Competition Heats Up in 3D Accelerators

Market Hungers for a Unified Multimedia Platform from Microsoft

by Yong Yao

This is the second of two articles on PC three-dimensional graphics. The first part (see **100103.PDF***) provides background for understanding 3D graphics. This article discusses 3D products from 3Dfx, 3Dlabs, Chromatic, Nvidia, Oak, Rendition, S3, Trident, and VideoLogic.*

With the PC market moving quickly to 3D graphics, a torrent of new products has appeared to improve performance on 3D software. Although 3D accelerators have been common for years in workstations, the new products share a common goal of bringing high-performance 3D to PCs without adding significant cost. As this market is just emerging, the new chips display a variety of technical and business models for achieving this goal. The key business challenge is to pull together the chip design, manufacturing, and sales.

From the technology side, 3D chip makers must be concerned about a variety of issues. Choosing the right DRAM technology is critical to achieving the required bandwidth. Overall rendering speed must exceed the performance of the host CPU and be competitive with other accelerators; performance should also scale well as the host processor or memory system is upgraded. To meet the required price point, roughly $250 for an add-in card, vendors must consider combining the frame buffer, texture buffer, and Z-buffer. Another cost-reduction tactic is to combine 3D acceleration with other standard functions, such as 2D and video acceleration and possibly audio.

The key issue remaining is software. There is no 3D software standard for PCs today. Some of the early players, such as Creative Labs, are trying to establish proprietary software schemes, and any developer selling chips today has to seek ISVs to write applications specifically for its chips. In short, vendors must have a plan to attract applications. For most, this plan is to develop drivers for Microsoft's Direct3D API and hope that Microsoft ships it soon and that ISVs code to this emerging standard.

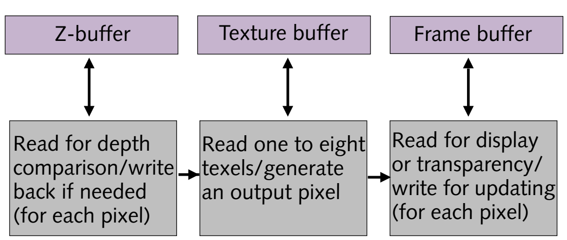

Figure 1. Data flow in a simplified rendering pipeline for 3D hardware accelerators in most of today's graphics workstations.

There is an urgent need for a unified multimedia software platform. OpenGL is a great 3D API, but it carries some overhead and runs slowly under Windows 95. Microsoft does not have a plan to fine-tune OpenGL for Windows. Although Silicon Graphics is working on its strategy for pushing the OpenGL technology to PCs, Direct3D is likely to prevent OpenGL from becoming the main 3D API for the PC. The overall multimedia software platform will be Microsoft's DirectX, which includes Direct3D. The problem is that Direct3D is late, and it is not clear how soon accelerator vendors will have drivers for it. Until we see whether Direct3D meets its goals, sales of PCs with 3D accelerators, and the number of 3D titles for PCs, will be limited.

Hardware Accelerators Share a Common Goal

Adding 3D capability to a PC is not just an incremental increase in functionality. Like the move from a text-only interface to today's graphical user interface (GUI), 3D is a feature that will enable a new set of applications. These applications include 3D animation, virtual reality, architectural walk-through, mechanical design, and, of course, games. Using 3D offers viewers more freedom, enabling them to go around and behind objects, much as in real life. But 3D graphics is an order of magnitude more complex and compute intensive than 2D graphics.

Although none of the 3D vendors covered are taking the same approach toward 3D hardware implementation, all share the same goal: to deliver Jurassic Park–like 3D performance while meeting PC price points. The design trade-off is between affordability and realism.

There are many ways to reduce the cost of a traditional workstation 3D subsystem without significantly impacting its graphics performance. For instance, Figure 1 illustrates the data flow in a simplified rendering pipeline for a typical workstation 3D accelerator. When performing texture-mapped 3D graphics, memory can be allocated in each of the three buffers: the frame buffer, texture buffer, and Z-buffer. For decent 3D graphics, each of these buffers must be 2M or larger. Even using commodity DRAM, these three buffers alone will cost more than $160, the cost of an entire 3D add-in card for the retail PC market, preventing this approach from being effective in the PC market.

Simply getting rid of these three buffers or combining them without careful design will not meet the 3D performance requirements. Products like those from Oak and VideoLogic address the high cost of memory by eliminating the need for an external Z-buffer without losing too much performance. Other products discussed below innovate in different ways to achieve the common goal.

Features in Common or for Differentiation

The products discussed in this article all have a glueless PCI interface and share the following 3D features:

- Alpha blending: zero, one, and source alpha
- Filtering: point-sampled and bilinear filtering
- Primitives: points, lines, and triangles
- Raster capabilities: two-op, line draw, and polygon fill
- Shading capabilities: flat and Gouraud shading
- Texture formats: 1-, 2-, 4-, and 8-bit palletized textures
- Textures: animated textures, texture mapping with perspective correction, and MIP mapping

Other common 3D features are double buffering and dithering. Table 1 lists additional 3D features that are not common among all the products covered.

