Bringing Photographs to Life with View Morphing

Steven M. Seitz

Department of Computer Sciences
University of Wisconsin, Madison, WI 53706
seitz@cs.wisc.edu, http://www.cs.wisc.edu/computer-vision

Abstract

Photographs and paintings are limited in the amount of information they can convey due to their inherent lack of motion and depth. Using image morphing methods, it is now possible to add 2D motion to photographs by moving and blending image pixels in creative ways. We have taken this concept a step further by adding the ability to convey three-dimensional motions, such as scene rotations and viewpoint changes, by manipulating one or more photographs of a scene. The effect transforms a photograph or painting into an interactive visualization of the underlying object or scene in which the world may be rotated in 3D. Several potential applications of this technology are discussed, in areas such as virtual reality, image databases, and special effects.

1 Introduction

We think of a photograph as a 2D snapshot of a person or a scene frozen in space and time. This notion, however, is now being challenged as a result of new image manipulation tools that can transform still images into compelling animations. These *morphing* techniques are already widely used in television and movies to produce special effects in which one object is fluidly transformed into another.

Transitions, however, are only one of a whole new realm of interesting applications for image morphing. Instead of using morphing to distort an image, for instance, we and others in our field have recently devised ways of using morphing to simulate three-dimensional motions and changes in perspective. This idea, which we call *view morphing*, makes it possible to change the effective camera viewpoint *after* a photograph has been taken. The ability to do this kind of photo postprocessing could simplify the task of the professional photographer by allowing an image's perspective to be fine-tuned in the lab. The same technique can also be applied to drawings and paintings. For instance a police sketch could be morphed to show how a crime suspect would appear from different viewpoints.

Taken to the limit, the ability to morph viewpoint allows one or more photographs to be transformed into an interactive, photorealistic visualization in which a user can virtually inspect a remote 3D object or visit a real scene by moving a virtual camera through the environment. The illusion of a real three-dimensional space is maintained by continually warping the images to correspond to the user's changing perspective. Current systems like Apple's Quicktime VR [1] provide a first step in this direction.

The ability to change viewpoint can be combined with other uses of morphing to create better special effects. To morph two faces using current methods [2, 3], the heads need to be roughly aligned or else severe distortions can arise during the transition. By compensating for changes in viewpoint, however, it is possible to create very effective morphs, even under large changes in object pose. The resulting

The support of the National Science Foundation under Grant Nos. IRI-9220782 and IRI-9530985 is gratefully acknowledged.

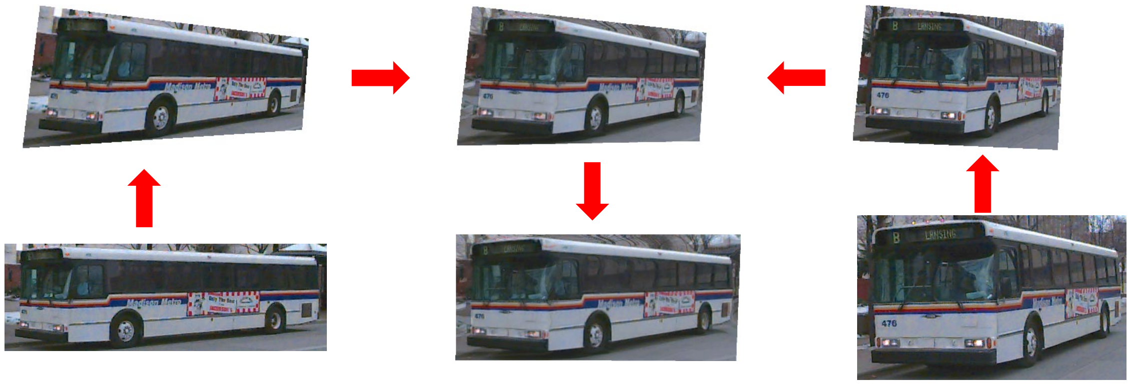

Figure 1: View morphing in three steps. Two original images of a bus are first prewarped to make the cameras parallel and then morphed to create an in-between view. The desired gaze direction of the morphed view is established with a postwarp operation. The arrows show the flow of the algorithm, from original to prewarped, then to morphed, and finally to postwarped.

transitions depict a simultaneous interpolation of shape, color, and pose, yielding morphs that appear strikingly three-dimensional.

