White Paper

Sinking in a Sea of Pixels— The Case for Pixel Fusion

David Hughes – Silicon Graphics, Inc.

Table of Contents

Abstract

Sophisticated virtual environments have been with us now for over a decade, and they have enhanced the working practices of organisations all over the world, whether finding oil or discovering new drugs. Virtual environments have proved effective at enabling people to make better insights, better decisions and communicate more effectively. However the mounting challenge is managing the continuing advance of data complexity, interfacing with increasingly heterogeneous systems generating visual information and needing to work effectively with colleagues, suppliers and customers in a distributed global working environment. This is actually starting to dilute the effectiveness of virtual environments and will ultimately result in complete information and complexity overload. In response, recent advances in visualisation computer systems architecture and design have enabled a concept called Pixel Fusion that potentially offers a practicable solution that can start to address these challenges and provide a common architecture for global visual interaction in a distributed networked, heterogeneous world. The concept of the pixel is explored as a base common working element and an architecture proposed, referred to here as the 'Pixel Fusion Engine', that exploits the potential of a scalable shared memory architecture to efficiently manage the ingestion, fusion and sharing of pixel data across numerous visual streams and output channels. Visual serving technology is integrated into this to provide an efficient pixel distribution capability, enabling not just collaborative interaction, but also major efficiencies in data management and sharing. It is also proposed that this architecture does not just provide a technical solution to data management and interaction in immersive visualisation environments, but also offers an operational infrastructure that has the potential to enhance the way organisations function and communicate on a global level.

The Success of Virtual Environments

Immersive virtual environments have become a common and vital tool for solving complex problems in a wide range of applications from finding oil to car design. These environments that have been with us for over a decade now have enabled teams of scientists, engineers and researchers to improve operational efficiency dramatically. A combination of a number of key factors has enabled this medium to make the contribution that it has. This has lead to dramatic returns on investment, improved time to insight as well as smoother workflow and operational efficiency. Also they have brought other important benefits such as detecting issues and faults in a design earlier in the design cycle (or at all) and improved communications and teamwork. In many cases they have altered the ways companies organise their procedures to enable multi-disciplinary teams to function more effectively and access larger amounts of more complex data in better ways. This has led to faster, better decisions with resultant cost and efficiency benefits as well as improved product output.

The process behind this phenomenon was elegantly summarised in an article in VR News[1] by Mike Bevan, editor, following an interview with the author at the first SGI Visualisation Summit, hosted at the Telenor Expo in Oslo. To summarise there are three key phenomena at work in an immersive virtual environment that can greatly amplify the productivity and creativity of groups in such environments: peripheral context, information density and group dynamics.

Peripheral Context

Peripheral context is the ability for the human brain to process vastly increased amounts of contextual information in the periphery of our vision, but in a more subconscious manner. This is a vital part of our process of natural comprehension and deduction, but is often necessarily hampered with conventional focal only displays, such as monitors or headsets. When the brain is supplied with both focal and peripheral feeds together (as in real life unless you have tunnel-vision or are wearing blinkers), our brains do an amazing job of piecing together the information. This builds up a combined picture that is greater than the sum of the parts.

Information Density

Information density is the ability to perceive both the 'big picture' as well as the detail within. Before computers, we used to do this all the time, with traditional analogue technologies such as blackboards, blueprints laid out on conference tables or pinned to walls. Even pinning flip chart sheets around the room as the meeting progresses is effective. This enables the team to focus on the current process while still being able to cross-reference to an earlier discussion or related piece of data that came earlier. This is challenging with a traditional computer workstation environment and until immersive virtual environments became available this process was near impossible in electronic form.

Group Dynamics

Group Dynamics in immersive environments have become one of the major contributory factors to operational efficiency. Again this harks back to pre-computer analogue days where team members would typically assemble around charts or plans to discuss and collaborate on a problem. Group dynamics in such an environment would be lively and interactive and virtual immersive environments can now provide similar benefits in a complex electronic environment with access to both vast amounts of complex data and computer processing power. However the natural human interaction benefits remain the same. This works well particularly in multi-disciplinary teams who might not have always enjoyed such good cross-departmental communications and thus fosters concurrent working.

