

Image Morphing

Stephen Mullens

Simon Notley

Abstract

Presented here is an investigation into the technique of *morphing*: a method of creating smooth transitions between two images. Morphing is a common technique in special effects for entertainment, but is becoming important in the fields of computer vision and psychology. Mesh warping and crossfading techniques are employed to morph between various still images. Morphing between full-motion film sequences is also considered and a basic example included. The algorithms used are described in detail including an analysis of the number of 'critical points' that must be identified for a successful morph and the possible problems that could be encountered. The mesh warping method, despite being one of the oldest techniques, is shown to be very effective.

1) Introduction	4
2) Morphing Techniques	5
3) Developing a Morphing Program	7
i) The Morphing Process	7
ii) The Interface	9
4) Photographic Morphing	10
5) Video Morphing	12
i) Morphing and Selection of Frames	13
ii) Alternate Video Technique	14
iii) Background Issues	14
iv) The Video Morph	15
6) Conclusions	16
7) References & Sources	16

CD-ROM Contents

The CD-ROM included with this report contains all the videos featured in the report in addition to demonstration videos showing the morphing interface in use.

To access the main menu of the CD-ROM, open *index.htm* using a web browser.

The contents of the CD-ROM are also available online at:

<http://www.ug-home.nottingham.ac.uk/~ppywson/morphing/index.htm>

1) Introduction

Morphing is the name given to an effect used frequently in film and television whereby an object – often a person – is seen to change smoothly into another object. The thing that makes morphing so striking is the seamless nature of the change. In a good morph, the object should appear well defined and physically reasonable throughout, gradually changing its shape to become something else.

Morphing was conceived in the late 1980s when computers were becoming powerful enough to handle images of the quality needed for use in the film industry. The first morphs seen in a commercial film were based on the mesh warping algorithm invented by Doug Smythe, discussed in some detail later.

Although developed for the screen, morphing has found uses outside of special effects. The results of morphing between faces have proved of interest to psychologists. The amount of information the human brain infers from the subtleties of facial features is enormous. The ability to bring about fluid changes and to somehow combine the salient features of two faces has opened up new fields of research in the areas of face recognition and human perception (see Rotshtein 2005).

In recent years, the psychological research based on the first morphing techniques has proved useful to a new generation of computer scientists, interested in developing automated face recognition algorithms (see Sinha 2005). More sophisticated morphing algorithms have also been adopted by computer vision researchers looking at the broader field of object recognition. In these methods, a computer will attempt to morph an observed object into a selection of objects stored in a database. Using a measure of how 'easy' it was to morph from the original to each database object, the computer will then decide which of the database objects it is seeing. Some success has been had with this approach and is documented by Singh (2000).

Although these newer areas are unquestionably exciting and offer much opportunity for future research, we limit ourselves here to study of the basic morphing algorithms developed by the likes of Smythe for the world of cinema. Initially we examine the methodology behind various morphing techniques, with particular focus on applying the mesh warping technique to both still images and video footage.

2) Morphing Techniques

In general, morphing algorithms consist of two processes: warping and combination or crossfading. If one was to think of the simplest possible way to produce a smooth transition from one image to another it is likely that a simple crossfade would be the result. At its most basic, crossfading involves creating a series of frames composed of linear combinations of the two original images. For a transition N frames in length between greyscale images I_i and I_f , the intensity of a pixel at (x,y) in frame n is given by

$$F(x,y,n) = \left(\frac{n-1}{N-1}\right)I_f(x,y) + \left(\frac{N-n}{N-1}\right)I_i(x,y) \quad \text{Eqn. 1}$$

The operation for a colour image is slightly more complicated, but most image processing software has built-in code to handle linear combinations of colour images in a way that appears natural to the eye. However, the results of this process are a long way from the smooth, organic transition described in the previous section. Figure 1 shows the results of applying a simple crossfade to two images of faces. Despite the original images being quite similar, the features are misaligned and the intermediate images just look like two faces superimposed, rather than one face as is the goal. One can imagine that the results would be considerably worse were the original images of radically different shape or size.

