Using entropy based mean shift filter and modified watershed transform for grain segmentation


Published in:
International Journal of Automation and Computing

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
© 2015 Institute of Automation, Chinese Academy of Science and Springer-Verlag Berlin Heidelberg
The final publication is available at Springer via http://dx.doi.org/10.1007/s11633-014-0841-2

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person’s rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Download date:12. Jan. 2019
Using Entropy Based MeanShift Filter and Modified Watershed Transform for Grain Segmentation

Kun Zhang¹, Minrui Fei¹ · Huiyu Zhou²

¹ School of Mechatronic Engineering & Automation, Shanghai University, Shanghai, 200072, PRC
²The Institute of Electronics, Communications and Information Technology, Queen’s University of Belfast, Belfast, U.K.

Abstract: Life science research aims to continuously improve the quality and standard of human life. One of the major challenges in this area is to maintain food safety and security. A number of image processing techniques have been used to investigate the quality of food products in the last decade. In this paper, we propose a new algorithm to effectively segment connected grains so that each of them can be inspected in a later processing stage. One family of the existing segmentation methods is based on the idea of watershedding, which has shown promising results in practice. However, due to the over-segmentation issue, this technique has experienced poor performance in various circumstances, such as inhomogeneous background and connected targets. To solve this problem, we present a combination of two classical techniques to handle this issue. In the first step, a mean shift filter is used to eliminate the inhomogeneous background, where entropy is used as a converging criterion. Secondly, a color gradient algorithm is used in order to detect significant edges, and a marked watershed transform is applied to segment the cluttered objects out of the previous processing stages. The proposed framework is capable of compromising among execution time, usability, efficiency and segmentation outcomes in analyzing ring die pellets. The experimental results demonstrate that the proposed approach is effective and robust.

Keywords: Robust segmentation, cluttered grains, mean shift filter, entropy, modified marked watershed transform

1 Introduction

Inspection of grains is a very important task for ensuring food safety and security. An inspection process consists of segmentation of grains and quality evaluation of each piece of grains. Currently, segmentation of grains is manually handled and hence time consuming. In the meantime, it is very difficult to maintain the quality of segmentation across different inspectors. With the development of image processing and machine vision techniques, the nondestructive inspection of grains can be achieved at a high speed and accuracy [1-3]. However, it is also recognized that the segmentation becomes a challenge in the presence of clutters or inhomogeneous backgrounds.

Traditional low level image segmentation processing approaches, for example, thresholding [4], region growing [5], clustering [6] and evolutionary algorithms [7], require a considerable amount of human-computer interaction in order to attain satisfactory results. Making these systems hand-free is difficult because of ambiguity, uncertainty, and variability of individual objects. Furthermore, inhomogeneous background or connected target can cause over-segmentation of images.

Many segmentation techniques have been presented in the literature in order to effectively separate grains. For example, Shatadal used an erosion method for segmenting cluttered grains [8]. Shatadal used ellipse fitting to segment touching grains [9]. Van den Berg developed an algorithm to detect characteristic grains edges along the detected contours [10].

These approaches are not effective in case a large number of grains are connected [11]. Connected grains are especially prominent in severe cases [12]. In the field of connected objects segmentation, watershed transform is one of the most commonly used techniques. Images’ gray distribution corresponds to geographic altitudes, and water basin corresponds to a separable region. The growing process starts from a local minimum, and each of the minimum points generates an individual region after transformation [13]. Watershed transform can be used to produce wide continuous edges and homogeneously closed regions [14]. This mechanism is consistent with human being’s perception. However, watershed methods suffer from the over-segmentation problem due to the presence of a large number of local minimums. These local minimum points are due to image noise. Meanwhile, its computation is of high complexity that makes it not feasible for real-time process. To our knowledge, many solutions have been proposed in order to reduce the noise’s effect on this technique:

1) Vincent proposed a fast watershed algorithm based on immersion simulation [15]
2) Bentsson proposed a robust watershed algorithm based on H-minima mark to overcome the problem of over-segmentation[16]
3) Gaoli proposed an improved marked watershed using a low-pass filter to eliminate fake local minima regions[17]

To solve these problems, we propose in this paper an improved Watershed transform combining mean shift filtering with entropy. Entropy is used as a stopping criterion to regulate the convergence of mean shift. Color gradient image is directly calculated in the RGB space to obtain higher accuracy than the methods using gray gradient calculation. After we extract markers and water basin information, over-segmentation can be greatly reduced. Experiments show that the proposed approach provides consistent and robust segmentation performance

¹ Manuscript received date; revised date
This work was supported by National Key Scientific Apparatus Development of Special Item(2012YQ150008,03), and this work is supported by Nantong Research Program of Application Foundation under Grant (BK2012030)
with satisfactory efficiency; in addition, our new approach is noise immune with good edge-location capability.

The rest of the paper is organized as follows: In section 2, as a well-known algorithm, we briefly introduce watershed methods, mean shift and the entropy concept. In section 3, we describe our segmentation algorithm into two parts. Finally, in section 4, we show the experimental results. Section 5 concludes this paper.

2 Preliminaries

2.1 Watershed using geomorphologic distance

Watershed transform is considered as a segmentation approach based on regions [17, 18]. There are two major classes of watershed transforms, one was proposed by Vincent & Soille, and the other by Meyer [19]. We focus on the Meyer’s algorithm, and the reader can refer to [20] for details.

We assume that $y$ is a gray image. The gradient $S(e)$ is defined as the maximal gradient and decent to its neighborhood in a low altitude. $S(e)$ is written as:

$$S(e) = \max_{v \in N(e)} \left( \frac{y(v)-y(e)}{d(e,v)} \right)$$

Where $n(e)$ is a numerical set of a neighbor pixel $e$, and $d(e,v)$ is the Euclidean distance, which is bound up with $edge(e,v)$. When $e = v$, $S(e)$ is equal to zero. The pixels whose neighborhoods are of higher gray levels than others will be assigned zero intensity. So we have a lower gradient where $e$ is a local minimum pixel. Eventually, the loss from pixel $e$ to the neighborhood $v$ is described as:

$$loss(e,v) = \begin{cases} S(e) \cdot d(e,v) & \text{if } y(e) > y(v) \\ S(e) \cdot d(e,v) & \text{if } y(v) < y(v) \\ \frac{1}{2} (S(e) + S(v)) \cdot d(e,v) & \text{if } y(e) = y(v) \end{cases}$$

The geomorphologic distance along $\pi = \{e_1, ..., e_n\}$ between $e_0 = e$ and $e_n = v$ is defined as:

$T^\pi(e,v) = \sum_{i=0}^{n-1} d(e_i, e_{i+1})loss(e_i, e_{i+1})$ (3)

Following the geomorphologic distance’s definition, we define the basin of the local minimum pixels as a data set, and it has smaller geomorphologic distances than the other local minimum values. Finally, the pixels data sets that do not belong to any of the basins are set to be the watershed.

The purpose of watershed transform is to extract watershed basins from an image, and its performance relates closely to a gradient image. Classical gradient methods will obtain many regional basins due to image noise. The result of the classical gradient based watershed transform methods includes a large number of tiny local regions, which severely affects the post-processing.

2.2 Mean shift filter

The following is the concept of the classical mean shift approach [21]: Let $x$ be a numerical sample of $n$ in a $d$ dimensional space. The basic mean shift is defined as

$$M_k(x) = \frac{1}{k} \sum_{i=-k}^{k} (x_i - x)$$ (4)

Where $g_k$ is a window with center $x$ and radius $h$. $k$ is the sample set number in $g_k$. $(x_i-x)$ is an relative offset of center $x$.

Eq. (4) is a monotonic form and not effective in practical applications. The Kernel based mean shift algorithm is to minimize the following function:

$$M_k(x) = \frac{1}{h} \sum_{i=-k}^{k} Q(\frac{x-x_i}{h})$$ (5)

$\alpha(x)$ is the self-impact factor, $Q(x)$ is a kernel function.