Thus the virtual immersive environment is a facilitative tool that dovetails with our natural social and physiological methods of working as human beings, but with the key difference that it operates in a digital domain with complex digital data sets.

One important aspect of all of these factors is that they are all human, social and ergonomic aspects and are very dependent on an effective user interface with the technology.

Issues of Increased Complexity

While the benefits of working in collaborative environments are well documented, the challenge that our customers have is to work efficiently in geographically dispersed environments. Many of the issues they face today involve bringing together teams virtually with connected sites. In the meanwhile problem size and data complexity increases. Any technology that enables users to interpret data quicker, gain additional insight, promote improved communications, foster collaboration and increase productivity should be of substantial interest. In summary, there is an operational and business need for technologies that enable seamless collaborative visualisation to become a reality.

As computer power grows and network capacity accelerates, the amount, size and variety of data that users are subjected to has increased dramatically. This is leading to an increasingly complex and heterogeneous workspace with a greater variety of modes of data access and information supply. Couple this with a continued acceleration of raw computer power that continues to enable larger data sets to be generated more frequently with higher degrees of complexity and the problem for the user becomes analogous to that of drinking from a waterfall rather than a water fountain. Couple this with increased network capability, where almost anyone can be contacted or make contact, and the user is presented with a dazzling array of options. This leads to a situation where too much varied information can be almost as bad as too little.

The Need for Ubiquitous Computing

However even in this increasingly complex environment, the fundamental requirements of a user in immersive environments still remain the same. These can be summarised in 3 main areas:

Access

Intuitive, transparent access to key resources is essential to enable efficient operations and good decision making. This means not just accessing the data and compute/visualisation resources but also the individuals and teams that enable the correct consultation and information provision. In 'Human Aspects of Computer-Supported Collaborative Working'[2], Kalaswsky describes how effective Virtual Collaborative Environments require spatial awareness, aural, visual and textual communication as well as social communications capabilities through a sense of presence and video and auditory communications. This combined access to key data and key people is referred to here as the 'People/Data Quorum'. Access also refers to interaction with the immersive environment itself and thus the quality of the user interface becomes vital to the effectiveness and resultant benefits of immersive environments.

Comprehension

This is the ability to effectively assess the problem at hand, whether it is the process of gaining insight into the meaning of a data set or the identification of an issue in a particular design. This is often a collaborative, interactive and iterative process. Increased data complexity and network connectivity presents both opportunities and challenges simultaneously. Modes of comprehension and insight are not automatic and the quality of the environment is extremely influential on the end result. Insight into a problem may never occur, unless the correct information is made available, presented correctly with the appropriate cross correlation from other team members. Complex, difficult-to-use interfaces will often work against the user to divert attention and context from the job in hand.

Productivity

Making better decisions faster is key to the operational effectiveness of organisations and is one of the main returns on investment for immersive environments. Improvements should seek to accelerate workflow and improve the decision making process to deliver tangible benefits with reduced costs, shorter time-scales, better products and improved communications and collaboration within an organisation.

Next Generation Immersive Environments

Prof. Guy Claxton describes in the book 'Hare Brain Tortoise Mind'[3] how the brain can operate in two distinct modes, one highly linear and analytical and the other less structured but more 'intuitive'. It is suggested that much of our insightful and creative work takes place in the latter mode. It also describes how it is difficult, if not impossible, to operate in both modes simultaneously and depending on circumstances, mood etc, the brain transitions from one state to the other. Highly analytical tasks such as grappling with a complex user interface (UI) or operating mode will tend to encourage transition to an analytical state, ie the Hare Brain. Smooth uncluttered, intuitive operation will encourage 'the intuitive flow' mode, the Tortoise Mind. One could suggest that an exceptionally well designed immersive environment that provides an intuitive, non-intrusive seamless user interface to an advanced immersive environment with access to all the relevant data and personnel is going to increase the prevalence of the intuitive state of mind. This is significant because it will enhance the way people work with very tangible results, thus increasing the probability of intuitive insight and creative behaviour. However it can be seen that this state of mind transition is a fragile thing and heavily influenced by the quality and non-intrusive nature of the user environment.