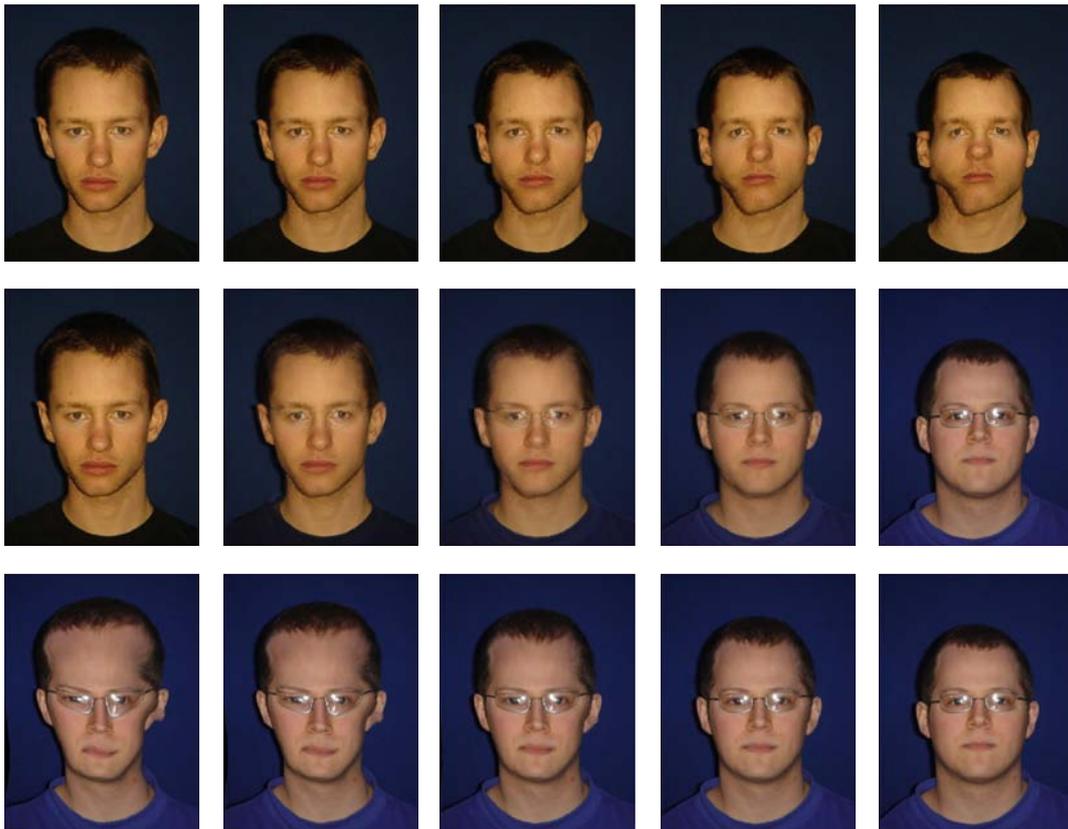
Fig. 1



A simple crossfade between two faces. The intermediate frames here are very indistinct with no clear features, a video of this sequence would simply appear blurred with no sense of motion.

Clearly, the key to making a crossfade look better is to try and match the shapes of the original images as precisely as possible. However, unless we can find a way of manipulating the images, we will be restricted to morphing between very similar objects, which rather defeats the object of morphing. Warping is the name given to the process of distorting the original images so the crossfade technique can be used successfully. There are several approaches to the problem of warping, documented concisely by Wolberg (1998). All the methods have a basic premise in common; they require selection of a set features that are equivalent in the two images (for example, the eyes, nose and mouth on a face). The warping algorithm will then work out a function that maps the initial features onto the final ones, using interpolation of some kind to determine the mapping of every pixel in the image. The main points of difference between the techniques are the way the features are identified and the method of interpolation.

Fig. 2



Morphing. The final morph is a crossfade between the two sequences created by warping one faces features to the position of the others. Note that in a perfect morph, all the features should be aligned in each frame. Whilst this is a very good morph, it is not perfect, the jaw particularly is misshapen.

Figure 2 shows the composition of a morphing sequence. The warping algorithm is applied incrementally to both images, creating a series of intermediate frames for each original image. These are then combined linearly as described above and compiled into a video to create the final morph sequence. The result is a transition where the key features of the original object are seen to move to match those of the final one whilst the colours and textures gradually interchange.

3) Developing a Morphing Program

i) The Morphing Process

In order to create morphs from scratch, we would need to design successful warping and crossfading algorithms. It was decided that the crossfading algorithm should be kept as simple as possible. Whilst it is possible to achieve interesting results by crossfading some areas of the image before others, it was decided to use a simple linear combination of the whole image as described in the previous section.

