

# Watercolour Rendering of Portraits

Paul L. Rosin and Yu-Kun Lai

Cardiff University, Cardiff, CF24 3AA, UK,  
Paul.Rosin@cs.cf.ac.uk, Yukun.Lai@cs.cardiff.ac.uk

**Abstract.** Applying non-photorealistic rendering techniques to stylise portraits needs to be done with care, as facial artifacts are particularly disagreeable. This paper describes a technique for watercolour rendering that uses a facial model to preserve distinctive facial characteristics and reduce unpleasing distortions of the face, while maintaining abstraction and stylisation of the overall image, employing stylistic elements of watercolour such as edge darkening, wobbling, glazing and diffusion.

**Keywords:** non-photorealistic rendering, watercolour, portraits



**Fig. 1.** An example of our watercolour portrait rendering. Important facial details have been preserved while other areas are more stylised.

## 1 Introduction

The goal of image-based non-photorealistic rendering (NPR) is to take an image and re-render it in an alternative style, generally for artistic effect. Target styles include natural media (oil painting, watercolour, pen and ink, etc.) and can also be categorised according to traditional artistic movements (impressionist, expressionist, pointillist, etc.). This paper describes an approach to NPR stylisation of portrait images using a watercolour effect. Watercolour provides an attractive abstraction, and has been used in several applications of NPR, such as the rendering of CAD models of interior decorating plans to provide a more aesthetic rendering [18] and for rendering 3D models for augmented reality [6], enabling a clear separation of original and inserted content.

Portraits are an important and long standing branch of art, which normally focuses on producing a representation of the sitter’s face. However, good portrait paintings or photographs go beyond capturing just a realistic representation, but attempt to express or reveal aspects of the nature of the sitter that may not be otherwise directly visible in real life. Current NPR algorithms do not aim to automatically modify the stylisation according to the personality of the sitter. Nevertheless, there remain challenges for performing stylisation. For instance, if a substantial amount of abstraction is desired, then the identity of the sitter is liable to become obscured as discriminating features are removed. In addition, the human visual system is particularly sensitive to artefacts when viewing faces.

The contributions of this work are to provide a watercolour rendering method that is targeted at portraits. In order to preserve distinctive facial characteristics a facial model is used, while a salience map controls the degree of abstraction elsewhere in the image. In addition, techniques for stylistic elements of watercolour such as wobbling, glazing and diffusion are introduced. Results are shown on images from the NPRportrait benchmark [22].

## 2 Related Work

**NPR portraits:** Given the importance of the portrait in art it is not surprising that there has been a substantial body of NPR work dealing specifically with portraits, and we describe some recent examples. Berger *et al.* [3] learned models of stroke parameters from training examples which consisted of line sketches at four abstraction levels that were provided by seven artists. The method by Rosin and Lai [21] first stylises the image with abstracted regions of flat colours plus black and white lines [15]. Facial features are localised in the input image. In the skin region of the face, shading is stylised, and reduced line rendering is applied to reduce visual clutter. In addition, the facial model is used to enhance the rendering of the eyes, lips, teeth and highlights. Zhang *et al.* [27] describe an approach to creating cartoon versions of faces. Facial components are detected in the input image, and are matched to a dictionary of cartoon stylised components. Support Vector Regression is used to train a composition model to ensure good layout of the selected components. Recent work [11] has applied style transfer to videos of faces, and uses a semantic segmentation guide to control the synthesis. In an attempt to analyse the performance of NPR portrait algorithms, Rosin *et al.* [22] recently introduced a benchmark data set of 40 images, split into two levels. The first level is constrained to contain frontal views of adult faces with neutral expressions, and uncluttered background. The second level is more challenging to algorithms, as it allows facial hair, a wider variety of facial expressions, more varied backgrounds, etc. They applied six stylisation algorithms, including three portrait-specific algorithms, to the benchmark, and demonstrated the benefits of using portrait-specific algorithms over general-purpose NPR algorithms. Namely, the former tended to use domain-specific information to preserve key details such as eyes and conversely to improve abstraction by removing semantically unnecessary details.

