

International Journal of Computer Science and Mobile Computing



A Monthly Journal of Computer Science and Information Technology

ISSN 2320-088X

IJCSMC, Vol. 4, Issue. 7, July 2015, pg.412 – 418

RESEARCH ARTICLE

ENHANCED DATA MINING USING FUZZY, IMPROVED BILATERAL AND DECISION BASED ALPHA TRIMMED MEDIAN FILTERS

Gagandeep Kaur¹, Er.Gursimranjeet Kaur²

¹Department ECE & PTU, INDIA

²Department ECE & PTU, INDIA

¹ deepgagandeepkaur@yahoo.in; ² gursimranvirk@yahoo.com

Abstract— Many techniques continue to be proposed so far to get rid of the noise through digital images with more optimistic method. Each technique has its very own drawbacks. Although Fuzzy Mean Median (FMM) has demonstrated promising results on the available techniques, given it utilizes the top features of data mining to get rid of mixed noise. This data mining method is employed to check which sort of noise occurs in the picture. But it has not yet considered the enhancements with the filtering techniques i. e. improved bilateral filtration system and decision based alpha trimmed median filtration system. Therefore this perform has proposed some sort of novel data mining and improved bilateral and also decision based alpha trimmed based filtration system. The comparison offers clearly shown that the proposed technique outperforms on the available one.

Keywords—Data mining, mean filter, median filter, FMM filter, alpha trimmed median filter, improved bilateral filter

I. INTRODUCTION

This Data mining is simply a powerful brand-new technology with amazing potential to generally help big companies target the absolute vital information from the details they've collected regarding the behavior of the customers can use and potential customers. It discovers details inside information that will questions and accounts can't effectively show.

Generally, data mining would be the technique of studying data from diverse perspectives and outlining it into beneficial information - information that enables you to improve revenue, reductions costs, or each. Data mining software program is obviously on the list of several analytical tools intended for analyzing data. It allows users to evaluate data from many different dimensions or facets, categorize it, and also summarize the associations identified. Technically, data mining would be the technique of obtaining correlations or patterns among an enormous amount of career fields in large relational listings [1].

Although this data mining remains with its infancy, companies in the wide selection associated with industries - such as retail, finance, health care, manufacturing transportation, and aerospace - are already utilizing data mining tools and techniques to get the most from historical data. With the use of pattern recognition engineering and statistical and mathematical processes to sift through warehoused info, data mining facilitates analysts recognize important facts, relationships, developments, patterns, exceptions and anomalies that could otherwise go unnoticed [2].

Data mining is definitely the practice of right away searching outlets of data for getting patterns and developments that exceed uncomplicated analysis. Data mining employs sophisticated mathematical algorithms

to segment the info and evaluate the exact probability of future events [3]. Data mining is also known as Knowledge Discovery in Data (KDD).

II. THE QUITE PROPERTIES OF THE DATA MINING TECHNIQUES

The Quite properties of the data mining techniques are as follows:

A. Programmed Finding of designs

Data mining is accomplished since they build models [4]. A model uses an algorithm to do something on a number of data. The notion linked with automatic discovery would be the term for the execution linked with data mining sorts. Data mining models permits you to mine the data where they may be constructed, but most types of models are common to new info. The process linked with applying a kind to new data is regarded as scoring.

B. Prediction on most likely outcomes

Many forms linked to data mining are usually predictive. For illustration, a model might actually predict income based on education and a few other demographic factors [5]. Predictions produce a good associated probability (How possible is this prediction for being true?). Prediction probabilities are called confidence (How assured may i be of these kinds of prediction?).

Some forms of predictive data seek generate rules, which is often conditions that mean settled outcome. As an illustration, a rule might specify that the person who includes a bachelor's degree and lives in a really certain neighborhood could have an income greater than the regional widespread. Rules have the actual associated support (What percentage on the population satisfies the exact rule?).

C. Creation linked with actionable information

Other forms concerning data mining understand natural groupings from the data. For illustration, a model might actually identify the segment on the population which has a income within the particular specified range, that particular features includes good driving a car report, and that leases a brand new car on the yearly basis [6].

