Removal of impulse noise in digital images with naïve Bayes classifier method

Cafer BUDAK1,*, Mustafa TÜRK2, Abdullah TOPRAK3

1Department of Electric-Electronic Engineering, Batman University, Batman, Turkey
2Department of Electric-Electronic Engineering, Frat University, Elazığ, Turkey
3Department of Electric-Electronic Engineering, Dicle University, Diyarbakır, Turkey

Received: 07.01.2014 • Accepted/Published Online: 25.11.2014 • Final Version: 15.04.2016

Abstract: A new method has been presented in this paper to remove randomly formed impulse noise in digital images. This method is one of the favorite learning approaches of the Bayes learning method and is frequently called the naïve Bayes classifier. It has especially been used more frequently in recent times in the field of signal processing. Prior to restoration of the noisy pixels of the image as is done here, the image is first separated into pieces, and then an associated learning set is formed for each piece using the noise-free pixels. These learning sets that are different for each piece are used in order to estimate the pixel that will replace the noisy one. The proposed method is both simple and easy to apply. Our comprehensive experimental studies show that our proposed method outperforms other filters that are very popular in the literature.

Key words: Impulse noise, naïve Bayes classifier, image restoration, machine learning

1. Introduction

Recent developments in image processing methods bring up new possibilities to increase human life quality in various fields. Impulse noise occurs during taking and sending digital images due to many natural and nonnatural reasons (digital recording errors, errors in communication devices, electromagnetic interference, synchronization errors, camera sensors, etc.) [1]. Impulse noise is classified into two different categories; the first type is salt and pepper noise with either a minimum or a maximum value (0 or 255), and the second type is random-valued impulse noise, which is uniformly distributed in the range of [0, 255]. Random-valued impulse noise is more difficult to handle due to the random distribution of noisy pixel values. Since the random noise values are continuously distributed in the range of the original pixels’ values, it is more difficult to determine which pixels have been corrupted [2]. Impulse noise significantly decreases the quality of the image and hence should be eliminated. Therefore, effective noise removal methods are still important in image processing [3]. Median-based nonlinear filters are used widely for the removal of impulse noise since they are fast and efficient. The main principle of median-based filters is to replace the gray-level value of every pixel by the median of its neighbors. A significant shortcoming of these filters is that they can cause defects in the originality of the images by ruining significant details in the image. Since they are usually implemented identically across the images, both noisy and noise-free pixels are modified. Consequently, some desirable details in the image can be removed [4]. An ideal filter is expected to preserve the details in the digital image while it effectively removes impulse noise. The objective of standard signal processing is to eliminate impulse noise without harming details of the image. To
solve this problem, filters of different structures were proposed for improving the filtering effect and preserving the details. Among these filters, the adaptive switching median [5] filter was shown to have good performance for identifying the pixels that are likely to be noise candidates, but with the increase of noise level, especially above 60%, the precision of noise detection decreases steeply. The improved switching median filter [6] uses four one-dimensional Laplacian operators to detect impulses and to separate them from edges. The boundary discriminative noise detection [7] algorithm is a very effective filter. However, it is very time-consuming in calculation, and its detection mechanism is not suitable for images that are contaminated with random-valued impulse noise. The directional weighted median filter [8] is another efficient impulse noise removing strategy. This filter makes use of the differences that exist between the current pixel and the neighboring pixels in four-edge directions. The purpose of this filter is to find out whether the current pixel is noisy or noise-free. A decision-based algorithm (DBA) [9] was used for restoration of images highly corrupted by impulse noise. The DBA uses a small $3 \times 3$ window having only neighbors of the corrupted pixel that have higher correlation. It provides more edge details, leading to better edge preservation. It shows promising results with lower processing time, but it degrades the visual quality of the image as the noise density is increased. To overcome this problem, an improved decision-based algorithm (IDBA) [10] was proposed. In this method, the noise candidate is replaced by the median or mean of neighboring pixels and since the uncorrupted pixels are not modified this algorithm is more effective. However, in this replacement process, the features of the image cannot be preserved. The improved fuzzy-based switching median [11] filter computes all the differences between each pixel and the central pixel (corrupted pixel) in a selected window and then calculates the membership value for the corrupted pixel based on the highest difference. However, in a real-time environment, the density of noise cannot be known in advance as the original image is not available, so choosing a fixed size window in a real-time application is an unrealistic assumption. An impulse noise detector was suggested based on the simple neuro-fuzzy (SNFF) [12] method. The proposed detector contains two identical neuro-fuzzy subdetectors combined with a decision maker. The internal parameters of the subdetectors are adaptively adjusted by training. Its algorithm requires a high computation time since it includes two 3-input neuro-fuzzy detectors and a decision maker. Moreover, the performance of the detector is significantly reduced for corrupted images with noise greater than 50%.

