Fuzzy Anisotropic Diffusion: A Rule Based Approach

Jialin Song, H. R. Tizhoosh
System Design Engineering Department, University of Waterloo
200 University Ave. West, Waterloo, ON N2L 3G1, Canada
j8song@uwaterloo.ca, tizhoosh@uwaterloo.ca

ABSTRACT
A fuzzy approach to anisotropic diffusion algorithm is proposed in this paper. In the preprocessing stage, the edginess and noisiness memberships are calculated. In the consequent fuzzy inference stage, the diffusion coefficients are approximated using the edginess and noisiness information to control the diffusion process. The advantages of this approach rely on its flexible and robust edge definition and the consideration of noise distribution. The proposed algorithm results in a better edge-preserving noise removal effect and faster diffusion process.

Keywords: Fuzzy Anisotropic Diffusion, Fuzzy Reasoning, Rule Base, Noise Reduction, Edge Enhancement

1. INTRODUCTION
The major tasks of low and intermediate level image processing are noise removal, edge enhancement and segmentation. There have been many approaches developed to implement these tasks. However, because the noise and the edges are both high frequency image components, most of the conventional approaches do not work well for edge-preserving smoothing of images corrupted with noise. With these filters, we always have to find a tradeoff between sharpening and blurring.

Since first proposed by Perona and Malik in 1990 [1], anisotropic diffusion has been developed and applied to different areas of image processing including edge enhancement, noise reduction and segmentation. The essential idea of diffusion approach is quite simple: deriving a family of increasingly smooth images by convolving the original image with a Gaussian kernel of increasing width to estimate the original image. To avoid blurring at the edges, instead of using the constant diffusion coefficients based on the original linear isotropic diffusion, an edge stopping function was proposed to estimate the diffusion coefficients, which ensures the diffusion process taking place mainly inside of the regions rather than at their boundaries and thus the smoothing happens only in the interior of regions without crossing the edges. This algorithm is called anisotropic diffusion or non-linear diffusion.

In this paper, a rule based fuzzy anisotropic diffusion is proposed. Instead of using the gradient as the edge factor in the anisotropic diffusion, the edge detectors are introduced into the diffusion processing to provide a more flexible and robust way to define the edges. A fuzzy inference system is employed to replace the edge stopping function to approximate diffusion coefficients, which can be expected to have more control on the diffusion processing. Also the consideration of noise distribution in the fuzzy inference makes the diffusion process faster and more robust.

2. BASICS OF ANISOTROPIC DIFFUSION
The basic idea behind the diffusion processing is to use a family of increasingly smooth version images \( u(x, y, t) \) of the original image \( u_0(x, y) \), indexed by diffusion parameter \( t \), to estimate the original image. This process can be viewed as the result of the image convolving with a Gaussian kernel of increasing width [1]:

\[
I(x, y, t) = I_0(x, y) \ast G(x, y, t)
\]

The anisotropic diffusion equation proposed in [1] is defined as follows:

\[
\frac{\partial I}{\partial t} = \text{div}(c(x, y, t) \nabla I) = c(x, y, t) \Delta I + \nabla c(x, y, t) \nabla I
\]

where \( c(x, y, t) = g(\|\nabla I(x, y, t)\|) \) is the diffusion coefficient. \( \nabla I \) denotes the gradient of the image. \( g(\cdot) \) is called edge stopping function, which is selected as a decreasing function of the gradient of the image. The initial condition is \( I(x, y, 0) = I_0(x, y) \).

The discrete version of anisotropic diffusion was proposed as follows [1]:

\[
I^{t+1}_s = I^t_s + \frac{\lambda}{|\mathcal{N}_S|} \sum_{p \in \mathcal{N}_S} g(\nabla I_s, p) \nabla I_s, p
\]

where \( \mathcal{N}_S \) represents the four neighboring pixels in North, West, South and East diffusion directions, \( |\mathcal{N}_S| \) is equal to the number of pixels in the neighborhood which is normally 4 except for at the boundaries. \( \nabla I_s, p = I_p - I_s \) is the discrete gradient in one of four diffusion directions.

