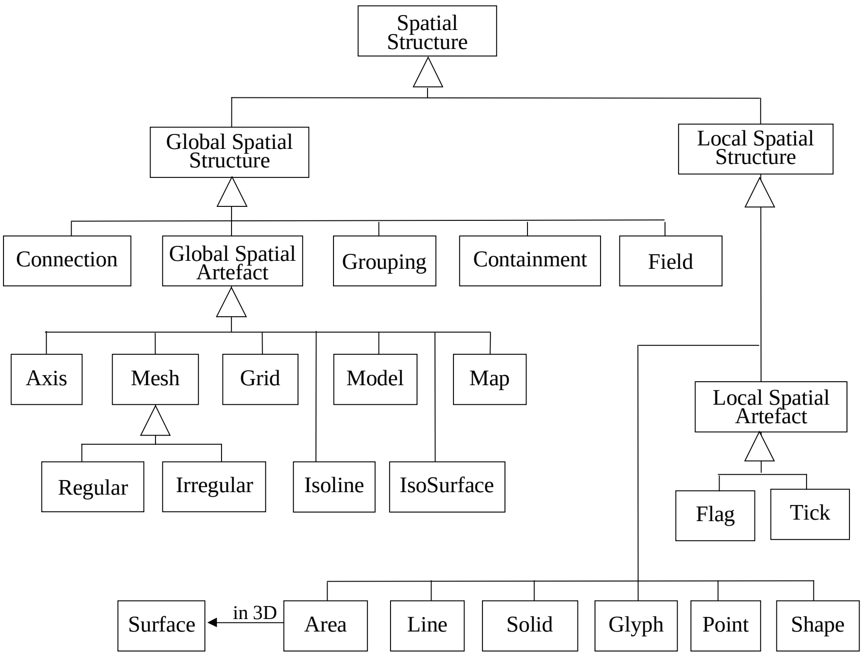


Figure 3-4 A UML diagram showing the types of spatial structure.

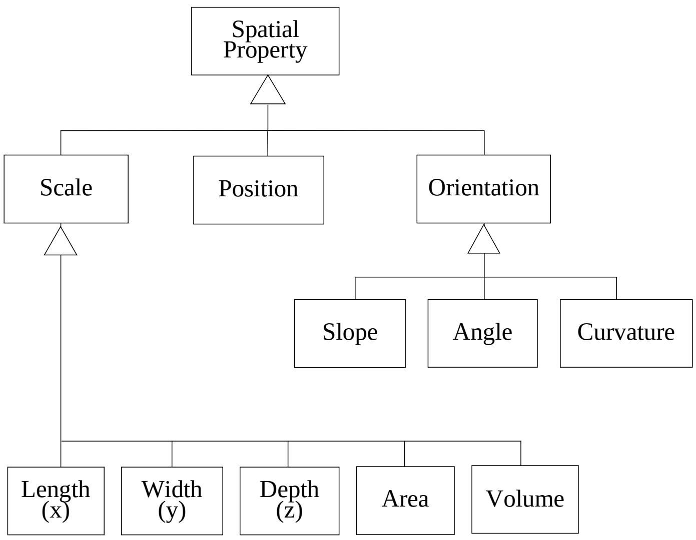
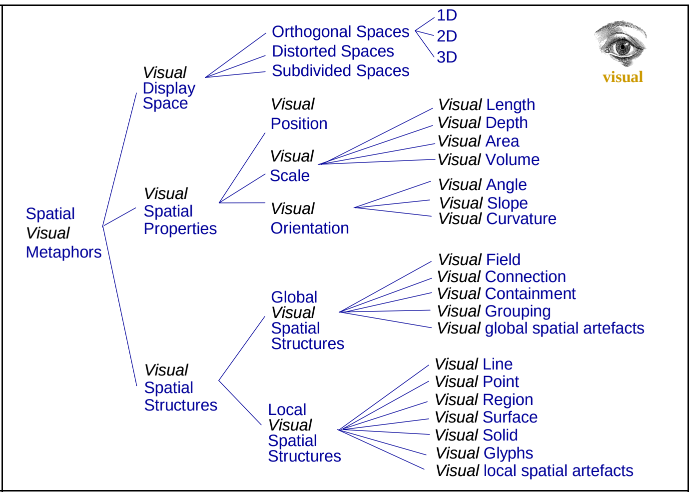


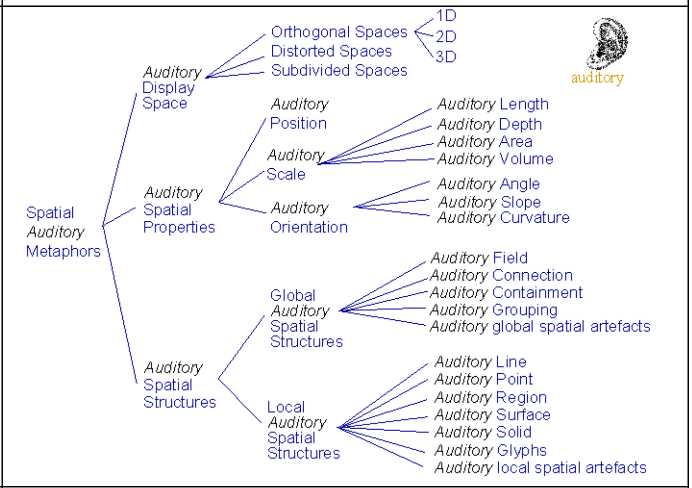
Figure 3-5 A UML diagram showing the types of spatial properties.

The example of the `scatterplot` was used to describe the high level design concepts used in spatial metaphors. This is an example from the domain of visual display; however, the concepts of spatial metaphors are general and can be applied to each of the senses (table 3-2). Chapter 3 contains a number of examples of the concepts for spatial visual metaphors (section 3.2), spatial auditory metaphors (section 3.3) and spatial haptic metaphors (section 3.4). However, before proceeding, section 3.1.2 introduces some more terminology used for defining the display space.

Spatial **Visual** Metaphors



Spatial **Auditory** Metaphors



Spatial **Haptic** Metaphors

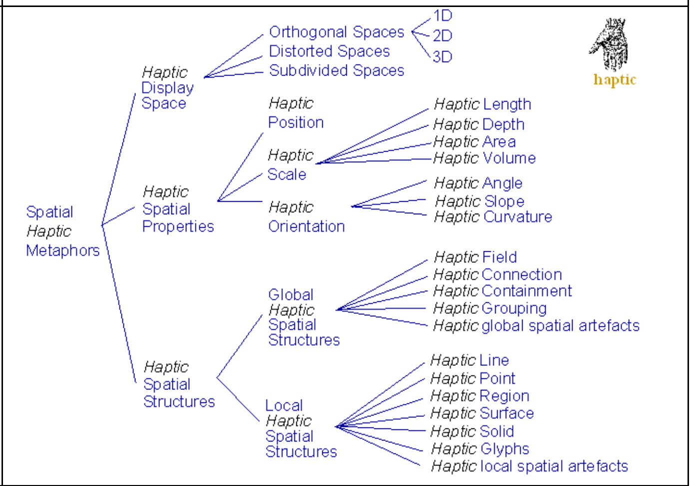


Table 3-2 The concepts describing spatial metaphors can be applied to each sense.

3.1.2 The Display Space

The display space is defined using the terms *type*, *metric* and *continuity*. The *type* of display space can be quantitative, ordinal or nominal. This idea of *type* is sometimes used in Information Visualisation to describe data attributes. Type has also been used to describe the *spatial substrate* for visual displays [Card, Mackinlay et al. 1999].[4] The *spatial substrate* described by Card, Mackinlay et al. is very similar to the concept described here as the visual display space.

A display space may also have a *metric*. This is related to the mathematical concept of a *metric space* [Jain 1993] and is a very important consideration when designing the information space. If a display space has a *metric,* it is possible to interpret the distance between any two points in that space. For example, being able to compare the distance between points on a `scatterplot` is often important for interpreting relationships. In general, the spatial properties of scale and orientation rely on a metric being defined as the user must interpret the distance between points.

Most display spaces are *continuous* in that there are no breaks in the space. The exception to this occurs with subdivided display spaces. In this case the display space is deliberately broken into segments to allow disparate elements to be displayed. For example, Bertin describes a matrix of `scatterplots` [Bertin 1981]. In this case the space is disjoint rather than continuous so there is no spatial relationship that can be inferred between the individual `scatterplots`.

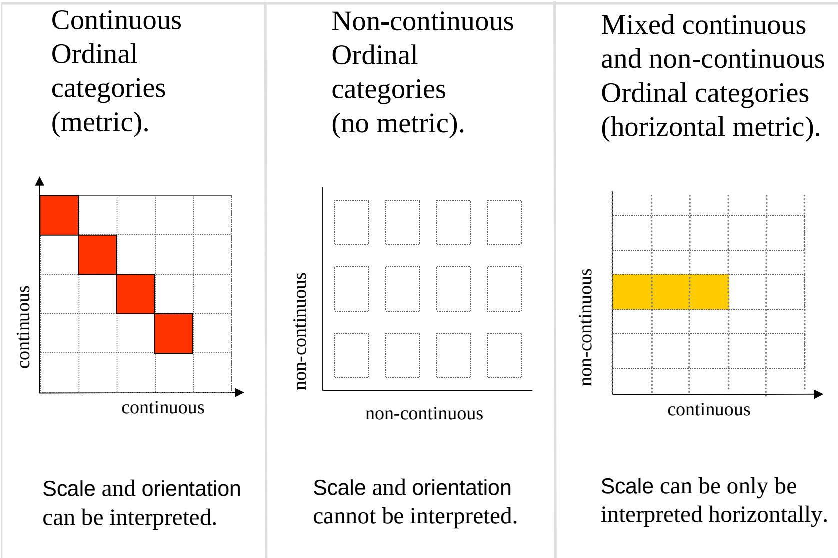


Figure 3-6 Some examples of continuous and non-continuous ordinal 2D visual spaces.

A space that is quantitative in type is defined by real numbers and is continuous (figure 3-6). Such a space also has a metric defined.

---

[4] There are a number of similarities between the concepts described in the MS-Taxonomy and the categorisation of the *visual* design space provided by Card, Mackinlay et al. [Card, Mackinlay et al. 1999]. There are also a number of subtle differences between the two approaches. The Card, Mackinlay et al. description of the visual design space has been extended to cover the full multi-sensory design space [Nesbitt 2001b]. This publication also compares the MS-Taxonomy and the extended Card, Mackinlay et al. taxonomy. A copy of the paper is included in Appendix C.

A space that is ordinal has been divided into ordered categories (figure 3-6) and is usually continuous. However, it is also possible to describe a non-continuous ordinal space that uses disjoint categories. For a display that uses a continuous ordinal space it makes sense to interpret the meaning of distance between spatial categories. With a non-continuous ordinal space the idea of a metric does not hold. In this case it is not appropriate to interpret distance as meaningful.

Nominal space is divided into non-ordered, categories and is non-continuous (figure 3-6). Nominal space does not describe a metric and so interpreting distance within these spaces is not meaningful.

3.2 Spatial Visual Metaphors

This section describes the concepts of spatial metaphors as they apply to the visual sense (table 3-3). Spatial visual metaphors relate to the visual perception of the scale of objects in space, the position of objects in space and the structure of objects in space. The important considerations in the design of spatial visual metaphors are (figure 3-7):

- the visual display space (section 3.2.1)
- the visual spatial structure (section 3.2.2)
- the visual spatial properties (section 3.2.3).

Due to the relative maturity of the field, it is not surprising that a number of generic designs for visual display have evolved. These generic designs describe a specific way of using visual display space, visual spatial structure and visual spatial properties for displaying abstract data. For example, a `histogram` is a type of generic visual design. These existing generic designs are summarised in section 3.2.4.

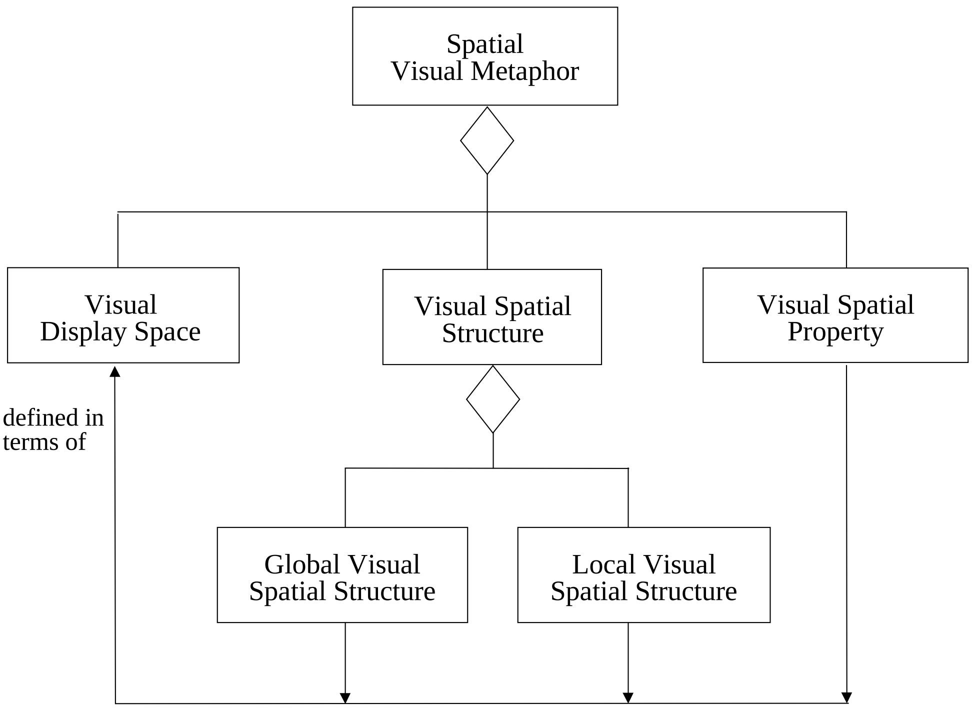


Figure 3-7 A UML diagram showing the components of the spatial visual metaphors.

Spatial Visual Metaphors (section 3.2)	Visual Display Space (section 3.2.1)	Orthogonal Visual Space	1D (section 3.2.1.1) 2D (section 3.2.1.2) 3D (section 3.2.1.3)
		Distorted Visual Space (section 3.2.1.4)	
		Subdivided Visual Space (section 3.2.1.5)	
	Visual Spatial Structure (section 3.2.2)	Global Visual Spatial Structure (section 3.2.2.1)	Visual Field Visual Connection Visual Containment Visual Grouping Visual Global Spatial Artefact
		Local Visual Spatial Structure (section 3.2.2.1)	Visual Line Visual Point Visual Region Visual Surface Visual Solid Visual Glyph Visual Local Spatial Artefact
	Visual Spatial Properties (section 3.2.3)	Visual Position	
		Visual Scale	Visual Depth Visual Length Visual Area Visual Volume
		Visual Orientation	Visual Angle Visual Slope Visual Curvature

Table 3-3 The sections in chapter 3 that describe the concepts of spatial visual metaphors.

3.2.1 Visual Display Space

Spatial visual metaphors rely on the definition of a visual display space. Mapping data to the position or scale of objects in space is a common strategy for spatial visual metaphors. For example, the length of a bar in a `histogram` can be interpreted as some abstract quantity. However, for the user to correctly interpret the scale as a quantity first requires the user to understand how the space has been defined. For example, in a `histogram`, one unit of visual space may correspond to twenty barrels of oil.

There are a number of ways to design the visual display space (figure 3-8). The space can be defined using orthogonal coordinates to define a 1D, 2D or 3D space. The use of an orthogonal 1D visual space is not common and is discussed in section 3.2.1.1. A lot of previous work in Information Visualisation has focused on using an orthogonal 2D visual space and this is discussed section 3.2.1.2. There is an accent in this thesis on using Virtual Environments. Virtual Environments typically use an orthogonal 3D visual space and this is discussed in section 3.2.1.3. A useful design technique in visual displays is to use a distortion of the visual display space and this is discussed in section 3.2.1.4. A common design approach is to subdivide the visual display space and to then tile or nest it with repeated elements. This approach is discussed in section 3.2.1.5.

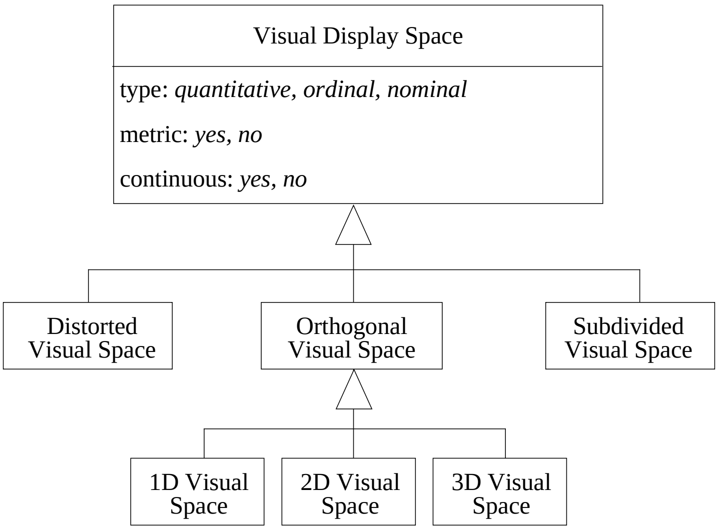


Figure 3-8 A UML diagram showing the components of the visual display space.

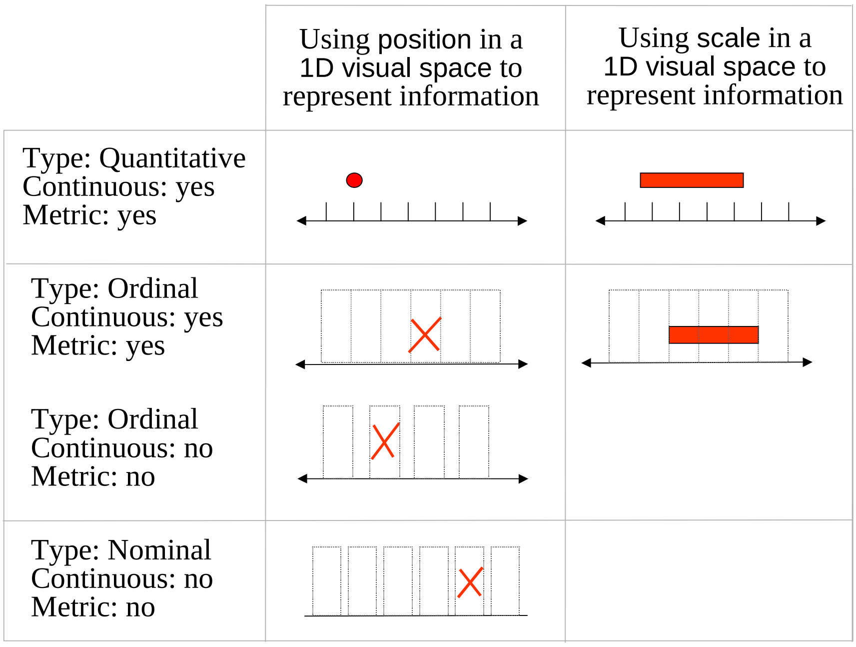


Figure 3-9 With 1D visual space, position or scale can be used to represent information.

3.2.1.1 Orthogonal 1D Visual Space

The 1D visual space is defined by a single axis (x) representing an abstract quantity. Though `1D lineplots` are used to illustrate number concepts in mathematics, using a 1D visual space is not a common technique in Information Visualisation. This is because a single dimension only allows a single data attribute to be represented (figure 3-9).

`Lifelines` is one application that makes use of a 1D visual space to display timelines [Freeman and Fertig 1995]. In this application, time is the abstract quantity represented by a single axis. The space is quantitative and continuous. The items of interest, in this instance, events from a person's life, are arranged, ordered by time, along this axis (figure 3-10). Because a metric is defined it is possible to interpret the distance between items as a measure of time.

An interesting use of ordinal 1D visual space is the work done by Keim and Kriegel [Keim and Kriegel 1994]. Their visualisation is designed to maximise the use of screen pixels. To achieve this, the 1D visual space is folded around itself to form a spiral arrangement (figure 3-11). One application of this technique is to show the results of a multi-dimensional database query. A single pixel is used to represent a single data item that matches the database query. Results are then laid out on a 1D line that forms a spiral. The data items are ordered in this spiral according to their relevance to the query. A similar technique has been used to display results from a document search [Cugini, Laskowski et al. 2000]. Icons representing documents are positioned along a spiral according to a weighting factor calculated by the number of matching keywords (figure 3-12).

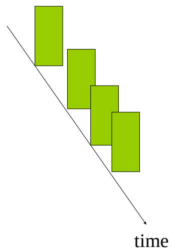


Figure 3-10 The Lifelines application displays a timeline in a 1D visual space [Freeman and Fertig 1995].

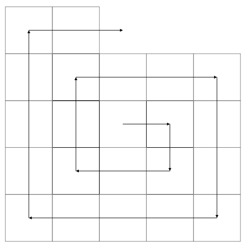


Figure 3-11 This display uses a spiral arrangement of pixels to encode a 1D visual space [Keim and Kriegel 1994].

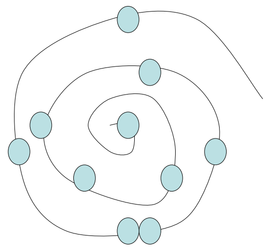


Figure 3-12 The results of a document search, displayed using a spiral layout [Cugini, Laskowski et al. 2000].

3.2.1.2 Orthogonal 2D Visual Space

The 2D visual space is defined by two orthogonal axes (x, y) representing different abstract quantities. This spatial arrangement allows the user to compare different abstract quantities on the x and y axes (figure 3-13) by using the ability of the visual sense to interpret spatial cues.

It is common to define this space by two quantitative axes. This ensures that a metric is defined over the space and allows for the distance between objects to be interpreted. However, it is also possible to use any combination of ordinal, nominal or quantitative axes to define the 2D visual space (figure 3-13). In some of these combinations the concept of a *metric* is not meaningful. For example, sometimes a metric is only defined in one of the two dimensions.

The `Data Map` is a common metaphor that combines cartography and statistics. It was one of the first 2D visual spaces used to show abstract data [Tufte 1983]. A `Data Map` has a fairly obvious real world analogy (figure 3-14). Devised in about 1800 by

William Playfair [Cleveland and McGill 1984], it transfers the idea of a map of real world space to a map of 2D abstract space. This step eventually resulted in the evolution of some useful organisations of abstract 2D visual space. In particular some techniques that have emerged into popular usage include the `2D lineplot`, `scatterplot` and `histogram` (figure 3-15).

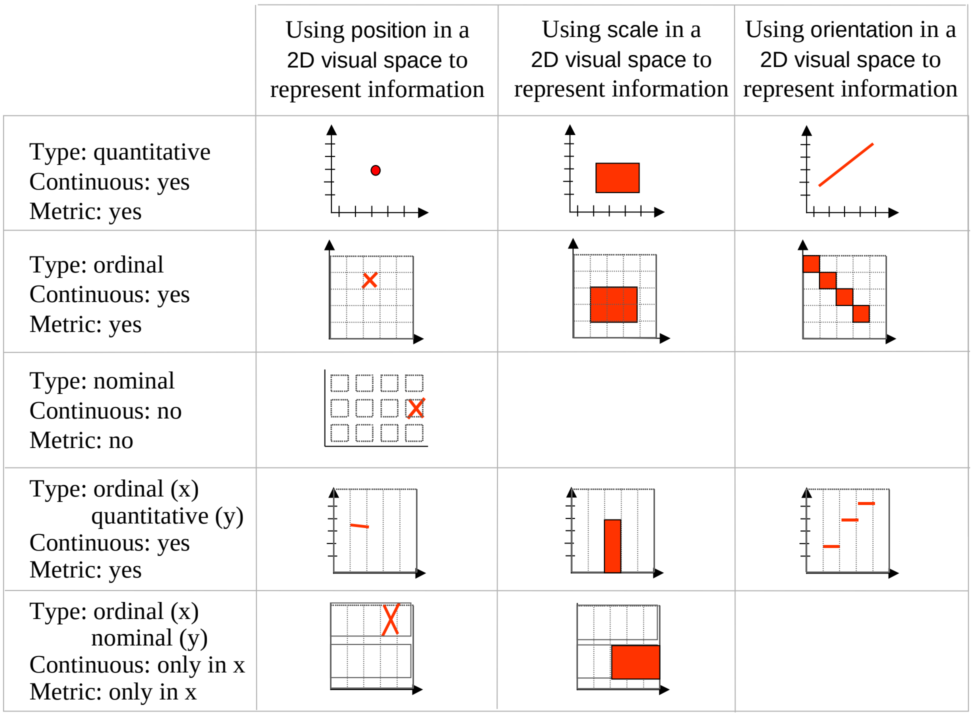


Figure 3-13 With 2D visual space; position, scale or orientation can be used to represent information. The space can be defined by using a combination of quantitative, ordinal or nominal quantities on the x-axis and y-axis.

`Scatterplots` use a 2D visual space that is typically constructed from two quantitative abstract variables. `Scatterplots` are a frequently used visualisation technique in many domains. For example, the applications called, `Spotfire` [Ahlberg and Wistrand 1995] (figure 3-18) and `Filmfinder` (figure 3-19) [Ahlberg and Shneiderman 1994] both use a `scatterplot` technique called a `Starfield Display` to present information.

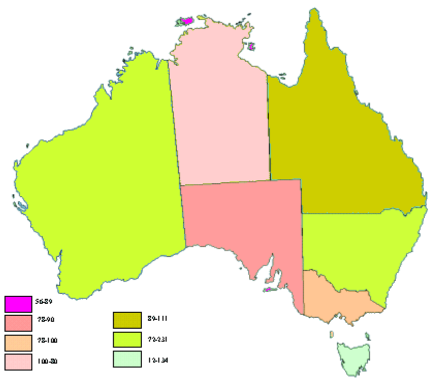


Figure 3-14 The Data Map. Statistical data is overlaid on geography.

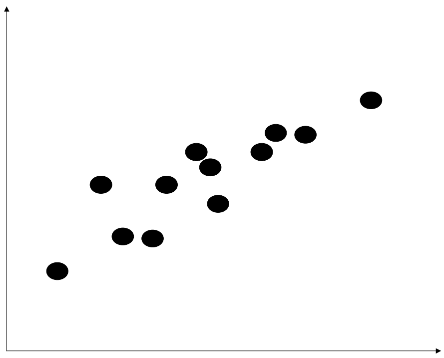


Figure 3-15 The scatterplot.

The 2D lineplot allows for two abstract data attributes to be compared in a 2D visual space [Bertin 1981]. A 2D lineplot typically defines the space in terms of two quantitative variables. This space has a metric and so allows quantitative relationships to be interpreted based on visual distance. For example, relationships between two attributes can be interpreted in terms of the visual scale and visual orientation of plotted lines in space.

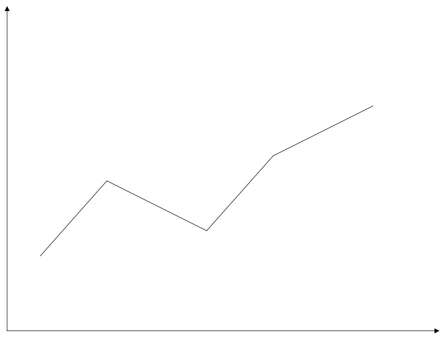


Figure 3-16 The lineplot.

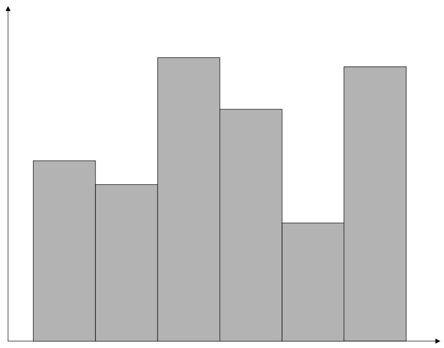


Figure 3-17 The histogram.

A number of different types of histograms can be constructed by using a combination of nominal, ordinal and quantitative variables to define each of the two axes. Commonly, the y-axis represents a quantitative variable, and the x-axis represents nominal categories. A bar marks each category along the x-axis. The scale of the bar in the y-direction then allows a quantity to be interpreted. This is a 2D visual space with a metric defined in the y-direction, but not along the x-direction. Therefore the vertical distance has meaning, but the horizontal distance does not.

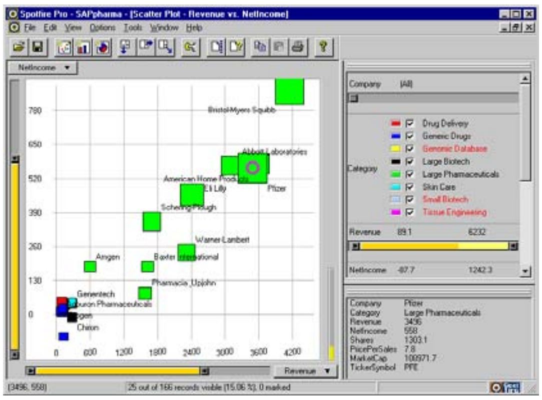


Figure 3-18 The Spotfire application [Ahlberg and Wistrand 1995].
Courtesy of: HCIL, University of Maryland http://www.cs.umd.edu/hcil/spotfire/

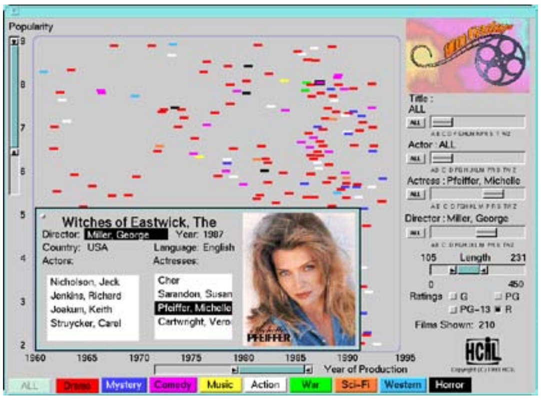


Figure 3-19 The Filmfinder application [Ahlberg and Shneiderman 1994].
Courtesy of: HCIL, University of Maryland http://www.cs.umd.edu/hcil/spotfire/

A variation of the lineplot and histogram is a technique called Parallel Coordinates [Inselberg 1997]. This system tries to transform multi-attributed data relations into useful 2D visual patterns. To define the 2D visual space the y-axis is defined by a quantitative variable. This could, for example, be simply a meaningful range of real numbers. Defining the y-axis is an important step as this is the variable that will link all the other data attributes. To define the x-axis a number of data attributes are arranged along the axis as categories. The data attributes are not usually ordered and so these categories are nominal. To complete the display a line is drawn for each instance of multi-attributed data. This line passes through the points described by the appropriate y-value for each of the x-categories.

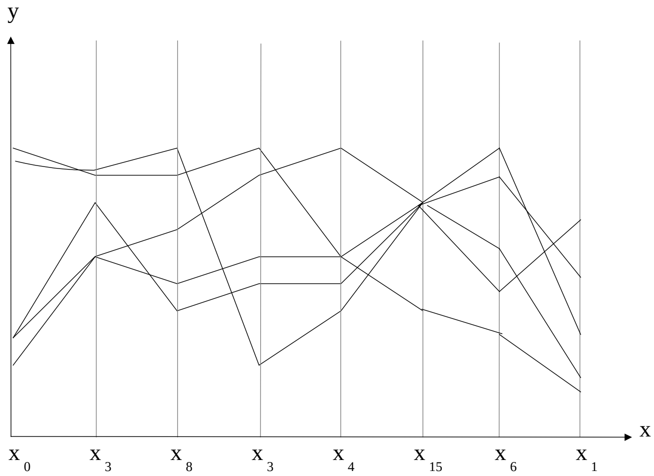


Figure 3-20 Parallel Coordinates [Inselberg 1997].

This 2D visual space used by `Parallel Coordinates` is like the arrangement of space used by some `histograms`. The y-axis defines a quantitative space and the x-axis defines a nominal space. With this arrangement of space the visual distance is only meaningful in the y-direction. One criticism of the `Parallel Coordinate` technique is that the visual orientation of lines may be misinterpreted as significant. However, since the 2D visual space does not have a metric defined in both the x-direction and y-direction, visual orientation is not meaningful. The visual orientation of the lines is simply an accident of the order in which the nominal categories are arranged along the x-axis.

3.2.1.3 Orthogonal 3D Visual Space

A 2D visual space can be extended to make a 3D visual space by adding a third orthogonal axis. Because each of the three axes can now be defined by either a quantitative, nominal or ordinal variable there are many different ways to define a 3D visual space (figure 3-21). Some applications arrange three continuous, quantitative variables on the three axes. With this type of 3D visual space a metric applies, and this allows the use of visual position, visual scale and visual orientation of objects to convey information.

Some general forms of 3D visual space have been described and these include `3D-Scatterplots` (figure 3-22), `3D-Lineplots` (figure 3-23), `3D-Surfaceplots` (figure 3-24) and `3D-Barplots` (figure 3-25).

An example of using a `3D-Surfaceplot` is provided by the `Themescape` application [Wise, Thomas et al. 1995]. This application presents information about a search of text documents using a surface in the 3D visual space (figure 3-26). The attributes that define the search also define the xy-space while the result of the search is interpreted in terms of the height above the xy-space. In another example of 3D visual space, Wright uses a quantitative 3D visual space to design a number of landscapes (figure 3-27). These landscapes use surfaces that are designed to assist in equity trading [Wright 1995].

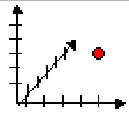
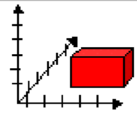
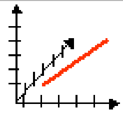
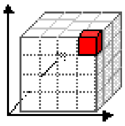
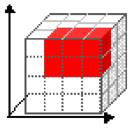
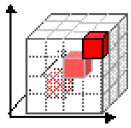
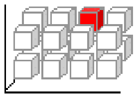
	Using *position* to represent information	Using *scale* to represent information	Using *orientation* to represent information
Quantitative 3D visual space. (continuous).			
Ordinal 3D visual space (continuous).			
Nominal 3D visual space (Non-continuous).			

Figure 3-21 With a 3D visual space, position, scale or orientation can be used to represent information. The space can be defined by using a combination of quantitative, ordinal or nominal quantities on the three axes.

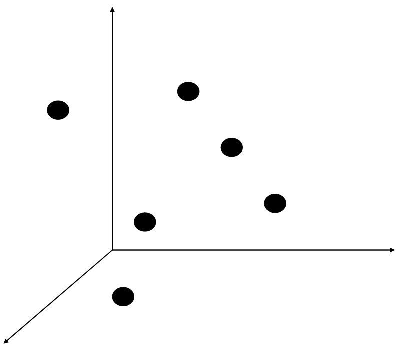


Figure 3-22 The 3D-Scatterplot.

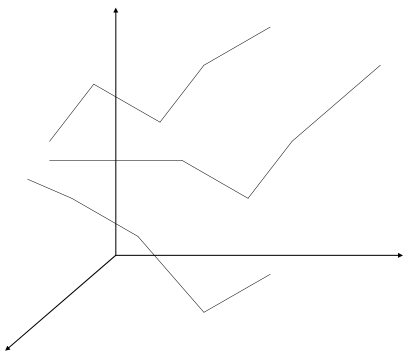


Figure 3-23 The 3D-Lineplot.

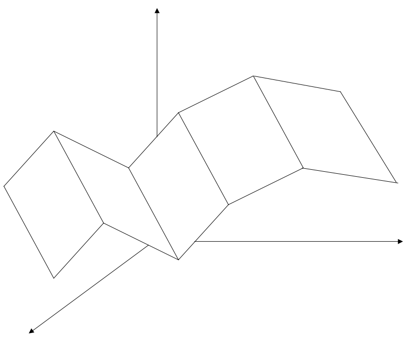


Figure 3-24 The 3D-Surfaceplot.

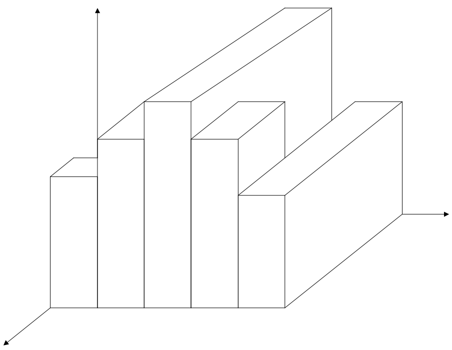


Figure 3-25 The 3D-Barplot.

A more novel example of using 3D visual space is found in the `Cityscape` [Russo Dos Santos, Gros et al. 2000]. This application uses a 3D visual space designed to look like a real world city (figure 3-28). For example, abstract data describing network performance is arranged into buildings (figure 3-29). The three dimensions of the `Cityscape` can be defined by a combination of both ordinal and quantitative variables. So, for example, if the z-axis is defined quantitatively, the height of the buildings can be interpreted. If the xy-space is defined by two ordinal properties then a metric may or may not be defined along the x-axis or y-axis. Normally design focuses on providing a metric for the xy-space as this allows visual scale or visual orientation of objects to be interpreted.

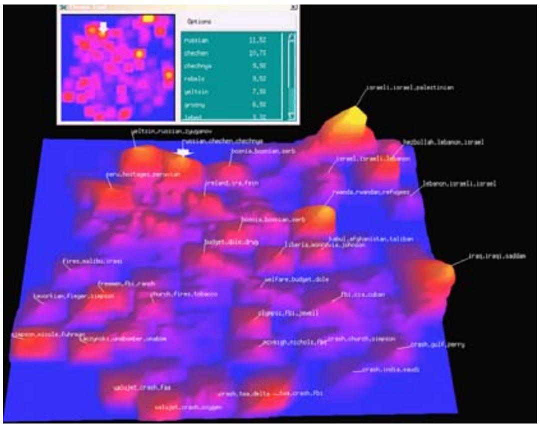


Figure 3-26 The Themescape application [Wise, Thomas et al. 1995].
Source: Pacific Northwest National Laboratory, USA http://www.pnl.gov/infoviz/

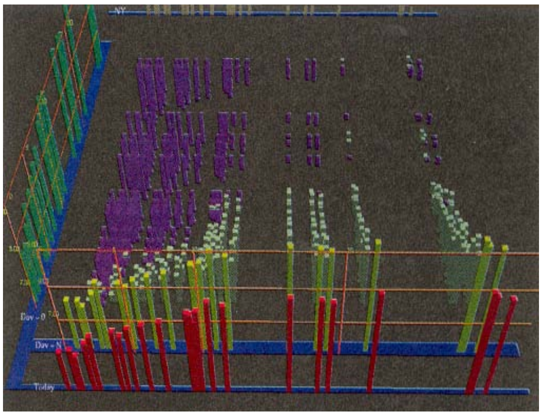


Figure 3-27 A 3D landscape for equity trading [Wright 1995].
Source: IEEE Information Visualization, 1995.

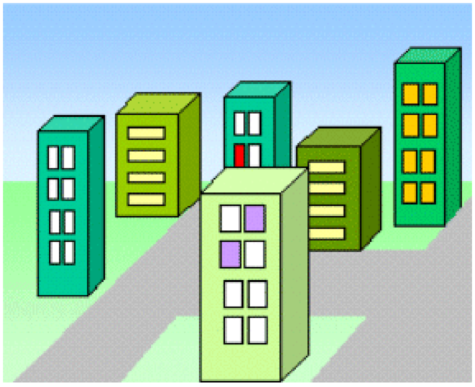


Figure 3-28 The Cityscape [Russo Dos Santos, Gros et al. 2000].

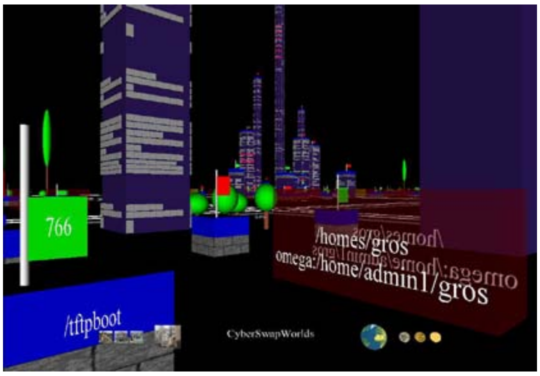


Figure 3-29 Network data as a Cityscape [Russo Dos Santos, Gros et al. 2000].
Courtesy of: Institut Eurécom, France.

3.2.1.4 Distorted Visual Space

Usually the visual space is defined consistently regardless of position in space. With distorted visual spaces the definition of space varies according to position. For example, the `Fisheye View` [Furnas 1981] provides a distorted definition of space.

For orthogonal visual spaces with a metric defined, the interpretation of distance is relatively simple as distances can be compared regardless of the position of objects in the display space. In a distorted visual space distance may need to be interpreted differently depending on the region of the display that an object occupies.

Just as the idea of abstract statistical spaces originated from maps, the idea of a distorted visual space also emerged from the field of cartography. Mercator projections were developed around 1569 to better visualise a map of the globe on a flat surface [Snyder 1993]. This particular distortion of space preserves shape and direction while introducing some size distortion. These size distortions were however acceptable for navigation. Most importantly the distorted visual space maintained relevant structural information such as relative direction and shape.

Distorting the abstract space of a `Data Map` has also proved useful. For example, one problem with statistical data overlaid on maps is that the area of a region can create a false impression as to the importance of the data. The area may be interpreted in a quantitative way although it is only an accident of geography. Large areas tend to dominate the view and take on importance simply because of their size [Tufte 1983]. One solution to this problem involves distorting the spatial map, for example, by relating distance to another quantity such as dollars, thus making area proportional to the abstract quantity under study [Dent 1975].

The distortion technique called `Fisheye Views` [Furnas 1981] aims to maximise the amount of visual display space devoted to the objects of interest. The central area of the display shows more visual detail and is called the *focus* of the display. The surrounding area of the display is less detailed and is described as the *context*.

Importantly, the `Fisheye Views` approach still maintains many visual spatial structures such as connection. It is the visual spatial structures that the user interprets for information. A metric is defined on the space. However, the interpretation of distance is difficult, as it requires some cognitive model of the spatial distortion. This is counter to a user's normal perception where space is uniform.

One specific use of `a Fisheye View` is for drawing `Hyperbolic Trees.` `Hyperbolic Trees` display graphical structures in the fisheye distorted visual space. `Hyperbolic Trees` have been used to display large organisational hierarchies in 2D visual space [Lamping and Rao 1994] (figure 3-30) and also web page connections in 3D visual space [Munzer and Burchard 1995] (figure 3-31). Some general approaches to distorting 2D visual space, such as orthogonal and radial projections, have been extended to 3D visual space [Carpendale, Cowperthwaite et al. 1997].

