Contents lists available at http://qu.edu.iq

Al-Qadisiyah Journal for Engineering Sciences

Journal homepage: https://qjes.qu.edu.iq

Enhance watershed algorithms using principal component analysis capabilities

Ahmed Fathy Ahmed Mohamed*, Fayez Wanis Zaki Solima, and Mohamed M Ashour

Electronics and Communications Engineering Department, Faculty of Engineering, Mansour University, Dakhlia, Egypt

ARTICLE INFO

Article history:
Received 05 December 2023
Received in revised form 15 January 2024
Accepted 20 February 2024

Keywords:
Image processing
Image segmentation
Principal component analysis
Watershed segmentation
Watershed algorithm
Over-segmentation

ABSTRACT

Traditional image segmentation algorithms have many drawbacks, such as over-segmentation and image distortion due to reflected light. The Watershed algorithm is one of the most popular image segmentation algorithms. Over-segmentation errors caused by overlapping targets in the image, as well as noise and glare, must be removed. In this article, we apply image processing using the watershed algorithm and propose to improve the algorithm based on principal component analysis. PCA is a popular technique for analyzing large datasets with many advantages per observation. PCA improves data interpretability while maximizing information content, enabling visualization of multidimensional data by finding image component gradients in a new space called the principle component that is unaffected by noise and reflected light. In contrast, the components mainly containing noise will eliminate with negligible information. This paper introduces three primary steps. The process involves applying the watershed algorithm to the image in the first phase, using the proposed approach (applying the watershed algorithm and suggesting an improvement based on principal component analysis) to the image in the second step, and comparing the outcomes of the two previous processes. Test results show that the suggested technique can achieve accurate and durable target shapes.

1. Introduction

Image processing plays an important role in many fields**,** such as image mining, medical image processing [1], and web mining [2-3]. Medical image analysis, in particular, is an area where Development in Image processing can improve the quality and speed of research. Image processing focuses on his two main areas: image enhancement and image analysis. Image enhancement techniques include a large number of literature searches discover a variety of applications with clear recognition of the need for improvement. It may also be helpful to analyze the image to better understand it. Among other things, techniques such as clustering and classification are considered. With recent technological advances, there are many applications of image processing, from ordinary image processing to image analysis.

However, Image segmentation can be thought of as dividing an image into regions with similar characteristics or as dividing an image into semantically meaningful parts. Techniques and Challenges of Image Segmentation [4]. The development of image segmentation technology is closely related to many fields and disciplines, e.g. autonomous vehicles [5], intelligent medical technology [6-7], image search engines [8], industrial inspection, and augmented reality.

Image processing focuses on two basic areas: image enhancement and image analysis. Image enhancement methods include various literature reviews highlighting various applications that clearly indicate the need for further development. Analyzing images to better represent them may also be helpful. Due to recent advances in technology, the application fields of

* *Corresponding author.*
E-mail address: ahmedfathyamer14@gmail.com (Ahmed Mohamed)

https://doi.org/10.30772/qjes.2024.145116.1055

Nomenclature:

2D	two dimensional	*PSNR*	Peak Signal-To-Noise & Ratio
3D	three dimensional	*SNR*	Signal-To-Noise Ratio
IMMSE	Image Mean-Squared Error	*SSIM*	Structural Similarity Index for measuring image quality Mean-Squared Error
LPG	local pixel grouping	*WA*	Watershed algorithm
MRI	Magnetic Resonance Imaging	*WSA*	Watershed Segmentation Algorithm
PCA	principle component analysis	*WSA-PCA*	Watershed Segmentation Algorithm pre-processed with Principal Component Analysis

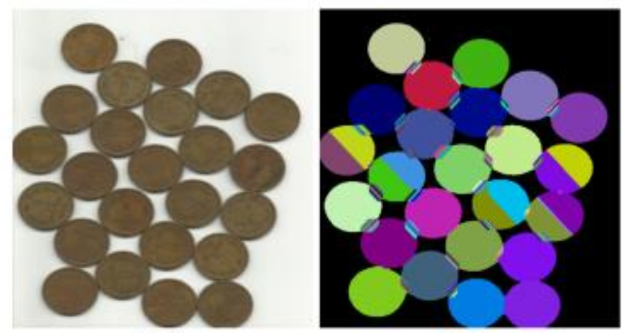
image processing are wide-ranging, from general image processing to image analysis. Regardless, image segmentation can be thought of as dividing an image into locations with comparable characteristics (color, surface, shape, etc.) or dividing an image into semantically significant parts.

1.1. Related work

No algorithm looks good for all types of images, nor are all algorithms equally good for a particular type of image. The image segmentation process remains a difficult and unsolved problem in image processing and computer vision. Many studies are comparing important image segmentation techniques, the most relevant works are: In [9], this article presents an algorithm for watershed color image segmentation based on morphological gradients. An improved watershed color image segmentation algorithm based on morphological gradients is proposed. In [10], "Review of Watershed Algorithm Implementations in Open Source Libraries", describes benchmark results of his six marker-driven open-source Watershed implementations for 2D and 3D image segmentation. In, [11], this article reviews the advancement in image segmentation methods systematically. According to the principles of segmentation and characteristics of image data, we mainly review the three important phases of image segmentation, namely classical segmentation, collaborative segmentation, and deep learning-based semantic segmentation. In, [12], in this article, we review various digital image processing techniques and their applications, such as remote sensing, medical image processing, and forensic research. In [13], this article gives reviews image segmentation techniques. In [14], this article describes the extraction of brain tumors from his MRI images using prominent image segmentation method in [15]. This work presents a hybrid approach to image region extraction, focusing on automated region proposal and segmentation techniques. In [16], this article explains that PCA is a statistical technique that simplifies data sets by reducing them to lower dimensions. PCA-based image noise reduction. This paper, proposes a noise reduction method using a new statistical approach called principal component analysis with local pixel grouping (LPG), repeats this procedure twice to further improve the noise reduction performance, and adaptively adjusts the noise level in the second stage.

1.2 Problem statement

For some images, it is not possible to set segmentation process parameters, such as thresholds, to extract all objects of interest from the background or each other without over-segmenting the data. To understand over-segmentation, assume that an object is perceived by neighboring objects in the background. It may be split into many sub-objects, making parts of the image unrecognizable. Over-segmentation increases the chance of extracting important contours, but this comes at the cost of producing many unimportant contours.