**Watercolour rendering:** Watercolour painting allows for many effects that artists use to good effect, such as dry-brush, edge darkening, backruns, granulation, flow patterns and colour glazing. Several interactive painting tools have been developed for creating watercolour renderings [7,8] that attempt to capture these effects. They use stroke based rendering along with physical models of the painting process. Such particle based models can effectively describe the flow of water and pigment on the paper surface, taking into account many factors such as the water velocity, viscosity, drag, brush pressure, pigment concentration and paper saturation. Of more relevance to our work is the simple and efficient filter based approach of Bousseau *et al.* [5]. Dispensing with the physical model they use instead a series of image processing operations that can also mimic some of the watercolour effects. Wang *et al.* [25] extended the filter based approach in a number of ways, incorporating saliency to control level of detail, colour adjustment based on training examples, and extended the range of watercolour effects. Another approach to stylisation is to use general style transfer methods which can then be applied for the specific instance of watercolour. A pioneering example was Image Analogies [13], which effectively learns a mapping from a pair of training images containing one scene in two styles, and that style transform is then applied to an independent test image. In the last few years deep learning approaches to style transfer have become popular [12], and have the advantage that only a single style image is required rather than a corresponding pair (which is generally not convenient and may not be possible). Yan *et al.* [26] describe a method for automatic photo adjustment using CNNs, and apply it (amongst other things) to generate a watercolour style based on training examples provided by a professional photographer using Photoshop. However, the training and test images are not truly watercolour style. A recent work attempts to combine both the Image Analogies and deep learning approaches [17], but is computationally expensive. Such example based approaches have some limitations. First, errors can occur if local textures are not distinctive. Second, transfer is based on a combination of style and content, and this can lead to undesirable content from the style image being transferred to the content image.

### 3 Our Approach

We follow the basic elements of Bousseau *et al.*'s [5] filter based approach, although the details are not identical. Following Bousseau *et al.* [5] our approach is not to produce a physically accurate watercolour simulation, but rather capture the important stylistic elements. Abstraction is performed by the following set of filtering steps: 1/ apply a small amount of Gaussian smoothing, 2/ apply a larger amount of median filtering, 3/ apply morphological filtering (opening and closing), 4/ apply a small amount of Gaussian smoothing. We prefer to avoid relying on accurate segmentation, as this is potentially unreliable.

The edge darkening effect (in which pigment migrates towards edges of areas of paint) is simulated by extracting edges from the abstracted image using the Sobel operator. Although not detailed in [5], [4] specifies that the edge map

is inverted and scaled, and then combined with the abstracted image using an overlay blend operation:

$$f(a, b) = \begin{cases} 2ab & \text{if } a < 0.5 \\ 1 - 2(1 - a)(1 - b) & \text{otherwise} \end{cases}$$

where  $a$  is the base layer value (the abstracted image) and  $b$  is the top layer value (the modified Sobel image). This combines the images so that the result is darkened where the base layer is dark and lightened where the base layer is light. The benefit of using the overlay blend rather than a multiply blend is that the former darkens the edges while preserving their colour, whereas the latter uniformly darkens the three colour channels, and so the darkening tends towards gray, which is less attractive.

Next, wobbling (distortion due to the paper’s granularity) is added (see section 3.2), and finally, paper and pigment textures are added using overlay blends.

### 3.1 Image Brightening

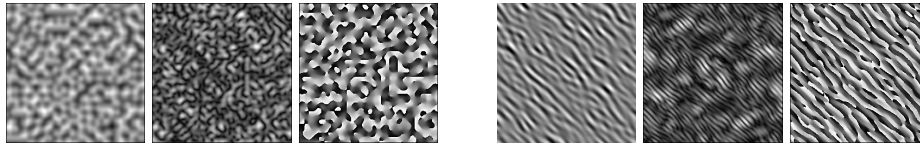
The image is brightened to make it more vibrant looking. This can go beyond scaling the intensities to fill the dynamic range, since blown out highlights are compatible with the watercolour style. However, this needs to be done in moderation, and therefore the scaling should be determined carefully. Moreover, for portraits the scaling should depend on the face rather than the background.

First, a morphological closing is performed on the image to remove small and medium sized bright spots from consideration. This stage controls the amount of blown out highlights that are created. A mask of the main facial area is created using the facial landmarks extracted from the image by OpenFace [2]. Content outside the face is removed according to the facial mask, and the maximum intensity  $I_{\max}$  in this version of the image is determined.

Two brightening operations are considered. The first applies to each pixel a global scaling of  $\min(1/I_{\max}, 1.5)$ ; the upper limit of 1.5 is set as we assume that images are taken under reasonably normal lighting conditions, and this avoids over-stretching that would otherwise occasionally occur due to dark skins. Clipping is applied to any overflowing values that may occur in the individual RGB colour channels; consequently, a colour shift may occur. The second operation reduces the scaling factor at each pixel, if necessary, so as to prevent any colour channel overflowing. While avoiding colour shift, the result shows a few blown out highlights, and moreover, can cause intensity variations to be removed, resulting in flat unattractive regions. We incorporate both approaches by taking a linear combination of the two colours  $\mathbf{c}_1$  and  $\mathbf{c}_2$ ,  $\alpha \mathbf{c}_1 + (1 - \alpha)\mathbf{c}_2$  with the weight  $\alpha = 0.75$ , which achieves a good compromise in which a reasonable amount of blown out highlights are created while colour shift is reduced, and also faintly retains the original colours in blown out areas.