The warping algorithm used here is based on *mesh warping*, the original morphing technique developed by Smythe. In this technique the features are identified simply by a series of corresponding points. Later techniques allow the user to select corresponding lines or even free-form shapes (Wolberg 1998, pg. 364), the outline of a head for example, but it was felt these presented too many technical challenges to be feasible for this investigation.

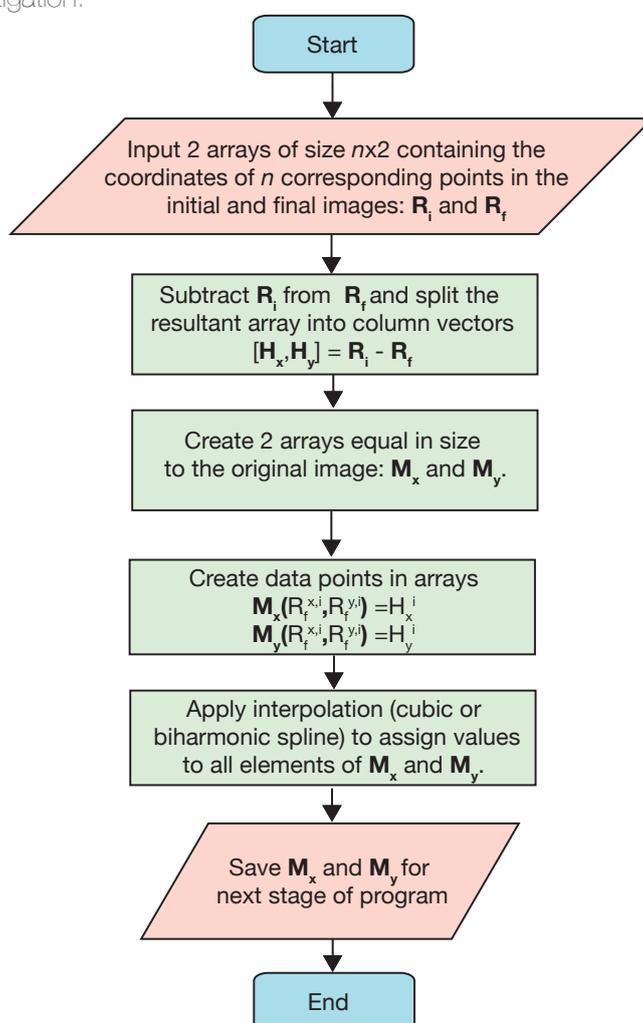
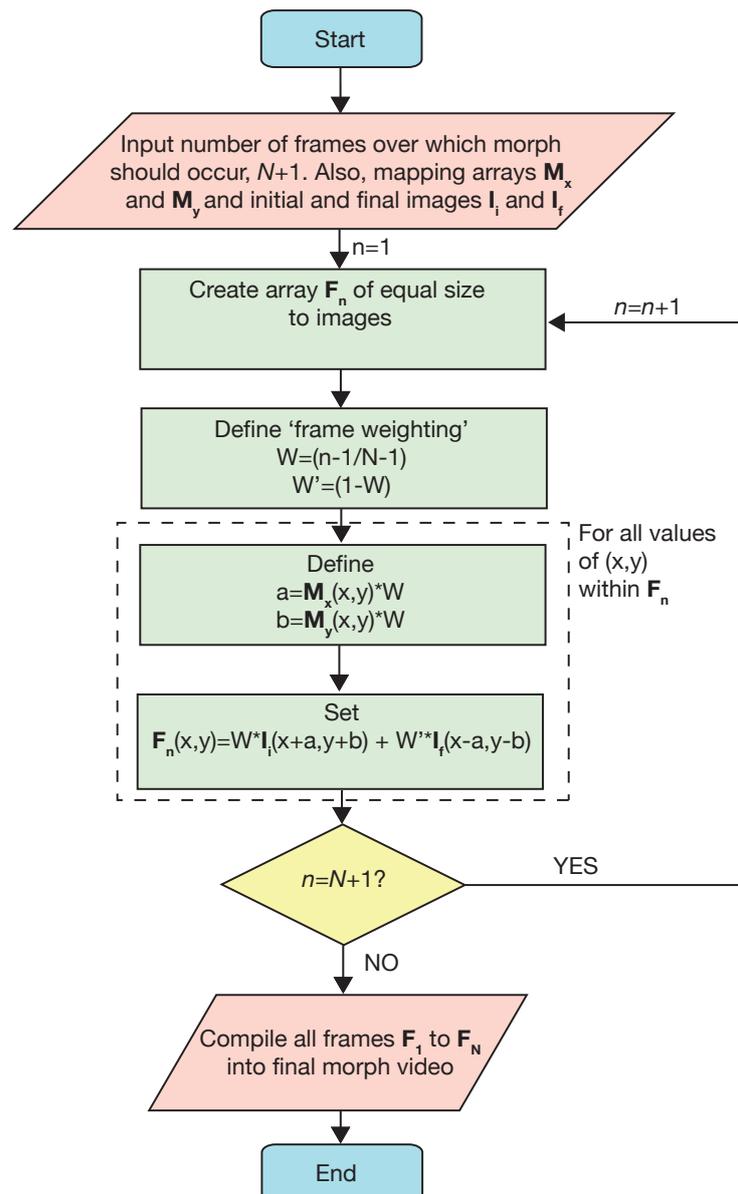


