I will begin with a summary of Tomasi and Manduchi's bilateral filter and then discuss how their approach can be extended into the joint bilateral filter, as described by Georg Petschnigg, which deals with flash / no-flash image pairs. Filtering is a principal technique in image processing and is typically used to enhance an image - this often refers to smoothing, sharpening or strengthening edges. In spatial filtering approaches, the intensity of a pixel is influenced by the intensities of neighboring pixels. The intuition behind this is that pixel intensities tend not to fluctuate greatly within a small region of an image, unless these pixels have been corrupted by noise. In an effort to reduce this noise, we can smooth the image by replacing the value of a pixel with the average of its neighbors (applying a mean filter [1]), although this causes blurring at edges where variations in pixel intensity are vast and deliberate.

PART A - Bilateral Filter

Bilateral filtering is a *non-linear* spatial filtering approach used to smooth an image whilst still preserving its edges. As in all spatial filtering techniques, a pixel's new intensity is determined by the properties of its neighbors. The *bilateral filter* [2] infers the level of influence individually held by each neighbor through considering both similarity and spatial proximity to the pixel at the centre of the neighborhood. The size of the neighborhood can be increased, although it is worth noting that pixels that are a large distance apart are unlikely to be relevant to one another. In general, neighbors that are further away from and dissimilar in intensity to a pixel have less influence over its final value. The similarity of two pixels p and p' in an image A, refers to the difference of their intensities, (A_p - $A_{p'}$), whereas spatial proximity refers to the euclidean distance between the coordinates of p and p', given by (p - p').

In order to calculate the output intensity of a pixel, p, we implement the following: ref [3]

$$A_p^{Base} = \frac{1}{k(p)} \sum_{p' \in \Omega} g_d(p'-p)\, g_r(A_p - A_{p'})\, A_{p'} \qquad k(p) = \sum_{p' \in \Omega} g_d(p'-p)\, g_r(A_p - A_{p'})\,.$$

where $k(p)$ is a normalisation term

The equation above describes how each pixel, p', in the neighborhood of p (denoted Ω) is considered when calculating the new intensity of p. The gaussian function, g_d, is applied to the distance between p and p' and g_r is applied to the difference of their intensities. Furthermore, the relative importance of spatial proximity and similarity is separately defined by the standard deviations of g_d and g_r, denoted by σ_d and σ_r respectively. The product of g_d and g_r can be described as the weight attributed to p' which is the combined influence of spatial proximity and similarity, therefore both values should be high if p' is to have significant influence over p. The normalisation term, k(p), is the sum of the weights of pixels in the neighborhood. Hence, dividing by k(p) ensures that the average intensity of pixels in A remains the same before and after filtering. The number of pixels that influence the intensity of p is determined by the size of Ω, however as Ω increases, it takes longer to apply the bilateral filter - especially to large images.

The gaussian function, g(x), is positive everywhere and symmetrical about the mean, μ, therefore when μ = 0, g(x) = g(-x). Hence, it is not necessary to take the absolute value of (A_p - $A_{p'}$) or (p - p'). The gaussian function is fundamental to the success of the bilateral filter as it enables us to decide the relative importance of spatial proximity and similarity. Defining σ_d and σ_r separately allows us to experiment with attributing more influence over A_p to one factor.

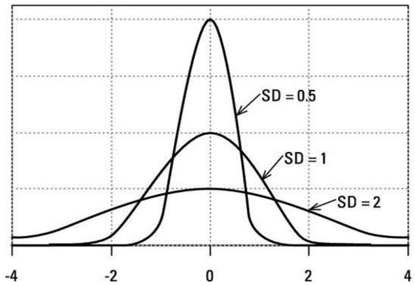


Image source: [4]

As standard deviation increases, the gaussian curve "flattens", lowering the peak of the bell-shaped curve. A low standard deviation indicates that more of the distribution lingers close to the mean while a high standard deviation suggests that the distribution is spread more widely across a range of x values. In the context of the bilateral filter, closer and more similar pixels will have a greater influence over the output intensity of p, denoted A_p, when the standard deviation is lower. The gaussian function outputs larger values when the input is close to the mean, and this effect is amplified as σ tends towards 0.