Products that integrate a 2D accelerator all deliver conventional Windows acceleration and provide the logic required for a standard VGA-compatible graphics card. Products that integrate a video accelerator all perform YUV-to-RGB color space conversion and image scaling, features that are useful for accelerating software-based MPEG video playback. Besides integration and rich feature sets, there are other implementation strategies for product differentiation.

One way to resolve the cost of multiple buffers is to perform Z-buffering without a Z-buffer. In general, Z-buffering provides simpler database generation by allowing interpolation of surfaces for hidden-surface removal to create more complex and realistic objects with fewer polygons than non-Z-buffering. For cost-sensitive systems, some 3D applications can run without the Z-buffer. For instance, most of today's games do not require a Z-buffer. But the best of both worlds is to support Z-buffering without paying the cost of a Z-buffer.

Another way to reduce cost is to hold the frame buffer, texture buffer, and Z-buffer in a single memory. That memory can then be allocated dynamically under software control. For example, if an application does not need a Z-buffer, the available memory can be used for textures or display.

Other products differentiate themselves by handling the setup part of the 3D geometry processing in hardware, which has several advantages. It removes the performance bottleneck in the geometry process, since in some cases the 3D setup can consume a great deal of computing power. This strategy also reduces PCI traffic, since instead of sending all 3D primitives over the PCI bus, only 3D vertices need to be transferred.

We believe more companies will realize the benefit of having a 3D hardware accelerator handle the setup part of the 3D geometry process. Designs that rely on the host processor to perform setup calculations for each triangle quickly become CPU-bound, as well as PCI-bound, when rendering speed increases.

Programmable 3D: Chromatic and Rendition

Programmability makes it easy to adopt new software without worrying about hardware obsolescence as standards evolve. Programmable processors, also called media processors, present OEMs with a new way of accelerating multimedia functions like 3D graphics, potentially offering more features for less cost. In contrast, a hardwired 3D architecture may suffer as APIs, data formats, and algorithm technology change over time.

Multimedia companies taking the programmable approach include Chromatic, IBM, Philips, and Rendition. Since IBM's Mfast *(see* **091601.PDF***)* and Philips' TM-1 *(see* **091506.PDF***)* are far from being available, this section discusses only Chromatic's Mpact and Rendition's Verite. Mpact is based on a VLIW (very long instruction word) architecture, whereas Verite is RISC-based. These two companies write their own software to perform 3D acceleration on their programmable engines. Essentially, the programmable engines provide the computing power required for this software to run. The functions of these products depend on how the software is written.

Mpact Leads Multimedia Integration

The Chromatic Mpact media processor *(see* **091404.PDF***)* is designed to efficiently process "natural data types" such as graphics, video, and sound. For these data types, the media processor can perform up to two billion integer operations per second, about ten times more than today's high-end Pen-

	3Dfx Voodoo	3Dlabs Permedia	Chromatic Mpact	Nvidia NV3	Oak OTI64311	Rendition Verite	S3 Virge	Trident T3D9695	VideoLogic PowerVR
Anti-aliasing	Yes	No	Yes	No	Yes	Yes	No	No	No
Atmospheric effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Via Apha	Yes
Memory required	2M	2M	2M	1M	1M	2M	2M	2M	2M
Memory type	EDO	SGDRAM	RDRAM	SGRAM	EDO	EDO	EDO	SGRAM	EDO
Sub-pixel positioning	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Texture compression	Yes	No	No	Yes	Yes	Yes	Yes	No	No
Texture buffer needed	Yes	No	No	No	Yes	No	No	No	Yes
Trilinear filtering	Yes*	No	Yes	No	Yes	Yes	Yes	Yes	No
Z-buffer needed for Z-buffering	Yes	No	No	Yes	No	No	No	No	No

Table 1. 3D feature comparison among the nine newly announced 3D products shows 3Dfx and Oak provide the most complete feature set, while 3Dlabs' Permedia and Trident's 9695 deliver only the bare necessities. *3Dfx uses an equivalent called advanced filtering. (Source: vendors)

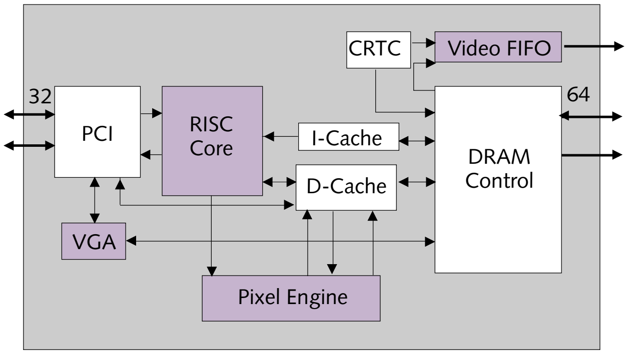

Figure 2. The internal block diagram of Rendition's Verite, a RISC-based programmable 3D accelerator, is similar to that of a highly integrated microprocessor, with on-chip cache, DRAM control, and a PCI interface.

tium processors. To fill a 792-bit-wide internal data path, which is linked via high-speed connections to different levels of on- and off-chip memory, the media processor relies on a 500-Mbyte/s (peak) Rambus interface.