### 3.2 Wobbling

One of the watercolour effects is wobbling, which is the distortion due to the paper’s granularity. Bousseau *et al.* [5] use the horizontal and vertical gradients



**Fig. 2.** Noise functions: Perlin noise (left block), Gabor noise (right block). For each block, left: one of the 2D noise maps used to generate the offset vector; middle: magnitude of offset vector; right: phase of offset vector.

of the paper to provide offsets in the rendered image, but do not describe their approach in detail. Doran and Hughes [10] found that a similar approach produced poor results, and therefore applied morphological smoothing to make the results more coherent. The jittering and smoothing was applied iteratively.

However, a limitation of the above approach is that it is difficult to control the wobble effect. Instead, to create wobbles we use spatially coherent noise. Given a 2D coherent noise function  $n(\mathbf{x})$  producing a single noise value at location  $\mathbf{x} = [x, y]$  we combine two instances to form an offset vector  $\mathbf{t} = [n_1(\mathbf{x}), n_2(\mathbf{x})]$  which is used to produce the warped image  $W$  from the unwarped image  $U$ :  $W[\mathbf{x}] = U[\mathbf{x} + \mathbf{t}]$ .

This is a general approach that has two main benefits. First, there are a range of types of coherent noise that have been developed for generating procedural textures, for instance the well known Perlin noise [19] and also more recent examples such as Gabor noise [14]. Second, such noise functions can be easily controlled, with parameters for frequency, magnitude, orientation, etc. It is consequently straightforward to manipulate these values to provide more extreme effects that go beyond watercolour simulation in order to generate highly stylised renderings. Wang *et al.* [25] used the Perlin noise approach in their watercolour system, but applied it to simulate hand tremor rather than paper roughness. Bousseau *et al.* [5] incorporated Perlin noise into their 3D model rendering pipeline, but not their 2D image-based pipeline. Figure 2 shows examples of both Perlin and Gabor noise and their application to create offset vectors  $\mathbf{t}$  and visualises their magnitude and phase.

### 3.3 Glazing

Colour glazing is a watercolour technique of adding thin washes of colour. It creates semi-transparent layers, which are then blended. We propose a simple image processing approach to simulate this. We first obtain a set of regions from the image using segmentation. For robustness, we use a state-of-the-art superpixel generation method [1], although alternative segmentation methods can also be used. Superpixels naturally involve oversegmentation, but with our method extra boundaries where regions on both sides have identical colour will remain unchanged during the process. For each region  $R_i, i = 1, 2, \dots, N_R$ , where  $N_R$

is the number of regions, we obtain an expanded region  $\tilde{R}_i$  using morphological dilation corresponding to the area affected by  $R_i$ , creating a layer with both solid and semi-transparent pixels. For each pixel  $p_k \in \tilde{R}_i$ , we work out its colour component value  $\tilde{v}_{k,i}$  (corresponding to the R, G, B channels), and weight  $\tilde{w}_{k,i}$  (where larger weight means more contribution in blending).  $\tilde{v}_{k,i}$  is defined using membrane interpolation by minimizing the following harmonic energy:

$$\min_{\tilde{v}_{k,i}} \int_{p_k \in \tilde{R}_i} \|\nabla \tilde{v}_{k,i}\|^2, \quad s.t. \quad \tilde{v}_{k,i} = v_k, \forall p_k \in R_i. \quad (1)$$

where  $v_k$  is the original pixel value for  $p_k$ .  $\nabla \cdot$  is the gradient operator, and in the image setting, is simply approximated using pixel differences. This can be easily solved by a sparse linear system. It keeps pixels in the region  $R_i$  unchanged, and propagates pixel values from the boundary of  $R_i$  to the pixels in the dilated region. To work out the weight  $\tilde{w}_{k,i}$ , denote by  $d_{k,i}^F$  the shortest Euclidean distance between the pixel  $p_k \in \tilde{R}_i$  and pixels in  $R_i$  (distance to “foreground”), and  $d_{k,i}^B$  the shortest Euclidean distance between the pixel  $p_k$  and pixels outside of  $\tilde{R}_i$  (distance to “background”), both of which are normalised to  $[0, 1]$  by linear scaling for pixels in  $\tilde{R}_i$ . The weight  $\tilde{w}_{k,i}$  is then defined as:

$$\tilde{w}_{k,i} = 1 - d_{k,i}^F / (d_{k,i}^F + d_{k,i}^B + \varepsilon) \quad (2)$$

which ensures that pixels in  $R_i$  have weight 1, and pixels on the boundary of  $\tilde{R}_i$  have weight 0.  $\varepsilon$  is a small value to avoid dividing by zero, and is set to  $10^{-5}$  in our experiments. The pixel value  $\bar{v}_k$  in the output image with glazing is simply obtained by the weighted sum of contributions of all the regions covering this pixel:

$$\bar{v}_k = \frac{\sum_{i, p_k \in \tilde{R}_i} \tilde{w}_{k,i} \tilde{v}_{k,i}}{\sum_{i, p_k \in \tilde{R}_i} \tilde{w}_{k,i}} \quad (3)$$

