Simple Art as Abstractions of Photographs

Peter Hall[*]
University of Bath, UK

Yi-Zhe Song[†]
Queen Mary University London, UK

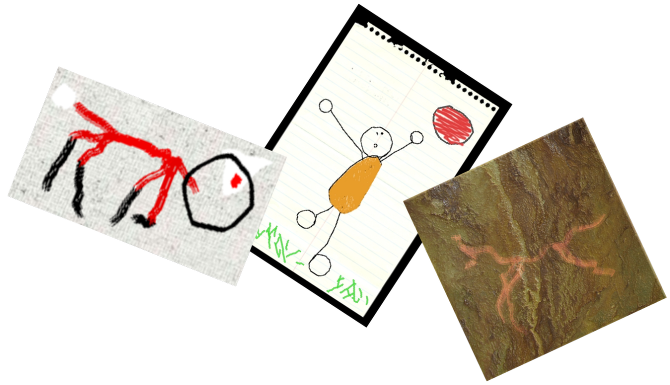


Figure 1: *Examples of Simple Art.*

Abstract

This paper shows that it is possible to semi-automatically process photographs into *Simple Art*. Simple Art is a term that we use to refer to a group of artistic styles such a child art, cave art, and Fine Artists as exemplified by Joan Miró. None of these styles has been previously studied by the NPR community. Our contribution is to provide a process that makes them accessible.

We describe a method that automatically constructs a hierarchical model of an input photograph, and asks a user to identify objects inside it. Each object is a sub-tree, which can be rendered under user control. The method is demonstrated using emulations of Simple Art. We include an assessment of our results against a set of norms recommended by a Cultural Historian. We conclude that producing Simple Art raises important technical questions, especially surrounding the interplay between computational modelling and human abstractions.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Display Algorithms

Keywords: Photographic Processing, Image Abstraction, Simple Art, Child-like Art,

*e-mail:pmh@cs.bath.ac.uk
†yizhe.song@eecs.qmul.ac.uk

CAe 2013, July 19 – 21, 2013, Anaheim, California.

ACM 978-1-4503-2203-4/13/07 $15.00

1 Introduction

Every child is an artist. The problem is how to remain an artist once we grow up, Picasso is reputed to have said. The problem of drawing as a child was the initial motivation behind this paper. The issue has received no attention we know of in the Computer Graphics literature, yet is clearly deemed important by an acknowledged genius of Fine Art.

The paper is a first step towards an automated system that is able to draw in a way similar to a child. More exactly, the paper describes the essential elements of a semi-automatic system that processes photographs into emulations of child art, cave art, Miró's art, and related pictures. There is no collective noun for this group, so for the purposes of convenience we use the term *Simple Art*. The term is intended to cover not just the examples in this paper, but also real world classes of art, not just child art, cave art, and Miró, but also Picasso, Lowry, Kadinsky, the art of aborigines in Australia, and other similar forms.

Simple Art has a defining characteristic: it depicts the connections between an object's parts. The importance of connectivity to Simple Art is most evident in stick figure drawings, where limbs and body are lines that connect hand to shoulder, often a circle for the head. The importance of connectivity is evident too in examples that show someone's head unattached from their body, which is not uncommon in children's drawings. In such cases the connection between head and body is recognised by the child but it is not drawn. This suggests that the connection is understood as an abstraction, and it relationship between parts that is of interest to this paper.

Simple Art appears simple to make: it takes such little skill to draw a stick figure that almost everyone over about 3 years old can do it. Yet appearances are deceptive in this case, adults cannot draw like children. However hard it is for an adult human to draw like a child, it is harder still for a computer. The underlying problem for a computer is that the child is manufacturing what seems to be an almost arbitrary visual representation of an object. Clearly, the representation is not fully arbitrary, otherwise the object would be unrecognisable. The interesting point is that it is difficult to imagine raising the degree of abstraction above that witnessed in children's art (and Simple Art) while still being able to recognise the image as belonging to a class of objects. We note that very young in-

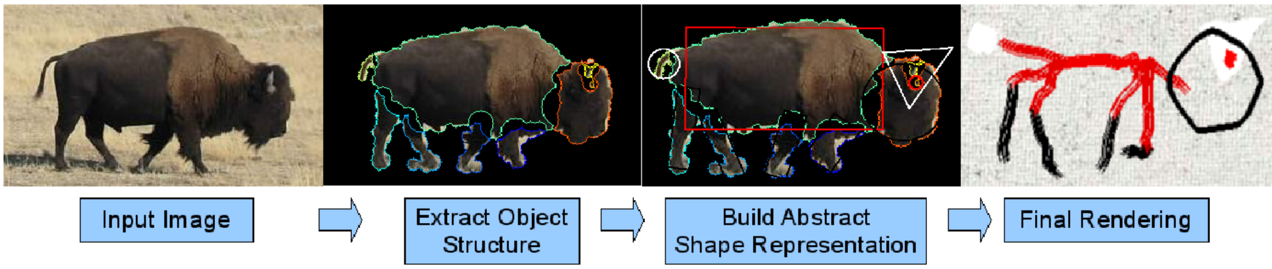


Figure 2: *Out method in a nutshell: An input photograph is processed a simplified graph structure, abstracted as simple geometric shapes, then rendered in some artistic style. (Bison copyright Big Sky Fishing.)*

fants are incapable of drawing recognisable objects, so the level of skill needed to produce Simple Art is considerable: it takes years to learn.