(a) (b)

Figure 1. Example of Over-segmentation Problem: (a) image; (b)segments.

In this case, pre-processing techniques should be used to remove noise, improve sharpness between subjects, and smooth image texture, which can make segmentation difficult. If these techniques are not sufficient, hyper-segmentation can be used as a preliminary processing step, followed by a clustering step to reconstruct objects into individual image events (objects). Each segment derived from an image is called an object, and the properties of the segment are called attributes or features. Applying standard watershed algorithms to real images results in over-segmentation, [17]. This occurs because, in the standard approach, the introductory marker places all neighborhood minima of the image. Any one of the local components, e.g. all neighborhood minima, could be the center of the catchment, but not all these minima are equally convincing. For example, some may be distorted by noise, slight changes in brightness, or auxiliary structures in the image [18]. Figure 1 shows an example of over segmentation problem.

2. Materials and methods

2.1 Watershed algorithm

Watershed segmentation belongs to the category of regional similarity. Watershed segmentation is a mathematical-morphological approach that draws similarities with real-life flood situations [19]. Watershed Segmentation [20-22], is an image segmentation algorithm that combines ideas about topography and regional growth. This algorithm displays the grayscale image as a topographic map, as shown in Fig. 2, [23]. Pixel areas with high grayscale values Represent Mountains, and pixel areas with low grayscale values represent lowlands. Once the contours are modeled by the watershed algorithm, water flows down the mountainside, forming a "lake". This "lake" in the photo is called the basin. When water levels rise in one catchment, water can flood other nearby catchments. If we build dams at the intersections of each basin, water will never overflow. These dam locations represent watersheds, which is a desirable outcome for image segmentation. Compared to some classical edge detection algorithms (such as Sobel, Canny, and other operators), [24] Watershed's algorithm has the advantages of low cost and high computational accuracy. It takes the gradient of an image as input and outputs a continuous edge line of width 1 pixel [25]. However, due to gradient noise, quantization errors, and subtle textures within the object, flat regions can often have many local "valleys" and "mountains". After the deformation of the basin, small spaces are formed and over segmentation occurs.

The performance of image segmentation methods based on watershed transformations is highly dependent on the algorithm used to compute the gradients of the segmented images. For watersheds, the ideal output of the gradient operator should be equal to the height of the input edge, or the difference between the pixel grays on either side of the edge, rather than the slope of the edge. Natural images rarely have ideal step edges, and typically have diagonal edges with relatively blurred edges. The traditional gradient operator returns the edge gradient for an edge of this type [26].

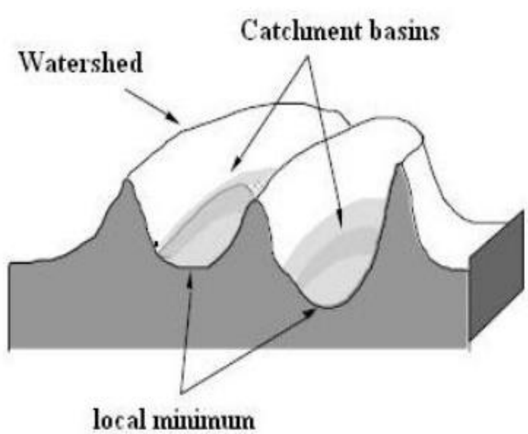


Figure 2. Modelling of contours by watershed algorithm

Segmenting an image according to the watershed algorithm can be considered a two-step process. First finds the watershed and then uses a fixed complement to find the watershed. Here we mainly focus on watershed segmentation as flood simulation. If you want to flood the ground from a minimum and prevent water from different sources from collecting, divide the image into two different groups: basins and watersheds, as shown in Fig. 3.

An algorithmic definition for watershed transformation with simulated flooding was given by Vincent and Soille [27].

Let $f: D \rightarrow N$ be a digital grayscale image, where h_{min} and h_{max} are the minimum and maximum values of f. Define a recursion in which the gray level h increases from h_{min} to h_{max}. In this recursion, the basin connected to the minimum value of f is continuously expanded. Let Xh represent the union of basin sets computed at level h. The connected component of the threshold set $Th+1$ at level $h+1$ is either a new minimum or an extension of the basin in Xh. In the latter case, the geodesic influence zone IZTh+1(Th) of Xh within Th+1, updated will be Xh+1. Let h_{min} represent the union of all area minima of height h.

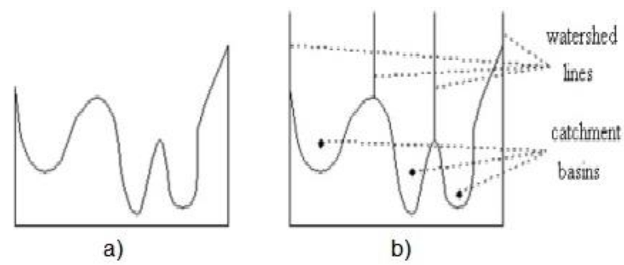


Figure 3. One-dimensional example of watershed segmentation. a) Gray level profile of image data, b) Watershed segmentation local minima gray level (altitude) yield catchment basins and the watershed lines.

While, if the Xh+1, then h_{min} represent the union of all area minima of height h.

Definition 1: Watershed by flooding.

$Xh_{min} = P \in D \; f_p = h_{min} = Th_{min}$

$Xh + 1 = h_{min} \cup IZTh + Th, \qquad h \in h_{min}, h_{max}$

The watershed f is a functin of Xh_{max} and DW_{shed} , $f = D/X \; h_{max}$

According to definition 1, pixels with grey value h' ⩽ h that are not yet part of a basin after processing level h is merged with the basin at the higher level h + 1. Pixels that are the same distance apart in a given iteration can be provisionally labeled as "watershed pixels" by assigning the label W to at least the two closest basins. An example of watershed transformation using flood simulation is shown in Figs 4. A, B, C, and D. The watershed label W is used to indicate the pixel of the watershed point (the smallest pixel in the input image is shown in bold as a marker). Note the dependence on connectivity. The example shown in Fig. 4 is of individual 7x7 images on a square grid with 8 connections. Because there are four local minima (the zeroes), there are four basins with pixels labeled A, B, C, and D.