Fig. 3

A flowchart showing the process of generating mapping arrays to describe the movement of every pixel based on the points selected by the user.

One of the major benefits of the mesh warping technique is the simplicity of the interpolation stage. The objective of this stage is to create a mapping detailing the movement of every pixel from the initial to the final image. Since the features are identified as single points, this mapping can be generated by the process of *scattered data interpolation*, a very common computational process with a wealth of effective techniques available. Figure 3 shows the process of creating a mapping.

The interpolation process used for the morphs presented in the report is 'biharmonic spline interpolation' (Sandwell 1987). This was chosen as it was observed to produce slightly better results than bicubic interpolation and is built in to the Matlab programming environment, used for this project.

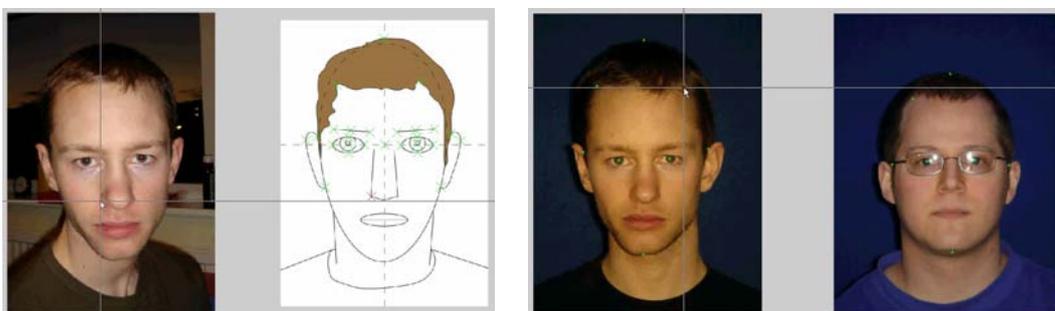


A flowchart illustrating the process of using mapping arrays to generate a whole sequence of frames that can be compiled to form a morph.

Once a mapping has been calculated, the process of generating a morph is quite straightforward. The mapping should simply be applied incrementally to shift the features in the original image to those in the final image and vice versa. These two sequences are combined via a linear crossfade to form the final morph. This whole sequence is described in more detail in Figure 4 for a greyscale image. The program in fact works in colour using the 'imlincmb' function in MatLab to perform the combination.

ii) The Interface

The interface is designed to allow the user to accurately and quickly select all the critical points on an object or pairs of objects. Two methods were used to select critical points, firstly a GUI which pushes the user through selecting a predefined selection of critical points and secondly a GUI which displays two images side by side then allows the user to specify common points within the images. *Fig. 5*



Left - A frame from the GUI which forces the user to select specific critical points

Right - A frame from the GUI which allows feature comparison between a pair of images

Figure 5 shows both of the interfacing techniques used to select critical points, the left GUI is useful for cataloguing a vast selection of images and tells the user where next to place a point whereas the right GUI is typically better for producing flawless morphs between two images. Both GUI techniques work on the same file input and output mechanisms, at start they ask for one or two image files respectively then once completed they output *.mat files. The *.mat file contains a vector of points which were selected and the original image, these files can then be opened by the morphing program.

The output from the morphing program is shown in Figure 6, the central window is constantly updated as new morph frames are produced and saved as tiff files to the hard disk.