Additionally, many nonlinear filtering methods described in the literature have good ability to remove noise successfully [13–20], but this does not mean that no further studies should be carried out on this topic. It is still possible to develop methods that give less damage to the image while removing noise effectively.

The main propose of this study is to evaluate and validate a new approach that can be used effectively to remove impulse noise using the naïve Bayes classifier. The new method presented here is an alternative to the filters that have been designed until now and is open for development. The proposed algorithm has the ability to restore impulse noise effectively. The appropriate pixel that has the most local similarity is used to replace the noisy pixels while keeping the details as well. Our method has been compared with other filters such as the DBA, IDBA, SNFF, center-adaptive weighted median filter (CWMF) [21], adaptive median filter (AMF) [22], decision-based switching median filter (DBSMF) [23], triangular-based linear interpolation with differential evolution (TLIDE) [24], and directional difference-based noise detector and adaptive weighted mean (NFT) [25], which are traditional in the literature, and better results have been obtained as a result of the experiments that have been carried out. Non-median-based filters in the literature also use median values when restoring noisy pixels. For example, filters that remove noise using artificial neural networks also use median values when training the network. Naïve Bayes, which is one of the classification model algorithms, has been selected to solve this problem here. This method has no relation with the median filter and it has a completely
unique structure. From this point onwards, the filter will be referred to as the naïve Bayes classifier filter (NBCF).

1.1. Noise models
A gray-scale image is symbolized by a two-dimensional array where a location (i, j) is a position in the image and named as a pixel or picture element. During this study, standard matrix notation is used for images. For example, when \( U \) is an image, \( U_{i,j} \) will represent the intensity value of \( U \) at the pixel location (i, j) in the image domain. Experiments were performed by varying the amount of noise. Generally, a noisy image can be modeled as:

\[
U_{i,j} = \begin{cases} 
    n_{i,j} & \text{with probability } p \\
    O_{i,j} & \text{with probability } (1 - p)
\end{cases},
\]

where \( p \) is the percentage of the amount of noise, \( n_{i,j} \) is the value of the impulse noise, and \( O_{i,j} \) is the original pixel value. There are mainly two types of noise models used in this study. They are:

- **Fixed impulse noise model**: \( n_{i,j} \) can take only one of the two values, which can be either 255 or 0, for an 8-bit image.

- **Random-valued impulse noise model**: \( n_{i,j} \) can take any value, which can be chosen uniformly from the range of \([0, 255]\), for an 8-bit image [26].

2. Naïve Bayes classifier
Naïve Bayes is a simple and fast statistical estimation algorithm that performs well in image and signal processing applications, both in terms of accuracy and computational time. Naïve Bayes has similarities with artificial neural networks and decision trees in some respects. A naïve Bayes classifier can be trained to classify patterns involving thousands of attributes and can be applied to thousands of patterns. Therefore, naïve Bayes is a preferred algorithm for text mining and other large classification problems.

The naïve Bayes classifier is a machine learning method with proven success [27]. It can predict class membership probabilities, such as the probability of a given sample belonging to a particular class. It basically protects the dependency of data features to a single specified class and performs the most appropriate classification. The naïve Bayes classifier assumes that the effect of an attribute value on a given class is independent of the values of the other attributes. This is known as the naïve Bayes independence assumption [28]. It is made to simplify the computation involved, and in this sense it is considered naïve. Even though this assumption is not valid in many problems, naïve Bayes mostly provides a very good classification performance [29]. Independence assumption and parameters for each feature are learned separately, and this makes learning easier, especially when the number of features is high [30].

