Genetic Paint: A Search for Salient Paintings

J. P. Collomosse and P. M. Hall

Department of Computer Science, University of Bath, Bath, U.K.,
{jpc | pmh}@cs.bath.ac.uk

Abstract. The contribution of this paper is a novel non-photorealistic rendering (NPR) algorithm for rendering real images in an impasto painterly style. We argue that figurative artworks are salience maps, and develop a novel painting algorithm that uses a genetic algorithm (GA) to search the space of possible paintings for a given image, so approaching an "optimal" artwork in which salient detail is conserved and non-salient detail is attenuated. We demonstrate the results of our technique on a wide range of images, illustrating both the improved control over level of detail due to our salience adaptive painting approach, and the benefits gained by subsequent relaxation of the painting using the GA.

1 Introduction

Paintings are abstractions of photorealistic scenes in which salient elements are emphasised. In the words of art historian E.H. Gombrich, "works of art are not mirrors" [1] — artists commonly paint to capture the structure and elements of the scene that they consider to be important; remaining detail is abstracted away in some differential style. This *differential level of emphasis* is evident in all artwork, from the sketches of young children to works of historical importance.

Processing images into artwork remains an active area of research within the field of non-photorealistic rendering (NPR). This paper presents a novel automatic NPR technique for rendering images in an impasto painterly style. Our approach contrasts with those before us in that we seek to emulate the aforementioned *differential emphasis* practised by artists — automatically identifying salient regions in the image and concentrating painting detail there.

Our algorithm makes use of a new image salience measure [2], that can be trained to select features interesting to an individual user, and which performs global analysis to simultaneously filter and classify low-level features to detect artifacts such as edges and corners. This enables us both to adaptively vary level of detail in painted regions according to their salience, and to vary brush stroke style according to the classification of salient artifacts. Further, we use a genetic algorithm (GA) to search the space of possible paintings for the given image, and so approach an optimal painting. A painting is deemed "better" if its level of detail coincides more closely with the salience magnitude of the original image, resulting in conservation of salient detail and abstraction of non-salient detail. Although we are not the first to propose relaxation approaches to painting [4, 5], our approach is novel in that we converge toward a *globally* defined minimum distance between salience and corresponding detail in the painting.

1.1 Related Work and Context

The development of automated painterly renderers arguably began to gain momentum with Haeberli's semi-automatic painting environments [6]. Fully automatic data dependent approaches were later presented, driven by heuristics based on local image processing techniques that estimated stroke attributes such as scale or orientation. Litwinowicz [7] employed short, linear paint strokes, which were clipped to thresholded edges. Treavett and Chen [8] proposed the use of local statistical measures, aligning strokes to axes of minimum intensity variance. A similar approach using chromatic variance was proposed in [9]. Hertzmann proposed a coarse-to-fine approach to painting [10] and was the first to automatically place curved (β-spline) strokes rather than dabs of paint. Our stroke placement algorithm is based firmly upon this technique. Gooch *et al.* [11] also use curved strokes fitted to skeletons extracted from locally connected regions.

A commonality exists between all of these algorithms; the attributes of each brush stroke are determined independently, by heuristics that analyse small pixel neighbourhoods local to that stroke's position. Rendering is, in this sense, a *spatially local* process. The heuristics typically seek to convey the impression of an artistic style whilst preserving content such as edges, and other artifacts contributing to the upper frequencies of the Fourier spectrum. Indeed, existing relaxation-based painting algorithms [4, 5] actively seek to maximise conservation of high-frequency content from the original image. Measures of variance [8, 9], or more commonly, simple edge detectors (such as Sobel) [7, 10] drive these heuristics. This results in a painting in which all fine detail is emphasised, rather than only the *salient* detail. Arguably this disparity contributes to the undesirable impression that such paintings are of machine rather than natural origin.

