The Art of Scale-Space.

J. Andrew Bangham, Stuart Gibson and Richard Harvey
School of Information Systems, University of East Anglia, Norwich,
NR4 7TJ, UK. Email:{a.bangham@uea.ac.uk}

Abstract

Artists pictures rarely have photo-realistic detail. Tools to create pictures from digital photographs might, therefore, include methods for removing detail. These tools such as Gaussian and anisotropic diffusion filters and connected-set morphological filters (sieves) remove detail whilst maintaining scale-space causality, in other words new detail is not created using these operators. Non-photorealistic rendering is, therefore, a potential application of these vision techniques. It is shown that certain scale-space filters preserve the appropriate edges of retained segments of interest. The resulting images have fewer extrema and are perceptually simpler than the original. A second artistic goal is to accentuate the centre of attention by reducing detail away from the centre. The process also removes the detail providing perceptual cues about photographic texture. This allows the 'eye' to readily accept alternative, artistic, textures introduced to further create an artistic impression. Moreover, the edges bounding segments accurately represent shapes in the original image and so provide a starting point for sketches.

1 Introduction

A photographer tends to choose uncluttered backgrounds and make careful use of focus to direct attention. Of course, lens blurring is both easy and effective for it exploits the natural and powerful way in which the brain rejects non-foveated regions of a scene (they are simply out-of-focus). The technique finds its way into rendering, digital art, advertising, and video through the Gaussian blur filter widely used to de-focus background material. But the method is rarely used by painters. Rather, they direct attention by selecting detail and manipulating textures and geometry.

By contrast, a painter starts with a blank canvas, adds paint and the more skilled knows when to stop. It is the progressive addition of detail that characterizes the process of producing representational art in which only some detail directly represents that in the original scene. It difficult to capture representational detail manually from three-dimensional (3D) scenes onto two-dimensional (2D) canvases, but this does not satisfactorily explain why trained artists limit the amount of detail they use. After all two dimensional, photographic quality, images have been traced for over five centuries by those projecting images onto surfaces using concave mirrors and lens [10]). But the evidence from the resulting pictures suggests that artists pick only those details that resonate with their artistic interpretation. They *choose* to ignore some objects and lots of detail. Painting is not photography.

BMVC 2003 doi:10.5244/C.17.58

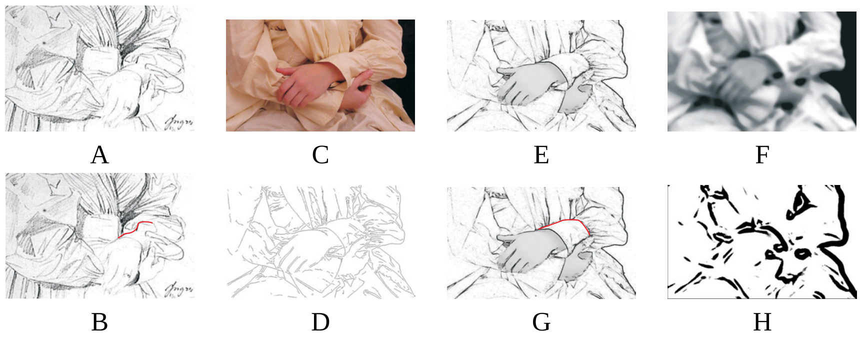


Figure 1: (A) Hockney [10] draws attention to how, in this Ingres drawing, “the cuff of the left sleeve is not followed ‘round the form’ as you would expect, but carries on into the folds”. (B) Red overlay indicates the relevant lines. (C) Photograph of a similar subject. Neither Canny (D) nor Sobel (E) edge filtering reveal the artistic line. (F) Shows in red the line line from (E) that follows the cuff rather than the light. Blurring reveals the large scale cuff-to-sleeve highlight (G) however it yields edges (H) with a graphical rather than the sketch-like appearance created by Ingres.

