SPEECH AND EXPRESSION : A COMPUTER SOLUTION TO FACE ANIMATION

Andrew Pearce, Brian Wyvill, Geoff Wyvill, David Hill

Department of Computer Science, University of Calgary.
2500 University Drive N.W.
Calgary, Alberta, Canada, T2N 1N4

Abstract

Animation which uses three dimensional computer graphics relies heavily on geometric transformations over time for the motions of camera and objects. To make a figure walk or make a liquid bubble requires sophisticated motion control not usually available in commercial animation systems.

This paper describes a way to animate a model of the human face. The animator can build a sequence of facial movements, including speech, by manipulating a small set of keywords. Synchronized speech is possible because the same encoding of the phonetic elements (segments) is used to drive both the animation and a speech generator. Any facial expression, or string of segments, may be given a name and used as a key element. The final animated sequence is then generated automatically by expansion (if necessary) followed by interpolation between the resulting key frames.

We present two alternative modelling techniques for constructing the face: a polygon mesh and a functional description using a technique called *soft objects*.

Résumé

L'animation par ordinateur en 3-dimensions compte fortement sur des transformations géométriques sur temps pour les mouvements de la caméra et des objets. Pour faire marcher une silhouette ou bouillir une liquide, il faut un contrôle de mouvement sophistiqué, qui n'est pas généralement disponible aux systèmes d'animation commerciaux.

Cet article fera la description d'une façon d'animer un modèle d'un visage humain. L'animateur peut construire une séquence de mouvements facials, y compris le discours, en manipulant une petite série de mots clés. Le discours est automatiquement synchronisé, parce que les mêmes éléments phonétiques contrôlent l'animation et la restitution vocale, N'importe quelle expression facialle, ou série de segments, peut être nommée, et utilisée comme élément clé. En dernier lieu, la séquence d'animation finale est créée automatiquement par l'expansion (si nécessaire) suivi par l'interpolation entre les images clées.

Nous présenterons deux techniques alternatives de construire le visage: à polygonale maille, et par une description fonctionelle utilisant des *objets moux*.

Introduction

Three dimensional animation using computer graphics often suffers from a lack of sophisticated motion. Current modelling techniques can produce realistic looking images, but are not suited to representing objects in motion. Nor do we have established ways to describe complex motion to the computer system.

The human face is a prime example of an object which moves in a very complicated way, that cannot be easily and convincingly controlled by simple geometric transformations in time, unless constraints are placed on the possible motion. The major work in this area was done by Fred Parke at the University of Utah [Parke 74] and later developed at New York Institute of Technology [Parke 82]. Parke uses a face built from polygons and identifies groups of polygons which can be changed according to a set of parameters to control facial expressions and features. A second approach [Platt 81] is to use a structure based model where the muscles to be moved are described. While simulating the underlying facial muscles allows for exact representations of wrinkles and face motions, an adequate facial model has not been fully developed using this representation. This is due both to the difficulty in encoding all of the facial muscles and the complexity of its motion due to the number of degrees of freedom allowed the animator.

Although much work has been done on the modelling of the face, synchronized speech animation is still effected through rotoscoping or related techniques.

The Graphicsland Animation System

The *Graphicsland* project group [Wyvill, B. 85a] at the University of Calgary has developed an organised collection of software tools for producing animations from models in three dimensions. The system allows the combination of several different kinds of modelling primitive [Wyvill et al 85b]. Thus polygon based models can be mixed freely with fractals [Mandelbrot 83, Fournier 82] and particles [Reeve 83] in a scene. Motions and camera paths can be described, and animations generated. Note that we do not include the use of a two dimensional "paint" system. Our objective is always to construct views of a full three dimensional model.

Our objective in this work was to introduce better techniques for motion control than commonly available and integrate them into *Graphicsland*.