From a computational point of view we want a model that can be rendered in many different ways. Different ways of rendering are needed to cope with the different range of mark making tools and methods that exist, and thereby add verisimilitude to an emulation. Our rendering engine allows the user to set the render to produce output in a particular sub-style, but the details are the rendering of the computer; these are stochastically controlled. In this we obtain a different but similar looking result each time from the renderer, even when the same input model is given.

Simple models are easy to make if a user takes full responsibility for their manufacture, but very difficult to create automatically from photographs. This is because the model must not only segment an object but also determine its parts and their connectivity. The problem of fully automated acquisition is beyond the scope of any algorithm that exists today. Our solution is a semi-automatic system that allows users to construct a model using just a few mouse clicks. The model output references semantic parts sufficiently well that examples of Simple Art can be produced using it. Figure 2 offers an illustration of our system.

2 Background

Non-photorealitic rendering is a wide area of study, stretching from Scientific Visualisation to sophisticated interactive media. We are interested in NPR from photographs (NPRP), and in particular producing artwork as output – a sub-genre we call Aesthetic Rendering, which influences the manner in which we will assess the quality of our work (see Section 5).

Non-Photorealistic Rendering from Photographs (NPRP) began in earnest in the early 1990s with semi-automated paint systems [Haeberli 1990; Salisbury et al. 1994], and continued with media simulation [Cockshott et al. 1992; Litwinowicz 1997; Hertzmann 1998; Curtis et al. 1997; Brooks 2007]. These papers belong to a much larger body of work that sets out to answer one question: *what does a stroke look like?'* and can be used in many systems. Another relevant question is *where should a stroke be placed?* A natural answer is "at the edges of objects", which is given by Liwinowicz [1997] and many others. Some ask both questions, as when the trajectory of strokes are fashioned so as to lay along the edges of objects [Hertzmann 1998; Kang et al. 2007]. Others use a filtering process, for example edge aware filtering over high dynamic range images can produce impressive results [Paris et al. 2011]. This paper is not about media, strokes or filtering, we are more concerned with shapes of regions and the way these regions relate to one another.

Creating art using segmented images has distinct advantages for NPRP, at least if the segmentation yields regions with some kind of reasonable semantic interpretation. The value of image segmentation to NPRP is recognised in the work of DeCarlo and Santella, who construct a hierarchy based on eye fixations [DeCarlo and Santella 2002] and in that sense is interactive. Others build hierarchies automatically and put them to artistic use [Bangham et al. 2003]. Non-hierarchical segmentations can also be found in the painterly work of Gooch *et al* [2002], the stained-glass of Mould [2003], and the Cubism of Collomosse and Hall [2003]. Segmentations have produced excellent watercolors [Bousseau et al. 2006], coloured sketches [Wen et al. 2006], manga art [Qu et al. 2008]; stylized black and white images [Mould and Grant 2008; Xu and Kaplan 2008], and mosaics [Orchard and Kaplan 2008; Huang et al. 2011].

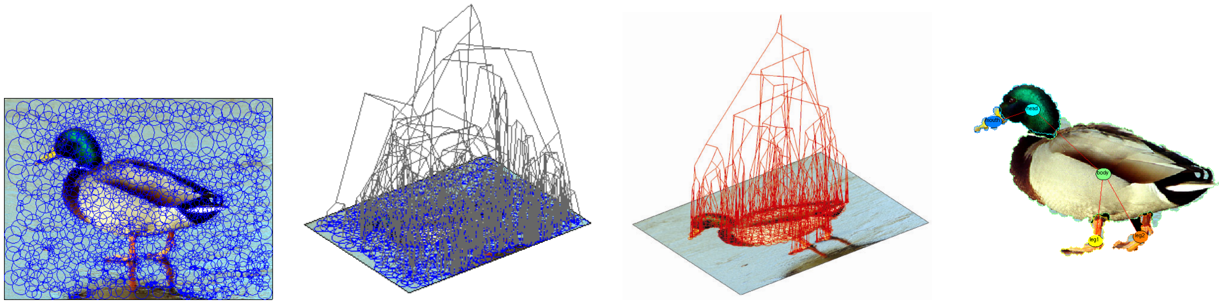
There is clear scope for introducing more abstract NPRP styles to the literature, and we are especially interested in styles used by Fine Artists such as Kandinsky, Mattise, Miró, and Picasso. These master artists often use simple geometric shapes and toplogical object structures in their artworks, and in this way they are similar to child art. In order to synthesise this type of art, we need to bring a much higher level of abstraction to images, such as representing objects via their structure, and abstracting image segments as pure geometric shapes. Shape abstraction has been used by Song *et al.* [2013], but structural abstraction (relation between parts) is new to NPRP, so far as we know. In short, rendering in the style of Simple Art is an interesting literature gap, one this paper sets out to address.

3 Method: From Photograph to Model, from Model to Painting

Image representation is the key to our aim. The model must be constructable from photographs, capture semantic structures, and versatile in terms of rendering options. We use a hierarchical segmentation as the base abstraction from which particular renderings arise. Although we explain how to create the hierarchy we use, we do so only for the sake of completeness – there is nothing special about our hierarchy and there are many alternatives in the Computer Vision literature to chose from.

There are two salient points about any hierarchical abstraction that is suitable for our purpose. The first is that it can be used to reproduce the underlying photograph to any desired degree of approximation, by including additional detail into the leaves of the hierarchical tree. Thus the nodes of the hierarchy close to root represent a large scale, coarse approximation, and ever greater detail is added close to the leaves. The second salient point is that the finer details are nested inside their parents. A typical hierarchy is shown in Figure 3.