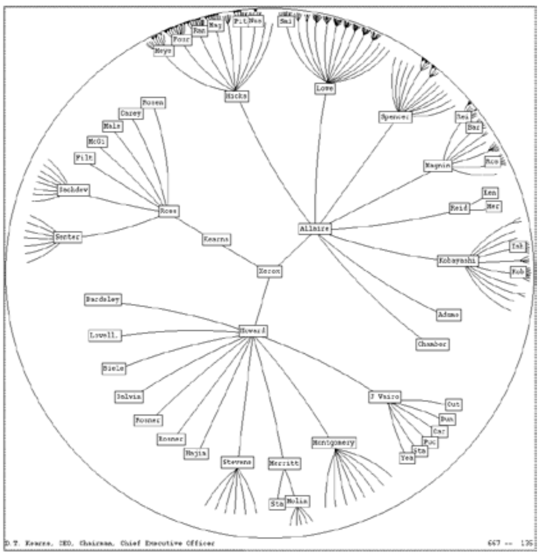


Figure 3-30 Hyperbolic Tree display of a large organisational hierarchy [Lamping and Rao 1994].
Source: http://www.acm.org/sigchi/chi96/proceedings/ videos/ Lamping/hb-video.html

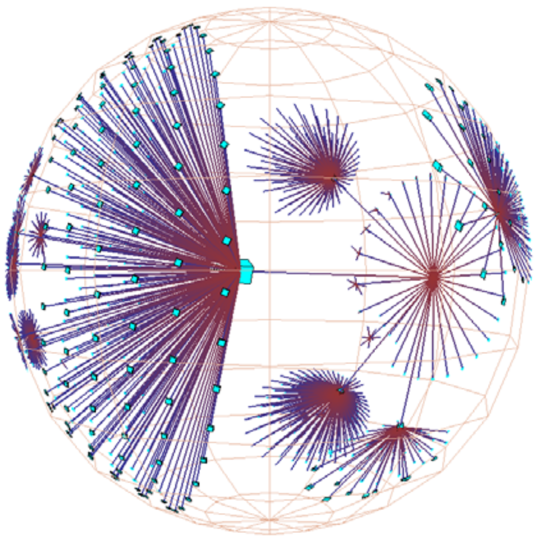


Figure 3-31 A 3D Hyperbolic Tree used to display a file system hierarchy [Munzer and Burchard 1995].
Courtesy of: University of British Columbia http://graphics.stanford.edu/papers/munzner_thesis

Another technique that uses a distorted visual space is the `Perspective Wall` [Mackinlay, Robertson et al. 1991] (figure 3-32). The motivation of this design is also to provide a *focus* and *context* display. Here the distortion of space matches closely with our visual perception. It has a wide field of view and linear perspective is used to that the wall is perceived as receding into the distance.

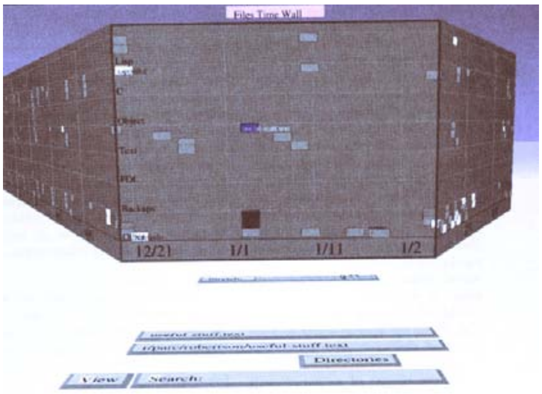


Figure 3-32 The Perspective Wall [Mackinlay, Robertson et al. 1991].
Source: User Interface Research, PARC
http://www.parc.xerox.com/istl/members/mackinlay

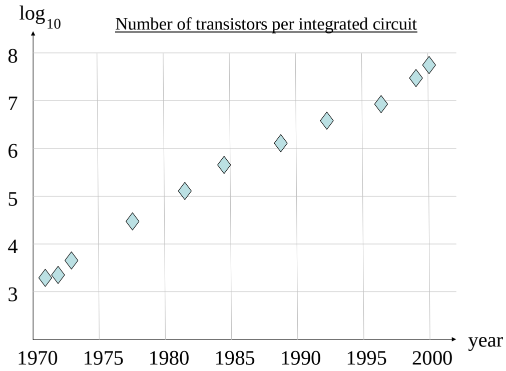


Figure 3-33 Using a logarithmic space to display data.

The concept of a distorted visual space is described in some places as a *view transformation* [Card, Mackinlay et al. 1999]. In this case the distortion is considered to be a lens that the user holds in front of the display. The focus here is on the user interaction rather than the model itself. Simpler distortions such as logarithmic spaces have also been described as *data transformations* [Card, Mackinlay et al. 1999]. The focus in this case is on preliminary processing of the data. However, neither of these approaches is incompatible with the concept of a distorted visual space. What is most important from the design perspective is that the concept of using a distorted visual space is considered during the design.

3.2.1.5 Subdivided Visual Space

Subdividing the space allows an extra dimension of data, such as, a nominal or ordinal data attribute, to be displayed[5]. Multiple copies of a single visual design are constructed using different data that varies in one attribute. These multiples are then arranged in subdivided space for comparison. If the extra data dimension being displayed is ordinal then the multiples are arranged in an ordered fashion.

Bertin describes using this method of subdivided visual space for constructing a series of `scatterplots` or `histograms` [Bertin 1981] (figure 3-36). Cleveland uses this method to construct a matrix of `scatterplots` [Cleveland 1994]. Tufte describes this methodology as tiling the 2D visual space using `small multiples` [Tufte 1990] (figure 3-35) and recommends this as one of the more effective methods for comparing data. The advantage of using repeated design elements is that they provide a familiar context that the user can rely on to make direct comparisons.

5 A subdivided visual space is similar to using nominal or ordinal categories to define the space. However, because it is a useful and commonly used technique, it is considered separately during design.

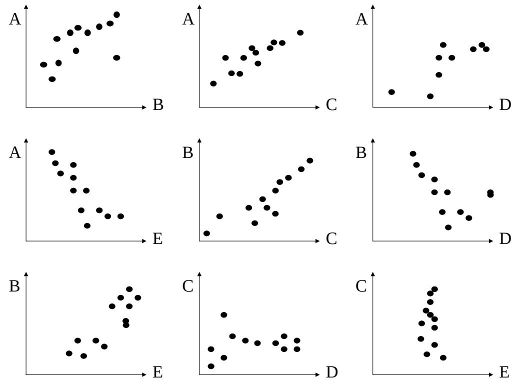


Figure 3-34 A series of scatterplots as described by Bertin [Bertin 1981].

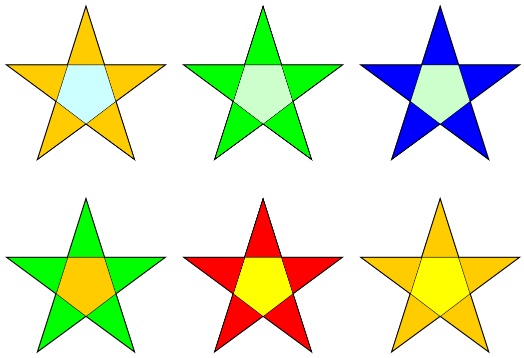


Figure 3-35 The use of, small multiples as described by Tufte [Tufte 1990].

Some designs take a hierarchical approach to defining the subdivided visual space. The technique of `Dimensional Stacking` organises the attribute space into hierarchical 2D visual space that are nested inside each other [LeBlanc, Ward et al. 1990] (figure 3-36). The `Treemaps` technique is designed to fill the available 2D visual space by a hierarchical partitioning of the screen [Shneiderman 1992]. `Treemaps` use the subdivided visual space to capture the structure of the underlying hierarchical data (figure 3-37). A further hierarchical nesting technique places histograms within histograms [Mihalisin, Timlin et al. 1991a]. This nesting is once again based on some hierarchical subdivision of the data (figure 3-38).

Nesting strategies have also been used to subdivide a 3D visual space. The `Worlds-within-Worlds` application uses multiple 3D subspaces nested within each other to display data [Feiner 1990] (figure 3-39). Likewise, the `InfoCube` application visualises hierarchical data as a series of nested 3D boxes [Rekimoto and Green 1993] (figure 3-40).

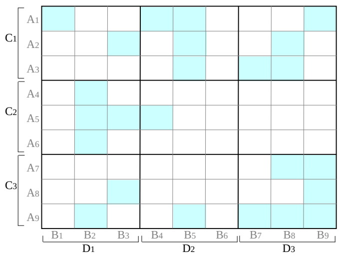


Figure 3-36 The technique of Dimensional Stacking [LeBlanc, Ward et al. 1990].

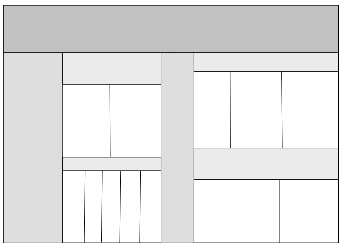


Figure 3-37 The Treemap technique [Shneiderman 1992].

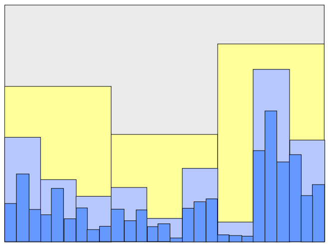


Figure 3-38 Nested Histograms [Mihalisin, Timlin et al. 1991a].

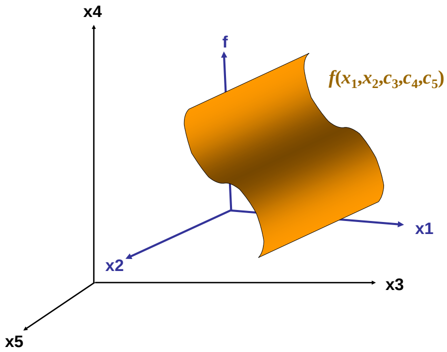


Figure 3-39 The Worlds-within-Worlds application [Feiner 1990].

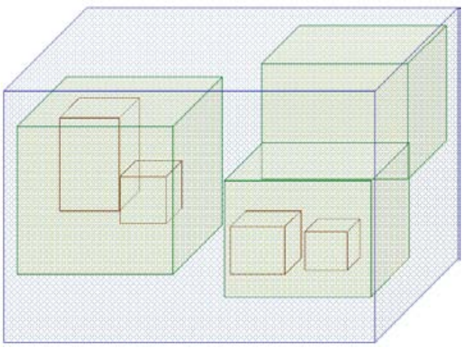


Figure 3-40 The InfoCube application [Rekimoto and Green 1993].

3.2.2 Visual Spatial Structures

Having defined a visual display space, objects can then be created to occupy this space. These objects are described as visual spatial structures. These structures represent the data and are interpreted in terms of the underlying space.

Section 3.2.1 considers the possible ways to define the visual display space. Section 3.2.2 now discusses the types of visual spatial structures that are used to display data within the visual display space. This discussion considers two levels of visual spatial structures. They are global visual spatial structures (section 3.2.2.1) and local visual spatial structures (section 3.2.2.2)[6].

Global visual spatial structure emphasises the relationships between objects. Local visual spatial structure emphasises the relationships within objects. Cleveland emphasises these two levels and notes that *“the power of a graphical display is that it allows us to summarise general behaviour and at the same time to examine details"* [Cleveland 1985].

3.2.2.1 Global Visual Spatial Structures

Global visual spatial structure refers to the types of relationships that occur between objects (figure 3-41). For example, both visual connection and visual containment[7] can imply relationships. The perception of structure can also occur based on the way objects are arranged in the visual display space. For example, objects that form visual groups may be interpreted as related. More complex relationships could be inferred from symmetric arrangements or the recognition of familiar forms such as squares or circles.

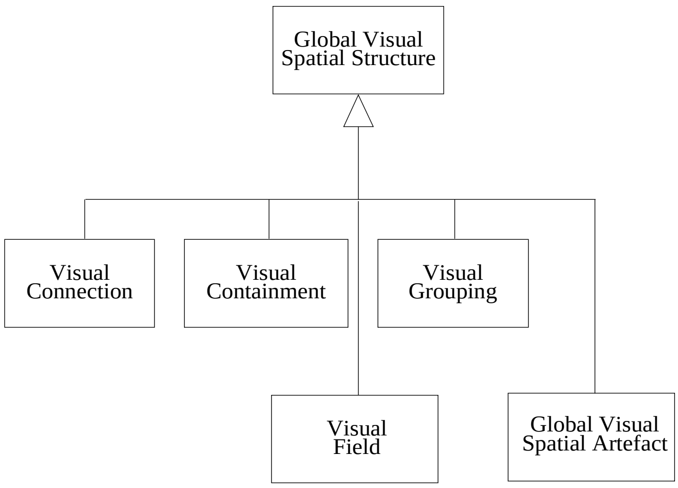


Figure 3-41 A UML diagram showing the types of global visual spatial structures.

---

[6] The division into *global* and *local* forms is somewhat arbitrary. The level of detail may simply depend on the user’s viewpoint. For example, if the user zooms towards a local visual spatial structure it may reveal details that, in the new context, are more *global*. The important point during design is that both *global* and *local* strategies for presenting information are considered.

[7] Visual connection is used in table 3-2 to show the hierarchy of visual spatial metaphors. The same hierarchy is represented by visual containment in table 3-3.

Predicting how global visual spatial structures are perceived by the user is quite complex. The *gestalt principles* [Goldstein 1989] help explain how users perceive the form of visual objects (figure 3-42). These principles include visual properties such as the familiarity (figure 3-43), the connection (figure 3-44), the similarity (figure 3-45), the simplicity (figure 3-46), the proximity (figure 3-47) and the continuity (figure 3-48) of visual elements.

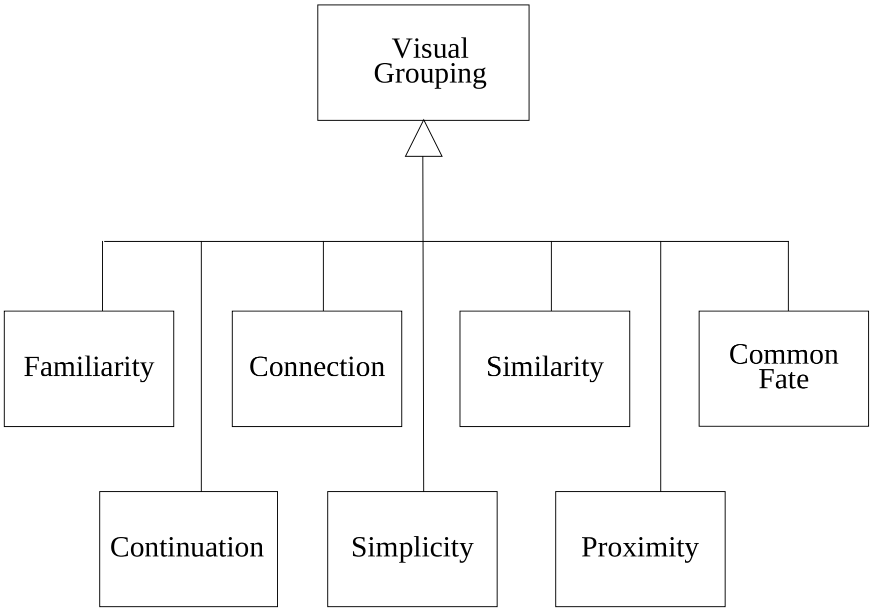


Figure 3-42 A UML diagram showing the types of *visual grouping*. These types are derived from the *gestalt principles*.

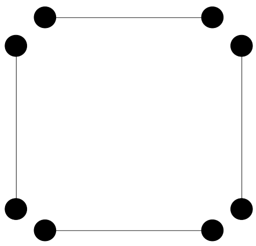


Figure 3-43 The *law of familiarity* states that we perceive structures that appear familiar such as a rectangle in this figure [Goldstein 1989].

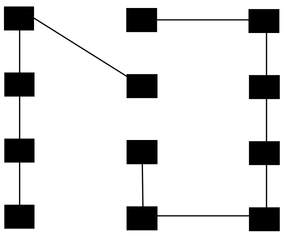


Figure 3-44 The *law of connectedness* states that we perceive connected objects as a single structure. In this figure we see two connected components [Goldstein 1989].

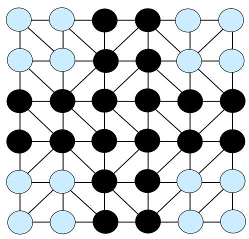


Figure 3-45 The *law of similarity* states that similar things, such as the black nodes in this graph tend to be grouped together [Goldstein 1989].

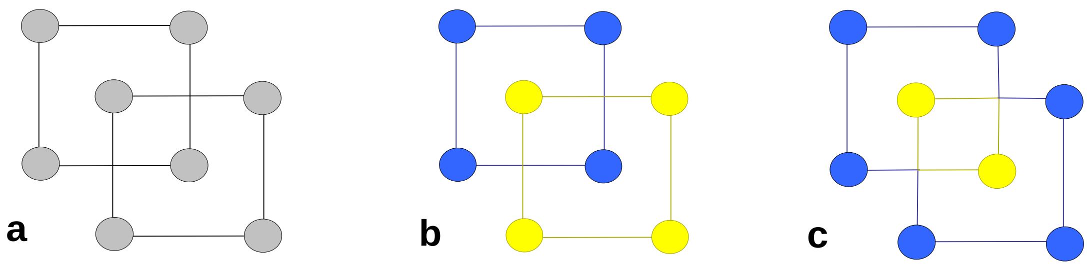


Figure 3-46 The *law of simplicity* states that we perceive an ambiguous structure [a] to be made up of simple shapes [b] and not the more complicated possibilities [c] [Goldstein 1989].

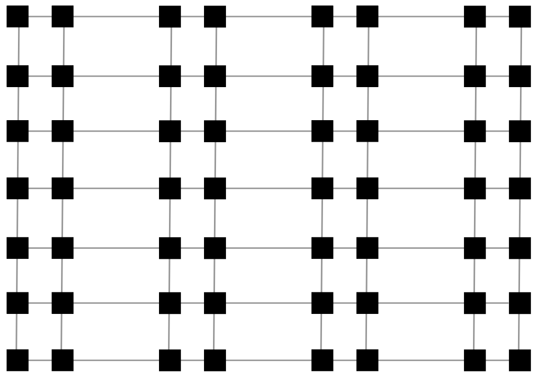


Figure 3-47 *The law of proximity states* that we perceive objects that are close in space to be part of the same structure. Here the spatial arrangement means we normally perceive four groups of paired columns even though it is possible to devise other groupings, for example, by row. [Goldstein 1989].

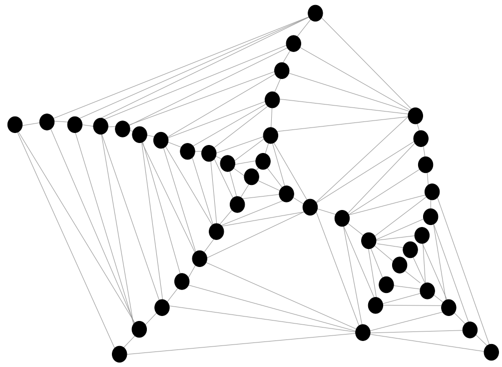


Figure 3-48 The *law of good continuation* states that we perceive smoothly curving lines and straight lines in smooth paths. In this graph we perceive the nodes as part of smooth curves and straight lines [Goldstein 1989].

The perceptual grouping of objects by a user to form global visual spatial structures can often be explained by the gestalt principles. For example, the visual grouping of points in a `scatterplot`, or the continuity of a line in a `lineplot` may be influenced by these principles. The gestalt principle of familiarity implies that a familiar-looking structure[8] is more likely to be perceived in a display.

A graph is a generic structure for describing abstract data. The domain of `Graph Drawing` specialises in producing visual representations of data represented as a graph [diBattista, Eades et al. 1999]. The data entities are typically shown as visual objects called *nodes*. Relationships are then drawn as lines or *edges* between the nodes. `Graph Drawings` are a well-known technique for capturing global visual spatial structure. `Graph Drawings` use visual connection to represent relationships.

`UML diagrams` are a specialised type of `Graph Drawing` used to model relationships in the domain of software engineering [Booch, Rumbaugh et al. 1999]. `UML diagrams` are used throughout this thesis to show links between concepts (for example, see figure 3-41).

In the `SeeNet` application, email traffic is represented by a `Graph Drawing` [Eick and Wills 1993]. In this example, the nodes of the `Graph Drawing` represent staff members and the edges linking the nodes represent the email traffic between different staff members (figure 3-49).

The `Narcissus` program uses a `Graph Drawing` in 3D visual space to display interconnections between sites on the World Wide Web [Hendley, Drew et al. 1995] (figure 3-50). The `Cone Trees` application [Robertson, Mackinlay et al. 1991] implements an interactive `3D Graph Drawing` for displaying hierarchical graphs (figure 3-51).

[8] Familiar-looking structures may also help orientate the user within a display. For example, this is the philosophy that motivates the `Cityscape` application [Russo Dos Santos, Gros et al. 2000] which provides intuitive structural cues to the user by adopting the familiar layout of a city.

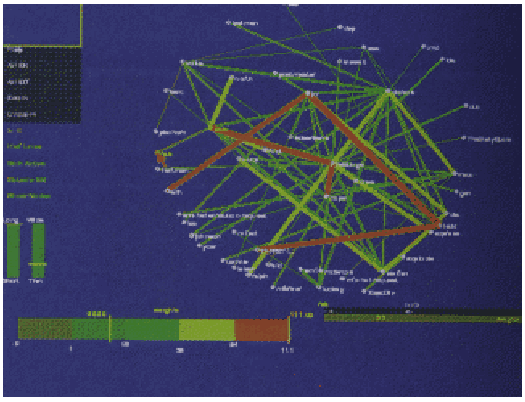


Figure 3-49 The SeeNet application displays email traffic [Eick and Wills 1993].
Source: http://archive.ncsa.uiuc.edu/SCMS/DigLib/text/technology/Visualization-Study-NSFNET-Cox.html

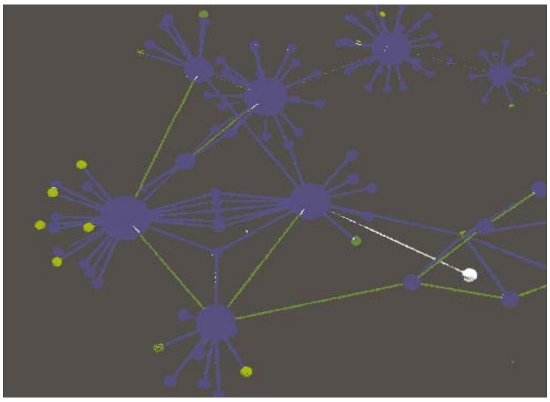


Figure 3-50 Narcissus displays nodes on the web [Hendley, Drew et al. 1995].
Source: [Card, Mackinlay et al. 1999]

For many Graph Drawing applications the visual display space is not defined and the actual visual position of nodes in space does not convey information. One exception to this is when a Graph Drawing uses geographical landmarks as nodes. For example, cities on a map were constrained to their normal geographical location in an application to help display the usage of telephone networks [Becker, Eick et al. 1990b] (figure 3-52). A further example is the use of Graph Drawings to display the route maps of a railway network (figure 3-53). In these displays, the train stations become the nodes and the track routes become the links[9]. In these cases the Graph Drawing is made in an abstract visual display space that distorts the real physical space. Although a metric is not defined on the space, the drawing maintains the useful spatial relationships between the nodes. For example, the ordering of train stations is maintained as is the relative position of the stations to natural landmarks.

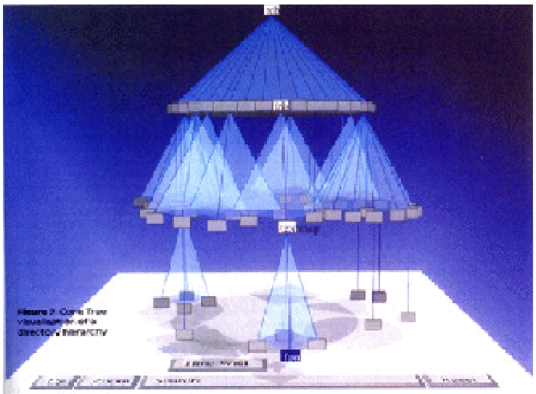


Figure 3-51 The Cone Trees application [Robertson, Mackinlay et al. 1991].
Source: [Card, Mackinlay et al. 1999]

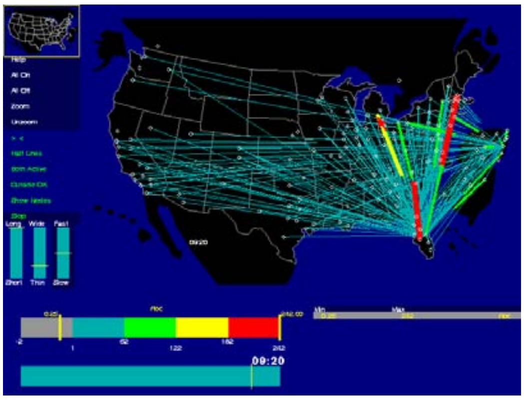


Figure 3-52 Telephone networks displayed on geography [Becker, Eick et al. 1990b].
Source: AT&T Bell Labs. http://cm.bell-labs.com/cm/ms/departments/sia/video-library/seenet.html

Graph Drawings primarily make use of visual connection to display information. However, an alternative to using visual connection is to use visual containment to represent the same relationships. For example, the structure of the MS-Taxonomy can be displayed using visual connection (figure 3-1). The same structure can be displayed using visual containment (table 3-3). The Treemaps technique also

9 In Appendix B this metaphor is used to display the abstract trains of thought connecting this thesis.

uses visual containment to display hierarchical graphs [Shneiderman 1992] (figure 3-37).

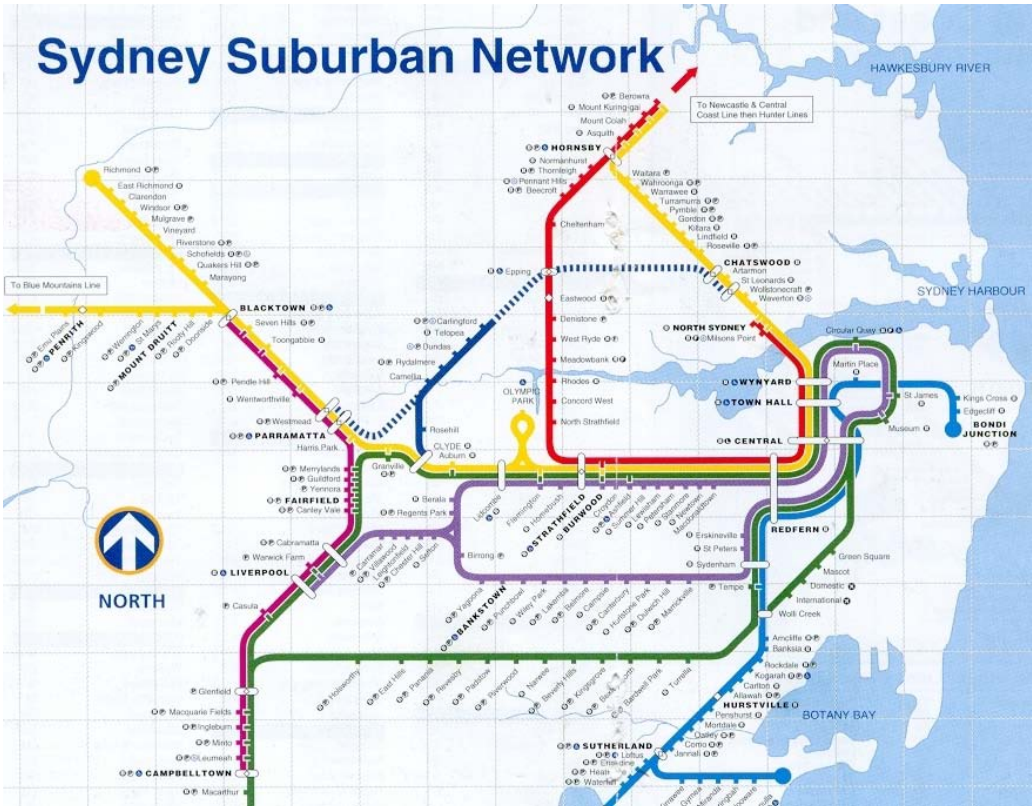


Figure 3-53 The map of the Sydney Rail network. This type of display was first developed by Harry Beck to show the London Underground network [Spence 2001].
Courtesy of: State Rail Authority of NSW.

The concepts of visual grouping, visual connection and visual containment apply to both 2D visual spaces and 3D visual spaces. However, when considering a 3D visual space the further concept of visual surfaces can be considered. Visual surfaces are frequently used in Scientific Visualization. For example, in a meteorological simulation of a cyclone, a 3D visual space contains visual surfaces defined from volumetric data. These visual surfaces display data with equal pressure [Hibbard, Uccellini et al. 1989]. Visual surfaces are also used in Information Visualisation as illustrated by the `Themescape` application [Wise, Thomas et al. 1995] (figure 3-54).

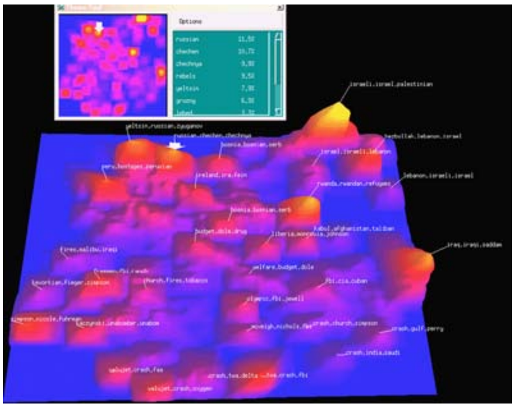


Figure 3-54 Using surfaces to display data in the Themescape application [Wise, Thomas et al. 1995].

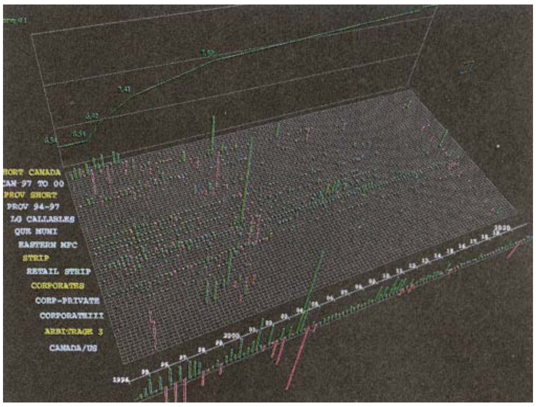


Figure 3-55 Wright uses *visual grids* and *visual axes* to help structure the *visual display space* [Wright 1995].

Source: Pacific Northwest National Laboratory, USA http://www.pnl.gov/infoviz/ **Source:** IEEE Information Visualization, 1995.

The visual field is a spatial structure. The spatial structure like the underlying data varies according to position in the display space. The use of a light field has been suggested as a way to encode abstract information [Healy, Booth et al. 1995].

Another type of global visual spatial structure is the concept of a global visual spatial artefact (figure 3-56). Global visual spatial artefacts do not display abstract data but rather act as aids for interpreting the data. For example, a visual grid is a type of global visual spatial artefact. A visual grid is not part of the data itself but can assist the user to interpret the visual position of other structures in the visual display space (figure 3-57). Visual axes are a commonly used global visual spatial artefact (figure 3-57). The landscapes designed by Wright use both visual grids and visual axes to help structure the visual display space [Wright 1995] (figure 3-55).

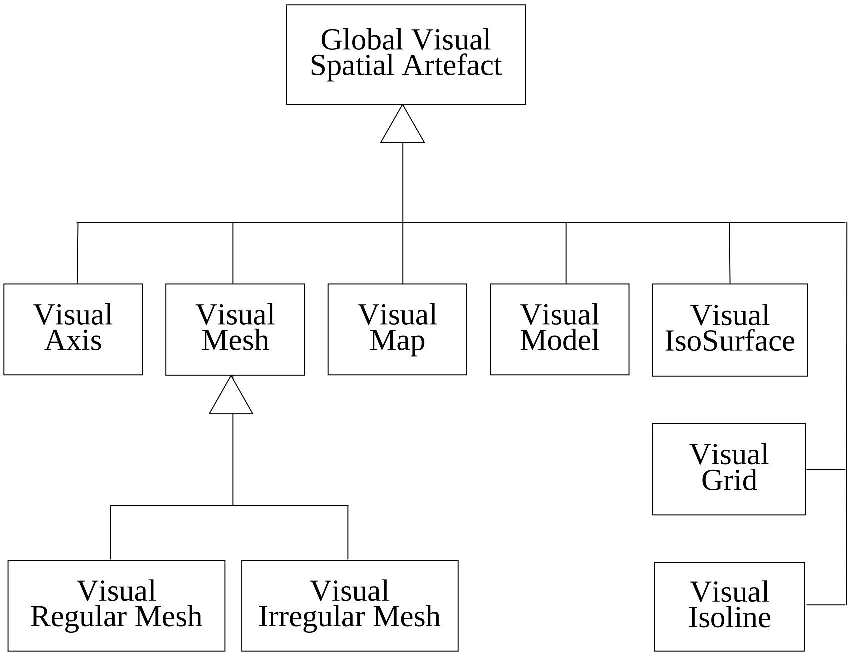


Figure 3-56 A UML diagram showing the types of global visual spatial artefacts.

Another useful global visual spatial artefact is the concept of a visual map. These can be used to overlay the data. `Data Maps`, for example, use a visual map to add structure to the display. Whenever data is collected from a specific longitude and latitude it can naturally be associated with a visual map. In these cases a visual map provides a useful reference for displaying the data. For example, this technique has been used to give context to a computer network [Becker, Eick et al. 1990a] and telephone traffic [Becker, Eick et al. 1990b] (figure 3-52).

Some other types of global visual spatial artefacts such as visual models (figure 3-57) are more frequently used in the domain of Scientific Visualisation. For example, the human body is often used as a model to provide a structural reference for medical data[10]. This technique is also used for displaying results from simulations such as flow

[10] This type of display was described as Virtual Hybrid World in chapter 1 because it overlays abstract data on real world structure.

modelling within a physical structure. Once again the visual model is used as a background to display the simulation results.

Visual meshes are also a common global visual spatial artefact used to display the results from physical simulations. For example, a regular visual mesh or irregular visual mesh may be used to display the geometry used in modelling a fluid dynamics problem (figure 3-57). Other artefacts such as visual isolines and visual isosurfaces[11] can be useful to enhance structure not readily apparent in the data (figure 3-57).

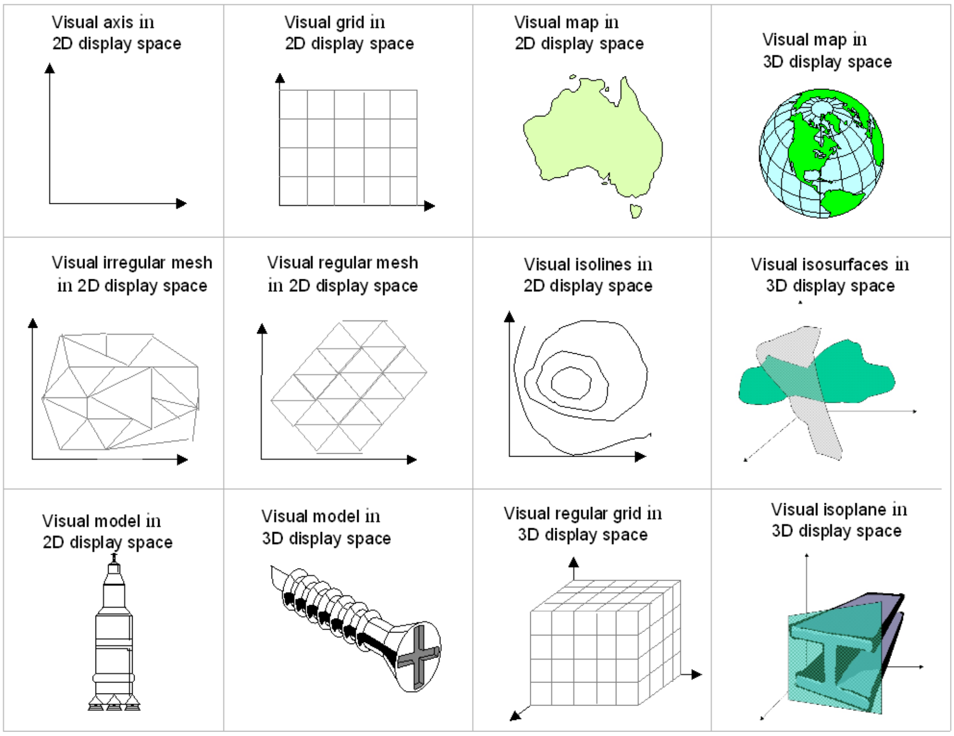


Figure 3-57 Examples of global visual spatial artefacts in 2D and 3D visual display space.

3.2.2.2 Local Visual Spatial Structures

Local visual spatial structure describes to atomic data objects or the types of relationships that occur within objects (figure 3-58). For example, shape details or relations between the sub-components of an object may convey information. Glyphs are an example of using local structure to provide an additional level of detail.

Local visual spatial structures typically represent instances of data in the visual display space. For example, in a `scatterplot`, visual points are used to mark occurrences of data at a visual position in the abstract space. Visual lines are used in `lineplots` to represent a relationship between the abstract variables represented on the visual axes. Visual regions are used to represent data in a `histogram`. In 3D visual space the concept of a visual region becomes a visual surface. In a `3D-surfaceplot`, a visual surface is used to represent relationships between three abstract variables. In a

[11] Concepts such as the visual mesh and isosurface are more familiar to the Scientific Visualisation community. However, the concepts of the MS-Taxonomy are very general and so apply to applications in this domain.

3D-Barplot a visual solid shaped like a bar is used to represent relationships in a 3D visual space.

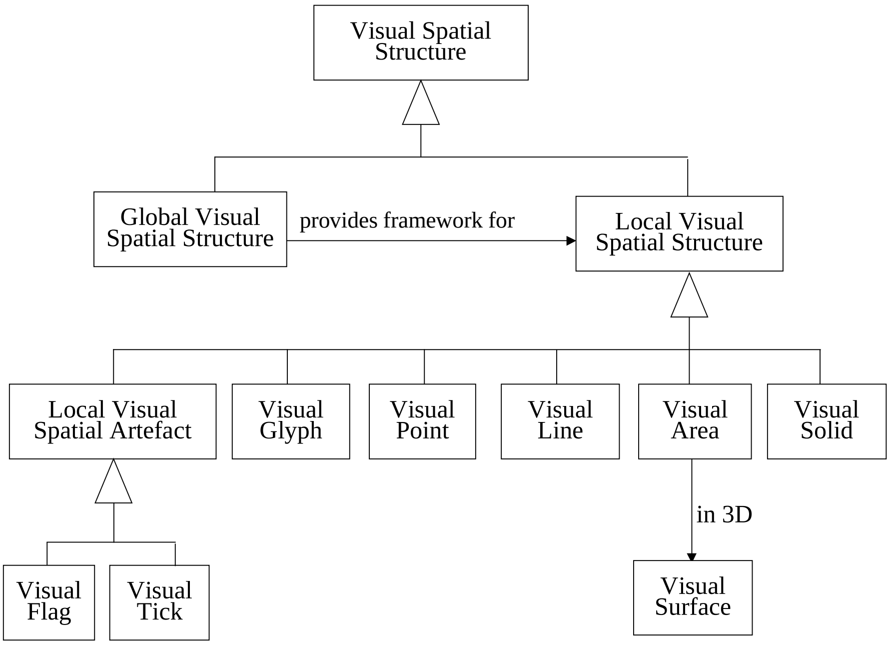


Figure 3-58 A UML diagram showing the types of local visual spatial structure.

In general any visual shape could be used to represent data. At this point the MS-Taxonomy distinguishes between two different uses of shape. When shape is used to encode a nominal category it is included under direct visual metaphors. This type of visual metaphor uses a limited set of shapes that act as visual properties. The shape property is directly interpreted to represent a category and is referred to in the MS-Taxonomy as direct visual shape (section 4.2.4).

A second use of shape is to directly encode data relationships in the local structure of the shape. This type of local visual spatial structure is referred to as a visual glyph. For example, Chernoff devised a visual glyph that used the relationships between elements of the human face to encode data attributes [Chernoff 1973]. These, so called `Chernoff Faces` replaced the points in a normal `scatterplot` with facial icons (figure 3-59). Another example of visual glyphs is called *stick figures* (figure 3-60). In this case the data attributes are mapped to the angle and length of limbs. Positioned in a 2D visual space the *stick figures* create textural patterns that can highlight characteristics of the data [Pickett and Grinstein 1988]. In a further example, the `Driftweed` method displays a time series of multivariate data as a collection of joined line segments, where each segment represents a data attribute [Rose and Wong 2000] (figure 3-61).

The concept of the visual glyph has also been extended to the 3D visual space in a design called the `3D Wheel` [Chuah and Eick 1997]. The `3D Wheel` was used to encode data attributes about software development (figure 3-62). The technique is designed to distinguish different types of trends by the tapering or sharpening point of the glyphs.

A display may also be augmented with local visual spatial artefacts. Local visual spatial artefacts are a particular type of local visual spatial structure that can assist with reading the display. Local visual spatial artefacts do not directly represent data. For example, a visual tick is a common local visual spatial artefact used to augment a visual axis. A visual tick helps to identify a specific visual position in the visual display space. More directly related to the data itself is the concept of a visual flag. This can be used to pinpoint a particular instance of data in the display. In the landscapes designed by Wright visual flags, called *signposts* are used to help mark visual positions in the 3D visual space [Wright 1995].

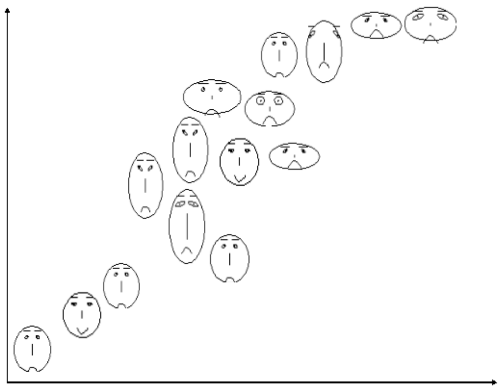


Figure 3-59 Using Chernoff Faces to represent points on a scatterplot [Chernoff 1973].

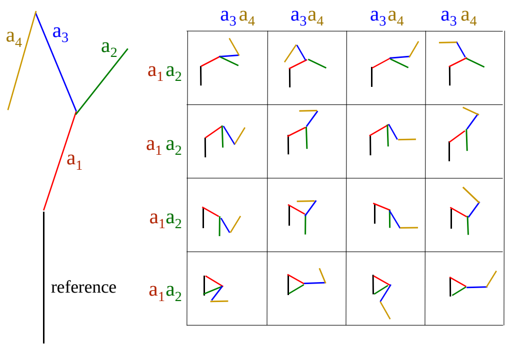


Figure 3-60 Stick figures use the angle and length of limbs to encode information [Pickett and Grinstein 1988].

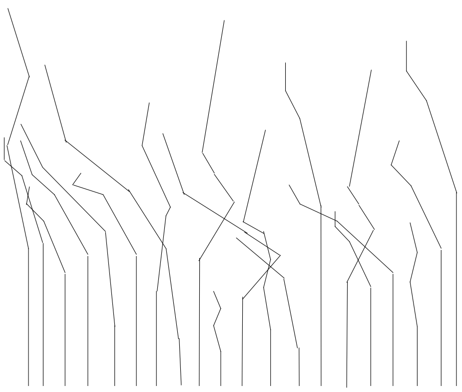


Figure 3-61 The Driftweed method [Rose and Wong 2000].

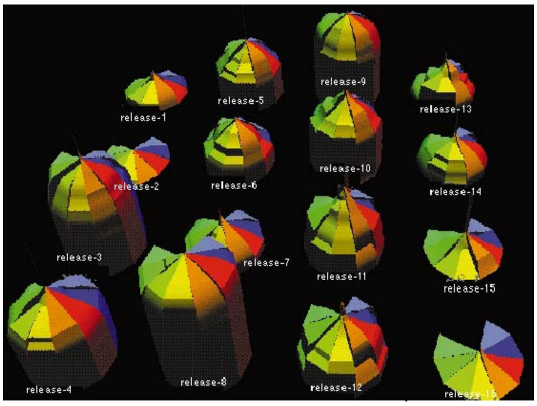


Figure 3-62 Visual glyphs in 3D visual space [Chuah and Eick 1997].
Source: IEEE Computer Graphics and Applications, July/August 1998

3.2.3 Visual Spatial Properties

Section 3.2.1 considered the possible ways to define the visual display space. Section 3.2.2 discussed the types of visual spatial structures that are used to display data within the visual display space. Visual spatial structures have properties that are defined in terms of the visual display space and these properties are described as visual spatial properties (figure 3-63). The user interprets visual spatial properties in terms of the visual display space. Visual spatial properties include:

- visual position
- visual scale
- visual orientation.