2.2 Principal component analysis (PCA)

Principal component analysis is a technique that reduces a case-by-variable data table to its essential features, called principal components. The principal components are linear combinations of the original variables that maximally explain the variance of all variables [28]. The new variables have the property that all variables are orthogonal. Principal component analysis highlights are valuable as a preprocessing step for the watershed algorithms Fig. 5. Principal Component Analysis PCA is a variance-focused approach that aims to reproduce the total variance of variables [29]. In this case, the components reflect both the common and Eigen variances of the variables. The goal of reducing noise using principal component analysis is to remove noise while preserving as many of the important characteristics of the signal or image as possible [30]. Capturing and transmitting images often introduces additional noise. The main purpose of using principal component analysis algorithms is to reduce the noise level while preserving the image characteristics. To achieve good performance here, the component analysis algorithm must first adapt to the image discontinuities. For this reason, the algorithms in this article are more suitable for watershed algorithms.

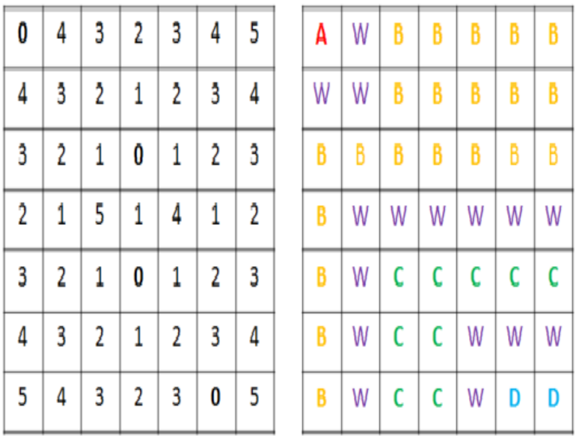


Figure 4. Watershed transform on the square grid with 8-connectivity, showing thick Watersheds. (a): original image; (b): result according to flooding

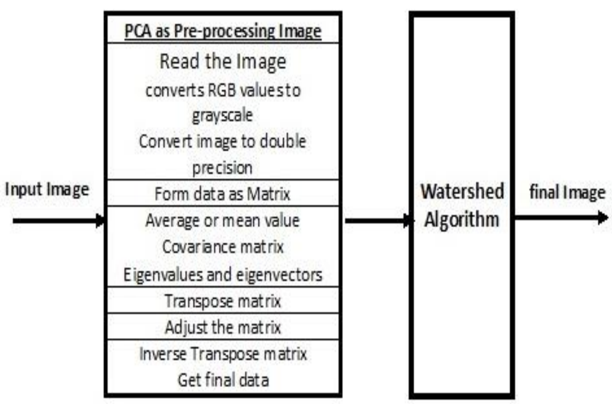


Figure 5. The principal component analysis is a pre-processing step before the watershed Algorithm.

The algorithm in this paper uses Principal Component Analysis (PCA) to convert the image from RGB to a new space, allowing you to avoid reflected light affecting the image. As shown in Fig. 6, the principal component analysis algorithm consists of two steps. The first step removes most of the noise and makes an initial estimate of the image, and the second step further refines the output of the first step [31]. Therefore, the de-noising step can be performed in the second round to improve the de-noising results.

Result: PCA New Reduced Coordinated: reduced
Form-> Data As Matrix (ataset, datamatrix)
Adjust-> MatrixToMean (datamatrix, meanadjusted)
Transpose Matrix (meanadjusted,)
Mult (meanadjusted, meanadjustedtransposed, correlation)
Eigenvectors -> And -> Eigenvalues (correlation, newcoords, eigenvalues)
Filter -> Out -> Unused Coordinated (newcoords, eigenvalues, reduced)
Return reduced

Figure 6. Two stage principal component analysis

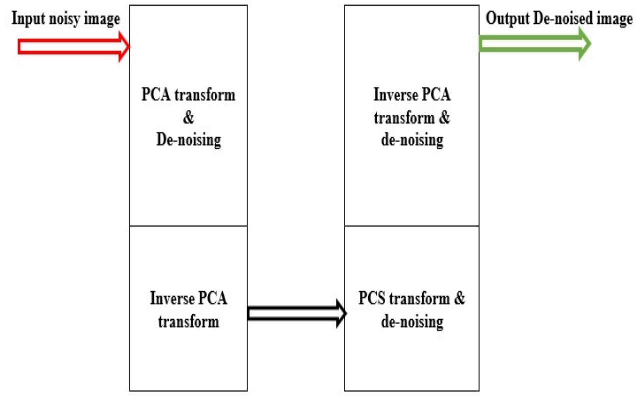


Figure 7. Pseudo code for watershed boundary finding algorithm

In this work, the principal components were analyzed into a region-based watershed algorithm. Based on this, this work proposes a gradient "peak enhancement" method to correct the weakened edge contours, Fig. 7.

2.3 Proposed method (watershed algorithm enhanced with principal component analysis)

There are two main steps as shown in Fig. 8, for the proposed approach of the watershed algorithm. The first stage is defined as the preprocessing stage and marked with the blue rectangular, here the principal component analysis was used as the preprocessing stage, (PCA) principal component analysis takes the input image and converts it to grayscale then converts it to double, this conversion determine the intensity of each pixel in the image then applies the maximum inter-class maximum variance algorithm to automatically threshold the final gradient image to make it ready to determine the eigenvectors and eigenvalues after taking the transpose, that tell us the degree of variation in the direction of its respective eigenvector. The eigenvector with the largest eigenvalue is the direction with the greatest variation, we call this eigenvector the first principal component. The first principal component helps to determine the most important data pixels in the image which are unaffected by noise and reflected light and helps to eliminate the component contains mostly noise with negligible information in them. Finally, we take the inverse transpose to get the final image from the treated mirror image that helps to improve the quality of the final image. The second stage is the watershed algorithm, marked by the red rectangle in below Fig. 8. In this step, the watershed algorithm receives the adjusted image from the previous stage and processes the noisy images using principal component analysis. Here, the watershed Algorithm begins searching for a safe background using morphological operations such as opening and dilation. Next, use a distance transform to find a safe foreground. We then take the unknown region, i.e. the region without foreground and background lines, and use it as a marker for the watershed algorithm. At long last compare the gotten comes about from the past two steps. Tests almost show up that exact and diligent target shapes can be fulfilled by the proposed methodology.

3. Results and discussion

3.1. Experimentation setup

The proposed technique was actualized in MATLAB computer program.

For this reason, MATLAB R2020b was utilized.

