Generation of Self-Referential Animated Photomosaics

Daryl D'Souza
School of CS&IT
RMIT University
Melbourne, Australia
daryl.dsouza@rmit.edu.au

Vic Ciesielski
School of CS&IT
RMIT University
Melbourne, Australia
vc@cs.rmit.edu.au

Marsha Berry
School of Creative Media
RMIT University
Melbourne, Australia
marsha.berry@rmit.edu.au

Karen Trist
School of Creative Media
RMIT University
Melbourne, Australia
karen.trist@rmit.edu.au

ABSTRACT

We describe the implementation of an art installation that generates animated photomosaics of the viewer. Photomosaics are target images composed of smaller images known as tiles. When a photomosaic is viewed from afar the detail of the tiles is lost and the target image is evident. Up close, the opposite occurs: the detail of the tiles is evident and the target image is lost. Our system uses a photo of the viewer as the target and miniatures of the viewer's face as tiles. Evolutionary search is used to find the best selection and arrangement of tiles. Each newly found best image is then used as the frame of a movie. The resulting animations start from a random arrangement of tiles and gradually the viewer's face emerges and is clearly visible, and then gradually de-materialises into a random pattern.

Categories and Subject Descriptors

J.5 [**Arts and Humanities**]: Fine Arts; H.5.1 [**Multimedia Information Systems**]: Animations

General Terms

Algorithms, Design, Human Factors

Keywords

Genetic Algorithm, Evolutionary Search

1. INTRODUCTION

Photomosaics are digitised refinements of traditional mosaics, which are composed of many pieces of glass, stone or ceramic tiles of similar shapes and sizes to make up larger images. An example of a photomosaic appears in Figure 1e. Viewed up close the digitised tiles or smaller images are clearly evident as in Figure 2, which is an enlarged segment of Figure 1e. From afar the tiles lose their significance and a new, larger image becomes visible as in Figure 1e.

MM'07, September 23–28, 2007, Augsburg, Bavaria, Germany.

This paper explores the animation of photomosaics. Animations are created by using some of the intermediate images created during the search for the best tiling as frames of a movie. For the search algorithm we have used, the process generates animations in which the subject of the image gradually emerges from a random image. Some still images from an animation are shown in Figures 1a-e. We have found that many viewers find such animations engaging and we think that they will be even more engaging if the viewer is the subject of the animation. (Other sample photomosaics appear on our website[1].)

Self portraits and portraits have fascinated artists for centuries. Recently the self portrait is becoming an ever more popular form of self expression. Karl Baden's recent work has explored this phenomenon. Feeney, a journalist with the Boston Globe, in his review of Karl Baden's photographic exhibition contemplates the recent phenomenon of serial self-portraitists. He claims that the rise of digital technology such as digital cameras and social software such as Flickr and YouTube have encouraged the popularity of self-portraits and that "Like blogs, the sites function as a record of one's existence, except they are visual rather than verbal and concerned with physical appearance rather than personal opinion [1]." Baden's exhibition is at Howard Yezerski Gallery, titled "Every Day: 2/23/87-2/23/07, Twenty Years – Ten Bucks," May 23 - 29, 2007. The exhibition comprises around 7,000 Baden self-portraits that Feeney describes as "a mosaic of me's". A search of `www.youtube.com` with the key words "Photo every day" gave over 20 videos which had been composed in some way from daily pictures taken over a number of months or years.

Our goal is to create a system that will function as an installation where users complete the art work. Users will enter a "booth", position themselves as directed and click the mouse. The system will take their photos and do all of the necessary processing to deliver a CD containing, as an mpeg file, an animated photomosaic in which the tiles are miniatures of the original photo. The animation will start