Here, the \( g(\nabla I) \) function acts as an edge stopping function. It gets its maximum value (=1) at constant signal regions where the gradients are small and falls to zero at the edges where the gradients are large. Black et al. [2] proposed an edge stopping functions based on robust statistics called Tukey’s biweight which is defined as follows:

\[
g(x, \sigma) = \begin{cases} 
\frac{1}{2 \sigma^2 - (x/\sigma)^2} & |x| < \sigma \\
0 & \text{otherwise}
\end{cases}
\]

where \( \sigma \) is called the scale parameter which should be selected to be smaller than the gradient at edges and larger than the gradient at noise.

Because the Tukey’s biweight edge stopping function can descend to zero while the gradient becomes larger than the scale parameter, it may be expected to get sharp edges. However, there...
are still at least two drawbacks with the proposed Tukey’s biweight edge stopping function.

1) One major drawback is related with the selection of the scale parameter $\sigma$. The proposed edge stopping function requires the scale parameter $\sigma$ to be selected in such a way that it should be larger than the gradient of the noise but smaller than the real discontinuities, the edges. This selection is to ensure the edges with gradient larger than $\sigma$ to be well preserved and the noise with gradient smaller than $\sigma$ to be smoothed. But in practice, when the gradients generated by some of the noise are comparable with those by the edges, it is not sufficient to separate them according to the strength of gradient. The selection of $\sigma$ becomes a trade-off between preserving the preferred edges and smoothing the noise. In such a situation, we have to either preserve both the high energy noise and the preferred edges or smooth them out together.

2) Another drawback is about the strength of smoothing at the noisy regions. Because the Tukey’s biweight function is a decreasing function of the gradient, it is obvious that the function will get smaller diffusion coefficients at the noisy regions than at the flat regions, thus slower diffusion speed at noise. Obviously, a slower diffusion speed at noise will result in more iterations to be required. As discussed in [4], a larger amount of iterations, on one side, is a time consuming process, and on another side, may finally make the edges blurred and damaged.

In this paper, we propose a fuzzy approach to anisotropic diffusion which can overcome the drawbacks discussed above. The anisotropic diffusion algorithm with Tukey’s biweight function described [2] will be used as a reference for performance comparison.

### 3. RELATED PREVIOUS WORKS

Fuzzy logic, as a powerful tool in representing and processing vagueness and human-like reasoning, was also applied to the anisotropic diffusion algorithm. Santiago Aja et al. [4] introduced an anisotropic diffusion filter controlled by fuzzy rules. They improved the anisotropic diffusion by calculating diffusion coefficients with fuzzy inference instead of the edge stopping function. The luminance difference between each pixel and its 8 neighboring pixels was calculated and the output diffusion coefficients were mapped in a decreasing way to the luminance difference from 0 to 1 in their fuzzy inference system. Their fuzzy rule base works in such a way that it makes the diffusion processing easier in constant signal areas and difficult at the zones with high gradients.

### 4. THE PROPOSED FUZZY ANISOTROPIC DIFFUSION

As mentioned before, for the anisotropic diffusion in [1] and [2], the scale parameter $\sigma$ has to be selected larger than the gradient at noise and smaller than the real edges. This means that the edges are simply defined by a selected constant number of the gradient. The pixels with gradient larger than the scale parameter $\sigma$ will be treated as edges and will be preserved. The remains are viewed as constant signal or noise and thus will be filtered out finally with a large enough number of iterations. This definition, however, does not work well for the image with overlapping in the strength of edges and noise. Also as discussed in [5], the geometry-driven anisotropic approaches work poorly in images with low contrast or low signal-to-noise ratio.

Suppose if we could get the better defined edge information and use this information to achieve more accurate control on the diffusion processing, we should be able to overcome the shortcoming mentioned above. Based on this idea, we propose a fuzzy anisotropic diffusion algorithm to approximate the diffusion coefficients to control the diffusion processing. The proposed fuzzy anisotropic diffusion processing includes three steps: pre-processing, fuzzy inference and diffusion iterations.

The behavior of the diffusion process is determined by the diffusion coefficients. In our fuzzy approach, instead of using a decreasing function (Eq.3) of the gradient as the edge stopping function (diffusion coefficient function), the edge and noise distribution information detected at the preprocessing step are fed into the fuzzy inference system as input to approximate the diffusion coefficients. As the outputs, each pixel of the image will get one corresponding fuzzy diffusion coefficient which will determine the diffusion process among itself and its neighboring pixels. Therefore, the strength of smoothing at the edges and the interior of the regions are controlled by fuzzy rules during the diffusion iterations according to the local image characteristics. The proposed Fuzzy anisotropic diffusion algorithm is explained in details in the following sections:

#### 4.1. Pre-Processing

The pre-processing step consists of two separated tasks, edge detection and noise detection, through which the exact edge and noise membership values are calculated from the input image. These membership values will be used to calculate the diffusion coefficient using fuzzy inference in the next step.