### 3.4 Pigment Diffusion

Diffusion of pigment tends to go from dark to light, and depending (amongst other factors) on the amount of water present,<sup>1</sup> a large amount of diffusion can be introduced for artistic effect. Many physically-based models and methods exist for simulating pigment diffusion [7]. Dong *et al.* [9] described a simple and efficient image-based approach for rendering Chinese ink paintings, which performed spattering (independent random movement of pixels) followed by median filtering. However, we did not find their approach effective for watercolour. Wang *et al.* [25] performed diffusion by first identifying regions that satisfy appropriate criteria on hue difference and intensity differences with their neighbouring regions; tighter thresholds are used for salient regions. Seeds are randomly scattered around each selected boundary on the light side only, and blurring is applied in the normal direction to the region boundary. A disadvantage is that

<sup>1</sup> For example, wet in wet is a watercolour technique in which paint is specifically applied to wet paper.

their approach requires accurate segmentation of both the image and saliency map. Also, their results consisting of blurred lines can look unnatural.

In this paper we only aim to generate small amounts of diffusion, and describe two simple methods for creating different effects. The first approach leverages the method for wobbling described in section 3.2 and creates a “blobby” effect. Let  $W_1$  be the watercolour image which has been stylised using the initial amount of wobbling, and  $W_2$  be a version of the watercolour image that is created with more wobbling that is roughly simulating diffusion. Then  $W'_2 = \min(W_1, W_2)$  limits the diffusion to go from dark to light. To prevent the diffusion component  $D = W'_2 - W_1$  appearing too distracting we apply a gamma correction to it in order to reduce strong densities whilst preserving weak densities, and the final result is given by  $W_1 + 0.5D^{0.7}$ .

The second approach produces a more uniform diffusion effect by applying erosion to the image to diffuse dark pigment. To avoid isolated dark groups of pixels having an undue effect they are removed first by applying an opening. As before, the diffusion component is extracted, and then a uniform blur is applied:  $D = \text{blur}(W_3 - W_1)$ , where  $W_3$  is the result after erosion. In order to limit the diffusion to go from dark to light,  $D$  is multiplied by the following weight map:  $\max(W_1 - \text{blur}(W_1), 0)$ . Before the diffusion  $D$  is added back to the watercolour image  $W_1$  an overlay blend with a texture map can be optionally applied to  $D$  to make it less regular and more natural looking.

Both these diffusion methods are controlled using the level of detail scheme described in the next subsection to avoid distracting effects in foreground regions.

### 3.5 Level of Detail

When applying NPR to portraits, the desire for significant abstraction in order to obtain distinctive stylisation is potentially at odds with the requirement that the portrait should remain recognisable and the need to avoid unpleasant artifacts on the face. For instance, given the sensitivity of the human visual system to facial artifacts it is probably unacceptable to render an eye such that it appears mishapen.

A means to satisfying the above goals is to incorporate several levels of detail, so that important aspects of the face are preserved at high detail, while less important elements in the image can be more highly abstracted. Recently several state-of-the-art portrait-specific NPR algorithms were applied to a benchmark set of portraits [22], and it was notable that all of them ensured that the eyes were the most carefully rendered parts of the image. In addition to performing rendering at a higher resolution for the eyes, other elements in the processing can ensure accuracy. For instance, in [20] morphological cleaning of lines was uniformly applied to remove clutter and noise. However, the size threshold for cleaning was reduced for the eye regions to avoid the possibility of removing critical detail. Another example is given in [21] where an artificial highlight was added to the eye in case the real highlight was absent due to extremely low contrast in the input image. In the more general area of photo enhancement special attention is also paid to eyes [23].

In this work we also incorporate differential processing of different elements in the image. We use three levels of detail which are controlled by saliency and face masks. The discriminative regional feature integration method (DRFI) [24] is used to provide the saliency map. DRFI segments the image into superpixels and uses a random forest regressor to map region features to saliency values. Since DRFI provides crisp region based saliency maps we also apply blurring so that the final rendering does not exhibit discontinuities between areas with different levels of detail. A sigmoid mapping is also applied to increase the separation in saliency values between foreground and background objects. For facial components we create our facial regions (eyes, eyebrows, nose, mouth) from the OpenFace landmarks. The regions are dilated to ensure that the regions are adequately covered (as the landmarks found by OpenFace may not be accurate), and blurred to provide soft masks.

In addition to rendering the eye area at higher resolution, they are further enhanced to ensure that they stand out. Contrast limited local histogram equalisation (CLLHE) is applied to the high resolution image before stylisation to enhance detail, increasing contrast in the eye regions. The original and equalised images are blended, with the weighting determined by the blurred eye mask.