Mpact does not handle floating-point computing; it relies on the host processor to execute floating-point calculations. Thus, its 3D performance is somewhat limited by the host CPU's floating-point capability. For example, Mpact cannot handle the setup part of 3D geometry processing, which requires floating-point operations.

For 3D, Chromatic adopts a unified buffer architecture, combining the frame buffer, texture buffer, and Z-buffer within a single memory, which is also used for 2D and video functions. The Mpact feature set is detailed in Table 1.

Chromatic calls its Mpact software Mediaware. It is modularized and can perform multiple functions: video, 2D, 3D, audio, fax/modem, telephony, and videophone. The Mediaware software modules enable these seven functions and provide various levels of capability within each function. The modules contain assembly-language microprograms for the media processor as well as for the operating system and needed drivers.

Chromatic has an interesting business model. The company relies on its semiconductor partners to manufacture and sell the Mpact chips. So far, Toshiba and LG Semicon are the only two vendors that have publicly announced support for the Mpact architecture. The agreements with its partners do not prevent Chromatic from bringing additional manufacturers on board. Chromatic will enjoy some royalty revenue from its partners' sales, but its main source of income will be from the sale of Mediaware modules directly to PC OEMs. These modules will be installed on a PC's hard disk, allowing an OEM to produce several models with varying multimedia functions, all using the same basic motherboard design.

Chromatic has begun alpha testing of both the Mpact media processor and the Mediaware software; Mpact-based motherboards are up and running today. Production-ready Mpact silicon and Mediaware modules are slated for 3Q96. Chromatic claims that five of the top six home PC makers worldwide are working on designs using the Mpact product. For the ISV community, besides making those developers aware of Mpact's capabilities, Chromatic depends on Microsoft to push various types of software development.

Verite Uses RISC-based Design

Instead of using a VLIW instruction set, Rendition adopted a RISC-style instruction set for its Verite product, illustrated in Figure 2. The Verite RISC core handles geometry setup and filtering functions, while its pixel pipeline handles pixel-level texture mapping, blending, and other 3D operations that must occur at hardware speeds. The key architecture differentiation is that Verite offloads the host processor from doing the setup part of geometry operation.

Verite is a single chip integrating the programmable 3D engine, a VGA-compatible 2D accelerator, and a digital video accelerator. An important piece missing in Rendition's design is audio capability. Most consumers interested in 3D will use their 3D machines for games. Without high-performance sound, these games will not be too exciting. Verite must depend on the host processor or an audio chip for sound processing.

Verite's internal RISC processor can issue multiple instructions per cycle and has an expanded graphics instruction set. These graphics instructions, coupled with the large register file (128 × 32 bits), significantly reduce the instruction count of graphics rendering algorithms.

The Verite RISC is similar to other RISC microprocessors. Its RISC engine has the following features:

- Interlocked 32-bit integer processor
- Single-word, multiple-operation instructions
- Linear 32-bit address space
- Combinatorial 32 × 32 multiply
- Delayed load with large register file
- Register file can be used as memory scratchpad
- Delayed non-annulling branches
- Load/store double-word, word, unsigned half-word, and unsigned-byte operands

Compared with Chromatic's Mpact, the Verite programmable engine is tuned mainly for 3D graphics. It has less flexibility in supporting other programmable features than the Mpact processor.

Like the Chromatic chip, the Verite processor combines the frame buffer, Z-buffer, and texture buffer. Unlike Mpact, which is based on RDRAM, the Verite media processor connects to a single bank of relatively inexpensive EDO DRAM. Because of its built-in DMA engine, Verite can also read textures from the PC's main memory.

The potential cost savings from these key features are attracting quite a number of multimedia board makers and title developers to the Verite platform. Creative Labs has committed to Verite for the next generation of its 3D Blaster entertainment platform; it has ported the Creative Graphics Library (CGL) API to Verite. Number Nine is also working on a Verite-based board. Software vendors like Domark, Mindscape, and Terminal Reality are all planning to release next-generation game titles that take full advantage of the Verite performance.

Hardwired 3D: 3Dfx, 3Dlabs, and Nvidia

Often a hardwired product has higher performance than an equivalent programmable one. A programmable product has advantages in flexibility, but only if the hardwired vendors cannot revise their designs quickly enough to keep up with the evolution of standard algorithms and APIs.