[1] `www.cs.rmit.edu.au/~vc/evolved-art/gallery`

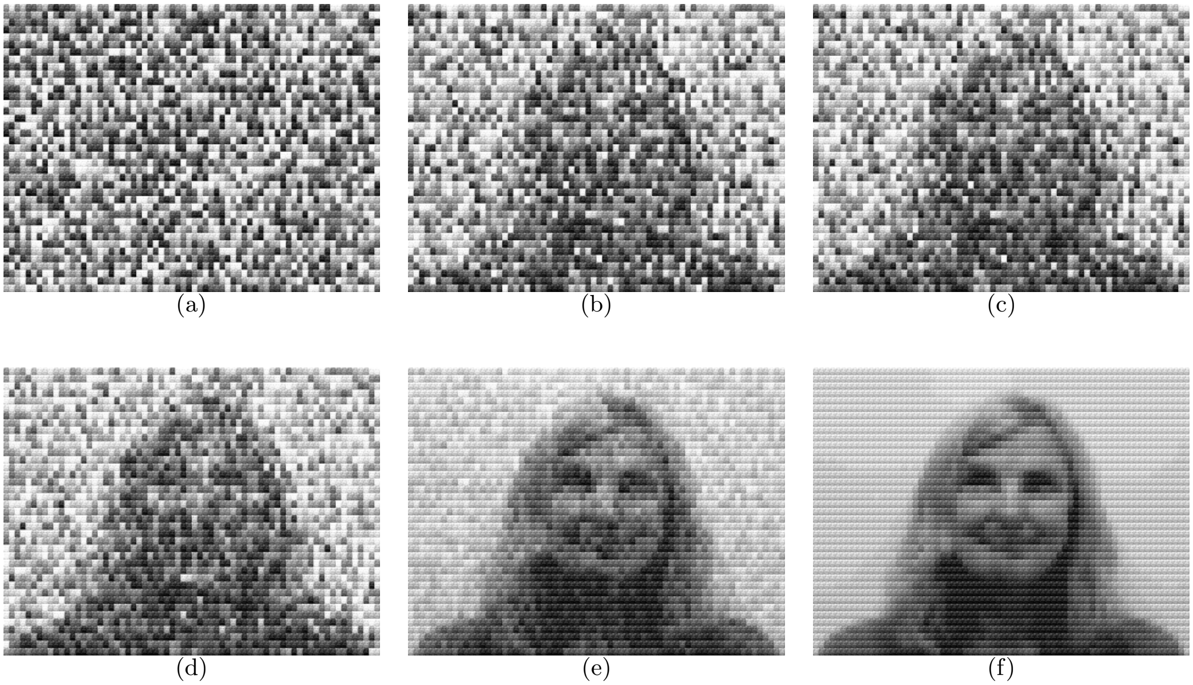

Figure 1: Examples of Evolved Images

Figure 2: Enlargement to natural size of a segment of Figure 1e

with a random collection of tiles, as in Figure 1a, gradually "materialise" the user as in Figures 1b-e and then run in reverse to "de-materialise" the user back to random.

2. RELATED WORK

Historically, the term *photomosaic* is attributed to the stitching together of juxtaposed pictures into a compound image. The idea of combining digital images into larger building blocks originated in a computer graphics system called *DominiPix* [5], in which pictures are constructed from sets of dominoes. Silvers [9] later conceived of the idea of dividing a target image into smaller regions and generating the best approximation to the target image, by comparing the regions to the target. Silvers' [8] company produces photomosaics on a commercial basis.

A quantitative method for generating photomosaics is proposed by Tran [13], in which two distance measures are evaluated according to *similarity* between the photomosaic and the target image, *granularity* of details in each individual tile, and *variety* of the selected tiles. Costs are measured on the basis of *running time* of the generation and *library size*, the number of tiles available for selection. Kim and Pellacini [3] use a mosaicing technique known as Jigsaw Image Mosaic (JIM), where tiles of arbitrary shapes are used to compose the final image. They use a framework in which they seek to minimise a "mosaicing energy function", to select the optimal tile configuration. The energy function penalises configurations that do not maintain the colour of the target image. Their work was inspired by the work of the Renaissance Italian painter, Guiseppe Arcimboldo [12], who invented a style of painting that used clusters of images of fruit, vegetables and other materials, to paint human faces. In another investigation, Li et al [6] use content-based image retrieval (CBIR) to support the generation of a series of destination artistic mosaics, from an arbitrary, original image. The CBIR image database contains 50,000 images sourced from the COREL database.

Animations using mosaic tiles produce interesting effects where the final picture is revealed through a process of gradual disclosure. These offer much to the artist and are already used by scientific data imaging, for example NASA produced an animation visualising the Amazon rainforest using mosaic tiles representing data collected from satellites [7]. This technique is used to render images and image sequences that are meaningful to the viewer. Research on animated mosaics that aim to gradually evolve mosaic frames towards the best approximation of the target image, focus on generating the best fit frame at each iteration. For example, Klein et al. [4] develop a distance measure to assess based on average colour and three-dimensional wavelet decomposition signatures in the colour space. Their matching process also dynamically selects smaller tiling sequences from a collection of candidate source videos, and supports the use of different tiling shapes. Finally, in other work on mosaic animation, Smith et al. [10]

explore among other things the generation of temporally coherent mosaics, that is, wherein the individual frames are chosen to present a coherent sequence in time.

