THE PROBLEMS OF COMPUTER-ASSISTED ANIMATION

Edwin Catmull
Computer Graphics Lab
New York Institute of Technology
Old Westbury, New York 11568

ABSTRACT

In the last few years several systems have been written for aiding in the conventional two-dimensional animation process. The main purpose of these systems has been to let the computer produce missing drawings based on extreme drawings produced by animators. While there has been some success and a great deal of optimism, the promise of higher output and quality using a computer has not been realized. The transition from simple drawings optimized for use on the computer to the complicated and detailed drawings of quality conventional animation has been much harder than expected. The principle difficulty is that the animators drawings are really two dimensional projections of three dimensional characters as visualized in the animators head, hence information is lost, ie. one leg obscures another. The problem of making a program infer the original object from its projections is akin to extremely difficult artificial intelligence problems. Efforts to overcome this by drawing skeletons or increasing the number of overlays require more manual intervention thereby offsetting the gains of using the computer. One way to analyze an approach is to determine the average time required of an artist or operator at all stages of animation for every frame. A second problem not generally recognized is that a production animation system requires the management of hundreds of thousands of drawings, hence data base management techniques not normally found in experimental animation systems.

Key words: computer animation, computer graphics, keyframe animation.

CR classification: 8.2

INTRODUCTION

There have been several systems written to aid in the process of animation [see references]. A summary of these has been presented in [2]. Some of these systems have been quite successful and have been used to produce short interesting films, most notably Hunger by Peter Foldes. However, there are some significant problems that remain.

One of the inescapable conclusions one draws after examining the process of conventional animation is that there is a good deal of tedium involved. This observation has led to the idea that the computer can be used to greatly speed up the process and make it cheaper. It is time to examine this idea.

There are several kinds of conventional animation and computer animation. Broadly speaking, in conventional animation there is a continuum from the limited animation of the Saturday morning cartoons to the full animation of the Disney Studios. This continuum is referred to as "character animation." Outside of this continuum are the animation "art" films with a great variety of style, quality, and methods.

With computer animation there are roughly three categories: Art and graphic animation, three-dimensional animation, and computer-assisted character animation, sometimes called "key-frame animation". Art and graphic films are highly varied and will not be part of this analysis. 3-d is one area where the computer offers some exciting possibilities [5,6,10] but also is not part of this paper.

The area to be explored here is the application of computers to character animation. Of course one might legitimately wonder why we are trying to mold the computer into an existing process. While that is not the topic of this paper it should be noted that character animation is a very powerful medium for story telling.

Since character animation is a continuum we should say something about the extremes of that continuum. At the low end, the action is very simple. Movement is more indicated than actually performed. There is a lot of repeat animation, cycles, sound effects to indicate action, etc. Actually, it is hard to imagine how it could be made much cheaper even with a computer. In any case most programmers interested in the problem aren't personally interested in producing low quality animation.

At the other end, the Disney animation is fluid and subtle. It is now so expensive that even the Disney Studios cannot afford to reproduce the quality of Pinocchio or Fantasia. Of course it is not necessary to live up to some past standard, but animators and interested computer professionals share a common goal of producing high quality films. The style may be different but it must be good.

It is useful to characterize animation by the kinds of transformations that can be performed on figures. An important concept is that of "inbetweening" where a person (called an "inbetweener") or a program draws a figure based on two extreme poses of a character. Most computer animation uses linear transformations and simple linear interpolation. The success of these systems has been extrapolated to use for full character animation. The extrapolation is not warranted. It is the thesis of this paper that there is a fundamental limitation in automatic inbetweening and that there is a method for analyzing proposed solutions to the problem. The limitation and the method of analysis are discussed in the latter sections of this paper.

THE CONVENTIONAL ANIMATION PROCESS

Conventional animation techniques have been developed to handle the organization of tens of thousand of drawings. The process is roughly a sequential pipeline with painting of the background going on in parallel. The steps are:

1. story written
2. storyboard laid out
3. sound track recorded
4. detailed layout
5. sound track reading
6. animate extremes
7. assistant animator draws some inbetweens
8. inbetweener draws remaining figures
9. camera department films drawings from paper for pencil test
10. Xerox copy or ink trace onto acetate
11. paint figures on acetate
12. check for errors
13. camera department for final photography
14. edit together

A movie is made up of sequences and a sequence is made up of scenes. A scene corresponds roughly to 5-30 seconds of uninterrupted action. All of the scenes that belong to the same part of the story make up a sequence.