D. Focus on significant data units and also databases.

Data mining may possibly derive actionable details from large quantities of data. As an example, a town planner might try a model that predicts income good demographics to create a plan for low-income property. A car letting agency might use model that will probably identifies customer segments to make a promotion emphasizing on high-value customers.

III. MIXED NOISE MODELS

A number of different types connected with mixed noise design [7] is commonly considered within the literature, such seeing that, blur and Gaussian (or impulse) noises; Poisson additionally Gaussian noise as well as Gaussian plus impulse noise. As noise is usually regarded as stochastic along with uncertain error, probability density function (PDF) may be used to explain noise. Gaussian noise is the most prevalent noise type with nature, Whose PDF (F) of a Gaussian random variable (n) is explained as:

$$F(n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(n-\varphi)^2}{2\sigma^2}}$$

Where φ the mean or (average) value of Gaussian is noise (n) and σ is the standard deviation of Gaussian noise. Generally, impulse noise suggests the noise pixel of image using the two kinds connected with extremeness gray beliefs, with gray value 0 representing light and gray value 255 representing dark-coloured. When da along with db are non-zero along with approximately equal, the particular impulse noise appears to be the pepper-and-salt particulates circulating in image randomly. Regarding 8 bit gray image with 256 gray grades, the PDF (D) associated with salt-and-pepper noise can be described as :

$$D(n) = \begin{cases} da, & m = 0 \\ db, & m = 255 \\ 0, & \text{otherwise} \end{cases}$$

Many images, for example those from radiography, comprise noise that complies with a Poisson distribution. The PDF with the Poisson distribution is given as:

$$D(s) = \frac{m^s}{s!} e^{-m}$$

where s is the quantity of occurrences of an events and m will be the expected number of occurrences within a given interval. In the particular literature, several studies introduced and discussed that this Poisson distribution approaches a Gaussian thickness function in the case of high volume of counts. Moreover, Miller in showed that this Gaussian approximation is actually surprisingly accurate, even for a fairly few counts.

Although varieties of mixed noise models are offered, it is quite challenging to inform exactly which specific form of mixed noise model fits specific loud images. Theoretically a mixed noise model could be summing upward independent types connected with basis noise designs. As a way to obtain better denoising effectiveness, specific filter will be needed for a certain mixed noise model. In the subsequent, we will give attention to the specific graphic with noise consists of Gaussian and salt-and-pepper noise. Image restoration is definitely an important task with image processing. The final idea is to estimate a great image I_{org} in the observed noisy graphic I_{imp} . Here, we assume that this observed noisy graphic is obtained by using the mixed noise model:

$$I_{imp} = D_{image}(I_{org} + G')$$

where G' is an additive Gaussian noise; D_{image} indicates the image destruction by salt-and-pepper noise.

IV. DETECTION OF NOISE BY USING FUZZY RULE

Focusing on the $G(1)$ rulebase, i.e. the group of rules having N_1 as the consequent, we conclude that the noise type is Gaussian noise. Similarly, if we obtain the $G(2)$ rulebase, i.e. the group of rules having N_2 as the consequent, the noise type is generally considered as salt and pepper noise and for $G(3)$ rulebase, the group of rules having N_3 as the consequent, then the noise type is Poisson noise and for $G(4)$ rulebase, the group of rules having N_4 as the consequent, then the noise type is speckle noise.

$$\begin{aligned} G(1) &= \text{IF } (X, M_i) \text{ THEN } (Y, N_1), \quad i=1,2,3 \\ G(2) &= \text{IF } (X, M_i) \text{ THEN } (Y, N_2), \quad i=1,2,3 \\ G(3) &= \text{IF } (X, M_i) \text{ THEN } (Y, N_3), \quad i=1,2,3 \\ G(4) &= \text{IF } (X, M_i) \text{ THEN } (Y, N_4), \quad i=1,2,3 \end{aligned}$$

When the consequents of the fuzzy rules are not identical, i.e., various combinations of $[N_1, N_2, N_3, N_4]$ {except $G(1), G(2), G(3), G(4)$ } appear in the THEN part, we define the type of noise as a mixed one. In such a case, the proposed fuzzy mean-median filter, improved bilateral filter as well as DBMF filter will be applied to remove all noises from images.