Selectively removing detail simplifies a photograph and is implicit in existing methods for producing painterly pictures. Systems for creating pen-and-ink drawings from existing images clearly remove both color and spatial detail and simultaneously add artistic detail in the form of pen strokes [20]. In the case of painting Haeberli *samples* the image, effecting a simplification. He then modulates and spreads the color over a larger region using a brush, an action that also adds technique detail by simulating the medium and stochastic brush strokes that are modulated by edge gradients to imply artistic interpretation [8]. By modelling the flow of water dragging pigments over paper Curtis [5] removes detail from the original photograph by a form of blur and simultaneously substitutes texture detail that replicates watercolor. Hertzmann starts by removing detail with large brushes and then uses finer brushes to selectively refine the picture where the sketch differs from blurred photograph, a form of multiscale removal of detail [9]. These methods sub-sample the source image either before or after smoothing: the standard way to remove detail and prevent aliasing. In this paper, however, we concentrate on another way to control the level of detail in a digital image. Scale-space filtering to both remove detail and uncover large scale image maxima (highlight) and minima (lowlights).

Chiarascuro (bright highlights and dark shadows) and its manipulation characterizes the work of many artist’s, since the renaissance. Hockney [10] draws attention to the way the line used by Ingres follows the light rather than the form (as evidence of optical assistance of which he gives many other examples). Figure 1(A) shows an extract from the original drawing of Madame Godinot 1829. (B) Shows the artist’s lines that, Hockney argues, follows the light. We illustrate the problem by analysing the photograph shown in (C). Conventional edge detectors (D and E) produce a prominent line along the back edge of the cuff (F): a boundary that was ignored by the artist. The problem lies with the local edge filter. Typically they have a small region-of-support that responds to the strong edges around the form and so cannot ‘see’ the larger picture (the Canny filter (D) is more complex but has related problems). Simplifying the image by blurring, Figure 1(G),

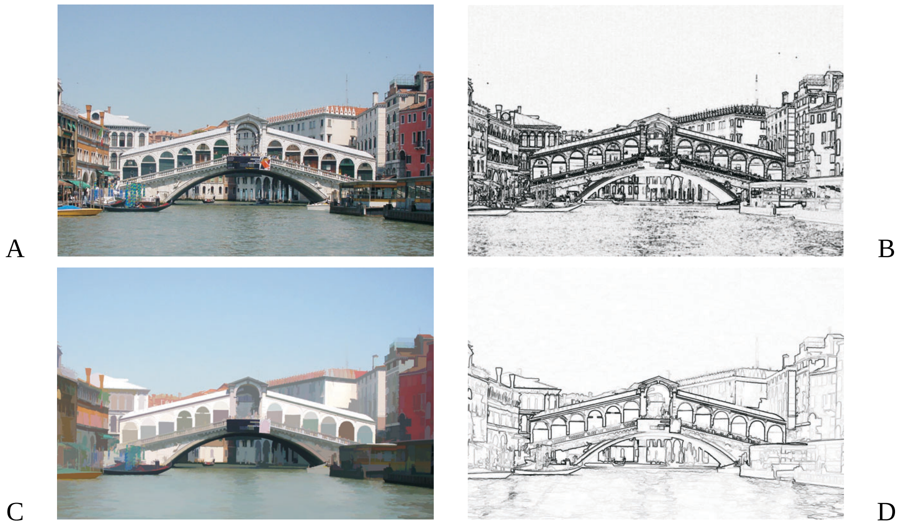


Figure 2: (A) Photograph and (B) associated edges. (C) Sieved to remove detail and (D) fewer edges make a more sketch-like picture.

increases the region-of-support and does both reveal the expected large scale highlight running from the cuff *into* the sleeve but it removes detail.

Thus Gaussian scale-space filters meet two requirements of a pre-processor for non-photorealistic rendering. As such it is used to segment images and create pictures where the artist's eye-gaze governs the level of detail rendered at different positions in the image [6]. Whilst pleasing, the results are limited in the range of styles they can support because such filters introduce significant geometric distortion reflected in the edges (H) that, whilst graphically interesting, do not form the basis of a sketch: important to the artist. Here we pursue alternative scale-space filters.

