Using Artificial Intelligence Techniques to Emulate the Creativity of a Portrait Painter

Steve DiPaola  
Simon Fraser University  
Canada  
sdipaola@sfu.ca

Graeme McCaig  
Simon Fraser University  
Canada  
graeme_mccaig@sfu.ca

We present three new machine learning based artificial intelligence (AI) techniques, which we have added to our parameterized computational painterly rendering framework and show their benefits in computational creativity and non-photorealistic rendering (NPR) of portraits. Traditional portrait artists use a specific but open human creativity, vision, technical and perception methodologies to create a painterly portrait of a live or photographed sitter. By incorporating more open ended creative, semantic and concept blending techniques, these new neural network based AI techniques allow us to better model the creative cognitive thinking process that human painters employ. We analyse these AI based methods for their operating principles and outputs that together along with our parameterized NPR modules can be relevant to the field of computational creativity research and computational painterly rendering.

Artificial intelligence; computation creativity; painterly rendering; computer graphics, cognitive science

1. INTRODUCTION

Traditional portrait artists use a specific but open human creativity, vision, technical and perception methodologies to create a painterly portrait of a live or photographed sitter. To create a portrait, a human portrait painter setups up the room/lighting, positions the sitter, interviews the sitter to understand/capture inner (personality, …) and outer resemblance, and at the same time merges these goals with how the artist wants to convey their own painting style in the trajectory of their painting career as well as striving for some personal and universal (cultural, political, philosophical) truth of the current world they live in. Balancing these goals (sitter inner/outer resemblance, artists goals, universal statements) an artist has a palette of choices of themes, brush style, colour plan, edge and line plan, abstraction style, and emotional narrative at their disposal to create the final painting.

Our research uses several artificial intelligence techniques including Genetic Algorithms, Neural Networks and Deep Learning neural networks in an attempt to begin to understand and emulate this creative process. For instance our Deep Learning Networks can understand how to balance or blend different aesthetic, conceptual, abstraction concerns at a semantic level. We review three new Artificial Intelligence techniques that we add to our Painterly Rendering software framework that are bringing more high level tools and insights to model the cognitive creative process of a portrait artist.

![Figure 1: Outputs of guided Deep Dream (c.) blending inputs (a.,b.) then put through our NPR system.](image)

2. COMPUTATIONAL PAINTERILY RENDERING

Non Photorealistic Rendering (NPR) is a computer graphics technique, which creates imagery with a wide variety of expressive styles inspired by painting, drawing, technical illustration, cartoons and mapping (Gooch & Gooch, 2001). This is in contrast to typical computer graphics, which focus on photorealism. NPR already has applications in video games, animation, movies, architectural and technical illustration, and rising fields such as computational photography, art therapy and virtual reality.

Many current computer painterly rendering systems rely on computer imaging approaches (e.g. edge detection, image segmentation) that model at the
physical level such as blobs, strokes and lines. Our novel painterly rendering research approach relies more on parameterizing a cognitive knowledge space of how a painter creatively thinks/paints.

Our cognitive painting system, Painterly (DiPaola 2007, 2009, DiPaola et al 2013) which models the cognitive processes of artists, uses algorithmic, particle system and noise modules to generate artistic colour palettes, stroking and style techniques. This paper will explore three new AI “thinking about what to paint first” pre-painting systems to our NPR framework.

Artists and scientists have different approaches to knowledge acquisition, usage and dissemination. This research work is one attempt to bridge these different fields. Our domain of inquiry is the creation and viewing of fine art portrait painting -- we are interested in elucidating cognitive and perceptual mechanisms or ‘cognitive correlates’ which correspond and relate to artists’ techniques and conceptions regarding fine art painting and then modelling the human process in software.

Here we are interested in the process of fine art painting, how it is created and perceived as well as the cognitive phenomena, which underlay it. Therefore the process of fine art painting is the behaviour we attempt parse from a ‘cognitive correlate’ perspective. We are motivated in part by the growing recognition that ‘Artists are Neuroscientists’: i.e. they have discovered valuable ways of understanding and working with human cognitive and perceptual mechanisms to achieve authorship techniques to convey desired experiences and narrative in their created art works (Cavanagh, 2005; Zeki, 2001). That human like a art paintings because our brain is stimulated by it more than say a photograph or real life and artists have intuited how to exploit these neural mechanisms in the brain by specific based down painterly knowledge. Whether artists are or act like a type of neuroscientist as many cognitive scientists like to state or more simply that artists have a passed down methodology and talent space where they use their eyes, perception and mind to produce a painting in a way that when analysed through a certain cognitive lens can benefit both the arts and cognitive sciences, is less the point. Hence, we explore art topics through a cognitive science perspective with the aim of enriching our understanding of both art practice (i.e. the act of fine art painting) and the underlying perceptual mechanisms.