3. THE SEARCH ALGORITHM

Our approach to generating a photomosaic requires a target image and a fixed set of tile images. The algorithm views the artist's canvas as a two-dimensional grid that is of the same size as the target image. A search, using a permutation-based genetic algorithm [2, 11], is used to select and arrange tiles on the grid so that the final arrangement is most similar to the target image. Such a heuristic search algorithm overcomes the inefficiencies of an exhaustive search.

A genetic algorithm operates on a population of potential solutions and uses operators based on biological selection, crossover and mutation to generate child potential solutions from parents. For each individual in a population there must be a way to evaluate how close it is to the desired or optimal solution. This is a measure of the $fitness$ of the individual and is computed by a $fitnessfunction$. The search is conducted as follows:

1. Generate a population of individuals at random.
2. Use the fitness function to determine the fitness of each individual in the population.
3. Compute the next population (generation) by
   (a) Select parents at random, but favouring fitter individuals
   (b) Apply a crossover operator to get 2 children
   (c) Randomly apply a mutation operator to the children
   (d) Place the children in the new generation
4. If no improvement for a number of generations stop
5. Go to 2

Since the selection process chooses fitter individuals, over a number of generations the individuals get better and better. After some number of generations it is not possible to find a better individual than the current best and the evolution can be stopped.

From an application point of view the evolutionary process can be considered as a generic 'engine' to which the user must provide two inputs: (1) The way that an individual, i.e. a potential solution, is represented, and (2) the fitness function. For the photomosaic problem, these are described in the next two sections.

3.1 Representation

A potential solution is an arrangement of tiles in a two dimensional grid that is the same size as the target image. If the target is of size $N \times M$ and the tiles are of size $n \times m$, we require that n divides N and m divides M. $N/n = C$ tiles are required to fill a row of a potential solution and $M/m = R$ tiles to fill a column. A potential solution is thus an $R \times C$ array of integers in which each integer is an index into the tile set provided. If t tiles are provided then each integer is between 1 and t. A chromosome is this matrix represented as a single vector in row major order.

Table 1: Genetic Algorithm Configuration

Parameter	Value
Population Size	120
Crossover Rate	0.7
Mutation Rate	0.0001
Elitism Rate	0.1
Max Generations	50,000
Selection	Proportional to fitness
Termination	50,000 generations
Replacement	Generational replacement

Crossover can only occur at points that are integer boundaries. We have used 1 point crossover[11], with two children being created from two parents. The crossover operator is applied with probability *Crossover Rate*. If crossover does not occur, the parents are simply copied to the children. Mutation is implemented by picking a random position in the array and replacing it with random integer between 1 and t. *Mutation Rate* is interpreted as the probability of changing a particular tile position.

3.2 Fitness Evaluation

For grey level images, fitness is evaluated by summing the differences in pixel grey levels between the candidate solution and the target (equation 1). For colour images the sum is taken over the red, green and blue components.

$$\sum_{i=1}^{N}\sum_{j=1}^{M}|target(i,j) - individual(i,j)| \quad (1)$$

This fitness algorithm allows us to obtain a mosaic that is as similar as possible to the target image.

3.3 Construction of photomosaic animations

Each time a new best image is found in an evolutionary run it is rendered and written as an image file. These images form the individual frames of a movie. To achieve the 'materialise' and 'de-materialise' effect described earlier we have copied the frames generated by the search algorithm, reversed the order, and constructed a movie file from the original and reversed frames. The animations are quite compelling when played in a loop.

Unfortunately there is a large variability in the available tools for constructing movie files from individual frames and for playing the movies. We found that the same animation file was rendered differently by different (media) players. Very pleasing artistic effects achieved with one player were lost with another. Renderings by some players were far less successful, indeed unacceptable. The speed of the computer processor also affected the rendering. We found that constructing MPEG files with PPMTOMPEG and viewing them with the VLC player on a PC, the MPLAYER player on a Solaris environment and the totem player on a linux environment generally worked well.

4. THE INSTALLATION

The installation consists of an iMac computer with a built in camera. The user stands in a designated position and clicks

Table 2: Configuration of the Installation

Parameter	Value
Target Image Size	525×700
Tile Size	15×20
Intensity difference between tiles	10
Write only every kth best	5
Average time to construct animation	6 mins
Number of frames per animation	200
Run time of an animation	1 min

the mouse to start the process. The subsequent processing is as follows:

1. A picture of the user is taken and used as the target.
2. The facial area of the target is cut out and saved as as the primary tile.
3. A set of tiles, at varying levels of light and dark, and scaled to the required size, is then generated.
4. The algorithm described in section 3 is used to get the sequence of best images.
5. The method described in section 3.3 is used to construct an mpeg file from the individual frames.
6. The animation is written to a CD or emailed to the user.