The use of visual position to display information is very common and has already been discussed in section 3.2.1 (see figure 3-13). For example, `Scatterplots` use the visual position of points to convey information (figure 3-22).

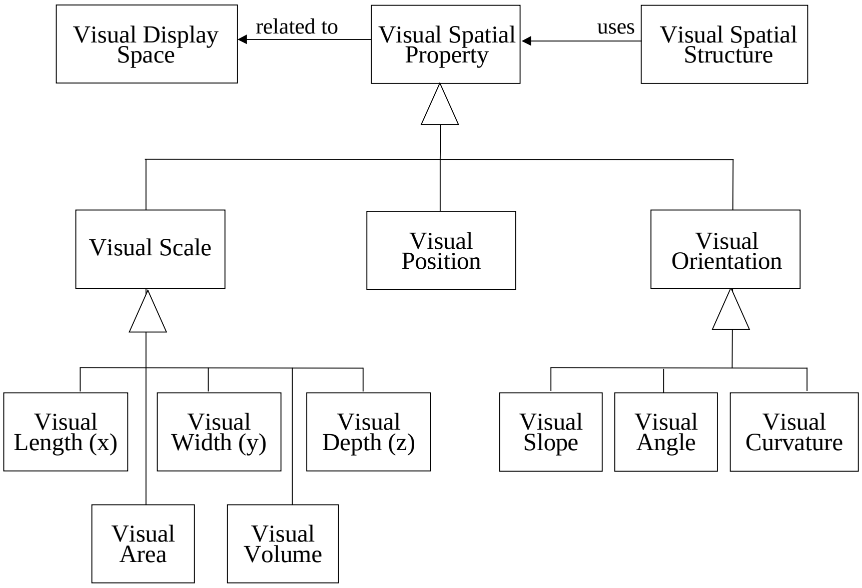


Figure 3-63 A UML diagram showing the types of visual spatial properties.

There are different types of visual scale and these include the linear properties of visual length, visual width and visual height. These linear properties can be used to represent quantitative data in the visual display space. In a 2D visual space the visual area may also be used to represent data. `Histograms` use the visual scale of bars to convey information (figure 3-17) and `pie charts` (figure 3-65) use visual area to convey information.

Visual orientation is frequently used in `lineplots` (figure 3-22) and `3D-surfaceplots` (figure 3-23) to convey information. There are different types of visual orientation and these include visual slope, visual curvature and visual angle. For example, both `Rose diagrams` (figure 3-64) and `pie charts` (figure 3-65) contain information that can be interpreted from visual angles in the display.

Note that a metric must be defined on the space when the distance between objects is used to convey information. In some cases a metric is not defined in all the visual dimensions. This restricts visual scale to only being interpreted in the dimensions with a metric. Some displays, such as those that use visual orientation to convey information must have a metric defined in all appropriate dimensions.

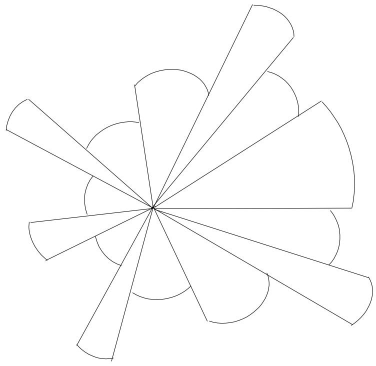


Figure 3-64 The Rose diagram.

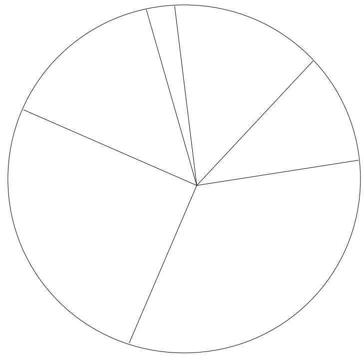


Figure 3-65 The pie chart.

3.2.4 A Summary of Spatial Visual Metaphors

The previous sections have reviewed the concepts that describe spatial visual metaphors. The important high-level components of spatial visual metaphors are the visual display space (section 3.2.1), spatial visual structures (section 3.2.2) and spatial visual properties (section 3.2.3). These components are an intuitive specialisation of the more general concepts of display space, spatial structure and spatial properties that describe a spatial metaphor. This is not surprising as vision is a spatial sense and of all the senses it is most adept at identifying spatial relationships [Friedes 1974].

The visual display space is continuous and can be designed to display quantitative, ordinal or nominal data. Visual spatial structure can be interpreted on both a global level and local level. Visual position is a very accurate way of representing data. as it can be accurately interpreted. Provided a metric is defined other visual spatial properties that rely on interpreting the distance between points are also accurate methods for displaying data. For example this includes visual scale and visual orientation.

The previous sections provided a number of examples of generic designs[12]. It is important that new design strategies consider these existing spatial visual metaphors. Existing designs provide an opportunity for reuse and offer a pragmatic approach to engineering solutions in new domains. An advantage of using established designs is that the user may already have experience with using that particular spatial visual metaphor.

A number of generic designs for 1D visual space are summarised in figure 3-66. Some generic designs for 2D visual space are captured in figure 3-67 and some generic designs for 3D visual space are shown in figure 3-68.

12 Generic information metaphors are also included in Appendix A.

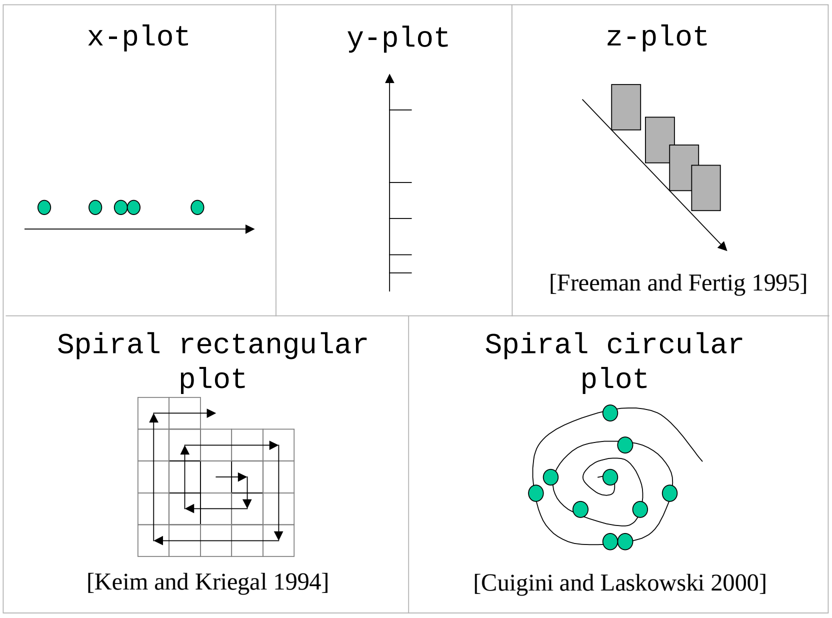


Figure 3-66 Generic designs using 1D visual space.

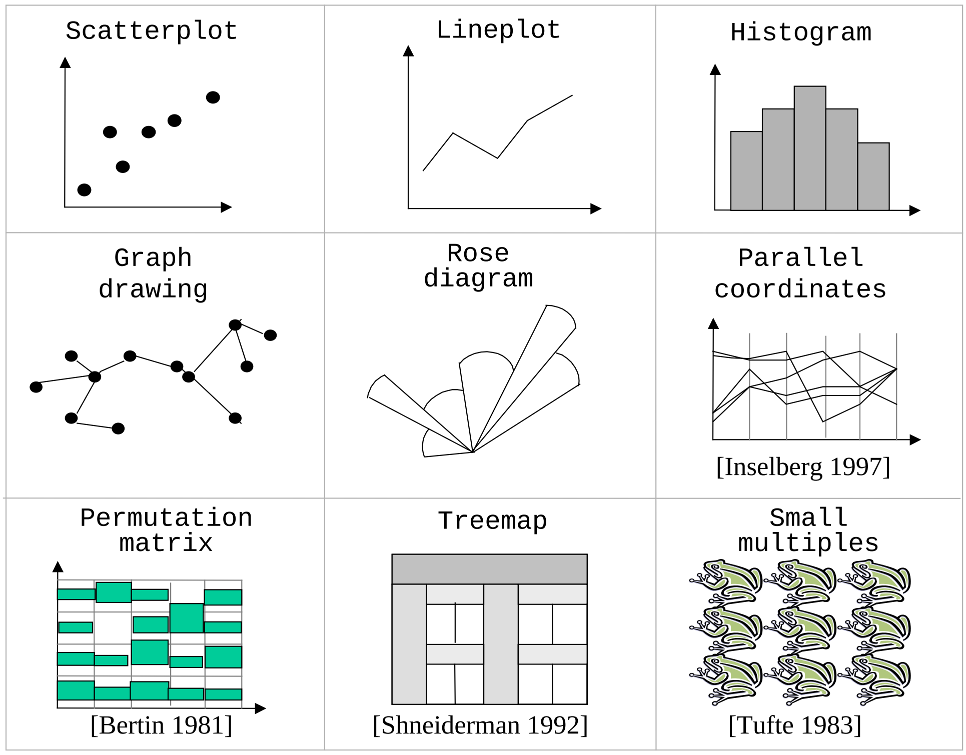


Figure 3-67 Generic designs using 2D visual space.

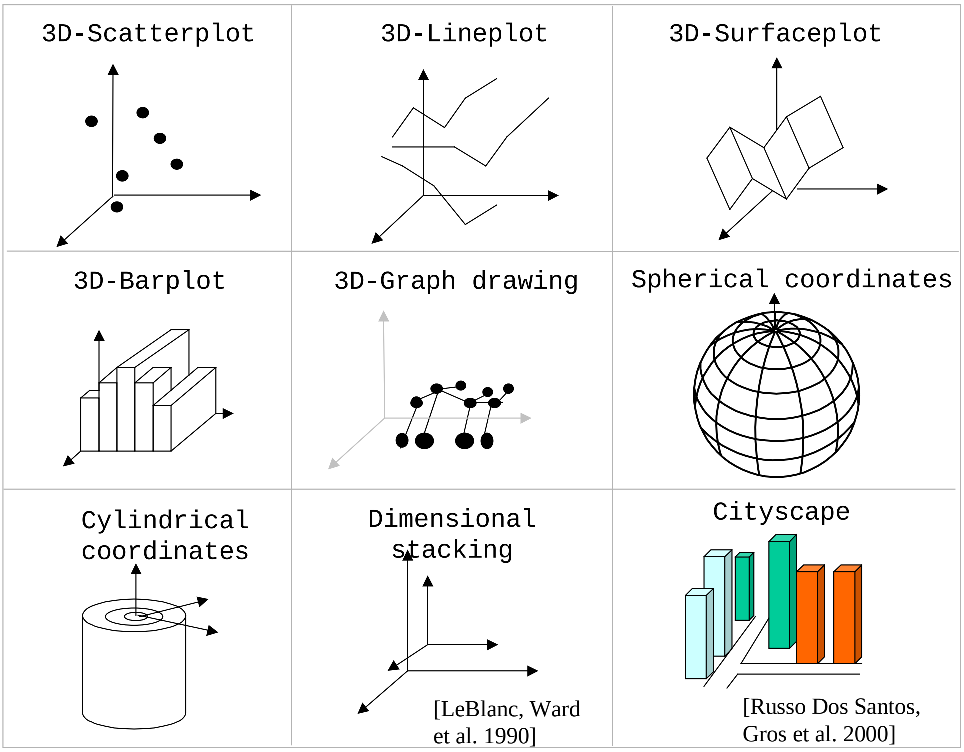
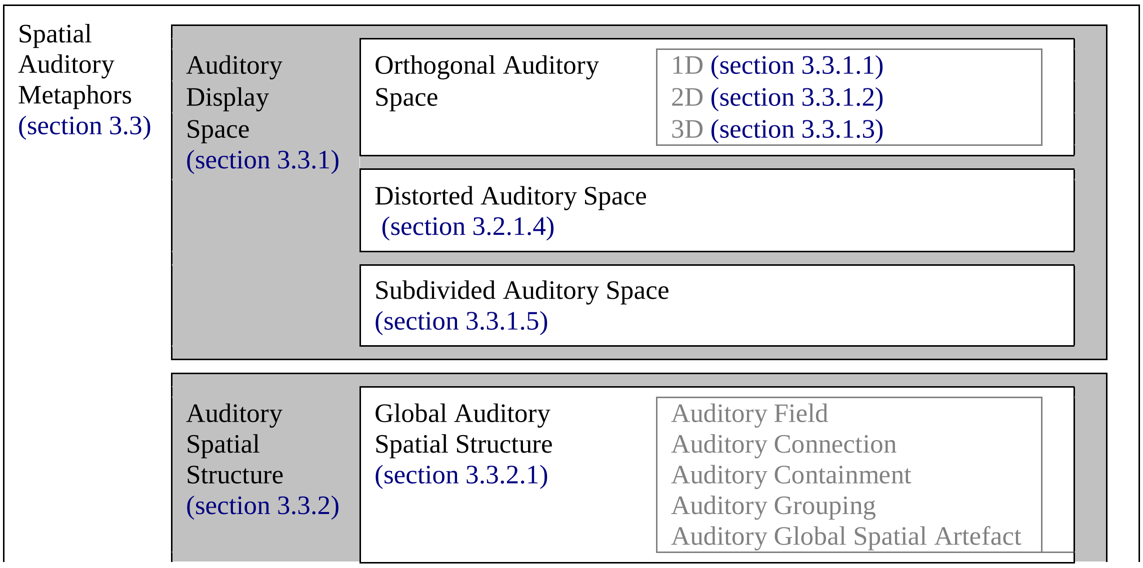


Figure 3-68 Generic designs using 3D visual space.

3.3 Spatial Auditory Metaphors

This section describes the concepts of spatial metaphors as they apply to the auditory sense (table 3-5). Spatial auditory metaphors relate to the auditory perception of the scale of sounds in space, the position of sounds in space and the structure of sounds in space. The important considerations in the design of spatial auditory metaphors are (figure 3-69):

- the auditory display space (section 3.3.1)
- the auditory spatial structure (section 3.3.2)
- the auditory spatial properties (section 3.3.3).



	Local Auditory Spatial Structure (section 3.3.2.1)	Auditory Line Auditory Point Auditory Region Auditory Surface Auditory Solid Auditory Glyph Auditory Local Spatial Artefact
Auditory Spatial Properties (section 3.3.3)	Auditory Position	
	Auditory Scale	Auditory Depth Auditory Length Auditory Area Auditory Volume
	Auditory Orientation	Auditory Angle Auditory Slope Auditory Curvature

Table 3-4 The sections in chapter 3 that describe the concepts of spatial auditory metaphors.

The following sections also review some previous work in Information Sonification in terms of auditory display space, auditory spatial structure and auditory spatial properties. Section 3.2.4 summarises the important points of designing spatial auditory metaphors.

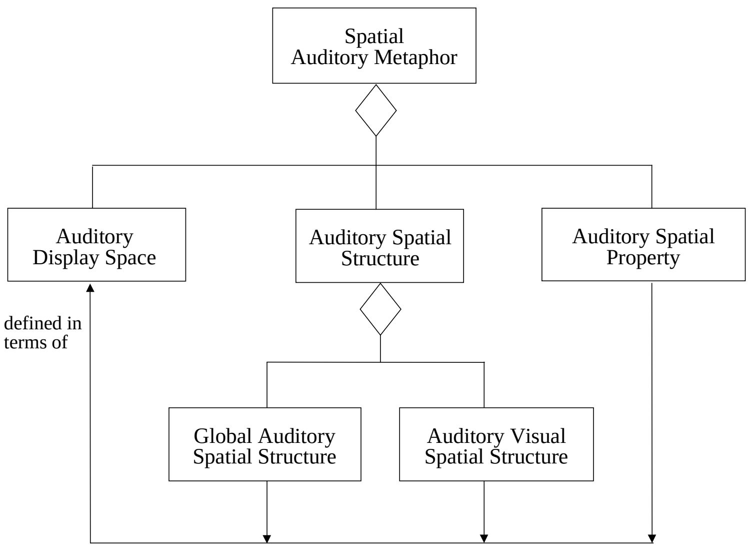


Figure 3-69 A UML diagram showing the components of the spatial auditory metaphors. Structure and properties are defined in terms of the auditory display space.

3.3.1 Auditory Display Space

Spatial auditory metaphors rely on the definition of an auditory display space. Mapping data to the position of sounds in space is a possible technique for displaying information. For example, an `auditory scatterplot` would use the position of sounds in space to mark points of abstract data. However, to correctly interpret the abstract data requires the user to understand how the space has been defined.

There are a number of ways to define the auditory display space (figure 3-70). The space can be defined using orthogonal coordinates to define a 1D, 2D or 3D space. The use of an orthogonal 1D auditory space is a reasonably common approach and is discussed in section 3.3.1.1. The use of an orthogonal 2D auditory space is discussed in section 3.3.1.2 and orthogonal 3D auditory space is discussed in section 3.2.1.3. Two useful design techniques used in visual displays are the use of a distorted space and a subdivided space. These approaches can also be applied to the auditory display space. Distorted auditory space is discussed in section 3.2.1.4 and subdivided auditory space is discussed in section 3.2.1.5.

While a variety of spatial visual metaphors have been used to display abstract data, there are fewer examples of where spatial auditory metaphors have been used. Indeed the use of spatial auditory metaphors for displaying abstract data has not been well explored. This is perhaps not surprising as vision can be thought of as a spatial sense, while hearing is principally a temporal sense.

Many applications that use sound to display abstract data simply generate sounds that have no spatial properties. However, what is most important for design purposes is that the ability of the auditory sense to resolve spatial cues is considered. For example, it should be noted that the auditory sense has limited spatial precision compared to the visual sense (see section 2.3.3). Despite this lower resolution the auditory display space should still be considered for design.

It is possible to define a continuous, quantitative auditory display space. This space would have a metric defined so that the distance between two sounds is meaningful to the listener. While this is possible, a more appropriate to use of auditory display space would be based around continuous ordinal categories. This type of mapping provides a better match to the lower spatial resolution of the auditory sense and can still have a metric defined. Another alternative is to subdivide the auditory display space into nominal categories. This definition of the space does not relying on the user being able to interpret a metric.

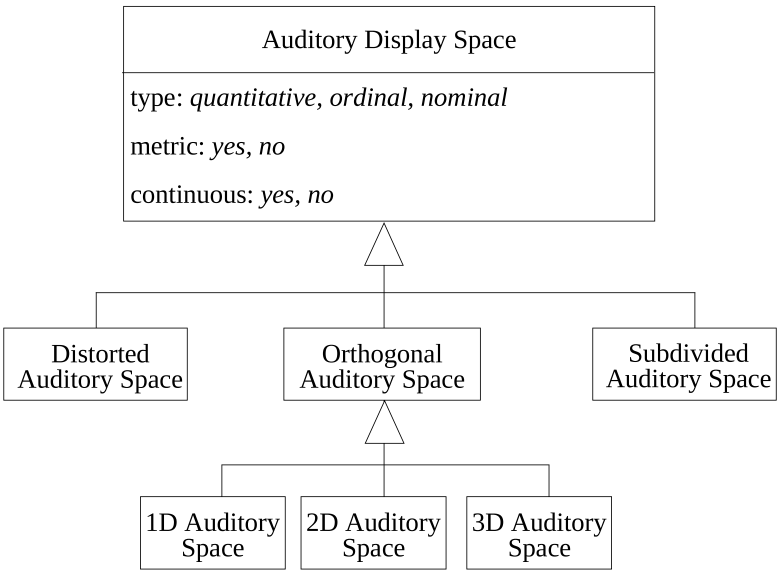


Figure 3-70 A UML diagram showing the components of the auditory display space.

3.3.1.1 Orthogonal 1D Auditory Space

The 1D auditory space is defined by a single axis (x) representing an abstract quantity. An orthogonal 1D auditory space is analogous to a line of sound. For example, a typical stereo display uses the balance of sound volume between the left and right speakers to display the sound along a continuum between the speakers. This technique is also called *panning* [Kramer 1994a]. The use of 1D auditory space is the most common technique for defining the auditory display space.

Madhyastha and Reed developed an auditory display of geographical information [Madhyastha and Reed 1994]. In this application, multi-attributed data was displayed for a number of cities with the longitude of the city mapped to the stereo balance. For example, cities with higher longitudes are heard to the right. This display uses a continuous quantitative 1D auditory space where auditory position encodes the data attribute of longitude. A similar definition of the 1D auditory space was used to display attributes of data recorded from chemistry experiments [Yeung 1980].

A very simple definition of the 1D auditory space is used in an application to look for communication patterns on a parallel computer [Madhyastha and Reed 1994]. The 1D visual space is divided into two nominal categories. The nominal categories in this case are the *send* or *receive* messages from the parallel processors. For example, *send* messages are played from the left and *receive* messages from the right.

3.3.1.2 Orthogonal 2D Auditory Space

The 2D auditory space is defined by two orthogonal axes (x, y) representing different abstract quantities. This spatial arrangement allows the user to compare different abstract quantities on the x and y axes by using the ability of the auditory sense to interpret spatial cues.

An orthogonal 2D auditory space extends the orthogonal 1D space by a further dimension. This is analogous to a plane of sound. To generate a 2D auditory space requires complex hardware, to create a field of sound, or complex software to integrate perceptual cues into the sound [Zacharov and Koivuniemi 2001] (see section 1.2.4). The added complexity of generating 2D auditory display spaces probably accounts for the relatively small number of applications that use such spaces.

In one application, Fernström and McNamara use a 2D auditory space to present nominal categories to the user [Fernström and McNamara 1998]. This auditory display is designed to assist browsing through a collection of tunes and searching for a target tune. The multiple sound streams, each with a different tune are positioned as categories in the 2D auditory space.

A similar use of 2D auditory space to present nominal categories is demonstrated by the application called `AudioStreamer`. The user can browse a selection of audio-news broadcasts simultaneously. Each particular news item is placed at disjoint auditory positions. The listener's head movement towards any of the sources is detected and results in an increase in the relative gain of the source [Schmandt and Mullins 1995].

A more complex use of 2D auditory space was developed for use on a wearable audio display called `Nomadic Radio` [Sawhney and Schmandt 1997]. This application uses the 2D auditory space to present dynamic message information. `Nomadic Radio` can map attributes of the message, such as, time of arrival onto the auditory position. For example, the message arrival time could be used to position messages in chronological order on a spatial clock laid out on a plane around the user's head. This defines the 2D auditory space in terms of 24 ordered categories, one for each hour of the day.

3.3.1.3 Orthogonal 3D Auditory Space

An orthogonal 3D auditory space extends the orthogonal 2D space by a further dimension. This is analogous to a volume of sound. Once again the added complexity of generating 3D auditory display spaces probably accounts for the relatively small number of applications that use such spaces. Although the use of a 3D auditory space sound for displaying abstract data has been accelerated by the explicit inclusion of 3D sound displays in many Virtual Environments such as the CAVE [Kaper, Tipei et al. 1999] and CyberStage [Eckel 1998].

A quantitative definition of the 3D auditory space was used to represent aircraft traffic in a `Traffic Collision Avoidance System` [Begault and Wenzel 1992]. In this system the perceived spatial location of the spoken word, “traffic” is analogous to the real-world location of the aircraft.

A nominal division of the 3D auditory space is used in an application that allows blind users to browse the internet [Roth, Petrucci et al. 1998]. An auditory element is simulated at an auditory position determined by the location of the corresponding visual element in the Web browser window. In this case the space has no metric defined, but rather is used to simply organise the auditory elements and convey structural information to the user.

3.3.1.4 Distorted Auditory Space

A distorted auditory space is not defined in a uniform fashion but rather the definition of space varies according to position. For example, in the visual domain the `Fisheye View` [Furnas 1981] provides a distorted definition of the visual display space. It is also possible to consider an `auditory fisheye(fishear?)`display.

It is difficult to find applications of distorted auditory spaces currently being used to display abstract information. However, the concept itself is not new and an experiment by Shinn-Cunningham and Durlach demonstrated that listeners can learn to interpret a distorted auditory space [Shinn-Cunningham and Durlach 1994]. This particular work tested the ability of users to perform a task with an *acoustic fovea*. The auditory display space was defined so that sounds directly in front of the user were pushed apart in space. At the same time, sounds at the periphery were compressed together[13]. This had the effect of increasing the display resolution in front of the user and reducing it to the side. Subjects were able to perform tasks better using the distorted auditory space.

The use of distorted auditory space is one area of spatial auditory metaphors that has not been well explored. This thesis describes a distorted auditory space that was implemented as part of an auditory-visual application. This application is described in detail in chapter 13.

3.3.1.5 Subdivided Auditory Space

A subdivided auditory space segments the available space into categories. Each segment of space can be defined differently. Alternatively, the same design can be repeated across the spatial categories. For example, the technique of `small multiples` is a very effective spatial visual metaphor for comparing a repeated visual element [Tufte 1990]. In the auditory domain repeated auditory elements can also be considered.

One advantage of using different spatial categories is that it helps the user to separate multiple sound sources[14]. This assists the listener focus on a single stream amongst other competing sound sources[15].

In the application called `Nomadic Radio`, the space was divided into 24 categories to help order messages by time [Sawhney and Schmandt 1997]. In work at Georgia Tech a 3D visual space was used in an environment called `Audio Rooms`. This application divided the auditory space based on rooms in a building [Mynatt 1994].

Once again, the use of a subdivided auditory space has not been well explored. In chapter 13 this thesis describes a subdivided auditory space used as part of an auditory-visual application.

3.3.2 Auditory Spatial Structures

Having defined an auditory display space, sounds can then be generated to occupy this space. These sounds can form auditory spatial structures that represent the data and are interpreted in terms of the underlying space. Section 3.3.1 considered the possible ways to define the auditory display space. This section considers two levels of auditory spatial structures. They are global auditory spatial structures (section 3.2.2.1) and local auditory spatial structures (section 3.2.2.2)[16].

---

[13] This is somewhat akin to the way visual space is defined in the application called the `Perspective Wall` [Mackinlay, Robertson et al. 1991].

[14] Although in the auditory domain, sounds can join together and form a single stream, thus being perceived to originate from a single location. The way separate sounds form streams is based not only on auditory position but also on a number of *gestalt principles* [Bregman 1990] (see section 3.3.2).

[15] This is sometimes called the *cocktail party effect* [Wenzel 1994].

[16] As with visual spatial structures the division of auditory spatial structures into *global* and *local* forms is somewhat arbitrary. The level of detail may simply depend on the user's location.

While it is relatively simple to choose an auditory display space and then organise sounds within this space hence, displaying information according to auditory position, it is much more difficult to design auditory spatial structures[17]. A fundamental problem with auditory representations is that sounds are transient and so do not provide an external representation on which the person can form a spatial mental model [Bennett and Edwards 1998]. Auditory objects have no persistence.

The idea of using auditory spatial structure for abstract display is not well explored. In fact sound has even been described as *"dimension free"* [Axen and Choi 1996]. While attempts have been made to represent complex structural geometry using sound properties they do not rely on the auditory spatial structure of sound. Axen and Choi, for example, use a symbolic mapping of geometry to the traversal of waves through the geometry. Events such as wave splitting, recombining and disappearing are used to signal global topological properties of the data. The goal is to produce audio output that bears a morphological relationship to the geometry. While this work has the goal of displaying spatial structure it does not attempt to make use of auditory spatial structures in sound, instead it uses direct sound properties and temporal structures to encode spatial structure. This demonstrates a major difficulty of using sound to display geometrical structure.

Kramer notes that the edges of an *audible object* are defined when a sound appears to move from one location to another or the subject moves through the sound [Kramer 1994a]. In each case, movement is required to determine the boundary of the auditory spatial structure. The spatial structure is being represented temporally as movement of the user or the sound requires time. While scanning a single object or looking back and forth between objects is a simple task visually, judging spatial extent by listening to multiple sounds moving from edge to edge is not [Kramer 1994a].

3.3.2.1 Global Auditory Spatial Structures

The global auditory spatial structure refers to the types of spatial relationships that occur between sounds (figure 3-71). For hearing, the perception of structure is usually based on the way that sounds are arranged in the auditory display space. For example, sounds that form auditory groups may be interpreted as related. Other spatial metaphor concepts such as connection or containment are unusual to consider in the auditory domain. However, these concepts can still be considered as possible options in auditory design.

---

[17] This may seem surprising, as the concept of structure is fundamental to musical composition. However, this type of structure refers to the auditory *temporal* structure of music. (Auditory temporal structures are discussed in chapter 5.)

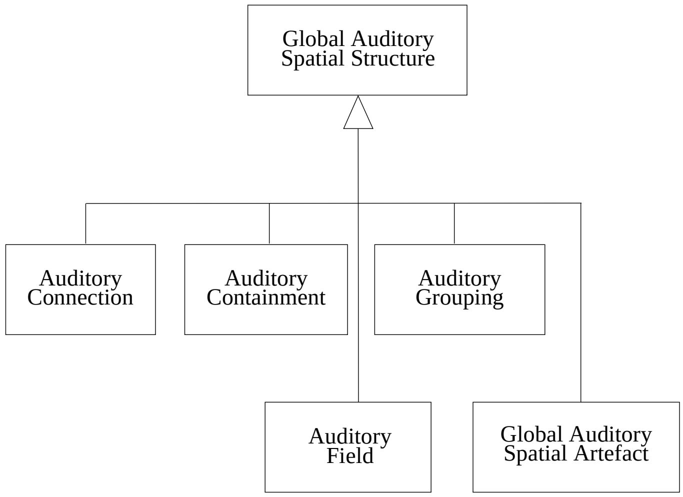


Figure 3-71 A UML diagram showing the types of global auditory spatial structures.

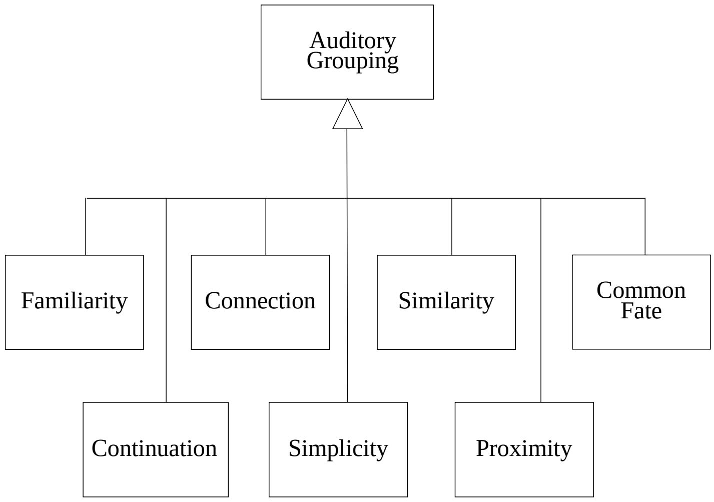


Figure 3-72 A UML diagram showing the types of auditory grouping.

Predicting how a listener might perceive global auditory spatial structures is quite complicated. Sounds from different spatial locations can join together and form a single perceptual stream [Bregman 1990]. Sounds that stream together are perceived to come from the same auditory position.

Bregman describes the way that a listener interprets a sound environment as *scene analysis* [Bregman 1990]. Scene analysis is determined in part by the *gestalt principles* [Goldstein 1989]. These principles have previously been used to help explain how users perceive the grouping of auditory objects [Williams 1994]. These *gestalt principles* are related to some auditory *spatial* properties. However they also relate to both auditory *direct* and auditory *temporal* properties (table 3-5, figure 3-72).

In general, when designing global *auditory* spatial structure it may be appropriate to use these principles to improve the separation of sounds or to encourage grouping. For example, to display an `auditory graph drawing,` it may be desirable to group some auditory nodes and separate others based on some data attribute.

Gestalt Principle	Impact on Sound Streaming
Law of Belonginess	A sound can only form part of one stream at a time. A sound is perceived in relation to the stream to which it belongs.

Law of Familiarity	If prior, well-known groupings of sounds are recognised they are grouped together.
Law of Similarity	Sounds that share attributes are perceived as related. An element in a stream may be captured by another stream with elements that are similar.
Law of Good Continuation	When a sound undergoes a smooth transition to another sound, those sounds are perceived as related.
Law of Proximity	Sounds that are close to each other are more likely to be grouped together.
Law of Common Fate	Sounds that undergo the same type of changes at the same time are perceived as related
The Law of Stability (Continuation)	When the user has achieved an interpretation of the sound that interpretation will remain fixed throughout slowly changing parameters until no longer appropriate

Table 3-5 The gestalt principles affect the way sounds are grouped in space [Williams 1994].

A number of application developers have experienced problems with the auditory display of structural information. Developers of an audio browsing environment for the web, found that one of the main difficulties was providing the spatial cues that normally provide a context for the information [Wynblatt, Benson et al. 1997]. In the visual domain, these context cues are often provided by the global spatial *visual* structure of the HTML documents. Other developers noted a similar problem when building internet browsers for the blind [Roth, Petrucci et al. 1998].

In section 3.2.2.1, `Graph Drawings` were described as an effective way for displaying relationships in data as global *visual* spatial structure. Attempts have also been made to display graphs using auditory displays. For example, Bennett and Edwards implemented an auditory display of node and link diagrams [Bennett and Edwards 1998]. The particular focus of this application was to present circuit diagrams. Experimental results showed that, while users could understand and internalise very simple diagrams, they generally failed to interpret more complex networks.

This example again illustrates the difficult nature of providing the listener with global *auditory* spatial structure. The outcomes illustrate that hearing is perhaps less effective for spatial metaphors than vision. This is not really surprising considered in light of the *Modal Specific Theory* [Friedes 1974] (section 2.3.4) or the fact that higher level neural processing of visual signals is based on spatial maps while hearing is based on tones [Goldstein 1989, p98] (section 2.3.2).

However, the difficulty of providing global *auditory* spatial structure does not mean that effective organisations of sound in space can not be designed. For example, the `Nomadic Radio` (section 3.3.1.2) uses a novel categorisation of the 2D auditory space, arranging messages about the user by using the structure of a clock face. Messages are arranged in space according to arrival time and are thus grouped according to this attribute. This is an excellent example of using auditory grouping (figure 3-72) to provide structure.

Auditory fields are another type of spatial structure. Fields are characterised as having a property that varies according to position in the display space. For example, sound fields have been used to encode abstract information about air flow in the cabin of a car [Eckel 1998]. In this application car designers can analyse the air flow from the air-conditioning ducts by moving through a sound field. The magnitude of air flow at any location is mapped to the sound level of a wind-like noise.

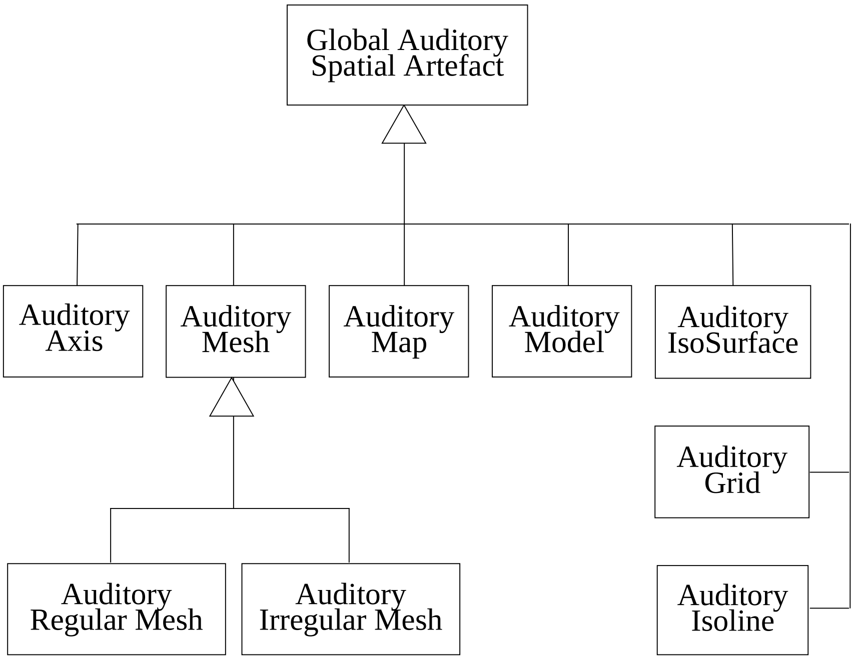


Figure 3-73 A UML diagram showing the types of global auditory spatial artefacts.

The concept of a global spatial artefact (figure 3-73) can also be considered for auditory displays. Global auditory spatial artefacts are sounds that are but not directly related to the data but designed to help the user interpret the display. For example, an `auditory grid` is analogous to a visual grid but displays reference lines as sounds.

There has been little work in using sound in this way. However the concept of an auditory map provides a good illustration of how a *global* auditory spatial artefact would work. For example, the application called `Cityscape` [Russo Dos Santos, Gros et al. 2000] used the global *visual* spatial structures of the city to display information. To enhance the *sights* of the city this application could be extended to include the *sounds* of the city. A user might then interpret proximity to busy 'roads' from the traffic sound.

3.3.2.2 Local Auditory Spatial Structures

The local auditory spatial structure refers to the types of spatial relationships that occur within sounds (figure 3-74). It is unusual to think of sounds with spatial structure and even more unusual to describe local spatial structure between the parts of a sound.

This type of display has not been well explored. This is not surprising, as it is very difficult to perceive spatially distinct sounds when they occur close together. This suggests that trying to interpret information from such structures would be extremely difficult. Though at least conceptually a notion like the auditory glyph is still feasible, even if the intuition is that it would be ineffective.

The most common use of local auditory spatial structure is to simply display all sounds as if they originate from a point in space. This concept is called an auditory point. More complex shapes suggested by auditory lines, auditory areas, auditory solids or auditory glyphs have not been used in information displays. However, experiments have tested the ability of subjects to recognise geometric shapes and alphanumeric characters [Hollander and Furness 1994]. In one experiment, subjects had to correctly identify an alphanumeric sound shape from 10 possible choices. Subjects were correct 33% of the time, performing significantly better than chance. In a second experiment, subjects had to correctly identify a geometric sound shape from 6 possible choices. Once gain subjects performed better than chance, correctly identifying the sound shape 42% the time. This indicates that the use of local spatial structure within sound may still be worth investigation.

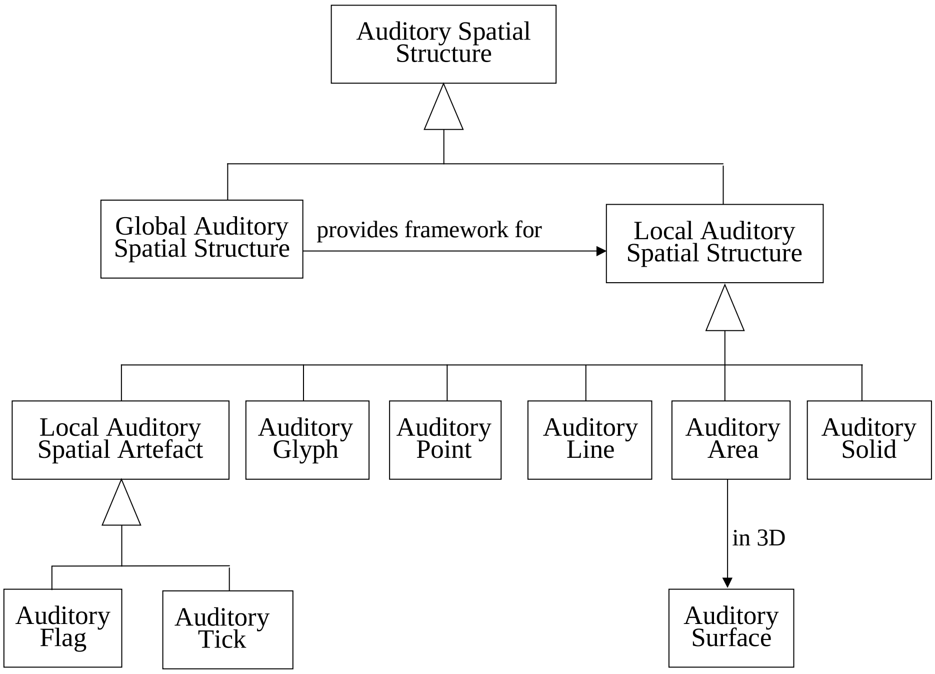


Figure 3-74 A UML diagram showing the types of local auditory spatial structures.

3.3.3 Auditory Spatial Properties

Section 3.3.1 considered the possible ways to define the auditory display space. Section 3.3.2 discussed the types of auditory spatial structures that are used to display data within the auditory display space. Auditory spatial structures have properties that are defined in terms of the auditory display space and these properties are described as auditory spatial properties (figure 3-63). The user interprets auditory spatial properties in terms of the auditory display space. Auditory spatial properties include:

- auditory position
- auditory scale
- auditory orientation.

For example, in `Nomadic Radio`, the position of a sound message in space encodes the message time [Sawhney and Schmandt 1997]. This applications displays sound as though it is generated from a point source. Hollander and Furness note that real sound sources are rarely point sources and usually have some scale or extent to them [Hollander and Furness 1994]. For example, the sounds of wind or the ocean waves have a size, although rarely is the auditory scale or for that matter auditory orientation ascribed much importance. Compared to spatial visual properties the use of spatial auditory properties to convey abstract information is unexplored.

The intuition is that spatial properties are not as effective when displayed to a temporal sense like hearing compared to a spatial sense like vision. However, the same concepts apply and if a design does use auditory scale or auditory orientation it should ensure that the auditory display space has an appropriate metric defined. Because the spatial resolution of hearing is low, the intuition is also that a categorical division of space would be the best way to define that metric (figure 3-76).

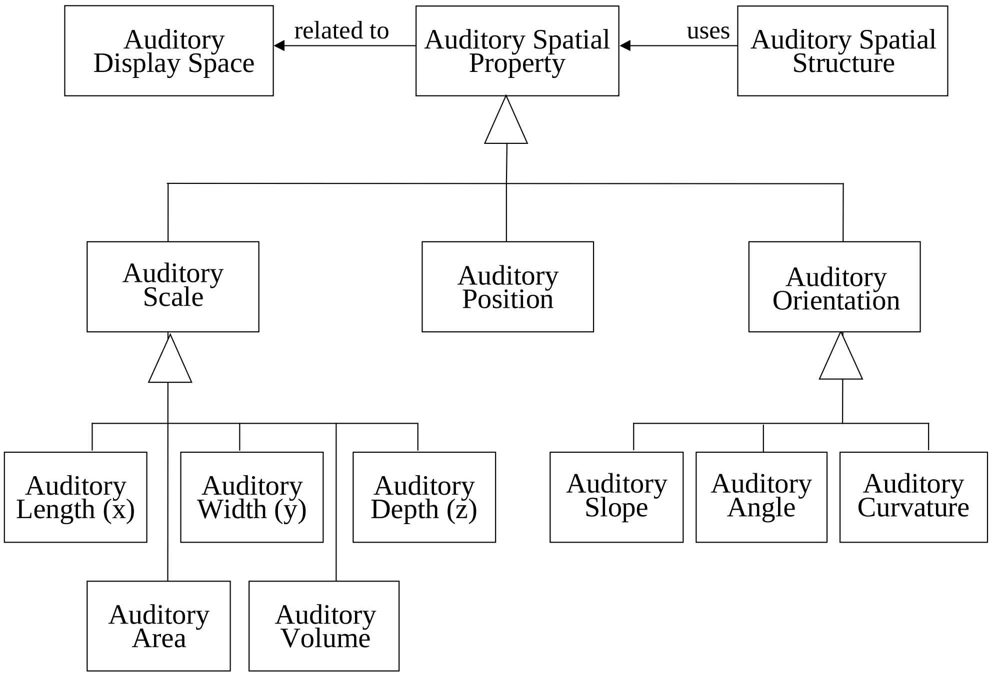


Figure 3-75 A UML diagram showing the types of auditory spatial properties.