Like Chromatic and Rendition, 3Dfx, 3Dlabs, and Nvidia are all fabless startups that aim to provide 3D solutions to the PC market. These three companies, however, are taking very different hardware approaches from those of Chromatic and Rendition. Their products are fixed in functions and features and depend on host-resident driver software to track changing APIs.

3Dfx Does Voodoo Graphics

The only “pure” 3D company today is 3Dfx. Its first product, Voodoo, does nothing but accelerate various 3D operations. The Voodoo chip set consists of two devices: Pixelfx and Texelfx. The Pixelfx chip includes a triangle engine, a pixel pipeline, and interfaces to PCI and a Z-buffer. The Texelfx chip performs texturing functions such as trilinear filtering, MIP mapping, transparency, and translucency; it connects directly to texture memory. The Voodoo chip set has some features, such as detailed and projected textures, texture morphing, and texture compositing, that are not offered by most existing PC 3D implementations. In terms of 3D performance, Voodoo is one of the highest-performing chip sets for PCs, as Table 2 shows.

Voodoo divides various 3D tasks among the host CPU and its two chips. It is the CPU’s responsibility to handle backface culling, transformations, lighting, viewpoint clipping, polygon clipping, and texture management. The Voodoo chip set deals with triangle setup, rasterizing polygons, pixel texturing, texture smoothing, texture combining, pixel fog, depth testing, and alpha blending. Like Rendition’s Verite, Voodoo offloads the host processor by taking care of triangle setup, one of the most compute-intensive parts of the 3D geometry operation.

Recognizing that gaining PC OEM business is tougher than selling to arcade and game-console vendors, 3Dfx takes a two-step approach. The company starts by winning designs in the arcade market, where its product’s cost is insignificant compared with the selling price of coin-operated arcade machines. For such systems, 3D performance is critical. Most of the company’s key contributors are ex-SGI employees who have experience designing high-performance 3D graphics. 3Dfx attracts the arcade-system vendors by allowing them to use high-end Pentium PCs to replace their expensive proprietary engines.

Another advantage to attacking the arcade market first is that the arcade is an existing market for 3D, whereas 3D is an emerging technology for PCs. In addition, most current high-end 3D titles are developed for arcade applications. Those title developers will have more incentive to develop 3D titles for 3Dfx’s platform, since, with slight modification, their products can work on both coin-op arcade machines and a growing base of multimedia PCs.

The second step is to push the Voodoo 3D architecture into the PC market. In the mainstream PC market, there is no room for standalone 3D chip sets; what PC makers need is a complete multimedia solution that must include at least 2D and video along with 3D. To address the incompleteness of the initial Voodoo product, 3Dfx recently established a

	3Dfx Voodoo	3Dlabs Permedia	Chromatic Mpact	Nvidia NV3	Oak OTI64311	Rendition Verite	S3 Virge	Trident T3D9695	VideoLogic PowerVR
2D acceleration	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Video acceleration	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Audio capability	No	No	Yes	Yes	No	No	No	No	No
On-chip RAMDAC	No	No	No	No	Yes	No	Yes	Yes	No
TV/VCR interface	No	No	No	No	Yes	No	No	Yes	No
Telephony	No	No	Yes	No	No	No	No	No	No
Fax/modem	No	No	Yes	No	No	No	No	No	No
Pixel-fill rate (million)	45-90	25	TBD	66	33	25	17	37.5	TBD
Polygon-rendering (50-pixels, textured)	655,000	500,000	TBD	500,000	660,000	150,000	340,000	750,000	550,000
Price for chip	$75	<$50	$25-$50	$45	TBD	$75	$40	<$40	<$50
Price for software	Free	Free	$10-$40	Free	Free	Free	Free	Free	Free
Price for board (est)	$350	$250	$150	$90	TBD	$300	$200	$200	$250
Packaging (PQFP)	208/240	256 (BGA)	240	208	208	240	208	208	208
Production date	June 96	2Q96	3Q96	Oct. 96	2H96	1Q96	Feb. 96	4Q96	3Q96
Contact number	(415) 934-2425	(408) 436-3456	(415) 254-0729	(408) 720-7132	(415) 962-9550	(415) 335-5900	(408) 980-5400	(415) 691-9211	(415) 875-0606

Table 2. Functions, performance, price, and availability among the nine newly announced 3D products. These figures show the 3Dfx Voodoo and Trident 9695 as the best performers, but the performance numbers could be misleading because some vendors may quote numbers with certain features disabled. TBD indicates data not available. (Source: vendors)