The naïve Bayes classifier is applied to learning processes in which each x sample is expressed by a feature vector and where the target function \( f(x) \) can take on any value among those in the \( V \) (target classification based on finite set of classes) limited value set. A training set is prepared for the target function. A new sample is expressed by \( \langle a_1, a_2, \ldots a_n \rangle \), and the feature vector is taken and the trainer is asked to estimate its class. The determination of the class of the new sample according to the naïve Bayes approach is the assigning of the most probable target value \( \nu_{MAP} \) (maximum a posteriori hypothesis) given the feature vectors that express the sample \( \langle a_1, a_2, \ldots a_n \rangle \).

\[
\nu_{MAP} = \arg \max_{v_j \in V} P(v_j | a_1, a_2, \ldots a_n)
\]
Eq. (1) can be rewritten as follows using the Bayes theorem:

$$\nu_{MAP} = \arg \max_{v_j \in V} \frac{P(a_1, a_2, \ldots, a_n|v_j)P(v_j)}{P(a_1, a_2, \ldots, a_n)}$$

(2)

$$\nu_{MAP} = \arg \max_{v_j \in V} P(a_1, a_2, \ldots, a_n|v_j)P(v_j)$$

(3)

The two terms in Eq. (3) are calculated based on training data. Each $P(v_j)$ value can be easily calculated by counting the $v_j$ frequency for each of the target values. It may cause a problem when the number of these terms is equal to the number of the products of possible samples and possible targets. Hence, each sample has to be seen in the sample space many times in order to acquire realistic results. The naïve Bayes classifier is based on the assumption that feature values are independent of the target value. In other words, the probability of the \(<a_1, a_2, \ldots, a_n>\) context is the product of the probability of each feature:

$$P(a_1, a_2, \ldots, a_n|v_j) = \prod_{i=1}^{n} P(a_i|v_j)$$

(4)

Thus, Eq. (3) could be rewritten as follows to define the naïve Bayes classification approach.

$$\nu_{NB} = \arg \max_{v_j \in V} P(v_j) \prod_{i=1}^{n} P(a_i|v_j)$$

(5)

$v_{NB}$ represents the output, that is the class of the naïve Bayes classifier. The number of different $P(a_i|v_j)$ terms that should be calculated using training data in a naïve Bayes classifier is equal to the product of the values of different features with the number of different target values. In short, the naïve Bayes learning method is a learning stage in which various $P(v_j)$ and $P(a_i|v_j)$ terms are calculated based on their frequencies in the training data. This calculation method is related to the learning hypothesis. This hypothesis is later used to classify each new sample by applying Eq. (5).

A visualized version of the classification process explained in Section 2.1 can be seen in Figure 1.

![Figure 1. Visualization of the classification task.](image-url)
2.1. Random-valued impulse noise detection process

Detecting a noisy pixel in a random-valued noise corrupted image is more difficult than detecting a fixed valued noise because the value of a noisy pixel can be much higher or lower than the value of the neighboring pixels. That is why conventional median filters do not perform well, especially with random-valued high noise rates.

\[
\begin{array}{c|c|c|c}
  & j-1 & j & j+1 \\
\hline
i-1 & a_1 & a_2 & a_3 \\
\hline
i & a_4 & A_{scan} & a_5 \\
\hline
i+1 & a_6 & a_7 & a_8 \\
\end{array}
\]

Figure 2. Pixel intensity and coordinates of a 3 x 3 window.

Our process of detecting the random-valued noise consists of two phases. In the first phase, at the beginning of iteration, the differences between the central pixel \( A_{scan} \) and the closest left and right pixels \( a_4 \) and \( a_5 \) should be calculated, which is windowed in Figure 2. The same process is applied for the closest top and bottom pixels \( a_2 \) and \( a_7 \). If the absolute difference between the left and right pixel values \( |a_4 - a_5| \) is greater than \( |A_{scan} - a_4| \) or \( |A_{scan} - a_5| \), or the absolute difference between the top and bottom pixel values \( |a_2 - a_7| \) is greater than \( |A_{scan} - a_2| \) or \( |A_{scan} - a_7| \), then this pixel will be regarded as a noisy pixel.