1.1 Simplification maintaining scale-space causality

In image processing the process of removing detail from a digital image emerged from studies on finding *salient*, edges [17]. The work with Gaussian filters lead to the, theoretically tidy, representation of images known as scale-space [11, 27, 14]. This important concept is seen as a requirement of image simplification systems since it guarantees that extrema in the simplified image are not artifacts of the simplification process itself. Computation systems that preserve scale-space causality are usually associated with Gaussian filters [1] and diffusion [18] in which the image forms the initial conditions for a discretization of the continuous diffusion equation: $\nabla \cdot (c\nabla f) = f_s$. If the diffusivity is a constant this becomes the linear diffusion equation, $\nabla^2 f = f_s$ which may be implemented by convolving the image with the Green's function of diffusion equation: a Gaussian convolution filter. Of course, care is needed when this equation is discretized [15] but, if it is done correctly, a scale-space with discrete space and continuous scale may be formed[1].

[1] Or with a discrete scale parameter if preferred.

Approximations to this Gaussian blur filter are common in image-editors and graphical rendering systems. The problem with blurring, when finding salient edges at large scales, is that edges wander away from the true edge and objects become rounded: a consequence of convolution, Figure 1H. It is better if diffusivity depends upon contrast, as in anisotropic diffusion, but computation then becomes lengthy and unwanted small scale detail with a high enough contrast may nevertheless be preserved. In other words, as with linear diffusion, there is an interaction between the intensity and scale of an object.

More recently the multiscale analysis of images has been explored in the field of mathematical morphology. Two rather different approaches to constructing a morphological scale-space have been suggested. In the first [22, 13] the image is either eroded or dilated using an elliptic paraboloid. As is often the case in morphology (and convolution filters) the shape of the structuring element (window) dominates over structure in the image. That said however, the brush like 'texture' introduced by the structuring element can be useful in digital art and is used in photo-editor plug-ins (Adobe Photoshop Gallery Effects).

The second approach uses those connected-set alternating sequential filters sometimes termed sieves [3]. Sieves [4] appear in a variety of guises but they have their starting point in connected-set graph morphology [19, 24, 25] and watersheds [21]. At small scale they filter out maximally stable extremal points [12] or detail and at larger scale, entire objects. Figure 2(C) confirms that fine detail is removed and that edges (D) of remaining features are well preserved. These edges are more sketch-like than those derived directly from the image (B) or from a Gaussian smoothed image Figure 1(H). This, and the more poster-like simplified image provides a reason to explore further how these scale-space filters can be used in non-photorealistic rendering.

2 Methods

We implement the sieve described in [3]. The algorithm first creates a list of all maxima and minima. These extrema are level 8- (or 4-) connected-sets of pixels that are then sorted by area. A scale decomposition progresses by merging all extrema of area 1 to the next most extreme neighbouring pixel(s), i.e. all extreme values are replaced by the value of the next most extreme adjacent pixel. If the segment remains an extremum it is added to the appropriate scale extremum list. The decomposition continues by merging all extrema of scale 2, 3 and so on. Thus, for example, by scale 100 there are no maxima (white) or minima (black) areas of less than 100 pixels. We use low-pass, band-pass and high-pass filters created by combinations of sieving operations.

The image is represented as a graph [23] $G = (V, E)$. The set of edges E describes the adjacency of the pixels (which are the vertices V). A pixel, x, is connected to its eight neighbours. A region, $C_r(G,x)$, is defined over the graph that encloses the pixel (vertex) x, $C_r(G,x) = \{\xi \in C_r(G) | x \in \xi\}$ where $C_r(G)$ is the set of connected subsets of G with r elements. Thus $C_r(G,x)$ is the set of connected subsets of r elements that contain x. For each integer $r \geq 1$ the operators ψ_r, γ_r, $\mathscr{M}^r$, $\mathscr{N}^r : \mathbf{Z}^V \to \mathbf{Z}^V$ are defined as $\psi_r f(x) = \max_{\xi \in C_r(G,x)} \min_{u \in \xi} f(u)$,, $\gamma_r f(x) = \min_{\xi \in C_r(G,x)} \max_{u \in \xi} f(u)$, $\mathscr{M}^r = \gamma_r \psi_r$, $\mathscr{N}^r = \psi_r \gamma_r$. $\mathscr{M}^r$ is a connected-set grayscale opening followed by a closing defined over a region of size r.