In Fig. 1 (left) we demonstrate that not all fine scale artifacts are salient; indeed in these images, salient and non-salient artifacts are of similar scale. Such examples make the case for some other measure of salience incontrovertible. When one speaks of the salience of image regions, one implicitly speaks of the importance of those regions relative to the image as a whole. It follows that *global* image analysis is a prerequisite to salience determination, rather than restricting attention to *spatially local* image properties.

Our desire to control level of detail in NPR is most strongly aligned with recent techniques, which appeal to user interaction to control emphasis. De Carlo uses a gaze-tracker [12] to guide level of detail in painting. Masks, specified manually or *a priori*, have also been used to interactively reduce level of detail [10, 13]. Yet the problem of automatically controlling painting emphasis remains; this paper presents a solution.

2 A Salience Measure for Painting

Salience is subjective; faces photographed in a crowd will hold different levels of salience to friends or strangers. User training is one way in which subjectivity can be conveyed to an automated salience measure, although current Computer Vision restricts general analysis to a lower level of abstraction than this example.
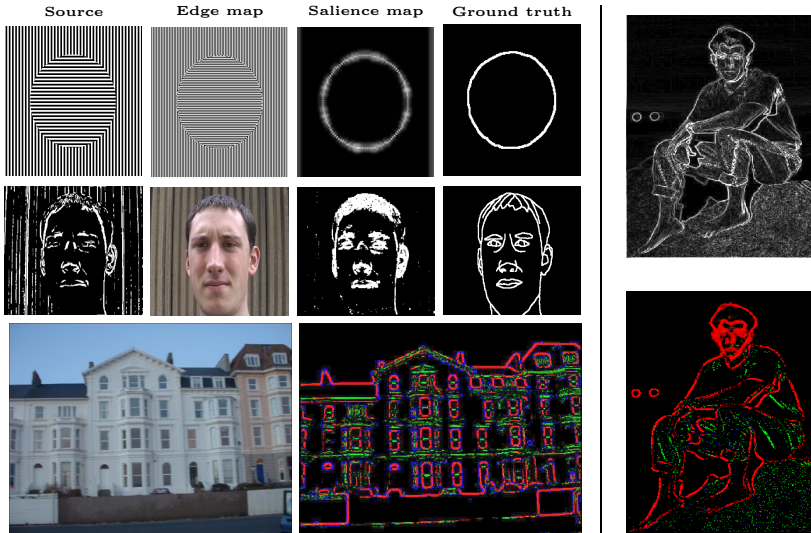
Fig. 1. Left: Examples of images edge detected, salience mapped, and a hand-sketched ground truth. We observe that the global, rarity based salience maps [2] are qualitatively closer to sketches, and can "pick out" the circle and face where local methods such as edge detection fail. The salience measure estimates salience magnitude and also classifies artifacts into trained categories (bottom row). Edges are red, ridges green, and corners blue. Right: Sobel edges (top) and salience map (bottom), corresponding to Fig. 6b. Salient edges are discriminated from non-salient high frequency texture.

We make use of a trainable salience measure, described more fully elsewhere [2], that combines three operators to estimate the *salience map* of an image — a scalar field in which the value of any point is directly proportional to the perceived salience of the corresponding image point. The first of the three operators performs unsupervised global statistical analysis to evaluate the relative rarity of image artifacts (after Walker *et al.*[14] who observe that salient features are uncommon in an image). Salient artifacts must also be visible, and a second operator filters detected artifacts to enforce this constraint. Finally, certain classes of artifact, for example edges or corners, may be more salient than others. This observation is accommodated by a third operator that is trained by the user highlighting salient artifacts in photographs. Signals corresponding to these artifacts are clustered to produce a classifier which may be applied to to estimate salience in novel images. This definition holds further advantage in that *classes* of salient features may be trained and classified independently.