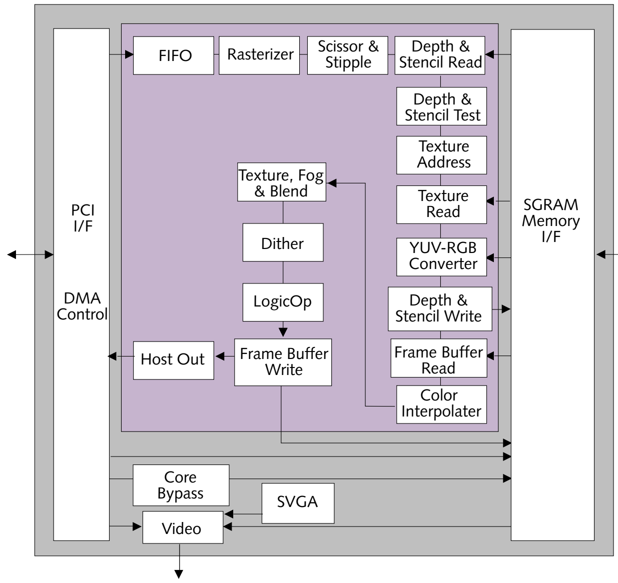

Figure 3. The Permedia internal architecture and graphics pipeline contains a variety of specialized function blocks.

partnership with Alliance Semiconductor. In the short term, this partnership will focus on ensuring compatibility between the Voodoo chip set and Alliance's ProMotion graphics and video controllers. In the long term, the joint effort aims to combine Alliance's DRAM technology, 2D graphics expertise, and PC market experience with 3Dfx's 3D technology into an integrated multimedia solution.

Permedia Aims at Mainstream PCs

Although 3Dlabs was formed from a buyout of Du Pont Pixel in April 1994, it has a long history. If one counts the original Du Pont development and its investment of more than $40 million, the firm has accumulated 10 years of experience in designing hardware and software for 3D graphics. The Permedia device is the company's third-generation product, following Glint and the 3D Blaster chip, which it designed for Creative Labs. Permedia is less expensive than Glint and faster than the 3D Blaster chip.

The Permedia feature set was refined based on feedback from more than 200 software developers for 3D Blaster. New features include chroma-keying and specular lighting. Permedia also adds both 2D graphics and video acceleration to the 3D Blaster feature set, and it replaces that chip's VL-bus with a PCI interface. The chip's internal architecture is illustrated in Figure 3.

The company picked SGRAM (synchronous graphics DRAM) as its memory technology with the anticipation that synchronous memory technologies will dominate PCs in 1997 and 1998, because these chips can scale to frequencies of 100 MHz and beyond. Permedia uses 50-MHz parts now but will enhance performance by taking advantage of faster SGRAMs as they become available.

Like Rendition's Verite, Permedia integrates a VGA-compatible 2D graphics accelerator and video-acceleration functions. Using SGRAM block fills, the Permedia 2D graphics engine can achieve a 2D fill rate of 1.6 Gbytes/s, which is much faster than Verite's rate. In addition, Permedia has a built-in DMA engine that can download textures and command streams from main memory via the PCI bus at rates up to 100 Mbytes/s.

3Dlabs has an interesting marketing strategy, which is to use Creative Technology, a $1.2 billion company, as its board-development partner to test out its technology before it pursues the general market. As one of the first companies to enter the PC 3D business, 3Dlabs needed a partner that could drive ISVs to provide custom support.

The company also licenses its technology to companies that want to integrate 3D into their system silicon, although 3Dlabs expects most of its revenue to come from chip sales. Its licensing support also includes channel, fab capacity, sales force, and technology exchange. To make this business model work, its chip designs are done completely in VHDL. They are modular and abstract.

3Dlabs will maintain two product lines: Glint and Permedia. Glint is for the professional workstation market, optimized for performance, where OpenGL is the main API. Permedia is for the mass consumer market, optimized for low cost and performance, where Direct3D is the emphasis.

The company has realized the benefit of moving the 3D setup operation from the host CPU to external hardware. It has been developing a separate geometry-engine chip for months now. The first silicon of this chip is now working in the lab. Future devices will implement a unified memory architecture, putting the texture buffer, Z-buffer, and frame buffer into the system's main memory. Dynamically allocating memory for various function blocks within the system from a central memory may be the ultimate way to resolve increasing demand for memory while keeping the overall system cost low.

Nvidia Supports Multiple Polygon Models

Among the products discussed in this article, Nvidia's NV1 is the only one being shipped in volume today. The NV1 design *(see* **090904.PDF***)* requires only a frame buffer, the same buffer used for the 2D graphics function. The Z-buffer is optional.

In Nvidia's architecture, textures are stored in system memory and for rendering are transferred across PCI via bus-mastering DMA. System memory is swapped out as needed by the application. Texture mapping starts with the texture and computes the destination pixels to which these source texels should be mapped. This algorithm, called forward rendering, accesses the source texels in consecutive order by scanline, making these accesses amenable to bursting over PCI. This method allows Nvidia to keep its textures in system memory and still maintain the bandwidth required by the internal rendering engine.

In contrast, traditional 3D texture-mapping algorithms start from the destination pixels on the display space, calculate the mapping to the texture space, read the texture, and apply it to the destination pixel. This algorithm, called inverse rendering, makes burst accesses difficult, because the accesses to texture memory are not in consecutive order. It is not feasible to store textures in system memory if there is no mechanism to get them to the rendering engine quickly.