The types of sieve known as *M*- or *N*-sieve are formed by repeated operation of the $\mathscr{M}$ or $\mathscr{N}$ operators that are also known as connected alternating sequential filters. An

A B C D E F

Figure 3: (A, C, E) Sieve red, green blue channels to scales 10, 100, 5000 respectively. The channels have extrema that do not necessarily overlap, this creates irregular areas particularly at large scales. (B, D, F) Convex hull colour sieve to scales 10, 100, 5000 respectively. At large scales this produces a clearer result.

M-sieve of f is the sequence $(f^{(r)})_{r=1}^{\infty}$ given by

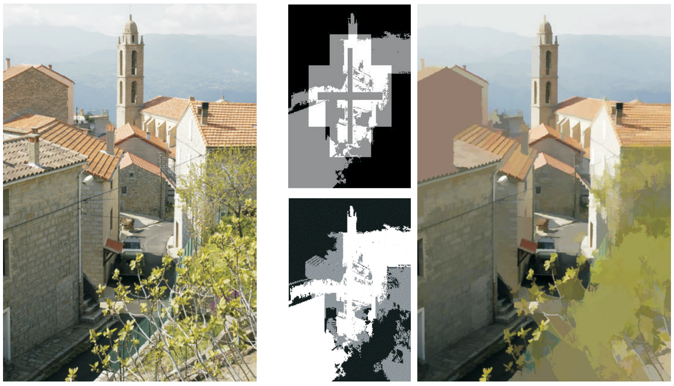
$$f^{(1)} = \mathscr{M}^1 f, \quad f^{(r+1)} = \mathscr{M}^{r+1} f^{(r)}, \quad r \geq 1 \tag{1}$$

The N-sieve is defined similarly. It has been shown how connected set openings can be performed in approximately linear time [26] using a modification to Tarjan's disjoint set algorithm and a similar implementation is used here for the alternating sequence of openings and closings that forms the sieve [2, 3] [2].

$f^{(r)}$ is a low-pass filter removing all extrema up to scale r. $f^{(1)} - f^{(r)}$ is a high-pass filter keeping all extrema from scale 1 to r and $f^{(s)} - f^{(r)}$ is a band-pass filter for extrema (granules) of scales between s and r.

The sieve requires two orderings. Level connected-sets are ordered by value and extrema are removed in order of scale. Where a pixel represents a triple, red, blue and green (RGB), there is no clear way of jointly ordering by value. This is addressed in two ways. We note that all three channels have a high correlation with brightness (unlike hue, saturation, value) and so the three colour planes are sieved independently, the RGB-sieve. The effect of removing detail can be seen by comparing Figure 2(A) and (C). It is evident that (B) has less detail yet, in contrast to alternative scale-spaces, the edges of large scale objects are preserved. This is, perhaps, more obvious in the edge images compare Fig-

[2]Note to reviewers: Aspects of the procedure are covered by patents



A B C

Figure 4: '(A) Photograph. (B) Top, the central cross designating the region of primary interest together with its border. Bottom, white segment indicates the region automatically selected to be foreground resolution, gray segment at middleground resolution, black at background resolution. (C) The union of foreground, middleground and background resolution image segments. Controlling the level of detail helps direct attention to the interest points in the centre.

ure 2B and D. The resulting image is both grayer than the original (extrema are removed) and the colours change slightly because they arise from the signals obtained from colour channels sieved independently. There is no link between a pixel and its colour.