$$G(5) = \{\text{IF } (X, \sum_{i=1}^4 M_i) \text{ THEN } (Y, \sum_{i=1}^4 N_i), \} \\ \{\text{Here } i=1, 2, 3, 4.\}$$

V. PERFORMNACE EVLAUTAION

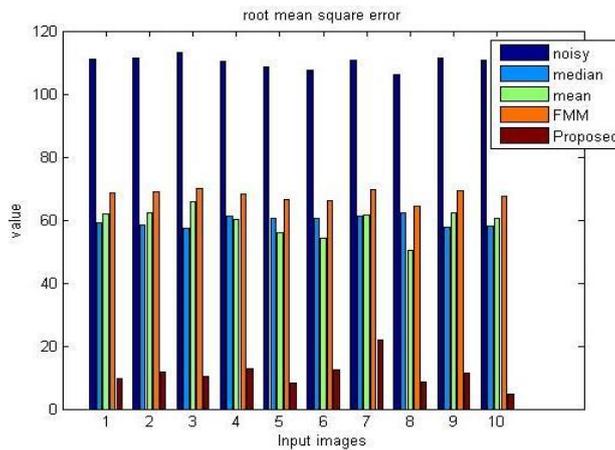
A. RMSE (Root Mean Square Error)

Root-mean-square error is a measure of the differences between values predicted by a model or an estimator and the values actually observed.

Table 1: This RMSE comparison table shows that the values of the noisy image and various filters such as median filter, mean filter, FMM (fuzzy-mean-median filter) filter and finally proposed filters (i.e. bilateral filter, DBMF and FMM filter).The values of proposed filters should be as low as possible than the existing filters.

Table 1: RMSE values of existing filters and proposed filters

Images	Noisy	Median	Mean	FMM	Proposed
1.	111.0946	59.0931	61.9758	68.6174	9.8124
2.	111.3149	58.4466	62.1932	68.8513	11.7649
3.	113.0221	57.2014	65.8635	70.1612	10.3443
4.	110.3268	61.1310	60.0999	68.4054	12.6879
5.	108.7704	60.3573	56.0625	66.5946	8.2873
6.	107.5965	60.4483	54.1018	66.0382	12.3478
7.	110.8197	61.1392	61.7495	69.7133	21.8154
8.	106.0896	62.3217	50.2991	64.5233	8.6463
9.	111.3104	57.8878	62.2334	69.1326	11.4346
10.	110.6571	58.2065	60.5805	67.6375	4.7553



Graph 1: RMSE graph show results of existing filters and proposed filters

Graph 1: This graph shows that the values of RMSE of proposed work as well as existing work. The values of RMSE of the proposed work show better results than the existing work because the value of RMSE is as low as possible.

B. Structure context

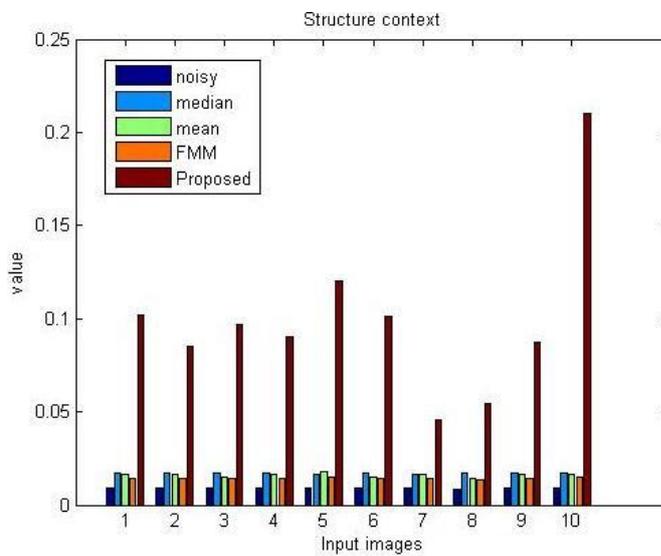
This measure effectively compares the total weight of an original signal to that of a coded or given. It is a global metric; localized distortions are missed. This measure is called as structural content and large value of structure context means that the image is of poor quality.