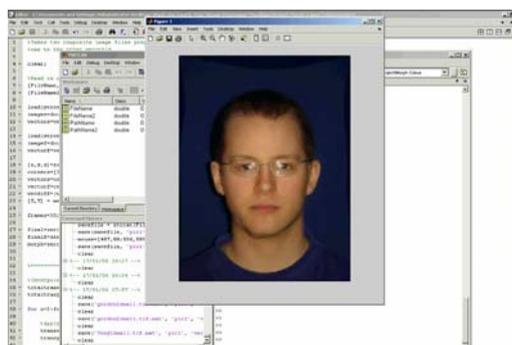
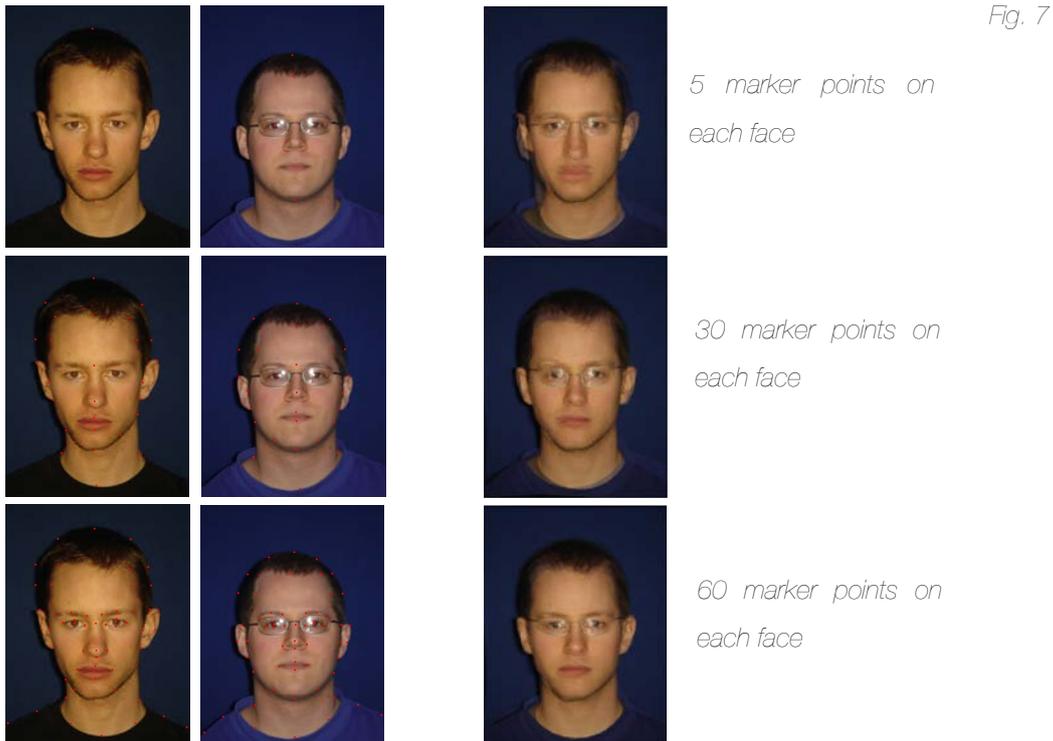


Fig. 6

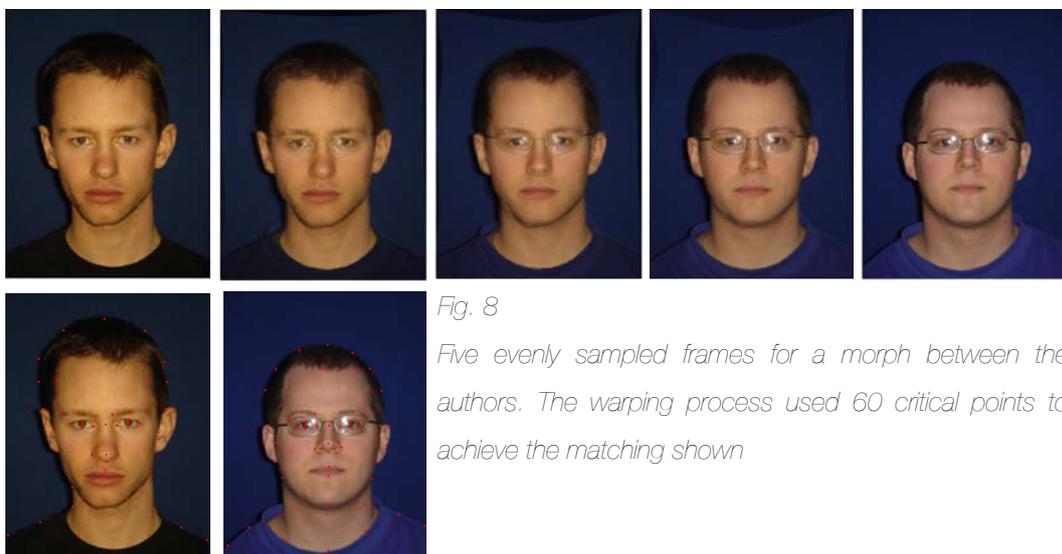
A screen shot of the morphing program as it morphs between the two images on the right hand side of Figure 5