This provides the bilateral filter with its characteristic ability to preserve edges as we can prevent two contrasting pixels from having an effect on the intensities of one another, such that a dark pixel will have negligible influence on the intensity of a neighboring light pixel. This is most effective at boundaries between *high contrast* pixels as the difference in intensity is vast therefore it is more obvious that they should not integrate. One application of this in MRI scans and medical imagery [5] as edge deterioration may result in misdiagnosis if image detail is not preserved.

In order to experiment with the parameters of the bilateral filter, I used the OpenCV function *bilateralFilter(src, d, sigmaColor, sigmaSpace)*, where src refers to the image upon which the filter is applied, d is the diameter of each pixel's neighborhood, and sigmaSpace and sigmaColor refer to σ_d and σ_r respectively. The bilateral filter is an effective image denoising technique, however it is crucial that filter parameters are suited to the image and the level of enhancement required. In order to access the source image and to write the filtered image to a folder, I used OpenCV's image reading and writing functionality.

The level of filtering required is often dependent on the image. For example, significant smoothing of test image 1 would result in loss of detail, whereas test image 2 would benefit from more smoothing in order to soften wrinkles and improve the appearance of the skin. I will investigate the effect of varying the parameters σ_d, σ_r and d when applying the bilateral filter to test 1 and test 2.

Image A

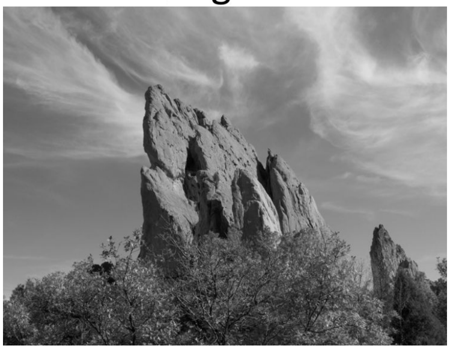


Image B

Image C

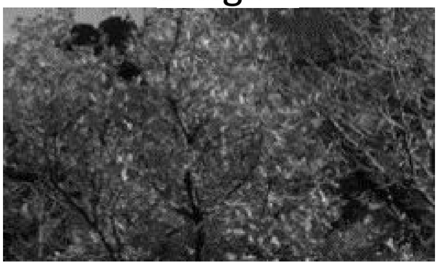


Image D

TEST IMAGE 1

As sigmaSpace (σ_d) becomes very large, the gaussian curve with standard deviation σ_d will flatten. This will essentially negate the importance of spatial proximity such that only $g_r(A_p - A_{p'})$ will contribute towards the output value of each pixel. Image B demonstrates the effect of this as σ_d was set to 10000, σ_r = 20, and d = 9. Edges in image B are sharper than those in the original (image A), although some detail has been lost in the clouds and trees.

I tweaked parameters in order to enhance the image, such that the sky was smoothed but still contained some natural detail and the foreground (rock and trees) retained most of the original definition. I found that this effect was achieved when the parameters were set to σ_r = 5, σ_d = 5 and d = 9 - as shown in image C. However, resource [3] suggests that σ_r should be set between 0.05 and 0.1 therefore I also experimented with σ_r = 0.075 (image D). For very small values of σ_r, p and p' can only influence the intensity of one another if their intensities are very similar. This effect is prominent in image D, especially in finely detailed regions such as the trees. Image D appears sharper than image C as there is a high contrast between dark and light colors which accentuates detail.

Image 1

Image 2

Image 3

TEST IMAGE 2

As sigmaColor (σ_r) becomes very large, the bilateral filter mimics a gaussian blur as $g_r(A_p - A_{p'})$ becomes almost constant over the intensity interval covered by the image [6]. Image 2 demonstrates this property as σ_r = 10000, σ_d = 20, and d = 9. It is clear to see that image 2 has been smoothed significantly, however the sharp edges present in the original (image 1) have not been preserved. Image 3 displays the effect of filtering with σ_r = 50, σ_d = 50 and d = 100. As the neighborhood (Ω) is large, pixels that are far away but still within the neighborhood are able to influence the new value of a pixel. This is compounded by the fact that σ_d and σ_r are both large, therefore neighboring pixels that are far away from and dissimilar to a pixel still have significant influence over it's intensity. The result of this is an ultra-smooth, cartoon-like rendition of the original image whereby the hair and skin have become block colors and the woman in the image no longer looks realistic.