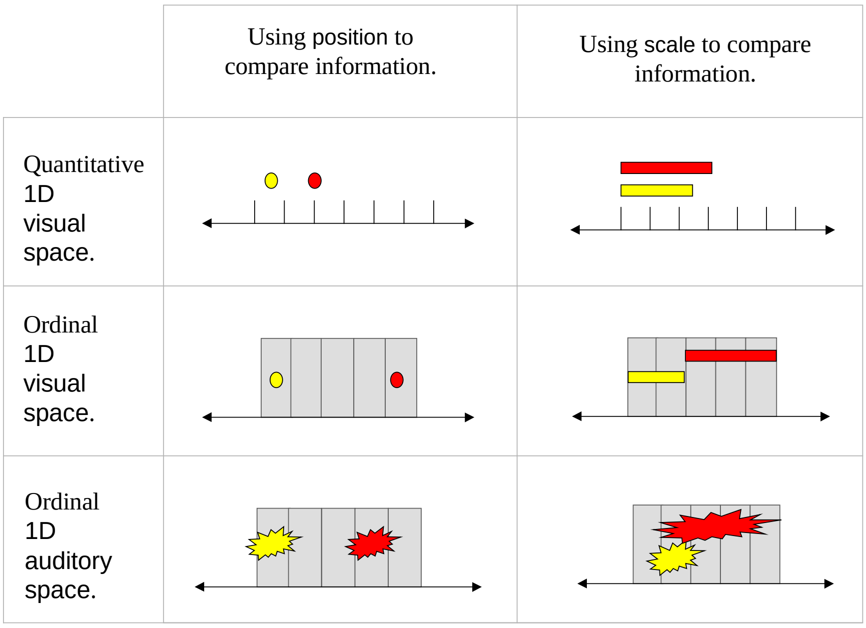


Figure 3-76 Comparing the use of position and scale in a 1D visual space and a 1D auditory space.

3.3.4 A Summary of Spatial Auditory Metaphors

The previous sections have reviewed the concepts that describe auditory visual metaphors. The important high-level components of auditory visual metaphors are the auditory display space (section 3.2.1), auditory visual structures (section 3.2.2) and auditory visual properties (section 3.2.3). These components are a specialisation of the more general concepts of display space, spatial structure and spatial properties that describe a spatial metaphor. However, the specialisation to the auditory sense is not as intuitive as the specialisation of the same concepts to the visual sense. This is not surprising as hearing is predominantly temporal and is more adept at identifying temporal relationships rather than spatial relationships [Friedes 1974].

The auditory display space is continuous and can be designed to display quantitative, ordinal or nominal data. However, auditory position is not a very accurate way of representing data. To overcome this problem a categorical subdivision of the auditory display space is more appropriate. Because of the problems with accurate identification of auditory position, other auditory spatial properties, such as auditory scale and auditory orientation are not recommended for displaying information.

A number of areas of the auditory design space have not been explored. These areas include the use of distorted and subdivided space. There are a number of opportunities for developing new types of auditory designs. In particular, adopting strategies from the better explored domain of spatial v*isual* metaphors could lead to new types of auditory displays. However, the lower resolution of the auditory display space needs to be considered when transferring these designs from the visual domain.

While spatial position is an effective way to display information, previous works have demonstrated the difficulty of displaying both global and local structures in auditory space. The suggested guideline is to display spatial structure visually. This raises the question in multi-sensory display of how are the visual display space and auditory display space related? In the real world, the reference framework for sounds is provided by the visual spatial structure. This same metaphor is suggested in the design of multi-sensory displays. That is, the auditory spatial structure should use the visual spatial structure as a framework for displaying the auditory information.

Tying the auditory spatial structure to the visual spatial structure does not preclude the use of different types of display space. The auditory display space may be different to the visual display space. One technique is to use an auditory display space with a different resolution to the visual display space. For example, in an application that provided both a visual and an auditory display of petroleum data a visual spatial structure was used for reference [Barass and Zehner 2000]. Data recorded from a petroleum well was mapped to a bar of colours. This bar showing attribute values at different depths in the well. The available resolution of the screen restricted the resolution of the visual display. The auditory display was used to display other parameters of the well data. As the user moved a cursor through the visual spatial structure showing the well, sounds were generated. These sounds represented a higher spatial resolution then the one provided by the visual display.

3.4 Spatial Haptic Metaphors

This section describes the concepts of spatial metaphors as they apply to the haptic sense (table 3-6). Spatial haptic metaphors relate to the haptic perception of the scale of objects in space, the position of objects in space and the structure of objects in space. The important considerations in the design of spatial haptic metaphors are (figure 3-77):

- the haptic display space (section 3.4.1)
- the haptic spatial structure (section 3.4.2)
- the haptic spatial properties (section 3.4.3).

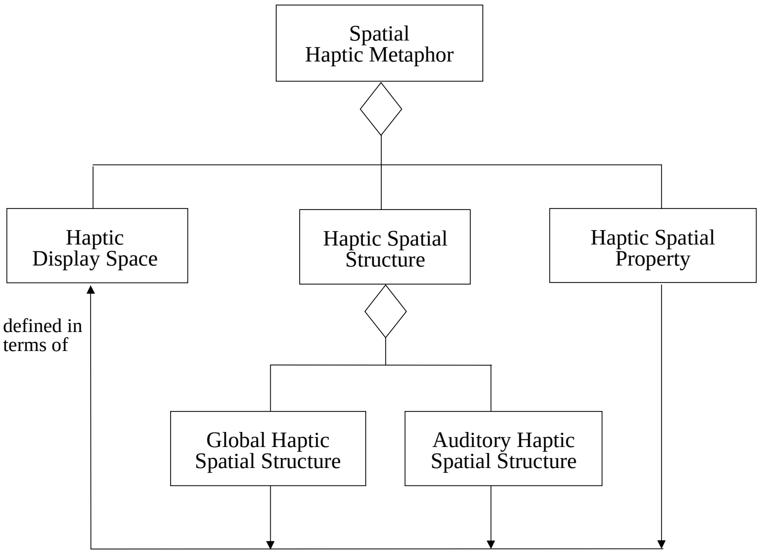


Figure 3-77 Spatial haptic metaphors describe the positioning of haptic structures in an organised space.

The following sections review some previous work using haptic displays[18]. This work is discussed in terms of the haptic display space, haptic spatial structure and haptic spatial properties. Section 3.2.4 summarises the important points of designing spatial haptic metaphors.

Spatial Haptic Metaphors (section 3.4)	Haptic Display Space (section 3.4.1)	Orthogonal Haptic Space	1D (section 3.4.1.1) 2D (section 3.4.1.2) 3D (section 3.4.1.3)
		Distorted Haptic Space (section 3.4.1.4)	
		Subdivided Haptic Space (section 3.4.1.5)	
	Haptic Spatial Structure (section 3.4.2)	Global Haptic Spatial Structure (section 3.4.2.1)	Haptic Field Haptic Connection Haptic Containment Haptic Grouping

[18] While a variety of spatial visual metaphors are used to display abstract data, there are fewer examples of where spatial haptic metaphors have been used. This is perhaps surprising as haptics, like vision can be thought of as a spatial sense. This scarcity of applications is probably due to the relatively recent availability of haptic displays and also the fact that current technology relies on temporal integration of spatial properties.

		Haptic Global Spatial Artefact
	Local Haptic Spatial Structure (section 3.4.2.1)	Haptic Line Haptic Point Haptic Region Haptic Surface Haptic Solid Haptic Glyph Haptic Local Spatial Artefact
Haptic Spatial Properties (section 3.4.3)	Haptic Position	
	Haptic Scale	Haptic Depth Haptic Length Haptic Area Haptic Volume
	Haptic Orientation	Haptic Angle Haptic Slope Haptic Curvature

Table 3-6 The sections in chapter 3 that describe the concepts of spatial haptic metaphors.

3.4.1 Haptic Display Space

Spatial haptic metaphors rely on the definition of a haptic display space. Mapping data to where objects can be felt in space is another possible technique for displaying information. For example, a `haptic scatterplot` would represent abstract data as points in space that the user could feel. Although to correctly interpret the abstract data requires the user to understand how the space has been defined.

The MS-Taxonomy defines the types of display space in terms of orthogonal, distorted and subdivided space. These types of space are applicable to the haptic sense, thus providing a number of ways to define the haptic display space (figure 3-78). The space can be defined using orthogonal coordinates to define a 1D, 2D or 3D space. The use of an orthogonal 1D haptic space is discussed in section 3.4.1.1. The use of an orthogonal 2D haptic space is discussed in section 3.4.1.2 and orthogonal 3D haptic space is discussed in section 3.4.1.3. Distorted haptic space is discussed in section 3.4.1.4 and subdivided haptic space is discussed in section 3.4.1.5.

It is possible to define a continuous, quantitative haptic display space. This space has a metric defined, so the distance between two objects is meaningful. However, The perception of haptic position is not as accurate as visual position (see section 2.3.3) and the lower resolution of the haptic display space needs to be taken into account. Therefore it is more appropriate to use a haptic display space based around continuous ordinal categories. This type of mapping provides a better match to the lower spatial resolution of haptics. Another alternative is to divide the haptic display space into nominal categories. This definition of the space does not rely on the user being able to interpret a metric.

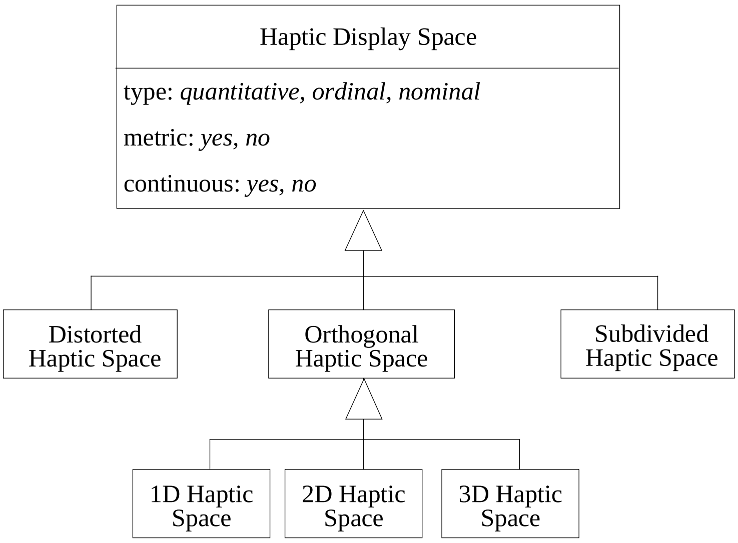


Figure 3-78 A UML diagram showing the components of the haptic display space.

3.4.1.1 Orthogonal 1D Haptic Space

The 1D display space is defined by a single axis (x) representing an abstract quantity. An orthogonal 1D haptic space is analogous to a haptic line. For example, in a `haptic lineplot` the user would be constrained to move along a line in 1D haptic space. Points along the line could be felt as barriers or walls of force through which the user must push.

No applications have explicitly been developed that use a 1D haptic space for displaying information. However, Fritz and Barner et al. describe the haptic line concept for scientific visualisation [Fritz and Barner et al. 1996] and some applications have implemented the concept. For example, using force to constrain the user to a line was used to implement `haptic slider widgets` in a user interface [Massie 1996].

3.4.1.2 Orthogonal 2D Haptic Space

The 2D haptic space is defined by two orthogonal axes (x, y) representing different abstract quantities. This spatial arrangement allows the user to compare different abstract quantities on the x and y axes by using the ability of the haptic sense to interpret spatial cues.

Some attempts have been made to use a 2D haptic space for displaying mathematical relationships. Force feedback was used to display both haptic lines and haptic points in this abstract space, thus producing the `haptic scatterplot` and `haptic 2D lineplot` [Fritz and Barner et al. 1996], [Ramloll, Yu et al. 2000]. For example, the use of a `haptic 2D lineplot` could help blind people to interpret abstract relationships [Sjöström 1997].

The `haptic 2D lineplot` and `haptic scatterplot` transfer common spatial visual metaphors into the haptic domain. These spaces have a metric and so allow quantitative relationships to be interpreted based on distance. The issue with transferring these concepts to spatial haptic metaphors is that the perception of haptic position is not as accurate as visual position. To compensate for this lower resolution, the use of a continuous ordinal space is more appropriate than a continuous quantitative space. This creates a space that is more discretised. For example, the line in a `haptic 2D lineplot` would feel more like a staircase than a straight line.

A number of applications have constrained the user to work within an abstract 2D haptic space. The `Nanomanipulator` was an early application of force feedback technology to assist users feel sub atomic resolution images from a scanning tunnelling microscope [Taylor, Robinett et al. 1993]. While the user could feel the height of artefacts on the image they were predominantly constrained to work along a plane. An interesting use of 2D haptic space was an application designed for remote collaboration on documents [Oakley, Brewster et al. 2000]. In this application the 2D haptic space is used to encode the proximity of remote document users. If a remote user is working on the same part of the document then a resistance force indicates their presence.

There have been a few efforts to augment the standard user-interface with haptic interaction. In this case the 2D haptic space corresponds to the 2D visual space of a typical computer user interface. For example, standard user-interface widgets, such as, buttons, sliders and menus are augmented with haptic feedback. These haptic widgets are designed to assist navigation or targeting [Eberhardt, Neverov et al. 1997], [Miller 1998], [Ruspini 1997]. However, when Oakley, McGee et al. tested a haptic scroll bar it was shown to provide no benefit for tasks involving searching and scrolling [Oakley, McGee et al. 2000].

In further work designed to enhance the user-interface the concept of planar haptic constraints were developed [Hutchins and Gunn 1999]. These force-based constraints can be employed to constrain users to a 2D haptic space while working in a 3D visual space. For example, Hutchins and Gunn developed planar shaped constraints (circle, rectangle) for use in a task, such as, planning mine shafts [Hutchins and Gunn 1999]. In an application designed to assist the interpretation of seismic data, the user was constrained to work on a haptic plane that sliced through the data [Nesbitt, Orenstein et al. 1997] (figure 3-79). This same idea of a 2D haptic space in a 3D visual space was used for the analysis of fluid flow in a blast furnace [Nesbitt, Gallimore 2001] (figure 3-80). In this example the 2D haptic space is continuous and quantitative and represents the fluid flow at each haptic position in space.

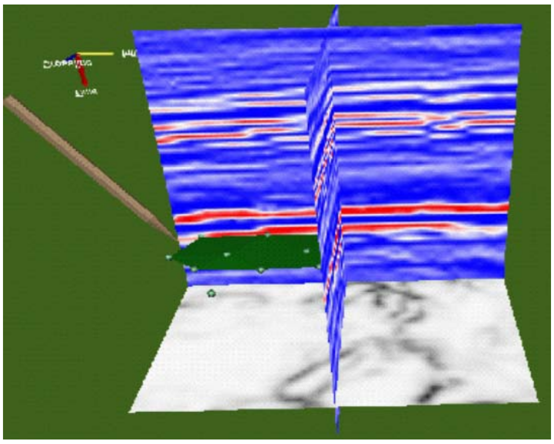


Figure 3-79 The user is haptically

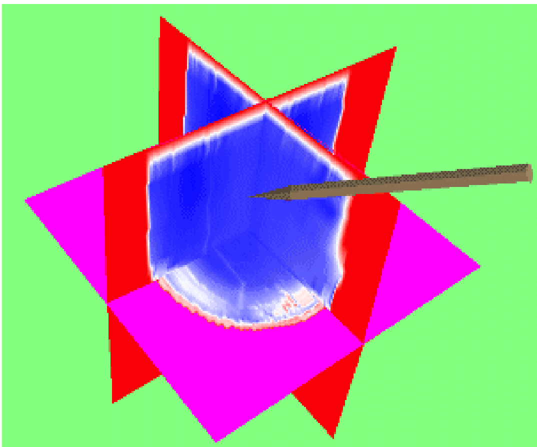


Figure 3-80 The user feels the fluid flow on a

constrained to a 2D plane for interpreting seismic data [Nesbitt, Orenstein et al. 1997]
Courtesy of: BHP Research, Newcastle, Australia

2D plane in a mathematical model of a blast furnace [Nesbitt, Gallimore 2001].
Courtesy of: BHP Research, Newcastle, Australia

3.4.1.3 Orthogonal 3D Haptic Space

A 2D haptic space can be extended to make a 3D haptic space by adding a third orthogonal axis. Because each of the three axes can now be defined by either a quantitative, nominal or ordinal variable there are many different ways to define a 3D haptic space.

While orthogonal 3D haptic spaces have rarely been used for displaying abstract data, it is relatively common for representing real world models. For example, a number of systems have used these spaces to augmenting real world structure. These include applications that provide a 3D haptic space for interaction with CAD models [Larsson 1997], electronics test consoles [Davidson 1996], anatomical organs [Mor 1996], clay sculptures [Massie 1996] and even Lego™ bricks [Young, Chen et al. 1997]. Furthermore, a number of applications have had the goal of producing physical-based simulators [Latimer and Salsibury 1997]. This trend is evident in the medical industry where haptics have been used in a needle insertion simulator, a surgical cutting simulator and a dental probe simulator [Reinig 1996].

There are some examples of defining a 3D haptic space when abstract data is overlaid on real world models. For example, an application used haptics to help segment medical image data (CT, EBT, MRI, Doppler, Ultrasound) [Giess, Evers et al. 1998]. In an application to assist in oil exploration, spatial location in the x and y dimension and time in the z dimension define a quantitative and continuous 3D haptic space [Nesbitt, Orenstein et al. 1997]. While exploring the seismic data, the user could feel important features such as faults in the 3D haptic space.

There has also been some interest in using 3D haptic spaces for Scientific Visualisation. For example, vector fields such as magnetism and fluid flow have been represented in a continuous, quantitative 3D haptic space [Fritz and Barner et al. 1996]. Force feedback devices display force vectors that can be easily be used to represent other vector fields in a 3D haptic space. Haptic devices have also been used to model real world dynamics, such as, momentum and inertia [Ruspini 1997]. One of the first uses of force to display information was the `GROPE` project at the University of North Carolina [Batter and Brooks 1972]. This display was designed to assist users in molecular docking studies. The 3D haptic space in this application is a quantitative continuous space. In this space the user feels the intermolecular forces at any location dependent on the structure of molecules.

3.4.1.4 Distorted Haptic Space

A distorted haptic space is not defined in a uniform fashion but rather the definition of space varies according to position. For example, in the visual domain the `Fisheye View` [Furnas 1981] provides a distorted definition of the visual display space. It is also possible to consider a `haptic fisheye (fishhand?)` display. In general, the use of a distorted haptic space is an area of spatial haptic metaphors that has not been explored.

3.4.1.5 Subdivided Haptic Space

A subdivided haptic space segments the available space into categories. Each segment of space can be defined differently. Alternatively, the same design can be repeated across the spatial categories. For example, Fritz and Barner produced an application that divides a plane into a 2D lattice [Fritz and Barner et al. 1996]. Data was then mapped onto the haptic height at each categorical location in the lattice. Once again this concept has not been explored for the haptic display of abstract data.

3.4.2 Haptic Spatial Structures

Having defined a haptic display space, objects can then be generated to occupy this space. These objects can form haptic spatial structures that represent the data and are interpreted in terms of the underlying space. Section 3.4.1 considered the possible ways to define the haptic display space. This section considers two levels of haptic spatial structures. They are global haptic spatial structures (section 3.4.2.1) and local haptic spatial structures (section 3.4.2.2).

The division into *global* and *local* structures is somewhat arbitrary. The level of detail may simply depend on the user's position. For example, if the user moves closer to a local haptic spatial structure it may reveal details that, in the new context, are more *global*. The important point during design is that both *global* and *local* strategies for presenting information are considered. During implementation the distinction is more important as the current hardware for haptic displays must be considered. Point-based force-feedback devices require temporal integration of spatial information. While these displays allow for interpretation of *local* haptic spatial structure they are less effective for interpreting *global* haptic spatial structure.

3.4.2.1 Global Haptic Spatial Structures

Global haptic spatial structure refers to the type of relationships that can be felt between objects (figure 3-81). For example, both haptic connection and haptic containment can infer relationships. The perception of structure can also occur based on the way objects are arranged in the haptic display space. For example, objects that form haptic groups may be interpreted as related. More complex relationships could be inferred from symmetric arrangements or the recognition of familiar shapes such as squares or circles.

As was the case with both the visual and auditory senses, predicting how haptic spatial structures are perceived is quite complex. However, the gestalt principles that help explain how users perceive the form of visual objects and group auditory objects can also help explain how haptic structures are perceived [Goldstein 1989, p308].

The terms used to describe spatial structures are fairly intuitive for the haptic sense. For example, the concepts of haptic containment and haptic connection are straightforward to understand. There is also an obvious link between visual structures and haptic structures. In an application designed for the visually impaired, users could feel the global structure of a painting [Sjöström 1997]. Each colour in the painting was identified by a different texture. This allowed users to feel the overall structure of the picture This application might be interpreted as using haptic containment to display structural information. It also illustrates how haptic grouping can apply. For example, objects with a similar surface texture are likely to be grouped together.

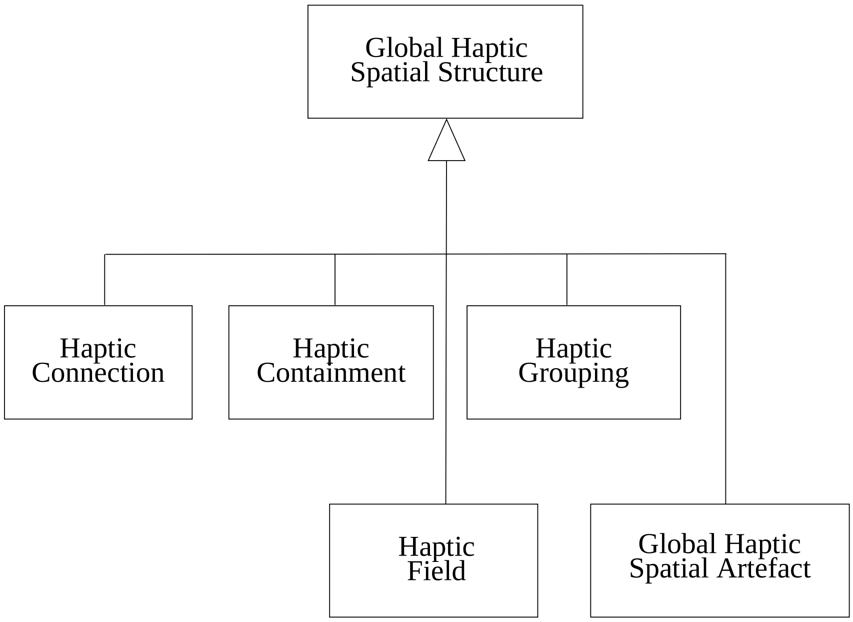


Figure 3-81 A UML diagram showing the types of global visual spatial structures.

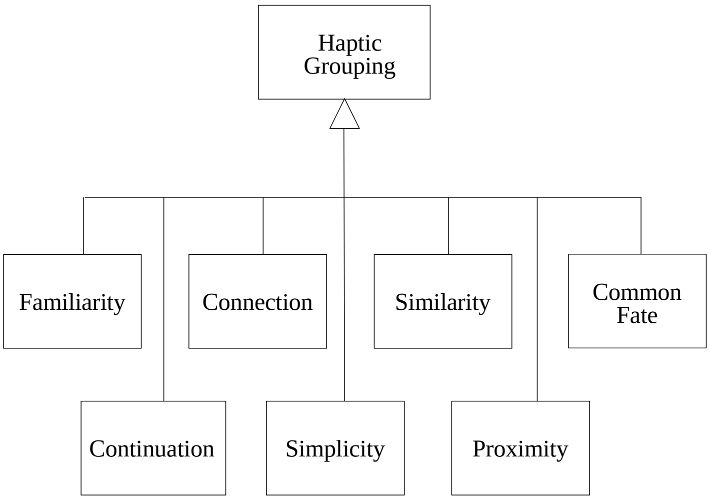


Figure 3-82 A UML diagram showing the types of haptic grouping. These types are derived from the gestalt principles.

A number of applications have been developed that provide the user with a haptic display of structure. For example, a system was developed to create `haptic 2D lineplots.` The haptic plots try to mimic a surface engraving [Yu, Ramloll et al. 2000]. In some preliminary experiments it was found that users could perceive a general structure of the plot. However there were also some problems found with the system. For example, users could not distinguish maximum and minimum points and were often confused when exploring more complex plots with multiple lines[19].

Force feedback has been used to provide additional global haptic spatial structure in the user interface [Miller 1998]. These haptic cues were designed as navigation aids for the X-Window System. Ridges and dimples were added to icons and menu items. Alignment guides were added to assist moving windows. However, no formal evaluation of these enhancements was reported.

In a prototype produced for petroleum exploration, users were assisted to feel the global haptic spatial structure representing a discontinuity between geological strata [Nesbitt, Orenstein et al. 1997]. This surface in the volume of seismic data is called a *horizon*. Interpreters mapping a horizon are constrained to the most likely surface as calculated from the underlying data. This has the effect of assisting the interpreter tracking along the horizon by following the haptic connection.

In an application from the related domain of mineral exploration haptics are used to assist the user understand the global haptic spatial structure of underground ore bodies [Veldkamp, Turner et al. 1998]. The model of such ore bodies is never exact but is constantly being extrapolated from the available data. Usually the data is only accurate at the sparse locations where samples have been drilled from the ground. The advantage that haptic display provides in this application is that the surface geometry can be flexible in some locations and yet fixed in others. Fixed structures are more accurate as they are based on fewer interpolations of the underlying data.
These two applications, described for mineral and petroleum exploration, both use the idea of *haptic constraints*. *Haptic constraints* use force-feedback to constrain the range of the user's movement. General constraint classes have previously been developed for geometric objects such as planes, boxes, spheres, cones and cylinders [Hutchins and Gunn 1999]. These primitives can be combined to describe more elaborate global haptic spatial structures. Haptic constraints are usually implemented as haptic fields and are discussed in more detail in section 3.4.3.

Avila and Sobierajski describe an application that helps users understand complex spatial arrangements by displaying global haptic spatial structures [Avila and Sobierajski 1996b]. For example, users can feel the surfaces representing a complex network of dendrites from nerve cells. The claim is made that the haptic display both assists navigation and provides the user with a better understanding of the geometry. However, no experimental evidence is provided.

In the `Legoworld` project [Young, Chen et al. 1998] a 3D environment was created that allows the assembly of Lego™ blocks. This includes the ability to examine the surface geometry along with the weight and inertia of blocks. The user can also feel the sensation of locking and unlocking blocks together. This example highlights the use of force to represent local haptic spatial structure spatial structure by direct contact with objects. More global haptic spatial structure is interpreted from the collisions that occur between objects.

---

[19] It is noted that single-point force-feedback makes interpretation of complex global haptic spatial structure difficult. The intuition is that being able to simultaneously feel multiple locations on the 2D haptic lineplot would better allow structural features such as maximum and minimum points to be identified.

A haptic tool called Dynadraw was developed to help sculptors create global haptic spatial structures [Snibbe, Anderson et al. 1998]. In an application called Griddraw, haptic display has been used to assist the user to draw straight lines [Snibbe, Anderson et al. 1998]. These applications again illustrate the concept of use haptic constraints. *Haptic constraints* can be a useful way of modelling both haptic containment and haptic connection.

The GROPE project allowed users to interpret molecular structure in terms of the forces between molecules[Batter and Brooks 1972]. By feeling a force model in addition to seeing the visual display of the molecule the user could gain an understanding of the structure of molecules, the location of molecules and the interaction of forces between atoms. SCULPT is a commercial molecular modelling package that allows the user to feel local haptic spatial structure, but also interpret global haptic spatial structure in terms of the global forces [Wanger 1998]. The force model is generated using virtual springs between the atoms and the cursor.

Haptic fields can be thought of as a type of spatial structure in the haptic display space. Fields display a direct property that varies according to position in the display space. The concept of a haptic field has frequently been implemented[20]. Some example applications of haptic fields include the modelling of molecular forces of attraction [Batter and Brooks 1972], fluid flow fields [Nesbitt, Gallimore 2001] and gravity-wells [Oakley, McGee et al. 2000]. The work of Durbeck, Macias et al. specifically tries to enhance Scientific Visualisations by allowing users to interact with vector fields representing gravity, pressure, force and velocity [Durbeck, Macias et al. 1998].

A common technique used in haptic interfaces is to implement constraints on the user's actions. These constraint forces are modelled as a haptic field. The user encounters the field at specific locations in the display. For example, in an application for human-assisted motion planning, haptic hints about suitable paths were provided to help the user solve a motion planning problem [Bayazit, Song et al. 1999]. Many types of constraints have been designed which can also provide structure. Another example is the interaction paradigm described as the magnetic metaphor [Noma, Kitamura et al. 1996]. In this interface the user is restricted to moving objects as though the objects were constrained like a magnet along a metal surface. Experiments have shown that such magnetic fields can assist the user in the accuracy of acquiring targets [Wall, Paynter et al.2002]

The alternative to constraining forces is to use a haptic field of attraction. Eberhardt, Neverov et al. implement gravity-like attraction fields [Eberhardt, Neverov et al. 1997]. These haptic fields are designed to assist the user in the task of button acquisition in a typical windows-icon-menu-pointer interface

Another type of global haptic spatial structure is the concept of a global haptic spatial artefact (figure 3-83). For example, a haptic grid is a type of global haptic spatial artefact. A haptic grid can assist the user to interpret the haptic position of information in the haptic display space. Haptic grids have been implemented to assist scaling and

[20] This is probably because there is a natural match between the use of current force-feedback displays and the concept of a field. Force feedback devices display a force vector to the user at each time step. This force vector is simply determined from the user's position in the display space.

navigation in a system for Scientific Visualisation [Fritz and Barner et al. 1996]. The same concept was also used to try and help users interpret `2D haptic lineplots` [Yu, Ramloll et al. 2000]. In fact, these haptic grids proved problematic as users had to count across gridlines to identify their haptic position. The same concept was used in a game for visually impaired players. The game is called `Submarine` and is based on the traditional game called `Battleship`. Players feel squares defined by a haptic grid when deciding on a location to drop a bomb [Sjöström 1997].

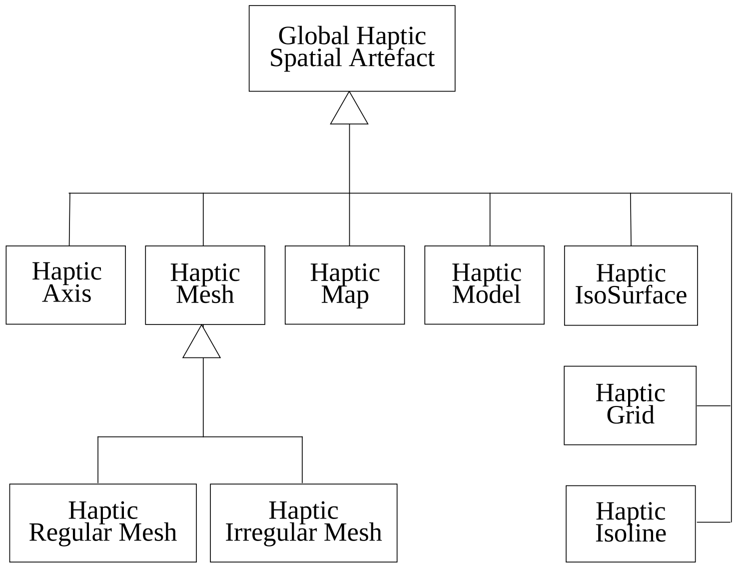


Figure 3-83 A UML diagram showing the types of global haptic spatial artefacts.

The concept of a haptic grid is directly related to the concept of a visual grid. Indeed many of the concepts that describe global *visual* spatial artefacts transfer intuitively to the domain of global *haptic* spatial artefacts. For example, a common structure used in Scientific Visualisation is the visual mesh. Fritz and Barner et al. provide for a haptic mesh to be constructed by connecting haptic points with haptic lines [Fritz and Barner et al. 1996]. The `inTouch` application also implements a haptic mesh, allowing the user to paint onto the mesh surface with an electronic paintbrush [Lin, Gregory et al. 1999].

The use of a haptic isosurface is illustrated in a number of applications. For example in an application to help explore seismic data, the user can select a plane through the data at some particular location in the abstract space [Nesbitt, Orenstein et al. 1997]. The user is then constrained to work in this plane. The effect is to constrain one dimension of the data space to a constant value. Furthermore, the user may also choose to automatically construct haptic isosurfaces that estimate geological faults.

3.4.2.2 Local Haptic Spatial Structures

The local haptic spatial structure refers to atomic objects or the types of spatial relationships that occur within objects (figure 3-84). For example, shape details or relations between the sub-components of an object may convey information. For a concept like the haptic glyph is included at this level of the MS-Taxonomy.

In general any haptically rendered shape could be used to represent data. At this point the MS-Taxonomy distinguishes between two different uses of shape. When a group of simple shapes are used to encode a set of nominal categories the concept is called direct haptic shape. The concept of direct haptic shape is included under direct haptic metaphors and is discussed in chapter 4. The concept of local haptic spatial structure encodes information using the relationships between the components of the structure. For example, with a haptic glyph information is inferred from patterns in the spatial relationships of the sub-elements. With direct haptic shape it is the overall perceived shape that is recognised perceptually.

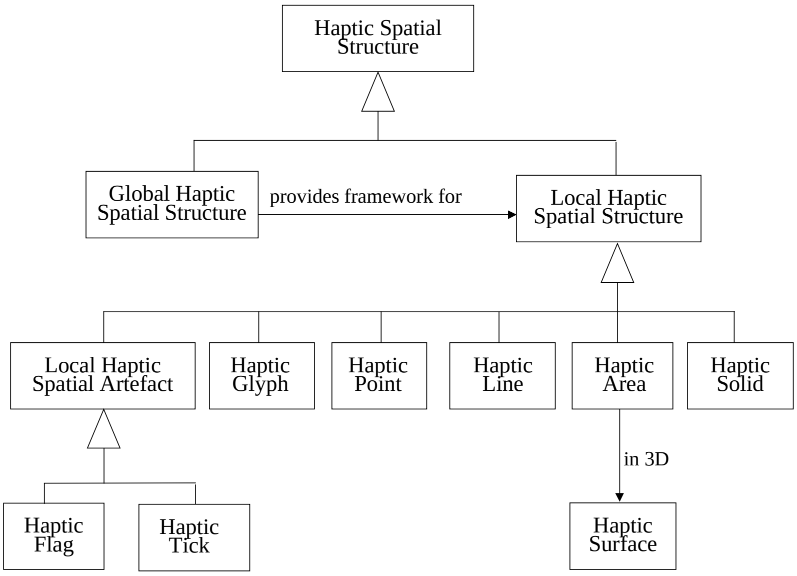


Figure 3-84 A UML diagram showing the types of local haptic spatial structure.

The use of local haptic spatial structure to display information is unexplored. However this technique is relatively common in applications that display real world models. Usually local *haptic* spatial structure is displayed in conjunction with global *visual* spatial structure. Haptic feedback provides accurate spatial cues about the position of objects in space. For example, making contact with an object determines its precise location in space. In a 3D visual display, the visual cues may only provide approximate information about the depth of an object in space.

There is some experimental evidence of the usefulness of combining local haptic spatial structures with global visual spatial structures [Srinivasan and Basdogan 1997]. Subjects were tested in terms of their ability to navigate through a maze using both visual and haptic feedback. Subjects performed best when the global *visual* spatial structure was augmented with the local *haptic* spatial structure of the maze [Srinivasan and Basdogan 1997]. This technique was also reported to be effective in an application designed to help segment surfaces from volumetric medical data [Giess, Evers et al. 1998].

3.4.3 Haptic Spatial Properties

Section 3.341 considered the possible ways to define the haptic display space. Section 3.4.2 discussed the types of haptic spatial structures that are used to display data within the haptic display space. Haptic spatial structures have properties that are defined in terms of the haptic display space and these properties are described as haptic spatial properties (figure 3-85). The user interprets haptic spatial properties in terms of the haptic display space. Haptic spatial properties include:

- haptic position
- haptic scale
- haptic orientation.

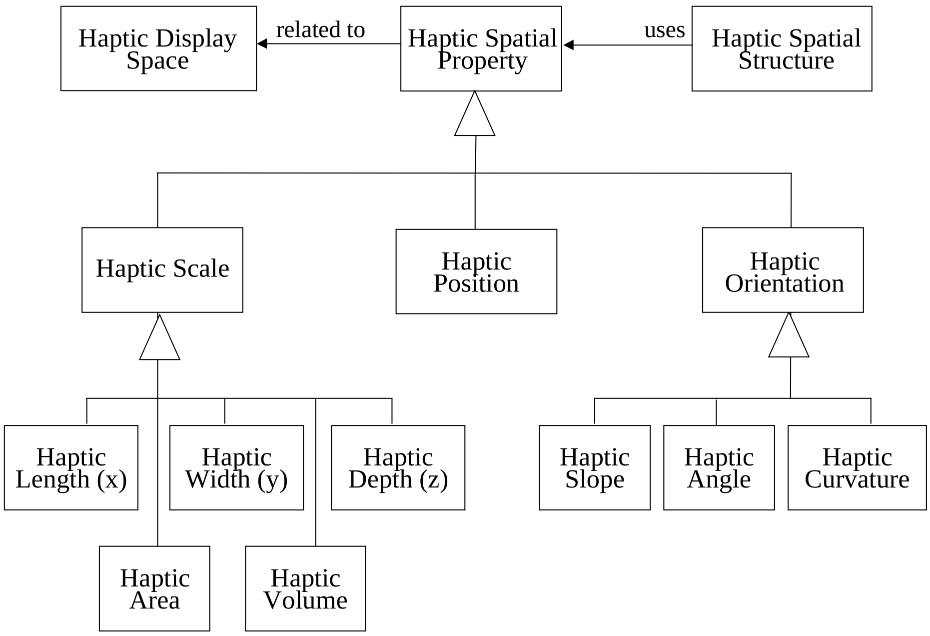


Figure 3-85 A UML diagram showing the types of haptic spatial properties.

Examples of using haptic position, haptic scale and haptic orientation to display information have already been mentioned in examples in section 3.4.1. It has been noted that a metric must be defined on the space for using some of these haptic spatial properties. It is also noted that compared to the visual senses the haptic sense is much less accurate at identifying the position of objects in space. The haptic system also tends to lose track of absolute spatial location [Sekuler and Blake 1994]. This makes it difficult to perform accurate tracking of haptic position. The intuition from this observation is that an ordinal division of the haptic display space is more appropriate for the design of spatial haptic metaphors.

Despite this, attempts have been made to use a quantitative haptic display space for transferring common spatial visual metaphors to the haptic domain. These include the `haptic 2D lineplot`, `histogram, Pie Charts` [Yu, Cheung et al. 2002] and `surfaceplots` [Sjöström 1997]. These applications illustrate the use of haptic position, haptic scale and haptic orientation to display information.

The difficulty of interpreting haptic orientation is illustrated in experiments designed to measure the ability of subjects to discern between lines separated by different angles [Riedel and Burton 2002]. Subjects were asked to determine if two haptic lines diverged, converged or remained parallel. Most subjects detected divergence in lines that were 15° removed from parallel. However, when detecting divergences of 5°, subjects performed no better than chance[21].

3.4.4 A Summary of Spatial Haptic Metaphors

Due to the relative immaturity of the field, it is not surprising that few generic designs for haptic information display have evolved. As the domain of spatial haptic metaphors is relatively unexplored there are a number of opportunities for developing new types of haptic designs. In particular, strategies from the domain of spatial *visual* metaphors can be transferred to spatial *haptic* metaphors[22].

The haptic sense integrates both spatial and temporal information [Friedes 1974]. The haptic display space is continuous and can be designed to display quantitative, ordinal or nominal data. However, when representing abstract information it is more accurate and faster to use visual position rather than haptic position [Sekuler and Blake 1990, p209]. To overcome this problem a categorical subdivision of the haptic display space can be used. The ability to detect haptic position accurately will also impact on the ability to accurately interpret then other haptic spatial properties such as haptic scale and haptic orientation.

How are the visual display space and haptic display space related? In the real world there is often a close level of co-operation between visual and haptic perception [Jansson, Fänger et al. 1998]. In the real world the reference framework is provided by the global visual spatial structure while finer detailed information can be obtained by local haptic spatial structure. Furthermore the haptic sense is usually dominated by vision if they are in conflict to each other. This is called *intersensory dominance* and experiments have shown that subjects base judgements on visual shape rather than haptic shape when the two senses are in conflict [Rock and Victor, 1994]. Visual information can also alter the haptic perception of object size, orientation and shape [Srinivasan and Basdogan 1997].

The haptic sense is predominantly associated with detection of movement [Friedes 1974]. The temporal and spatial nature of haptics makes this sense an intuitive choice for displaying vector fields. Another concept that is intuitive to the domain of haptics is the idea of constraints. These constraints correspond to barriers or walls in the real world and can be thought of as limits or ranges in abstract data.

As already noted with spatial auditory metaphors, the display spaces for each sense can have different resolutions. For example, it is possible to consider a design space at three levels of resolution perceived by the visual, auditory and haptic senses. Hence the resolution of the haptic display space may be finer than the visual display space.

Finally it is noted that current point-based force feedback devices requires temporal integration to interpret haptic spatial structure. Normally our haptic sense integrates information across a wider spatial extent. For example, small objects are usually manipulated by holding them with both hands and using several fingers. While for

---

[21] There were also some unusual anomalies reported. Subjects performed significantly worse interpreting lines with a vertical orientation compared to lines with a horizontal orientation.

[22] When transferring concepts from the domain of visual display to haptic display, the lower resolution of the haptic sense needs to be considered.

exploration with the Phantom™ there is never more than a single point of contact between the observer and the object [Jansson, Fänger et al. 1998].

Because of the availability of the Phantom™ most of the current work in haptic display has been implemented using this device. This device relies on *temporal integration* of shape information. The importance of using *spatial integration* to discern shape information has been illustrated in previous studies [Kirkpatrick and Douglas 1999]. Kirkpatrick and Douglas examined the hand gestures that subjects made to capture shape information. The subjects took seven times longer to capture shape information with a single finger compared to using their hands. This study illustrates that integrating spatial information over time (temporal integration) is not as effective as perceiving spatial structure in a single gesture (spatial integration).

Thus current devices are more appropriate for displaying local haptic spatial structure or confirming global visual spatial structure. The intuition is, that with the advent of more sophisticated, spatialised, haptic displays the viability of spatial haptic metaphors for displaying abstract information will improve.

3.5 Conclusion

This chapter has discussed in detail the concepts that describe spatial metaphors. These are general concepts that can be specialised for each of the senses. This has been illustrated by a detailed description of the spatial metaphor concepts for spatial *visual* metaphors (section 3.2), spatial *auditory* metaphors (section 3.3) and spatial *haptic* metaphors (section 3.4). The concepts of spatial metaphors are very intuitive for both the visual and haptic sense and less intuitive for the auditory sense. This is perhaps not surprising given that hearing is described as a temporal rather than spatial sense.

Previous work was discussed using the concepts of the spatial metaphors. No existing displays were uncovered that could not be described by the MS-Taxonomy. This supports the validity of the concepts that make up spatial metaphors. For interpreting abstract data in terms of space, the concepts of spatial metaphors successfully describe the multi-sensory design space. These concepts also provide multiple levels of complexity for considering spatial information displays.

During the review some observations have been made about the effectiveness of using different spatial metaphors. These problems and suggestions are incorporated into the MS-Guidelines for spatial metaphors described in chapter 7. Some generic spatial visual metaphors were identified however there are no examples of commonly used (generic) auditory and haptic displays.