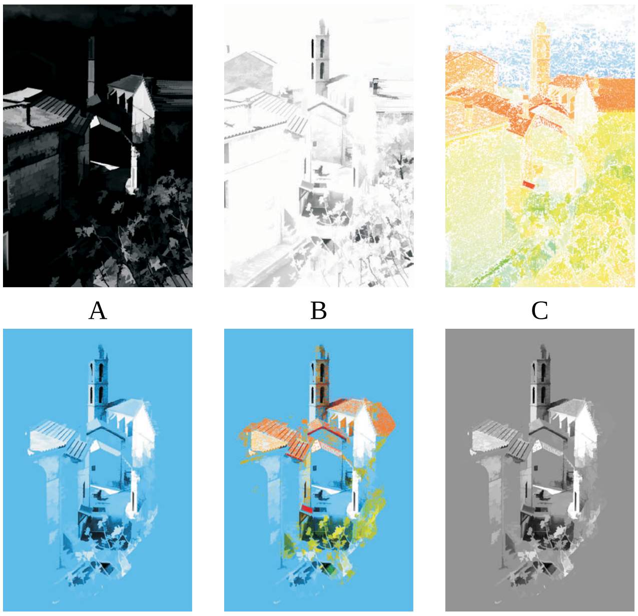
A new alternative is the convex 'colour sieve' which follows from a geometric interpretation of the colours of a region and its neighbours. A convex hull is fitted to points in the region projected into colour space. All points that lie on the convex hull itself are extreme [7] and those enclosed are not extreme. This provides an ordering - the distance from the convex hull. This definition is tidy because many typical colour transformations such as gamma correction and linear transformations affect the geometry but not the topology of the convex hull and the extrema inherit the invariance properties.

To simplify the image we merge smaller regions into larger ones without introducing additional extrema by merging to the neighbour with the closest Euclidean distance. Neighbouring regions with identical colour distances are further ordered by computing the difference of their luminance $L = (r+g+b)/3$ and further tiebreaks are achieved by ordering by their G,R and B values. The merging is repeated iteratively until idempotence.

3 Results

Figure 3 compares the two sieve implementations and shows that the low-pass colour-sieve produces colours that are less washed-out than the low-pass RGB-sieve, particularly at large scales where the colour-sieve produces a 'poster' effect more effectively than quantisation (commonly used in paint packages). The colours are also more faithful to the original as they have been selected not computed. However, the high order-complexity of the current colour-sieve means that the RGB-sieve is used in the remaining Figures.

We now consider how high- and low- lights extracted from the image using a high-



A B C D E F

Figure 5: '(A) Band-pass highlights displayed against a black backround. (B) Band-pass lowlights displayed against a white backround. (C) Band-pass HSV saturation channel used to select colours, such as the red roofs, that stand out from backround. (D) Highlights and lowlights incorporated into a picture. (E) Replacing the chroma of (D) with colours from (C) adds colour highlights (visible in colour prints, pdf only) that are not visible in the luminance image (F).

pass sieve can be incorporated into non-photorealistic image renderings. The grayscale image is band-pass sieved, $q = f^{(s)} - f^{(r)}$, to find the associated scale highlights, where $q > 0$, $y = q$, else $y = 0$. Figure 5A shows the result. Likewise, lowlights where $q < 0$, $y = 1+q$, else $y = 1$, Figure 5B. Combined by painting them onto a mid-tone background, Figure 5D, the effect is similar to chalk and charcoal. Colour highlights are located at a particular range of scales by sieving the HSV saturation channel and using this to control the chroma, $q_{hue} = f^{(s)}_{hue} - f^{(r)}_{hue}$, where $q_{hue} > t$, $hue = q_{sat}$, $sat = q_{sat}$, $val = 1$, else $hue = 1$, $sat = 1$, $val = 1$, where t is a threshold that can be adjusted by the user[3]. The colour highlights have been painted by *replacing* the NTSC chroma values on the canvas with those from the colour highlights, Figure 5E[4]

Interestingly, as Livingstone [16] points out, by colouring the canvas with the complement (the two colours sum to gray) of, for example, the red roofs an optical illusion is created. The effect is to make the colour appear more interesting that it otherwise might be for two reasons. Firstly, it challenges the viewer's vision system (and monochrome display devices) because the NTSC grayscale (perceptual luminance) does not change even when the chrominance does: all trace of the colour change vanishes in an NTSC

[3]For digital artists the convention, that user-adjustable thresholds should be avoided, is not relevant.