Image 1

Image 4

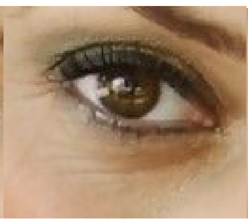
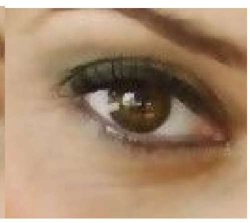


Image 5



Images 4 and 5 focus on the appearance of the under eye wrinkles (the full images can be viewed in the appendix). I experimented with various parameters, however I found that maintaining definition in the hair whilst still softening wrinkles was very difficult. This effect is prominent in image 5 which was filtered using the parameters σ_r = 80, σ_d = 80 and d = 5. Whilst the wrinkles have almost disappeared and the skin is sufficiently smoothed, fine details (such as eyelashes and eyebrows) have been over-blurred. Despite this, the high contrast between the white of the eye and the black makeup means that edges around the eye have been preserved quite well. I found that applying the bilateral filter with parameters σ_r = 35, σ_d = 160 and d = 5 produced good results which are displayed in image 4. The appearance of the wrinkles has been reduced and some detail in the skin and hair has been preserved - making her look realistic.

As shown in the images above, even after carefully adjusting the parameters, the basic bilateral filter tends to either lose detail or fail to sufficiently smooth the image in certain regions.

PART B - Joint Bilateral Filter

The *joint bilateral filter* [3] is an extension of the bilateral filter which uses two images taken in quick succession - a flash image and a no-flash / ambient image. The contrasting properties of flash and no-flash images enable a prominent application of the joint bilateral filter known as *ambient image denoising*, whereby fine detail is reintroduced to images taken in low-light environments. Flash images are relatively noise-free and provide a good estimate of high-frequency detail, although they do not preserve the natural illumination in an ambient environment. Conversely, no-flash images do not display the level of detail found in flash images and are often noiser - despite their ability to capture ambience. In order to combine the strengths of the flash / no-flash image pair, the joint bilateral filter enhances the ambient image, A, with detail found in the flash image, F.

In order to calculate the output intensity of a pixel, p, we implement the following: ref [3]

$$A_p^{NR} = \frac{1}{k(p)} \sum_{p' \in \Omega} g_d(p' - p)\, g_r(F_p - F_{p'})\, A_{p'} \qquad k(p) = \sum_{p' \in \Omega} g_d(p' - p)\, g_r(F_p - F_{p'})$$

where A^{NR} is the noise reduced version of A and $k(p)$ is a normalisation term

Like the bilateral filter, the new intensity of each pixel, p, is a non-linear combination of the intensities of its neighbors. The joint bilateral filter also considers both spatial proximity and similarity of one pixel to another, such that neighboring pixels that are far away from and vastly different in intensity to the pixel at the centre of the neighborhood have little influence over its new intensity. Two gaussian functions, g_d and g_r are used to determine the relative importance of each factor. The distance function g_d is computed from the ambient image, however the edge-stopping function, g_r, is computed from the flash image as edges in image F are typically sharper and easier to detect. F is considered to be a good estimator of the high frequency content in the ambient image, hence this allows us to reinforce the edges of image A even if they are hardly present.

The joint bilateral filter is particularly effective when the input image is not guaranteed to provide reliable edge information - this could be due to noise or poor image quality. Useful properties of the joint bilateral filter such as edge reinforcement and noise removal mean that it is used in many varied applications. One of these applications is upsampling used in tone mapping [7]. Tone mapping methods are used to compress the intensity values of a high dynamic range (HDR) image, such that it can be mapped to a low dynamic range display. In this application, a joint bilateral filter is applied to upsample a low resolution solution computed for a downsampled version of the high resolution image.