Using morphing to simulate viewpoint changes is ambitious because it requires physical correctness—one must take into account laws of optics and the camera projection process for the illusion of a real scene to be convincing. The difficulty is compounded by the fact that the geometry of the underlying scene is generally not known. Nevertheless, we and other researchers in graphics and computer vision [1, 4, 5, 6, 7, 8, 9] have made significant advances in this area, to the point where practical morphing-based scene visualizations are now becoming a reality.

In this paper we describe our work in view morphing and discuss several potential applications of this technology to problems in virtual reality, image databases, and special effects.

2 View Morphing

Image morphing is often described as a way of blending the shape and color of two images to produce a smooth and interesting transition. Current methods work by having a user point out pairs of corresponding features or regions in the two images. From these features a correspondence map is computed, defining how pixels should move during the course of the transition. In the warp phase, in-between images are generated by moving each pixel along a fixed path (most commonly a straight line) between its positions in the two original images. Pixel colors are also interpolated during the course of the transformation.

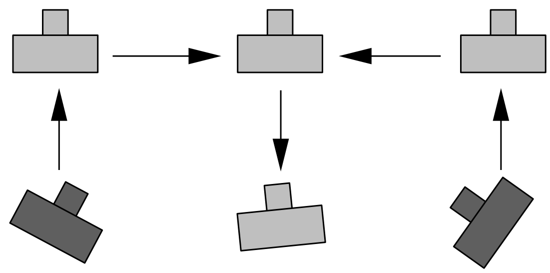

Figure 2: The warp steps shown in Fig. 1 correspond to the above sequence of camera transformations. The original cameras are shown in dark gray and the virtual ones in light gray.

In order to simulate viewpoint changes with morphing methods, it is critical that the warp step preserve the apparent shape of the 3D object or scene whose images are being morphed. In other words, pixels should be moved in such a way that the in-between images represent physically correct perspective views of the scene. Our approach to solving this problem is to decompose the morph into three simpler warps that are each easy to compute. The component warps also correspond to camera operations as follows:

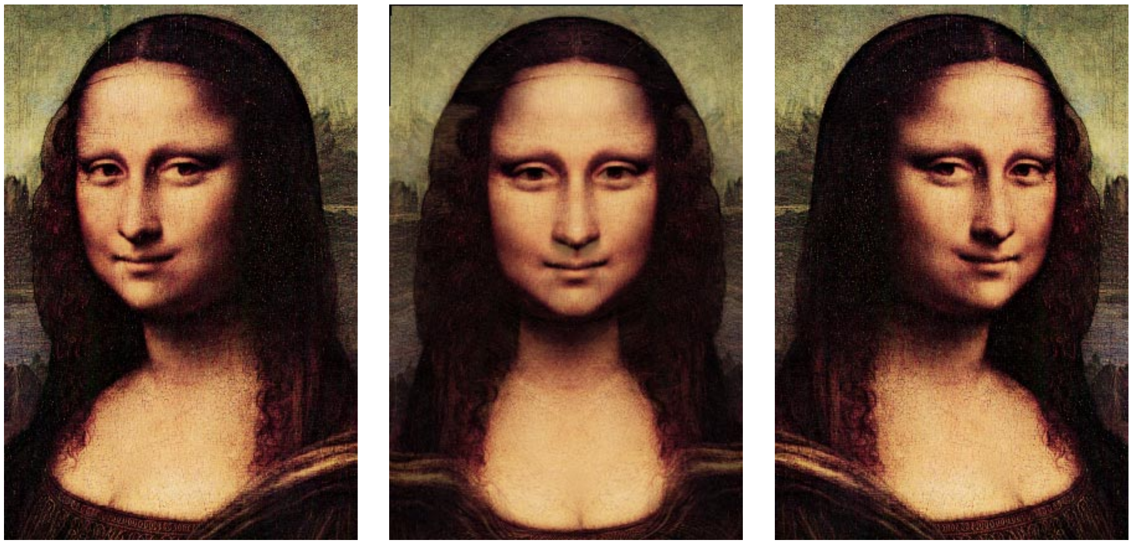

Figure 3: Mona Lisa view morph. Morphed view (center) is halfway between original image (left) and its reflection (right). Although the two sides of the face and torso are not perfectly symmetric, the morph conveys a convincing facial rotation.