We generate Simple Art from such hierarchy by visualising both the *structure* of the hierarchical tree and the *shapes* of the nodes. We traverse only a short way down the tree, because fine detail is not



(a) Primitive Features (b) Image Hierarchy (c) Object Hierarchy (d) Object Structure

Figure 3: *Extracting Object Structure*

important to our purpose. Here “visualising” means re-presenting the tree in some artistic form. Exactly which artistic form is a matter for the user, who is able to pick and chose styles appropriate the art they wish to create. We rely on prior literature to generate marks. What is of interest to us, and what is new here, is the fact we make use of structure in a direct way.

3.1 From Photograph to Model

There are many ways to produce a hierarchy from a photograph. The method explained below has the merits of simplicity and sufficient utility, other methods may fare better in tests against human produced ground truth at the cost of complexity. Since we are not advocating any one hierarchy but are including a description for completeness, we will keep our account brief.

Our method of model building has three main steps: (1) detect primitive features, (2) build these features into a hierarchical description, (3) parse the description into objects. The result is a hierarchical description of objects in an image that is of value in producing Simple Art. We now outline each step in turn.

3.1.1 Detect Primitive Features

We convolve an image $f(x)$ with Difference of Gaussian (DoG) filters of increasing scale, σ; with $x \in \Re^2$. This gives a scale-space signal at every pixel,

$$h(x, \sigma) = f(x) * (G(x|\sqrt{2}\sigma) - G(x|\sigma)), \tag{1}$$

in which $G(x|\theta) \propto \exp(-1/2|x|^2/\theta^2)$ is a Gaussian. Extrema in this signal, $h(x, \sigma)$ are known to correspond to features in the image [Mikolajczyk and Schmid 2004], and the scale at which the extrema occur give the size of the feature. We use the scale, σ of the first extremal value to fix the radius of a feature at a pixel, x. Since we determine a radius for every pixel we have, in effect, a large collection of overlapping discs.

We do not need all of these discs, we keep only the salient. Initially we discard any disc that is not wholly contained by the image. Next, a more sophisticated process picks only those which contain the rarest patterns, on the grounds that important features do not occur very often. For example, a disc placed at random on a picture is most likely to cover a homogeneous region, otherwise segmentation based on homogeneity would not work. The disc is less likely to cover edges, and complex patterns such as eyes would be covered even less often.

The association of rarity and saliency has been used before in NPRP [Collomosse and Hall 2003], we adapt their idea to our context. First we make a feature vector for each disc, then keep only the rarest features. For example, SIFT descriptors [Lowe 2004] give a feature vector of fixed length. A single eigenmodel fitted to their distribution in feature space is sufficient to determine the probability. This allows discs to be thrown away using a greedy algorithm. The least probable (most rare) disc is picked and kept, and all discs with centres inside this one are discarded; then the next least probable is kept. Now iterate until the image is covered or no discs remain. Figure 3 shows an example output.

3.1.2 Build a Hierarchical Description

Once in possession of a manageable sized collection of leaf-primitives we continue by merging them. We define two regions to be neighbours if they overlap by at least one pixel. The average colour is stored with each primitive. We merge every pair of neighbours (i, j) using a measure shown to minimize error in color [Haris et al. 1998]:

$$e(i, j) = \frac{N_i N_j}{(N_i + N_j)^2} |\mu_i - \mu_j|^2 \tag{2}$$

in which μ_i, μ_j are mean colours; and N_i, N_j are the number of pixels; in regions i and j, respectively. The average colour is updated for the new regions. This continues recursively until just one region remains at the top of the binary tree.

The tree is an image description. Every node corresponds to a continuous region in the image. Some of these regions (nodes) are meaningful objects, other regions (nodes closer to the leaves) are object parts. So objects — and, importantly, their structure — are embedded in the tree, the next step is to parse the tree into objects.

3.1.3 Parsing into Objects

Each node in the tree potentially expands into a semantic objects, or into an object part. Computer Vision allows for the automatic detection of some objects, in particular those for which a model of some kind has been constructed (for example as a histogram of visual words; as in the the so-called bag-of-words family of classifiers). However, the general case remains an open problem, therefore we allow a user to identify objects and salient parts (such as the head of a person) with a few mouse clicks, typically less than five.

User interaction means that false hierarchies can be amended, and that user preference is made possible, as people tend to interpret objects in different ways. The user clicks to locate an object against its background, and to identify a few of its major parts. Since the underlying image description is a graph these clicks are sufficient to identify complete subgraphs, which can be filtered using a graph theoretic measure (graph energy) to automatically break