#### 4.1.1 Edge detection: One of the important subjective factors of noise quality judgment is the sharpness at edges. In [1], the gradient is used as kind of edge detector. The edges are simply defined as those pixels with gradient larger than the selected scale parameter $\sigma$. However, it is not accurate and does not work well under the low contrast and low signal-noise-ratio situations. It can be expected that if we can find a better way to define the edges, the result should be better as well.

In our fuzzy approach, instead of using gradient as the measurement of edges, some robust edge detectors or their variations become applicable to calculate the degree of edginess. This gives us a more flexible and robust way to measure and handle the edges. The edge detector is applied to the input image to generate an edge map. The edginess membership value of each pixel is calculated as the corresponding normalized edge map value. In our experiments, some simple edge detectors such as Sobel, Prewitt, Roberts, Laplacian of Gaussian, Canny and a fast fuzzy edge detector [7] are available and the selection of edge detectors depends on the characteristics of the images. There is no restriction on the selection of edge detectors. However, since the quality of the estimation of edge and noise affects the performance of the fuzzy inference and thus the whole algorithm directly, the edge detector should be able to detect the true edges or the edges preferred. In other word, the detected edge map should include all of the edges required. The setting of the parameters for the edge detector depends on the requirements and the types of the applications. The detected edge values are normalized into the interval of 0 and 1 to represent the degree of edginess.

To improve the performance of edge detection and get more accurate edge information, some pre-edge-detection processes or post-edge-detection processes are also applicable. Those
optional pre-edge-detection processes and post-edge-detection processes are as follows:

1) Pre-edge-detection processes: these operations are applied before the edge detection to get better edge representation.
   - Gaussian filtering: apply Gaussian filter to the image before the edge detection. Notice that this Gaussian smoothed version of input image is only for edge detection and will be not fed to the diffusion iterations. The purpose to apply Gaussian filtering is to reduce the noise sensitivity of edge detectors to minimize the number of noise remain on the edge map.
   - Contrast enhancement: for some low contrast images, the adaptive or non-adaptive contrast enhancement operators can be applied to the input image to get better performance on edge detection.

2) Post-edge-detection processes: these operations are applied after the edge detection to remove the noise left on the edge map.
   - Weighted median filter (WMF) or other filters to remove noise: since the edges are more likely to be connected, the weighted median filter or similar filters can be use to remove those isolated pixels from the edge map which are more likely to be generated by noise.
   - Weighted median filter to detect details: human observers subjectively prefer sharper images with a little noise over noise-free blurred images. In fact, researches show that human vision system is less sensitive to noise close to strong edges than those in the constant signal regions [6]. To avoid blur on the details of image, it is reasonable to protect a small neighborhood with details, which is usually edge concentrated area, from smoothing. The weighted median filter is used to detect those neighborhoods with details.
   - Gaussian filter to protect details: those small neighborhoods with details detected by WMF will be convolved with a Gaussian kernel to give some weights to the neighboring pixels close to edges on the edge map for less smoothing during diffusion iterations.

We applied the Sobel edge detector to the cameraman image corrupted by speckle multiplicative noise (variance 0.04). The left picture of Fig.4.1.1 is the edge map of noisy cameraman image detected with Sobel operator from matlab. The white pixels on the map indicate there are edges, thus will be preserved from smoothing. There is obviously some noise left on the edge map and some of the edges are corrupted. To improve the performance of edge detection by Sobel operator, a Gaussian filter is applied on the image before the edge detection to reducing noise. WMF was applied to remove the isolated pixel, which is considered as generated by noise, from the edge map after edge detection. The small neighborhoods with details, e.g. the camera, the tripod and the human’s face on the image, are convolved with Gaussian filter to give the 3x3 neighboring pixels around the center edge pixel a small positive value to protect them from smoothing. The modified version of edge map has less noise and less edge corruption as shown on the right image of Fig.4.1.1. For the segmentation task, a thin edge is better because we only want to preserve the edge rather than the details in the neighborhood. The values of the edge map are normalized to the interval of 0 and 1 to represent the edginess membership values.