This trainable salience measure is well suited to our NPR painting application for two reasons. First, the salience maps produced have been shown to be measurably closer to human figurative sketches of scenes than edge maps and a number of other prescriptive salience measures [3]. Second, the ability to estimate both the salience and the classification of image artifacts simultaneously allows us to vary stroke style according to the class of artifact encountered (Fig. 2). We begin by applying the salience measure to the source image; obtaining both

a salience map and a classification probability for each pixel. An intensity gradient image is also computed using Gaussian derivatives, from which a gradient direction field is obtained. The source image, direction field, salience map and classification map are used in subsequent stages of our painting algorithm.

3 Painting as a Search

Our observations of artists lead us to assert that the level of detail in a painting should closely correlate with the salience map of its source image. In this sense, the *optimality criterion* for our paintings is a measure of the strength of this correlation (defined in subsection 3.2, step I). We treat the painting process as a search for the "optimal" painting under this definition. Our search strategy is genetic algorithm (GA) based. When one considers the abstraction of a painting as an ordered list of strokes [6] (comprising control points, thickness, etc. with colour as a data dependent function of these), the space of possible paintings for a given source image is very high dimensional, and our optimality criterion makes this space extremely turbulent. Stochastic searches that model evolutionary processes, such as GAs [15], are often cited among the best search strategies in such situations; large regions of problem space can be covered quickly, and local minima more likely to be avoided [16, 17].

Our algorithm accepts as input a source image I; paintings derived from I are points in our search space. We begin by initialising a fixed size population of individuals. Each individual is single point in our search space, represented by an ordered list of strokes that, when rendered, produces a painting from I. Having initialised the population, the iterative search process begins. We now describe the initialisation and iteration stages of the search in detail.

3.1 Initialising the Painting Population

We initialise the search by creating an initial population of fifty paintings, each derived from the source image via a stochastic process. We now describe this derivation process for a single painting.

A painting is formed by compositing curved spline brush strokes on a 2D canvas of identical size to the source image. We choose piecewise Catmull-Rom splines for ease of control since, unlike β-splines (used in [10, 11]), control points are interpolated. A collection of "seed points" are scattered over the canvas stochastically, with a bias toward placement in more salient regions. Brush strokes are then grown to extend bi-directionally from each seed; each end grows independently until halted by one or more preset criteria. Growth proceeds in a manner similar to [10] in that we hop between pixels in the direction tangential to intensity gradient. The list of visited pixels forms the control points for the stroke. However, noise forms a component of any real image, and hop directions are better regarded as samples from a stochastic distribution. We have observed this noise to obey the central limit theorem [18], and so model this distribution
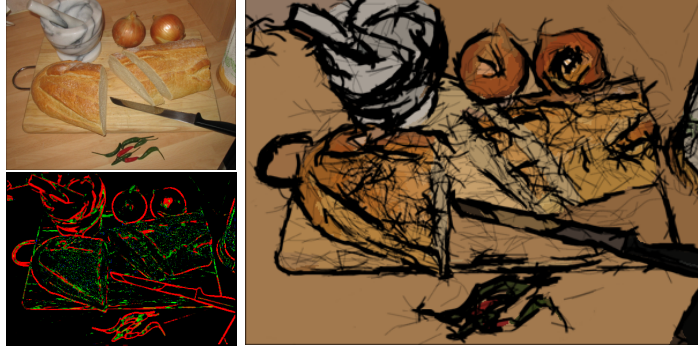
Fig. 2. Left: a still-life composition and corresponding salience map. Right: a loose and sketchy painting, exhibiting differential stroke style determined by local feature classification. Edges are drawn with hard, precise thick strokes; ridges with a multitude of light, inaccurate strokes. Rendered prior to the relaxation step of subsection 3.2.

as a zero centred Gaussian, $G(0,\sigma)$; we determine σ empirically (see subsection 3.1.1). Given a locally optimal direction estimate θ we select a hop direction by adding Gaussian noise $G(0,\sigma)$. The magnitude of the hop is also Gaussian distributed; $G(\mu',\sigma')$, both parameters being inversely proportional to the local value of the salience map. The growth of a stroke end is halted when either the curvature between adjacent pixels, or the difference between the colour of the pixel to be appended and the mean colour of visited pixels exceeds a threshold. This method initially yields a sub-optimal trajectory for the stroke with respect to our optimality criterion. However, for a "loose and sketchy" painting this is often desirable (see Fig. 2).