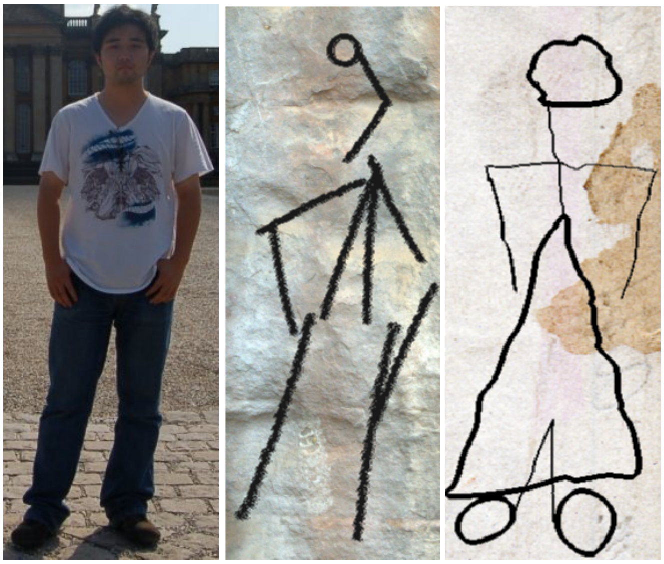


Figure 4: *Left, a source photograph. Middle, a stick-figure that visualises the arcs of the hierarchical model. Right, the same model rendered with shape simplification on nodes.*

objects into parts. The graph model can be filtered using Laplacian graph energy [Song et al. 2010], which can reduce its complexity by an order of magnitude with no loss of expressive power.

The result is a collection of objects, each modelled with a hierarchy, as in Figure 3. Hierarchies are of value for many problems in Computer Vision, including recognition and tracking, but we use them in a novel way – the purpose of making Simple Art.

3.2 From Model to Artwork

Now that we have a hierarchical model of objects that can be expressed as a graph of nodes and arcs, $G = (N, A)$. Our purpose is to visualise this graph in a manner appropriate to some chosen style, in this case as one the variants covered by Simple Art. As mentioned above, the term *Simple Art* is one we use in this paper as a short-hand for several groups of artwork. In this paper we focus on people and animals by (i) simplifying the shape of objects and objects parts, and (ii) emphasising connections between the parts of an object. These elements of abstraction map conveniently onto the nodes and arcs of a graph. Taken to an extreme, both shapes and connections are represented only by lines, as shown by the stick-man in Figure 4. In this example of Simple Art, the only shape not draw as a point is the head, which is drawn as a circle; otherwise the artwork is composed entirely of visualisations of arcs.

3.2.1 Depicting Structure

The ability to depict stick figures in an effective way relies on an ability to detect salient nodes within a hierarchical model, and then to draw lines to represent arcs connecting parent to child. Salient node detection is possible using the same mechanism as hierarchical simplification, which is to say via analysis of the Laplacian graph energy. Song *et al.* [2010] provides a detailed account, but briefly local minima in an energy function that indicates local graph complexity is used to signal salient nodes. Salient nodes tend to be parents that connect to children arranged in some regular form; a regular polygon, for example.

Once the salient nodes in an object have been identified it is relatively straight forward to depict structure. The easiest method of all is to draw straight lines that connect the centroids of parent/child nodes. It is only marginally more complicated to jitter the centroids using (for example) Gaussian noise. In Figure 4 we jitter both ends of each line, independently of all other lines so as to create a more broken appearance.

In principle we could continue depicting arcs in ever more sophisticated ways, perhaps to emulate the scribbled and over-drawn lines that so often inhabit children's drawings. However, such an emulation is beyond the scope even of Picasso (hence his famous quote), and since depicting connection is only one of the two characteristics of Simple Art, we now turn to the second: depicting shape.

Figure 5: *Fitting Shapes to Object Parts: robust convex hull, ellipses, and rectangles. (Source photograph taken from Berkeley's publicly available database.)*

3.2.2 Depicting Shape

The tendency to depict objects using just a few shape primitives such as circles (ellipses), squares (rectangles), and triangles is witnessed in the art of Picasso, Miró, Kadinsky, Matisse and many other important Western artists. It is also observed in the art of non-Western traditions as well as older forms of art in the West, including Greek, Egyptian, and cave at. Children too tend to draw and paint using just a few shape primitives. We wish to follow the example set by all of these artists.

Our aim is to classify the shape of nodes in our hierarchical description, based on a method described in [Song et al. 2013], where a more detailed description can be found. The central idea is to use a library of canonical shape primitives, the canonical shape that is the best fit to a binary mask of a node is used to depict that node. The remainder of this section briefly outlines the operation of this classifier.

We begin by obtaining a binary mask for a node via the union of discs that cover its features, as discovered in Section 3.1.1. Next we fit each canonical shape to that mask. Finally we select the optimal shape based on error measurements; that is we classify the shape of the mask. If no canonical shape makes a good fit we use a robust convex hull to overcome problems caused by spiky protrusions and large indentations [Rosin and Mumford 2006]. Figure 5 offers examples of shapes fitted to a decomposed face.

Each canonical shape has its own particular fitting method, to ensure as snug a fit as possible. We fit ellipses, rectangles, and triangles using methods described by Voss and Süße [1997]. Once each shape has been fitted we choose the best amongst them. It is not sufficient simply to find the shape of smallest error, because examples can be constructed for which the best shape model, as judged by a human, has larger error; the situation is analogous to using RMS error to measure the quality of a decompressed image. So we decide which particular shape is optimal via a classifier that uses statistics of the error distribution: mean, deviation, skew, and kurtosis. The classifier we use is a C4.5 decision tree [Quinlan 1993]. This

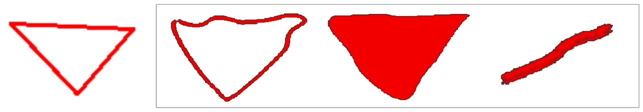


Figure 6: *A pure geometric shape and its three rendering styles.*

classifier partitions a data set (here a collection of 4-dimensional vectors, each describing the fitting error using statistical moments) into smaller sets, then again and so on. The most discriminative terms are at the top of the tree.