Table 2: This SC comparison table shows that the values of the noisy image and various filters such as median filter, mean filter, FMM (fuzzy-mean-median filter) filter and finally proposed filters (i.e. bilateral filter, DBMF and FMM filter).The values of proposed filters should be as high as possible than the existing filters.

Table 2: SC values of existing filters and proposed filters

Images	Noisy	Median	mean	FMM	Proposed
1.	0.0090	0.0169	0.0162	0.0145	0.1017
2.	0.0090	0.0172	0.0161	0.0145	0.0850

3.	0.0088	0.0175	0.0151	0.0142	0.0967
4.	0.0090	0.0172	0.0162	0.0145	0.0904
5.	0.0092	0.0166	0.0179	0.0150	0.1206
6.	0.0088	0.0175	0.0151	0.0142	0.1013
7.	0.0090	0.0163	0.0162	0.0144	0.0458
8.	0.0087	0.0173	0.0141	0.0138	0.0547
9.	0.0090	0.0173	0.0161	0.0145	0.0875
10.	0.0090	0.0169	0.0164	0.0147	0.2102



Graph 2: SC graph show results of existing filters and proposed filters

Graph2: This graph shows that the values of SC of proposed work as well as existing work. The values of SC of the proposed work show better results than the existing work because the value of SC is as high as possible.

C. NCC (normalized cross co-relation)

These measures measure the similarity between the two different images, hence, they are complementary to the difference based measures. As difference measure and correlation measures complement each other, minimizing Distance measures are maximizing correlation measure.

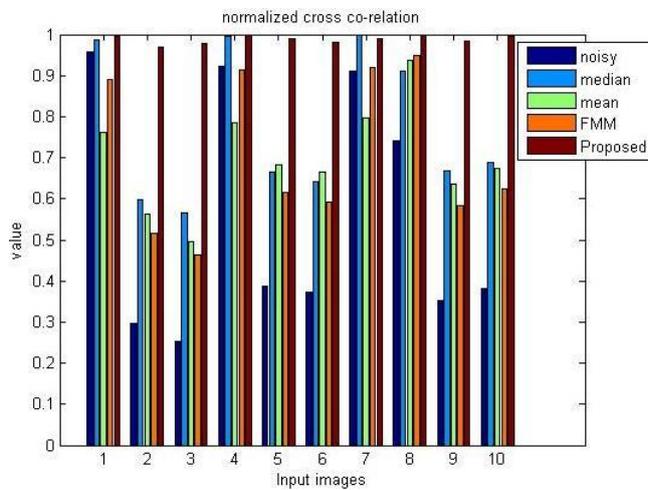
$$NCC = \frac{\sum \sum [I(i, j) * o(i, j)]}{\sum \sum [I(i, j) * I(i, j)]}$$

Table 3: This NCC comparison table shows that the values of the noisy image and various filters such as median filter, mean filter, FMM (fuzzy-mean-median filter) filter and finally proposed filters (i.e. bilateral filter, DBMF and FMM filter).The values of proposed filters should be as high as possible than the existing filters.

Table 3: NCC values of existing filters and proposed filters

Images	Noisy	Median	Mean	FMM	Proposed
1.	0.9584	0.986	0.7627	0.8913	0.9966
2.	0.2981	0.5982	0.5618	0.5152	0.9681

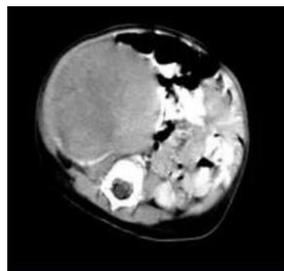
3.	0.2539	0.5671	0.4954	0.4646	0.9775
4.	0.9221	0.9945	0.7858	0.9138	0.9975
5.	0.3883	0.6652	0.6832	0.6159	0.9891
6.	0.3733	0.6412	0.6660	0.5914	0.9818
7.	0.9112	0.9991	0.7956	0.9189	0.9891
8.	0.7419	0.9102	0.9375	0.9502	0.9974
9.	0.3512	0.6690	0.6350	0.5846	0.9827
10.	0.3817	0.6884	0.6737	0.6238	0.9948



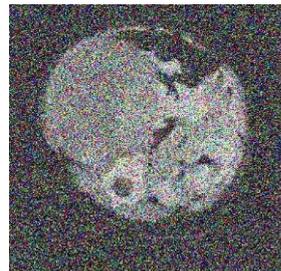
Graph 3: NCC graph show results of existing filters and proposed filters

Graph 3: This graph shows that the values of NCC of proposed work as well as existing work. The values of NCC of the proposed work show better results than the existing work because the value of NCC is as high as possible.