- **Prewarp**: warp the two original images so that they correspond to a pair of *parallel views* (see Fig. 2) from cameras that share a common image plane.

- **Morph**: move pixels linearly between their starting and ending positions. This corresponds to moving a virtual camera on the line between the two original cameras.

- **Postwarp**: warp the morphed image to change the gaze direction of the virtual view as desired.

This sequence of operations is shown in Figs. 1 and 2. A nice feature of this three-step method is that it does not require that the two original images be of the same object. For instance, a user can view morph two different faces with the same amount of work that it would take for two images of the same person. A second advantage is that it's not necessary to know the positions of the cameras from which the images were taken. View morphing is therefore powerful enough to be applied to any two images, regardless of source or content. Details on how the warps are computed can be found in [10].

3 Applications

In this section we describe three types of potential applications for view morphing that we are currently investigating. The main idea is to exploit the ability to synthesize changes in viewpoint to solve problems in virtual reality, image databases, and special effects.

3.1 Image-Based Scene Visualization

A nice property of photographs is their realism; producing images of photographic quality and detail with 3D graphics techniques is difficult and very time consuming, in comparison. This feature has inspired an exciting new set of *image-based rendering* tools like Apple's Quicktime VR. These tools allow a user to interactively view a real scene, such as the Eiffel Tower or the inside of a new Saab, by looking in any direction from a fixed vantage point. The visualization provides a strong sense of presence and realism that is achieved by warping and compositing photographs of the real environment. The results are especially promising because a convincing 3D effect is achieved without actually constructing a 3D model of the scene, thereby avoiding a difficult and error

Figure 4: Photo normalization. An image database of differently oriented faces (top) can be normalized using view morphing so that the faces are all horizontally front-facing (bottom).

prone model reconstruction task.

The next generation of image-based rendering tools will take these capabilities a step further by allowing the user to change vantage point in addition to gaze direction, enabling far more immersive visualizations. Using these tools it will be possible to create virtual walk-throughs of a museum, for example, in which a student can navigate freely and continuously through the rooms using a mouse or a 3D input device. Ultimately, photorealistic visualizations of entire cities or regions will be feasible, where a user can virtually walk down a street and examine both the outsides and insides of museums, shopping centers, cathedrals, and so forth. These tools will also incorporate multimedia—walking into a concert hall might invoke an internet audio and video broadcast of the current performance.

Another interesting use for this technology is inspection of historical objects which now exist only through old photographs or drawings. For instance, a present day tourist could virtually visit a historical building, long ago destroyed, by using an image-based rendering tool to process and warp a set of photographs taken at a time when the building was still standing. The same techniques can be applied to paintings and drawings. For instance, applying our view morphing method to an image of the Mona Lisa yields a new set of images showing the Mona Lisa from several alternative viewpoints, as shown in Fig. 3. Viewed interactively, the images show a strikingly three-dimensional Mona Lisa turning her head from right to left. Keep in mind that a 3D model is never actually constructed; the effect is achieved entirely by means of the 2D image warping operations outlined in Section 2.

3.2 Photo Correction

While view morphing usually requires images from two different cameras, objects with certain symmetries can be view morphed from a single image. For instance, the Mona Lisa morph in Fig. 3 was computed by first creating a mirror image of the painting and morphing the original and mirrored views. This process exploits the fact that human faces are approximately bilaterally symmetric.

The ability to artificially change the viewpoint of individual images suggests using view morphing as a means for *photo correction*. A portrait taken slightly off-center could be view morphed into a front-on view in the lab. Similarly, a professional photographer could interactively fine-tune an image's perspective after inspecting the developed photograph. These

Figure 5: View morphing between faces of two different people produces a simultaneous interpolation of facial shape, color, and pose, giving a strong sense of three-dimensionality.

capabilities would add increased flexibility in the image acquisition process and provide the photographer with a powerful new set of image postprocessing tools.

Photo correction also has potential uses in face recognition and image database applications. Future home security systems may use face recognition technology in addition to traditional key-based locks. At the most basic level, current face recognition techniques work by calculating how well an observed face correlates with a database of reference images of known faces. To achieve better accuracy, these systems should compensate for differences in face orientation between the observed and reference images. One way of doing this is to normalize the images, using the procedure in Fig. 3, to convert both observed and reference images to frontal views.