There are a number of decisions to be made in configuring the installation. These mostly relate to quality of the video and run time. The final video needs to fit on a computer screen without the need to scroll. The tiles need to be big enough to be easily discernible on the screen, but small enough that the evolved mosaic can be clearly recognised. Using every best image in the movie can give animations that are too slow to maintain viewer interest. The evolutionary search can be quite slow because of the frequent need to write a large image file, and the user can get impatient while waiting for the evolutionary search to finish. These two problems can be ameliorated by only writing out every kth best image. We have found that the values shown in the top half of Table 2 to be a reasonable compromise. Using these values gives animations with the characteristics shown in the bottom half of Table 2.

5. CONCLUSIONS

Our goal in this work was to develop an installation that would generate interesting and engaging animated photomosaics with an image of the viewer as the target and miniatures of the viewer's face as tiles. By using an evolutionary search to find the best fitting arrangement of tiles, and using each (nth) new best individual as a frame of an animation, we have created successful animations. The major reason for getting interesting animations is the genetic search. Using a classical exhaustive search to find the best matching tile in each position in the target would result in very boring animations, even though the final mosaic might be slightly better.

Further work is planned, at least in areas involving search and colour. Currently, the search takes about six minutes and some viewers become impatient with the wait. Furthermore, our animations are grey level; in future we would like to extend them to colour.

6. ACKNOWLEDGMENTS

Irwan Hendra and Austin Wood, at various times, worked on the mosaic programs and associated scripts. Michael Cassidy implemented the integrated tool to generate animated photomosaics, from these programs and scripts.

7. REFERENCES

[1] M. Feeney. *For 20 years, photographer had a daily obsession: Himself.* The Boston Globe, http://www.boston.com/ae/theater_arts/articles/2007/05/22/for_20_years_photographer_had_a_daily_obsession_himself/, Visited 27/05/2007.

[2] R. I. Haupt and S. E. Haupt. *Practical Genetic Algorithms.* John Wiley and Sons, New York, 1998.

[3] J. Kim and F. Pellacini. Jigsaw image mosaics. In *Proceedings of the ACM-SIGGRAPH International Conference on Research and Development in Graphics and Interactive Techniques*, pages 657–663, San Antonio, Texas, 21–26 July 2002.

[4] A. Klein, T. Grant, A. Finkelstein, and M. Cohen. Video mosaics. In *Proceedings of the 2nd International symposium on Non-photorealistic animation and rendering*, pages 21–28, Annency, France, 21–26 July 2002.

[5] K. Knowlton. DominoPix, US Patent No. 4,398,890, Representation of Designs. http://www.knowltonmosaics.com/, http://www.metron.com/DominoPix/, 1983. Visited 02-Nov-06.

[6] X. Li and Y. Yuan. Artistic mosaic generation. In *Proceedings of the 3rd International Conference on Image and Graphics*, pages 528–531, Hong Kong, 18–20 December 2004.

[7] NASA. Mapping the Amazon: Mosaic tiles animation. http://svs.gsfc.nasa.gov/vis/a000000/a002400/a002405/, Visited 06-Oct-06.

[8] Runaway Technology. Company web site. http://www.photomosaic.com/rt /highlights.htm, Visited 06-Oct-06.

[9] R. Silvers and M. Hawley. *Photomosaics.* Henry Holt and Co., Inc., New York, NY, USA, 1997.

[10] K. Smith, Y. Liu, and A. Klein. Animosaics. In *Proceedings of the ACM-SIGGRAPH International Conference on Research and Development in Graphics and Interactive Techniques*, pages 201–208, Los Angeles, CA, 29–31 July 2005.

[11] R. E. Smith, D. E. Goldberg, and J. A. Earickson. SGA-C: A C-language implementation of a simple genetic algorithm, Sept. 05 1991. http://citeseer.ist.psu.edu/341381.html.

[12] C. Strand. *Hello, Fruit Face!: The Paintings of Giuseppe Arcimboldo.* Prestel, 1999.

[13] N. Tran. Generating photomosaics: an empirical study. In B. Bryant, G. Lamont, H. Haddad, and J. Carroll, editors, *Proceedings of the 1999 Symposium on Applied Computing*, pages 105–109, San Antonio, Texas, 28 February–2 March 1999.