\[
\begin{align*}
  F_1 &= |a_4 - a_5| \\
  F_2 &= |A_{scan} - a_4| \\
  F_3 &= |A_{scan} - a_5| \\
  F_{11} &= |a_2 - a_7| \\
  F_{22} &= |A_{scan} - a_2| \\
  F_{33} &= |A_{scan} - a_7|
\end{align*}
\]

\[
A_{scan} = \begin{cases} 
  \text{Noisy pixel,} & \text{if (} F_2 \text{ or } F_3 > F_1 \text{) or (} F_{22} \text{ or } F_{33} > F_{11} \text{)} \\
  \text{Noise free pixel,} & \text{otherwise}
\end{cases}
\]

The first stage sometimes cannot successfully detect highly corrupted pixels. In this case the second phase may be applied, where the difference between the diagonal pixels should be compared. The noisy pixel is detected in the same manner as explained in the first phase.

\[
\begin{align*}
  F_a &= |a_3 - a_6| \\
  F_{aa} &= |a_1 - a_8| \\
  F_b &= |A_{scan} - a_3| \\
  F_{bb} &= |A_{scan} - a_1| \\
  F_c &= |A_{scan} - a_6| \\
  F_{cc} &= |A_{scan} - a_8|
\end{align*}
\]

\[
A_{scan} = \begin{cases} 
  \text{Noisy pixel,} & \text{if (} F_b \text{ or } F_c > F_a \text{) or (} F_{bb} \text{ or } F_{cc} > F_{aa} \text{)} \\
  \text{Noise free pixel,} & \text{otherwise}
\end{cases}
\]

Pixels that are found to be noisy should be marked. The values of corrupted pixels will not be used in the restoration phase.

2.2. The Use of Naïve Bayes Classifier as Filter

After the noisy pixels in the image are detected, our algorithm passes to the second stage, i.e. the restoration stage. The restoration process is only applied to the noisy pixels. In order to specify the pixel values that are to be restored, the naïve Bayes classifier will be used. The proposed overall algorithm is illustrated in Figure 3.
The algorithm in Figure 3 can be described in the following steps:

- First the image is separated into pieces.
  The sample image given in Figure 4a is separated into pieces as shown in Figure 4b.

- A training set is formed from each noise-free piece for classification.
  It should be noted that the number of pieces that the image will be separated into and the number of samples that will be taken from each piece for the training set depend on the designer. Although the sensitivity of the filter can be improved by separating the image into more pieces or taking more samples from these pieces for the training set, it will cause the calculation time to increase drastically, which is shown in Figure 5.

- The noisy pixel is selected in a $3 \times 3$ window.
• Then the probabilities of the neighboring noise-free pixels that will replace the noisy pixel are calculated. A noisy pixel can take one of the 256 different values in the range [0–255]. However, calculating the probability for each of these 256 values increases the processing time. Therefore, only the probabilities of 8 neighboring pixels are calculated. The probabilities of the remaining pixels are equal and lower when compared with the neighboring pixels. Thus, hundreds of nonneighboring pixels with low probabilities are skipped and only neighboring pixels with high probabilities are calculated.

• Finally, the neighboring pixel with the highest probability is replaced with the noisy pixel. Iteration is carried out accordingly.

![Figure 4. The separation of the image into pieces in order to form the training sets: a) corrupted image, b) image separated into pieces.](image)

![Figure 5. PSNR and processing time plots for different pieces and samples of brain MRI corrupted with 40% random-valued noise: a) PSNR (dB) vs. number of samples, b) time (s) vs. number of samples](image)

3. Experimental results and discussion

The suggested NBCF method has been explained in the previous section. In this section, the popular filters used recently to remove impulse noise are compared with our suggested method. All experiments were applied
in MATLAB. The Lena and Brain MR [31] images, which are popular in the literature, have been used as the test images in our experiments. Each of the test images used in the experiments was divided into nine sections and tests were performed by taking 2500 samples from each section. Noisy images were enhanced using both traditional and modern methods for comparison. The restoration quality between the enhanced image and the original image was evaluated using the peak signal to noise ratio (PSNR) method and mean squared error (MSE) [32] criteria, which are well-known in the literature.

\[
MSE = \frac{\sum_{ij} (r_{ij} - x_{ij})^2}{M \times N} \tag{6}
\]

\[
PSNR = 10 \log_{10} \left( \frac{255^2}{MSE} \right) \tag{7}
\]

Here \( r_{i,j} \) is the original image, \( x_{i,j} \) is the enhanced image, and \( M \times N \) is the image size.