The degrees of freedom available from each of the many hops combine to create a range of stroke loci, at least one of which will result in the maximum conservation of salient detail. The combination of these optimally positioned strokes comprises the optimal painting, and it is by means of breeding the fittest paintings to create successively superior renderings, that we later search for such a painting in our iterative process. This process can out-perform stroke placements based purely on local estimates of direction.

3.1.1 Calibration for image noise

The choice of σ significantly influences stroke growth, and later the relaxation process. A value of zero forces degeneration to a loose and sketchy painting system; a high value will lengthen the relaxation process unnecessarily and also may introduce unnecessary local minima. We propose a one time user calibration process to select σ as follows.

The user is asked to draw a window around an image region where direction of image gradient is perceived to be equal; i.e. along which they would paint strokes of similar orientation. Gradient direction within this window is sampled, and σ computed as twice the unbiased standard deviation of the sampled angles.

We typically obtain similar σ values for similar imaging devices, which allows us to perform this calibration very infrequently. A typical σ ranges from around 2 to 5 degrees. This variation allows between 12 and 30 degrees of variation per hop which, given the number of hops per stroke, yields a wide range of stroke loci. This adds credence to our argument for a relaxation process taking into account image noise; local variations due to uncompensated image noise will likely cause inaccurate stroke placements in single iteration painterly systems [7, 10, 9, 11].

3.1.2 Rendering a Painting

At this stage we may render one of the paintings in our initial population to produce a "loose and sketchy" painting (Fig. 2). Alternatively we may proceed to the iterative search stage of subsection 3.2 to locate a more optimal painting — each iteration also requires paintings to be rendered to evaluate fitness. We now describe how paintings are formed from individuals in the population.

A painting is formed by scan-converting and compositing its list of curved spline brush strokes. Stroke thickness is set inversely proportional to *stroke salience*; taken as the mean salience over each control point. Stroke colour is uniform and set according to the mean of all pixels encompassed in the footprint of the thick paint stroke. During rendering, strokes of least salience are laid down first, with more salient strokes being painted later. This prevents strokes from non-salient regions encroaching upon salient areas of the painting. The ability of our salience measure to differentiate between classes of salient feature (e.g. edge, ridge) also enables us to paint in context dependent styles. In Fig. 2 the classification probability of a feature is used as a parameter to interpolate between three stroke rendering styles *flat*, *edge* and *ridge*.

3.2 Iterative Relaxation by GA

Genetic algorithms (GAs) simulate the process of natural selection by breeding successive generations of individuals through cross-over, fitness-proportionate reproduction and mutation [17]. In our implementation such individuals are paintings; their genomes being ordered lists of strokes. We now describe a single iteration of the GA search, which is repeated until the improvements gained over the previous few generations are marginal (the change in both average and maximum population fitness over a sliding time window fall below a threshold).

I. Fitness and Selection. The entire population is rendered, and edge maps of each painting are produced using by convolution with Gaussian derivatives, which serve as a quantitative measure of local fine detail. The generated maps are then compared to a precomputed salience map of the source image. The mean squared error (MSE) between maps is used as the basis for determining the fitness of a particular painting; the lower the MSE, the better the painting. In this manner, individuals in the population are ranked according to fitness. The bottom 10% are culled, and the top 10% percent pass to the next generation; this promotes convergence. The top 90% percent are used to produce the remainder of the next generation. Two individuals are selected stochastically with a bias