The use of spatial visual metaphors for abstract data display is more mature than both spatial *auditory* metaphors and spatial *haptic* metaphors. This provides an opportunity to transfer existing ideas from spatial *visual* metaphors to both spatial *auditory* metaphors and spatial *haptic* metaphors.

Chapter 4

Direct Metaphors

Green-eyed Butterfly, Blue-eyed Dragonfly (2002)

"Human subtlety will never devise an invention more beautiful, more simple or more direct than does Nature, because in her inventions, nothing is lacking and nothing is superfluous." [Leonardo DaVinci]

Chapter 4

Direct Metaphors

4.1 Introduction

In the real world a great deal of useful information is perceived directly from the properties of sights and sounds. For example, a sound may have a certain loudness or pitch. Objects in the real world may be recognised on the basis of visual properties such as colour or on haptic properties such as surface texture. Real world information is often interpreted in terms of properties like pitch, colour and texture. Abstract information can also be interpreted in terms of these direct properties.

An important distinction between spatial metaphors and direct metaphors is that direct metaphors are interpreted independently from the perception of space. While the concepts of spatial metaphors apply generally for each sense this is not true for direct metaphors. There is very little intersection, for example; between the low level concepts of direct visual metaphors and the low level concepts of direct auditory metaphors. This is not surprising as direct metaphors relate to the properties that the individual sensory organs can detect.

Direct metaphors are concerned with direct mappings between the properties perceived between each sense and some abstract information. Direct metaphors are specific for each sense. This chapter describes the three classes of direct metaphors (table 4-1):

- *Direct visual metaphors* concern the perceived properties of pictures.
- *Direct auditory metaphors* concern the perceived properties of sounds.
- *Direct haptic metaphors* concern the perceived properties of touch.

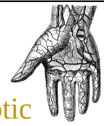
		SENSORY DISPLAY MODES		
		 visual	 auditory	haptic 
METAPHOR CLASSES	SPATIAL METAPHORS (chapter 3)	Spatial Visual Metaphors (section 3.2)	Spatial Auditory Metaphors (section 3.3)	Spatial Haptic Metaphors (section 3.4)
	DIRECT METAPHORS (chapter 4)	**Direct Visual Metaphors (section 4.2)**	**Direct Auditory Metaphors (section 4.3)**	**Direct Haptic Metaphors (section 4.4)**
	TEMPORAL METAPHORS (chapter 5)	Temporal Visual Metaphors (section 5.2)	Temporal Auditory Metaphors (section 5.3)	Temporal Haptic Metaphors (section 5.4)

Table 4-1 The focus of chapter 4 is on the three classes of *direct metaphors*.

Direct metaphors consider the following design concepts (figure 4-1):
- spatial structure
- direct properties.

The concept of spatial structures was discussed in the chapter 3. Spatial structures are a component of spatial metaphors that can be used to convey information. These structures can be encoded with additional information by using a directly perceived property of any sense. For example, colour can be used with a visual display, pitch with a sound display or surface texture with a haptic display.

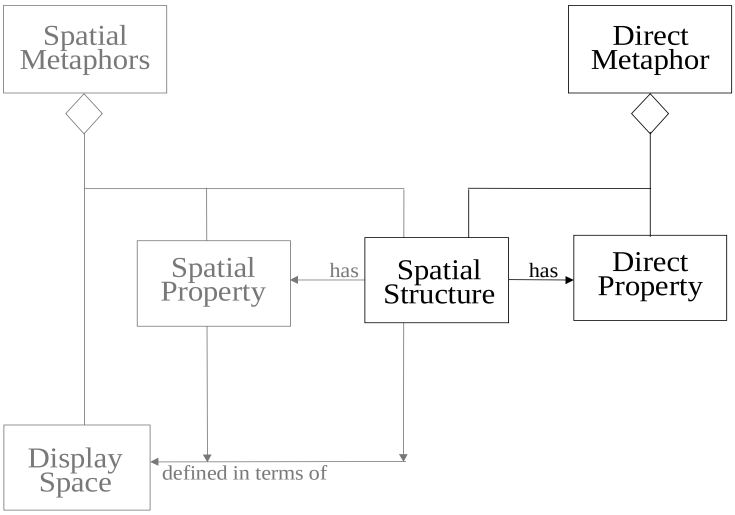


Figure 4-1 A UML diagram showing the high-level components of direct metaphors.

The key component of direct metaphors is the direct property used to convey the information. In terms of design, the effectiveness of a direct metaphor is independent of the display space and the spatial structure. However, there needs to be some consideration for the size of the spatial structure. For example, very small areas of colour may not be visible to the user. Likewise, it may be difficult for a user to investigate the haptic surface texture of a very small surface. Because spatial structures are discussed in the chapter 3 they will not be discussed again. This chapter focuses on the direct properties perceived by each sense.

The ability to accurately interpret direct properties varies between senses and properties. The direct properties for each sense are discussed in detail in this chapter. In general, the perception of all direct properties is of insufficient accuracy to allow accurate judgement of quantitative values. This suggests that direct properties should only be used to encode ordinal or nominal categories of data. Because direct properties such as colour or pitch are continuous they can easily be mapped to continuous data. However, it should not be assumed that a user is capable of interpreting exact data values represented as direct properties.

The MS-Taxonomy distinguishes between direct visual, direct auditory and direct haptic metaphors. At a low-level of the hierarchy, the concepts do not abstract across the senses (figure 4-2). This makes it difficult for direct metaphors to be directly compared between senses. For example, it makes little sense to compare the ability of the visual and auditory sense at judging the *pitch* of sounds.

However, at a higher level, the concept of a direct property does apply across the senses. Therefore at a conceptual level the designer can consider substituting one direct property with another. For example, the direct visual property of visual texture could be substituted with the direct auditory property of timbre or the direct haptic property of compliance.

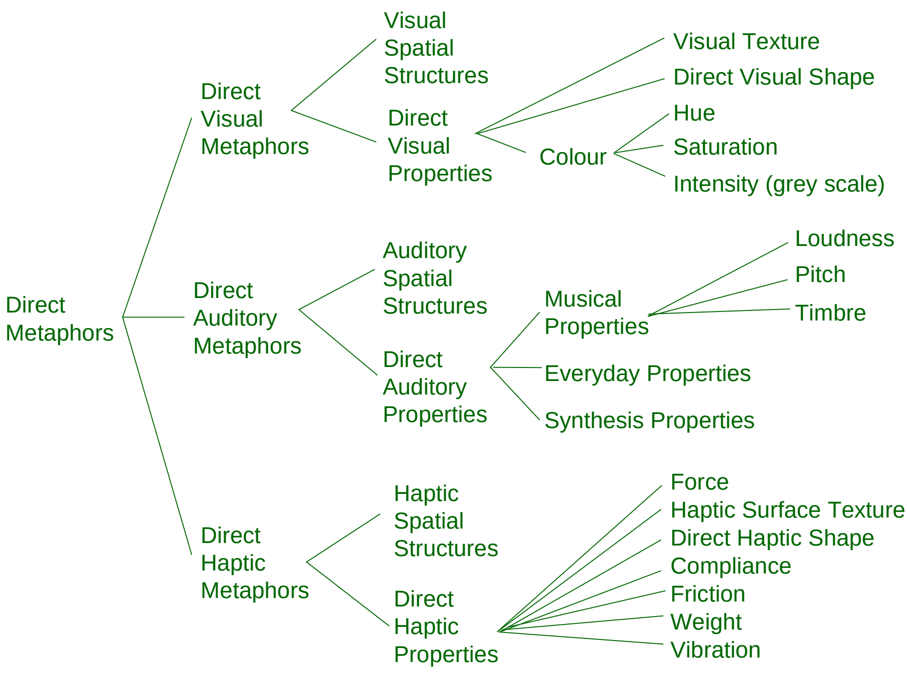


Figure 4-2 The hierarchy of concepts describing direct metaphors.

The classification used in the MS-Taxonomy for direct metaphors is intuitive. For direct metaphors it seems natural to discuss the interpretation of direct visual, auditory or haptic properties. Indeed, apart from the immature field of haptic display, these direct metaphors have frequently been used for the display of abstract information.

The concepts that describe direct metaphors are shown in figure 4.1. This chapter reviews the concepts for direct *visual* metaphors (section 4.2), direct *auditory* metaphors (section 4.3) and direct *haptic* metaphors (section 4.4). This review includes many concrete examples of abstract information display[1]. It provides a summary of existing literature and helps to explain the concepts that describe direct metaphors.

4.2 Direct Visual Metaphors

Direct visual metaphors use direct mappings from the attributes of data to the perceived properties of sight. These properties include colour hue, colour saturation and visual texture. Figure 4-3 shows the different types of direct visual properties.

Using direct visual properties to represent information has been well studied. Bertin described the basic properties of visual objects as *retinal properties* [Bertin 1981]. Bertin's *retinal properties* include the scale and orientation of objects. These concepts are dependent on the visual space and so are included in the MS-taxonomy as visual spatial metaphors (chapter 3). In this thesis, direct visual properties are independent of the visual space and so do not include scale and orientation. However, Bertin's other *retinal properties* are all concepts within direct visual properties[2]. They are:

- colour - hue
- colour - saturation
- colour - intensity (grey scale, value)
- visual texture
- direct visual shape.

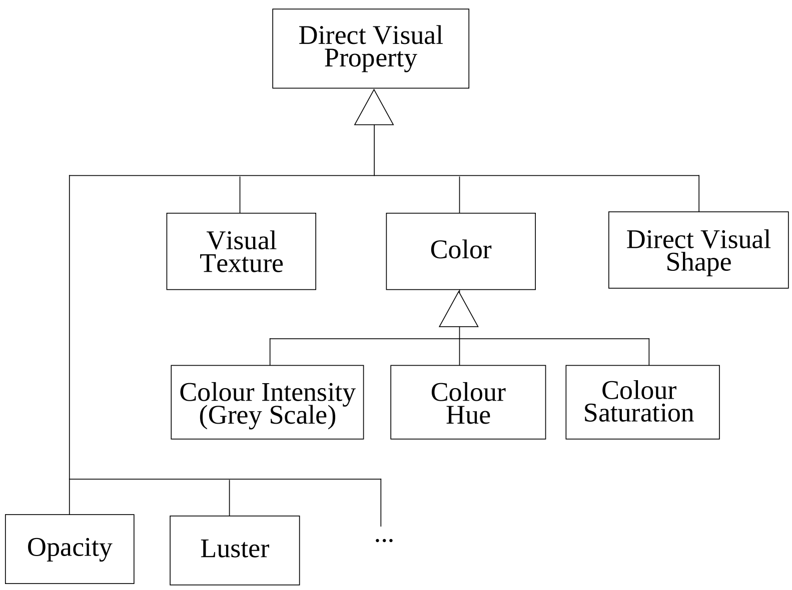


Figure 4-3 Direct visual metaphors use the direct visual properties of objects.

---

[1] A summary of the applications discussed in this chapter is available in Appendix A.

[2] The Card, Mackinlay et al. description of the visual design space also uses Bertin's *retinal properties*. It makes the further distinction of properties that undergo *automatic visual processing* [Card, Mackinlay et al. 1999]. Such visual properties can be interpreted perceptually and do not require cognitive processing. The Card, Mackinlay et al. description of the visual design space has been extended to cover the full multi-sensory design space and then compared with the MS-Taxonomy [Nesbitt 2001b] (See Appendix C).

The general concept of colour is discussed in section 4.2.1, along with hue and saturation. Colour intensity is discussed in section 4.2.2. Section 4.2.3 discusses the direct visual property of visual texture and section 4.2.4 discusses direct visual shape. Some more exotic direct visual properties have been suggested including luster [Healy, Booth et al. 1995] and opacity [MacEachren 1995] but they are not discussed here.

4.2.1 Colour - Hue and Saturation

It is estimated that we can perceive about two million different colours [Goldstein 1989, p134]. The range of colours that can be perceived can be described using different models. Two commonly used models are the RGB (Red, Green, Blue) colour space and the HSV (Hue, Saturation, Value) colour space (figure 4-4). The MS-Taxonomy uses the concepts of hue, saturation and intensity (value) to define three different direct visual properties.

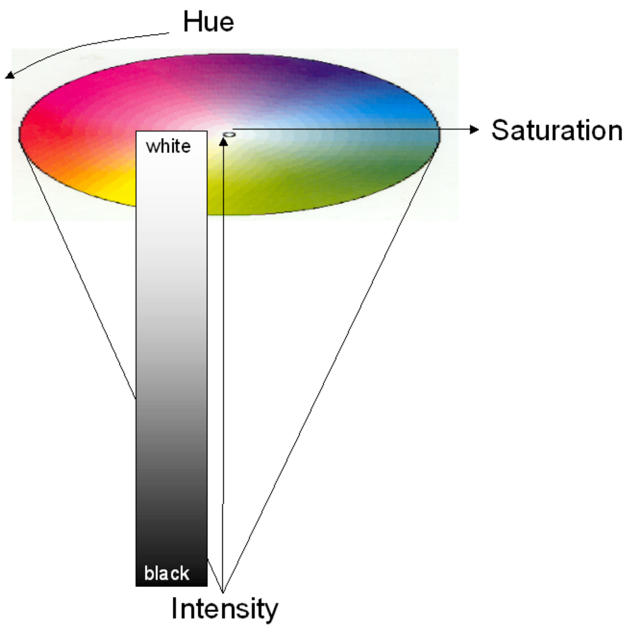


Figure 4-4 The HSV colour space is represented by an inverted cone.

The direct visual properties of hue, saturation and intensity are closely related and should not be considered independent. Indeed colour itself is often used to encode information without regard to whether hue, saturation or intensity is being used.

The perception of colour is complex and is known to be very dependent on the context of surrounding colours [Itten 1970] (figure 4-5). An object's characteristic colour also influences our perception of its colour [Goldstein 1989, p161]. For example, a user may perceive an apple shape to be redder than it actually is. The colour of objects can also have other effects on the display. For example, the size of objects can be perceived differently depending on their colour and their background colour [Keller and Keller 1993].

The colour space is continuous and can be mapped to a continuous range of data. For example, the spectrum naturally forms a continuum of hues from violet to red (figure 4-6). The continuous nature of the colour space makes it suitable for mapping to quantitative data. However, since no two viewers can be certain of perceiving the same colour, it is not effective for fine distinctions or exact measurements [Keller and Keller 1993]. Indeed any viewer's perception of colour is both uncertain and complicated [Tufte 1990]. Therefore, if a task requires accurate identification of data values it is more appropriate to use a categorical division of the spectrum for encoding data ranges (figure 4-6).

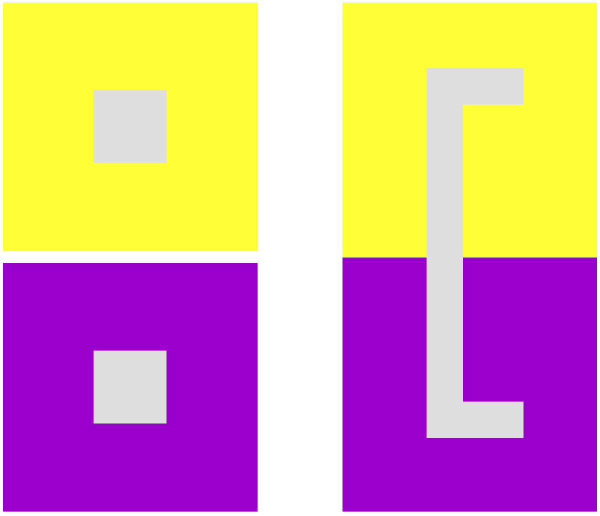


Figure 4-5 The perceived colour of foreground objects depends on the background colour [Itten 1970]

In `VisDB`, Keim and Kriegel visualize the results of a database query [Keim and Kriegel 1994]. Hue is mapped to a single continuous parameter called the *relevance factor*. This factor represents the relevance of each result to the initial data base query. A continuous variation of hue is also used in the `SeeSys` system. In this application hue is used to encode the size of components in a software system [Baker and Eick 1995].

One issue with the use of hue to encode data is that there is no natural ordering of hues [Card, Mackinlay et al. 1999]. The spectrum is better described as a *colour wheel* [Card, Mackinlay et al. 1999]. This circular nature of hues (figure 4-6) makes it useful for displaying continuous data that is also circular [Card, Mackinlay et al. 1999]. However, because there is no natural ordering of hues, this direct property is not recommended for displaying ordinal relationships. For example, in a map showing tidal variations for different parts of the world, a circular colour map is utilized. High tides are shown in red and low tides in purple, with a smooth variation around the colour wheel used to depict tides between the high and low mark [Card, Mackinlay et al. 1999].

However, hue is a very useful direct property for representing nominal relationships. For example, hue is often used to classify data and help users locate objects. Experiments in visual search and recognition tasks have shown that colour can assist the subjects speed and accuracy [Hardin 1990]. Users can detect about 24 steps of hue [Trumbo 1981].

In a financial decision support system to assist in tracking bonds in a portfolio, hue is used to distinguish different types of bonds [Wright 1995]. This is a typical use of to distinguish between nominal classes. The same approach is used in the `Filmfinder` application [Card, Mackinlay et al. 1999]. Points in a `scatterplot` are coded using different hues to indicate different film genres. So, for example, light blue points are westerns, pink points are comedies and white points are action movies. In `SemNet`, a `3D graph drawing` of knowledge bases, hue is used to encode different relationships between the nodes [Fairchild, Poltrock et al. 1988].

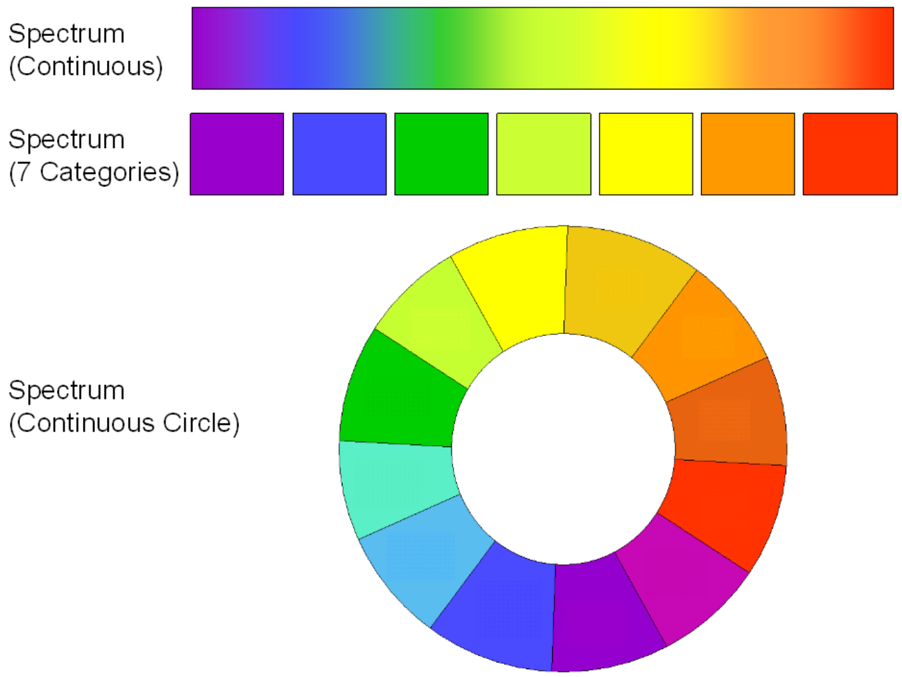


Figure 4-6 Hue is continuous and can be divided into categories. Hue is sometimes depicted in a colour wheel that can be used to represent a circular range of data.

When displaying nominal categories it is important to choose colours that are well delineated so that classes will not be misinterpreted. Maps of railway networks, such as the map of the London Underground, provide good examples of how distinct colours can be used to distinguish different classes (figure 4-7).

Another useful group of colours for specifying nominal categories is the *focal* colours. Compared to other colours the *focal* colours are the most readily picked out of piles of colour chips by young children, and they are more easily remembered and more quickly named by both adults and children [Hardin 1990]. There are 11 *focal* color categories: black, white, gray (middle) and red, yellow, blue, green, orange, purple, pink and brown [Rosch 1975] (figure 4-7). Note that non-*focal* colours such as red-pink, purple-pink, purple-blue, blue-green, yellow-green, yellow-orange, orange-red have been shown to be equally effective in searching for a known colour target [Boynton and Smallman 1990].

Hue may also have a natural association with some data. For example, in a model of global climate, the surface temperature of the ocean is mapped to spectral bands with blue representing cooler regions and red the warmer zones [Keller and Keller 1993, page 67]. In another application blue (cold) is related to low energy and red to high energy (hot) [Keller and Keller 1993, page 66].

These examples illustrate how two different colours can be used to create a continuous colour space (figure 4-8). In this case low data values are highlighted by the blue regions and high data values by the red regions. When only the end values of the data range are of interest, a third colour such as grey can be used to display data between the two ranges (figure 4-8). This is a typical colour map used to analyse petroleum seismic data [Nesbitt, Orenstein et al. 1997].

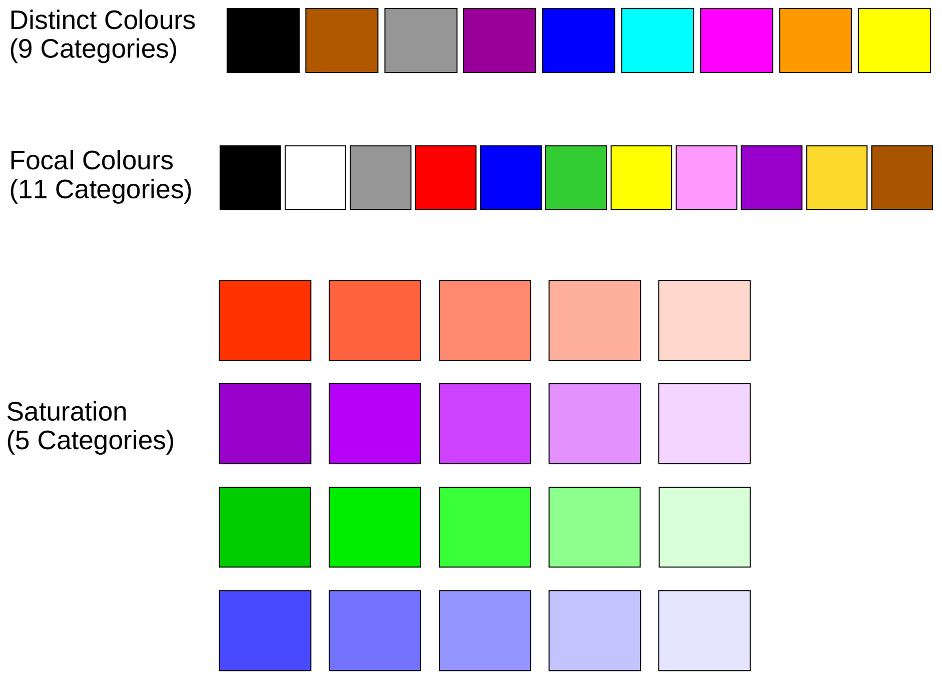


Figure 4-7 A group of distinct colours can be used to represent nominal categories. This figure shows the distinct colours used on the London Underground map to distinguish routes. Focal colours also form readily identified categories [Hardin 1990]. To specify ordinal categories a variation in the saturation or hue can be used.

Sometimes it is appropriate to highlight threshold values in the data. This can be achieved by creating a discontinuity in the colour space (figure 4-8). For example, in a `scatterplot,` where colour is used to display velocity, the possible values range from –1 to +1. A discontinuity is created around data with a value of 0. This creates a sharp contrast in the colour map and thus serves to highlight points at which the velocity changes sign [Keller and Keller 1993, page 110].

Even more complex encodings using both hue and saturation have been used (figure 4-8) For example, in a visualization of a shock front from the fluid dynamics domain, a matrix of colours formed by hues and saturation is used. In this example, density is mapped to hue, with vorticity mapped to saturation [Keller and Keller 1993, page 46].

Although hue has no ordering, the direct property of saturation is naturally ordered. Gradually increasing or decreasing colour saturation has been recommended as an effective way to show a continuum of change [Keller and Keller]. Once again it is not possible for users to accurately identify quantitative values mapped to saturation. However, this continuous range can be divided to create ordinal categories (figure 4-7). The number of steps of saturation that an observer can detect is about five [Trumbo 1981].

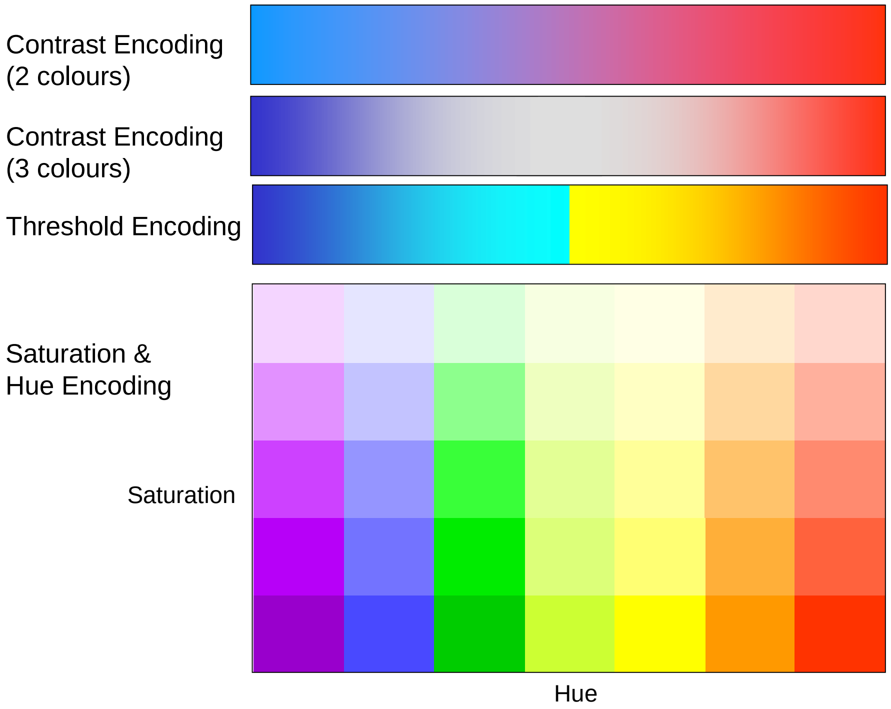


Figure 4-8 Some different strategies for mapping colour to data.

4.2.2 Colour - Intensity (Grey Scale, Value)

The *value* attribute of the HSV colour space is also known as intensity. Because intensity corresponds to a continuous colour scale between black and white it is also called *grey scale* (figure 4-9). Intensity is an ordered continuous property and can be used to represent continuous data or ordinal categories. However, once again, users cannot be expected to distinguish exact data values on a continuous scale and dividing intensity into a number of distinct categories is appropriate (figure 4-9). There is an implied ordering from light to dark and so it is not suitable for representing nominal categories.

In the `InfoCrystal` tool, visual glyphs are used to represent subsets of information [Spoerri 1993]. These visual spatial structures are augmented using intensity to represent continuous data, such as the number of elements represented by each visual glyph. In another example, different designs for nuclear reactors are compared [Beddow 1990]. Multiple attributes that impact on the design decision are displayed in a matrix and coded using three simple ordinal categories. Black represents a desirable value for an attribute, grey represents a neutral value and white signifies an undesirable value.

Intensity can also be combined with hue to create an ordered property (figure 4-9). For example varying the intensity of a single hue such as red can create a continuous colour space ranging from dark red to light red (figure 4-9). This continuous range of colour can also be divided into categories to present ordinal information. Once again this colour map has an implied ordering from light to dark and should not be used for displaying nominal data categories as the user may misinterpret an ordering.

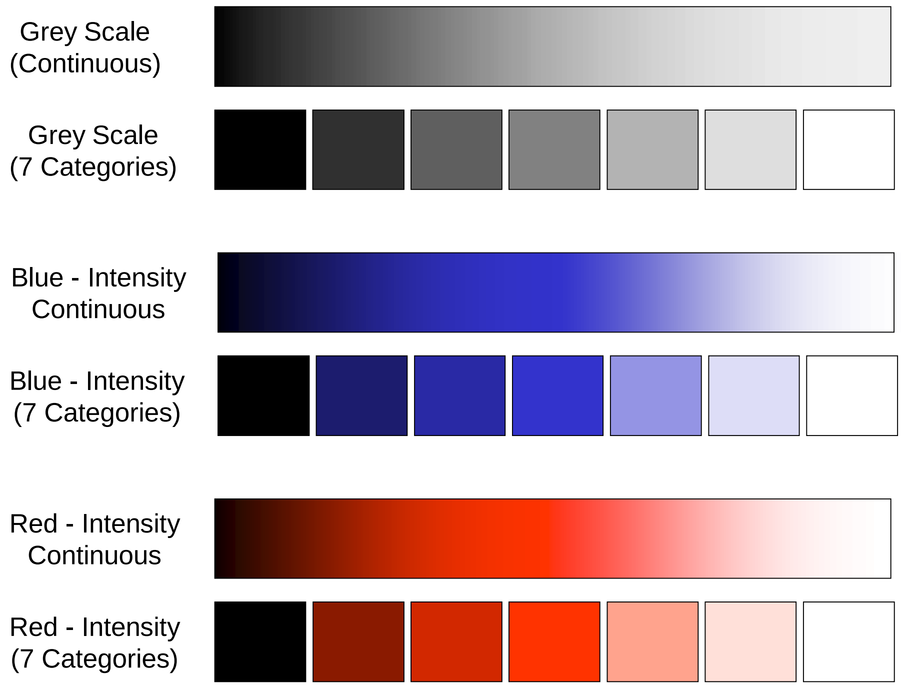


Figure 4-9 Intensity can be used to represent a continuous range of values or divided into categories. Hue can be combined with intensity to produce different colour maps.

In the `Treemap` application, hierarchical subdivisions of the screen space are used to represent a hierarchical graph [Johnson and Shneiderman 1991]. Colour with different hue and intensity can then be used to fill the subdivisions based on other data attributes. For example, a `Treemap` was used to display file systems where the hue of each subdivision was coloured according to file type and the intensity was based on the files date of modification. Thus all red boxes may represent text files with dark red boxes corresponding to recently opened files and light red boxes indicating files that have not been accessed recently.

4.2.3 Visual Texture

Visual texture is a direct property associated with the size, shape and arrangement of elements on an object's surface (figure 4-10). There is no intuitive mapping from visual texture to quantitative data. Some visual textures can be ordered from rough to smooth on the basis of *spatial frequency*[3]. However, it is not always easy to identify an ordering with visual texture (figure 4-10) and the best use of this direct property is to represent nominal categories.

For example, in a visualisation of search results, visual texture was used to identify which of the search criteria were satisfied [Spoerri 1993]. In a system designed for searching an image database the user can navigate a display using shapes with visual textures [Tian and Taylor 2000]. In this example, the visual textures are extracted from features in the original image.

Closely related to the use of visual texture are visual patterns (figure 4-10). This direct property can also be used for distinguishing nominal categories. For example, the bars

[3] *Spatial frequency* is the number of times an element repeats in space [Sekuler and Blake 1990, p153].

of a `histogram` can be encoded with different visual patterns (figure 4-11). One potential problem with visual patterns is that they may produce unwanted visual anomalies called Moiré effects [Tufte 1983] (figure 4-11).

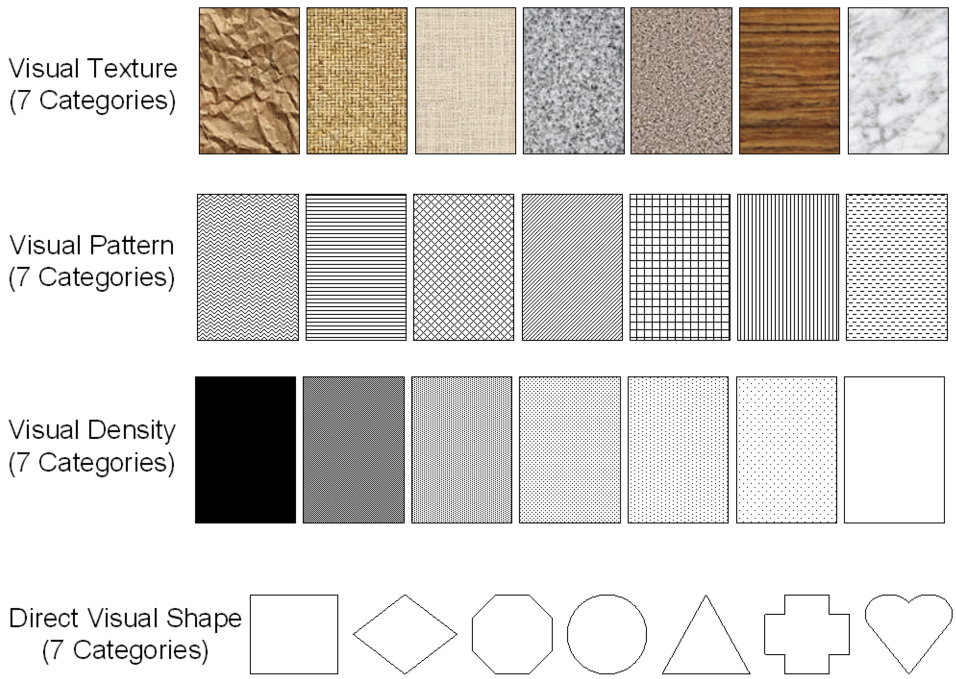


Figure 4-10 Visual texture, visual pattern and direct visual shape can be used to represent nominal categories. Visual density can be used to encode ordinal categories.

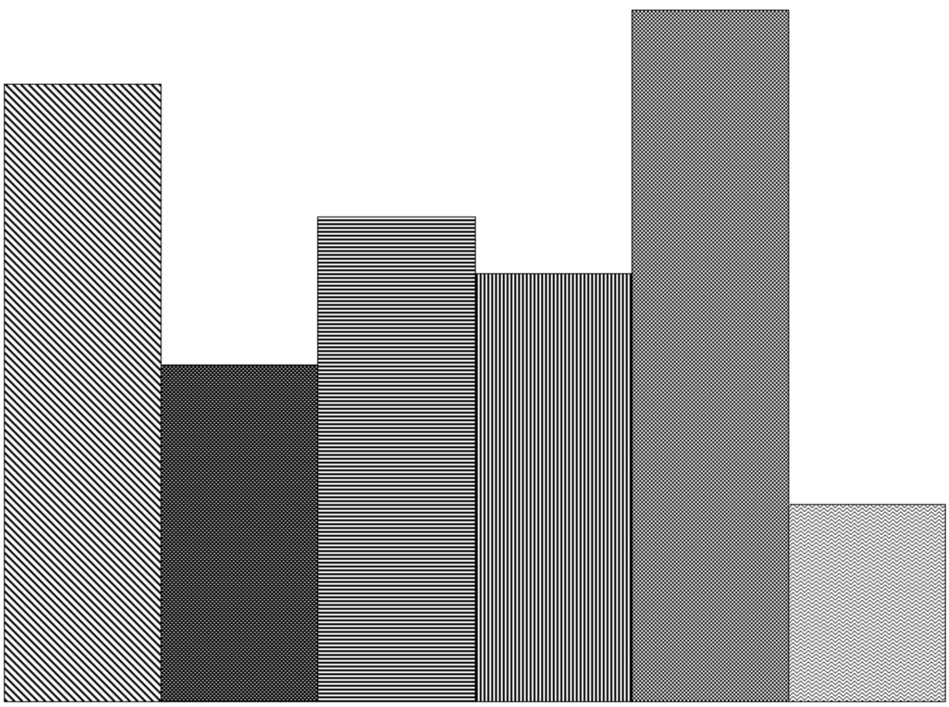


Figure 4-11 On this histogram, visual patterns distinguish nominal categories. The visual anomalies are called Moiré effects.

Visual density is another direct property that is related to both colour intensity and visual texture (figure 4-10). However, unlike visual texture, visual density does have an ordering from dense to sparse and can be used for representing ordinal categories. One previous application of visual density is for representing ranges of statistical data on a `data map` (figure 4-12).

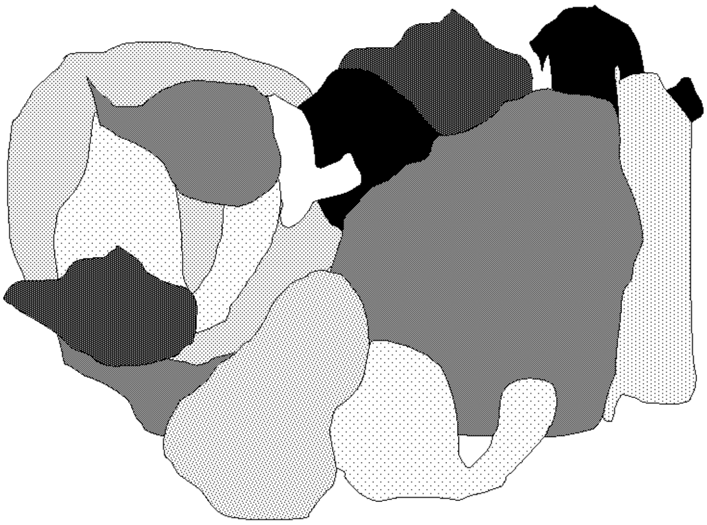


Figure 4-12 On this data map visual density is used to represent a number of ordinal categories.

4.2.4 Direct Visual Shape

Direct visual shape is an unordered direct property that can be used to represent nominal categories. Because direct visual shapes have no implied order they are not suitable for ordered classes (figure 4-10).

The concept of direct visual shape is used to describe forms that are directly recognised. It is known that some neurons in the cortex respond directly to both simple forms and some more complex shapes [Goldstein 1989, p111]. This type of form perception is distinguished from visual glyphs. In visual glyphs data relationships are conveyed in the substructure of the shape. With direct visual shape the form is identified with a particular class.

In an application for analysing supply chains, direct visual shapes depict nominal categories [Chuah, Roth et al. 1995]. For example, rectangular bars are used to represent centres in need of supplies and cylinders are used to indicate the supply centres [Chuah, Roth et al. 1995]. In the `InfoCrytsal` application a range of direct visual shapes are used to encode the number of criteria satisfied in a search [Spoerri 1993]. For example, a circle represents one criteria satisfied, a rectangle represents two criteria satisfied and a triangle represents three criteria satisfied

In an application called SequenceWorld, direct visual shapes are used to represent genetic sequences. The display is designed to assist users perform search tasks in a genetic database [Rojdestvenski, Modjeska et al. 2000]. In another example, a document analysis identifies different document types by a variety of 3D crystal shapes [Börner 2000].

4.3 Direct Auditory Metaphors

Direct auditory metaphors use direct mappings from the attributes of data to the perceived properties of sound. The use of direct auditory properties for representing abstract data is a relatively recent field of study. Many of the perceived properties of sound are not well understood [Madhyastha and Reed 1994]. The direct auditory properties are less generally agreed on than the *retinal properties* of Bertin [Blattner, Papp et al. 1994]. The commonly used properties of sound include loudness, pitch and timbre. Figure 4-13 shows the different types of direct auditory properties. Figure 4-14 shows the relationship between the physical properties of sound and the perceived properties.

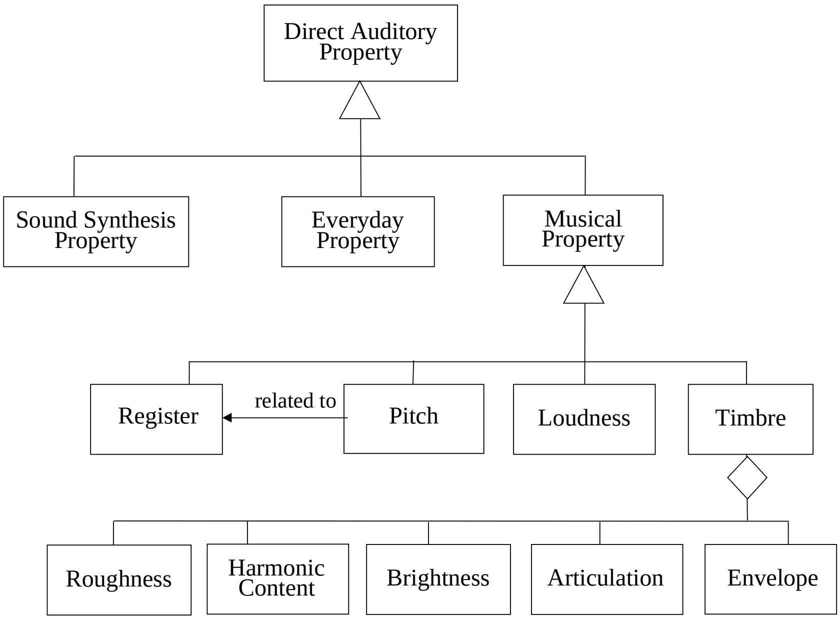


Figure 4-13 A UML diagram showing the types of direct auditory properties. The direct auditory properties are usually described in terms of musical properties.

The direct auditory properties are not independent or orthogonal. For example, the pitch of the sound affects the perceived loudness of the sound. Furthermore, both pitch and loudness are not equally prominent to the listener [Bly 1994]. Bly describes it as a significant challenge to determine how to map data to an audio representation and notes that currently, there is no known method for determining a best mapping [Bly 1994].

The direct auditory property of pitch is discussed in section 4.3.1. Loudness is discussed in section 4.3.2 and timbre in section 4.3.3. These direct auditory properties have also been referred to as *musical properties* [Gaver 1994]. This is because they are properties that are interpreted by directly listening to the qualities of the sound itself. This contrasts with the concept of *everyday listening* where sound properties are interpreted in terms of the objects and events that generate the sounds [Gaver 1994]. *Everyday properties* are discussed in section 4.3.4.

Some sound synthesis algorithms have been developed for displaying abstract data. In some cases the sound properties can be described in terms of the parameters that define these algorithms. Sound synthesis properties are discussed in section 4.3.5.

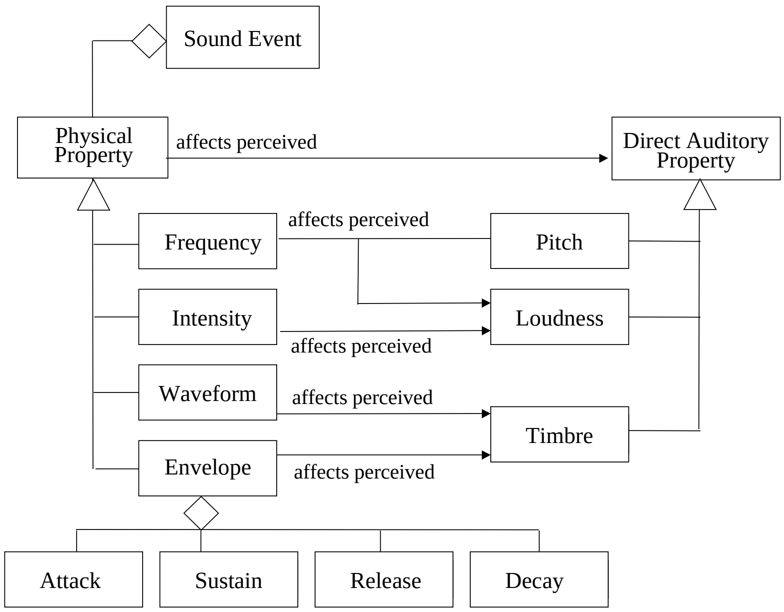


Figure 4-14 A UML diagram showing the relationship between the physical properties of a sound event and the direct auditory properties.

4.3.1 Loudness

The direct auditory property of loudness is closely related to the *intensity* or *amplitude* of the original sound event (figure 4-15). The perceived loudness is dependent upon both the pressure of the sound wave (measured in decibels) and the frequency of that wave [Goldstein 1989]. For example, in low intensity sounds both the higher and lower frequency tones may not be as loud as midrange frequencies. This is an important consideration when mapping abstract data to the property of loudness. Sounds of equal intensity but with different frequencies may not be perceived to have the same loudness. The relationship of perceived loudness, sound pressure and frequency is complex and is usually described by *audibility curves*. (figure 4-15).