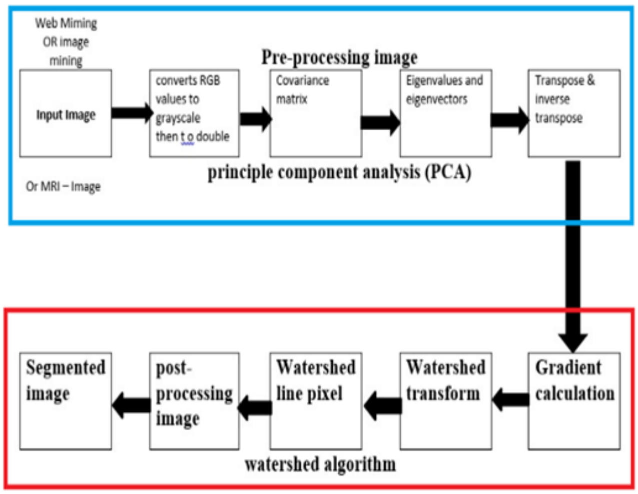


Figure 8. Block diagram of the proposed watershed algorithm

4. Results and discussion

4.1. Experimentation setup

The proposed technique was actualized in MATLAB computer program. For this reason, MATLAB R2020b was utilized. The image-handling toolkit was utilized to supply basic picture preparation capacities. The proposed demonstration was assessed by executing it in MATLAB. Furthermore, the obtained images were analyzed using MATLAB techniques to determine the SSIM (Structural Similarity Index for measuring image quality), IMMSE (Image Mean-Squared Error), PSNR (Peak Signal-To-Noise & Ratio), and SNR (Signal-To-Noise Ratio). The execution environment is displayed in Table 1.

Table 1. Implementation environment**.**

System specification	Values
Processor	Intel Core i5
Memory	4 Gb
MATLAB version	MATLAB R2020b
Operating system	Windows 10 (64-bit)

4.2. Performance evaluation

Over the past few decades, issues relating to image segmentation and methods of segmenting images, especially for the watershed algorithm have become issues of concern in Image processing, particularly with the emergence of a set of methodological perspectives and tools within the field of segmentation. Benefit researchers in the field of image segmentation and methods of applying; this contributes to closer links, and this was confirmed by the discussions presented by the scientific journals in this field and their scientific communities.

Actually, the reality refers to the lack of curing the over-segmentation problem, and therefore the current paper was to examine in depth and detail the nature of over-segmentation issues in watershed segmentation algorithm, in the light of research methodologies and approaches using the principal component analysis as a preprocessing step to the watershed algorithm. The article is based on designing a proposed method framework that reduces and removes image noise and avoids the effect of reflected light in the image that leads to over-segmentation and causes loss of some areas that may contain important targets in the image by using the principal component analysis as a preprocessing step before applying the watershed algorithm to obtain a final clean and clear segmented image.

4.2.1. The first example is a noisy coin image

The first example is a noisy coin image as shown in Fig. 9. This image contains a total of 10 targets represented by 10 coin images and is therefore carefully selected to demonstrate the advantages and performance of the proposed method. This image contains his two elements that are the main causes of over-segmentation, represented by the last two elements or targets #9 and #10. The 9th target shows an example of a target with reflected light (e.g. Targets exposed to overflight that make the target glare), the 10th and final target represents a target affected by noise and distortion due to poor lighting conditions and is very close to the background color. Use the image in Fig.10 and mark it as the original image. Applying the watershed algorithm as described in Watershed Segmentation in the previous related literature. This example shows how to use watershed segmentation to separate touching objects in an image. The watershed transforms find "catchment basins" and "watershed ridge lines" in an image by treating it as a surface where light pixels are high and dark pixels are low. Segmentation using the watershed transform works well if you can identify, or "mark," foreground objects and background locations. Marker-controlled watershed segmentation follows this basic procedure:

- *Compute a segmentation function*: This is an image whose dark regions are the objects you are trying to segment.
- *Compute foreground markers*: These are connected blobs of pixels within each of the objects.
- *Compute background markers*: These are pixels that are not part of any object.

Modify the segmentation function so that it only has minima at the foreground and background marker locations.

Compute the watershed transform of the modified segmentation function. First step the algorithm starts to Read the Image and Convert it to Grayscale.

The second step uses the Gradient Magnitude as the Segmentation Function, this will compute the gradient magnitude. The gradient is high at the borders of the objects and low (mostly) inside the objects. The third step marks the Foreground Objects, A variety of procedures could be applied here to find the foreground markers, which must be connected blobs of pixels inside each of the foreground objects. In this example, using morphological techniques called "open-ing-by-reconstruction" and "closing-by-reconstruction" to "clean" up the image. These operations will create a flat maximum inside each object Fig. 11. The number of targets in the original image (noisy coin image) is 10, and these targets are visible to normal human eyes. The obtained image from the Watershed Segmentation algorithm described in the previous related literature is shown in Fig.11. The obtained image does not recognize targets #9 and #10, only eight targets appeared in the Watershed image (Watershed's noisy coin image) as a result of the segmentation of the original image by the Watershed algorithm. In this case, the watershed algorithm will produce an inaccurate watershed image. This error is due to noise in the image that affects the percentage of gray levels and causes over-segmentation. Item #9 was not detected due to over-light exposure. Item #10 was not detected due to poor light conditions. Both objects produce noise and distortion regardless of whether there is too much or too little light.

Figure 9. Noisy coin image

Figure 10. Noisy coin image (original image).

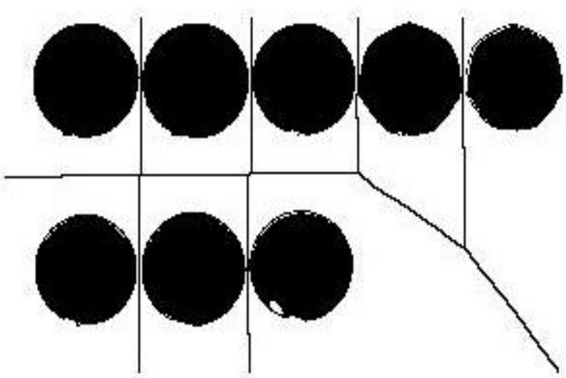


Figure 11. Watershed coin image

These noises and distortions caused disturbances in the basin lines, which caused missing targets in the segmented image (targets #9 and #10 are missing due to noise in the original image).