Nvidia has admitted that if it could market NV1 all over again, it would stress that the NV1 part accelerates triangles as well as curves. Because the company put some initial emphasis on its propriety quadratic texture map (QTM) algorithm, which is similar to NURBS, many ISVs believe NV1 uses a proprietary algorithm, something most ISVs don't want to support.

Actually, NV1 can perform well using regular polygons like triangles and rectangles. The difference is that the NV1 architecture can handle one additional class of primitives —curves—that cannot be rendered by other current 3D implementations. Applications that use the curve support of the QTM algorithm have found better photorealism than those running with that algorithm disabled.

But of the software vendors Nvidia is working with today, none is planning to use the QTM algorithm. Most current software models curved surfaces using many triangles. To improve appearance without using QTM, software can modify the light intensity of pixels so the polygons appear curved and smooth. This method has the advantage of working on a variety of 3D chips, unlike the proprietary QTM.

Nvidia Moving to Second Generation: NV3

Like the other 3D startups, Nvidia has a unique business model. The company arranged for SGS-Thomson to allocate foundry capacity to Nvidia's accelerators, granting SGS an exclusive right to sell the DRAM version of NV1 while Nvidia sells the VRAM version. Nvidia began shipping NV1 products last November. The company claims 100,000 units have shipped, as of 1/31/96. The company has sold more of its VRAM version than SGS-Thomson has of the DRAM version, because the early adopters of the new chip care more about performance than cost.

Sega has already bought into Nvidia's story. Its Virtua Fighter game software was demonstrated at Comdex '95 running under Windows 95. So far, the NV1 is the only chip available that accelerates all the APIs contained in Microsoft DirectX, using only 1M of memory.

The NV1 chip provides 2D acceleration, 3D graphics, wavetable MIDI synthesis, MPEG video acceleration, and a precision digital/analog joystick interface. No host CPU intervention is needed for the joystick, and with a Pentium-133 CPU, the audio engine takes only 1% of the host's power for multitrack simultaneous MIDI and digital mixing.

The built-in DMA and 600-Mbyte/s on-chip packet bus together deliver the data stream required by the multimedia engine. The chip maximizes 2D scrolling performance beyond the refresh rate of the monitor at any resolution. One drawback is that the NV1 is not quite Sound Blaster-compatible. For compatibility, an external device is needed.

Nvidia is now working on its next-generation media accelerator, NV3, which will offer 100% Sound Blaster compatibility. The key feature of NV3 is not its Sound Blaster compatibility, but its superior performance (see Table 2). The NV3 part also integrates the custom RAMDAC that is a separate chip for NV1, reducing system cost. NV3 is a functional superset of NV1, maintaining 100% software compatibility with NV1. Recently, Nvidia announced a partnership with Lexicon to deliver Lexicon's home theater sound into the PC using the NV3 product. The company hopes to ship the NV3 this fall.

Home-grown 3D: Oak, S3, Trident, VideoLogic

These four companies have been providing 2D graphics, video, and other peripheral chips to the PC market for years. With the increasing demand for 3D in PCs, these companies cannot afford to stand still. If they do, they will not only miss many 3D business opportunities but will also lose their existing business to those aggressive 3D startups. The advantages these companies have are their 2D/video technologies and their existing customer bases, as well as their knowledge of the PC business. Their disadvantage is that some of them lack 3D design experience.

This section briefly discusses the OTI64311 from Oak, S3's Virge, the T3D9695 from Trident, and PowerVR from VideoLogic and NEC. Like the Nvidia chips, some of these devices combine graphics and video acceleration with 3D features. We describe these chips as media accelerators.

Oak Launches Its First 3D Chip

Oak, the world's biggest supplier of CD-ROM controllers, also produces high-quality audio and video products. The company recently assembled a 30-person team for its 3D business unit. Its 3D designers come from Kubota, SGI, Sun, and Martin Marietta. Some key technologies were acquired from Kubota, which recently exited the workstation graphics business. In this way, Oak has quickly established its own 3D knowledge base.

The company's first 3D chip, the OTI64311, will combine 2D and video with 3D graphics rendering. Leveraging the company's 2D and MPEG technologies, the OTI64311 is a highly integrated single-chip media accelerator. The chip has a port for connecting directly to an external NTSC/PAL decoder for TV and VCR output and also includes a 170-MHz RAMDAC and dual clock generators. This level of integration should greatly reduce system cost and the board space needed for multimedia acceleration.

The OTI64311 incorporates several innovative memory and computation techniques to enable high-performance 3D graphics with standard EDO DRAMs. The key architectural feature of the OTI64311 is the way Z-buffering is implemented. With a proprietary memory-efficient algorithm, the OTI64311 can perform a normal Z-buffering operation without paying the Z-buffer cost. The 3D features of the OTI64311 media accelerator are listed in Table 1.