The work we present here attempts to use strong analysis of the artistic painterly process through a lens of new scientific understanding of human cognition to create a cognitive knowledge based painterly NPR software toolkit that can have both wider range and improved results compared to many current NPR techniques. By limiting the investigation to fine art portraits (as opposed to all painterly art forms) and how they vary from their photography analogues, the system can use strong knowledge (e.g. salience) of portraits and faces to inform the system.

3. PORTRAIT PAINTING AND SALEINCE
Fine art painting involves making subjective decisions about what aspects of the source are particularly noticeable, important, prominent or salient. Painterly salience involves emphasising details using painterly techniques. Later we will discuss how Neural Network based AI techniques such as Deep Learning (DL) which we have begun to employ, give us a more semantic tootset to work with such as conceptual blending, salience and creative goals (e.g. style, abstraction resemblance). For example, given the goal of creating a regal but approachable portrait of the Bishop sitting in front of them, an artist considers what elements or regions of the portrait should stand out, what methods to use and modifies these regions using specific artistic manipulations or conceptual blends that exert a psychological impact on viewers. Cognitive researchers believe that some of the techniques used by artists include the following:

- “simplify, compose and leave out what’s irrelevant, emphasizing what’s important”. (DiPaola 2007),
- “direct the viewer’s attention to the relevant content and to influence their perception of it” (Santella & Decarlo, 2002),
- create “abstractions of photorealistic scenes in which the salient elements are emphasized” (Vanderhaeghe, 2013).

There has been much debate amongst psychologists, vision scientists, image processing experts and even NPR researchers about the true meaning of salience and how to measure it. Santella and Decarlo (2002) tracked a casual viewer’s gaze through a scene in a NPR painting (with an eye tracker) to determine what part of the scene needed to be emphasized. Our view is that painterly salience (emphasising a style or details via painterly techniques) is authored not by the viewer but by the artist -- to emphasise what she believes she wants to convey. It is a goal-oriented endeavour. In a landscape scene, for example, with mountains and a stream in the background and two horses and a tree in the grassy foreground - what should be emphasised? The tree or perhaps one or both of the horses? Zeng and Zhao have attempted to deal with this problem in NPR by creating hierarchical salient parse trees of a scene. In the above example, the mountain, stream, background,
grass, tree and horse objects would be mapped by attention priority in a salient priority tree structure form. Then emphasis and filtering painterly techniques may be used to complete the painting, say with one house at the top of the priory tree (Zeng et al., 2009). Since we take a knowledge-based approach, we again limit the problem space to one specific genre with considerable knowledge data: fine art portraits. A portrait contains the beginning of a hierarchical tree structure of salience. The approach to salience taken here is to start with portraiture and then expand to other genres as techniques are developed and refined. Portraiture makes more extensive use of cognitive science research than other art forms such as landscapes (research that extends considerably beyond face-specific knowledge rules).

In this work, we attempt to collect and use both general artistic methodology and (when appropriate or needed) specific portrait knowledge (what we call face semantics, i.e., the fact that eyes are more salient than hair). Hence, our painterly NPR system includes processes to deal with face semantics, style considerations, semantic and conceptual blending of ideas, all in the scripting systems that allow known portrait rules to be exploited.

4. PAINTERLY CREATIVE PROCESS

From a process standpoint, artists investigate something of interest about a scene or sitter they have chosen to paint, and come up with a 'visual narrative' of this scene that includes how it will be expressed using the language and techniques of painting, and how the painting will deviate from the real scene. Factors that may be taken into account during this process include (1) the content of the scene, (2) their career trajectory (e.g. techniques and innovations they have experimented with, advice they have received from mentors, and so forth) mixed with (3) personal feelings and intuitions as they arise in the moment about such factors as what will invite exploration or provide closure. Much of an artist’s time is spent re-assessing what is now there in front of them, and re-evaluating what to do next based on that data, a process that Gabora (2010) refers to as honing through re-iterated context-driven actualization of potential.