The focus is on solving this problem using principal component analysis (PCA). PCA is the central point through which prepared to see complex datasets in a more accessible way. (PCA) is a broadly utilized procedure in data examination and dimensionality reduction and also in noise reduction. It makes a difference in recognizing the preeminent essential plans and reducing the complexity of high-dimensional data by securing its principal data. Compress images using principal component analysis (PCA) to reduce the effects of noise and reduce dimensionality. To have a comprehensive understanding of how PCA works and how it benefits the analysis of multidimensional data and decreases the effects of noise and distortion. The solution will start by explaining the key steps included in PCA as follows:

Step 1: Convert the data to numeric information, as PCA is fundamentally suited for numerical information. Taking after, standardize the information. Standardization is vital since PCA is sensitive to the scales of factors. Standardization changes the information so that each variable highlights a cruel of and a standard deviation of 1. This ensures that all variables are on the same scale.

Step 2: Compute the covariance matrix of the standardized information. The covariance matrix represents the relationships between variables. The covariance matrix is a square matrix, with each component representing the covariance between two variables.

Step 3: Compute the Eigenvalue and Eigenvector, after computing the covariance matrix, the following step is to discover the Eigenvalues and eigenvectors of this matrix. These are crucial in deciding the principal components. This step creates various eigenvalues and eigenvectors as the number of factors within the information.

Step 4: Sorting and selecting principal components the eigenvalues represent the amount of variance within the data that each eigenvector explains. To reduce the dimensionality, sort the eigenvalues in descending order. The eigenvector corresponding to the largest eigenvalue explains the most variance and is the first principal component. The second largest eigenvalue compares to the second principal component, and so on. Ordinarily, selecting a subset of the top eigenvalues/eigenvectors that explain most of the variance within the data while reducing the dimensionality. Can decide on the number of principal components to keep based on a variance-explained threshold (e.g., 95% of the total variance).

Step 5: Data Transformation to reduce the dimensionality of the information while keeping the main data, create a projection matrix using the chosen eigenvectors (principal components). This matrix represents the transformation needed to project the information into the new reduced-dimensional space to get the new information within the principal component space.

Step 6: Interpretation and Examination Once the information is transformed, it can interpret the principal components and their relationships to the initial variables. This is pivotal for understanding the most significant patterns in the information. By holding the most significant variance and discarding less important variance, PCA can offer assistance reduce the impact of noise in the data.

Now, let's start preprocessing the image using principal component analysis as shown in Figs 12 a, b. The image in Fig 12-a shows a grayscale image of the original image (noisy coin image) before it was processed. The illustration in Fig12-b shows the grayscale image of the original image (noisy coin image) after it has been processed by principal components analysis. In image Fig 12-b, an illustration of the improvement due to our direct modeling of the principal components analysis. Principal component analysis shifts the image to a new space which improves the threshold line because the noise in the image has been reduced using principal component analysis and the grayscale in the image is approximated to the minimum by reducing dimensional properties of the principle component analysis

Figure 12-a. Grey scale coin image

Figure 12-b. Coin image preprocessing with principal component analysis

Now we apply the watershed algorithm after preprocessing the image with the principal component analysis function Fig. 13. The image in Fig. 13 shows the segmented image (noisy coin image) with the Watershed algorithm after applying Principle Component Analysis. The segmented image with the Watershed algorithm after applying Principle Component Analysis clearly shows the number of targets that completely match the number of targets in the original image (noisy coin image) both are ten targets. So far we have achieved the desired goal of this article, which is a comparison between two images that have been segmented by Watershed algorithm. The first image is directly segmented by Watershed algorithm. And the second image is segmented by Watershed algorithm after processing with principal component analysis. Principal component analysis preprocessing was a very successful one, as it facilitated the detection of all desired targets with high efficiency, due to the noise reduction characteristics, making the image clearer and more effective, as well as the direction of feature reduction, allowing to show the limits of the target more effectively.

4.2.1.1. Compare the results of the segmented images generated by the watershed`s segmentation algorithm and proposed algorithm (WSA-PCA).

This section explains the comparison between segmented images generated by the watershed`s segmentation algorithm and the proposed algorithm (WSA-PCA) based on pre-processing the images by principle Component Analysis before entering the watershed`s segmentation algorithm. MATLAB software is used to implement the two algorithms used to achieve the desired results, which will be discussed and analyzed in this section.

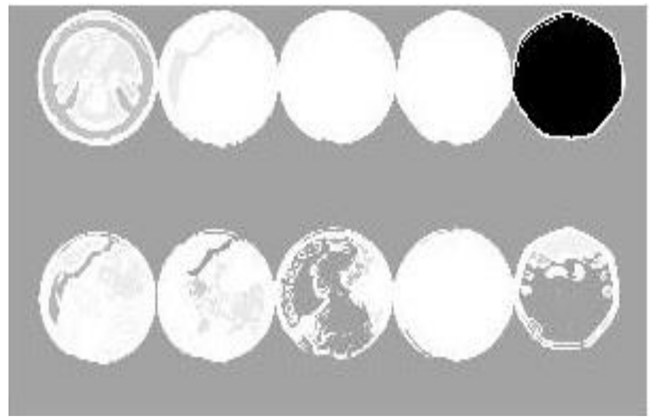


Figure 13. Watershed coin image after applying principle component analysis

Image segmentation results an amount of over-segmentation depend on the pre-processing stage. It is necessary to find out optimal solution for different types of images. Principle Component Analysis handles the over-segmentation problem by de-noising, repair and modification the given image. A set of different images are tested for segmentation qualities by applying the Principle Component Analysis techniques.

Below Table 2: shows change in SSIM (Structural Similarity Index for measuring image quality), IMMSE (Image Mean-Squared Error), PSNR (Peak Signal-To-Noise & Ratio) and SNR (Signal-To-Noise Ratio). Resultant Structural Similarity Index for measuring image quality The SSIM index is a decimal value from -1 to 1, where 1 indicates perfect similarity, 0 indicates no similarity, and -1 indicates perfect anti correlation. In this study the SSIM Value is 0.8307 for the image after applying the pre-processing stage with the Principle Component Analysis, this leads to SSIM Value 0.5699 in gradient image and produce SSIM Value 0.6760 in final watershed image. This SSIM values means the Proposed Method enhance the final segmented image.