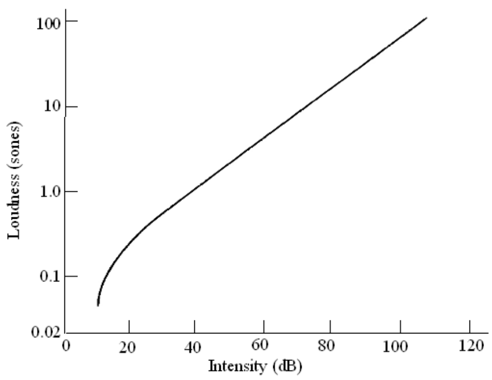


Figure 4-15 The relationship between sound intensity and perceived loudness for a 100 Hz tone [Goldstein 1989, page 353].

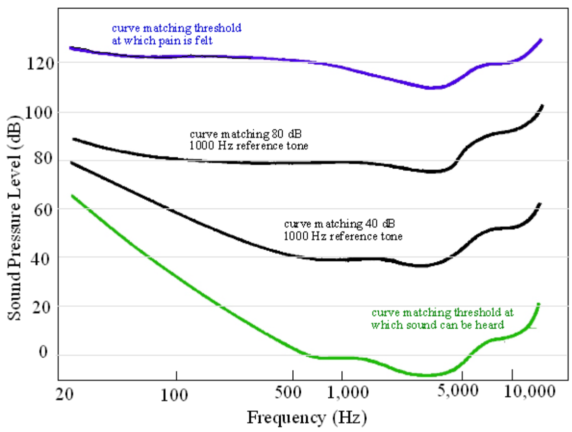


Figure 4-16 Audibility curves describe the relationship between sound frequency, sound intensity and perceived loudness. Each curve indicates the intensity (dB) required to create the same perceived loudness for sounds at different frequencies [Goldstein 1989, page 351].

Loudness is a continuous property and can be used to display a continuous range of quantitative data. However, the user should not be expected to accurately identify a loudness level. For this reason loudness may be better suited to displaying categories. Loudness has a naturally ordering from quiet to noisy, and Kramer describes a natural metaphorical association so that louder is perceived as more [Kramer 1994b]. Therefore loudness can be used for displaying ordinal categories.

Despite the natural association of loudness levels with high and low data values there is a single fundamental issue with using loudness to display data. That is, the intensity of the displayed sound is often under the control of the user. If the user reduces the intensity it may completely remove quiet sounds from the display. It may also have other unexpected effects. For example, the audibility of high and low frequency sounds could be reduced if the sound intensity is reduced.

Another issue to be aware of is that if loudness is mapped to a data attribute, then unexpected data values can result in the sudden occurrence of a loud sound. These alarming sounds may be useful if the intent is to attract the user's attention. However, loud sounds can also be irritating and distracting if the corresponding data points are not significant.

Madhyastha and Reed used loudness to display the cost of housing in a comparison of five different cities [Madhyastha and Reed 1994]. This application uses the natural ordering of loudness so that more expensive houses sound louder. Bly displayed six-dimensional random normal data using a number of direct auditory properties including loudness [Bly 1994].

The `Sonnet` application is designed for debugging of software programs [Jameson 1994]. Modulations of loudness are used to help the user track progress through software functions. For example, a note may sound as a function is being entered. Gradually this note reduces in loudness to indicate progress through the function. So a note that did not decrease in loudness may indicate that the function is not completing as expected.

Jackson and Francioni produced an auditory display for tuning algorithms on parallel computers [Jackson and Francioni 1994]. On parallel computers any idle processor indicates inefficiency in the way that resources are being used. In this application, any idle processor displays a sound. The loudness of the sound is proportional to the length of the processors idle period. So processors that are idle for long periods make loud sounds and these loud sounds attract the user's attention.

4.3.2 Pitch

The direct auditory property of pitch refers to how low or high a tone sounds. Simple tones contain sound of a single frequency and in this case the perceived property of pitch is closely related to the frequency of the sound. Hence, low frequency tones sound low in pitch and higher frequency tones sound high in pitch. However, most sounds are more complex than pure tones and are made up of multiple frequencies. In this case the perceived pitch is related to what is called the *primary* or *fundamental frequency* [Goldstein 1989] of the sound.

Related to the concept of pitch is the concept of register (figure 4-13). Register can be though of as a range of notes. The keys on a piano ascend from lower pitches to higher pitches; however these notes can also be ordered in registers or octaves. So the notes (A, B, C, D, E, F, G) repeat each octave. For example, all A notes have a similar sound. This quality of sameness across registers is called *tone chroma* [Goldstein 1989]. The similarity of notes with the same *tone chroma* can be used to display hierarchical information. For example, both the note and the register can be used to describe two data attributes.

Pitch is a continuous property and can be used to display quantitative data. However, it is unlikely that users will accurately interpret pitch. The ability to recognise the exact pitch of a sound is described as *perfect pitch*. This is a quite rare ability even for musicians [Sekuler and Blake 1990]. Musicians, however, do commonly develop what is described as *relative pitch*. This allows them to identify the difference between two notes or what is called *tonal intervals*. In general people are better at recognising differences between pitches than absolute pitches. For example, melodies are recognised because of the relation between the pitch of notes rather than their absolute pitch [Sekuler and Blake 1990]. One way to assist users in interpreting pitch is to provide a reference tone. This approach has been described as using a *beacon* [Kramer 1994b].

Despite the problems of users recognising exact pitch, this direct auditory property has often been used to display quantitative data. Bly used pitch to display a continuous data attribute from a statistical data set [Bly 1994]. Pitch is also used to display a quantitative variable in the `sonic graph` [Mansur, Blattner et al. 1985]. The `sonic graph` displays a series of pitches over time. For example, in a xy-plot, the x-axis is represented using time, while the y-axis is represented by pitch[4]. The `sonic graph` has been shown to be effective as it relies on the user distinguishing changes in pitch rather than absolute pitch values. In another application pitch was used to signal blood pressure in an auditory model of fluid flow through an artificial heart [McCabe and Rangwalla 1994]. Fitch and Kramer also used pitch to display blood pressure in an application for monitoring a patient in an emergency room [Fitch and Kramer 1994].

Because users cannot accurately interpret pitch it may be better mapped to categories. For example, an alternative is to represent the data values in ranges and map these to categorical ranges in pitch. Pitch is a naturally ordered direct auditory property. Indeed pitch describes the perception we have of sound being ordered from low to high tones. Hence pitch is well-suited to representing ordinal categories.

There are a number of examples where pitch has been used to represent ordinal categories. Jackson and Francioni suggest a range of pitches for representing the percentiles from 0-100% [Jackson and Francioni 1994]. In this case low notes correspond to low percentages and high notes to high percentages. Madhyastha and Reed use pitch categories to encode the processor numbers in a parallel computer [Madhyastha and Reed 1994]. In this application to monitor computer performance, messages sent between processors have their origin and destination encoded by the pitch.

[4] Note that although the sonic graph and traditional lineplot both display the relationship between two abstract variables they are very different in terms of the MS-Taxonomy. The sonic graph uses time to display the data and so is a temporal metaphor while the traditional lineplot is a spatial metaphor.

Kramer has suggested some natural metaphorical associations for pitch. For example, a higher pitch relates to more, a higher pitch relates to up and a higher pitch relates to faster [Kramer 1994b].

There are also some further issues with using pitch to represent data. Changing the intensity of a sound can change the perceived pitch of the sound. This relationship is complex. At less than 2000 Hz an increase in intensity increases the perceived pitch. At greater than 3000 Hz an increase in intensity decreases the perceived pitch [Brewster, Wright et al. 1994]. The perceived pitch is also dependent on the timbre of the sound. For example, sounds that are bright in timbre (have more high harmonics) are perceived as higher in pitch than less bright sounds of the same fundamental frequency [Stuart 1996].

4.3.3 Timbre

Timbre is a direct auditory property that is difficult to define. A dictionary definition describes timbre as *"tone colour or a quality of a sound that distinguishes it from other sounds of the same pitch and volume"* [Wilkes and Krebs 1990]. For example, timbre is the property of sound that makes it possible to distinguish between different musical instruments. When a violin and a guitar produce a sound that is the same pitch and loudness it is still possible to determine which instrument made which sound. This is due to the different timbre of the sounds made by different instruments. It is also the property of timbre that allows us to recognise familiar voices on the telephone.

Timbre is not continuous and so cannot be used to represent quantitative data. Timbre can however be used to represent nominal categories. Since timbre has no natural ordering it is not suitable for encoding ordinal data.

Timbre is itself a complex property which is made of a number of components (figure 4-13). These include the *harmonic content* and the *attack* and *decay* rate of the sound. The *harmonic content* refers to the spectrum of different frequencies and their magnitudes that combine to create the sound (figure 2-42). The *attack* refers to the way the sound commences and the *decay* to the way it declines at the end. The overall shape of the sound from *attack* to *decay* is also known as the *envelope*.

Models of multidimensional timbre space have been developed [Williams 1994]. Models of timbre have also been used to represent some quantitative data. Bly, for example, used the *envelope* of the timbre to help display normally distributed data [Cohen 1994]. Jameson used modulations in the timbre to display debugging information in an audio-enhanced software debugger [Jameson 1994].

A more typical use of timbre is to distinguish nominal categories of data. For example, an auditory display was developed to present information about five different cities [Madhyastha and Reed 1994]. The types of climate in each city are categorised and each one is represented by a different timbre. In another example, Jackson and Francioni use different timbres to distinguish between different types of events in an auditory display designed to study communication between parallel processes [Jackson and Francioni 1994]. Fitch and Kramer used timbre to represent carbon dioxide levels for physiological monitoring [Fitch and Kramer 1994].

4.3.4 Everyday Properties

The interpretation of loudness, pitch and timbre is what Gaver describes as *musical listening* [Gaver 1994]. The MS-Taxonomy uses these musical properties of loudness, pitch and timbre to describe direct auditory metaphors. However, there are other ways to describe direct auditory properties. For example, Gaver contrasts *musical listening* with *everyday listening*. Where m*usical listening* is concerned with hearing the properties of sound, e*veryday listening* is concerned with hearing the properties of the sound source. For example, when a container is tapped a number of questions can be asked. What material is the container made of? Does the container have contents? Is the container full or empty? How hard was the container struck? Thus interpreting the sound reveals the properties of the sound source. This is *everyday listening*.

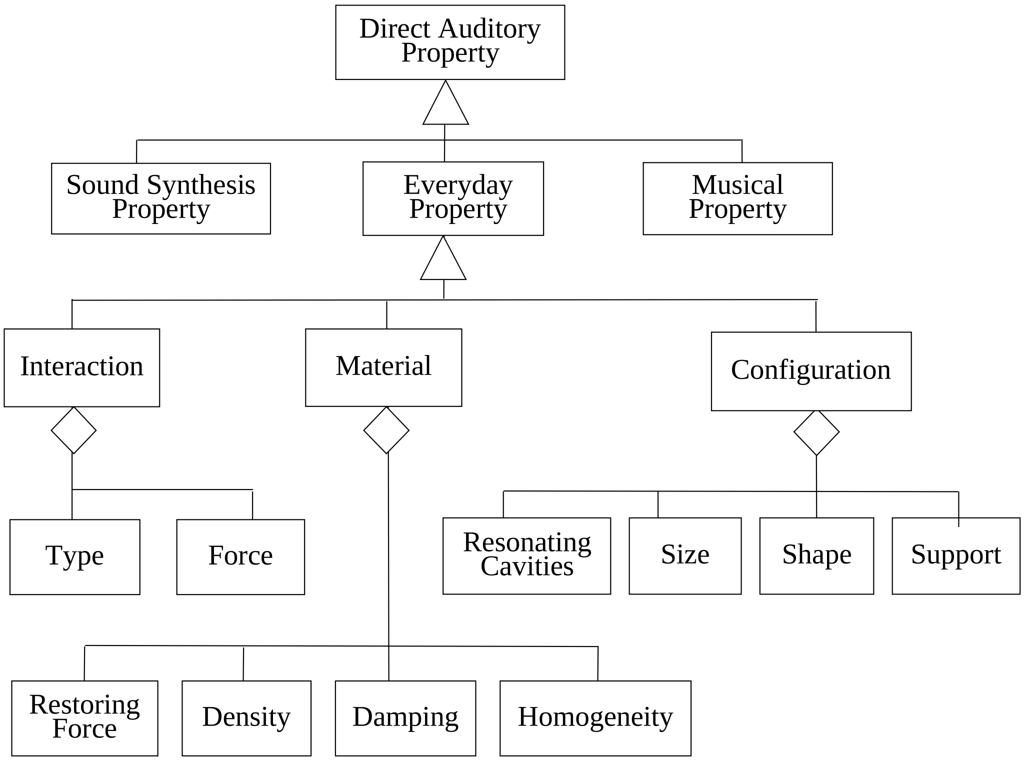


Figure 4-17 Everyday properties describe sounds in terms of the sound source.

Gaver developed `auditory icons` as sound events that allow abstract data to be mapped to the properties of the sound producing event rather than the properties of the sound [Gaver 1986]. For example, selecting a file in a computer screen may sound like an object being tapped. The object's material could be used to identify the type of the file and the size of the file could be mapped to another parameter such as the size of the object [Gaver 1994].

`ShareMon` is an application that uses `Auditory icons` to inform a user about background events on a computer [Cohen 1994]. File sharing activities are characterised by a walking sound. However the actual sound also depends on the amount of CPU time being used for file sharing. At higher levels of file sharing the display produces a jogging sound. At very high levels of file sharing the display produces a running sound.

Gaver characterises everyday sounds according to the interaction with a source object. The source objects are described in terms of their material and configuration (figure 4-17). One issue with using this type of mapping is that appropriate algorithms for synthesising the sounds need to be developed. With this aim in mind, Gaver has developed synthesis algorithms for a number of types of sound. These include impact sounds, bouncing sounds, breaking sounds, spilling sounds, scraping sounds and machine sounds [Gaver 1994].

4.3.5 Sound Synthesis Properties

The MS-Taxonomy uses the perceptual properties of loudness, pitch and timbre to describe direct auditory metaphors. The previous section discussed an alternative description called everyday properties. Another way to consider this part of the design space is to describe it using the actual parameters used to synthesise the display sound.

When generating display sounds, properties such as loudness, pitch and timbre can generally be controlled. For example, sampled or synthesised sounds can be altered in intensity or frequency to adjust loudness and pitch. While sound samples can usually only be adjusted by a few parameters, some synthesized sounds can be controlled by multiple parameters [Wenzel 1994].

For section 4.3.4 described everyday sounds. Gaver describes a number of sound synthesis algorithms for producing natural sounds that rely on mapping data to the attributes of the sound event [Gaver 1994]. A number of other sound synthesis algorithms have also been developed that may have numerous parameters controlled by the data. For example, Kramer suggests creating a sound field with input data controlling randomness, density, loudness and timbre of sound [Kramer 1994b]. Other examples of sound synthesis methods include frequency modulation, non-linear distortion and granular synthesis [Smith, Pickett et al. 1994].

For example, *granular synthesis* is a sound synthesis algorithm that produces a characteristic sound of a Geiger counter. Data attributes can be used to control *grain frequency* and *density* to synthesise the familiar Geiger counter sound. While it was originally used to display radiation levels, this synthesis algorithm has been adapted to display other types of abstract data. For example, Barass and Zehner use this technique for displaying petroleum exploration data [Barass and Zehner 2000].

The MS-Taxonomy chooses to describe direct auditory metaphors using the perceptual dimensions of pitch, loudness and timbre. However a designer of sound displays may which to consider everyday properties or sound synthesis properties as adjuncts to the design space.

4.4 Direct Haptic Metaphors

Direct haptic metaphors use direct mappings from the attributes of data to the perceived properties of the haptic sense. These properties include surface texture, force and compliance. Figure 4-18 shows the different types of direct haptic properties that are principally associated with the *tactile* sense. Figure 4-19 shows the different types of direct haptic properties that are principally associated with the *kinaesthetic* and *force* sense. Some of the direct haptic properties, such as compliance and friction, require the combined perception of *tactile, kinaesthetic* and *force* stimuli.

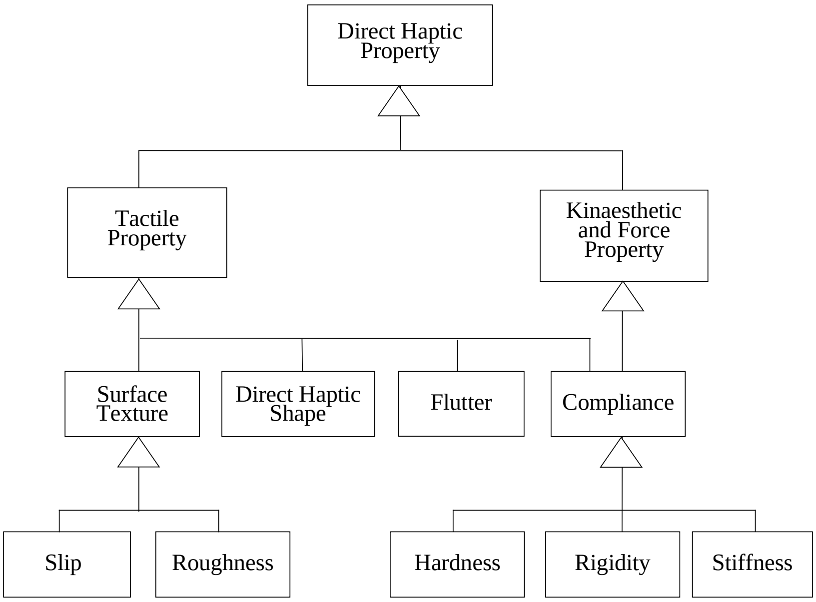


Figure 4-18 Direct haptic properties associated with tactile stimuli.

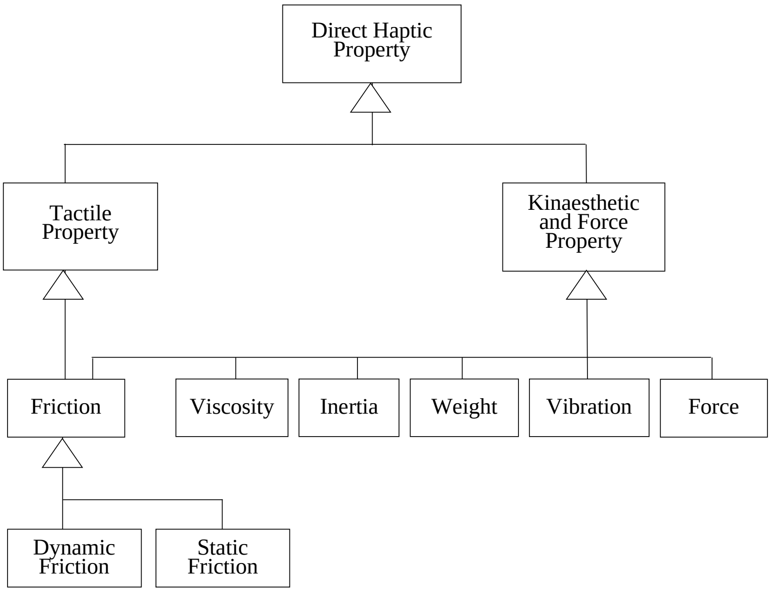


Figure 4-19 Direct haptic properties associated with kinaesthetic and force stimuli.

To display direct haptic properties a force feedback device such as the Phantom™ [Salsibury and Srinivasan 1997] can be used. These devices have only been commercially available since about 1993. Because of the recent nature of the technology there are very few examples where direct haptic properties have been applied to the display of abstract data.

Force feedback interfaces display a spatial and temporal array of forces that mimic a range of haptic sensations. For example, forces displayed to a finger point can mimic a combination of *tactile, kinaesthetic* and *force* stimuli.

The following direct haptic properties are considered for displaying information:

- force (section 4.4.1)
- surface texture (section 4.4.2)
- direct haptic shape (section 4.4.3)
- compliance (section 4.4.4)
- viscosity and friction (section 4.4.5)
- inertia and weight (section 4.4.6)
- vibration and flutter (section 4.4.7)

4.4.1 Force

The direct haptic property of force is ordered and continuous. However, it is not possible for users to distinguish forces accurately. To allow a user to distinguish between two forces requires about a 7% difference between them [Srinivasan and Basdogan 1997] or a step size of 0.5 Newtons [Durlach and Mavor 1995]. This suggests that force is most appropriate for displaying ordinal categories of data.

The direct haptic property of force is often associated with the concept of a field[5]. For example force fields were used to represent a flow field that was simulated by the drag forces on a small probe [Ottensmeyer and Salsibury 1997]. A force field was also used to represent the fluid flow in a blast furnace [Nesbitt, Orenstein et al. 1998].

Forces can also be used to represent constraints in data. For example this approach was used to help geologists model bodies of ore [Veldkamp, Turner et al. 1998]. Constraining forces were also derived from the data in an application designed to help geophysicists interpret features in petroleum exploration data [Nesbitt, Orenstein et al. 1997].

4.4.2 Surface Texture

The direct haptic property of surface texture is an ordered sensation from smooth to rough. The haptic sense is very accurate for discriminating between different surface textures and can outperform vision [Jansson, Fänger et al. 1998]. The *tactile* sense is used to feel much of the richness of fine-grained texture. A force feedback display can only simulate a gross range of textures [Oakley, McGee et al. 2000]. However surface texture provides the opportunity to represent a selection of ordered sensations from smooth surfaces (slip) to coarse surfaces (roughness). However, users should not be expected to accurately interpret a texture quantity and so surface texture is most appropriate for ordinal categories of data.

To our knowledge, surface texture has not been previously used to display abstract data. However, the `Sandpaper` system was developed for experimenting with feeling surface texture. Subjects were able to arrange a selection of simulated sandpaper surfaces in order of roughness [Minsky, Ouh-yong et al. 1990]. Green also developed a system that could simulate rigid textured surfaces. The aim of this work was to allow for remote texture sensing of soil and rocks [Green 1997].

---

[5] In the MS-Taxonomy a field is a spatial structure. Fields are characterised by a direct property that varies according to position in the display space. A force field is characterised by forces that vary according to haptic position in the haptic display space.

4.4.3 Direct Haptic Shape

The concept of direct haptic shape describes shapes that can be directly recognised by touch. This is in contrast to larger shapes that may require some time to explore and interpret. Therefore direct haptic shapes require minimal temporal integration and do not involve interpreting relationships between parts of the structure.

Direct haptic shapes are closely related to the concept of direct visual shapes. They are both unordered direct properties that can be used to represent nominal categories. Because direct haptic shapes have no implied order they are not suitable for ordered classes.

Caution needs to be taken if both direct haptic shape and direct visual shape are used in the same display. It has been shown that the visual representation of an object can affect the perceived haptic shape of the object [DiFranco, Beauregard et al. 1997].

The *tactile* sense can detect very small shapes. For example, a two micron high dot on smooth glass and or points which are 1 millimetre apart can be detected [Durlach and Mavor 1995]. A fingertip can normally detect direct haptic shapes in the order of a few millimetres. Moving the fingertip can allow surface shapes to be quickly integrated over time and is effective for small shapes.

Direct haptic shape has not previously been used to represent abstract data. However, haptic shape cues were used by Miller to augment a user-interface [Miller 1998]. Small haptic dimples were used to mark each item on the user-interface menu and tiny ridges separated the screen widgets. In a very early application of haptic technology called the `nanoManipulator`, users could feel small deviations in a surface shape [Taylor, Robinett et al. 1993] representing data recorded from a scanning probe microscope.

4.4.4 Compliance

Compliance can be described as the hardness of a surface, or the stiffness or rigidity of an object. This direct haptic property is ordered and continuous. A user cannot discriminate exact values of compliance and about a 7% change [Srinivasan and Basdogan 1997] is required to represent categories. Because this property is perceived as ordered it is most appropriate for displaying of categories of ordinal data.

The visual form of the object affects the perceived compliance of an object. For example, subjects perceived springs of equal stiffness to be different based on their visual models [Srinivasan, Beauregard et al. 1996]. In another experiment subjects misjudged the stiffness of objects when required to rank objects whose apparent visual order did not match their haptic order [DiFranco, Beauregard et al. 1997].

Compliance has not been used abstract information display. However, compliance is often used for simulated for real world models such as CAD surfaces [Larsson 1997]. It is also used to help model interaction with anatomical surfaces in medical simulators. For example, in a surgical trainer the contact between tissues and a needle are modelled with compliance [O'Toole, Playter et al. 1997].

4.4.5 Viscosity and Friction

The direct haptic properties of viscosity and friction are similar and can be modelled with force feedback. Viscosity and friction are both ordered and continuous properties. Again, it is not possible to accurately interpret different values of viscosity and friction. For example, a 12% difference is required to interpret different values of viscosity [Srinivasan and Basdogan 1997].

Friction depends on both the properties of the object and the surface over which the object is moving. Dynamic friction depends also on the way the surfaces interact when the object is moving. Viscosity is closely related to dynamic friction. Viscosity, static friction and dynamic friction have previously been modelled for applications in medical training [Huang, Qu et al. 1998].

Ruspini, Kolarov et al. developed a haptic rendering framework to allow display of friction and viscosity [Ruspini, Kolarov et al. 1997]. Viscosity display has also been developed for surgical simulators [Mor 1996]. In terms of abstract data display, Aviles and Ranta proposed using viscosity to represent attributes of geological data such as porosity [Aviles and Ranta 1999]. In a haptic simulation of soil, the user could move different shaped plough blades through soil with a set of properties that included friction [Green and Salsibury 1998].

4.4.6 Inertia and Weight

The direct haptic property of weight is again ordered and continuous and cannot be estimated accurately. To allow a user to distinguish between two weights requires about a 20% difference between them [Srinivasan and Basdogan 1997].

Inertia is the tendency of a body to resist acceleration. The weight of an object and its inertia are closely related. Both weight and inertia are most appropriate for displaying ordinal categories of data.

The visual model also impacts on the perception of weight. For example, in an experiment where subjects sorted objects by weight the subjects make systematic errors in discriminating objects of similar weights when their size was not related to their weight [Von der Heyde and Häger-Ross 1998]. Hence bigger objects were judged to be heavier than smaller objects of the same weight.

The use of weight and inertia for displaying abstract data has not been explored. However, in a system where users could assemble virtual Lego™ blocks, the users could perceive the weight of the blocks [Young, Chen et al. 1997]. For example, when multiple blocks were joined together the user perceives a weight determined by the number of blocks. Inertia was used to help users position a cutting plane through a volume of data [McLaughlin and Orenstein 1997]. In this application, a cutting plane through the data could be in two states, either stationary or moving. When the plane was stationary the user had to overcome a level of inertia to make it move [McLaughlin and Orenstein 1997].

4.4.7 Vibration and Flutter

The direct haptic property of vibration is ordered according to the frequency and amplitude of the stimulus. Vibrations can be detected with an amplitude from 1 micron to several millimetres [Howe 1996] and with a frequencies between 300-10,000 Hz [Durlach and Mavor 1995]. Flutter is the sensation caused by a light tapping on the skin and has very low frequencies of 3-40 Hz [Howe 1996]. This sensation is differentiated from vibration because a different anatomical mechanism is used to detect it.

Vibration requires about a 10% change in the number of pulses per second before a change in stimulus is detected [Durlach and Mavor 1995]. While both flutter and vibration are continuous they cannot be judged accurately. Hence these direct haptic properties are recommended for presenting ordered categorical data.

Vibration has previously been used to represent the differences between two data values at the same spatial location [Nesbitt, Orenstein et al. 1998]. Discrepancies between two abstract data values at the same location in space were mapped to vibration. The user could search the space for large discrepancies by feeling for areas of large vibration. In another application vibration was used to warn users [Oakley, Brewster et al. 2000]. This application allowed remote users to collaborate on a document. A user working in one area of the document would feel a small vibration to indicate that another user was approaching.

4.5 Conclusion

This chapter has discussed in detail the general concepts that describe direct metaphors. The concepts of direct *visual* metaphors (section 4.2), direct *auditory* metaphors (section 4.3) and direct *haptic* metaphors (section 4.4) were discussed. These concepts of the MS-Taxonomy were explained with reference to previous work. Note that there are few examples of the use of direct *haptic* metaphors for abstract data display.

Direct metaphors map data directly to a sensory property. These properties are directly related to the receptors that make up each sensory channel. It is not possible for users to make accurate judgements about sensory properties. The accuracy varies between direct properties. Many direct properties are continuous and ordered and can be used for displaying quantitative data. However, it cannot be assumed that a user will make an accurate judgement of the value of a property. Therefore, it is more appropriate to use ordered properties for displaying ordinal data. The exceptions are those direct properties that have no ordering and these are better suited for displaying nominal data.

During the review or each metaphor category, some observations are made about the effectiveness of using different direct metaphors. These problems and suggestions are incorporated into the MS-Guidelines for direct metaphors described in chapter 9.

Chapter 5

Temporal Metaphors

9-2-5 (1974)

"Time is a sort of river of passing events, and strong is its current; no sooner is a thing brought to site then it is swept by and another takes its place, and this too will be swept away." [Marcus Aurelius]

Chapter 5

Temporal Metaphors

5.1 Introduction

In the real world a great deal of useful information is dependent on the perception of time. For example, a pedestrian crossing a busy road is required to interpret the amount of time between vehicles. The rate and frequency of traffic may also impact on the pedestrian's decision of when to cross. Temporal concepts like duration, rate and frequency can also be used to encode abstract information.

Temporal metaphors relate to the way we perceive changes to pictures, sounds and forces over time. The emphasis is on interpreting information from the changes in the display and how they occur over time[1]. This chapter describes in more detail the three classes of temporal metaphors (table 5-1):

- *Temporal visual metaphors* concern the way pictures change with time.
- *Temporal auditory metaphors* concern the way sounds change with time.
- *Temporal haptic metaphors* concern the way haptic stimuli change with time.

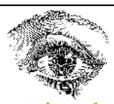
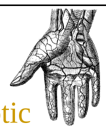
		SENSORY DISPLAY MODES		
		 visual	 auditory	haptic 
METAPHOR CLASSES	SPATIAL METAPHORS (chapter 3)	Spatial Visual Metaphors (section 3.2)	Spatial Auditory Metaphors (section 3.3)	Spatial Haptic Metaphors (section 3.4)
	DIRECT METAPHORS (chapter 4)	Direct Visual Metaphors (section 4.2)	Direct Auditory Metaphors (section 4.3)	Direct Haptic Metaphors (section 4.4)
	TEMPORAL METAPHORS (chapter 5)	Temporal Visual Metaphors (section 5.2)	Temporal Auditory Metaphors (section 5.3)	Temporal Haptic Metaphors (section 5.4)

Table 5-1 The focus of chapter 5 is on the three classes of temporal metaphors.

[1] All the senses require some amount of time to interpret a stimulus. This is very fast for vision, while with hearing most sounds are more prolonged events with a temporal structure. This means that temporal metaphors are not completely distinct from direct and spatial metaphors. The conceptual difference is that temporal metaphors relay how temporal changes convey information. They therefore involve the user's perception of events in time.

The MS-Taxonomy distinguishes between temporal visual, temporal auditory and temporal haptic metaphors. The general concepts that describe temporal metaphors are independent of sense. It is simply the ability of each sense to perceive changes over time that need to be considered for visual, auditory and haptic metaphors. Because the concepts abstract across the senses it is possible for temporal metaphors to be directly compared between senses[2]. For example, the ability of the visual sense to identify a visual alarm event can be compared with the ability of hearing to identify a sound alarm.

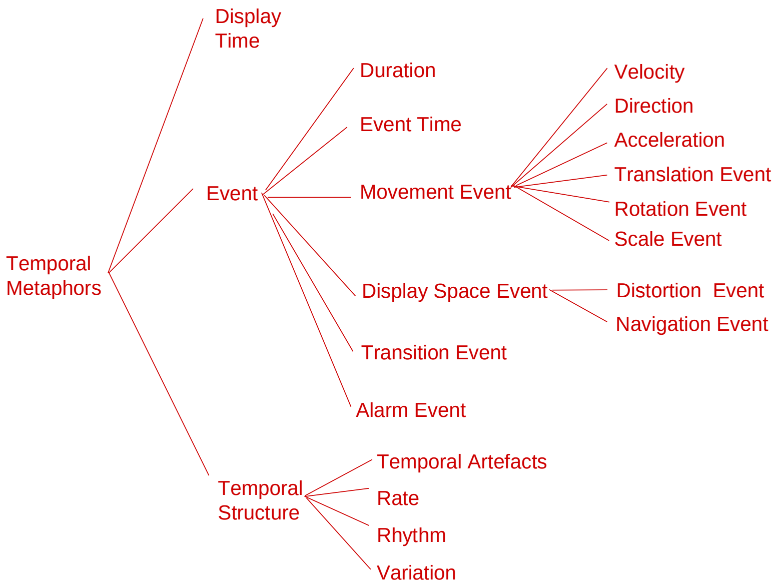


Figure 5-1 The hierarchy of concepts describing temporal metaphors.

The concepts that describe temporal metaphors are shown in figure 5.1. To understand temporal metaphors requires an overview of the high-level concepts and a detailed understanding of low-level concepts. This chapter first introduces the high-level concepts that describe temporal metaphors (section 5.1.1). To clarify the concepts, a more detailed discussion is included for each of the three senses, visual (section 5.2), auditory (section 5.3) and haptic (section 5.4). This detailed discussion provides a review of existing literature[3] and helps to explain the abstract concepts that describe temporal metaphors.

5.1.1 Temporal Metaphors

The design space for temporal metaphors can be described using the following general concepts (figure 5-2):

- the display time
- an event
- the temporal structure.

[2] Temporal metaphors can involve changes to spatial and direct properties over time. It is straightforward to compare the ability of each sense at judging changes in a spatial property. However, direct properties are very dependent on the sense and so are harder to compare.

[3] A summary of the applications discussed in this chapter is available in Appendix A.

Temporal metaphors are composed of events that occur within the *display time* (figure 5-2). The *display time* provides the temporal reference for the data events that are displayed. This is analogous to the way a metronome is used in music to provide a background measure of time. The display time is not usually considered as part of the design space, but simply assumed to be constant. However, it is possible to consider the display time[4] during the display design. For example, changing the display time could speed up or slow down the rate at which data is displayed.

Events have two main properties, the event time and the duration of the event (figure 5-3). Both the event time and event duration are interpreted in relation to the display time. These events affect changes to the visual, auditory or haptic display. It is these changes and the timing and duration of these changes that are interpreted by the user as information.

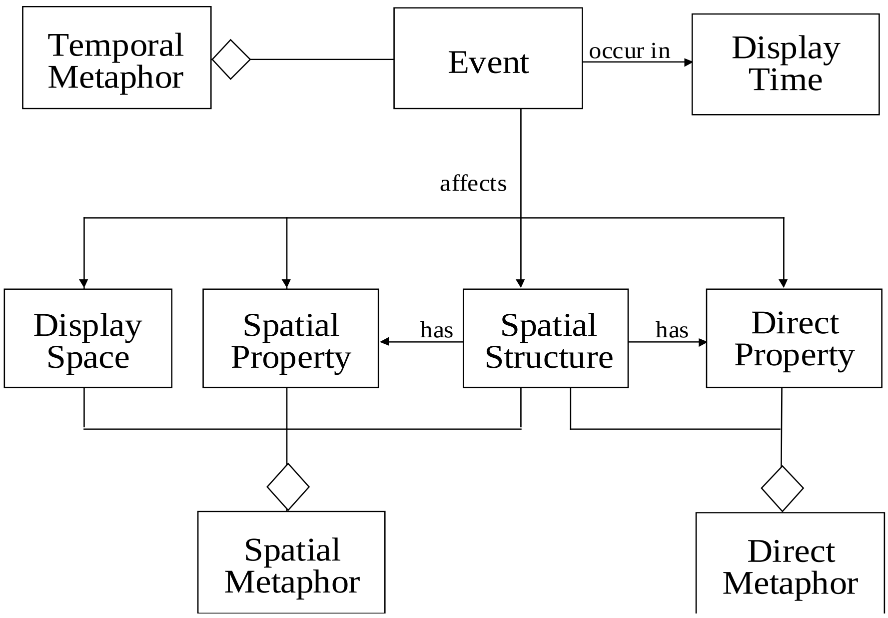


Figure 5-2 Temporal metaphors require time and involve a modification to either a direct metaphor, a spatial metaphor or both a direct metaphor and spatial metaphor.

An event can affect a change to the display space, a spatial property, the spatial structure or a direct property in the display[5]. This allows events to be categorised by reusing many of the concepts described for spatial metaphors and direct metaphors (figure 5-2). The MS-Taxonomy defines the following types of event (figure 5-4):

- a display space event
- a movement event
- a transition event
- an alarm event.

Display space events cause a change to the perceived display space (figure 5-5). For example, a distortion event can change the metric at a location in the display space. A navigation event can affect a change in the user's position in the display space and is usually associated with user interaction. Although this thesis does not attempt to cover

[4] The notion of a display time is analogous to the concept of the display space described for spatial metaphors. This suggests that a concept such as *distorted display time* could be used.

[5] Complex temporal metaphors can affect both spatial and direct properties at the same display time.

the design of user interactions, events, such as navigation are included for completeness.

Movement events[6] are related to changes in spatial properties of structures and can be characterised by properties such as direction, velocity and acceleration (figure 5-4). Distinct types of movement events include; translation events, rotation events and scale events. Translation events involve a change to the spatial property of position. Rotation events involve a change to the spatial property of orientation. Scale events cause a change to the spatial property of scale.

The other types of events are transition events and alarm events (figure 5-5). Transition events cause a slow change to either spatial structures or direct properties. By contrast alarm events cause a very sudden change to either spatial structures or direct properties.

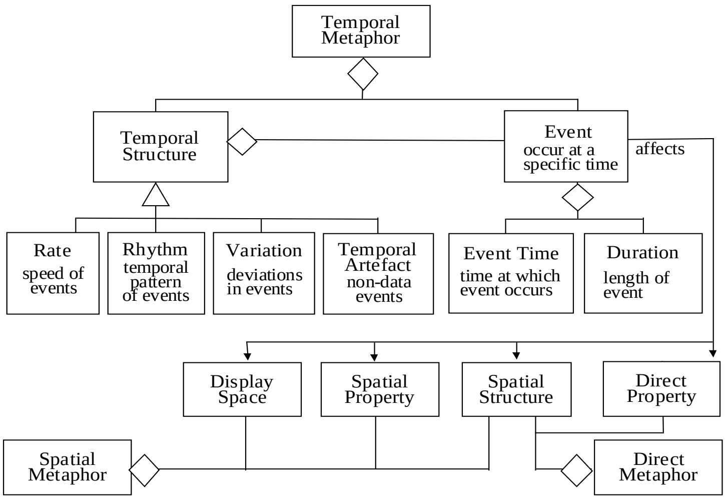


Figure 5-3 Temporal metaphors are usually composed of a number of events that have some temporal structure.

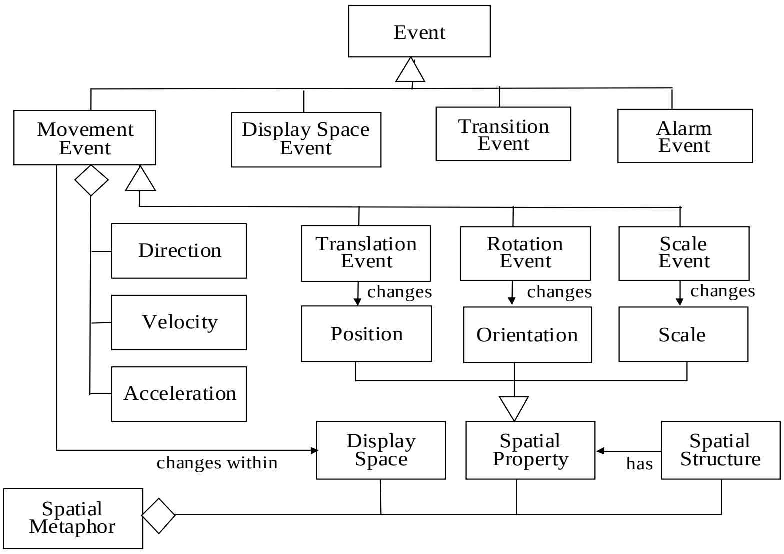
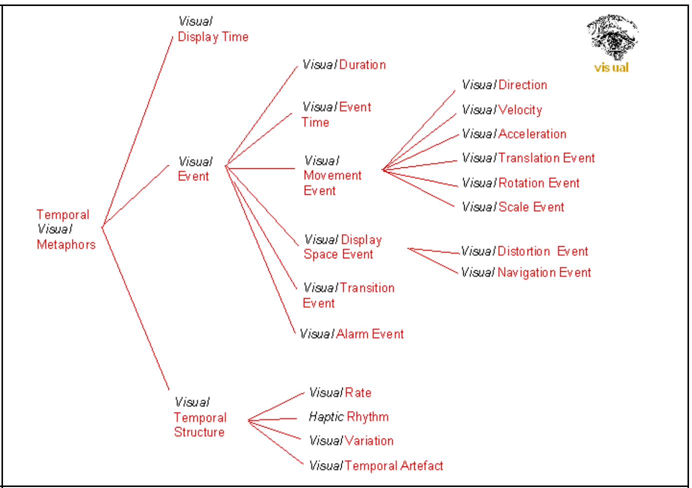


Figure 5-4 Movement events may have properties of direction, velocity and acceleration. Movement events are defined in terms of the spatial properties of position and orientation.

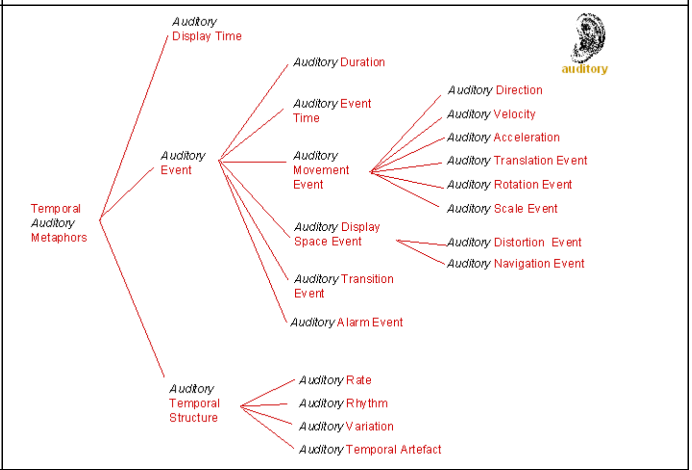
6 Movement events are defined in terms of changes that occur to elements in the display. These events could also be considered from the user's perspective and described as user interactions. However, the MS-Taxonomy focused on designing data models and not user interactions, therefore events are described from the perspective of the display.

A user may interpret information based on a single event. For example, a visible object changing position may be interpreted in terms of the old position and the new position, as well as the speed of movement. However, information may also be interpreted based on patterns that occur in a sequence of events. This is described as temporal structure. Types of temporal structure include the rate of events, the rhythm of events and the variations between events (figure 5-3). While these concepts are generally well described in the domain of music they are less commonly associated with information displays for the other senses.