As described above, a major application of the joint bilateral is ambient image denoising. The introduction of the flash image allows us to denoise and enhance the ambient image, producing something that looks more like the real scene. Illumination from the flash increases the signal to noise ratio (known as SNR), therefore the edges are more defined in the flash image. Edges in the flash image can be used as estimates of the less visible edges in the ambient image.

In my implementation of the joint bilateral filter, each RGB color channel is filtered separately with the same standard deviation parameters [3]. In addition, I decided to use only odd diameters to ensure that each neighborhood has a defined centre. I will investigate the effect of varying the parameters σ_d, σ_r and d of the joint bilateral filter when it is applied to the flash / no flash image pair test 3a and test 3b.

Image 6 Image 7 Image 8

Image 7 displays the effect of setting σ_d = 100, σ_r = 15 and d = 15. As σ_r is small, the black detailing on each vase is relatively well preserved as only pixels with similar intensities are able to interfere with one another. This image is very smooth, although its edges are blurry and texture in the background has been lost. Both standard deviation parameters in image 8 were made very large (σ_d = 10000, σ_r = 10000 and d = 15). In this instance, neither spatial proximity nor difference in intensity gives a neighboring pixel more influence than any other pixel in the neighborhood. Hence, the bilateral filter almost becomes a mean filter as each neighboring pixel is considered equally when determining the new intensity of a pixel.

I experimented with varying each parameter in the joint bilateral filter, however it quickly became clear that setting σ_r to 5 was producing good images. Image 9 displays the effect of setting σ_d = 1000, σ_r = 5 and d = 9. In this image, vase details are clear and the edges are well-defined despite the fact that a significant amount of noise has been removed. When σ_r was increased beyond 5, some of this detail began to fade therefore I experimented with decreasing σ_r. In image 10, σ_r was set to 2 and in image 11, σ_r was set to 0.5 - sigmaSpace and the diameter remained at 1000 and 9 respectively. Images 10 and 11 both contain more noise than image 9 - this is most visible on the body of the vases. However, they have each retained more texture in the background and sharper edges. I believe that image 10 looks the best as it achieves a good balance between noise reduction and blurring.

In conclusion, the joint bilateral filter is better able to preserve detail while reducing noise than the basic bilateral filter. However, a limitation of both approaches is that they are nonlinear, hence even the simplest implementation requires convolution in the spatial domain which can be very slow when σ_d is large. As the flash image is used as an estimator of the ambient image, one limitation of the joint bilateral filter is that encounters with flash shadows may cause inconsistencies in blurring of the ambient image. This happens because shadows in the flash image are not present in the ambient image.

References

[1] Griffin, L. (2000). Mean, median and mode Filtering of images. *Medical Imaging Science Interdisciplinary Research Group, King's College, London, UK*, [online] pp.2995-3004.

[2] C. Tomasi and R. Manduchi, "Bilateral Filtering for Gray and Color Images", *Proceedings of the 1998 IEEE International Conference on Computer Vision*, Bombay, India.

[3] Georg Petschnigg, Maneesh Agrawala, Hugues Hoppe. Richard Szeliski, Michael Cohen, Kentaro Toyama. Digital photography with flash and no-flash image pairs. *ACM Trans. Graphics (SIGGRAPH)*, 23(3), 2004.

[4] Patrick Weaver, Project Manager (2016), http://projectmanager.com.au/can-you-use-standard-deviation-in-project-management/

[5] Yu-Ju Lin and Herng-Hua Chang, "Automatic Noise Removal in MR Images Using Bilateral Filtering Associated with Artificial Neural Networks," International Journal of Pharma Medicine and Biological Sciences, Vol. 4, No. 1, pp. 39-43, January 2015.

[6] Paris, Sylvain. (2007). A gentle introduction to bilateral filtering and its applications. Pp.9-12

[7] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele, “Joint bilateral upsampling,” ACM Trans. on Graphics, vol. 26, no. 3, July 2007.

APPENDIX

Image 1

Image 2

Image 3

Image 4

APPENDIX

Image 5

APPENDIX

Image A

Image B

APPENDIX

Image C

Image D

APPENDIX

Image 6

Image 7

Image 8

Image 9

Image 10

Image 11