The resulting mean squared error, MSE, is used to check how close the estimate or prediction is to the actual value. The lower the MSE, the closer the prediction is to the actual value. This is used as a model evaluation measure for regression models, with lower values indicating better fit. In this study the IMMSE error Value is 0.7*104 for the image after applying the pre-processing stage with the Principle Component Analysis, this leads to IMMSE error value 1.479*104 in gradient image and produce IMMSE error Value 1.008*104 in final watershed image. This IMMSE error values means the Proposed Method enhance the final segmented image. The resulting maximum signal-to-noise ratio (PSNR) is the ratio between the maximum possible performance of an image and the power of disturbing noise that affects the quality of its representation. To estimate her PSNR of an image, we need to compare that image to an ideal, clean image with the maximum possible performance. PSNR is most commonly used to estimate the efficiency of things like compressors and filters. The higher the PSNR value, the more efficient the corresponding compression or filtering method. In this study the PSNR Value is 19.6124 dB for the image after applying the pre-processing stage with the Principle Component Analysis, this leads to PSNR value 6.4298 dB in gradient image and produce PSNR Value 8.0956 dB in final watershed image. This PSNR values means the Proposed Method enhance the final segmented image.

Table 2. SSIM, IMMSE, PSNR, and SNR for noisy coin image

	Psnr value	Immse error	Ssim value
Adjust image before and after applying the PCA	PSNR = 19.6124 SNR = 17.7469	710.9597	0.8307
Gradient image before and after applying the PCA	PSNR = 6.4298 SNR = 4.9203	14794.5063	0.5699
Watershed image before and after applying the PCA	PSNR=8.0956 SNR=7.2236	10081.3393	0.6760
explain	Peak Signal-To-Noise Ratio and Signal-To-Noise Ratio	Mean-Squared Error	Structural Similarity Index for measuring image quality

4.2.2. Second example is for a brain tumor image

One of the most important applications of the watershed algorithms is medical image segmentation. The application of the watershed algorithm in medical image segmentation is one of the most successful efficient algorithms, but unfortunately, the noise affecting the image causing image over-segmentation, is a drawback for these algorithms. Therefore, it is necessary to look for solutions that are quick and easy to apply. This paper presents one such possible solution, which has been proven to be successful in pre-processing the image by principal component analysis before applying the watershed algorithm. Fig. 14 shows an example for medical image of a brain tumor.

Our second example is for a brain tumor image Fig. 14. Fig.14-a showsan example of the original brain tumor image. This image contains a brain tumor in the left lobe of the back of the head. We will apply the proposed algorithm to this image and show the advantages and capability of the proposed method. Apply the watershed algorithm as described in Watershed Segmentation. This algorithm start to convert the image to grayscale values. Next, apply the Morphological process and continue with the watershed algorithm process to check the results of this algorithm. Fig. 14-b shows the input image of the watershed algorithm. Figure 14-c shows the image that was pre-processed by principal component analysis and then inputted into the watershed algorithm. The image in Fig. 14-b shows a grayscale image of the original image (brain tumor image) before it was processed. The illustration in Fig. 14-c shows the grayscale image of the original image (brain tumor image) after it has been processed by principal components analysis. In image Fig. 14-c, an illustration of the improvement due to our direct modeling of the principal components analysis. Principal component analysis shifts the image to a new space which improves the threshold line because the noise in the image has been reduced using principal component analysis and the grayscale in the image is approximated to the minimum by reducing dimensional properties of the principle component analysis, that make the image was clearer because the threshold line is automatically drawn to clearly show the tumor area.

Fig.15 shows the gradient image for brain tumor images segmented by the watershed algorithm. Fig. 15-a shows the original image of the brain tumor. Fig. 15-b shows the gradient image of the watershed algorithm, it shows some confusion due to the black targets at different positions in the gradient image. Fig. 15-c shows a gradient image that was preprocessed by principal component analysis and then inputted into the watershed algorithm. The principal component analysis preprocessed image is more effective because it marks the tumor at a specific location in the whole brain image.

Now we apply the watershed algorithm after preprocessing the image with the principal component analysis function Fig. 16. The image in Fig. 16-c shows the segmented image (brain tumor image) with the Watershed algorithm after applying principle Component Analysis. Figure 16-a shows the original image of the brain tumor. Fig. 16-b shows a segmented image of the watershed algorithm, it shows some confusion as the black target on the whole white image confuses the reader about the actual location in the brain. Fig. 16-c shows a segmented image that was preprocessed by principal component analysis and then entered into the watershed algorithm. Preprocessed principal component analysis is more effective because it marks the tumor at a specific location in the whole brain image, making the image easier to read and simpler to read.

So far we have achieved the desired goal of this article, which is a comparison between two images that have been segmented by Watershed algorithm. The first image is directly segmented by Watershed algorithm. And the second image is segmented by Watershed algorithm after preprocessing with principal component analysis. Principal component analysis preprocessing was a very successful one, as it facilitated the detection of the desired targets with high efficiency, due to the noise reduction characteristics of it. Principle component analysis PCA detected the classical edges in the area-based watershed algorithm, improved the weak edge contour, corrected and strengthened it that make the image clearer and more effective.

4.2.2.1 Compare the results of the segmented images generated by watershed`s segmentation algorithm and proposed algorithm (WSA-PCA).

This section explain the comparison between segmented images generated by watershed`s segmentation algorithm and proposed algorithm (WSA-

PCA) based on pre-processing the images by Principle Component Analysis before entered to the watershed`s segmentation algorithm. MATLAB software is used to implement the two algorithms used to achieve the desired results, which will be discussed and analyzed in this section.

Image segmentation results in an amount of over-segmentation depending on the pre-processing stage. It is necessary to find out the optimal solution for different types of images. Principle Component Analysis handles the over-segmentation problem by de-noising, repairing, and modifying the given image. A set of different images is tested for segmentation qualities by applying the principle component analysis techniques.

Below Table 3 shows changes in SSIM (Structural Similarity Index for measuring image quality), IMMSE (Image Mean-Squared Error), PSNR (Peak Signal-To-Noise & Ratio), and SNR (Signal-To-Noise Ratio). Resultant Structural Similarity Index for measuring image quality The SSIM index is a decimal value from -1 to 1, where 1 indicates perfect similarity, 0 indicates no similarity, and -1 indicates perfect anti-correlation. In this study the SSIM Value is 0.8195 for the image after applying the pre-processing stage with the principle component analysis, this leads to SSIM value of 0.6548 in the gradient image and produces SSIM value of 0.6072 in the final watershed image. This SSIM values means the proposed method enhances the final segmented image.