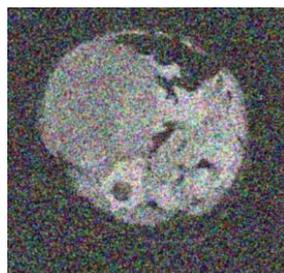
VII. THE SIMULATION RESULTS



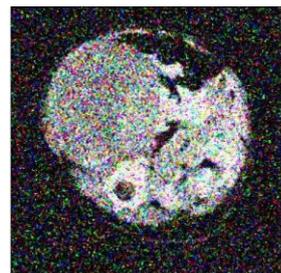
(a) Input Image



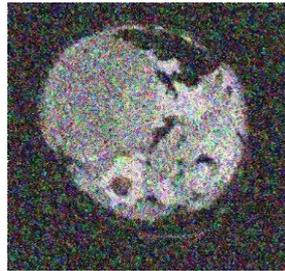
(b) Noisy Image



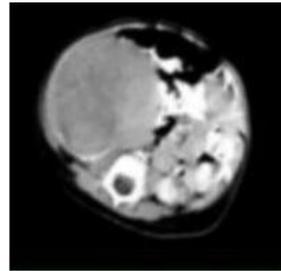
(c) Mean Filter



(d) Median Filter



(e) FMM Filter



(f) Proposed Filter

This simulation results shows better quality of the image by using proposed filters (i.e. FMM Filter, improved bilateral filter and decision based alpha trimmed median filter).

VIII. CONCLUSIONS

This Fuzzy Mean Median (FMM) has demonstrated promising results above the existing techniques, since it utilizes the popular features of data mining to get rid of mixed noise. Nevertheless it has not regarded the enhancements on the filtering techniques (i. e. improved bilateral filtration system and decision primarily based alpha trimmed median filter. Therefore this function has proposed any novel data mining and improved bilateral in addition to decision based alpha trimmed based filtration system. The proposed technique have been designed and implemented inside MATLAB using image processing toolbox. Various experiments have been done by considering various sort of images. The comparison features clearly shown which the proposed technique outperforms above the available one.

REFERENCES

- [1] Ou, Zenggui. "Data structuring and effective retrieval in the mining of web sequential characteristic." In Electronic and Mechanical Engineering and Information Technology (EMEIT), 2011 International Conference on, vol. 7, pp. 3551-3554. IEEE, 2011.
- [2] Zhu, Xiaodong, Hengshan Wang, Hongcheng Gan, and Chunchang Gao. "Construction and management of automatic reasoning supported data mining metadata." In Business Management and Electronic Information (BMEI), 2011 International Conference on, vol. 5, pp. 205-210. IEEE, 2011.
- [3] Farid, Dewan Md, and Chowdhury Mofizur Rahman. "Novel class detection in concept-drifting data stream mining employing decision tree." In Electrical & Computer Engineering (ICECE), 2012 7th International Conference on, pp. 630-633. IEEE, 2012.
- [4] Cao, Longbing. "Social security and social welfare data mining: An Overview "Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 42, no. 6 (2012): 837-853.
- [5] Madraky, Abbas, Zulaiha Ali Othman, and Abdul Razak Hamdan. "Hair data model: A new data model for Spatio-Temporal data mining." In Data Mining and Optimization, 2012 4th Conference on, pp. 18-22. IEEE, 2012.
- [6] Farid, Dewan Md, and Chowdhury Mofizur Rahman. "Novel class detection in concept-drifting data stream mining employing decision tree." In Electrical & Computer Engineering, 2012 7th International Conference on, pp. 630-633. IEEE, 2012.
- [7] Wang, Yongfu, Gaochang Wu, Gang Sheng Chen, and Tianyou Chai. "Data mining based noise diagnosis and fuzzy filter design for image processing." *Computers & Electrical Engineering* 40, no. 7 (2014): 2038-2049.