In a color image of $n \times n$ pixels, each pixel corresponds to a 5 dimension vector $R, G, B, X, Y$. Due to the independence of space and color information, the kernel function is obtained via Eq. (6)

$$Q(x) = \frac{1}{\sum_{i}^{5} q(||x-x_i||^2)}$$ (6)

Where $x$ is the spatial position of an image pixel; $x^*$ is the color information of the pixel; $g_s$ is the spatial window with center $x$ and radius $s$; $g_r$ is a color window with center $x$ and radius $r$.

2.3 Entropy

In the image processing, entropy is defined as:

$$F(t) = -\sum_{t=0}^{x-1} q(t) \log_2 q(t)$$ (7)

Where $A$ is an image with $log_2(0) = 0$; $q(t)$ is a gray value with a probability. Within a homogeneous local area, the minimum value of inhomogeneity can be found using entropy. The theoretical probability of $q(t)$ is the one used in a homogeneous local area. In practice, due to image noise, the entropy value cannot equal to zero. Thus, if we try to apply entropy to a measurement of confusion, it can be used as a down stopping rule for the iteration of a mean shift algorithm.

Entropy can be used to reduce noise in each local area, and when the entropy value reaches a stable value, the entire image tends to be more homogeneous than the original image.

3 Proposed Algorithms

The classical marked watershed transformation algorithm usually consists of the following four steps:

Step1: low-pass filtering to a color image
Step2: calculation of the gray image gradient $VI$
Step3: extraction of markers in $VI$ with $H - minima$ and imposing them as local region minima on the gradient image
Step4: taking watershed transform of the marked gradient image.

Based on the classical watershed transform, we generate our proposed approach.
1. Using the entropy based mean shift filter to improve the image edge identification.
2. Calculate the color image gradient $\nabla I$
3. Apply $H - \text{minima}$ parameters based on basin information

Fig. 1 illustrates the proposed algorithm. In this section, we show two improved algorithms. The first one is for image filtering and the second one is for the watershed segmentation process, which is the extension of the standard filtering approach.

Algorithm 1: entropy based mean shift filtering algorithm

Let $x_i$, $i = 1, ..., n$ be the input image. Let $O_i$, $i = 1, ..., n$ be the filtered image. Pixel $p \in x_i$, $p = (R, G, B, x, y) \in \mathbb{R}^5$. Let $ent.0$ be the entropy initial value, $ent.1$ be the next iteration of $ent.0$, and $ent.2$ be the next value of $ent.1$. Let $erras$ be the absolute value of the difference between the first two iterations. Let $edset$ be the thresholding as the algorithm’s stopping criteria. Our algorithm comprises the following steps:

1. Initialize $h = 1$, $y_{h,0} = p_h$, $ent.2 = 1$, $erras = 1$, $edset = 0$.
2. While $erras > edset$ then
   2.1 Filter the image using mean shift. Store the result in $O^{K1}$.
   2.2 Calculate entropy from the $O^{K1}$ and store the result in $ent.1$.
   2.3 Entropy is used to calculate the absolute difference which is obtained in the previous step; $erras = |ent.1 - ent.2|$.
   2.4 Update the parameters; $ent.1 = ent.2$ and $O^{K+1} = O^{K1}$.
   2.5 Continue to apply mean shift which is carried out till the image entropy converges.
3. Store at $Z_i$ which is calculated as $Z_i = (x_i', y_i')$. Where $x_i'$ is the spatial information and $y_i'$ is the color range information.

Algorithm 2: Modified marked watershed transform:

Let $M_1$ be the gradient image’s mean value, $M_2$ be the gradient image’s local minimum mean value and $M_3$ be the gradient image’s local maximum mean value. Let $S$ be the local minimum region area, $\nabla S$, be the local minimum region ratio between an area and a depth. $a(S)$ and $a(\nabla S)$ are the sequence of $S$ and $\nabla S$, respectively. $S'$ and $\nabla S'$ are the middle values of $S$ and $\nabla S$.