Table 3. SSIM, IMMSE, PSNR, and SNR for brain

	Adjust the image before and after applying the PCA	**Gradient image before and after applying the PCA**	**Watershed image before and after applying the PCA**	**explain**
Ssim value	0.8195	0.6548	0.6072	Structural Similarity Index for measuring image quality
Immse error	623.7482	13678.1074	15272.7354	Mean-Squared Error
Psnr value	PSNR = 20.1807 SNR = 17.8707	PSNR = 6.7705 SNR = 6.2169	PSNR=6.2916 SNR=6.2530	Peak Signal-To-Noise Ratio and Signal-To-Noise Ratio

The resulting mean squared error, MSE, is used to check how close the estimate or prediction is to the actual value. The lower the MSE, the closer the prediction is to the actual value. This is used as a model evaluation measure for regression models, with lower values indicating better fit. In this study the IMMSE error Value is 0.623*104 for the image after applying the pre-processing stage with the Principle Component Analysis, this leads to the IMMSE error value 1.367×10^4 in the gradient image and produce Ithe MMSE error value $1of\ .527 \times 10^4$in the final watershed image. This IMMSE error values means the Proposed Method enhances the final segmented image. The resulting maximum signal-to-noise ratio (PSNR) is the ratio between the maximum possible performance of an image and the power of disturbing noise that affects the quality of its representation. To estimate her PSNR of an image, we need to compare that image to an ideal, clean image with the maximum possible performance.

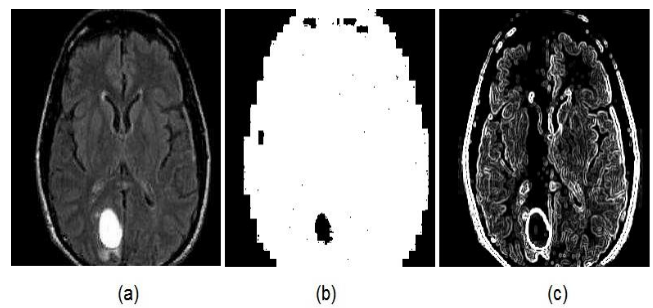


Figure 14. Original brain tumor image. (b) Adjusted watershed algorithm image, (c) adjusted PCA image segmented by the watershed algorithm.

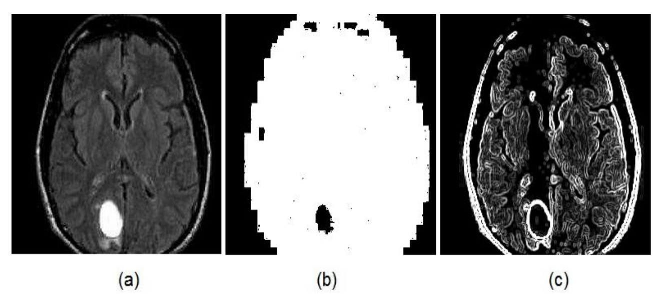


Figure 15. Original brain tumor image (b) gradient watershed algorithm image (c) Gradient PCA image segmented by watershed algorithm

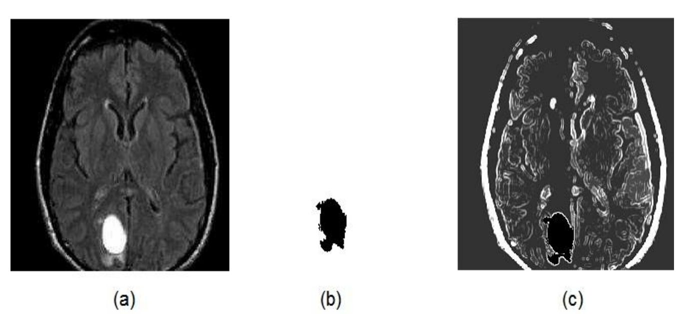


Figure 16. Original brain tumor image (b) segmented watershed algorithm image (c) Segmented PCA image segmented by a watershed algorithm.

PSNR is most commonly used to estimate the efficiency of things like compressors and filters. The higher the PSNR value, the more efficient the corresponding compression or filtering method. In this study the PSNR Value is 20.18 dB for the image after applying the pre-processing stage with the Principle Component Analysis, this leads to PSNR value 6.77 dB in the gradient image and produces PSNR Value of 6.2916 dB in the final watershed image. This PSNR value means the Proposed Method enhances the final segmented image.

Authors' contribution

All authors contributed equally to the preparation of this article.

Declaration of competing interest

The authors declare no conflicts of interest.

Funding source

This study didn't receive any specific funds.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

5. Conclusions

The image segmentation algorithms have many studies in this work, such as over-segmentation and image distortion due to reflected light. The image processing using the watershed algorithm was applied to improve the algorithm based on principal component analysis. The new processing shows that the technique can achieve accurate and durable target shapes.

REFERENCES

[1] Haq, I.; Ullah, N.; Mazhar, T.; Malik, M.A.; Bano, I. A Novel Brain Tumor Detection and Coloring Technique from 2D MRI Images. Appl. Sci. 2022, 12, 5744. https://doi.org/10.3390/app12115744.

[2] Santoshachandra Rao Karanam, Y. Srinivas and M. Vamshi Krishna, Study on image processing using deep learning techniques, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2020.09.536.

[3] G Sumanth Prasad. Implementation of Image Segmentation Algorithms in Digital Image Processing using MATLAB. International Journal for Interdisciplinary Sciences and Engineering Applications. IJISEA - An International Peer- Reviewed Journal 2020, Volume 1 Issue 1.

[4] Yu, Y.; Wang, C.; Fu, Q.; Kou, R.; Huang, F.; Yang, B.; Yang, T.;Gao, M. Techniques and Challenges of Image Segmentation: A Review. Electronics 2023, 12, 1199. https://doi.org/10.3390/electronics12051199

[5] Anwesh, K.; Pal, D.; Ganguly, D.; Chatterjee, K.; Roy, S. Number plate recognition from enhanced super-resolution using generative adversarial network. Multimed. Tools Appl. 2022, 1–17.

[6] Jin, B.; Cruz, L.; Gonçalves, N. Deep Facial Diagnosis: Deep Transfer Learning from Face Recognition to Facial Diagnosis. IEEE Access 2020, 8, 123649–123661.