There are two kinds of paper forms used to control the flow of drawings through the pipeline:

1) The route sheet

Every scene is listed with its length, vital statistics, and the name of the person in charge of the various stages. This allows the director to quickly determine the status and location of a scene.

2) The exposure sheet

The exposure sheet has a line on it for every frame in the film. Each line indicates the dialogue for that frame, the order of all figures, the background, and camera position. The exposure sheets are grouped according to scenes.

Some of the procedures used in the process are arbitrary but most are grounded in necessity. For example, an arbitrary procedure is that film is measured in feet rather than seconds or frames. On the other hand, the exposure sheets and routing sheets are needed to control the large amounts of information that must be passed around. The value of this sheet is almost totally ignored in most computer animation research.

THE COMPUTER IN THE ANIMATION PROCESS

There are several steps in the conventional animation process for which we would like to use the computer:

1. Inbetweening
2. input of drawings
3. coloring
4. composition and photographing or videotaping
5. background painting
6. sound track reading
7. pencil test
8. exposure sheets

Each of these steps merits comment:

1) Inbetweening

This is a complex topic and will be discussed in the next section.

2) Input of drawings

Figures can either be drawn directly on the tablet, traced in, or scanned in. Tracing and scanning correspond to inking and Xerox copying in the conventional procedure except that we may not have to trace as many figures if automatic inbetweening can be performed. The quality of tracing is dependent on the accuracy and ease of use of the pen/tablet system. Scanning is dependent on the resolution of the scanner and the algorithms used to extract information. In either case, overhead has been added to the conventional ink or copy.

3) Coloring

An operator indicates what color each area is to receive. The figures are colored by some area filling program. Of course at this stage we must take steps to ensure that we do not see "jaggies" in the final image. Coloring is a process in which the computer offers a distinct advantage over the hand operation. It is not without problems however. The coloring of a scanned-in image requires clean images and some hand touch up.

4) Composition

This is normally done either at an animation camera stand or in an optical house. With the computer we can do this in a frame buffer (a large random-access memory that can store a picture) before sending the picture to the film or video recorder. Full generality almost requires that a red-green-blue frame buffer be used with 8 bits for each color component. This is not obvious to most people, but it is an important capability to have for good picture quality.

The programs must include capabilities for zoom and pan of the components of the picture. This implies resolutions that are greater than that of the display.

5) Background painting

The software system for painting backgrounds using tablet, color monitor, and framebuffer is one of the most successful software packages at NYIT. It was written by Alvy Ray Smith and is in full time use by the professional background artist Paul Xander (See figure 2). The chief difficulties are that pictures consume large amounts of disk space and one has to wait longer than is desirable to get the pictures from the disk.

6) Sound track reading

The sound track is recorded on magnetic film with sprocket holes. Someone must determine the frame numbers of every single significant sound event in the movie which must be matched to action, ie. the location of every consonant in the dialogue. This is time consuming and error prone.

Since the dialogue is known and the track is usually clean, it is not unreasonable to think that an interactive system could be written to speed up the track reading. This is clearly a research topic at this point.

Digital sound equipment could also be used to synthesize sound or to fix errors by expanding or contracting sound on a tape.

7. Pencil test

Animators need to check the action in their scenes. They do this by "flipping" their drawings for immediate viewing and by photographing the pencil drawings to get an accurate check of the timing. The latter takes a long time because the drawings must be photographed and the film developed.

With the computer, the artist can get real-time play back of a scene as soon as the figures are entered or synthesized. A frequently overlooked facet of this process is that the animators still need to hear the corresponding sound track in sync with the pictures.

8. Exposure sheets

One can easily think of the exposure sheet as a data base management system. It is a natural for implementation on a computer and has been virtually ignored.

There have been several attempts at languages to control animation which are very useful for graphic animation but which are not applicable to conventional animation. A good analogy would be music: the conventional method for writing music is to write all the notes down. This method has very little in the way of iteration control. It is basically straight-ahead sequential with everything specified. Similarly, animators want total detailed control. Of course there are cycles and repeats but they can be handled quite nicely with the exposure sheets.

INBETWEENING

The task of producing inbetweens has been the focal point of most computer-assisted character animation systems. The problem is much harder than previously thought.

The problem seems simple enough: given figures A and C, find the correspondence between the figures such that we can produce an interpolated figure B dependent on the correspondence. This problem might be likened to another well know problem: given Russian and English, find the correspondence between them such that we can transform a sentence from one language to the other. On the face of it the problem seems quite simple but we now know that it isn't. Similarly there are subtle problems with inbetweening which require intelligence to resolve.