A detailed description of the watercolour processing pipeline follows:

**Brightening** – Apply rescaling to the original input image (denoted by  $I_O$ ).

**Levels of detail** –  $I_O$  is resized to double and half size versions ( $I_D$  and  $I_H$ ).

**Eye enhancement** – Apply CLLHE to enhance detail in the eye regions of  $I_H$ .

**Abstraction** – The initial abstraction steps of the watercolour effect are applied at the three levels of resolution to produce  $A_O$ ,  $A_D$  and  $A_H$ .

**Edge darkening** – Edge darkening is directly applied to  $A_O$  and  $A_D$  only, and  $A_D$  and  $A_H$  are then resized back to the input image size. Applying edge darkening with the above procedure to  $A_H$  resulted in the edge darkening appearing unnatural and unpleasant as the edges became too thick. Therefore, instead the edges from  $A_O$  were blended with  $A_H$  to perform edge darkening.<sup>2</sup>

**Wobbling** – Only a small amount of wobbling is applied to  $A_D$  in order to preserve the important details as much as possible. The magnitude and frequency of the spatially coherent noise are successively doubled and halved for  $A_O$  and  $A_H$  so as to make the stylisation of  $A_H$  more distinctive.

**Pigment Diffusion** – Compared to  $W_1$ , the increased amount of wobbling in  $W_2$  is achieved by doubling the magnitude of the noise. In addition, multiple octave separated scales of noise are added to  $A_O$  and  $A_D$ , as listed in table 1. Following that “blobby” diffusion, the uniform diffusion step is applied.

**Granulation** – For increased stylisation a strong pigment granulation or other effects can be applied to  $A_H$  using an overlay blend with an appropriate scanned texture map.

**Blending levels of detail** – First the saliency map  $S$  is used to blend the lower two resolution images  $B_1 = S \cdot A_O + (1 - S) \cdot A_H$  and this blend is then further

<sup>2</sup> Computing edges from  $A_H$  after it is resized to the original  $I_O$  size gives similar results, but the edges are still slightly thickened, and the additional more detailed mid-scale edges retained by the preferred method improve the watercolour effect.



Level of Detail	Standard Wobble	Increased Wobble
$A_D$	$\{f, m\}$	$\{f, 2m\}$
$A_O$	$\{\frac{f}{2}, 2m\}$	$\{f, 2m\}, \{\frac{f}{2}, 4m\}$
$A_S$	$\{\frac{f}{4}, 4m\}$	$\{f, 2m\}, \{\frac{f}{2}, 4m\}, \{\frac{f}{4}, 8m\}$

**Table 1.** Frequencies and magnitudes of coherent noise applied to make the standard wobble effect and the increased amount of wobbling used to produce the pigment diffusion effect.

blended with the facial mask  $F$  to produce the final blend of the three levels of detail  $B_2 = F \cdot A_D + (1 - F) \cdot B_1 = F \cdot A_D + S \cdot (1 - F) \cdot A_O + (1 - F) \cdot (1 - S) \cdot A_H$ . This weighting ensures that the eye and eye brow regions are preserved at the highest resolution, salient features (apart from the eye and eye brows) are rendered at approximately mid resolution, while the remaining non-salient areas of the image are rendered towards the lowest level of detail.

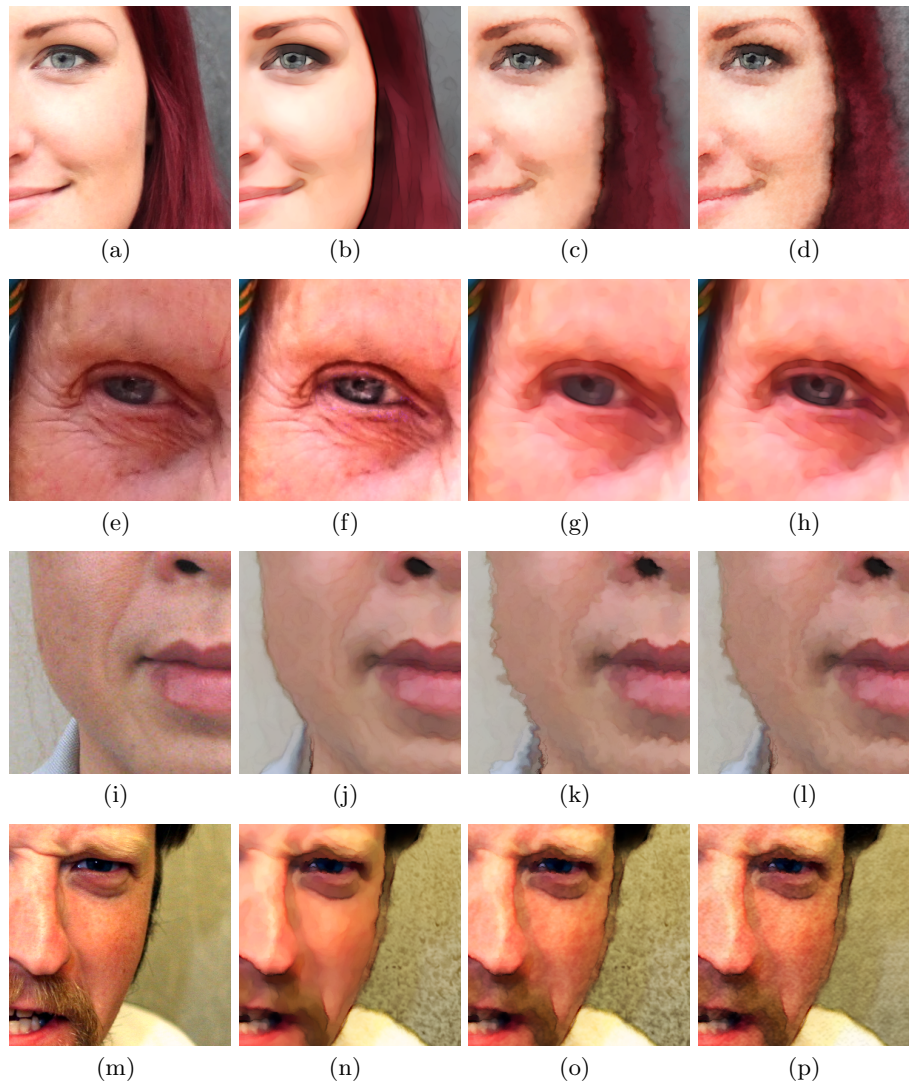
**Glazing** – Glazing is applied on the single output image  $B_2$  after the levels of detail have been combined to create  $B_3$ . In order to preserve the facial features these two renderings are combined using the facial mask to produce  $B_4 = B_2 + F \cdot B_3$