Some artists stick tightly to their basic formula or plans, while others deviate widely based on the source and the narrative. As they begin, they exploit a limited tonal and colour range where they can re-scale and re-centre the relative range of all aspects of the source image and source masses (tonally, tonal balance, concepts, themes, colour, colour balance, detail, edging, shapes, …) in ways they perceive fit their style and narrative goal. They work through the painting, taking anywhere from hours to days to weeks. They frequently make mental comparisons of how different elements or regions of the painting appear relative to one another. They usually start with large tonally limited masses (i.e., two or three values) that they tone sample mentally. Some artists squint to help both blur out details and remove colour importance. Then they shift to another region of the painting. As they continue, they increasingly focus in on smaller details i.e., smaller brush strokes, and more colour choices. These details (e.g. tonal range) are not only re-scaled and re-centred from the source but typically are picked using different aesthetic rules from those used initially, such as a warm/cool remapping. These decisions fall into place depending on what commitments they have already made or how the initial potentiality of the yet-to-be painted has been actualized thus far.

At some point in this progressively refined process, they cycle into a 'see, think, paint' loop, deciding what area to work on next, sampling a cognitive region (like minded tone sampled) from the source, thinking about it -- remapping it conceptually, and through the craft of stroking on canvas, commit this new conception to paper. It is this “think” area our new AI techniques are most interested in. To model in your head as a pre-act to the implementation stave of painting. With each cycle, some of the potentiality of the previous cycle gets actualized, and simultaneously, potentialities for future iterations arise. They cycle through this process repeatedly until the semantic area they are working on (e.g., the lower face) is complete. They may then make a more detailed pass involving progressive refinement of the area. Light is the primary consideration in deciding how the painting unfolds (in the form of tone sampled shapes), although artists also consider volume to a lesser extent (i.e., stroke up the volume of the nose) and content areas (working in the background or a cast shadow or for an eye, I do something different and specific).

We have attempt to model this process in Painterly our NPR toolkit, which uses a knowledge-based approach to painterly rendering to create a wide variety of computational paintings styles based on source portrait photograph and semantic knowledge maps (DiPaola et al., 2010; DiPaola, 2009; DiPaola, 2007). The knowledge rules were sourced/encoded by categorizing traditional ‘artistic painter process’ and linking the findings to theories from human vision, colour and perception, as well as semantic models of the face.

Painterly has contributed to research understanding in the cognitive nature of art and vision science, mainly with in empirical techniques (i.e. eye tracking studies) that allow images to be varied in systematic ways while still being judged as plausible works of art (DiPaola et al., 2010; DiPaola, 2009; DiPaola,
Using Artificial Intelligence Techniques to Emulate the Creativity of a Portrait Painter
Steve DiPaola, Graeme McCaig

2005). However it is still difficult for researchers/users to script the scores of parameters to make a strong painterly recipe. Painterly has two main sections. The first ‘Thinker’ section mimics the cognitive high level painterly process deciding progressively detailed passes and the cognitive blobs (shapes) that painters work in. Next the ‘Painter’ section implements the Thinker’s plans in low-level variables of brush’, size, length and transparency per pass and per cognitive blob. Lastly, the painter section uses a colour system which translates tonal value into final colour based on the semantic regions (eyes, clothes, etc.) of the destination painting. While our Painterly system has been used for both stages (thinker-painter), we have begun to update the thinker stage which is concerned with modelling the perceptual and creative process of human artists and painters, with new AI based “inner” modules (the thinker) that then output to the NPR outer or implementation modules (the painter). This follow sections of this paper mainly deal with 3 new AI based “thinker” techniques that we combine with our NPR “painter” output. These new AI modules better emulate the cognitive semantics we have described a human painter develops in the artistic painterly process. Currently the modules are separate interchangeable systems where several different AI thinker systems create intermediate output that is then further refined by our main NPR implementation systems for stroke based painterly rendering. These new AI “thinker” modules or systems that now can give us a greater level of high level semantic and concept blending processes to better emulate a cognitive approach to the thinker stage.

