Data-Driven Photorealistic Faces for Virtual Actors

Frederic Pighin
Tim Hawkins
Jonathan Cohen

February 5, 2002

1 Introduction

One of the ”holy grails” of computer graphics is to create human faces that look and act realistically. However, due to the staggering complexity of the human face, this goal has remained elusive. Traditional attempts at creating realistic faces have either involved a great deal of efforts by talented artists or detailed mathematical simulations. These approaches have yet to produce results that could trick us into thinking we are looking at a real person’s face and not a computer-generated image. A new approach is putting this goal within reach: recording the appearance of a real person to create a digital replica. While a photograph or a video capture the appearance of a subject it does not give the freedom to change the view point or the light falling on the person’s face as we would expect for a digital actor evolving in a synthetic environment. Computer graphics is starting to address this issue by recovering from a set of images the three-dimensional surface of a person’s face, its motion during a performance and its appearance under different lights. In this paper, we offer a glimpse at some of these techniques that will be used tomorrow to bring digital humans to life.

2 Geometry

The first step in acquiring a realistic model of a human face is to capture its three-dimensional shape. This ”3D scanning” step is usually performed by analyzing two images of the face taken from different orientations. If we know two corresponding locations on these images (e.g. the tip of the nose) and we also know the position and the orientation of the cameras in space, we can then trace a ray from each camera through the corresponding points. The intersection of these rays provides the 3D location of the point. If this process, called ”triangulation” is applied to a dense set of correspondances between the two images, the result is a 3D set of points representing the face’s surface. Computing a set of precise correspondances between two images is a trivial task for a human observer but a formidable one for computers who have no knowledge of how faces look like. To solve this problem, a pattern is projected on the face. The pattern is designed so that it creates a map on the face that allows a computer to uniquely identify each point on the images. 3D scanning is a powerful technique that can produce high resolution models accurate to within a fraction of a millimeter. It does have limitations, however. A fundamental issue with scanning real people is that people have a hard time remaining motionless, and scanning can take several seconds. An important engineering goal is to design scanners that are as fast as possible. Another major problem is occlusion, which refers to the fact that the scanner can only see one side of an object at a time. For example, to scan a face we usually scan the left side and the right side. Then we need to combine these two scans into a single scan. There still may be parts of the face that were not seen in either scan, called ”holes,” that need to be filled in.

3 Motion

Because facial expression plays a key role in non-verbal communication, telling a compelling story with digital actors requires precise facial choregraphies. Thanks to a set of techniques called ”motion capture” the motion of a performer can be recorded and transfered to a digital alter-ego, thus simply letting the human performer puppeteer the digital one. Motion capture allows to automatically record the three-dimensional position of a set of facial features (mouth, eyes ...) during a performance. The first step in this process is to record the performance from different angles using a set of cameras, then the facial features are located in each recorded image, finally the 3D locations of the features are computed through triangulation. Often a special makeup such as a set of brightly colored dots is applied to the performer’s face to highlight the features and help their localisation. However this heavy makeup can impact the quality of the performance. A promising alternative is to build a mathematical model that provides knowledge about the appearance of the features. The computer can then compare the model to different areas of the image and locate the features without using intrusive makeup. Using these motion capture techniques very accurate motions of the face can be recorded.

4 Reflectance

Once the geometry and motion of a subject have been captured, producing a final computer rendering of a performance requires some way of deciding how to shade each point on the model. This is typically done by assigning a texture to the face, where the texture captures the changing color of the skin for different areas of the face. Usually, each point of this texture is then brightened or darkened in accordance with the amount of virtual light falling on the model. As with geometry and motion, a realistic texture can be captured by taking photographs of a real person’s face. Usually, only a few photographs are taken–enough to make sure that each point of the face is seen in at least one or two. The texture map for the face is then taken directly from the photographs. This technique can produce fairly realistic results, but it misses some important subtleties of the appearance of skin. For example, skin exhibits spatially varying specular reflection, with highlights on a the forehead, for instance, being sharper than those on the cheeks. Skin also exhibits subsurface scattering, which makes skin seem suffused with light and gives it a ”soft” appearance. These qualities can not be captured by just a few photographs, but can be captured by taking many photographs while moving a light source to different positions relative to the person’s face. This then allows a realistic rendering of the person under whatever virtual lighting condition is desired, with all of the subtle interactions of light and skin accurately reproduced.

5 Conclusion

Recording the three-dimensional appearance of an actor is an important step towards the creation of realistic digital actors. Even more important is a set of tools that let a storyteller manipulate this appearance to tell a specific tale. Armed with the appropriate recording and editing tools, tomorrow’s digital artists will be able to populate digital worlds with truly realistic characters. One of the ”holy grails” of computer graphics is to create human faces that look and act realistically. The domains of application of such photorealistic faces is very wide. The digital effects industry often needs to create virtual replica of an actor for instance to have him appear in a very dangerous situation as a digital stuntman. In human computer interaction, researchers strive to design interfaces that would allow us to interact with computers like we communicate with human beeings. In teleconferencing the ability to render realistic faces is an essential component of model-based coding.

However, this goal has remained elusive for several reasons. First, the mechanisms that underly the face appearance and motion are extremely complicated. The appearance of the face is determined by how light bounces between multiple layer of skins resulting in subsurface scattering. The face is set in motion through the combined actions of ten different muscle groups. Moreover, a lot of the difficulties in creating realistic digital faces come from our well-honed ability to observe and interpret the faces and expressions of the people around us. This ability makes us very capable at noticing the slightest deviation from reality when observing the image of a human face.

Traditional attempts at creating realistic faces have either involved a great deal of efforts by talented artists or detailed mathematical simulations. The artistic approach is inherently limited by the amount of effort it takes to recreate by hand the realistic three-dimensional appearance of the human face. Although generations of talented artists have created very believable portraits and sculptures of human faces, creating a realistic facial model that reflects light and moves in a believable way is a much more challenging task. The mathematical approach has its limitations too. Although, there are realistic