**Pigment and paper variations** – Irregularities in appearance due to blotches caused by uneven pigment density and the paper texture are added as two overlay blends to  $B_4$ . When blending the pigment blotches,  $B_4$  is set as the base layer so as to create variations in the colour density. Subsequently the paper blend sets the paper texture image as the base layer, which ensures that the texture is visible even in light and dark areas.

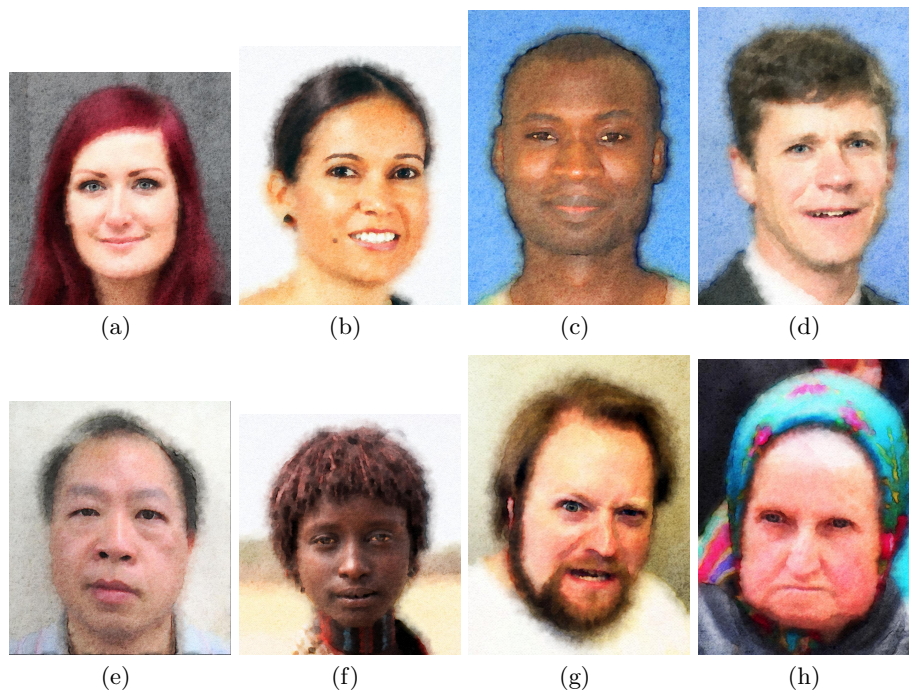
Figure 3 demonstrates some intermediate stages of the rendering pipeline. The top row shows that the combination of three levels of detail have ensured that while the hair has become significantly abstracted the eyes are only lightly stylised, and have retained their detail (see figure 3b). Likewise, diffusion is limited in the eye area (figure 3c). The second row shows how the input image (figure 3e) is brightened and the local contrast in the eye region enhanced (figure 3f). Although the difference between the abstracted results with (figure 3g) and without (figure 3h) eye enhancement is subtle, it is important to ensure that the eyes are rendered clearly. The third row shows how the two levels of wobbling ( $W_1$  and  $W_2$  in figure 3j & k) are combined to form blobby diffusion into the dark area only (figure 3l). Several texture overlays are added: strong pigment granulation in the background (figure 3n), pigment blotches (figure 3o), paper texture (figure 3p).