[4]PDF version of the paper.

A B

Figure 6: '(A) Figure 4 textured using a photograph of a watercolour wash and pencil cross-hatching (B).

grayscale print, Figure 5F. Exactly how Figure 5E appears on the printed page depends on the printer software. For readers able to see colour this in colour, Figure 5E plays to another colour illusion. The sharp boundary between the complementary colours enhances perceived brightness [16].

The brushwork in Figure 5A-F places the centre of attention in the centre of the picture by leaving the periphery free of detail. This is typical of many paintings. We, therefore, devise an algorithm that automatically selects a central region to be rendered in more detail than a middleground which, in turn is set against a backround with low detail. In other words, an algorithm that creates foreground, M_f, and middleground, M_m masks.

The process is outlined in Figure 4B. The idea is to create masks that exactly follow the boundaries of objects in the image and which place M_f in the centre and M_m around it. Each mask is created separately. The image is sieved to a scale, s, quantised by an amount q and the flat zones labelled. Those zones that intersect the innermost darker-cross and the pale-cross are then marked as shown by the white segments, Figure 4B. It white segment has an area A. The part of the pale-cross not covered by the marked zones has an area $\overline{A}$. We then search for a scale, s, and quantisation q that minimises difference between the areas, $A - \overline{A}$. An exhaustive search of only a few s and q suffices. Typical masks for M_f and M_m are shown in Figure 4B bottom panel and they have been used to combine images created by RGB-sieving to three scales, Figure 4C. The result is more detail towards the centre of the image helps draw the viewers attention. The changes of scale are subtle since they follow the boundaries of objects in the image rather than some externally imposed mask.

Notice that many of the areas in Figure 4C are flat because texture, fine detail, has been removed by the sieve. This creates an opportunity to replace the original texture with another as an artist might do by using paint or pencil. The image was mixed with a photograph of a simple watercolour wash (not shown) (multiplication rather than addition), Figure 6A, produces a distinctly watercolour like result. Mixing the same texture with the original is much less effective (easily achieved in Photoshop) because the underlying original detail leaves old texture cues intact. A more extreme example is shown in Figure 6B. Here, a photograph of an area of pencil cross-hatching is mixed with Figure 4C. Unlike Figure 6A however, each labelled level set in Figure 4C is filled with a segment of the cross-hatched picked from a random position in the texture image. In other words each of the objects is hatched separately. This is most clearly seen in the large flat areas top left and bottom right. Superimposing the edges completes the effect.

4 Conclusion

The sieve, particulary the convex-hull colour, algorithm is a useful starting point for non-photorealist rendering of photographs. It provides the digital artist with access to a choice of images with differently scaled detail. Unlike blurring, the system simplifies without distorting edges thus the edges provide a useful starting point for creating sketches. The large scale level sets it creates provide a mechanism for segmenting the image into regions that, by having different amounts of detail, create a centre of attention. It is data-driven rather than dependent on a pre-defined geometry. Band-pass sieves also allow artistically important high-, low- and bright coloured highlights to be found. Thus in Figure 7A the sieve removes small scale detail and the highlights are now treated in a way that is redolent of Figure 1A with edges that follow the light Figure 7B. We do not attempt to map photographs directly into art: the artist is still essential. Rather the aim is to provide the digital artist with tools. Further automatation might include object recognition to create ways of improving composition and tools to balance colour composition.

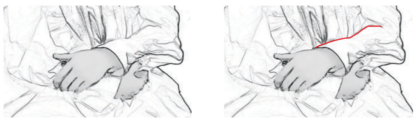
A  B

Figure 7: (A) Sobel edges after sieving RGB Figure 1C to scale 2000: the line carries into the folds. (B) Red line indicates the line.

References

[1] J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda. Uniqueness of the Gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8:pp 26–33, 1986.

[2] J. A. Bangham, P. Chardaire, P. Ling, and C. J. Pye. Multiscale nonlinear decomposition: the sieve decomposition theorem. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18:518–527, 1996.