[7] Zhao, M.; Liu, Q.; Jha, R.; Deng, R.; Yao, T.; Mahadevan-Jansen, A.; Tyska, M.J.; Millis, B.A.; Huo, Y. VoxelEmbed: 3D Instance Segmentation and Tracking with Voxel Embedding based Deep Learning. In Proceedings of the International Workshop on Machine Learning in Medical Imaging, Strasbourg, France, 27 September 2021; Volume 12966, pp. 437–446.

[8] Yao, T.; Qu, C.; Liu, Q.; Deng, R.; Tian, Y.; Xu, J.; Jha, A.; Bao, S.; Zhao, M.; Fogo, A.B.; et al. Compound Figure Separation of Biomedical Images with Side Loss. In Proceedings of the Deep Generative Models, and Data Augmentation, Labelling, and Imperfections, Strasbourg, France, 1 October 2021; Volume 13003, pp. 173–183.

[9] Wu, Y.; Li, Q. The Algorithm of Watershed Color Image Segmentation Based on Morphological Gradient. Sensors 2022, 22, 8202. https://doi.org/10.3390/s22218202

[10] Anton S. Kornilov and Ilia V. Safonov, An Overview of Watershed Algorithm Implementations in Open Source Libraries. J. Imaging 2018, 4, 123. J. Imaging 2018, 4, 123; www.mdpi.com/journal/jimaging. doi:10.3390/jimaging4100123.

[11] Yu, Y.; Wang, C.; Fu, Q.; Kou, R.; Huang, F.; Yang, B.; Yang, T.; Gao, M. Techniques and Challenges of Image Segmentation: A Review. Electronics 2023, 12, 1199. https://doi.org/10.3390/electronics12051199

[12] Silpa Joseph, IMAGE PROCESSING TECHNIQUES AND ITS APPLICATIONS: AN OVERVIEW, Vol-4 Issue-3 2018 IJARIIE-ISSN(O)-2395-4396, 8745 www.ijariie.com 2168.

[13] H. P. Narkhede, Review of Image Segmentation Techniques, Vol-4 Issue-3 2018 IJARIIE-ISSN(O)-2395-4396, 8745 www.ijariie.com 2168. International Journal of Science and Modern Engineering (IJISME) ISSN: 2319-6386, Volume-1 Issue-8, July 2013

[14] Shubha Majumder, Brain Tumor Extraction from MRI images using Prominent Image Segmentation Methods, Spring 2018

[15] Jardim, S.; António, J; Mora, C. Graphical Image Region Extraction with K-Means Clustering and Watershed. J. Imaging 2022, 8,163. https://doi.org/10.3390/jimaging8060163.

[16] Y. Murali Mohan Babu, 2Dr. M.V. Subramanyam, 3Dr. M.N. Giri Prasad. PCA based image denoising. Signal & Image Processing: An International Journal (SIPIJ) Vol.3, No.2, April 2012. http://DOI:10.5121/sipij.2012.3218.

[17] National Research Nuclear University MEPhI, Kashirskoye sh. 31, 115409 Moscow, Russia; J. Imaging 2018, 4, 123; doi:10.3390/jimaging4100123 www.mdpi.com/journal/jimaging.

[18] Safonov, I.V.; Mavrin, G.N.; Kryzhanovsky, K.A. Segmentation of Convex Cells with Partially Undefined Edges. Pattern Recognit. Image Anal. 2008, 18, 112–117.

[19] Y. Chang, X. Li, "Adaptive Image Region Growing", IEEE Trans. On Image Processing, Vol. 3, No. 6, 1994.

[20] Wu, Y.; Li, Q. The Algorithm of Watershed Color Image Segmentation Based on Morphological Gradient. Sensors 2022, 22, 8202. https://doi.org/10.3390/s22218202.

[21]. Khan, Z.; Yang, J.; Zheng, Y. Efficient clustering approach for adaptive unsupervised colour image segmentation. Sensors 2019, 13, 1763–1772. [https:// doi/10.1049/iet-ipr.2018.5976].

[22] Kurmi, Y.; Chaurasia, V. Multifeature-based medical image segmentation. Sensors 2018, 12, 1491–1498. [http://doi.org/10.1049/iet-ipr.2017.1020].

[23]. Ahmad EL ALLAOUI1 and M'barek NASRI11 - LABO MATSI, ESTO, B.P 473, University Mohammed I OUJDA, MOROCCO. The International Journal of Multimedia & Its Applications (IJMA) Vol.4, No.3, June 2012. DOI :10.5121/ijma.2012.4301.

[24]. Anuj Kumar, Dr. Umesh Chandra, Comparative Analysis of Image Segmentation using Edge-Region Based Technique and Watershed, International Journal of Latest Technology in Engineering, Management & Applied Science (IJLTEMAS) Volume VII, Issue V, May 2018.

[25] Amandeep Kaur, Aayushi, Image Segmentation using Watershed Transform, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-4, Issue-1, March 2014.

[26] Farheen K. Siddiqui, An Efficient Image Segmentation Approach through Enhanced Watershed Algorithm, Computer Engineering and Intelligent Systems www.iiste.org ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online), Vol.4, No.6, 2013.

[27] L. Vincent and P. Soille. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans. Patt. Anal. Mach. Intell. 13(6):583-598, 1991.

[28]. Michael Greenacre, Patrick J. F. Groenen, Trevor Hastie, Alfonso Iodice d'Enza, Angelos Markos, and Elena Tuzhilina, Principal Component

Analysis, Nature Reviews Methods Primers, January 2023.
[29] Ayad Al-Rumaithi (2023). Mode Decomposition using Principal Component Analysis (https://www.mathworks.com/matlabcentral/fileexchange/71878-mode-decomposition-using-principal-component-analysis), MATLAB Central File Exchange. Retrieved June 26, 2023.
[30] Y. Murali Mohan Babu, 2Dr. M.V. Subramanyam, 3Dr. M.N. Giri Prasad, PCA based image denoising, Signal & Image Processing: An International Journal (SIPIJ) Vol.3, No.2, April 2012, DOI : 10.5121/sipij.2012.3218.
[31] Lei Zhang , Weisheng Dong , David Zhang , Guangming Shi, “Two-stage image denoising by principal component analysis with local pixel grouping” Elsevier-Pattern Recognition,vol-43, 2010, 1531-15