A few images from a hypothetical image database and their normalized versions are shown in Fig. 4. This figure also shows artifacts that can arise as a result of this normalization process. View morphing an image with its mirror reflection produces an image which is horizontally symmetric. Asymmetries in lighting or shape can produce unnatural-looking effects. The rightmost image in Fig. 4 is problematic because of two strong asymmetries—the lighting is one-sided and much of the right side of the head is occluded. As a result, the lighting and some facial contours in the normalized image appear distorted.

3.3 Better Face Morphs

Interpolating different faces to create intriguing special effects is one of the most popular applications of image morphing. Face images are difficult to synthesize convincingly because our eyes are highly attuned to fine details in facial shape and color; even minor errors or artifacts are instantly recognizable. For the same reason, good facial morphs are very striking and convey an especially strong sense of realism.

Producing good morphs between two faces is something of an art. It helps if the faces are similar in some ways (e.g., presence of facial hair or eye glasses) but different enough to create an interesting transformation. One key requirement is similarity in viewpoint; current image morphing techniques can produce severe distortions when two faces are oriented in significantly different directions. For this reason, faces in images to be morphed should be oriented in roughly the same direction. This restriction not only limits the flexibility of current image morphing methods, but makes three-dimensional effects like rotations and translations difficult to achieve.

Using view morphing to compensate for changes in orientation eliminates the need for aligning faces and introduces a whole new range of three-dimensional effects to facial morphs. View morphing faces of two people in different poses produces a simultaneous interpolation of shape, color, *and* orientation, creating an image transition that appears highly three-dimensional. An example of this effect is shown in Fig. 5.

4 Conclusions

Still images can be brought to life by moving and blending pixels in creative ways. Using intelligent morphing techniques that obey optical laws, we have developed ways of adding three-dimensional motions, such as scene rotations and viewpoint changes, to conventional still images. The effect transforms a photograph or painting into an interactive visualization of the underlying object or scene in which the world may be rotated in three dimensions.

One major advantage of this technology is that it produces animations and visualizations that appear *photorealistic*, as a result of using real photographs instead of artificial 3D models. This realism opens up exciting new applications in virtual reality, image databases, face recognition, special effects, and other areas that require images of photographic quality and detail.

A current limitation is that the approach does not cope well with changes in visibility, i.e., features that appear and disappear at different viewpoints. In particular, view morphing a frontal view of a face with a profile view does not produce acceptable results. Extending the approach to accommodate major visibility changes, such as 360 degree rotations, is an area of current research.

We are currently focusing on building image-based rendering systems that can handle many more images and viewpoints. Our goal is to design a system that could create an interactive visualization of an entire building by processing images taken from numerous points throughout the interior and from the outside. The images could be acquired by carrying a hand-held video camera while walking through the environment.

References

[1] S. E. Chen, "Quicktime VR — An image-based approach to virtual environment navigation," in *Proc. SIGGRAPH 95*, pp. 29–38, 1995.

[2] T. Beier and S. Neely, "Feature-based image metamorphosis," in *Proc. SIGGRAPH 92*, pp. 35–42, 1992.

[3] G. Wolberg, *Digital Image Warping*. Los Alamitos, CA: IEEE Computer Society Press, 1990.

[4] S. E. Chen and L. Williams, "View interpolation for image synthesis," in *Proc. SIGGRAPH 93*, pp. 279–288, 1993.

[5] S. Laveau and O. Faugeras, "3-D scene representation as a collection of images," in *Proc. Int. Conf. on Pattern Recognition*, pp. 689–691, 1994.

[6] L. McMillan and G. Bishop, "Plenoptic modeling," in *Proc. SIGGRAPH 95*, pp. 39–46, 1995.

[7] T. Werner, R. D. Hersch, and V. Hlavac, "Rendering real-world objects using view interpolation," in *Proc. Fifth Int. Conf. on Computer Vision*, pp. 957–962, 1995.

[8] R. Szeliski, "Video mosaics for virtual environments," *IEEE Computer Graphics and Applications*, vol. 16, no. 2, pp. 22–30, 1996.

[9] D. Beymer and T. Poggio, "Image representations for visual learning," *Science*, no. 272, pp. 1905–1909, 1996.

[10] S. M. Seitz and C. R. Dyer, "View morphing," in *Proc. SIGGRAPH 96*, pp. 21–30, 1996.