Thus the challenge for the next generation of Immersive Virtual Environments is to not only address the rising tide of data quantity and complexity, but to also become more invisible, more accessible and easier to use in the process. Examples of initial projects such as the Continuum[4] from EVL and AGAVE[5] have provided advanced research environments for combining multiple visual devices into a physically integrated workspace that is networked to other facilities.

Ubiquitous Computing

Examining three of the major computer technologies, CPU power, Graphics rendering and disk storage, one can see how the emphasis has changed over the last decade and a half. Initially concern was centred on the raw power of individual components; how fast a CPU runs, the power of a graphics pipeline or performance of disk access speed. Next ways were sought to meet the rising demands of problems by exceeding the 'laws' of individual component performance, such as Moore's Law[6], by merging multiple entities together. Thus what emerged were technologies such as multiprocessing, scalable graphics and disk array technologies enabling greater than individual component performance and capacity. Now, more recently, we see a drive towards ubiquitous technologies that address the heterogeneous and distributed nature of computer systems. Not only do we have powerful scalable computer systems, but we now also need transparent, seamless ubiquitous access to these systems regardless of user location or computer brand. Hence we now see the growth of such technologies such as GRID computing or SGI® CXFS™[7] data access technology. However in the area of visualisation, it is still usually necessary to be physically present near the rendering engine to participate in a visualisation session.

Not Just Ubiquitous but Intuitive

One of the goals of the next generation of Immersive Virtual Environments needs to be able to address the challenges that face us in today's rapidly changing electronic world, while balancing the very human needs of users that do not evolve as quickly. Humans still have brains that need to be coaxed into working efficiently. Humans still need to communicate with fellow humans using the basic tools and techniques that we have evolved over the millennia with eg eye contact, body language, and emotion. Fabri, Moore and Hobbs describe how the role of emotion in human interaction is key and remains pertinent to immersive environments[8]. Also the last few metres of interface is always analogue, whether it be a pixel resolved to light photons on a screen or a hand gesture to controlling a 3D wand. We are still fundamentally analogue beings and thus the final digital to analogue interface to the user is crucial.

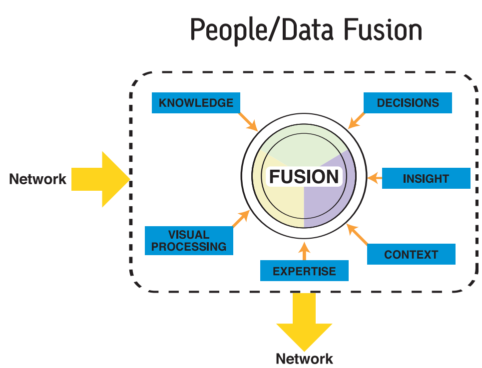
Intuitive human communication remains vital to effective working in virtual environments and ideally these would be as easy and natural to use as talking to a person sitting in the same room. Many advances in video conferencing systems have been made with both commercial technologies and academically developed IP based installations, such as VRVS[9] and the widely deployed Access Grid[10].

Research projects such as 'Office of the Future'[11], the EU funded VIRTUE[12] and the National Tele-immersion Initiative (NTI)[13] have also explored the use of integrated video conferencing in 3D in dedicated environments. However true integration of remote users into a 3D operational workspace in an open-systems architecture that is sufficiently robust for commercial and industrial use remains elusive.

The Pixel Fusion Environment

The Pixel Fusion Environment (MFE) is proposed as a scalable architecture that can address the requirements of the next generation Immersive Virtual Environments. In summary the challenges it is intended to address can be summarised as:

- Fuse large numbers of heterogeneous pixel streams into a common workspace
- Scale to meet the growth in data sizes and complexity
- Provide ubiquitous access to data, remote displays, people and visualisation resources
- Enable and maintain 'intuitive mind-flow' by providing a common virtual workspace that enables an intuitive user interface for ease of operation and interaction.
- Natural human to human interaction, while interacting with data – establish a people/data quorum



The MFE is based upon a single underlying premise of the pixel as the basic building block and a 3-stage process of Ingestion, Fusion and Distribution.