Temporal **Visual** Metaphors



Temporal **Auditory** Metaphors



Temporal **Haptic** Metaphors

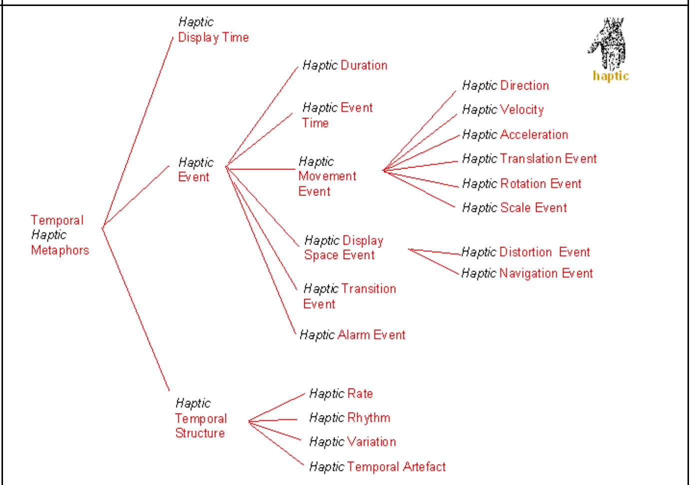


Table 5-2 The concepts describing temporal metaphors can be applied to each sense.

The concepts that describe temporal metaphors are general and can be considered for each of the senses (figure 5-5). Chapter 5 reviews the concepts for temporal visual (section 5.2), auditory (section 5.3) and haptic metaphors (section 5.4).

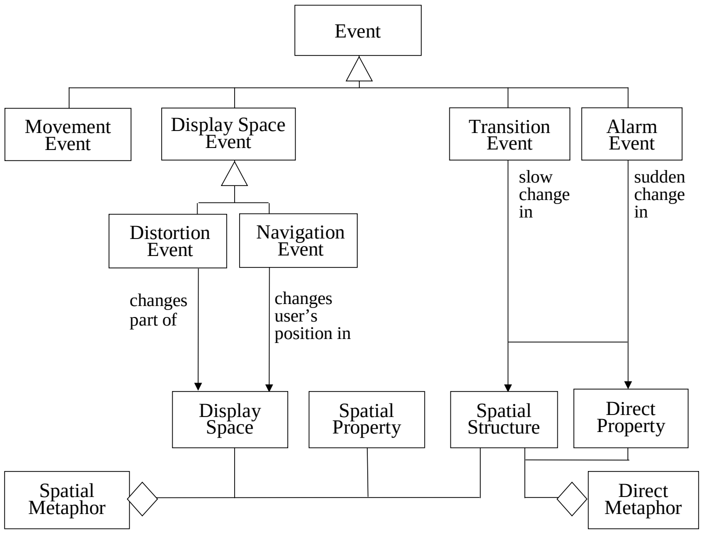


Figure 5-5 The different types of events.

5.2 Temporal Visual Metaphors

This section describes the concepts of temporal metaphors as they apply to the visual sense (table 5-3). Temporal visual metaphors relate to the visual perception of changes to the display space, the structure of objects and the spatial and direct properties of visual objects and the way they occur over time[7]. The important considerations in the design of temporal visual metaphors are (figure 5-6):

- the visual display time
- visual events
- visual temporal structures.

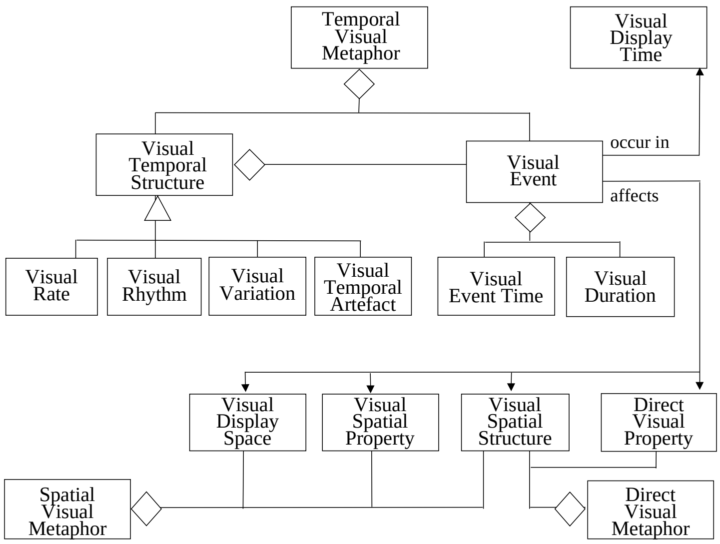


Figure 5-6 A UML diagram showing the components of temporal visual metaphors.

[7] A temporal visual metaphor displays information that is interpreted in terms of the changes to a spatial visual metaphor or a direct visual metaphor.

The visual display time refers to the reference time that is used to judge the timing of events. Most displays keep this reference time constant, although it is possible to speed up or slow down time. For example, slow motion is often used to display real world events. This allows the viewer to better view the sequence of events.

Temporal visual metaphors use visual events to display information (figure 5-6). Visual events occur at a specific moment in the visual display time. This is described as the visual event time. Visual events also have a visual duration. The use of visual events to display information is discussed in Section 5.2.1.

Temporal visual metaphors can also display information using a series of visual events (figure 5-6). A sequence of visual events can have a visual temporal structure. Types of visual temporal structure include visual rate and visual rhythm. The use of visual temporal structure to display information is discussed in Section 5.2.2.

Temporal Visual Metaphors (section 5.2)	Visual Display Time (section 5.2)		
	Visual Event (section 5.2.1)	Visual Duration (section 5.2)	
		Visual Event Time (section 5.2)	
		Visual Movement Event (section 5.2.1.1)	Visual Direction Visual Velocity Visual Acceleration
			Visual Translation Event Visual Rotation Event Visual Scale Event
		Visual Display Space Event (section 5.2.1.2)	Visual Distortion Event Visual Navigation Event
		Visual Transition Event (section 5.2.1.3)	
		Visual Alarm Event (section 5.2.1.4)	
	Visual Temporal Structure (section 5.2.2)	Visual Rate	
		Visual Rhythm	
		Visual Variation	
		Visual Temporal Artefact	

Table 5-3 The sections in chapter 5 that describe the concepts of temporal visual metaphors.

5.2.1 Visual Event

There are a number of different types of visual events and these include visual movement events, visual display space events, visual transition events and visual alarm events (figure 5-7, figure 5-8).

Both the visual event time and the visual duration can be used to encode information. For example, data that is naturally recorded as events that occur in time, or data that changes over time, can be displayed using a series of visual events. The timing and duration of the real event can map naturally to the visual event time or visual duration.

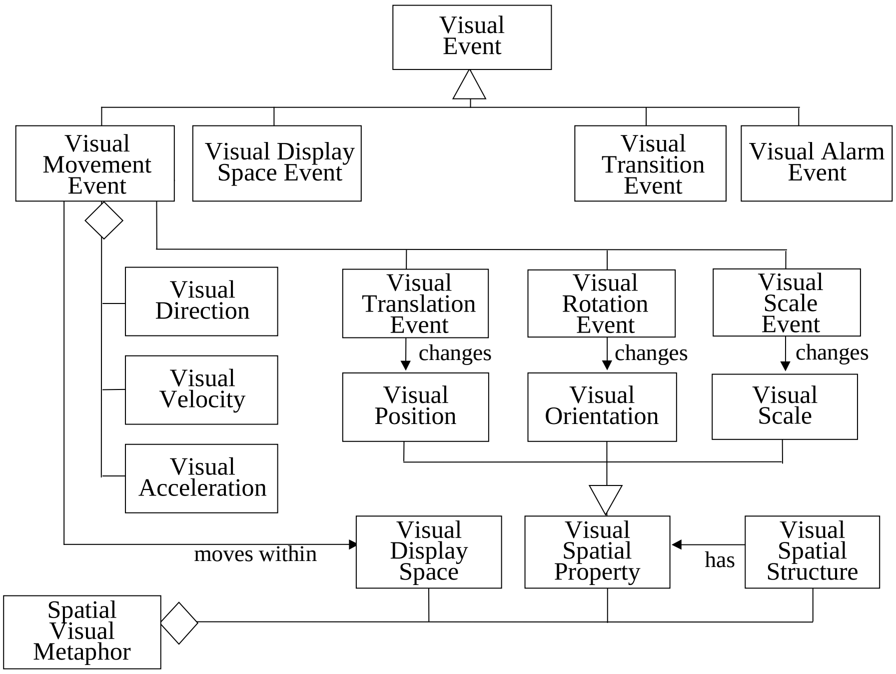


Figure 5-7 A UML diagram of visual movement events.

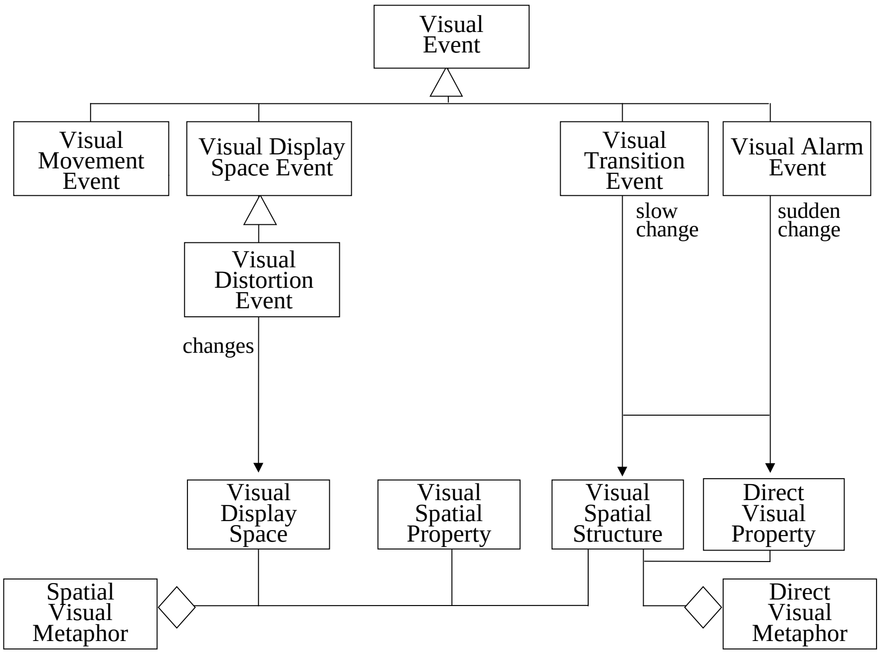


Figure 5-8 A UML diagram showing the types of visual events.

5.2.1.1 Visual Movement Events

A simple use of visual movement events is to assist in the understanding of complex visual spatial structures. The perceptual cues of motion parallax, accretion and deletion (see chapter 2) help in the interpretation of visual spatial structure.

Virtual Environments usually provide a number of mechanisms that allow objects to be moved or the user's viewpoint to be changed [Stuart 1996]. For example, in an application that displays petroleum exploration data, parts of the geological model, such as faults, can be selected and then translated or rotated [Harding, Loftin et al. 2000]. Wright produced a number of 3D views of the security market. Users could move these objects in 3D space to gain a better understanding of the relations displayed [Wright 1995]. As Wright notes, one problem with 3D is that occlusion along with distortions from the perspective view can make it difficult to compare objects. Allowing the user to move the objects in space alleviates these problems. In the `Cone Tree` [Robertson, Mackinlay et al. 1991] application a hierarchical tree is displayed in 3D. When the user selects a node in the graph, the tree rotates until the selected node is brought to the front. This rotation event helps the user maintain an understanding of the visual spatial structure.

Another simple use of visual movement events is to assist in the understanding of large displays of data. This is also a common technique in Virtual Environments where navigating through large worlds is required. For example, Robertson, Card et al. developed a number of interactive techniques to allow a user to move through large, 3D visual spaces [Robertson, Card et al. 1993]. These include the `Walking Metaphor`, the `Point of Interest Logarithmic Flight` and `Object of Interest Logarithmic Manipulation`. All these techniques allow for targeted exploration of the large data space. In the `Galaxy of News` system, the user can navigate through a network of points representing the relations between news articles [Rennison 1994].

One common approach from Scientific Visualisation that is used to assist in the interpretation of 3D volumetric data is to move a 2D cutting plane though the data. This slicing technique is has been used with medical diagnostic data [DeFanti, Brown et al. 1989] and geological data [Harding, Loftin et al. 2000].

The properties of visual movement events can also be used to directly display data. This is a technique used in *Scientific Visualisations*, where the direction, speed and acceleration of moving objects may directly convey information from physical simulations. For example `streamlines` [Giles and Haimes 1990] use moving arrows and `particle tracing` uses moving spheres to display fluid flows [Sadarjoen, van Walsum et al. 1994].

More abstract uses of movement events have been to indicate nominal state information. For example, Spence describes an agent visualisation where the agent is represented by a circle. The circle rotates when the agent is active at some task [Spence 2001]. In the `3D-Room` application an opening door signals that a user has moved to a new room [Robertson, Card et al. 1993].

Visual scale events are a particular type of movement event where the visual scale changes. Wright uses this technique to display real-time financial data [Wright 1995]. For example, bars representing share price change size as the underlying market fluctuates. In a tool for designing electric circuit diagrams the sensitivity of components is displayed as circle size [Spence and Apperley 1977]. The size of the circle changes as the designer tests the circuit over a range of frequencies.

5.2.1.2 Visual Display Space Events

Changes of the visual display space are usually associated with user interactions rather than changes to data. For example changing the view of the visual display space occurs when the user zooms. Many applications allow zooming to points of interest in the data or moving to predetermined points of view [Wright 1995]. For example, zooming is a feature of the `SeeNet` program that visualises network data [Becker, Eick et al. 1995].

In the `Hyperbolic Browser`, a distorted space is used to display a `graph drawing` [Lamping and Rao 1996]. In this application the user can interactively change the point of focus of the display. This causes a slow transition to a view with the new distorted space.

5.2.1.3 Visual Transition Events

A visual transition event is a slow change that occurs to either visual spatial structures or direct visual properties. In some cases both the visual spatial structure and direct visual properties may change at the same time. For the visual sense this type of event is better known as *animation*.

Mihalsin, Timlan et al. use a visual transition event to display abstract data with multiple attributes [Mihalisin, Timlin et al. 1991b]. In this case regular increments of time are associated with a regular change in one data attribute. The affect of this changing data attribute on the other data attributes is reflected in the change to the visual display. Using a similar technique, Chen displays the nodes of a `graph drawing` as they undergo a visual transition event. For example, in a drawing of an author co-citation network, the direct visual property of colour represents the publication field. During a series of regular steps through the different publication fields the nodes change colour. At any time step all authors in the network who share the same related area are highlighted [Chen 2000].

Visual transition events can also be a natural way to represent data that is evolving over real time. For example, Wright used live data feeds of parameters such as liquidity to display stock trading information [Wright 1995]. Changes to both direct visual properties such as colour and spatial visual structure such as shape produced animations of the security market. This showed how the underlying data changed with time.

While visual transition events may be appropriate to show how data changes in real time it is also possible to change the visual display time. For example, the visual display time can be slowed down or speeded up when appropriate. In the `SeeNet` program, data showing network usage can be animated as it changes between time periods. This application provides the user with a slider control to alter the speed of the animation [Becker, Eick et al. 1995].

Many systems are dynamic. Visual transition events are often useful for understanding how abstract processes or systems evolve over time. For example, animation techniques have also been applied to demonstrate how abstract concepts, such as computer algorithms, proceed. The sort algorithm has can be animated to use as a teaching aid

[Baecker, 1981]. However, there is little empirical evidence that algorithm animations are useful [Kehoe, Stasko et al. 1999].

Animations of a `Treemap` have been suggested as a way to study evolving hierarchical structures [Johnson and Shneiderman 1991] such as a financial portfolio. In the `TennisViewer` program an animation of the `Treemap` is designed to allow analysis of phases in a tennis game [Jin and Banks 1997].

Visual transition events are also important to help users understand how the visual spatial structure changes over time. In chapter 3 it was noted that graph drawings provide a good model of abstract relationships. As these relationships change a visual transition event can help the user understand the changes to the structure [Robertson, Card et al. 1993]. A sudden change between structures requires the user to make a cognitive effort to understand the new structure. This same principle is used to assist the user understand structural changes in the `Cone Tree` [Robertson, Mackinlay et al. 1991] and the `Perspective Wall` [Mackinlay, Robertson et al. 1991]. A number of techniques for automating animation between different graph drawings have previously been developed [Friedrich 2002].

5.2.1.4 Visual Alarm Events

A visual alarm event is a sudden change that occurs to either visual spatial structures or direct visual properties. A visual alarm event is designed to attract the user's attention and signal properties such as urgency and the nature of the event. For example, the use of pop-up windows for error messages is a common visual alarm event when using a windows interface.

The effectiveness of visual alarm events has been investigated for applications such as air traffic control [Athènes, Chatty et al. 2000]. Direct visual properties such as colour, shape and opacity have been shown to affect the reaction time of the user. Some guidelines for using the direct visual property of colour in visual alarm events have been suggested [Cardosi and Murphy, 1995]. For example, red is used to indicate warning or danger, yellow is used to indicate caution and green is used for normal events or to notify the user of a ready status. Spatial visual properties such as scale have also been shown to affect the time it takes for a user to react to a visual alarm event [Athènes, Chatty et al. 2000].

It is also known that peripheral vision is sensitive to movements [Goldstein 1989, p274]. Hence a sudden visual movement event to the periphery can also attract the user's attention.

5.2.2 Visual Temporal Structures

When a sequence of visual events is displayed the visual temporal structure can also be interpreted as information. It is perhaps more usual to think of using temporal structures in the auditory domain. For example, in music, the concepts of rate, rhythm and variations such as melody, inflection or dynamics are well established. However, this does not preclude the use of these temporal structures visually.

The types of visual temporal structure are:

- visual rate
- visual rhythm
- visual variation
- visual temporal artefact.

Visual rate describes the frequency at which events occur. *Visual rhythm* describes a regular pattern of events, where pattern is perceived by a change to some visual property. *Visual variation* describes a more complex pattern of events such as deviations in rhythm or rate over time. *Visual temporal artefacts* are unrelated to the data but are used to assist in the interpretation of data events. For example, displaying a visual temporal artefact at regular time intervals can assist the user keep track of time.

There are not many examples of where visual temporal structures have been explicitly designed for displaying information. One exception is the encoding of properties for visual alarm events. In this case varying the visual rate at which messages blink provides an effective way to encode urgency [Athènes, Chatty et al. 2000].

While there are not many examples of explicitly designing temporal visual structures, many data displays that evolve over time rely on the user's ability to detect these types of temporal patterns. For example, in a display of the stock market [Wright 1995], the visual rate displays the liquidity and volatility of the market. More complex visual variations may also be noticed. For example, perhaps a group of stocks all fall in price together. This might be reflected in the way that they change colour.

The case study developed in chapter 13 of this thesis also incorporates the real time display of stock market data. One focus of this display is to allow a financial analyst to identify useful temporal visual structures from the display.

5.2.3 A Summary of Temporal Visual Metaphors

The previous sections have reviewed the concepts that describe temporal visual metaphors. The important high-level components of temporal visual metaphors are the visual display time (section 5.2), visual events (section 5.2.1) and visual temporal structure (section 5.2.2).

The concept of a visual display time is not usually considered. Time is usually considered constant. However, speeding up and slowing down frame rates is a common technique used in film and television and has been used for abstract data displays [Becker, Eick et al. 1995]. The analogous concept from spatial metaphors is the visual display space. This suggests that introducing distortions into the display time may be a useful approach to explore.

Visual events are an intuitive specialisation of the more general concept of a temporal event. Data that is recorded as a series of events maps naturally to visual events. Movement events can assist with the interpretation of 3D structure. Navigation events are useful for exploring very large data sets. Both movement events and navigation events are closely related to user interaction. Visual transition events are more commonly described as animation and have often been used to display the way spatial structures evolve over time. Visual alarm events have effectively been used in the user interface for computer systems.

The explicit design of visual temporal structures for displaying information is not common. However, a simple strategy for designing a temporal display is to use data that naturally changes over time. The data attributes are mapped to visual properties and the properties evolve over time as the data changes. This type of display may allow the user to uncover useful visual temporal structures.

5.3 Temporal Auditory Metaphors

This section describes the concepts of temporal metaphors as they apply to the auditory sense (table 5-4). Temporal auditory metaphors relate to the auditory perception of changes to the spatial and direct properties of sounds and the way they occur over time[8]. The important considerations in the design of temporal auditory metaphors are (figure 5-9):

- the auditory display time
- auditory events
- auditory temporal structures.

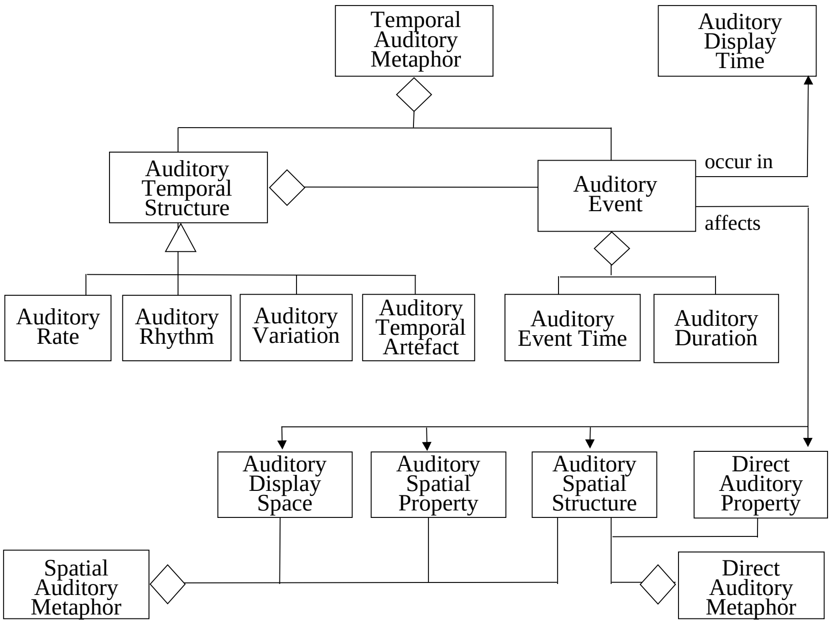


Figure 5-9 A UML diagram showing the components of temporal auditory metaphors.

The auditory display time refers to the reference time that is used to judge the timing of events. For example, the metronome is used to mark the reference time in music. Most displays keep this reference time constant, although it is possible to speed up or slow down time. The auditory resolution of temporal events is quite high [Kramer 1994a, Stuart 1996] and in general the auditory display time can be sped up without losing resolution.

8 A temporal auditory metaphor displays information that is interpreted in terms of the changes to a spatial auditory metaphor or a direct auditory metaphor.

Temporal auditory metaphors use auditory events to display information (figure 5-10). Auditory events occur at a specific moment in the auditory display time. This is described as the auditory event time. Auditory events also have an auditory duration. The use of auditory events to display information is discussed in Section 5.3.1.

Temporal auditory metaphors can also display information using a series of auditory events (figure 5-9). A sequence of auditory events can have an auditory temporal structure. Types of auditory temporal structure include auditory rate and auditory rhythm. The use of auditory temporal structure to display information is discussed in Section 5.3.2.

Temporal Auditory Metaphors (section 5.3)	Auditory Display Time (section 5.3)		
	Auditory Event (section 5.3.1)	Auditory Duration (section 5.3)	
		Auditory Event Time (section 5.3)	
		Auditory Movement Event (section 5.3.1.1)	Auditory Direction Auditory Velocity Auditory Acceleration
			Auditory Translation Event Auditory Rotation Event Auditory Scale Event
		Auditory Display Space Event (section 5.3.1.2)	Auditory Distortion Event Auditory Navigation Event
		Auditory Transition Event (section 5.3.1.3)	
		Auditory Alarm Event (section 5.3.1.4)	
	Auditory Temporal Structure (section 5.3.2)	Auditory Rate	
		Auditory Rhythm	
		Auditory Variation	
		Auditory Temporal Artefact	

Table 5-4 The sections in chapter 5 that describe the concepts of temporal auditory metaphors.

5.3.1 Auditory Event

Auditory events have two important properties, the auditory event time and the auditory duration. Both of these properties can be used to encode information. For example, data that is naturally recorded as events that occur in time, or data that changes over time can be displayed using a series of auditory events. The timing and duration of the real event can map naturally to the auditory event time or auditory duration.

Jackson and Francioni used sound events to monitor the performance of software running on parallel systems [Jackson and Francioni 1994]. Program events such as *send* and *receive* messages were mapped to sound events with different tones When executing, the relative timing, duration and pattern of sound events allowed the program to be monitored. This system provided auditory visual feedback of program performance that was superior to a purely visual or auditory display.

Kramer developed a number of scenarios, including `Loudness Nesting`, which displays data by breaking a sound event into a number of pulses [Kramer 1994b]. The pulse speed and the duration of each pulse can convey information as well as the attack and decay times of the pulse.

There are a number of different types of auditory events and these include auditory movement events, auditory display space events, auditory transition events and auditory alarm events (figure 5-10, figure 5-11).

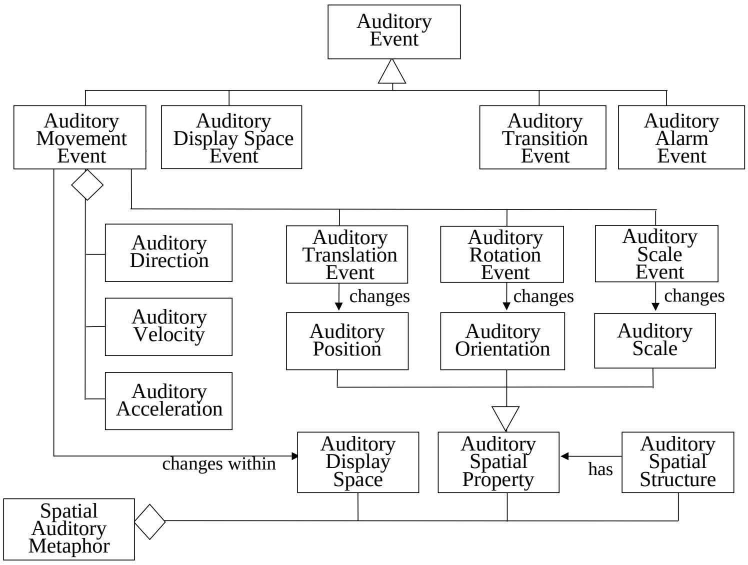


Figure 5-10 A UML diagram of auditory movement events.

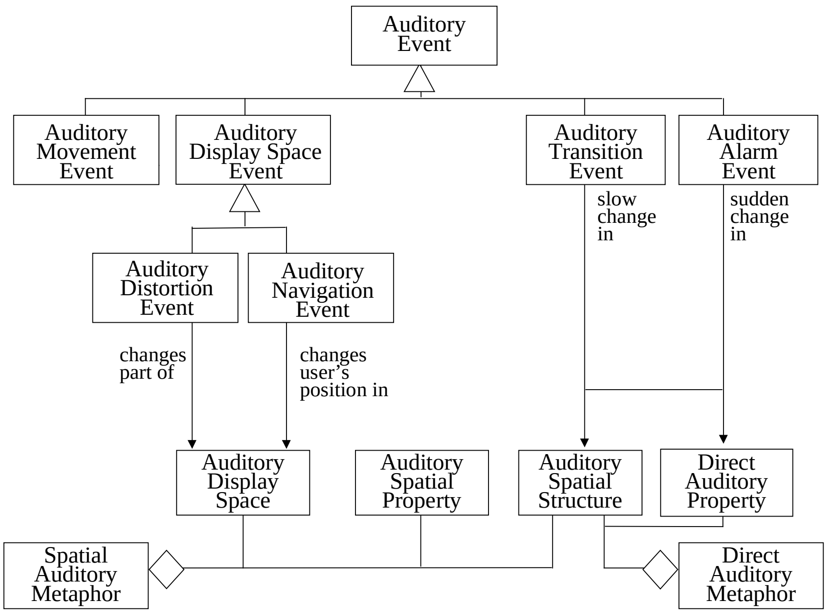


Figure 5-11 A UML diagram showing the types of auditory events.

5.3.1.1 Auditory Movement Events

Auditory movement events can assist in the interpretation of auditory position. Two of the primary cues for auditory position are the difference in loudness and arrival time of the sound at each ear. Sounds that remain equidistant from each ear can introduce some ambiguity about the location of the sound stimulus [Stuart 1996]. This ambiguity can be resolved by moving the head or if the sound itself is in motion (figure 5-13). Virtual Environments usually provide a number of mechanisms that allow sounds to be moved or a change in the user's head position to be incorporated into the display.

Some applications have used auditory movement events to explicitly display information. One example is in a study that uses auditory-guiding for instrument flying [Kramer 1994a]. In this case, turns of the are indicated by sweeping a sound from one side of the user to the other. For example, turning from left to right would be indicated by a sound sweeping from left to right.

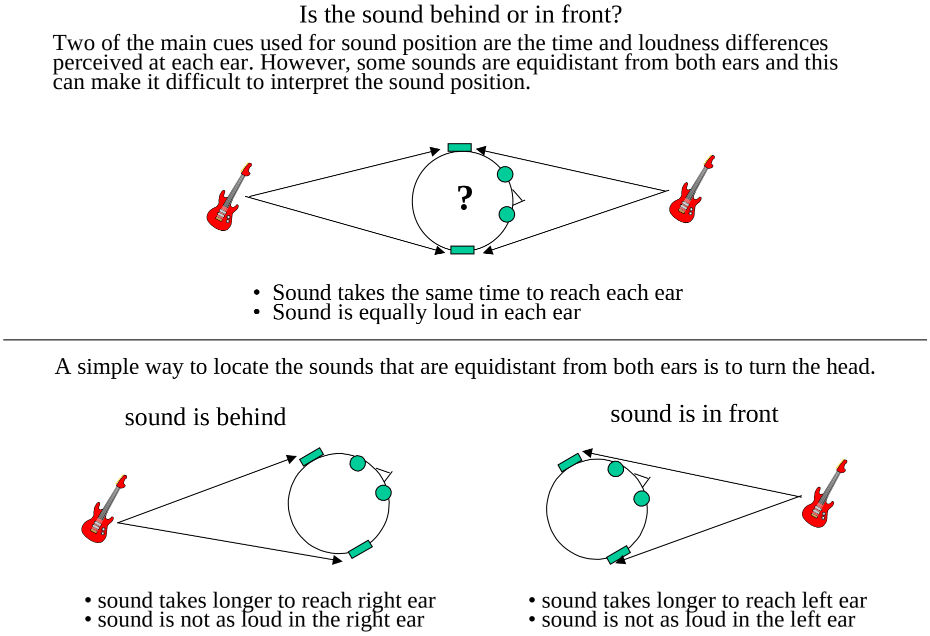


Figure 5-12 Movement events allow more accurate interpretation of auditory position.

5.3.1.2 Auditory Display Space Events

Changes of the auditory display space are usually associated with user interactions[9]. For example, changing the view of the auditory display space occurs when the user navigates. In a user-interface designed for blind people, Mynatt allowed the users to navigate the screen by receiving auditory feedback [Mynatt 1994]. A similar approach was used to assist navigation on a map [Blattner, Papp et al. 1994]. As the user moved over items in the map, auditory information was displayed. For example, if the map showed a group of buildings the user would hear attributes, such as the security access of each building.

---

[9] Once again, a detailed discussion of user interaction is not included in this thesis.

5.3.1.3 Auditory Transition Events

An auditory transition event is a slow change that occurs to either an auditory spatial structure or a direct auditory property. Alternatively both the auditory spatial structure and direct auditory properties may change at the same time. However, the use of auditory spatial structures is not common, and so most auditory transition events involve a slow change to direct auditory properties, such as pitch, timbre or loudness.

One important aspect of transitions between sounds is that they may not be noticed [Cohen 1994]. Rather, it is sudden changes to a sound signal that attract attention. For example, a constant audio signal may easily fade into the background. Gestalt theory describes this as the *principle of stability* [Williams 1994]. It states that once a user develops an interpretation of a sound then it tends to remain fixed. Any slowly changing parameters are not noticed until the original interpretation becomes completely inappropriate [Williams 1994].

Despite this problem, some systems have been designed to describe transitions within a continuous sound space. Scarletti describes two different systems [Scarletti 1994]. One system allows transition between two states, which is heard as the transition between two vowel sounds. A more complex system allows transitions between eight states, each represented by different timbres.

Displaying a sequence of auditory events is a technique used to create the `sonic graph`[10] [Mansur, Blattner et al. 1985]. The `sonic graph` uses transitions in pitch to represent each point along the x-axis. The auditory event time represents position on the y-axis. This display provides the user with a pitch signals that continuously change over time. `Sonic scatterplots` also display points using a series of auditory transition events with varying pitch [Madhyastha and Reed 1994]. The `sonic histogram` displays the value of each category over time by using transitions in both pitch and timbre [Scarletti 1994].

One important role for auditory transition events is to monitor a constant stream of events. For example, in the `SonicFinder` application, a monitoring sound, in the form of pouring water, was displayed when copying a file. The pitch of the pouring sound would gradually change to indicate the amount of copying that had been completed [Gaver 1994].

`Sonnet` is an audio-enhanced software debugger to hear events in the execution of a computer program [Jameson 1994]. This system tries to provide an overview of the program function in sound. A single note may be sounded to indicate the start of a function. The progress of the function is then monitored by modulations in the volume or timbre of the note. The note stops playing when the function is completed.

5.3.1.4 Auditory Alarm Events

An auditory alarm event is a sudden change that occurs to either auditory spatial structures or direct auditory properties. An auditory alarm event is designed to attract

[10] Although this is designed to function like a visual lineplot, the sonic graph is a temporal metaphor rather than a spatial metaphor and so the two designs are difficult to compare.

the user's attention and signal properties such as the urgency and nature of the event. The auditory sense is particular effective at identifying sudden changes to a constant sound signal [Cohen 1994].

Auditory alarm events also tend to be very insistent. They tend to not only alert the user but also orientate the user in a particular direction. Furthermore, *auditory* alarm events can be detected more quickly than *visual* alarm events [Wenzel 1994]. A number of guidelines have been suggested for the design of auditory alarm events [Patterson 1982, Deatherage 1992].

Becker, Eick et al. use auditory alarm events in conjunction with visual animations to display networks in the `SeeNet` program. After a certain time period the animation repeats and an alarm bell is used to indicate this [Becker, Eick et al. 1995]. Mynatt developed an auditory user interface called `Mercator`. In this interface the user is informed of an unexpected state change by the sound of an auditory alarm [Mynatt 1994]. In a telerobotic system certain events, such as, a command being issued were signalled by an auditory alarm event [Wenzel 1994].

`ShareMon` is a software system designed to inform a user about background events while the eyes are busy with another task [Cohen 1994]. Users are notified that someone else is logging in or logging off. This system used auditory alarm events for both warnings and status monitoring. This is similar to the approach used by the `SonicFinder` application where sounds act as alerts, indicating basic events such as selecting files or resizing windows [Gaver 1994].

The `SonicFinder` is an example of the technique called `auditory icons` [Gaver 1994]. Gaver developed `auditory icons` as analogies to everyday sounds. Just as everyday sounds signal events in the real world, `auditory icons` signal events in a computer environment. Another example of this technique is the model of a soft drink bottling factory [Gaver 1994]. A particular `auditory icon` identifies each machine in the factory. Alarms or alerts are associated with each machine to indicate problems with the equipment. These sounds allowed users to visually inspect specific machines while still monitoring the whole factory using the sound display.

5.3.2 Auditory Temporal Structures

When a sequence of auditory events is displayed the auditory temporal structure can also be interpreted as information. The types of auditory temporal structure are:

- auditory rate
- auditory rhythm
- auditory variation
- auditory temporal artefact.

It is fairly common to describe temporal structures in the auditory domain. For example, in music, the concepts of rate, rhythm and variations such as melody, inflection or dynamics are well established. Indeed many applications of auditory display use sound in this way. A major strength of auditory perception is the ability to recognise patterns that occur in sounds over time [Kramer 1994b].

Bargar describes the application of music composition techniques to the design of auditory displays for abstract data [Bargar 1994]. He describes using the musical auditory variations of *inflection*, *dynamics* and *melody* along with auditory rhythm for displaying information. *Inflection* breaks sound into phrases by changes to the direct auditory properties, such as loudness or timbre. *Dynamics* uses changes in loudness levels to display information. *Melody* is a sequence of notes or pitches that repeat.

Mayer-Kress, Bargar et al. used temporal auditory structures for displaying chaotic systems [Mayer-Kress, Bargar et al. 1994]. They noted that a simple low level auditory mapping has insufficient frames of reference for clarifying complex events such as those that occur in chaotic systems. They proposed that the organisation principles of music might help reveal patterns in scientific data. Both chaos and music have perceived order and disorder and elements can recur and have detailed complexity. Music provides a way to encode relations that change locally but also relations that change over longer periods.

The application called `Personal WebMelody` [Barra, Cillio et al. 2001] mixes traditional musical soundtracks with system-generated music to display information to the network administrator. Quinn mapped complex information such as climate data to a symphony [Quinn 2001]. Tran and Mynatt developed the `Music Monitor` to display dynamic data using musical variations such as auditory rate [Tran and Mynatt 2000]. Gaver also describes the use of auditory rate to display information about the activity level of factory equipment [Gaver 1994]. Vickers and Alty mapped structures in computer programs to musical motifs and developed principles for designing these mappings [Vickers and Alty 1998]. Leplâtre and Brewster developed musical structures to assist hierarchical navigation of menus [Leplâtre and Brewster 1998].

The technique known as `earcons` uses structured sounds to construct hierarchical messages [Blattner, Papp III et al. 1994]. The sounds have a layered structure allowing complex events and states to be interpreted from a message. For example, variations in the rhythm, pitch and timbre of the sound event might be used to indicate three levels of hierarchical information. In an example application, a series of menus were associated with different `earcons` [Brewster, Wright et al. 1994]. A menu is identified by the sound's timbre and the item on the menu by the sound's rhythm. Brewster, Wright et al. ran a series of experiments using these `earcons` and devised a number of guidelines for their design [Brewster, Wright et al. 1994].

Kramer explores the use of a hierarchy of musical parameters for creating a high dimensionality in the sound display [Kramer 1994b]. Kramer suggests a number of different approaches including `Loudness Nesting`, `Pitch Nesting` and `Brightness Nesting`. For example, `Loudness Nesting` makes particular use of auditory rate and auditory variations. The displayed sound is broken into a number of pulses. The pulse speed and the duration of each pulse can convey information as well as the attack and decay times of the pulse. A second level of information is created by adding a loudness waveform to the sound. This creates a second tier of time events which Kramer calls `clusters`. The speed of these secondary sequences or clusters conveys information as do the variations of overall volume between each cluster.

While there are a number of examples of explicitly designing temporal auditory structures, many displays simply evolve over time. These displays also rely on the user's ability to detect temporal patterns. For example, McCabe and Rangwalla use this approach to directly playback events that represented the solution of a fluid dynamics problem [McCabe and Rangwalla 1994]. In a simulation of an artificial heart pump, pitch was used to indicate pressure at one location in the pump. A tapping sound was used to signal a blood cell entering a turbulent area and a bass drum sound was used to identify a valve opening or closing [McCabe and Rangwalla 1994]. The user could listen and watch the simulation evolve over time. The sound was particularly useful for determining the frequency at which blood cells crossed into a turbulent vortex. The other sounds provided good temporal cues for exact timing of the visual events.

Fitch and Kramer developed an eight variable auditory display designed to monitor physiological parameters of a patient in an emergency room [Fitch and Kramer 1994]. The patient’s heart rate was mapped to the auditory rate of a heart sound while respiratory rate was mapped to the rate of modulation on a breathing noise.

Audification techniques collect data over time and treat this data as samples from a sound wave. For example, audification is used by Hayward to play recorded earthquake data [Hayward 1994]. Seismic data has similar properties to sound waves and playing them as sound is useful in analysis and monitoring of seismic signals. The user can basically listen to the seismic events and extract their properties. The playback speed of the seismic recordings can be much faster than the actual collection rate for the data. This is an example of increasing the auditory display time.

In the discussion on spatial structures in chapter 3, a number of spatial artefacts, such as axes and grids were described. Temporal artefacts are analogous to spatial artefacts except they occur over time. They are non-data items that can assist the user interpret information in the data. For example, Hayward used this concept when playing back the earthquake signals [Hayward 1994]. Regular sounds are overlaid on the seismic data to mark time. These marks help the listener interpret the auditory rhythm and auditory rate and were particularly useful when the data was played back at different speeds. Another example, the `SeeNet` program, uses a click to mark regular time periods in an animation of network data [Becker, Eick et al. 1995]. The same idea of repeating sounds in data was suggested by Kramer who describes them as `beacons`. `Beacons` are sounds that are associated with distinct auditory events in the data. These sounds help the user to orient themselves within the data [Kramer 1994b].

One problem that occurs with `sonic graphs` [Mansur, Blattner et al. 1985] is that users find it difficult to accurately determine pitch. To alleviate this problem, Scarletti describes the use of a constant frequency tone to act as an axis. By comparing this reference tone with the pitch of the data, the user can detect when the data crosses the axis [Scarletti 1994]. Scarletti develops this concept further to create a sonic grid. When data hits a level of the grid, filters are used to create a ringing sound [Scarletti 1994].

5.3.3 A Summary of Temporal Auditory Metaphors

The previous sections have reviewed the concepts that describe temporal auditory metaphors. The important high-level components of temporal auditory metaphors are the auditory display time (section 5.3), auditory events (section 5.3.1) and auditory temporal structure (section 5.3.2).

The concepts of temporal metaphors are very intuitive when described for the auditory sense. This is not surprising as hearing is usually identified as a temporal sense [Friedes 1974]. Indeed many of the concepts described in temporal auditory metaphors have been developed within the field of music. The intuition is that the both the terminology and the skills of musical composition can be transferred to the domain of abstract data display.

One issue with sounds derived from data is that they can be unpleasant to listen to. For example, Van Scoy tried to generate music that would allow the user to find hidden clusters of data points [Van Scoy 2000]. The sounds were generated from data about a basketball game and employed a complex mapping of pitch, register and length of beats to generate the music. Unfortunately few people recognised the sound as music and most found the tones repetitious and the chords dissonant.

Sound is useful for monitoring real time data. Audio fades nicely into the background but users are alerted when it changes [Cohen 1994]. It is also possible to compress the display time and still maintain a suitable temporal resolution [Stuart 1996]. However, one problem with using sound to display multi-attributed data is that the direct auditory properties are not orthogonal. So, for example, the sound frequency can affect the perceived loudness of sounds.

An alternative way to increase the amount of data displayed is to increase the number of auditory streams. Just as a visual scene can display multiple spatial elements in parallel, an auditory display can display multiple temporal elements in parallel. This concept is described in music as *polyphony* and can also be used to display multiple streams of data. For example, sound was used to monitor multiple processes on a parallel computer [Madhyastha and Reed 1994]. Furthermore a listener can choose to selectively listen to one particular stream of sound. This is described as the *Cocktail Party Effect* [Cohen 1994].

The ability to divide sounds into separate streams is called *auditory scene analysis* [Williams 1994]. The cues that are used to perceive different streams are quite complex and governed by similar principles of gestalt as those that govern the grouping of visual elements. These cues incorporate the knowledge that sounds that have the same proximity in time or undergo the same kind of changes at the same time are usually related [Williams 1994].

While displaying data in parallel streams is a technique that can increase the amount of data displayed, there are also difficulties with this approach. Sound streams may mask each other and it is difficult to directly compare the temporal relations in separate streams [Bregman 1990]. It is also difficult to ensure that separate data displays will not stream together. To ensure good separation the sound sources should be spatially disjoint and possess different timbres [Stuart 1996]. Multiple streams are used in the case study described in chapter 13.

Temporal auditory metaphors provide some advantages over visual temporal metaphors. Other objects do not occlude sounds. Therefore, an object associated with the sound does not have to be in the field of view for the user to be aware of it. Sounds act as good alarms and can help orientate the user's vision to a region of interest. Auditory signals can often be compressed in time without loose of detail. Because of the high temporal resolution of the auditory sense, events can still be distinguished. The temporal resolution of hearing is as low as a few milliseconds [Kramer 1994a, Stuart 1996].