We trained our classifier using (errors from) training shapes. Each training shape was labeled, so training was supervised. We tested the quality of the classifier by asking humans to agree or not with the classifications on a test set. This is a standard approach to training, giving a confusion matrix indicating the likelihood that a test shape will be correctly classified. We repeat this several times, stopping when the confusion matrix converged to a steady state. The result is a classifier that agrees with human classification of shapes about 90% of the time.

Shape classification allows us to depict nodes in the hierarchy in a way commensurate with Simple Art. It is possible to produce artwork on the basis of shape classification alone [Song et al. 2013], but here we use a combination of simple shapes to depict nodes of a graph and lines to depict arcs, as in Figure 4

3.3 Rendering Control

Rendering controlled by the user, who decides the overall style. The user sets parameters such as colour and level of noise for stochastic processes that (*e.g.*) "wobble" lines are they are drawn. They also decide the broad style of the rendering, such as to emulate a specific artist, or to output art with more child-like qualities, as explained next.

In accordance with the main characteristics observed from many of the abstract paintings from Miro (see "Birds, 1973 " and "Day Break, 1968" for example), we offer three different ways to render nodes in an abstract object model: (i) distorted contours of uniform colours, (ii) distorted regions of uniform colours; (iii) long and curly strokes. Figure 6 illustrates the three different styles a node can be rendered into. Each node can be rendered separately and later composited into the picture (we traverse the tree from root to leaf). We will now describe how shapes can be rendered into the above three different styles.

We "wobble" shape edges using the Flash and Hogan line model [Flash and Hogan 1985] which offers "the smoothest motion to bring the hand from an initial position to the final position in a given time" and has been successfully applied in creating realistic pencil lines [AlMeraj et al. 2009]. We then add Gaussian noise to sparsely sampled coordinates along the line to form the final trajectory. The trajectory is a suitable basis for many stroke-emulation techniques.

The user may choose to render regions not as a shape but as a line. In this case we simply apply a medial axis transform to a shape then render its longest axis using the line model above.

4 Results

Now that we can model objects as a hierarchy of shapes it is possible to render them in many different ways. The Figures in this paper demonstrate a range of possibilities we have used, which should not be regarded as a limiting set but as examples we happen to prefer. All of the source photographs used at the author's website.

One of the images in our teaser, Figure 1, shows a horse. In this case rendering is intended to represent a rock carving. The same source was used to make a cave painting and also an ink-wash drawing, both in Figure 7. We have also animated this horse in a style that flickers, see the supplementary material. Such flicker was a deliberate choice because we wanted to emulate the style of a particular TV show aimed at children ("Rhubarb and Custard") that we enjoy.

Joan Miró is one of our favourite artists, and has influenced two of the examples we show here. These are the bison in the teaser (Figure 1) and a piece that is a compound of several objects, seen in Figure 8.

We have attempted to emulate children's art too, as in the skater-turned-ball-player in the teaser, the duck in Figure 7, the drawings in Figure 4, and a second skater in Figure 7. That same second skater is also used in the large Miró emulation, Figure 8, which also includes a cat and an eagle. The bearded man in Figure 7 used the face portrait as a source in Figure 5. It is not intended to be any particular style, but is something we constructed for our own pleasure.

5 Discussion: Appreciating the Results

A question often raised with regard to aesthetic rendering, and NPR more generally, is "how should the work be assessed?" There is no single answer to this question. Isenberg points out that NPR is produced for different purpose, such as scientific visualisation, and should therefore be assessed with proper regard paid to its intend purpose [Isenberg 2013]. We are operating in the domain of Aesthetic Rendering where calls for user studies or Turing test experiments are not uncommon. We believe neither of these a suitable and so chose to use the assessment criteria laid out by Hall (Computer Scientist) and Lehmann (Cultural Historian) [2013] who suggest six norms be used.

Norm 1: The aesthetic quality of any NPR artwork is to be assessed in relation to all artwork, including human art. In general, the aesthetics of a particular work is not open to user studies – aesthetics is not democratic. Equally, aesthetics should not be subject to the Turing test, because knowing how a piece is made is often a component of aesthetic value. Nor is aesthetics a test of beauty as in "looking good", especially when an artwork references the real-world – images of war may be grotesque yet have high aesthetic value. In our case there are no such references, which necessarily limits aesthetic quality to "looking good". In this case we appeal to Shakespeare's observation that beauty lies in the eye of the beholder, so leave judgement to the reader.

Norm 2: It is common for NPR authors to claim their output emulates a school or an artist; the validity of this claim must be tested against real world examples. The claim here is that our output belongs to the class of Simple Art. If we accept that Simple Art is characterised by the role of connectivity between an object's parts, then we have met the claim – but of course run the risk of falling foul of the fallacy of stipulation. However, we can avoid that trap by discussing Simple Art a little more.

There are many examples of Simple Art that we have not tried produce, the Art of aborigines in Australia, and traffic road-signs are just some instances amongst many. It was possible to select only a few representative instances, but the lack of these other examples limits our claim nonetheless. In particular there are cases in Simple Art in which connectivity is not rendered – as when a head is disconnected from its body. This does not mean connectivity it is not important; without it a viewer would not associate the two parts, it just means that the connection has not been depicted. We have shown no such examples.

Also, connectivity is important to art forms that cannot be classified as Simple Art. Dali, for example exploits connectivity (more exactly, implied rather than rendered connections) in paintings such as *Metamorphosis of Narcissus* to deliberately create visual ambiguities and metaphors. Such a level of sophistication is well beyond this paper. Some art of Bridget Riley depends on the relationship between parts, her dot paintings comprise only isolated black circles on a white background. Dodgson [2008] gives an interesting account of Riley's art from a Computer Science point of view.