## 4 Experiments

This section shows results of the watercolour rendering. All results of the proposed method were obtained using fixed parameter settings, with the exception



**Fig. 3.** Intermediate steps in watercolour rendering. Note that different rows show different steps. Top row: **overall process**, (a) section of original image, (b) basic Bousseau *et al.* [5] filtering, (c) wobbling and uniform diffusion effect, (d) addition of textures; Second row: **face and eye enhancement**, (e) section of original image, (f) contrast limited local histogram equalisation (CLLHE), (g) filtering applied to original image, (h) filtering applied to CLLHE image; Third row: **diffusion**, (i) section of original image, (j) standard amount of wobbling, (k) increased amount of wobbling, (l) combination to produce “blobby” diffusion effect; Bottom row: **texture overlays**, (m) section of original image, (n)-(p) addition of pigment granulation, blotches, paper texture.



**Fig. 4.** Watercolour stylisation using the proposed method of images taken from levels one and two of the NPRportrait benchmark [22].

of figure 5d, which for comparison shows the effect of increasing the level of abstraction of the background. Note that since the watercolour effect is quite subtle the images need to be viewed at a large scale, otherwise the different elements of the stylisation will not be apparent.

Figure 4 shows results from both levels one and two of the NPRportrait benchmark [22]. Effective stylisation has been performed while maintaining the subject’s identity and avoiding any facial artifacts.

Additional results and comparisons with other image-based watercolour rendering methods are provided in figure 5. The output from our implementation of the method by Bousseau *et al.* [5] is shown in figure 5b. While the overall effect is reasonable, there is loss of detail on the face, particularly noticeable at the eyes. In contrast, our default watercolour stylisation (figure 5c) not only preserves these details, but provides more stylistic effects. Even with a version of our watercolour rendering that increases the amount of abstraction of the lower two levels of detail (figure 5d) the face is still well preserved, and artifacts avoided. Gabor rather than Perlin noise has been used to perform the wobbling, which can be seen most clearly as the oriented streaking of the background. Two results from the commercial app “DeepArt”, which is based on the deep learning



**Fig. 5.** Examples and comparison of watercolour rendering; inset images show the sources for style transfer. (a) target image, (b) Bousseau *et al.* [5], (c) our default watercolour stylisation, (d) our watercolour stylisation with Gabor noise and increased abstraction of background, (e) & (f) DeepArt [12], (g) CNNMRF [16], (h) Image Analogies [13], (i) Waterlogue, (j) BeCasso.

approach to style transfer by Gatys *et al.* [12], are shown in figure 5e & f. Some elements of the source watercolour examples have been captured, but even though well matched source images are provided, figure 5e contains significant facial artifacts while the face in figure 5f appears washed out. Another deep learning style transfer approach is demonstrated in figure 5g. The CNNMRF [16] approach gives good results for some styles, but it can be hard to control. Again, some aspects of the watercolour style have been well captured, but mismatches and inappropriate transfer have caused undesirable artifacts in the face. Figure 5h shows the result of using the watercolour image pair from the original Image Analogies [13] paper. The colours in the face are overly flattened while the highlights in the original become an unnatural gray. Moreover there are many local discontinuities due to texture transfer from widely separated source locations. A result from the commercial app Waterlogue is shown in figure 5i. The overall effect is attractive, but the eyes would benefit from less abstraction, and the excessive diffusion of the lips into the mouth is unwanted. The effect of unpainted white patches can be effective, but in this example their dense occurrence in the background is overly distracting. Finally, figure 5j shows the watercolour result from the BeCasso commercial app, which incorporates the wet in wet technique from Wang *et al.* [25]. This appears attractive in some areas, but distracting in others such as the teeth and nose.

## 5 Conclusions

An image-based non-photorealistic rendering technique for producing a watercolour effect has been developed, incorporating several stylistic elements of watercolour, i.e. edge darkening, wobbling, glazing and diffusion. Multiple levels of detail are employed, and controlled by weight maps generated from a fitted facial model as well as a general purpose saliency model. This ensures that distinctive facial characteristics are preserved and facial artifacts are avoided. A consequence of this approach is that the stylisation of the face may be too conservative, which we will address in future work.