4) Photographic Morphing



The initial and final faces of a morph and the middle frame from the morph which combines both. The number of critical points on the faces is increased for each row resulting in a clearer middle morph image

Figure 7 shows the effect of varying the number of critical points used on each face to define a mapping. It is expected that the warping function will work more effectively with more points and this has been clearly shown to be the case. The five-point image is clearly not physical, most features appear in pairs which do not overlap precisely. It should be noted that where the markers were placed on the image the features do overlap, the warping function away from these points becomes inadequate. As the number of marker points is increased to 30 then 60, the agreement of features slowly improves resulting in a near perfect overlay in the case of 60 points.



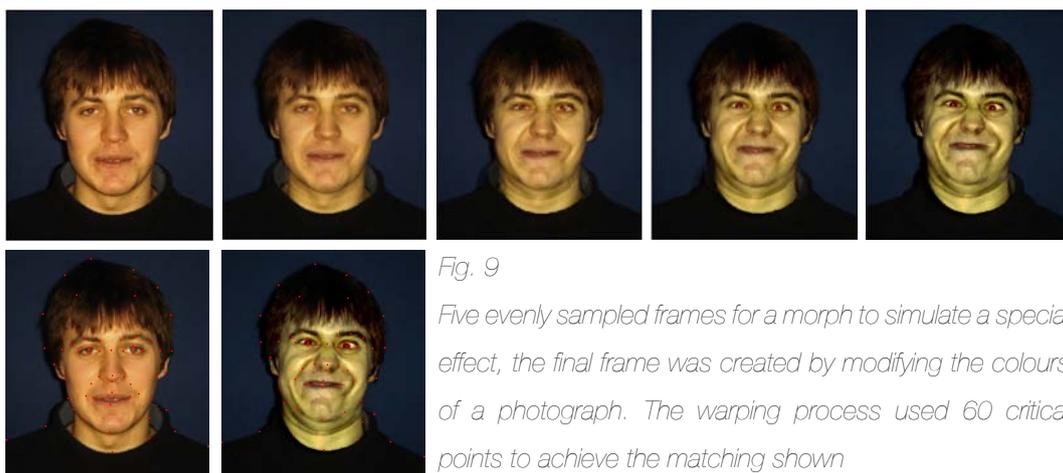


Fig. 9

Five evenly sampled frames for a morph to simulate a special effect, the final frame was created by modifying the colours of a photograph. The warping process used 60 critical points to achieve the matching shown

Figures 8 and 9 are relatively simple morphs with no extreme changes of detail, in both cases 60 critical points were sufficient to allow near perfect transitions between the frames of a video with little apparent double imaging (warping errors). The visible warping errors in both image occurred around the intersection of clothing and the neck, this is due to insufficient critical points being placed in these areas.



Fig. 10

Five evenly sampled frames for a morph between the current prime minister Tony Blair and Gordon Brown. The warping process used 114 critical points to achieve the matching shown

Figure 10 shows a significantly more complicated morph which required 114 critical points to minimise warping errors, some errors still exist around the neck and during the formation of Gordon Brown's teeth. From this it is clear that the number of critical points is a function of how complicated the morph needs to be. When morphing drastically different shaped faces more points will be needed so that the warping function better understands the shapes.

An interesting problem in Figure 10 was the formation of Gordon Brown's hair, it was required that it morph into view instead of just being intensity-blended into the image. The movement of hair was achieved by placing critical points at the ends of Gordon Brown's hair then matching it to points on Tony Blair's hair. Extra points were then needed above the eyebrow to limit the warping effect of the hair on the eyebrow. Despite the level of complexity of this morph, the program produces this morph at approximately 1 frame

per second on an AMD64 3200+, running Windows XP64 with MatLab 7.1, a total run time of less than two minutes for a 100 frame morph.