The visual performances of different filters applied to a brain MRI corrupted with 40% random-valued impulse noise can be seen in Figure 6. In Figures 7 and 8, the performances of these filters are plotted in PSNR and MSE values vs. different noise densities, respectively. Our filter has been tested especially with high noise

![Image of restoration results for brain MRI corrupted by 40% random-valued impulse noise](image-url)

**Figure 6.** Restoration results using various filters for brain MRI corrupted by 40% random-valued impulse noise: a) original image, b) corrupted image, c) DBA [9], d) NFT [25], e) IDBA [10], f) CWMF [21], g) AMF [22], h) DBSMF [23], i) SNFF [12], j) TLIDE [24], k) proposed NBCF.
rates where the noise densities in the plots range from 10% to 95% with an incremental step of 10%. The plots clearly show that our proposed filter outperforms the other filters by giving the highest PSNR and the lowest MSE value.

As a final illustration, Figure 9 is given for a subjective comparison of various filters performing the restoration process of the Lena image with 60% random-valued impulse noise applied in this case. The visual performance of our proposed filter is shown in Figure 9, where it can be easily seen that the details of the image have been preserved while almost all noisy pixels are removed.

Similarly, the plots in Figures 10 and 11 show that our suggested filter again achieves a significant performance. The proposed NBCF exhibits excellent visual performance compared to other well-known filters and these results are of high importance for impulse noise removal in images.
3.1. Computational cost

Brain MR and Lena images were used in the performance evaluation of our algorithm. The tests were performed on an Intel Core i3 4 GB RAM PC. The computational cost is calculated in terms of processing time, which is given in the Table. It can be seen that the processing time of the proposed NBCF method is lower than that of DBA [9], SNFF [12], CWMF [21], AMF [22], DBSMF [23], TLIDE [24], and NFT [25] and higher than that of IDBA [10].

<table>
<thead>
<tr>
<th>Image-noise density</th>
<th>Filters/runtime (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DBSMF</td>
</tr>
<tr>
<td>Brain MRI -40%</td>
<td>9.65</td>
</tr>
<tr>
<td>Lena image - 60%</td>
<td>9.83</td>
</tr>
</tbody>
</table>

Table. Comparison of computational costs
Figure 10. PSNR (dB) values for different filters on the Lena image.

Figure 11. MSE values for different filters on the Lena image.

At the same time, the processing time of the compared algorithms for Brain MRI with different noise densities is plotted in Figure 12. The proposed filter is seen to be the fastest in terms of processing time among others, excluding IDBA.

4. Conclusions

In a digital image, it is very important to determine an impulse noise as well as to remove the impulse noise. In this paper, the suggested detector determines noisy and noise-free pixels and then a training set is composed of noise-free pixels. Then the noisy pixels are restored using the proposed NBCF method instead of calculating the median values of the neighboring pixels. The NBCF method has been presented for the first time in this paper and is a good alternative among other methods that have been proposed recently to remove impulse noise. Beside its simplicity and success, the method has a disadvantage. The image is separated into pieces and a learning set is prepared using the noise-free pixels of each piece, which increases processing time. The NBCF method has shown an excellent visual performance, not only for lower noise rate but also for high noise.
rate. It also performs well in most regions including noisy edge areas. The NBCF method only requires an
easy probability calculation and can be successfully applied to all image types with high noise rates without
changing any parameters.

Expanded simulation and experimental results have shown that the suggested algorithm suppresses noise
as much as the best known methods with a high PSNR performance and a low MSE value, and it has been seen
that the details of the image are perfectly preserved.

Acknowledgment

The authors would like to thank three anonymous reviewers for their constructive and helpful comments, which
improved this paper greatly.

References

2010; 20: 1010-1018.
17: 587-590.
[7] Ng PE, Ma KK. A switching median filter with boundary discriminative noise detection for extremely corrupted


The USC-SIPI Image Database, University of Southern California [Online], October 2012.