5. AI NEURAL NETWORK PAINTERLY SYSTEMS

We now look at three new systems we have added to our NPR framework. All use the AI technique called neural networks. Classified in the AI domain of machine learning and cognitive science, neural networks (NNs) are a family of techniques inspired by biological neural networks (the central nervous in the brain). These interconnected “neurons” are layered at different levels with approximate functions and numeric weights that can be tuned based on experience, making neural nets adaptive to inputs and capable of learning. Two of our systems use a large or deep level of layers and are therefore called Deep Learning (DL) neural networks.

5.1 Planer or Cubism based Abstraction

Our first AI thinker system takes on planar (e.g. cubist style) abstraction. Many AI creativity and NPR systems strive for abstraction by simply varying parameters like stroke length to achieve a level of abstraction. This is a worthy way to create rendered abstraction but leaves out cognitive abstraction where the artist is using abstraction in deeper ways for instance to concept blend different thoughts or meanings. Our first attempt at this deeper “thinker” abstraction is with deep or severe planar abstraction. This neural network system uses a regressor based image stylization which has been influenced by blends of genetic algorithms and hill
climbing but rephrases the problem as a much speedier machine learning technique. With this neural network AI technique, pixels of a source image are treated as a learning problem: so the system takes the (x,y) position (each pixel) on a grid and learns to predict the colour (the pixel colour) at that point using regression to (r,g,b). The image information is encoded in the weights of the network. In figure 2 we see a series of planar abstractions from a photo source from our system.

The Python based system allows us to control the level from very severe abstraction (i.e. 3 planes) all the way through mild abstraction with 100’s of planes as in figure 2. It should be noted that this “thinking” AI phase first outputs less painterly low resolution simple planes that have one colour per plane, simulating not a finish painting at this point but thinking model about abstracted planes in the artists mind. To make the finished thinker-painter art pieces in figures 2 and 3, we then take this inner abstraction output and send it through our Painterly NPR system where artistic distortion, human level variation and brush stroking complete the “painting on canvas” look of the final work. This is more evident in Figure 3 where several different levels of first, AI planer abstraction are completed with several different NPR recipes of painterly rendering. Our research lab is now in performing user studies where we testing users reaction to a range of abstractions from this AI output. Our early test results show that users both prefer the abstractions to the original photography they were created from in terms of a work of art, as well are able to identify correct aesthetic descriptors (e.g. cheerful, trust, awe, scary, …) that people see the original source photograph in even severely abstracted versions. This preliminary study data supports that this AI based abstraction technique seems to have some qualities of human based abstract art – that the output holds the qualities inherent to the original sitter but within a simpler aesthetic form.

5.2 Deep Dream and Deep Style Techniques

Deep Dream (Mordyintsev et al. 2015) and Neural Artistic Style (Gatys et al 2015) are two techniques for modifying images through a process of analysis and search involving Deep Convolutional Neural Networks (DCNNs) (Krizhevsky 2012, LeCun 1998). DCNNs fall within the rapidly growing field of “Deep Learning” research (Bengio 2013). We adopt the shortened name “Deep Style” for ease of discussion of the technique in (Gatys et al 2015), abbreviating Deep Dream and Deep Style as DD and DS respectively.

DCNNs are typically trained on large datasets of images in order to build up a multi-level, feature-based re-encoding system, in which low-level features represent local features such as lines and curves while high-level features represent more abstract, composite visual “concepts” such as “spoked wheel pattern” or “animal hind-leg shape”.

This method of representing images in a multi-layer network with increasing abstraction is thought to bear resemblance to the way the human brain processes visual perception (DiCarlo, 2012). This structure facilitates performance in discrimination/classification tasks such as recognizing objects in an image as belonging to a certain learned category; however, as found with DD and DS, it is also possible to use DCNNs generatively, creating images which emphasize certain features or feature-layers of an image, or combine the features of one image with features from a second image to create an output image sharing qualities of both.

These generative abilities resonate with the idea from Neuroaesthetics (Zeki, 2001) that a possible role and motivation of art is for audiences/artists to reveal or stimulate the neural mechanisms of perception – we can view the different low- and high-level feature encodings within a DCNN as different perspectives on the essence of an image as analysed within a brain. In a different paper (McCaig et al 2016), we have examined how the combination of image features amounts to a computational model of visual concept blending and relates to Computational Creativity as a field.