To illustrate, an animator drew figures 1a and 1c as extremes. The two drawings were given to an assistant animator who produced figure 1b. The drawings would next be given to an inbetweener who would produce three drawings between 1a and 1b and three between 1b and 1c. This example was taken from a real animation situation and is not unusual. A close examination of the drawings will show that automatic inbetweening of full animation is a formidable task.

The principal difficulty is that the animators' drawings are really two-dimensional projections of the three-dimensional characters as visualized by the animator, hence information is lost. For example, one leg obscures another. It is the loss of information which severely limits automatic inbetweening. A person can infer the original object from the drawing because he knows what the original model is, i.e. he understands what a leg is. In order for a program to "understand" a drawing it must contain a model of a character that corresponds to the model in the animator's head. While such a program is not inconceivable, it is akin to difficult artificial intelligence problems.

There are several possible approaches:

1) Try to infer the missing information from the line drawings.
2) Require the animators or program operators to specify the missing information by editing.
3) Break the characters into overlays.
4) Use skeletal drawings.
5) Use 3-d outlines and centerlines.
6) Restrict the class of animation that may be drawn.

Each approach needs comment:

1) Infer information

This has already received comment. It should be further noted that human inbetweeners also make mistakes.

2) Editing

The operator must specify the correspondence of the lines and the hidden lines. One must also ensure that during the process of interpolation, the hidden lines become visible in a correct manner. The process if further complicated by extremes where there is no correspondence at all! (See figure 1.) In this case more extremes are required. Animators trained in the use of the sytems will know where to put extremes but the overall gain is reduced both because more extremes are required and because an editor must spend time on each drawing.

3. Break characters into overlays

This means that the arms, legs, and body may be on separate levels. Someone (probably not the animator) must do the separation. Although this approach takes care of several kinds of problems (ie. the arms and legs in a walk cycle) it doesn't take care of the rotating head.

4) Skeletons

Several people have suggested skeletons as an approach for easing both input and inbetweening [4,8]. The key idea is that a correspondence is established between a fully drawn character and a skeletal outline. It is then only necessary to animate the skeleton. This method handles limited body movement but is not really adequate for changes in feature, expression, the manipulation of nonrigid figures or cloth, or motion involving hidden surfaces and three-dimensional rotation. These capabilities are important for good character animation.

5) 3-d outlines and centerlines

It may be possible to extend the skeleton idea. For a large number of views of a figure a corresponding outline with centerline could be drawn. This would then define a pseudo-three-dimensional character. The animator would draw the outline and centerline following certain rules. The program would determine the appropriate outline from the character definition and hence the character. This is strictly a research topic at this time.

6) Restrict the animation

Just avoid troublesome figures or poses. Of course simple figures are also easier to produce conventionally so we may lose a substantial portion of the gain.

In practice, most systems (with one exception) require that the operator specify the correspondence between the lines of two extremes. If there is not a one-to-one correspondence then a new extreme is required. The one exception is that of Stern [11] whose system automatically determines correspondence from scanned-in images. In all cases, the problems of incorrect matchup and manual intervention remain.

ANALYZING PROPOSED SOLUTIONS

There are many variations on the above methods. What may seem like a good approach frequently requires some preparation or fixup. In order to evaluate any approach we must introduce the concept of "touch-time."

Touch-time is the amount of time that some operator or artist must spend working on a figure in some process. The total touch-time would then be the total of the times spent on all steps of the pipeline. An analysis then requires that we find the average total touch-time per frame. If the computer can perform inbetweening then the average total touch-time goes down. If any method requires some touchup, we must include the touch-up time. The processes included are tracing or figure entry, scanning, coloring or painting, error correction, etc.

Solutions to the various problems of computer-assisted character animation should include analyses of the ramifications of that solution. The full extent of any ramifications cannot be fully understood until the method is tried in the environment for which it is intended. Extrapolation of results will yield incorrect conclusions.

The throughput of the animation pipeline is dependent on touch-time per frame, machine time per frame, the number of stations for artists and operators, and the amount of equipment for processing the images. It is not a simple matter to just state that the computer makes animation cheaper and faster. Cost-effective computer-assisted character animation has yet to be demonstrated.

Another part of any analysis should be a statement regarding the class of characters to be handled. The ease with which any algorithm can be used is dependent on the complexity of characters and the quality of movement. It is necessary to understand any restrictions on characters before one can evaluate the approach used to animate them.

