A Gaussian filter is a linear spatial noise reducing filter. For each pixel it takes a weighted sum of the intensities of its surrounding pixels and multiplies it by the intensity of that pixel to produce a new intensity. For each pixel the defined area of surrounding pixels is referred to as the neighbourhood of a pixel.

This is done in practice through convolving the image with a Gaussian function. A matrix is formed with the value of each element being calculated using the Gaussian function (below) of the Euclidean distance to the centre of the matrix and a supplied value for the standard deviation. The matrix is then normalised through dividing by sum of elements. Resulting in the property that the weights sum to one to ensure the product is within range of possible intensities. The matrix is then applied to the neighbourhood for each pixel with the results summed to produce the new intensity for that pixel. This is done for each colour channel before the channels are merged to produce the resulting filtered image.

$$I_p^{output} = \frac{\sum_{p' \in \Omega} g(|p - p'|) I_{p'}}{\sum_{p' \in \Omega} g(|p - p'|)} \qquad g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

While the Gaussian filter does reduce noise through smoothing, it also reduces detail due to equally smoothing all parts of the image with the same weighting without taking into account any edges since only the spatial information is used with each intensity. This often results in Gaussian blur with a high enough standard deviation, an effect that can be useful artistically but not as a product of attempting noise reduction.

A bilateral filter is a more effective noise reducing smoothing filter which expands upon the Gaussian filter. Unlike Gaussian it is edge preserving and non-linear which means convolution cannot be used. A bilateral filter uses not only the spatial distance but also the difference in the intensities of the pixels. The filter uses an increased weight for pixels of similar intensities and a decreased weight for pixels that have a large difference in intensity.

The weight for each pixel in the neighbourhood is now calculated through the spatial and intensity differences passed through the Gaussian function with a separate standard deviation for each type of difference. The sum of each of these weights multiplied by a pixel intensity is normalised through dividing the sum by the sum of the weights. This process is applied to all pixels in each channel of the image but convolution can no longer be used since the weight now also depends on the individual difference in intensities rather than purely on the distance.

This has the effect of now having an increased weighting for similar in intensity pixel and a decreased weight for those drastically different in intensity. This means that edges are more likely to be conserved as those beyond on edge are more likely to be a different intensity and so have a decreased affect upon the filtered intensity.

One limitation the filter is the staircase effect. Due to the filter removing detail, as it is a smoothing filter, this can then create areas of close to or completely constant intensity. These areas then produce new contours due to the clear difference in intensity instead of their being a gradient. This affect can be useful artistically to create cartoon-esque images for examples but in the interest of preserving image detail and removing noise this effect is a limitation. Another is gradient-reversal, the filter preserves edges but does not always preserve the gradient at these edges.

The filter can also be used for videos. Applying the filter through the time axis will filter pixels that occur in the same locations across frame that are the same colour. Since bilateral only uses pixels of the same colour for filtering it means that if another object takes up that space in the frame it won't be used in filtering the previous object.

1.1

1.2

1.3

2.1

2.2

2.3

2.4

D = diameter of the neighbourhood

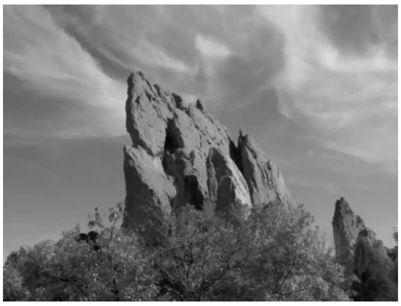
S_1 = standard deviation for the colour space

S_2 = standard deviation for the coordinate space

In the format of $D/S_1/S_2$

1.1 10/10/10
1.2 50/10/10
1.3 10/100/10
1.4 10/10/1000

2.1 10/50/10
2.2 50/50/10
2.3 10/250/10
2.4 10/50/5000



1.4

The bilateral filter has three parameters excluding the input image and the resulting image. The first being the size of the neighbourhood /mask used in the filter. As this increases there is a greater number of pixels involved in the filtering and so a greater smoothing takes place sections that are larger than the neighbourhood that are unbroken by edges. Areas of the image that are contained within edges that were already smaller than the neighbourhood do not undergo further particularly visible smoothing than before as the increase includes pixels that are outside of an edge. These have a minor weighting in the sum and so do not have a major effect that section. This can be seen in test1 where an increase in neighbourhood results in further visible smoothing of the sky but not a visible change elsewhere in the image where sections covered by the neighbourhood were already somewhat bounded by edges.

The second parameter is the standard deviation or sigma value for the Gaussian function used for the difference in intensity which is for the colour space. As this increases the filter acts more and more like a Gaussian filter due to the weightings of different intensities becoming more similar. This the spatial difference becomes the defining feature for extremely high values for this parameter. The filter that acts more closely to a Gaussian filter resulting in a more visible Gaussian blur across the image as the parameter increases.

