THE VIDEO COMPUTER: IMAGE COMPUTING IN THE STUDIO

Alvy Ray Smith
Pixar
San Rafael, California

ABSTRACT

A general-purpose *video computer* is proposed which combines many studio or post-production functions, now available only in separate pieces of equipment, and extends their functionality. The ideal machine is described and the current state of the idea is given. The restrictions of realtime and broadcast day are compared. The conclusion is that video computers are already a viable idea within the broadcast-day-turnaround criterion and that the hardware exists as so-called *image computers,* general-purpose digital computers for the class of computations on images. Consequently software houses could immediately begin preparing applications on video computers for the broadcast video market.

published:
SMPTE JOURNAL
97:3 Mar 1988
207-208

INTRODUCTION

A look around a modern broadcast video studio or post-production house reveals a collection of machines made possible by the ongoing digital technology revolution:

- A digital transformation box (zoom, rotate, perspective)
- A digital effects box (page turn, flag wave, etc.)
- A digital wipe machine (perhaps inside a switcher)
- A digital paint program
- A digital font generation box
- A digital standards conforming box
- A digital matting box
- A digital polygon rendering box
- And others

In general, a different piece of hardware is purchased - at considerable expense - for each function. Each of these boxes may be thought of as a *special-purpose digital computer*.

An old notion in computing, now widespread in the general population, is that of the *general-purpose digital computer*. This is, of course, one machine that may be programmed to simulate all special-purpose machines. Since these fully programmable machines are commonly available, and since all the functions listed above have in fact been programmed on them, why are video studios still purchasing expensive special-purpose computers? The answer has several parts:

Computing Bandwidth

The broadcast video market has high computing bandwidth demands. A frequently stated requirement is that a function be computed in realtime, which for NTSC video means a new picture must be computed every thirtieth of a second. Assume a video frame nominally contains 512 by 512 pixels. This is near enough to 640 by 480, 512 by 488, and other popular video raster representations for estimates. So a video frame contains .25 megapixels. Further assume that each pixel is represented by three bytes - one each for RGB. Then a frame consists of .75 megabytes nominally. Furthermore, assume that to compute an interesting byte requires execution of 100-1000 computer instructions. Then the computing bandwidth required for realtime video applications is 75-750 million instructions per thirtieth of a second, or 225-2250 million instructions per second (mips). The only general-purpose digital computers which can approach even the low end of this range are the million-dollar supercomputers. Broadly affordable general-purpose machines are in the .1-10 mips range. So special-purpose machines have been required in video thus far. The video computer described below is a solution between these two extremes powerwise, but at the lower end pricewise.

Programming Talent

A typical studio lacks programmers. This might be changing, but it is probably safe to say that programmers capable of implementing, say, polygon rendering, are not generally to be found in the video studio or post-production house. Special-purpose boxes obviate programmers, of course, their "programs" being hardwired in. Any programmable box would most likely be accepted by this market only if applications software packages were readily available for the desired functions. Advanced houses could add functionality with their own programmers but the typical houses would simply buy functions as they could afford them, and never hire programmers.

Algorithm Availability

Some of the algorithms necessary for implementing the desired functions have not been generally available, so programmers would have nothing to program anyway. These include some of the most sophisticated functions - such as full-blown polygon and patch rendering, with antialiasing, texture, bump, and environment mapping, motion blur, follow focus, depth of field, shadows, reflections, refractions, etc. Again, packages marketed by software houses specializing in video computing programs would be the natural solution.

THE VIDEO COMPUTER - A DEFINITION

Another old idea in computing is that speed can be obtained by trading off against generality. The special-purpose boxes listed above are an extreme application of this idea. The ability to implement a vast number of functions at 1 mips is traded for implementing just one function at 100 mips. The *video computer* is neither fully general-purpose nor fully special-purpose. It is a general-purpose machine for video computations. As such, it is an instance of the larger idea of *image computer* which is a general-purpose computer for image computations, including all video computations as a subset.

How do image computations - hence video computations - differ from ordinary computations? First of all, they *are* ordinary computations in the sense that they can - and have been - implemented (but slowly) on ordinary general-purpose computers. Second, they are distinguished from all other computations in several obvious regards:

- The natural data type is the pixel - not the byte or word
- The pixels are organized in a rectangular array, or raster
- Color is a natural attribute of each pixel

The interesting point is that these seemingly simple specializations make possible - when traded off against full generality (e.g., income tax computation) - a cost-effective image computer, or video computer, capable of realizing all the desired functions very fast.

The Ideal Video Computer

The ideal machine would implement all desired functions in the video market in realtime and would be broadly affordable - comparable in price to just one of today's special-purpose boxes at say $50,000 to $100,000 at introduction. This really is the key statement of the ideal. It could be amplifed to include ease of programming, ease of interfacing, ease of use, etc., but these will follow the essential requirements of speed and cost. The pressures of the marketplace would combine with the continuing decrease in cost per mips to drive the price down from the introduction price, while maintaining functionality.