The Pixel as Building Block

The pixel is a powerful thing. A few of these together have the power to save a life, start a war, discover a new world, topple a government or influence a nation. They can be used to communicate, educate and even assist in solving highly complex problems. The pixel is the basic building block of electronic visualisation. However the pixel is agnostic. It really doesn't carry much information at all and nothing about what it is representing, it is just colour and perhaps some depth and transparency information. It could be likened to an atom or an Internet Packet; it isn't until it is packaged up with others and presented in context that it makes any sense. However like an Internet packet, knowledge of what it represents is not required to process it. In fact this gives tremendous freedom and flexibility and enables a common infrastructure to be designed that can transparently manage packets regardless of whether they represent an email, an image or a video stream. This approach of a relatively simple scalable architecture for managing these elements highly efficiently regardless of content or context is what has enabled the Internet to grow into what it is today. It also means that the Internet can be a highly heterogeneous environment with many different devices and brands, but they can all inter-operate leveraging a few simple standards. In a similar way the Pixel Fusion Environment seeks to provide a similar scalable architecture, but with the pixel as the basic building block element and a server architecture that facilitates flow of information around the visual network.

The main building block for a Pixel Fusion Environment is the pixel server or here referred to as a 'Pixel Fusion Engine'. It has 3 primary functions:

Ingestion

Pixel data is read into the system regardless of location or source, whether this is supplied by a local, physically connected device, an Ethernet network connection or natively generated by a local image generation capability. As far as the server is concerned they all end up in the same place in a common memory space where pixels are located and stored in a combined buffer space. From here they can be selected, processed and managed and subsequently displayed or redistributed.

Fusion

The process of Fusion is where pixel data is brought together into a Common Operating Picture or now commonly referred to as a Common Relevant Operating Picture (CROP). This implies an additional degree of filtering and intelligence applied to the pixel streams to present pertinent information to the user(s) as required. Intuitive organisation of the pixel information here is vital to ensure that substantial streams of rapidly changing information can be assimilated and assessed quickly. Intuitive interaction and data navigation tools enable rapid selection and filtering of key elements as well as the ability to seamlessly interchange detail and context. Pixel stream selection, organisation, manipulation and interaction are key to the fusion process.

Fusion takes place in a common virtual workspace with interactive control of positioning and selection of pixel streams. Full interactive control of devices and source streams is performed virtually from an intuitive user interface. This interface can also be shared and controlled remotely.

Distribution

This is not just about input and fusion processing, but also the need to share and collaborate with other users in remote locations. This is essential to gain the full benefit of immersive environments. Many of the Immersive Virtual Environment facilities that have been built to date are largely individual operations and make limited or no use of any distributed capability for day to day operational use. However the difference between a distributed group of facilities inter-operating and an individual facility in stand-alone operation is substantial. This collective phenomenon can be extraordinarily powerful, and is largely under-exploited in a production environment. In Rick Steven's presentation, Scientific Workspaces of the Future[14], it is proposed that benefits of such distributed group working environments of 'Radical Collocation' can produce benefits of 2 to 10x productivity gains, higher quality output, reduced changes over time and faster, better decisions. However a critical mass of deployment has yet to be reached, network operations are a key goal for the future of Immersive Environments. Limited adoption to date is largely due to the complexity of establishing and seamless and ubiquitous connectivity.Pixel Fusion presents an opportunity to manage this effectively by leveraging recently developed 'visual serving' technologies that enable multichannel, real time OpenGL® visualisation pixel streams to be shared and controlled remotely in a largely client agnostic environment in a Visual Area Network[15] or 'VAN'. Using a combination of Ingestion, Fusion and Distribution, sophisticated virtual environments could be shared and inter-operated with fairly minimal setup. Moreover these sessions would have the ability to embed fused input streams into combined streams being served back out to remote locations. This means that remote users would be able to collaborate and control applications and systems served from the source location.