Overall though, the examples we show do bear at least passing resemblance to their real art counterparts. It is true that when people first see the child's drawing in the teaser (Figure 1) they are ready to believe it is a genuine children's drawing: they show surprise when told it is computer generated. This is anecdotal evidence, but it is nonetheless evidence in favour of our Simple Art being successful as art. It is tempered by the fact that our result and real child art have never been compared side-by-side. We invite the reader to compare our Miró style results to the real thing.

Norm 3: Media emulation is needed to more completely approximate any school or style. In this case we rely on the work of others to emulate media and strokes. We found no algorithms capable of scribbling as a small child, but equally no adult human than can emulate a child either; even Picasso struggled. We therefore we use Norm 2 to claim our contribution is a first approximation to a genre previously unavailable to NPRP.

Norm 4: Non-photorealistic rendering that depends wholly on humans – paint-boxes – must be assessed against the background of all human produced art. A fully automated system that uses photographs as source is impossible given today's state of the art. Therefore we opted for a semi-automatic approach and it is therefore relevant to compare and contrast our results with other NPR works that also employ semi-automatic methods. Readers are the best placed to do that, as in Norm 1.

Norm 5: The elegance of the underlying algorithms and system design is important for computer based art; this is comparable to the interest Cultural Historians show in the manner of production. We have provided a simple semi-automatic algorithm that leads to a model that explicitly denotes connectivity, and is versatile in that it can be rendered in many different ways. Rendering methods include shape classification, and drawing arcs as lines of some kind. The elegance of our approach is an aesthetic matter that we must leave to the judgement of the reader.

Norm 6: The novelty of any research is an essential ingredient and so is of obvious interest here. Given we know of no other NPR work aimed at producing Simple Art, we claim novelty.

6 Conclusion

Our results demonstrate that it is possible to semi-automatically process photographs into Simple Art. The method depends on an object model that is easy for users to construct and render. Construction requires just a few mouse clicks to identify an object and some of it parts, rendering means to set parameters such a as colour and noise level. Thus we have contributed to an identified literature gap by emulating child art, cave art, and Miró from photographs.

The key technical lessons are (i) that hierarchical descriptions are powerful, and (ii) that (some) Simple Art requires rendering at least structure and – preferably – abstracted shape too. The first lesson is only to confirm again what is widely accepted, the second is new so far as we know. Others have abstracted shape [Song et al. 2013; Huang et al. 2011], but rendering using graph arcs is new to NPR so far as we know (although not to Scientific Visualisation).

Whether our model supports the production of other forms of Simple Art is open; one avenue for future work. Perhaps more interesting would be classes of real world Simple Art that cannot be described using a hierarchy. A hierarchy will struggle with art such as that of Bridget Riley, so our model will not easily extend to all forms of art.

In conclusion: (i) rendering the arcs of hierarchy is the key technical novelty needed to support the production of Simple Art. This is true even if particular models used for rendering were crafted wholly by a human, rather than semi-automatically from a photograph. (ii) The results are reasonable imitations of some classes of Simple Art, the rendering style is subject to user control so that different classes can be emulated. However any art that cannot be described as a hierarchy is out of reach. (iii) The many open questions regarding modelling art work, photographs, and both at once make interesting future work.

References

ALMERAJ, Z., WYVILL, B., ISENBERG, T., GOOCH, A. A., AND GUY, R. 2009. Automatically mimicking unique hand-drawn pencil lines. *Computers & Graphics 33*, 4, 496–508.

BANGHAM, J., GIBSON, S., AND HARVEY, R. 2003. The art of scale-space. In *British Machine Vision Association*, 569–578.

BOUSSEAU, A., KAPLAN, M., THOLLOT, J., AND SILLION, F. X. 2006. Interactive watercolor rendering with temporal coherence and abstraction. In *NPAR*, 141–149.

BROOKS, S. 2007. Mixed media painting and portraiture. *IEEE TVCG 13*, 5, 1041–1054.

COCKSHOTT, T., PATTERSON, J., AND ENGLAND, D. 1992. Modelling the texture of paint. *Computer Graphics Forum 11*, 3, 217–226.

COLLOMOSSE, J. P., AND HALL, P. M. 2003. Cubist style rendering from photographs. *IEEE TVCG 4*, 9, 443–453.

CURTIS, C., ANDERSON, S., SEIMS, J., FLEISCHER, K., AND SALESIN, D. H. 1997. Computer-generated watercolor. In *ACM SIGGRAPH*, 421–430.

DECARLO, D., AND SANTELLA, A. 2002. Stylization and abstraction of photographs. *ACM Trans. Graph. 21* (July), 769–776.

DODGSON, N. A. 2008. Regularity and randomness in bridget riley's early op art. In *Proceedings of the Fourth Eurographics conference on Computational Aesthetics in Graphics, Visualization and Imaging*, Eurographics Association, 107–114.

FLASH, T., AND HOGAN, N. 1985. The coordination of arm movements: an experimentally confirmed mathematical model. *The journal of Neuroscience 5*, 7, 1688–1703.

GOOCH, B., COOMBE, G., AND SHIRLEY, P. 2002. Artistic vision: Painterly rendering using computer vision techniques. In *NPAR*, 83–90.