We have implemented both DD (github.com/google/deepdream) and DS (github.com/fzliu/style-transfer), using the Caffe deep learning framework (Jia, 2014), as modules within our AI-based painting software toolset, currently using them as a pre-processing stage which simulates an artists’ imagination and perception, transforming an image before it is sent to the second, artistic stroke-placement phase.

We now compare and demonstrate the operation of DD and DS within our system. Deep Dream (Mordyintsev et al. 2015) has two basic modes of operation. The mode we might call “free hallucination” begins with a source image, and uses backpropagation and gradient ascent to gradually transform the image pixels in order to emphasize the most strongly-activated features from a certain user-selected network layer. This results in the emphasis of pre-existing shapes and patterns as well as the appearance of hallucinated patterns in which the network gravitates towards “seeing” patterns it has learned to recognize. Figure 4 presents results from our DD implementation running in free hallucination mode.

Deep Dream also has a guide-image mode, which again uses backpropagation and gradient ascent, to analyse the strong features from one “guide” image and emphasize the best-matching features from a
second source image by transforming the pixels in this second image. In Figure 5 we show how the algorithm transfers visual attributes from one image to another, depending on the network layer used for feature comparison. In Figure 6 we show further examples of how different guide images can affect different visual attributes.

Deep Style (Gatys et al 2015) is somewhat similar in concept to Deep Dream’s guide-image mode. However, notably, it initializes its output image starting from random noise, and then optimizes the image based on multiple network layers simultaneously. The output image is optimized to closely match features at a certain higher layer/layers with a “content” input image (capturing the semantic object identity and placement from that image). It simultaneously is optimized to match a correlation-based metric on features from multiple layers with a secondary “style” image (capturing colours and textures from that image). Due to this separation of content and style, and the multi-layer technique which tends to closely capture the look of specific style image fragments, DS been found to be quite successful at painting style transfer; i.e. applying the colourist and brush stroking style from a painting to a new source image to create, e.g. an “artificial Rembrandt forgery”. Fig 7 shows two examples of DS used to transplant painting style.

Figure 4: Deep Dream and our NPR system creating different techniques depending on the layer.

Figure 5: Deep Dream where patterning from the source butterfly (a) creates output based on a lower (b) or higher (c) network level. Source photo by Ano Lobb.

We plan to further explore the possibilities for dynamic interaction between the perception/imagination phase and the stroke-painting phase. In this current system we apply the DD/DS module to the source photo first, followed by our NPR painting phase. Therefore all figures except 5 & 6 are completed with treatment with the Painterly cognitively-inspired painting module.

Figure 6: Deep Dream techniques using a) geometric and b) bee guided images and the results. Green photo by Goran Konjevod.

Our Painterly module which is an extension to our cognitive painting system, Painterly (DiPaola 2009) which models the cognitive processes of artists based on years of research in this area, uses
algorithmic, particle system and noise modules to generate artistic colour palettes, stroking and style techniques. It is in the NPR subclass of stroke based rendering and is used as the final part of the process to realize the internal DCNN models with stroke based output informed by historic art making. Specifically in this example, aesthetic advantages include reducing some noisy artefacting of the generated DCNN output via cohesive stroke based clustering as well a better distributed colour space.

![Image of Rembrandt and Freud paintings](image)

**Figure 7:** Deep Style using Rembrandt (a.,b.) and Freud paintings (c.,d) as style guides.

6. CONCLUSION

We have demonstrated how Machine Learning techniques such as Neural Networks and Deep Learning can play an important part in bringing higher level of creative semantics to painterly rendering and computational creativity. We have begun to use them to blend high level visual, sematic and creative concepts based on their wider context and associations which contribute to new forms of cognitively-relevant models of artist creativity and imagination. We have examined three new AI sub-systems, analysing their operating principles and outputs that together along with our parameterized NPR modules can be relevant to the field of Computational Creativity research and computational painterly rendering.

6. ACKNOWLEDGEMENTS

We would like to thank Liane Gabora, Sara Salevati, Daniel McVeigh and Jon Waldie for their thoughts and contributions as well as NSERC and SSHRC funding agencies.

7. REFERENCES


Le, Q. V., Ranzato, M., Monga, R., Devin, M., Chen,