4.1.2 Noisiness detection: The noise distribution is used in the fuzzy inference system to calculate the diffusion coefficients at each pixel of non-edgy regions. We expect the diffusion strength to be relatively stronger at noisy regions than the flat regions. The degree of noisiness for each pixel uses the standard deviation which is calculated as the absolute difference between the pixel intensity and the local mean of its neighborhood, Eq.4. A 5x5 or 7x7 window may be appropriate to calculate the local mean of the image.

\[
\mu_{n,s} = \frac{I_s}{\eta_s} \sum_{p \in \eta_s} I_p
\]

where \(I_s\) is the intensity of the center pixel and \(\eta_s\) denotes the neighboring pixels. The resulting \(\mu_{n,s}\) values are normalized to the interval of 0 and 1 to represent the noisiness membership.

4.2. Fuzzy inference

Our fuzzy inference system can be defined as a process of mapping from the input edge and noise information to the output diffusion coefficients, using the theory of fuzzy logic. The basic idea is quite simple: if there is an edge, then do not smoothing, or if there is no edge and it is noisy then do strong smoothing. In such a way, the edges will be well protected and the smoothing happens only in the interior of the regions rather than crossing the edges. Our fuzzy inference system consists of 2 inputs, 48 fuzzy rules and 1 output. The fuzzy rules will be like “if the edginess is low and the nosiness is high, then the diffusion coefficient is high”. In the fuzzification step, the antecedents, the input edge and noise information, are mapped into the interval of 0 and 1 to represent the degree of edginess and noisiness respectively. After “and” operation, the results are applied to the consequent, which is also known as implication. The implication operation truncates the output fuzzy set using “min” operator according to the degree of the antecedents. The aggregation operation truncates the output fuzzy set using “one fuzzy set” for each output and through defuzzification processing to get a single number as the output of fuzzy inference. In our system the centroid method is used for defuzzification. Our fuzzy rule base has 48 fuzzy rules which are shown in Table4.2.1 and Fig.4.2.1. The fuzzy rules are defined as “if the edginess is E0 and the noisiness is N0, then the coefficient is C2”. This fuzzy rule base ensures the output is decreasing with the degree of edginess and increasing with degree of noisiness.
After defuzzification, the fuzzy coefficients $C_{fuzzy}(I_E, I_N)$ are used to control the diffusion processing during the iterations. The output image is calculated in an iterative approach as follows:

$$I_{s+1} = I_s + \lambda \sum \min(C_{fuzzy}(I_{E,S}, I_{N,S}), C_{fuzzy}(I_{E,P}, I_{N,P})) W_i S_p$$

where the gradient $\nabla I_S, P = I_P - I_S$, $I_E, I_N$ are degree of edginess and noisiness respectively. $\lambda$ is called diffusion step parameter, which is to control the diffusion speed. The bigger $\lambda$ is selected, the faster the diffusion process. $\eta_S$ represents the four neighboring pixels in North, West, South and East directions. $|\eta_S|$ is equal to the number of pixels in the neighborhood which is normally 4 except for pixels at the boundaries. Each pixel has its own fuzzy diffusion coefficient. We take the min operation,

$$\min(C_{fuzzy}(I_{E,S}, I_{N,S}), C_{fuzzy}(I_{E,P}, I_{N,P}))$$

of the fuzzy diffusion coefficients between the center pixel and its neighboring pixel to get the smaller one as the diffusion coefficient in that direction. The min operator is used to ensure that the diffusion process will not occur in the corresponding direction as soon as one pixel in that direction reaches an edge. The number of iterations is determined by the purpose of application. A small number of iterations will leave more details on the image while a large number of iterations will result in a stronger smoothing effect.

5. Experiments

To study the performance of our fuzzy anisotropic diffusion on noise smoothing and edge enhancement, we applied our approach to the 256×256 grayscale cameraman image corrupted by the speckle multiplicative noise. The Sobel edge detector is employed in the preprocessing step of our fuzzy approach to calculate the degree of edginess. The diffusion coefficients calculated by our fuzzy approach and the conventional anisotropic diffusion are displayed in Fig.5.1.2. From the map of the output fuzzy diffusion coefficients, Fig.5.1.2, it is easy to see that the coefficients at the edges and the regions with details are in black color which means fuzzy diffusion coefficients are small at those areas. The coefficients are also relatively small at the dark areas, which are the less noisy regions because the multiplicative noise corrupted image is less noisy in those low intensity regions, e.g. the cloth and hair of the cameraman. The output fuzzy diffusion coefficients are relatively large at the bright noisy regions as expected.