The final parameter is the standard deviation or sigma value for the Gaussian function used for spatial difference. As this increases the weights for pixels at difference distance become increasingly more similar and within edges greater smoothing as long as the pixels are similar enough in their intensity the importance of distance within the neighbourhood decreases as this increases. This effect is not a strong as the increase in neighbourhood as can be seen by comparing the relevant images.

Joint bilateral filter

There are often times when a scene that is illuminated in low level lighting needs to be photographed. This low level lighting creates an ambiance and key to the scene but also produce a low signal-to-noise ration when recorded. The low illumination makes it difficult to capture enough information to accurately photograph the scene. One solution is to use flash when recording the scene which has a higher signal-to-noise ratio and often increases the amount of surface detail recorded. The downsides are the flash disrupts the natural lighting of the scene, removing whatever ambience was there and producing harsh distracting shadows.

A method to overcome this is to use a joint bilateral filter which is an extension of the bilateral filter. A flash and non-flash image are taken in quick succession and both are used by the filter to smooth and reduce the noise in the non-flash image while preserving the edges. The idea is that the flash image will have more accurately recorded the composition of the scene and have well-defined edges for the objects contained within due to the higher signal-to-noise ration caused by the increased illumination. The non-flash image on the other hand will have maintained the natural lighting and ambience of the scene. The filter is very similar to the bilateral filter with the non-flash pixel intensities still being replaced by a weighted sum. The spatial difference and pixel intensities being multiplied by the weight are still from the non-flash image but the alteration is that the intensity difference used in the filter is from the flash image. Denoising photographs of originally low level lighting is a particularly common application for the joint bilateral filter.

This is just one example of the concept and process of the filter. The general idea is that there are two images, one which will be filtered and one which contains further or separate information that the filter will use to guide its process upon the first image.

Another application is using the joint bilateral filter for depth map refinement. When recording depth information about a space input depth values often contain speckle and non-Gaussian noise and edge boundaries can sometimes be poorly defined. The filter can be used with an RGB image as the guide image with the depth map as the filtering target, basing the weights off the colour or intensity values within the image. The small non-Gaussian noise can be removed through smoothing and due to the edge preserving process of the filter the depth map can be fitted around the objects to strongly define edge boundaries within the map. This process can also result in poor definition being shifted to the outside of the depth map and certain types of noise could be transferred from the guide image onto the depth map. The filter can also be implemented to fill spatio-temporal holes with a depth map similar to how bilateral filters are used to filter video.

The left image was filtered with d = 3, S_1 = 0.01 and S_2 = 8 while the right was d = 11, S_1 = 0.01 and S_2 = 8.It can clearly be seen that increasing the neighbourhood size results in more effective smoothing and more noise removed from the image while maintaining the edges of the objects within. A larger area being used for each weighted sum for each pixel, which will mean that the amount of noise contained within the neighbourhood will further become a minority. This in turn further removes noise and improves smoothing since the more pixels that are taken into account surrounding a pixel the more accurate the approximation for that intensity will be.

The standard deviation for the colour space has been kept low as with bilateral as this increases the filter becomes closer and closer to acting like a Gaussian filter and so results in more and more blur. This is due to the distribution becoming more and more even for different intensities which means that an increasing part of the weight depends on the spatial difference rather than the intensity difference which is approaching a Gaussian filter. It is kept low to prevent blur and ensure that the edges are preserved. With the flash image having accurately recorded scene a high standard deviation is not required since there should be a clear intensity difference and so we can be more discerning with our intensity distribution. The standard deviation has been kept slightly higher for the special difference because of the fact that we have the flash image to discern which pixels will be appropriate and so we want to increase the number of pixels that could be included knowing that the intensity difference will do most of the work for selection. This is within reason since closer pixels are still more likely to similar so S_2 has been set at a reasonable mid point.

[1] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe and K. Toyama. "Digital photography with flash and no-flash image pairs." In ACM Trans. on Graphics, vol. 23, no. 3, pp. 664-672, 2004.
[2] E. Eisemann and F. Durand. "Flash photography enhancement via intrinsic relighting." In ACM Trans. on Graphics, vol. 23, no. 3, pp. 673-678, 2004.
[3] T. Matsuo , N. Fukushima Y. Ishibashi Weighted Joint Bilateral Filter with Slope Depth Compensation Filter for Depth Map Refinement
[4] H. Winnemoller S. C. Olsen B. Gooch Real-Time Video Abstraction
[5] P. Kornprobst, J. Tumblin , F. Durand. (2009). Bilateral Filtering: Theory and Applications. Foundations and Trends in Computer Graphics and Vision
[6] P. Kornprobst, J. Tumblin , F. Durand , S. Paris A Gentle Introduction to Bilateral Filtering and its Applications
[7] P. Kornprobst, J. Tumblin , F. Durand , S. Paris Bilateral Filtering: Theory and Applications
[8] Numerical-tourscom. 2018. Numerical-tourscom. [Online]. [10 December 2018]. Available from: http://www.numerical-tours.com/matlab/denoisingadv_8_bilateral/