S3 Ready to Ship Virge in Volume

Virge stands for Video and Rendering Graphics Engine; it is S3's first hardware accelerator to include 3D support. Virge is a single chip that consists of 3D functions plus the company's popular Trio64 graphics/video chip. The 3D features included in Virge are listed in Table 1.

The chip uses either EDO DRAM or VRAM for the graphics buffers. S3's multiplexed buffering allows Z-buffering without the cost of dedicated Z-buffer memory. This patented buffering scheme can share memory with the frame buffer in the system while minimizing additional memory requirements. A 220-MHz RAMDAC is also integrated into the Virge chip.

Virge is pin-for-pin compatible with S3's Trio64V+, one of the world's most widely implemented 64-bit 2D graphics controllers. This compatibility enables add-in-card developers and system makers to upgrade to 3D acceleration without redesigning their boards. Virge also incorporates the company's proprietary bus for a direct connection to live video or an MPEG decoder.

The business model for S3 has been to go after big customers such as Compaq and Microsoft. The company claims more than 15 hardware manufacturers and more than 10 game developers are supporting its Virge product.

Trident Announces Third Generation: T3D9695

Trident is one of the world's leading graphics/video chip suppliers. Its product strategy has always been to go after the volume market with the best balance of performance and price. Trident gained its experience from its first 3D chip, the T3D2000, which was announced in June 1995. It launched a software developer program that provides developers with evaluation boards based on the T3D2000 as well as documentation and software-development tools for porting to its 3D technology. The T3D9695 is the third generation of products with 3D-acceleration functions. The second-generation chip, the T3D9692, is currently selling.

The newest device integrates 2D, 3D, and video acceleration along with a VMI port for connecting an MPEG decoder and a video port for connecting an NTSC/PAL encoder for TV and VCRs. All these features are packed in a 208-pin PQFP package and manufactured using 0.5-micron CMOS. The key differentiators of the T3D9695 are:

- Pin-compatibility with its popular predecessors, such as the T3D9692 and TGUI9680, provides designers with the flexibility to use a single board design for multiple graphics solutions.
- Like Chromatic's Mpact, it uses a single buffer as frame buffer, texture buffer, and Z-buffer.
- Its high integration and low power dissipation are well suited to multimedia notebook PCs.
- The TV interface helps certain TV-related multimedia consumer products.
- It can work with UMA core-logic chip sets from Opti, SiS, UMC, and VIA, which will help bring basic 3D functions into low-cost PCs.

Trident expects OEMs to build a graphics, video, and 3D accelerator board for less than $200 using the T3D9695. The company is considering reducing the price of future products to the $25 range by cutting some little-used features, such as CAD support.

It is worth mentioning that the T3D9695 is one of the best video-quality chips among those covered in this article. Trident's proprietary edge-smoothing algorithm enables both horizontal and vertical interpolation for the video stream. The chip supports motion video from a CD-ROM at up to 30 frames per second without the need for additional hardware accelerators.

VideoLogic Teams with NEC for PowerVR

VideoLogic has been providing video chips and add-in video cards to the PC market. It is now entering the 3D race with semiconductor giant NEC as its partner. Under its agreement with NEC, VideoLogic is responsible for design and development while NEC handles production and sales. Both companies put effort into product design and marketing. The two have teamed with Namco, a leading arcade-system and game-title developer, to bring arcade-quality software to the PC platform. Namco's first game title running on PowerVR was demonstrated at Fall '95 Comdex. NEC says its sales target is two million PowerVR chips installed in PCs worldwide by the middle of 1996.

PowerVR is the 3D architecture originated by VideoLogic. The key to the architecture is abolishing the Z-buffer and reducing the memory-bandwidth required. The algorithm developed by VideoLogic is based on a so-called infinite plane (surface based) that does hidden-surface removal on the fly. This method sidesteps memory-bandwidth limitations that arise from accessing the Z-buffer.

Because PowerVR has lower bandwidth requirements, it is possible to render large polygons, which are required for some special effects. For example, shadows and searchlights are fundamentally large objects, as they represent light rays traveling long distances. A full shadow, where any object can cast shadows on any other object, and real-time searchlights bring a new dimension to games. In addition to eliminating the cost of a Z-buffer, PowerVR removes the bottleneck associated with accessing such off-chip memory.

Although a frame buffer and a texture buffer are still required in the PowerVR design, their bandwidth requirements are greatly reduced. Each pixel requires an access to the frame buffer only twice: once to be written to the frame buffer, and the second time for being transferred to the display. This is true even for transparent pixels.

Using a "deferred" texturing approach, the PowerVR chip reduces both texture- and frame-buffer bandwidth.