SUMMARY

The production of a good animated film is a major human effort. More advances must be made before the computer will be highly useful in such an effort. Automatic inbetweening has not been a panacea. While the tone of this paper has been to stress the negative side of computer animation, it is felt that significant progress can be made only after coming to a better understanding of the problem.

The Computer Graphics Lab at the New York Institute of Technology was set up to develop computer animation. The lab has developed an animation system (TWEEN), a painting system, and a scan-and-paint system. Figures 3,4 illustrate some kinds of images produced for animation. Current work includes 3-d animation and design.

One intention of this paper is to put computer-assisted character animation is proper perspective based on analysis and experience. While on the face of it the results appear somewhat gloomy, it is only so if one relies just on this kind of animation. Other kinds of computer-assisted animation were not considered for this paper, however in practice we are marrying character animation with other computer graphics techniques. The computer is a "natural" for creating images of three-dimensional objects and for manipulating graphic images. Merging this capability with character animation can enhance its story telling capability.

ACKNOWLEDGEMENT

The Computer Graphics Lab was conceived and sponsored by Dr. Alexander Schure, president of New York Institute of Technology. Lance Williams, Alvy Ray Smith, Garland Stern, David DiFrancesco, and Ephraim Cohen provided critical reading of the first draft. John Gentilella drew the extremes for figure two. The analogy to music was suggested by Alan Kay.

REFERENCES

1. Baecker, Ronald M., Picture-driven animation, SJCC 1969, p273.

2. Burtnyk, N., and Wein, M., Computer Animation, Encyclopedia of Computer Science and Technology, Vol. 5, 1977. Pub. Marcel Dekker, N.Y. pp. 397-436.

3. Burtnyk. N, and Wein, M., Computer generated key frame animation, Journal of the SMPTE, Vol. 80, Mar 1971, pp. 149-153

4. Burtnyk. N, and Wein, M., Interactive skeleton techniques for enhancing motion dynamics in key frame animation, CACM Oct. 1976

5. Catmull, E., A system for computer generated movies, Proceedings ACM Annual Conference, August 1972, pp.422-431.

6. Csuri, Charles, Realtime film animation Annual, report to NSF, 1972-73, Computer Graphics Research Group, Ohio State University.

7. Hackathorn, Ronald J., Anima II: a 3-d color animation system, Siggraph proceedings number 2, 1977, pp. 54-64.

8. Kitching, A., Computer animation - some new ANTICS, Br. Kinematography Sound Television J., 55(12), pp 372-386 December 1973.

9. Levoy, Marc, A color animation system based on the multiplane technique, Siggraph proceedings number 2, 1977, pp. 54-64.

10. Park, F.I., Animation of faces, Proceedings of ACM Annual Conference, Vol. 1, 1972.

11. Stern, Garland, "GAS - A system for computer aided keyframe animation", PhD dissertation, Univ. of Utah, 1978.

12. Tuori, Martin I., Tools and Techniques for computer-aided animation, Masters thesis 1977, Dept. of Comp. Sci., U. of Toronto.