A Broadened Notion of the Ideal

The ideal video computer - as just described - cannot yet exist. The realtime restriction is the problem in 1987. Closer scrutiny of this "requirement" yields interesting results. Observation of the actual use of so-called realtime machines reveals that it is often the case that their use is slowed by preparation time or multiple passes.

For example, a digital transformation box might be used to rotate a logo while flying it along a curved arc. The box could indeed perform this function in realtime - i.e., as fast as the controlling knob, joystick, trackball, or other device were moved. But the reality of human control is that we are jerky, change our minds, get distracted, etc. and therefore have to practice realtime moves. Or, if the computer itself is used to make the movement - say, under spline control - then several paths will probably be tried before one is selected as desirable.

Another cause of slowdown can be the fact that to get a desired effect, a special-purpose box might be used in several passes and then the results composited in yet another pass or series of passes. For example, suppose a scene is mapped onto the eight surfaces of a double pyramid which spins and then unfolds so that the eight triangles spin off screen along eight separate paths. One possible way to accomplish this is to choreograph eight separate paths, one for each triangle and its matte, then composite them in appropriate juxtaposition and front-to-back ordering using the mattes for correct combination.

None of this is to deny that realtime is important. It is rather to point out that perhaps a more useful time measure in many instances is the *broadcast day*. The idea is that a machine may be sufficiently fast if an effect or other result can be obtained the very day it is conceived so that it can be aired that evening.

Thus, in the multi-pass example above, it would probably be sufficient to have a computer which could compute and composite all eight triangles in each frame rapidly - but not in realtime - and do it all in one pass and in, say, half an hour for a 30-second shot, including recording. Or eighty triangles.

Where Is the Concept Today?

An image computer already exists which meets the broadcast day requirement for those functions already programmed. This is the Pixar Image Computer which computes at 40 mips. It is in fact a video computer since it generates broadcast quality NTSC or PAL compatible RGB as standard outputs. It also falls within the introductory price range mentioned above.

Since some kind of realtime feedback is very useful for interactive design, a useful frontend for today's video computer is a realtime vector machine which can be used for realtime preview of effects using outlines to represent the final forms.

At this writing, it is the software applications programs which are still largely missing, but early examples of software houses already contributing to this marketplace are Alias Research Incorporated, Symbolics Inc., TASC/WSI, and Wavefront Technologies, providing applications for the Pixar Image Computer.

A GLIMPSE AT THE NEAR FUTURE

Consider the following image computations:

• A first-rate rendering of a complex 3-D object - a car, a spacecraft, a tree, a landscape, a face - with multiple light sources, full antialiasing, transparency, full hidden surface removal, bump mapping, texture mapping, and environment mapping - is generated at supercomputer speeds. An entire spline-controlled animation is generated easily within a broadcast day. The object contains 100,000 polygons - or 1,000,000 polygons.

• A weather report shows roiling 3-D clouds casting shadows as they move across a bas-relief representation of the northeastern United States. A hurricane forms off the coast.

• A production assistant wraps a frame onto a disk and spins it offscreen right, revealing the scene behind, or twists a picture about its midline with the twist proceeding from left to right in an accelerating motion to ''throw'' the audience's attention that direction. There are no jaggies anywhere. Both effects are unavailable on special-purpose boxes. All standard rotates, scales, and perspectives are of course available too, as are all standard wipes.

• The full generality of a so-called ''paint program'' is invoked by the staff artist for a quick graphic or an exquisite painting for a background.

• A reactor accident occurs in the Ukraine. A quick call to a national earth resources database gets the studio a satellite photo of the area taken since the accident. The area of interest is magnified, the edges are sharpened, contrast is enhanced - all with standard image processing routines. Text is added and the result goes on the air.

• A difficult blue-screen matte is pulled using techniques not available without digital computations localized to the pixel level.

The important point is that: **One box - a video computer, or image computer - does all of these.** An effects box, paint computer, matting box, wipe box, transformation box, image processor, polygon-rendering computer are folded into one - and still other functions can be imagined and performed.

IMPLICATIONS

A partial listing of the implications explicit and implicit in the video computer concept is:

- One box instead of many
- Repeatable procedures - change one small thing but repeat everything else
- Graphics and effects production brought in-house
- Broadcast-day turnaround
- A natural growth and upgrade path - add functionality as needed
- Opportunities for vendors to provide advanced software applications
- Opportunities for vendors and customers to customize a look

Once general-purpose computation has entered the studio full-force, then all the powerful concepts which have evolved for the office and university become applicable to the studio: Operating systems, database managers, networks, communications standards, common interfaces, archiving, expert systems, etc.

AUTHOR

Dr. Alvy Ray Smith is CoFounder and Vice President of Pixar, a company specializing in digital visualization products and services. He has founded during his career, with Dr. Edwin Catmull, three centers of computer graphics excellence: the Computer Graphics Laboratory at the New York Institute of Technology in the mid 70s, the Computer Division of Lucasfilm Ltd. in the early 80s, and Pixar, a year ago. Alvy has published many technical papers in computer graphics, created or directed some of its best-known pictures and animations, and managed some of the most innovative people in the field, in hardware and software. He received his Ph.D. from Stanford in 1970.