Fig. 3. Relaxation by GA search. Detail in the salient region of the 'dragon' painting sampled from the fittest individual in the 1st, and 70th generation of the relaxation process. Strokes converge to tightly match contours in salient regions of the image thus conserving salient detail.
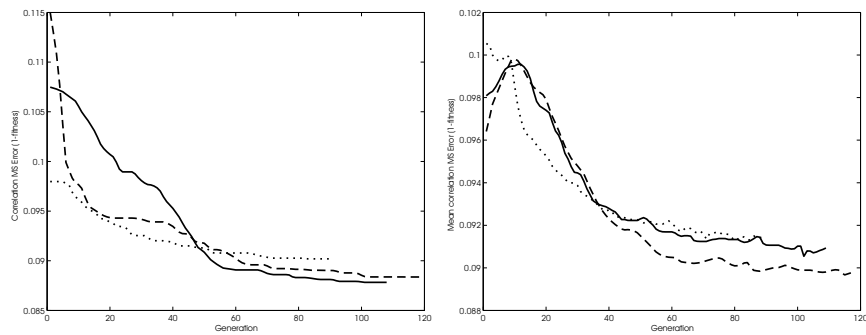
Fig. 4. Left: Three runs of the relaxation process; dotted line corresponds to the model (Fig. 6a), dashed line the dragon (Fig. 3) and solid line the truck (Fig. 6g). MSE of the fittest individual is plotted against time. Right: MSE averaged over each generation

to fitness, and bred via *cross-over* and *mutation* to produce a novel offspring for the successive generation. This process repeats until the population count of the new generation equals that of the current generation.

II. Cross-over. Two difference images, A and B, are produced by subtracting the edge maps of the parents from the salience map of the source image. By computing the binary mask $A > B$, and likewise $B > A$, we are able to determine which pixels in one parent contribute toward the fitness criterion to a greater degree than those in the other. Since the primitives of our paintings are thick strokes rather than pixels, we apply dilation to both masks. Strokes seeded within the set regions in each mask are copied from the respective parent to the new offspring. A binary AND operation between masks yields mutually preferred regions, for which the contributing parent is decided arbitrarily.

III. Random Mutation. A "temporary" painting individual is synthesised as described in subsection 3.1. A binary mask is produced containing several small discs of stochastic number, location and radius. Strokes seeded within set

Fig. 5. Detail from Fig. 6g, region *A*. Using our adaptive approach, salient detail (sign-post) is conserved, and non-salient texture (shrubbery) is abstracted away. Left: original. Middle: existing approach [7]. Right: our proposed GA approach.

regions of the mask are substituted for those in the temporary painting. On average, mutation occurs over approximately 4% of the canvas area.

Implementation Notes. The evaluation step is the most lengthly part of the GA process, and rendering is farmed out to several machines concurrently. In our implementation we distribute rendering over a small heterogeneous (Pentium III/UltraSPARC) cluster. The typical time to render a 50 painting generation at 1024×768 resolution is on average 15 minutes over 6 workstations. Relaxation of the painting can therefore take in the order of hours, but significant improvements in stroke placement can be achieved, as can been seen in Fig. 3.

4   Results and Discussion

We present a gallery of rendered paintings in Fig. 6. The painting of the model in Fig. 6b converged after 92 generations. Thin precise strokes have been painted along salient edges, while ridges and flats have been painted with coarser strokes. Observe that non-salient high-frequency texture on the rock has been attenuated, yet tight precise strokes have been used to emphasise salient contours of the face. In the original, the high frequency detail in both regions is of similar scale and magnitude; existing painterly techniques would, by contrast, assign both regions equal emphasis. With current techniques, one might globally increase the kernel scale of a low-pass filter [10] or raise thresholds on Sobel edge magnitude [7] to reduce emphasis on the rock. However this would cause a similar drop in the level of detail on the face (Fig. 6a). Conversely, by admitting detail on the face one would unduly emphasise the rock (Fig. 6c). We automatically differentiate between such regions using a perceptual salience map (Fig. 1) – contrast this with the Sobel edge field where no such distinction can be made.