With small number of iterations, Fig 5.1.3, 5.1.4, for the output image of fuzzy anisotropic diffusion, the information of the image details is better preserved, the edges are enhanced and the other regions get smoother than the conventional anisotropic diffusion algorithm. After a large number of iterations, Fig 5.1.5.a, with our fuzzy anisotropic diffusion, the edge information still remains on the image while the other areas get smoother. On the contrary, for the conventional anisotropic diffusion, we face a trade-off of leaving some of the high energy noise remain on the output image or destroy some of the weaker edges, see Fig.5.1.5.b, 5.1.5.c. As we mentioned in the previous sections, the proposed Tukey’s biweight edge stopping function requires the scale parameter $\sigma$ to be selected in such a way that it should be larger than the noise but smaller than the real discontinuities, the edges. This selection is to ensure the edges larger than $\sigma$ to be well preserved and the noise smaller than $\sigma$ to be smoothed. But in practice, the gradient of some noise is selected, the faster the diffusion process.

4.3. Diffusion iteration

To study the performance of our fuzzy anisotropic diffusion on noise smoothing and edge enhancement, we applied our approach to the 256×256 grayscale cameraman image corrupted by the speckle multiplicative noise. The Sobel edge detector is employed in the preprocessing step of our fuzzy approach to calculate the degree of edginess. The diffusion coefficients calculated by our fuzzy approach and the conventional anisotropic diffusion are displayed in Fig.5.1.2. From the map of the output fuzzy diffusion coefficients, Fig.5.1.2, it is easy to see that the coefficients at the edges and the regions with details are in black color which means fuzzy diffusion coefficients are small at those areas. The coefficients are also relatively small at the dark areas, which are the less noisy regions because the multiplicative noise corrupted image is less noisy in those low intensity regions, e.g. the cloth and hair of the cameraman. The output fuzzy diffusion coefficients are relatively large at the bright noisy regions as expected.
(a) Our FAD approach
(b) Anisotropic diffusion

Fig. 5.1.1 (a) The original cameraman image
Fig. 5.1.1 (b) Image with Speckle noise (variance=0.04)
Fig. 5.1.2 maps of output diffusion coefficients for noisy "cameraman" image

(a) Our fuzzy AD, $\lambda = 0.25$
(b) AD with $\lambda = 0.25, \sigma = 0.25$

Fig. 5.1.3 Resulting images after 5 iterations
Fig. 5.1.4 Resulting images after 20 iterations,

Fig. 5.1.5 (a) Resulting images after 200 iterations
Our fuzzy anisotropic diffusion with $\lambda = 0.25$

Fig. 5.1.5 (b) Resulting images after 200 iterations
Anisotropic diffusion with $\lambda = 0.25, \sigma = 0.25$

Fig. 5.1.6 MSE of resulting images by FAD and AD

Fig. 5.1.5 (c) Resulting images after 200 iterations
Anisotropic diffusion with $\lambda = 0.25, \sigma = 0.3$
To demonstrate the performance on the edge enhancement, the resulting images of the IC image corrupted by Gaussian additive noise (Fig.5.1.7) are shown in Fig.5.1.8 and Fig.5.1.9. The resulting images of our fuzzy anisotropic diffusion seem to be much smoother with less noise in the flat areas and sharper in the edgy regions than the conventional anisotropic diffusion. The mean square error (Fig.5.1.10), MSE shows that the fuzzy approach have a much faster diffusion speed and gets its minimal MSE in fewer number of iterations than the anisotropic diffusion. From the curve of MSE, we can also conclude that the resulting images of the fuzzy approach always have smaller MSE than those of the anisotropic diffusion.

6. CONCLUSION

In this paper, a fuzzy anisotropic diffusion algorithm is proposed. Instead of using the gradient, the edginess and noisiness memberships are calculated in a preprocessing stage. The edge detectors provide a much more flexible and robust way to define the edges. The integration of noise distribution within a fuzzy inference system provides a much faster diffusion speed than the conventional way. The experimental results show that the proposed fuzzy anisotropic diffusion has both better visual effects and smaller mean square error than the conventional anisotropic diffusion proposed in [2]. As our future work, we will study the diffusion processing for the images with low contrast or low signal-noise-ratio.

7. REFERENCES