## References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. PAMI* 34(11), 2274–2282 (2012)
2. Baltru, T., Robinson, P., Morency, L.P.: OpenFace: an open source facial behavior analysis toolkit. In: *Winter Conf. on Applic. of Computer Vision*. pp. 1–10 (2016)
3. Berger, I., Shamir, A., Mahler, M., Carter, E., Hodgins, J.: Style and abstraction in portrait sketching. *ACM Trans. Graph.* 32(4), 55:1–55:12 (2013)
4. Bousseau, A.: Watercolor tutorial. Tech. rep., Grenoble University (2006), [maverick.inria.fr/Membres/Adrien.Bousseau/watercolor\\_tutorial/tutorial1.pdf](http://maverick.inria.fr/Membres/Adrien.Bousseau/watercolor_tutorial/tutorial1.pdf)
5. Bousseau, A., Kaplan, M., Thollot, J., Sillion, F.X.: Interactive watercolor rendering with temporal coherence and abstraction. In: *Symp. NPAR*. pp. 141–149 (2006)

6. Chen, J., Turk, G., MacIntyre, B.: Watercolor inspired non-photorealistic rendering for augmented reality. In: ACM Symp. on Virtual Reality Software and Technology. pp. 231–234 (2008)
7. Curtis, C.J., Anderson, S.E., Seims, J.E., Fleischer, K.W., Salesin, D.H.: Computer-generated watercolor. In: ACM SIGGRAPH. pp. 421–430 (1997)
8. DiVerdi, S., Krishnaswamy, A., Mäch, R., Ito, D.: Painting with polygons: A procedural watercolor engine. *IEEE Trans. TVCG* 19(5), 723–735 (2013)
9. Dong, L., Lu, S., Jin, X.: Real-time image-based Chinese ink painting rendering. *Multimedia Tools and Applications* 69(3), 605–620 (2014)
10. Doran, P.J., Hughes, J.: Expressive Rendering with Watercolor. Master’s thesis, Brown University (2013)
11. Fišer, J., Jamriška, O., Simons, D., Shechtman, E., Lu, J., Asente, P., Lukáč, M., Šykora, D.: Example-based synthesis of stylized facial animations. *ACM Trans. Graph.* 36(4), 155 (2017)
12. Gatys, L.A., Ecker, A.S., Bethge, M.: Image style transfer using convolutional neural networks. In: Proc. CVPR. pp. 2414–2423 (2016)
13. Hertzmann, A., Jacobs, C.E., Oliver, N., Curless, B., Salesin, D.H.: Image analogies. In: ACM SIGGRAPH. pp. 327–340 (2001)
14. Lagae, A., Lefebvre, S., Drettakis, G., Dutré, P.: Procedural noise using sparse Gabor convolution. *ACM Trans. Graphics* 28(3), 54–64 (2009)
15. Lai, Y.K., Rosin, P.L.: Efficient circular thresholding. *IEEE Trans. Image Processing* 23(3), 992–1001 (2014)
16. Li, C., Wand, M.: Combining Markov Random Fields and convolutional neural networks for image synthesis. In: Proc. CVPR. pp. 2479–2486 (2016)
17. Liao, J., Yao, Y., Yuan, L., Hua, G., Kang, S.B.: Visual attribute transfer through deep image analogy. In: ACM SIGGRAPH. pp. 120:1–120:15 (2017)
18. Luft, T., Kobs, F., Zinser, W., Deussen, O.: Watercolor illustrations of CAD data. In: Eurographics. pp. 57–63 (2008)
19. Perlin, K.: An image synthesizer. *ACM Siggraph Computer Graphics* 19(3), 287–296 (1985)
20. Rosin, P.L., Lai, Y.K.: Artistic minimal rendering with lines and blocks. *Graphical Models* 75(4), 208–229 (2013)
21. Rosin, P.L., Lai, Y.K.: Non-photorealistic rendering of portraits. In: Proceedings of the Workshop on Computational Aesthetics. pp. 159–170 (2015)
22. Rosin, P.L., Mould, D., Berger, I., Collomosse, J., Lai, Y.K., Li, C., Li, H., Shamir, A., Wand, M., Wang, T., Winnemöller, H.: Benchmarking non-photorealistic rendering of portraits. In: Proc. Expressive. pp. 11:1–11:12 (2017)
23. Shu, Z., Shechtman, E., Samaras, D., Hadap, S.: EyeOpener: Editing eyes in the wild. *ACM Trans. Graph.* 36(1), 1:1–1:13 (2017)
24. Wang, J., Jiang, H., Yuan, Z., Cheng, M.M., Hu, X., Zheng, N.: Salient object detection: A discriminative regional feature integration approach. *Int. J. Computer Vision* 123(2), 251–268 (2017)
25. Wang, M., Wang, B., Fei, Y., Qian, K., Wang, W., Chen, J., Yong, J.H.: Towards photo watercolorization with artistic verisimilitude. *IEEE Trans. TVCG* 20(10), 1451–1460 (2014)
26. Yan, Z., Zhang, H., Wang, B., Paris, S., Yu, Y.: Automatic photo adjustment using deep neural networks. *ACM Trans. Graph.* 35(2), 11:1–11:15 (2016)
27. Zhang, Y., Dong, W., Ma, C., Mei, X., Li, K., Huang, F., Hu, B.G., Deussen, O.: Data-driven synthesis of cartoon faces using different styles. *IEEE Trans. Image Processing* 26(1), 464–478 (2017)