Unlike the conventional rendering system, where every polygon is textured, the PowerVR approach textures only final visible pixels. This technique makes the memory bandwidth independent of the polygon rendering rate. The texture buffer will be accessed one to eight times depending on the complexity of the texture mapping: point mapping requires only one access, whereas trilinear MIP mapping requires eight accesses.

The PowerVR design has the following advantages:

• Cost savings due to elimination of the Z-buffer

• Low memory bandwidth requirements because of the elimination of the Z-buffer, updates of only visible pixels, and the deferred texture approach

• Memory bandwidth requirements independent of polygon-per-second performance figures, since not all intermediate pixels and texels are considered

NEC will manufacture three chips using the PowerVR technology. An image synthesis processor (ISP) and a texture shading processor (TSP) will be used in arcade systems. Performance can be increased by adding more ISPs. For PC applications, another chip, called PCX1, combines one ISP, one TSP, and a PCI interface into a single die. The performance figures of the PCX1 chip are listed in Table 2. A future single-chip solution is also planned for console applications.

Choosing the Right Solutions for 3D Applications

With a dozen 3D graphics accelerators about to appear on the market and the lack of a unified software platform, picking the right 3D solution isn't easy. A 3D board manufacturer or a multimedia PC maker must take the following issues into consideration while choosing its 3D strategy:

Soft 3D. The Pentium-166 is already in the market. Even with the Pentium-133, Microsoft has recently found that under Direct3D, PCs are capable of performing many 3D tasks without any 3D hardware acceleration. Intel will roll out its long-awaited P55C in the second half of this year. With its MMX instructions *(see* 100301.PDF*)*, larger internal cache, higher clock frequency, and pipeline enhancements, the P55C processor can certainly bring host-based 3D performance to another level.

We expect the P55C will be introduced in 4Q96, followed quickly by Klamath, a low-cost Pentium Pro with MMX, in early 1997. Depending on the type of CPU and the expected shipment date of the PC under consideration, software-only 3D may provide adequate performance, eliminating the cost of a hardware 3D accelerator.

Integration. If the design goal is high integration without relying on add-in cards for multiple functions, the OEM should consider Chromatic's Mpact, Nvidia's NV1, and Trident's T3D9695. The Mpact media processor and Mediaware modules support 2D, 3D, audio, MPEG video, fax/modem, telephony, and videophone. The NV1 media accelerator supports 2D, 3D, audio, and video. Products based on the NV1 will have a migration path to the next level of performance and integration using the NV3. The T3D9695 supports 2D, 3D, video, and a TV interface, but not audio. The T3D9695 was designed for applications such as multimedia notebooks and TV-related consumer products.

Performance. If high performance is desired to differentiate its products, the OEM should consider 3Dfx's Voodoo, 3Dlabs' Permedia, Nvidia's NV3, Oak's OTI64311, and Rendition's Verite. These products are now either in their engineering-sample stages or their final verification stages. Testing or simulation results have been promising. The companies that produce these products have engineering teams that are capable of designing high-performance 3D products. None, however, has been benchmarked in production systems.

Cost. If the target market segment is price-driven, the system maker should consider products from Chromatic, Nvidia, Oak, Trident, and VideoLogic. Although the Mpact media processor needs a frame buffer, a texture buffer, and a Z-buffer, these three buffers are physically in a single unified memory. Because of Mpact's high integration, products based on it can get rid of many discrete components.

Because of their DMA engines and forward texture mapping, Nvidia's NV1 and NV3 eliminate the need for a dedicated texture buffer. Although they do not have fax/modem, telephony, and videophone support, the NV1 and NV3 media accelerators integrate four key multimedia functions: audio, video, 2D, and 3D.

The Oak chip's architecture supports Z-buffering without the cost of a Z-buffer. Both Oak and Trident have a track record for providing low-cost solutions to the PC market. Since the PowerVR architecture gets rid of the Z-buffer completely, it is a good solution for 3D add-in cards. It is not as good for a system solution, however, because of its lack of support for 2D and video functions.

Risk. As mentioned above, PC 3D graphics is an emerging technology. Most products discussed in this article have not yet been fully debugged. Comparatively, 3Dlabs' Permedia, Nvidia's NV1, S3's Virge, and Trident's T3D9695 are lower in risk. The NV1 is in production, and Virge has been proved to work. Permedia and the T3D9695 are third-generation products for 3Dlabs and Trident, respectively, reducing their risk.

Chromatic, 3Dfx, and Rendition are close to shipping their 3D products, but these companies are startups that have not yet produced any volume products. Although Oak and VideoLogic have been in business many years, the 64311 and PowerVR are the companies' first 3D products.

With so many companies working on competing products, it is likely that at least some will deliver attractive 3D accelerators in time for volume PC shipments later this year. Assuming Direct3D meets its objectives, software developers will be able to add exciting 3D capabilities to their games and other applications, spurring consumer demand for 3D performance. A plethora of new products and new companies will emerge as 3D becomes prevalent in PCs by the end of 1997. M