We present a still-life in Fig. 6e which achieved convergence after 110 generations. In Fig. 6f we present a similar painting prior to relaxation, demonstrating differential rendering style; strokes with a high probability of being edges are darkened to give the effect of a holding line. Further examples of level of detail adaptation to salience are given in Fig. 6g. In region *A*, the salient sign is empha-

Fig. 6. A gallery of paintings illustrating application of our algorithm. Higher resolution electronic versions of all our paintings are included in the material accompanying this paper.

sised whilst non-salient texture of the background shrubbery is not (see Fig. 5). For the purposes of demonstration we have manually altered a portion of salience map in region B, causing all detail to be regarded as non-salient.

All of our experiments have used populations of 50 paintings per generation. We initially speculated that population level should be set in order of hundreds to create the diversity needed to relax the painting. Whilst convergence still occurs with such population limits, it requires, on average, 2 to 3 times as many iterations to achieve. Such interactions are often observed in complex optimisation problems employing GAs [17]. We conclude that the diversity introduced by our mutation operator is sufficient to warrant the lower population limit.

As regards rendering, we might choose to texture strokes to produce more realistic brush patterns — however, we have concentrated on stroke placement rather than media emulation, and leave such implementation issues open. We believe the most productive avenues for future research will explore both new fitness functions and alternative uses for salience measures in image-based NPR.

References

1. Gombrich, E.H.: Art and Illusion. Phaidon Press Ltd., Oxford (1960)
2. Hall, P.M., Owen, M., Collomosse, J.P.: A trainable low-level feature detector. In: Proc. Intl. Conf. on Pattern Recognition (ICPR). (2004) 1:708–711
3. Hall, P., Owen, M., Collomosse, J.P.: Learning to detect low-level features. In: Proc. 15th British Machine Vision Conf. (BMVC). Volume 1. (2004) 337–346
4. Sziranyi, T., Tath, Z.: Random paintbrush transformation. In: Proc. 15th Intl. Conf. on Pattern Recognition (ICPR). Volume 3., Barcelona (2000) 155–158
5. Hertzmann, A.: Paint by relaxation. In: Proc. Comp. Graph. Intl. (2001) 47–54
6. Haeberli, P.: Paint by numbers: abstract image representations. In: Proc. ACM SIGGRAPH. Volume 4. (1990) 207–214
7. Litwinowicz, P.: Processing images and video for an impressionist effect. In: Proc. ACM SIGGRAPH, Los Angeles, USA (1997) 407–414
8. Treavett, S., Chen, M.: Statistical techniques for the automated synthesis of non-photorealistic images. In: Proc. 15th Eurographics UK Conf.. (1997) 201–210
9. Shiraishi, M., Yamaguchi, Y.: An algorithm for automatic painterly rendering based on local source image approximation. In: Proc. ACM NPAR (2000) 53–58
10. Hertzmann, A.: Painterly rendering with curved brush strokes of multiple sizes. In: Proc. ACM SIGGRAPH. (1998) 453–460
11. Gooch, B., Coombe, G., Shirley, P.: Artistic vision: Painterly rendering using computer vision techniques. In: Proc. ACM NPAR (2002) 83–90
12. DeCarlo, D., Santella, A.: Abstracted painterly renderings using eye-tracking data. In: Proc. ACM SIGGRAPH. (2002) 769–776
13. Hertzmann, A., Jacobs, C.E., Oliver, N., Curless, B., Salesin, D.H.: Image analogies. In: Proc. ACM SIGGRAPH. (2001) 327–340
14. Walker, K.N., Cootes, T.F., Taylor, C.J.: Locating salient object features. In: Proc. 9th British Machine Vision Conf. (BMVC). Volume 2. (1998) 557–567
15. Holland, J.: Adaptation in Natural and Artificial Systems. U. Michigan (1975)
16. de Jong, K.: Learning with genetic algorithms. Machine Learning **3** (1988) 121–138
17. Goldberg, D.: GAs in Search, Optimization, and Machine Learning. Add.W. (1989)
18. Collomosse, J.P.: Higher Level Techniques for the Artistic Rendering of Images and Video. PhD thesis, University of Bath, U.K. (2004)