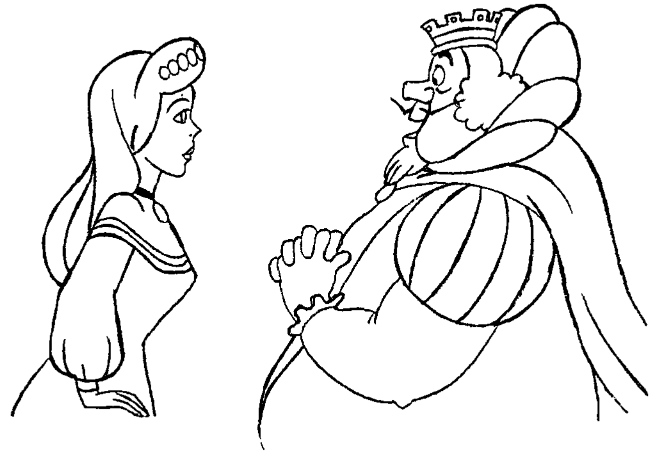

figure 1a

figure 1b

figure 1c

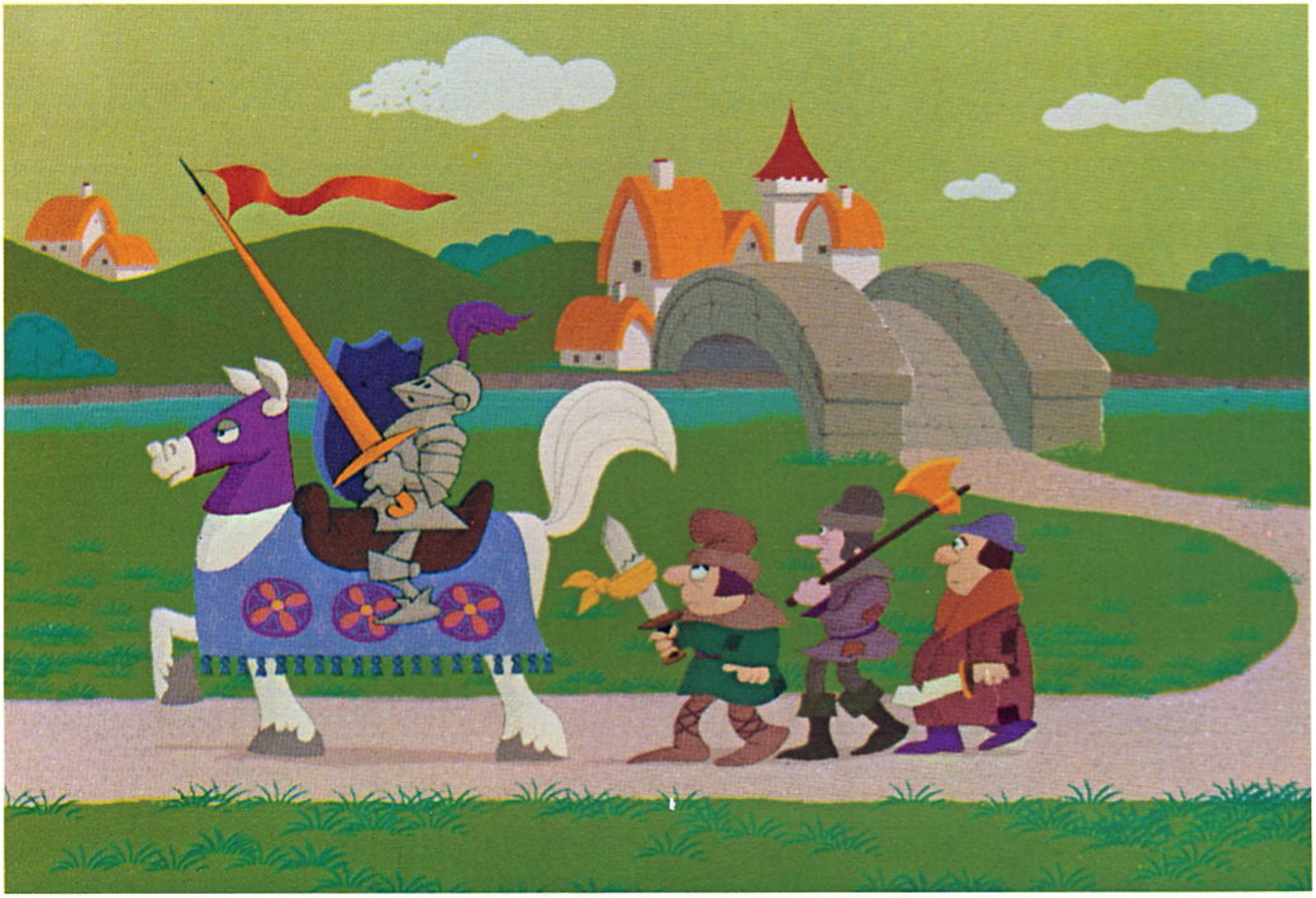

figure 2

figure 3

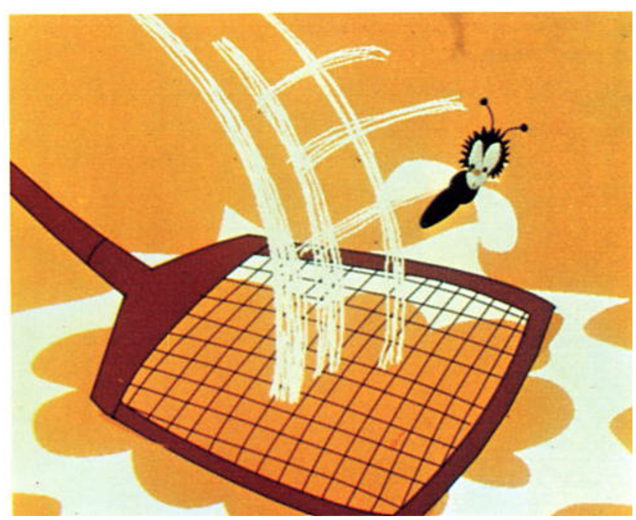

figure 4