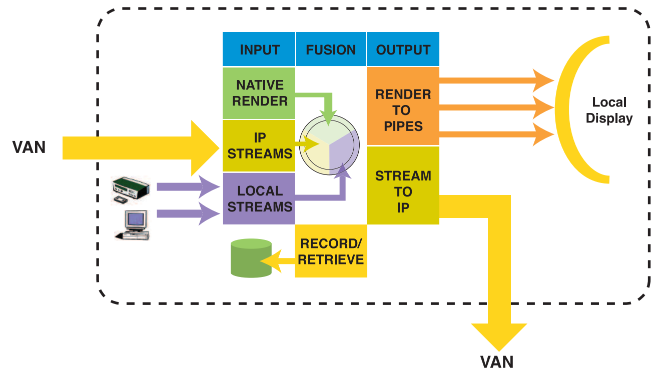
The distribution function also encompasses the ability to mirror, record and replay and subsequently embed multichannel immersive sessions at remote locations. Thus this enables 'nested fusion', where individual application windows and whole sessions can be consecutively embedded into other remote sessions. This introduces the concept of an immersive environment within an immersive environment and so on.

Building a Pixel Fusion Environment - the Pixel Fusion Engine

MFE Server Architecture

To enable a Pixel Fusion system architecture on a traditional high-speed (>=100 BaseT) network, a processing server is required that can perform the tasks of Ingestion, Fusion and Distribution. This server is referred to here as a Pixel Fusion Engine. This acts as a server, processor and router of pixel data on the network, but can also perform additional local functions such as image generation, applications serving and numerical processing as well as providing the user operating environment for management of the Common Relevant Operating Picture (CROP).

Pixel Fusion/VAN Routing Architecture



Pixel Fusion Engine Technical Criteria

Recent advances in hardware and software have made MFE now possible and economically viable.

Key enabling technologies in the 3 main areas of function are required as follows to enable the construction of an effective Pixel Fusion Engine:

Ingestion

- High performance digital media capability for ingesting real time pixel streams in high quality
- Scalable 'media bus' architecture enabling parallel ingestion of high numbers of multiple streams
- High performance, scalable, shared memory single system image architecture (SSI) to buffer and composite pixel streams
- Scalable multi-processing CPU architecture required to process pixel streams in real time

Fusion

- High performance, scalable, shared memory single system image architecture to buffer and composite pixel streams and to efficiently render multichannel outputs
- Scalable multi-processing CPUs to process pixel streams in real time and support rendering tasks
- Scalable real time graphics architecture to process, fuse, composite and present Common Operating Picture
- Software APIs to support real time multipipe rendering, and pixel/media processing

Distribution

- Visual Serving hardware and software to enable high performance remote display and control of fused media streams
- Scalable graphics architecture to support local and remote rendering
- A shared memory single system image architecture again to act as the hub for pixel flow through and out of the system
- High performance scalable storage systems to enable real time session record and replay of fused sessions both locally and remotely (via Visual serving technology)
- Seemingly homogeneous, but actually heterogeneous shared storage architecture enabling transparent data and file sharing across the Visual Area Network

It can be seen from the above that an extensible, scalable, transparent MFE architecture lends itself particularly well to a scalable, shared memory architecture. A fundamental requirement is that to perform efficient and flexible fusion, a common, expandable, high performance buffer space is required to provide a common location for management, composition, rendering, fusion, processing, storage and subsequent redistribution of pixel data streams. This avoids unnecessary duplication and copying of pixel steams from one part of the system to another.

Combining all of these capabilities into a single system architecture provides the basis for the Pixel Fusion Engine, which drives pixel flow on the network and powers the local Immersive Environment. This enables Pixel Fusion to take place enabling the user to access, interact with, process and share all types of media information and actively participate in the decision making process.

The Visual Portal

Generic client terminals and facilities can also participate with the network of Pixel Fusion servers, but provide no processing capability themselves. This introduces the concept of a 'Visual Portal' which is highly generic, uses off the shelf commercially available hardware but enables the Visual Portal user to access the full power of computing, visualisation and distributed data storage on the Visual Area Network. Thus they become a fully active participant in the Pixel Fusion Environment with minimal hardware and software infrastructure. This is very powerful, because it enables geographically independent access to all that MFE offers with no movement of data files, installation of applications or complex configuration, all from a generic low cost facility.