5) Video Morphing

A video consists of a series of frames which are played in quick succession to produce fluid motion. This section will focus on morphing two moving faces together, as both faces execute a similar movement. The basic process of video morphing is shown in Figure 11.

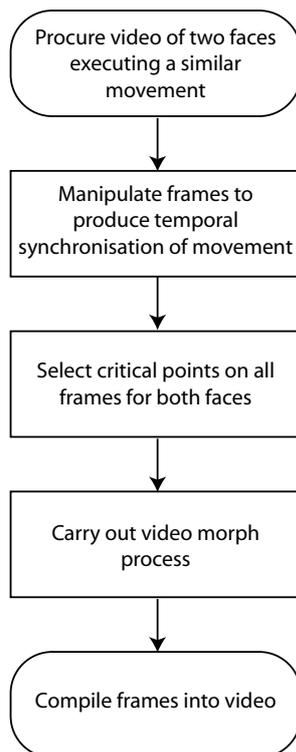


Fig. 11

The fundamental video morphing process

The hardest part of the video morphing process is the acquisition of two video sources whose frames represent the same or similar part of a motion. Although the morphing algorithm is easily capable of blending two out-of-sync frames, when these frames were combined into video, the result would not be a smooth motion. A technique to produce two roughly synchronised videos is to record two people carrying out a similar movement at approximately the same speed, frames can then be sampled from both videos at a specific rate as to ensure both videos will have an identical number of frames.

To select critical points the user has to look at every frame and identify the points of interest. Poor selection of points at this stage will result in misalignment of the two morphed images and poorly defined edges. It should be noted that critical points only need to be matched between opposing frames in the video sources and not between sequential frames; extra critical points can easily be added to specific pairs of frames if the subjects motion has induced erratic movement of specific facial features.

As part of the process for selecting critical points there has to be a compromise between the number of points identified and the time available. Ideally the user could identify upwards of 100 points on each frame to match up features, but this would take a significant amount of time when dealing with tens of frames. For the purpose of proof of principle a modest 50 points were included on a total of 62 frames. A selection of easily identifiable points were chosen to be used for all frames in the video morph, this distribution is seen in Figure 12.



Fig. 12

Typical placement of points used for every frame in a video morph, note only major features covered

i) Morphing and Selection of Frames

Figure 13 shows a diagrammatical representation of a video morphing process, essentially the process is just a series of conventional image to image morphs which are then filtered to produce a video.