HAEBERLI, P. 1990. Paint by numbers: abstract image representations. In *ACM SIGGRAPH*, vol. 4, 207–214.

HALL, P., AND LEHMANN, A.-S. 2013. Dont measure – appreciate! NPR seen through the prism of art history. In *Image and Video-Based Artistic Stylisation*. Springer, 333–351.

HERTZMANN, A. 1998. Painterly rendering with curved brush strokes of multiple sizes. In *ACM SIGGRAPH*, 453–460.

HUANG, H., ZHANG, L., AND ZHANG, H.-C. 2011. Arcimboldo-like collage using internet images. In *ACM Transactions on Graphics (TOG)*, vol. 30, ACM, 155.

ISENBERG, T. 2013. Evaluating and validating non-photorealistic and illustrative rendering. In *Image and Video-Based Artistic Stylisation*. Springer, 311–331.

KANG, H., LEE, S., AND CHUI, C. K. 2007. Coherent line drawing. In *NPAR*, 43–50.

HARIS K. AND EFSTRATIADIS, S. AND N. MAGLAVERAS, AND A.KATSAGGELOS. 1998. Hybrid image segmentation using watersheds and fast region merging. *IEEE Trans. Image Processing 7*, 12, 1684–1699.

LITWINOWICZ, P. 1997. Processing images and video for an imp ressionist effect. In *ACM SIGGRAPH*, 407–414.

LOWE, D. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision 1*, 91–110.

MIKOLAJCZYK, K., AND SCHMID, C. 2004. Scale & affine invariant interest point detectors. *Intl. J. of Comp. Vis. 60*, 1, 63–86.

MOULD, D., AND GRANT, K. 2008. Stylized black and white images from photographs. In *NPAR*, 49–58.

MOULD, D. 2003. A stained glass image filter. In *Eurographics workshop on Rendering*, 20–25.

ORCHARD, J., AND KAPLAN, C. 2008. Cut-out image mosaics. In *Non-photorealistic Animation and Rendering*, 79–87.

PARIS, S., HASINOFF, S. W., AND KAUTZ, J. 2011. Local laplacian filters: edge-aware image processing with a laplacian pyramid. In *ACM SIGGRAPH 2011 papers*, ACM, New York, NY, USA, SIGGRAPH '11, 68:1–68:12.

QU, Y., PANG, W.-M., WONG, T.-T., AND HENG, P.-A. 2008. Richness-preserving manga screening. In *SIGGRAPH Asia*, 1–8.

QUINLAN, J. 1993. *C4.5: Programs for Machine Learning*. Morgan Kauffman.

ROSIN, P., AND MUMFORD, C. 2006. A symmetric convexity measure. *Computer Vision and Image Understanding 103*, 2, 101–111.

SALISBURY, M. P., ANDERSON, S. E., BARZEL, R., AND SALESIN, D. H. 1994. Interactive pen-and-ink illustration. In *ACM SIGGRAPH*, 101–108.

SONG, Y.-Z., ARBALAEZ, P., HALL, P., AND LI, C. 2010. Finding semantic structures in image hierarchies using laplacian graph energy. In *European Conference on Computer Vision*, 694–707.

SONG, Y., PICKUP, D., LI, C., ROSIN, P., AND HALL, P. 2013. Abstract art by shape classification. *Transactions on Visualization and Computer Graphics*.

VOSS, K., AND SÜSSE, H. 1997. Invariant fitting of planar objects by primitives. *IEEE TPAMI 19*, 1, 80–84.

WEN, F., LUAN, Q., LIANG, L., XU, Y.-Q., AND SHUM, H.-Y. 2006. Color sketch generation. In *NPAR*, 47–54.

XU, J., AND KAPLAN, C. S. 2008. Artistic thresholding. In *NPAR*, 39–47.