Finally it is noted that all sounds have some temporal component. A sound stimulus is perceived by interpreting changes that occur in air pressure over time. Even a single sound event, such as a bottle breaking, contains a complex temporal pattern that is perceived over a short period of time. However, with temporal auditory metaphors the focus is on using the changes that occur in sound events to represent abstract information.

5.4 Temporal Haptic Metaphors

This section describes the concepts of temporal metaphors as they apply to the haptic sense (table 5-5). Temporal haptic metaphors relate to the haptic perception of changes to the display space, the structure of objects and the spatial and direct properties of haptic objects and the way they occur over time[11]. The important considerations in the design of temporal haptic metaphors are (figure 5-13):

- the haptic display time
- haptic events
- haptic temporal structures.

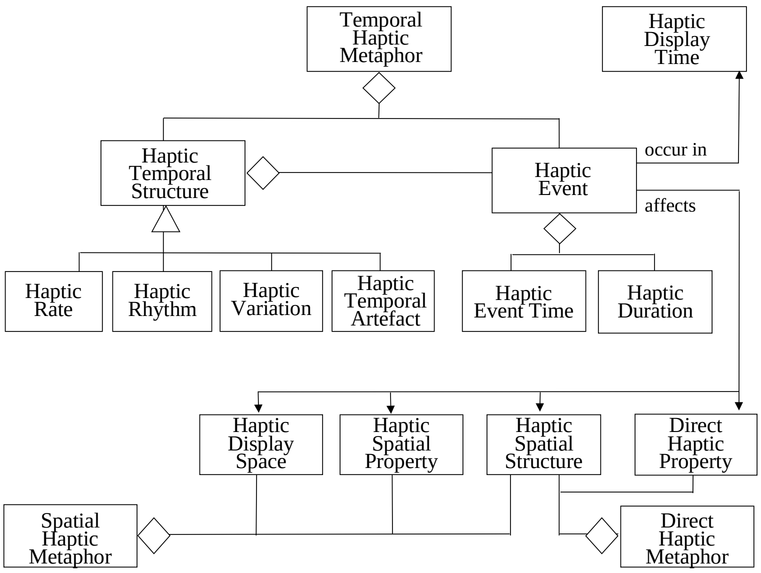


Figure 5-13 A UML diagram showing the components of temporal haptic metaphors.

As of 2002, the number of applications developed to use haptic display for interpreting abstract data is very small. Many of the concepts of temporal metaphors have not been explored. For example, the concept of haptic display time refers to the reference time that is used to judge the timing of events. No displays have been built that allow the haptic display time to be altered. However, it is possible to consider speeding up or slowing down time. The haptic resolution of temporal events is quite high. For example it is possible to perceive frequencies of 10,000 Hz [Stuart 1996].

[11] A temporal haptic metaphor displays information that is interpreted in terms of the changes to a spatial haptic metaphor or a direct haptic metaphor.

Temporal haptic metaphors use haptic events to display information (figure 5-13). Haptic events occur at a specific moment in the haptic display time. This is described as the haptic event time. Haptic events also have a haptic duration. The use of haptic events to display information is discussed in Section 5.4.1.

Temporal haptic metaphors can also display information using a series of haptic events (figure 5-9). A sequence of haptic events can have a haptic temporal structure. Types of haptic temporal structure include haptic rate and haptic rhythm. The use of haptic temporal structure to display information is discussed in Section 5.4.2.

Temporal Haptic Metaphors (section 5.3)	Haptic Display Time (section 5.3)		
	Haptic Event (section 5.3.1)	Haptic Duration (section 5.3)	
		Haptic Event Time (section 5.3)	
		Haptic Movement Event (section 5.3.1.1)	Haptic Direction Haptic Velocity Haptic Acceleration
			Haptic Translation Event Haptic Rotation Event Haptic Scale Event
		Haptic Display Space Event	Haptic Distortion Event Haptic Navigation Event
		Haptic Transition Event (section 5.3.1.2)	
		Haptic Alarm Event (section 5.3.1.3)	
	Haptic Temporal Structure (section 5.4.2)	Haptic Rate	
		Haptic Rhythm	
		Haptic Variation	
		Haptic Temporal Artefact	

Table 5-5 The sections in chapter 5 that describe the concepts of temporal haptic metaphors.

5.4.1 Haptic Event

It is possible to theorise a number of types of haptic events. These include haptic movement events, haptic display space events, haptic transition events and haptic alarm events (figure 5-14, figure 5-15). Concepts such as haptic display space events have not yet been explored for abstract information display.

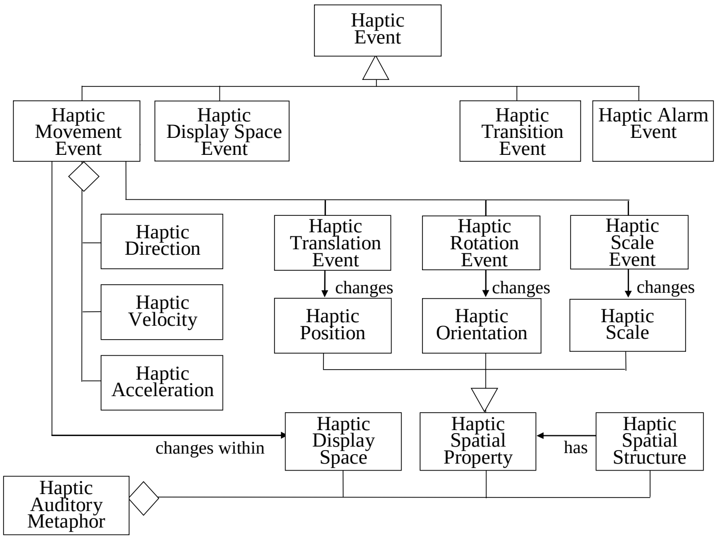


Figure 5-14 A UML diagram of haptic movement events.

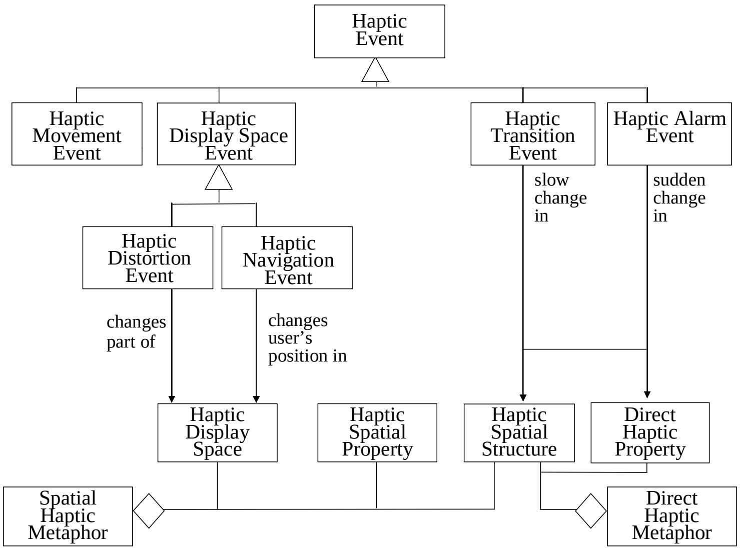


Figure 5-15 A UML diagram showing the types of haptic events.

5.4.1.1 Haptic Movement Events

The haptic sense has been characterised as most adept for interpreting movement [Friedes 1974]. It has also been noted that haptic movement events contribute significant information about the spatial map [Massie and Salisbury 1994].

A simple use of haptic movement events[12] is to assist the interpretation of complex spatial structures. This principle has been used in a number of displays. For example, in the `Legoland` application users can explore the surface geometry as they assemble virtual `Lego`™ models [Young, Chen et al. 1997]. Latimer and Salsibury developed some simple physically based models that allowed the user to move, rotate and deform

12 Haptics is bilateral in nature and generally involves some force being applied to the user and some force being received from the user. For example, the user's hand exerts a force that moves the display device, while the display device exerts some force that moves the users hand. Therefore movement events are often tightly coupled with navigation events. However, this section does not try to make a sharp distinction between the two concepts.

objects [Latimer and Salsibury 1997]. Avila and Sobierajski reported that this enabled better understanding of the spatial arrangement of complex 3D structure. In this particular case a visual display of nerve dendrites was augmented with haptic feedback [Avila and Sobierajski 1996b].

Durbeck, Macias et al. developed a system to enhance scientific visualisation that exploits the bilateral nature of haptic feedback. This system allows the user to navigate within vector fields that represented gravity, pressure, force, current, or velocity. Within these fields the user experiences appropriate haptic movement events. The direction and velocity of these events are determined by the underlying vector field [Durbeck, Macias et al. 1998].

5.4.1.2 Haptic Transition Events

A haptic transition event is a slow change that occurs to either haptic spatial structures or direct haptic properties. Alternatively both the haptic spatial structure and direct haptic properties may change at the same time. Once again this concept remains relatively unexplored for abstract information display.

Haptic transition events were implemented to display fluid flows [Nesbitt, Gallimore et al. 2001]. The user positions the force feedback stylus in a location in the flow field. and feels the transition along a flow path. However, even for simple patterns such as a unidirectional or spiral flow users had difficulty identifying the flow pattern. Crossan, Brewster et al. 2000 used haptic transition events to train veterinary surgeons [Crossan, Brewster et al. 2000]. In this application a simulator guided the user with a series of previously recorded movements.

5.4.1.3 Haptic Alarm Events

A haptic alarm event affects a sudden change to a haptic spatial structure or direct property. Like an auditory alarm event, a haptic alarm event is designed to attract the user's attention and signal properties such as the urgency and nature of the event.

A few applications have used sudden changes to vibration to signal a haptic alarm event. For example, in an application providing remote collaboration between users on a document, remote users were notified when other users approached the same area by a sudden vibration [Oakley, Brewster et al. 2000]. In another application, users navigated a model and could identify errors by sudden changes to vibration levels [Nesbitt, Orenstein et al. 1998].

5.4.2 Haptic Temporal Structures

When a sequence of haptic events is displayed the haptic temporal structure can also be interpreted as information. The types of haptic temporal structure are:

- haptic rate
- haptic rhythm
- haptic variation
- haptic temporal artefact.

In chapter 2 the haptic sense was characterised as both spatial and temporal. While there have been few attempts to use haptic temporal structure to display information it is not difficult to transfer many concepts from the auditory domain. Like the auditory sense, the haptic sense is also adept at distinguishing temporal variations in a signal. In music, concepts such as rhythm and dynamics are well established. For example, a series of haptic events could use haptic rhythm to encode information. Alternatively, the musical concept of dynamics could be displayed haptically using a change in force intensity.

One common temporal pattern detected by the haptic sense in the real world is vibration. This involves an interpretation of both the rate and amplitude of forces. A wide range of frequency of forces can be perceived, from fine vibrations at 5,000-10,000Hz up to very coarse vibrations of 300-400Hz [Stuart 1996]. While it is possible to consider using haptic rate explicitly to encode information no attempts have been made to display abstract information this way.

5.4.3 A Summary of Temporal Haptic Metaphors

The previous sections have reviewed the concepts that describe temporal haptic metaphors. The important high-level components of temporal haptic metaphors are the haptic display time (section 5.4), haptic events (section 5.4.1) and haptic temporal structure (section 5.4.2).

The use of temporal haptic metaphors has not been well explored. However, the concepts of temporal metaphors seem intuitive when described for the haptic sense. This is not surprising as the haptic sense is characterised as both a spatial and temporal sense [Friedes 1974].

The haptic sense has many similarities with hearing and it is possible to transfer concepts from this domain. For example, sounds act as good alarms and haptic alarm events may prove an effective tool in displays. Very complex temporal structures have been developed in music and these could be transferred to the haptic domain. The auditory sense has the ability to monitor a constant signal and detect sudden changes over time. In a similar way haptics could be used to monitor real time data. As with sound the temporal resolution of the haptic sense is very high and it may be possible to speed up a signal without loss of resolution.

Finally it is noted that all haptic sensations have some temporal component. Even a single haptic event, such as, a touching a surface provides a complex temporal pattern that is perceived over a short period of time. However, with temporal haptic metaphors the focus is on using the changes that occur in haptic events to represent abstract information.

5.5 Conclusion

This chapter has discussed in detail the concepts that describe temporal metaphors. These are general concepts that can be specialised to each of the senses. This has been illustrated by describing the concepts for temporal *visual* metaphors (section 5.2), temporal *auditory* metaphors (section 5.3) and temporal *haptic* metaphors (section 5.4).

The concepts of temporal metaphors are intuitive for all three senses and most developed for the auditory sense. Indeed the field of music has developed a complex language to describe temporal *auditory* metaphors. This is perhaps not surprising given that hearing is probably the principal temporal sense.

Previous work was discussed using the concepts of temporal metaphors. This served to illustrate the concepts and show the validity of the concepts that make up temporal metaphors. During the review some observations have been made about the effectiveness of using different temporal metaphors. These problems and suggestions are incorporated into the MS-Guidelines for temporal metaphors described in chapter 9.

No generic temporal metaphors were identified however a common design approach is to map data attributes to a particular property of the display and let it evolve over time. This is a simple approach for data that naturally changes over time or is recorded as events. Because vision is so adept at interpreting spatial relations, temporal visual metaphors provide an effective mechanism for displaying changes to spatial properties or spatial structures.

The use of temporal haptic metaphors is largely unexplored. While there are more examples of temporal visual metaphors and temporal auditory metaphors there is an opportunity to further transfer auditory concepts to the haptic and visual domain. In particular the use of temporal auditory structures is well explored in the field of music and it may be possible to transfer this knowledge to visual and haptic displays.

Temporal metaphors can also be considered as temporal changes that occur to spatial or direct metaphors. Therefore when designing temporal metaphors the design principles of direct and spatial metaphors need to be considered. For example, if a temporal metaphor involves interpreting a change to a direct property, such as colour, than the change needs to exceed an appropriate threshold to be noticed[13].

A further consideration with the design of temporal metaphors is the general perception of time. For example, comparing events or perceiving relations between events requires that past events be held in short term memory. There is a limit of seven on the number of items that can be held in short term memory [Miller 1957]. Another general aspect of perception that can influence the interpretation of temporal metaphors is known as *perceptual constancy* [Goldstein 1989, p 157]. When a slow change occurs to a sensory signal it may not be perceived. These types of perceptual principles are incorporated into the MS-Guidelines.

[13] This is related to what is described as the just noticeable difference (JND). This terminology comes from perceptual science and is briefly discussed in chapter 6.

Chapter 6

General Guidelines

"Design is choice. The theory of the visual display of quantitative information consists of principles that generate design options and that guide choices among options. The principles should not be applied rigidly or in a peevish spirit; they are not logically or mathematically certain; and it is better to violate any principle than to place graceless or inelegant marks on paper. Most principles of design should be greeted with some skepticism, for word authority can dominate our vision, and we may come to see only through the lenses of word authority rather than with our own eyes." [Tufte 1983]

Chapter 6

General Guidelines

6.1 Introduction

The MS-Taxonomy is a structured group of concepts that describes the multi-sensory design space for abstract data display. By abstracting concepts it is possible to hide some of the complexity of the design space. Understanding the MS-Taxonomy can afford the designer some knowledge of possible designs. However, it is also appropriate to provide more direct assistance in design decisions.

One way to assist with design decisions is to provide guidelines. This chapter introduces the MS-Guidelines and the rationale behind them. The MS-Guidelines forms a group of guidelines intended to help in designing multi-sensory displays of abstract data. They are organised by using the structure and concepts of the MS-Taxonomy.

In software engineering terms the MS-Taxonomy is akin to a *framework.* A *framework* is a generic software structure composed of general components. A framework may be used to realise specific applications [Nierstrasz and Meijler 1995]. The MS-Taxonomy is a generic structure composed of generic design concepts and can be used to realise specific designs. As Nierstrasz and Meijler note, *"an important complement to any framework consists of documentation and guidelines that aid developers to achieve a mapping from the problem domain to the abstractions provided by the framework"* [Nierstrasz and Meijler 1995]. This is also one important aim of the MS-Guidelines. That is, to assist designers to make appropriate mappings from the abstract data domain to the concepts of a multi-sensory display.

The notion of *following guidelines* is ubiquitous. For example, there are guidelines available on topics as diverse as teaching Braille [Myrna 1981], conserving nature [Hussey 1991], manufacturing smallgoods [Zemanovic 1992], analysing hazards [NSW 1992], providing medical care [Darzins 1995, Gunn and Jackson 1991], writing reports

[Blicq 1987] and closing urban roads [NSW 1979]. There are even guidelines to assist aspiring professors [Lipstreu and Doi 1963].

There are a number of ways that guidelines can assist with the design of multi-sensory displays and these include:

- guiding a process (section 6.1.1)
- capturing previous experience (section 6.1.2)
- providing structured knowledge (section 6.1.3)
- providing both general and specific principles (section 6.1.4)
- hiding complexity from the designer (section 6.1.5)
- communicating good solutions (section 6.1.6)
- evaluating the design (section 6.1.7)

6.1.1 Guiding a Process

Some guidelines are general, such as Johnson's guidelines for teaching mathematics [Johnson and Rising 1972]. Other guidelines are more specific, such as the guidelines for dumping packages of radioactive waste at sea [OECD Nuclear Energy Agency 1979]. However, in both cases the guidelines aim to assist users follow a process and to ensure the quality of the outcome [Humphrey 2000]. In chapter 10 of this thesis a process for designing multi-sensory displays is developed. A goal of the MS-Guidelines is to assist the designer follow this process and to produce a higher quality final design.

Looking more specifically at the field of design, areas such as architecture have frequently use guidelines to assist the design process. For example, there are general guidelines such as, "*Guidelines for design of low-rise buildings subjected to lateral forces*" [Gupta and Moss 1993] and "*Guidelines for energy efficient housing in the tropics"* [Aynsley 1999]. More specific aspects of architectural design are also covered, such as *"Guidelines for stair safety*" [Archea 1979].

Using guidelines to assist engineering design processes is also well established. It is not uncommon to find guidelines for designing both hardware and software. There are general guidelines, such as the *"Human Engineering Design Considerations for Cathode Ray Tube-Generated Displays"* [Banks, Gertman et al. 1992]. Quite specific guidelines have been developed, for example, to assist in the design of auditory alarms in the work place [ISO 1986], to help design user interfaces for control rooms in nuclear power plants [U.S. Nuclear Regulatory Commission 1981] or for developing software for a specific computer platform [Apple 1987]. Once again the motivation for providing guidelines for engineering design is to assist users follow a complex process and to try to ensure a level of quality in the outcomes.

6.1.2 Capturing previous experience

Designing user-interfaces is certainly a complex process and often the business success of a computer system relies on the quality of its interface. Not surprisingly, guidelines to assist in designing user interfaces are often proposed. For example, guidelines have been suggested for designing data displays [Smith and Mosier 1986], user-interfaces [Brown 1988], screen messages [Shneiderman 1982] and application screens [Galitz

1989]. Shneiderman notes, *"a guidelines document can help by promoting consistency among multiple designers, recording practical experience, incorporating the results of empirical studies, and offering useful rules of thumb"* [Shneiderman 1992a].

However, even the idea of guidelines to assist with the design of abstract data displays is not new. For example, a number of guidelines have been suggested for both visual display [Tufte 1983, Keller and Keller 1993] and auditory display [Kramer 1994a, Patterson 1982, Deatherage 1992]. It is appropriate to incorporate these existing guidelines where possible rather than invent new ones.

To capture previous experience, objective findings and useful hints is another goal of the MS-Guidelines. The design of multi-sensory displays encompasses a wide range of disciplines. So the MS-Guidelines are developed and extracted from a variety of sources. These include the fields of perceptual science, human computer interaction, information visualisation, haptic research, sonification and user-interface design. Although some new guidelines need to be developed that describe design principles in terms of the new MS-Taxonomy (Section 3.5).

6.1.3 Providing structured knowledge

It is not an aim of the MS-Guidelines to propose another set of completely new guidelines. Rather the MS-Guidelines collect existing knowledge and order it in a useful way. This ordering is achieved by using the structure of the MS-Taxonomy.

It is also expected that knowledge in the field of abstract information display will expand over the future years. Hence it is necessary to consider that the MS-Guidelines will also expand. By using the generic structure of the MS-Taxonomy, new guidelines can always be incorporated at the appropriate level.

6.1.4 Providing both general and specific principles

One problem with guidelines is that they can be hard to interpret [Mahemoff and Johnston 1998]. As noted some guidelines are very specific and detailed while others are more general and abstract in scope. Specific guidelines are precise but are usually numerous. For example, Smith and Mosier provide a very detailed list of almost 1000 guidelines for interface design [Smith and Mosier 1986]. The sheer number of guidelines can make it difficult to find the right guideline for any situation. As Wright and Fields note, to be tractable, guidelines need to be relatively small and thus they tend to be general [Wright and Fields et al. 1994]. So, general guidelines are often few in number but they may be so abstract that they are open to too much interpretation. For example, Tufte recommends that the display should *"focus on displaying the data"* [Tufte 1983].

However, both specific, detailed guidelines and abstract, general guidelines can be useful in design. Sometimes very specific guidelines can assist with a fine-tuning display performance, while more general principles may assist in setting the overall direction of design.

The approach of the MS-Guidelines is to use a number of levels of complexity and abstraction. These levels have already been defined within the structure of MS-Taxonomy. There are general guidelines about designing spatial visual metaphors. However, there are also more detailed guidelines for lower level concepts in the MS-Taxonomy. For example, the recommended frequency for auditory alarm events is specified. By using the structure of the MS-Taxonomy the MS-Guidelines are indexed by the relevant design concept. The different levels of the MS-Taxonomy allow the designer to choose guidelines for a general display concept or guidelines that target a very specific concept.

6.1.5 Hiding complexity from the designer

As discussed in chapter 1 the design of multi-sensory displays is a very complex domain and this is another reason suggested for using guidelines [Wright and Fields et al. 1994]. For example, Rasmussen and Vicente use a detailed model of human information processing to manage error in user inputs to software systems [Rasmussen and Vicente 1989]. However, they argue that this model is too difficult for software engineers to understand. To solve this problem they simply extract from their model some human factor guidelines for the software designers to use.

The MS-Guidelines work in the same way to help hide the complexity of some domains. For example, the MS-Guidelines include findings from perceptual science. However, it is not expected that the designer will require detailed knowledge of human perception to be able to apply the guidelines.

6.1.6 Communicating good solutions

User interface designers have found that some design problems often occur over and over again. When a good solution to a common problem has been devised it is desirable to reuse this solution. The issue however, often becomes how to communicate the solution amongst user interface designers. Guidelines have been suggested as a way of overcoming this communication issue [Granlund, Lafrenière et al.1999]. It is desirable that the MS-Guidelines act as an appropriate means of communicating solutions to the common problems that arise when designing multi-sensory displays.

6.1.7 Evaluating the design

A final motivating factor for developing guidelines is to act as a means of evaluating the process outcomes. It is noted that guidelines also provide a useful method for evaluating software applications [Balbo 1995]. For example, the specific recommendations of Smith and Mosier can be used to evaluate the user interface [Smith and Mosier 1986]. In another example, Bastien and Scapin developed ergonomic criteria for evaluating software [Bastien and Scapin 1993]. To ensure the MS-Guidelines are used in evaluation they are incorporated into the design process described in chapter 10.

6.1.8 Motivations for the MS-Guidelines

In summary, there are a number of motivations for developing guidelines. They can summarise accumulated knowledge in the domain of multi-sensory display of abstract data. They can hide the complexity of this cross-discipline domain. They can direct the design process. Helping the designer make design decisions. Guidelines can improve the quality of final designs and help to communicate and encourage reuse of good design solutions. Finally guidelines can assist in the evaluation of the design outcomes.

To help achieve these aims the MS-Guidelines uses the structure of the MS-Taxonomy. This structure allows guidelines to be associated with very specific concepts of design and higher-level abstract concepts. This should improve communication of the concepts and also help manage complexity. The design process to be developed in chapter 10 also uses this structure. This allows the MS-Guidelines to be closely integrated with steps of the process and should help the designer follow the process. The MS-Guidelines are also used to help evaluation of the design outcomes.

Like all guidelines the MS-Guidelines are not intended to direct an automatic process. Some of the MS-Guidelines are general principles or philosophies of design. In some cases they simply act as reminders of research results of abstract information display. In other cases they are very specific. Most importantly the MS-Guidelines both steer the design process and act as checkpoints to help evaluate the design.

6.2 Structure of MS-Guidelines

The MS-Guidelines are structured into a number of levels (figure 6.1). At the highest level are three groups of general guidelines:

- General guidelines for perception (section 6.3)
- General guidelines for information display (section 6.4)
- General guidelines for the MS-Taxonomy (section 6.5).

At the next level are guidelines for each of the nine metaphor categories of the MS-Taxonomy. Chapter 7 contains guidelines for the three types of spatial metaphors:

- Guidelines for spatial visual metaphors (chapter 7, section 7.2)
- Guidelines for spatial auditory metaphors (chapter 7, section 7.3)
- Guidelines for spatial haptic metaphors (chapter 7, section 7.4).

Chapter 8 contains guidelines for the three types of direct metaphors:

- Guidelines for direct visual metaphors (chapter 8, section 8.2)
- Guidelines for direct auditory metaphors (chapter 8, section 8.3)
- Guidelines for direct haptic metaphors (chapter 8, section 8.4)

Chapter 9 contains guidelines for the three types of temporal metaphors:

- Guidelines for temporal visual metaphors (chapter 9, section 9.2)
- Guidelines for temporal auditory metaphors (chapter 9, section 9.3)
- Guidelines for temporal haptic metaphors (chapter 9, section 9.4)

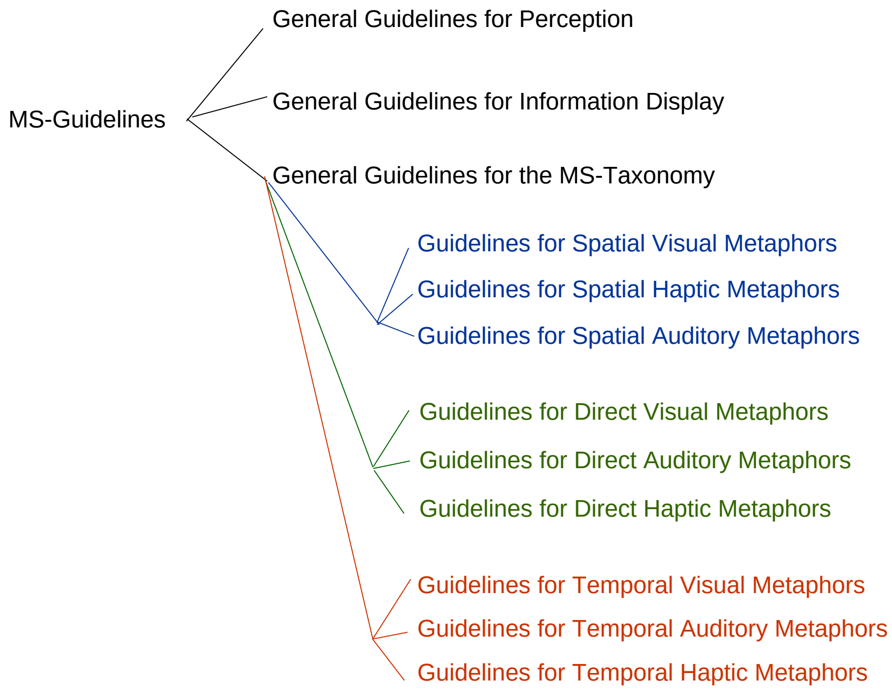


Figure 6-1 The MS-Guidelines consists of both general high-level guidelines and more specific low-level guidelines structured by the MS-Taxonomy.

6.3 General Perceptual Guidelines

As described in chapter 1 the motivation for this work was to assist in the design of Human Perceptual Tools for data mining. In general the display of information to any of the sense requires an understanding of how a physical stimulus such as sound is perceived by the user. Therefore to design multi-sensory displays requires a general understanding of some perceptual principles.

Note that a deeper discussion of these perceptual guidelines can be found in any general psychology text on human perception. Many of the guidelines presented here are adapted from Goldstein's book [Goldstein 1989]. Sekuler and Blake discuss similar material in some detail [Sekuler and Blake 1990].

GP-1 Perception is shaped by neural processing and physiology.
Information displays use the properties of a physical or *distal stimulus* to display information. The perception of this stimulus is not a direct automatic process. Perception is the end result of complex steps, many of which are not available to our awareness. The outcome from this indirect process (figure 6-2) is based on the properties of sensory receptors and neural processing [Goldstein 1989, p 65].

Figure 6-2 The perceptual process. All stages of the process affect perception.

GP-1.1 Neural maps assist spatial perception of touch and vision.
There is an explicit mapping of space into the way neural cells are organised in the visual cortex and touch cortex [Goldstein 1998]. For example, a stimulus on adjacent locations on the retina generates responses in adjacent nerve cells in

the visual cortex. The implication is that spatial visual metaphors and spatial haptic metaphors may be more effective than spatial auditory metaphors.

GP-1.2 Neurones respond to specific influences.
Many neurones are responsive to very specific stimuli. For example, a nerve may only fire to edges of a particular orientation or objects of a specific length or shape. The implication is that we may recognise a very specific stimuli generated by the display [Goldstein 1989].

GP-1.3 There are parallel pathways of perception.
There are parallel neural pathways for each sense. These pathways are specialised to process information about specific characteristics of the stimulus. For example, different visual qualities such as colour, movement and depth activate different groups of neurones in the brain. Similarly there are numerous parallel pathways that transmit auditory information from the cochlea towards the brain, while four channels of information originating from the four types of receptors in the skin are use to transmit touch information [Goldstein 1989, p6]. The implication is that information from the display can be processed in parallel.

GP-1.4 Perception is influenced by individual physiology.
Physiological processes influence what we perceive. For example, the rates of rods and cone dark adaptation and the regeneration of rod and cone visual pigments influence visual perception. The lack of cone convergence and high cone acuity account for the high resolution of foveal vision. Colour phenomena such as colour mixing, afterimages, simultaneous contrast have been connected to the properties of the cone receptors or the firing of opponent-process neurones [Goldstein 1989]. The implication is that individual physiology may influence the information perceived from a display.

GP-2 Perception is approximate.
Our perception does not always accurately match the physical stimulus. For example, a light that remains the same intensity becomes brighter during dark adaptation or two identically coloured squares appear different when they are surrounded by different coloured backgrounds [Goldstein 1989, p64]. For example, with touch, when two points that are close together touch the skin it may feel like a single point [Goldstein 1989, p65]. The implication is that a stimulus generated by an information display may not be perceived precisely.

GP-3 Perception is influenced by cognitive processes.
The perceiver's thoughts, past experiences and the meaning of the stimulus influence perception. For example, what an observer pays attention to is determined by factors such as the observer's interests and the demands of the task. The implication is that a user's performance with a perceptual display will still depend on the individual's cognitive processes.

GP-3.1 Perception is influenced by expectations.
Visual, auditory, and haptic perceptions are influenced by what we expect to see, hear and feel [Goldstein 1989].

GP-3.2 Perception is influenced by knowledge.
Perception can be influenced by knowledge that a person brings to the situation, and thought processes that are based on this knowledge [Goldstein 1989, p205].

GP-3.3 Perception may be influenced by recognition.
Recognition of a stimulus is generally the next step after perception and implies that a stimulus has been identified. However, perception may not always occur before recognition. Some recognition may occur before a perception is completely formed [Goldstein 1989, p 25]. The implication is that recognising something in a display may change the perception of the stimulus.

GP-3.4 Perception is influenced by attention.
The process of seeking out stimuli is called *attention* [Goldstein 1989, p118]. We take an active role in perception by seeking out stimuli that are of interest. Attention is important because it directs our senses to stimuli that we want to perceive and because it influences the way information is processed. Attention can enhance perception of stimuli to which we are concentrating on and decrease awareness of stimuli we are ignoring. This should, for example, be considered in testing multi-sensory displays as attention may bias performance with a particular sense.

GP-3.5 Perception is influenced by context.
Our perception of a stimulus often depends on the context within which the stimulus is perceived. For example, simultaneous colour contrast and lightness contrast affect colour perception [Goldstein 1989]. The implication is that the perception of information in part of a display cannot be assumed independent of other parts.

GP-4 Perception remains constant.
This guideline is derived from what is known *perceptual constancy*. Although the stimulus received by our sensory receptors may change, our perception of the stimulus remains constant [Goldstein 1989, p157]. For example, when looking at an object under changing lighting conditions we perceive the colour to be constant. The implication is that small changes to displayed properties may not be identified.

GP-5 Perception can be biased towards one sense.
This is known as *sensory bias* and refers to the fact that one sense may dominate another sense when there is a conflict between two or more senses [Welch and Warren 1980]. There is a strong tendency of the perceptual system to produce a single perceptual experience that is coherent. An example of this is the *ventriloquist effect*. In a ventriloquist act, the audience perceives the location of the sound source to be the moving dummy's mouth. In this case the visual sense biases the auditory sense. The implication of this is that a display to one sense may bias the display to another sense.

GP-5.1 Attention can affect sensory bias.
Sensory bias depends on the relative allocation of attention. The amount of attention can be influenced by the appropriateness of using a sense for a particular task [Welch and Warren 1980]. The implication is that a user's performance with a particular modality in a multi-sensory display can be influenced by the attention they allocate to that modality. This may need to be considered when comparing modality performance.

GP-5.2 Learning can affect sensory bias.

Individuals have innate predispositions or learned responses to allocate relatively more attention to a modality that is most appropriate for a given perceptual event [Welch and Warren 1980]. The implication is that different individuals may allocate attention to different modes of a display.

GP-6 Perceptual responses have thresholds.

Two frequently used measures of perception are the *absolute threshold* and the *difference threshold*. The *absolute threshold* is the smallest amount of stimulus energy necessary for an observer to detect a stimulus [Goldstein 1989, p8]. The *difference threshold* is the smallest difference between two stimuli that a person can detect. [Goldstein 1989, p10] The *difference threshold* is also known as *the Just Noticeable Difference* (JND).

GP-6.1 Weber's Law

This law states that:

$$\frac{JND}{S} = K$$

where *JND* is the Just Noticeable Difference, *S* is the stimulus and *K* is a constant. The ratio of JND to standard stimulus is constant over a wide range of intensities. This law holds true for most senses as long as the stimulus is not too close to the absolute threshold [Goldstein 1989, p11]. When designing information categories this law provides a guide to how much change in stimulus is required for each category.

GP-6.2 Steven's Power Law

This law states that:

$$P = KS^n$$

where *P* is magnitude, *K* is a constant, *S* is stimulus intensity and *n* is a real number. The implication is that the relationship between the intensity of a stimulus and our perception of its magnitude follows the same general equation for each sense [Goldstein 1989, p12-13].

GP-7 Perception groups small elements into larger elements.

This is demonstrated by the gestalt perceptual grouping principles where visual elements are grouped by properties such as familiarity and similarity. In a similar way the principles of auditory scene analysis indicate how various sounds become perceptually grouped with one another. Perceptual organisation also occurs with touch. For example, grouping allows the patterns of raised dots of Braille to be identified as a single letter during haptic exploration. The implication is that individual components of a data display will be perceived to have global structure whether it exists of not.

GP-8 Seven is a magic number.

There is an often-quoted limit on our capacity for processing information held in short term memory. As a rule of thumb the 'magic' number is 7 (± 2) [Miller 1957]. So, for example, it is desirable to limit the number of data categories that the user must understand to seven.

6.4 General Guidelines for Information Displays

GD-1 Emphasise the data.
The display should focus on displaying the data, avoiding artefacts and distortion while maximising the amount of data displayed [Tufte 1983].

GD-2 Simplify the display.
Where possible the display should be simple. It is better to complicate the mapping from data to display then to use an overly complex display. For example, Bly suggests that reducing the data to a single variable greatly simplifies the representation in sound displays [Bly 1994].

GD-3 Design for a task.
The task for which the display is required should steer the design process. The task should be used to check the success of the resulting design. This guideline has been recommended for auditory display by Kramer: *"System design should be heavily task dependent*" [Kramer 1994b]. It has also been suggested by Tufte for visual display: "*Displays should serve a reasonably clear purpose*" [Tufte 1983].

GD-4 Iterate the design process.
A quality process should follow standard quality steps of plan, do, check, and act. The iterative nature of this process with several check steps helps ensure a quality display. As Tufte notes, "*Revise and edit*" [Tufte 1983].

> **GD-4.1 Avoid designer bias.**
> The designer of a display become poor judges of the display the more they use it. Designers evolve into expert users by the time the display is built. Having become well acquainted with the display model it is perceived as effective and intuitive. This perception may be misleading. As Scarletti notes, *"Listening to the data in multiple ways is essential if one is to minimise the distortion and bias inherent in any one choice of representation"* [Scarletti 1994].

6.5 General Guidelines for the MS-Taxonomy

MST-1 Use each sensory modality to do what it does best.
The Modal Specific" theory states that each modality has distinct patterns of transduction. So, each sense has unique sensory and perceptual qualities that are adept with certain kinds of complex information [Friedes 1974]. Because of this the various sensing modalities are differentially well suited for different kinds of perceptual tasks [Welch and Warren 1980]. With complex information, specific modalities are adept with certain kinds of spatial or temporal patterns [Friedes 1974]. There are three primary categories of patterns in sensory information: one involves space, one involves time and the other involves movement (space and time). For each of these categories (space, time, movement) one of the senses (visual, auditory, haptic) is the most adept modality [Friedes 1974].

> **MST-1.1 Vision emphasises spatial qualities.**
> Vision is well suited to finding spatial patterns. The visual sense is most accurate at judging spatial relationships.

MST-1.2 Hearing emphasises temporal qualities.
Hearing is a primarily temporal sense as sound evolves over time. Hearing is well suited to judging temporal relationships.

MST-1.3 Haptics emphasises movement.
Haptic perception involves the integration of spatial and temporal patterns. Haptic sense can act on the environment, unlike vision and audition [Durlach and Mavor 1995].

MST-1.3.1 Point force-feedback only provides temporal information.
Most current haptic displays only provide force feedback to a single point, and so are not as useful for providing the full range of haptic spatial cues.

MST-1.3.2 Tactile displays are not readily available.
Most current haptic displays only provide force feedback, which only mimic a limited range of haptic sensations.

MST-2 Use the spatial visual metaphor as a framework for the display.
Augment the visual spatial structures with sound and haptics. While different modalities are adept with different types of information, they are all integrated about a common spatial model. The visual spatial model is designed first and this is used as a framework for the other sensory displays.

MST-3 Increase the human-computer bandwidth.
The ultimate aim is to display more attributes of data. This requires different information to be presented to the users different senses. This type of display is characterised as *complementary display* [McGee, Gray et al. 1998, Pao and Lawrence 1998]. McGee, Gray et al. described the three types of multi-sensory display as:

- *Conflicting display*s result in performance that is worse with the combined display than with a single modality. In this case contradictory information is displayed to each sense.
- *Redundant displays* result in performance that is the same with the combined display and a single modality display. Although users may report a reduction in workload or an increase in confidence. This approach has also been described as *cooperative rendering* [Pao and Lawrence 1998]. In these types of display the same information is displayed to each sense.
- *Complementary displays* result in performance that is superior with the combined display compared to a single modality. In this case different information is displayed to each sense.

MST-3.1 Use complementary display.
This maximises the amount of data that can be displayed (GD-1).

MST-3.2 Avoid redundant display.
Do not present the same information to each sense as this increases the complexity to the display without increasing the human-computer bandwidth.

MST-3.3 Avoid conflicting display.
Ensure that information on each sense does not produce conflicts. This is difficult to predict as cross modality effects are difficult to determine. Users may match a stimulus from one sense with another. For example, it has been shown that subjects draw relationships between colours and tones with people typically matching lighter squares with higher pitches and darker squares with lower pitches [Sekuler and Blake 1994].

MST-4 Consider sensory substitution.
With complex information the senses are adept in their preferred modality. However, if the information is simple then different modalities are equivalent and the senses can substituted for each other. In this case there may be opportunities to adapt the spatial mappings of vision and the temporal mappings of audition to alternative senses.

MST-4.1 Adapt spatial visual metaphors to spatial auditory metaphors.
Spatial visual metaphors have been well explored and can be transferred to the auditory domain provided the reduced spatial location ability and temporal nature of hearing is considered.

MST-4.2 Adapt spatial visual metaphors to spatial haptic metaphors.
Spatial visual metaphors can be transferred to the haptic domain provided the reduced spatial location ability of haptics is considered. Touch has previously been used as a substitute for vision. For example, the `Optacon` was developed to use a number of small vibrators to provide a display of photographs or text [Goldstein 1989, p209].

MST-4.3 Adapt temporal auditory metaphors to temporal visual metaphors.
Temporal auditory metaphors have been well established in the domain of music and these principles and can be transferred to the visual domain for encoding temporal patterns.

MST-4.4 Adapt temporal auditory metaphors to temporal haptic metaphors.
Temporal haptic metaphors have not been well explored; however, concepts from temporal auditory metaphors may be transferred to the haptic domain for encoding temporal patterns.

6.6 Conclusion

This chapter has introduced the motivation for developing the MS-Guidelines. It also introduced some general high-level guidelines for designing multi-sensory displays. These include some general guidelines for perception that are summarised in table 6-1. They also include some general principles for designing information displays that are summarised in table 6-2. Finally some general guidelines for using the MS-Taxonomy were outlined and these are summarised in table 6-3.

Chapter 7 describes lower-level guidelines for spatial metaphors. Chapter 8 describes lower-level guidelines for direct metaphors. Chapter 9 describes lower-level guidelines for temporal metaphors. In chapter 10 a process for developing multi-sensory displays will be developed that incorporates the MS-Guidelines. These guidelines are evaluated during a case study and the results of this evaluation presented in chapter 14.

GP-1 Perception is shaped by neural processing and physiology. GP-1.1 Neural maps assist spatial perception of touch and vision. GP-1.2 Neurones respond to specific influences. GP-1.3 There are parallel pathways of perception. GP-1.4 Perception is influenced by individual physiology.
GP-2 Perception is approximate.
GP-3 Perception is influenced by cognitive processes. GP-3.1 Perception is influenced by expectations. GP-3.2 Perception is influenced by knowledge. GP-3.3 Perception may be influenced by recognition. GP-3.4 Perception is influenced by attention. GP-3.5 Perception is influenced by context.
GP-4 Perception remains constant.
GP-5 Perception can be biased towards one sense. GP-5.1 Attention can affect sensory bias. GP-5.2 Learning can affect sensory bias.
GP-6 Perceptual responses have thresholds. GP-6.1 Weber's Law GP-6.2 Steven's Power Law
GP-7 Perception groups small elements into larger elements.
GP-8 Seven is a magic number.

Table 6-1 A summary of the general guidelines for perception.

GD-1 Emphasise the data.
GD-2 Simplify the display.
GD-3 Design for a task.
GD-4 Iterate the design process. GD-4.1 Avoid designer bias.

Table 6-2 A summary of the guidelines for information display.

MST-1 Use each sensory modality to do what it does best. MST-1.1 Vision emphasises spatial qualities. MST-1.2 Hearing emphasises temporal qualities. MST-1.3 Haptics emphasises movement. MST-1.3.1 Point force-feedback only provides temporal information. MST-1.3.2 Tactile displays are not readily available.
MST-2 Use the spatial visual metaphor as a framework for the display.
MST-3 Increase the human-computer bandwidth. MST-3.1 Use complementary display. MST-3.2 Avoid redundant display. MST-3.3 Avoid conflicting display.
MST-4 Consider sensory substitution. MST-4.1 Adapt spatial visual metaphors to spatial auditory metaphors. MST-4.2 Adapt spatial visual metaphors to spatial haptic metaphors. MST-4.3 Adapt temporal auditory metaphors to temporal visual metaphors. MST-4.4 Adapt temporal auditory metaphors to temporal haptic metaphors.

Table 6-3 A summary of the guidelines for the MS-Taxonomy.