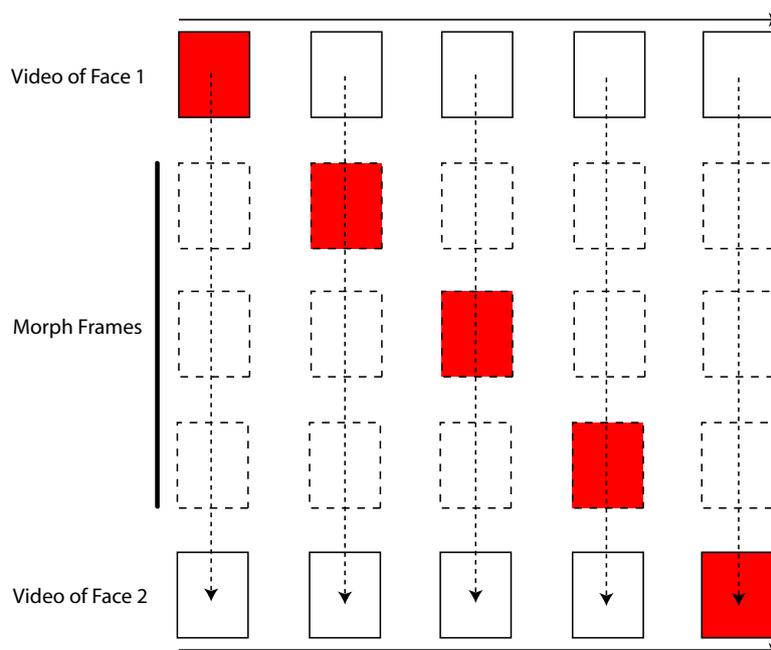


Fig. 13

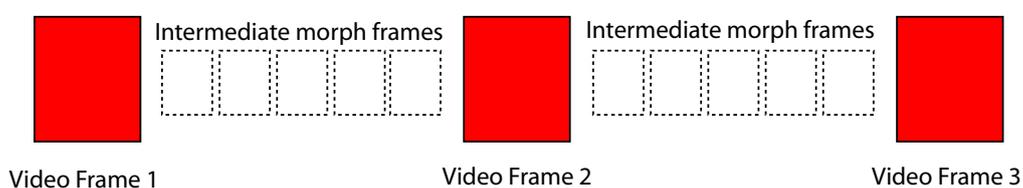
Illustration of the video morphing process. The vertical arrows represent morphs from each starting frame to its corresponding final frame. Selection of the filled frames will produce a video of the same length of the original video but slowly executing a morph from one face to the other.

ii) Alternate Video Technique

A faster but less physical video morphing technique is to create a video morph with few frames covering the motion, then using the normal morphing process to interpolate frames between the video frames. For the case of a mouth opening and closing there are five critical frames, e.g. closed, half open, open, half open, closed. If a video morph was made using the technique shown in Figure 13 then the output would be a five frame jumpy video.

Now critical points should be selected on the five frames making sure that all points in the frames correspond to each other. A conventional morph can now be used between each adjacent pair of numbers to fill gaps, this process is shown in Figure 14 for the case of a three video frames.

Fig. 14



The required steps to smooth out a jumpy 3 frame video morph

iii) Background Issues

Background detail in the videos can produce unpleasant distractions from the morph itself. Figure 15 shows how backgrounds with inherent symmetry and strong lines can rapidly descend into non-physical scenes. To prevent these background distortions, extra critical points could be added to the brickwork in all frames, however this would be wasting time on an insignificant feature. The background just needs to be suppressed or the videos taken with a constant flat background, i.e. a blue screen.

Fig. 15



Frames sampled from a basic video morph. Note that although the background was roughly in the same place for both videos, the mapping functions used for the morph cause massive distortions.

iv) The Video Morph

Fig. 16



Seven frames evenly spaced throughout the video morph which shows the authors morphing between one another.

Figure 16 shows frames from a 31 frame video morph, the morph was chosen to vary linearly with frame number so that at the middle frame there would be 50% of each person. In principle the morph has worked and produced physically reasonable faces on the connecting frames, however edges on the outer features of the head are often blurred and are seen as a pair of double images, essentially the morph has not accurately warped one face into the other. Poor blending of faces is the result of not enough marker points being used on the original images, more points would have forced things like the t-shirt to hold position between the frames and not appear as two shapes slowly blending past one another. There are some minor double images on the right ear and left side of the neck, as with the t-shirt these could be remedied by more marker points in those areas.

6) Conclusion

We have seen that a mesh warping based morphing algorithm is easily capable of producing good quality morphs for still images. We have also shown that it has considerable potential for use in full-motion video. Indeed the limiting factor in this case seems to be the patience of the user and the quality of the original footage.

The program developed has clear potential for extension, particularly in the area of the point-selection interface. Having to pick every pair of points individually rapidly becomes laborious and the mapping accuracy is improved with every point. The first stage of developing an improved point-selection method would be to allow the user to choose lines or shapes rather than points. The program could then generate points within these shapes and the same algorithms could be applied.

Despite these minor limitations the results of this short investigation are very promising. The results produced all demonstrate the smooth organic transition between images required, the quality of the results is particularly encouraging given the short runtime of the program. Overall the development of the program can be regarded as a success and forms a solid basis for further development.

7) References & Sources

Rossion, Is sex categorization from faces really parallel to face recognition?, *Visual Cognition*, 2002, 9 (8), 1003–1020

Rotshtein, Treves, Henson, Driver, Dolan, Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain, *Nature Neuroscience* 8, 107-113 (01 Jan 2005)

Sandwell, Biharmonic Spline Interpolation of GEOS-3 and SEASAT Altimeter Data, *Geophysical Research Letters*, 2, 139-142, 1987.

Singh, Papanikolopoulos, *Pattern Recognition* 33 (2000) 1683-1699

Wolberg, Image morphing: a survey, *The Visual Computer* (1998) 14:360-372

All figures used in this report were created by the authors, the morphing program was constructed in MatLab 7.1 (Unix and PC versions) and all further image processing was carried out using Adobe Photoshop CS2 for Macintosh.