[3] J. A. Bangham, R. Harvey, P. Ling, and R. V. Aldridge. Nonlinear area and volume scale-space preserving filters'. *Journal of Electronic Imaging*, 5(3):283–299, July 1996.

[4] J. A. Bangham and S. Marshall. Image and signal processing with mathematical morphology. *IEE Electronics and Communication Engineering Journal*, 10:117–128, 1998.

[5] C. J. Curtis, S. E. Anderson, J. E. Seims, K. W. Fleischer, and D. H. Salesin. Computer-generated watercolor. In *Proceedings of SIGGRAPH 97*, pages 421–430, Los Angeles, California, August 1997. ACM SIGGRAPH, ACM Press/Addison-Wesley. ISBN 0-89791-896-7.

[6] Doug DeCarlo and Anthony Santella. Stylization and abstraction of photographs. In *SIGGRAPH'2002 Proceedings*, pages 769–776. ACM, 2002.

[7] S. Gibson, Richard Harvey, and Graham D. Finlayson. Convex colour sieves. In *Proc. Scale-Space'2003*, pages 550–563, 2003.

[8] P. E. Haeberli. Paint by numbers: Abstract image representations. In *Proceedings of SIGGRAPH 90*, pages 207–214, August 1990.

[9] Aaron Hertzmann. Painterly rendering with curved brush strokes of multiple sizes. In M. E. Cohen, editor, *Proceedings of SIGGRAPH 98*, pages 453–460, Orlando, Florida, July 19-24 1998. ACM SIGGRAPH, ACM Press/Addison-Wesley. ISBN 0-89791-999-8.

[10] D. Hockney. *Secret Knowledge*. Routledge and Kegan Paul, 2001.

[11] T. Iijima. Basic theory of pattern normalization (for the case of a typical one-dimensional pattern. *Bulletin of the Electrotechnical Laboratory*, 26:368–388, 1962.

[12] M. Urban J. Matas, O. Chum and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *BMVC*, volume 1, pages 384–393, Cardiff, 2002.

[13] Paul T. Jackway and Mohamed Deriche. Scale-space properties of the multiscale morphological dilation-erosion. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18:38–51, 1996.

[14] J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[15] Tony Lindeberg. Scale-space from discrete signals. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12:234–254, 1990.

[16] Margaret Livingstone. *Vision and Art: the biology of seeing*. Harry N. Abrams, New York, 2002.

[17] D. Marr. *Vision*. W. H. Freeman and Company, New York, 1982.

[18] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

[19] P. Salembier and J. Serra. Morphological multiscale image segmentation. In *SPIE Visual Communication and Image Processing*, pages 620–631, 1992.

[20] M. P. Salisbury, S. E. Anderson, R. Barzel, and D. H. Salesin. Interactive pen-and-ink illustration. In A.S. Glassner, editor, *Proceedings of SIGGRAPH 94*, pages 101–108, Orlando, Florida, July 1994. ACM SIGGRAPH, ACM Press/Addison-Wesley.

[21] P. Soille and L. Vincent. Determining watersheds in digital pictures via flooding simulations. In M. Kunt, editor, *Visual Communications and Image Processing '90*, volume SPIE-1360, pages 240–250, 1990.

[22] R. van den Boomgaard and A. Smeulders. The morphological structure of images: the differential equations of morphological scale-space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(11):1101–1113, November 1994.

[23] L. Vincent. Graphs and mathematical morphology. *Signal Processing*, 16:365:388, 1989.

[24] L. Vincent. Morphological area openings and closings of greyscale images. In *Workshop on Shape in Picture*. NATO, 1992c.

[25] L. Vincent. Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. *IEEE Trans Image Processing*, vol. 2:pp 176–201, 1993.

[26] M. H. F. Wilkinson and J. B. T. M. Roerdink. Fast morphological attribute operations using tarjan's union-find algorithm. In L. Vincent J. Goutsias and D.S. Bloomberg, editors, *Int Symposium on Mathematical Morphology*, pages 311–320. Kluwer, 2000.

[27] A. P. Witkin. Scale-space filtering. In *8th Int. Joint Conf. Artificial Intelligence*, pages 1019–1022. IEEE, 1983.