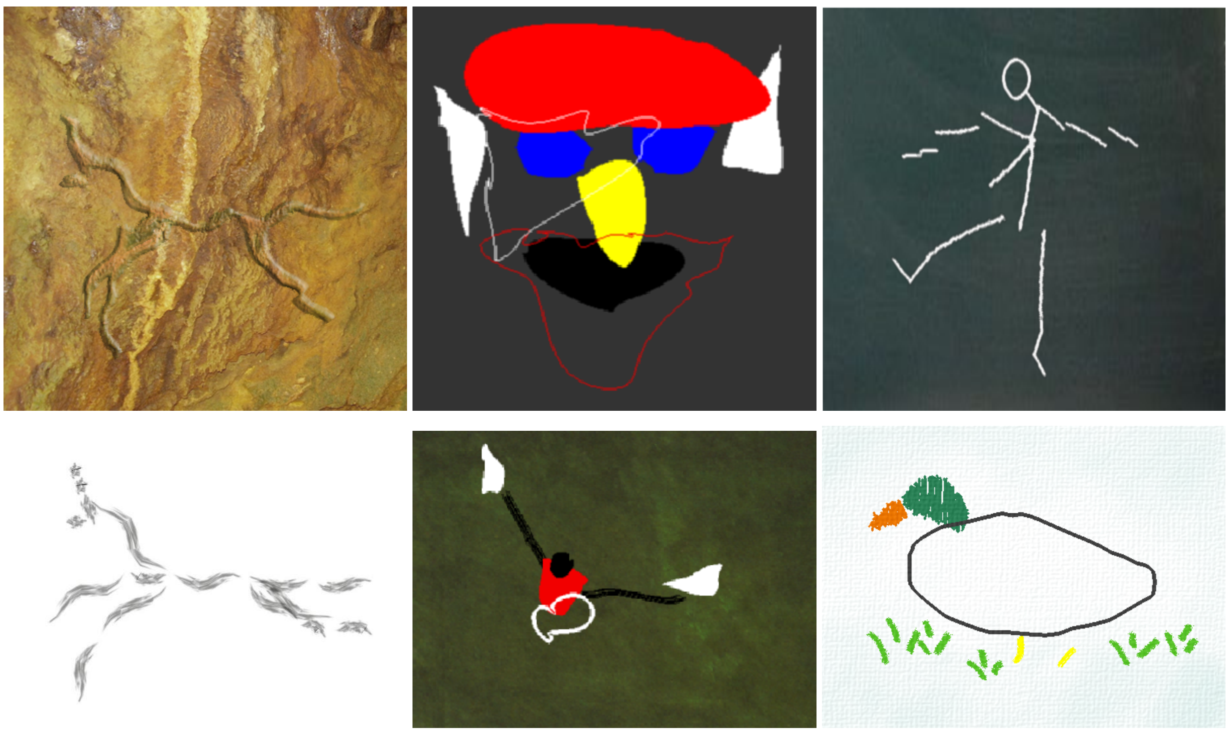


Figure 7: *A Variety of Examples: cave horse; bearded man; chalk skater; washed horse; flying eagle; duck. See the author's website for source images.*

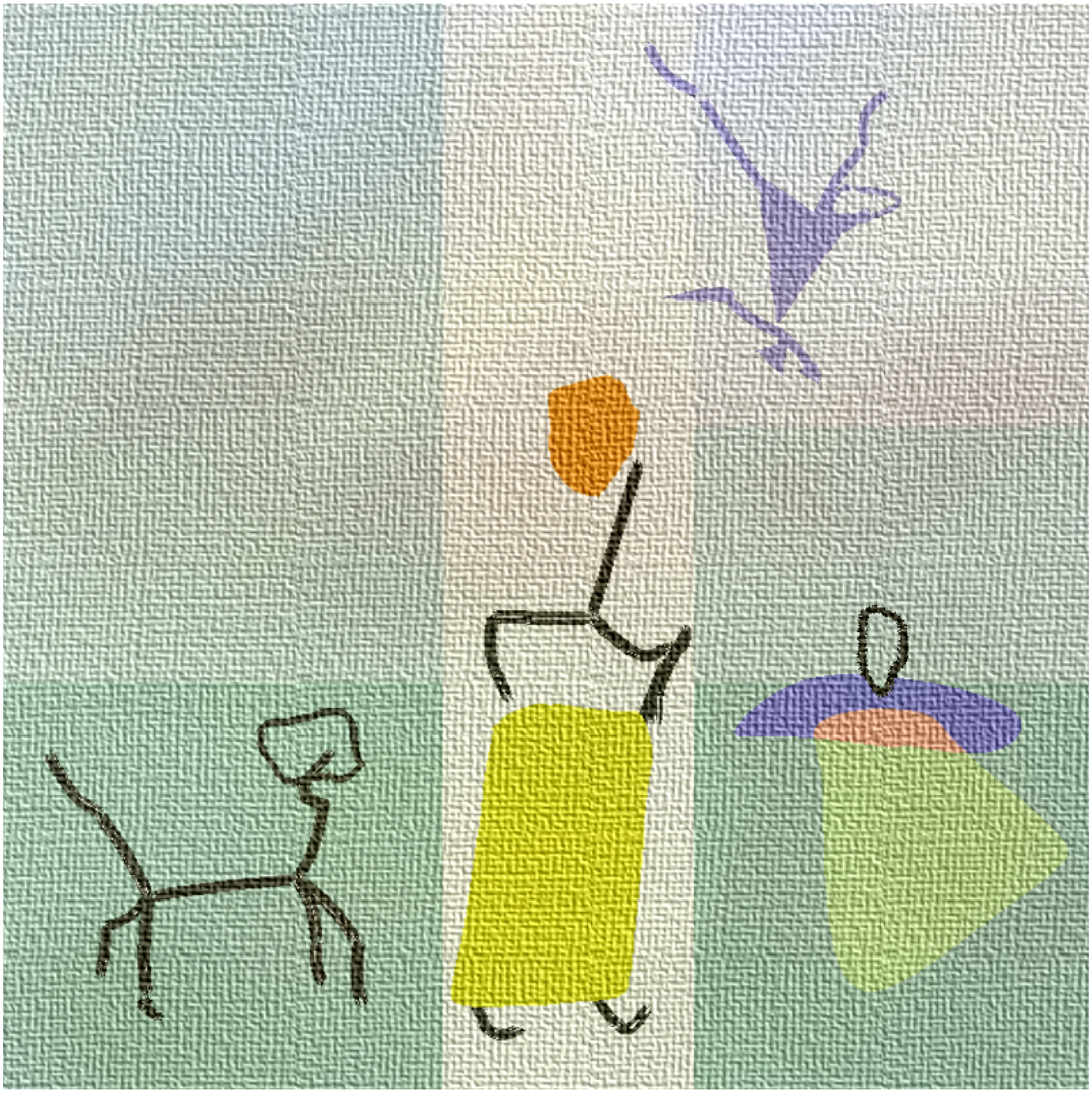


Figure 8: *An imitation Miró comprising several models constructed from the photographs of a cat, a bird, an ice skater, and a man. See the author's website for source images.*