Evolving an Enhanced Workflow Paradigm

So what would working in a network supported by a Pixel Fusion Environment be like? Imagine being able to access any application or device resident to the network, anywhere as part of your own work session. Imagine being able to collaborate with other users with any application and cross correlate it with your own data and share that remotely too. Imagine being able to conference in other users into your common workspace regardless of their location or communications technology and integrate them into the overall workflow as if they were in the room. Add this to the ability to use advanced fusion techniques to enable 'personal interaction' capturing intent and social communications to reproduce a natural round table working environment. Session record and replay functions would enable local or remote groups to perform post mortems, tele-briefings and remote training sessions. This would also enable asynchronous working either in a common location or spread around the globe in different time zones. For example this might apply to a chain of design studios in Milan, LA and Tokyo working around the clock on a key automotive project. All of this while leaving the data effectively static and using the computer power near the data and moving only pixels around the visual network.

Concluding Remarks

Pixel Fusion has the capacity to enable seamless working in a heterogeneous work environment with a variety of different clients, enabling intuitive access to larger and more complex data. To users, at each node in the network, this would appear homogeneous, offering tremendous flexibility and ease of use.

This results in immersive environments that are easier to use, accelerating workflow and improving productivity, enabling groups of users to manage more data, more complex data and more diverse visual data. This will lead to greater insight, faster better decisions and improved cross-organisation communications and collaboration.

References

[1] VR News, ISSN 13360-3485 Volume 7 Issue 4: May 1998
[2] *http://www.avrrc.lboro.ac.uk/index.html*
[3] Hare Brain, Tortoise Mind: Why Intelligence Increases When You Think Less – by Guy Claxton, ISBN: 1857027094, published by Fourth Estate
[4] *http://www.evl.uic.edu/cavern/continuum/indexmain.html*
[5] Leigh, J., Dawe, G., Talandis, J., He. E., Venkataraman, S., Ge, J., Sandin, D., DeFanti, T. A., AGAVE : Access Grid Augmented Virtual Environment, Proc. AccessGrid Retreat, Argonne, Illinois, Jan, 2001.
[6] *http://www.intel.com/research/silicon/mooreslaw.htm*
[7] *http://www.sgi.com/products/storage/tech/file_systems.html*
[8] The Role of Emotion in Collaborative Environments - Marc Fabri (Leeds Metropolitan University) – WACE 2002.
http://www-unix.mcs.anl.gov/fl/events/wace2002/material/wace-02-papers/fabri-role.pdf
[9] VRVS: “Virtual Room Video-Conferencing System”, Home Page, *http://www.vrvs.org/About/index.html*
[10] *http://www.accessgrid.org/*
[11] Raskar, Welch, Cutts, Lake, Stesin, Fuchs. “The Office of the Future: A Unified Approach to Image-Based Modelling and Spatially Immersive Displays”, proceedings SIGGRAPH 98, Orlando Florida
[12] VIRTUE *http://www.ece.eps.hw.ac.uk/~mtc/research.html*
[13] Amela Sadagic, Herman Towles, Jaron Lanier, Henry Fuchs, Andries van Dam, Kostas Daniilidis, Jane Mulligan, Loring Holden and Bob Zeleznik, “National Tele-Immersion Initiative: Towards Compelling Tele-Immersive Collaborative Environments”, presentation given at Medicine meets Virtual Reality 2001 conference, January 24-27, 2001, Newport Beach, California
[14] Scientific Workspaces of the Future – Rick Stevens – WACE 2002 -
http://www-unix.mcs.anl.gov/fl/events/wace2002/material/wace-02-talks/talk-rick-swof.pdf
[15]*http://www.sgi.com/visualization/van/whitepapers/index.html*

Corporate Office
1500 Crittenden Lane
Mountain View, CA 94043
(650) 960-1980
www.sgi.com

North America +1 800.800.7441
Latin America +55 11.5185.2860
Europe +44 118.912.7500
Japan +81 3.5488.1811
Asia Pacific +1 650.933.3000

3804 [06.2006] J14909