$sum$ is the number of a local minimum region, and $r$ is the threshold of the maximum number of the segmented region.

Then, the segment algorithm comprises the following steps:
1. Calculate color image gradients according to literature[22]
2. Extract markers in the gradients, and adopt $H$ that has been defined as
   $$H = \alpha \cdot (M_3 - M_2), \quad \alpha = \frac{(M_3 - M_1)^2 + (M_1 - M_2)^2}{(M_3 - M_2)^2}$$
3. If the $sum$ is larger than $2r$, we obtain first $2r$ of $S$ and $\nabla S$. Finally, update $S$ and $\nabla S$.
4. Apply the watershed transform
5. End

4 Experiments and analysis

The proposed algorithm has been evaluated against a number of pellet images. The pellet samples are collected with the help of a company located in Shandong Province of China with 70 images which are captured using a Basler CCD sensor (A601f) and the images are then resized to be of 640*480 pixels. Experiments have been carried out on the ring die pellet images with different conditions to evaluate the feasibility and efficiency of the proposed algorithm. These experiments were carried out on an Intel dual-core 3.0GHz PC with 4 GB RAM. The first two of the computations were performed with Matlab 7.0, and the other experiments were performed with visual studio 2010.

4.1 Analysis of major effect issues

As mentioned above, inhomogeneous background and cluttered objects are the major issues affecting the outcome of image segmentation. Firstly, the experiment shows the segmentation results of the traditional algorithms in different situations. The results are shown in Fig. 2, and
Fig.3. Fig. 2(a) is the origin image collected in a homogeneous background and non-clutter environment, and Fig.3 (a) is of inhomogeneous background with a noisy and cluttered environment. Fig. 2(b), Fig.3(b), Fig.2(c), Fig.3(c), Fig.2(d), and Fig.3(d) show the segmentation results from several state of the art methods: Canny, Sobel and improved Canny. From the comparisons, we observe that these classical segment methods perform poorly in the challenging conditions.

4.2 Analysis of the entropy based mean shift filter

In image denoising, a classical low-passing filter can be used to suppress high frequency noise [23-25]. However, it is hard for them to preserve the edges of images due to the mixture in some frequency bands. Here, we use discrete Fourier energy density spectrums to illustrate the outcomes of different filters. Discrete 2D Fourier transform is based on the following definitions:

\[
F(u,v) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n) \exp \left( -2\pi i \left( \frac{mu}{M} + \frac{nv}{N} \right) \right) \tag{8}
\]

\[
u = 0,1,\ldots, M-1, v = 0,1,\ldots, N-1
\]

Inverse Fourier transform is defined as follows

\[
f(m,n) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(U,V) \exp \left( 2\pi i \left( \frac{mu}{M} + \frac{nv}{N} \right) \right) \tag{9}
\]

\[m = 0,1,\ldots, M-1, n = 0,1,\ldots, N-1
\]

For the convenience of calculation, Eq.(8) can be turned into Eq.10

\[
F(u,v) = \frac{1}{M} \sum_{m=0}^{M-1} \left[ \frac{1}{N} \sum_{n=0}^{N-1} \exp \left( -2\pi i \left( \frac{2nu}{N} \right) \right) f(m,n) \right] \exp \left( -2\pi i \left( \frac{2mu}{M} \right) \right) \tag{10}
\]

\[u = 0,1,\ldots, M-1, v = 0,1,\ldots, N-1
\]

The item in the square bracket corresponds to the one-dimensional Fourier transform of line m . The one-dimensional Fourier transform is easy to be calculated by standard fast Fourier transform (FFT). Each row is calculated by Fourier transform, followed by the calculation of one-dimensional Fourier transform for each column.

The result of 2D Fourier transform is a complex frequency spectrum. The value of the real part and the image part of the frequency spectrum are very large, and sometimes they can reach millions. For the convenience of data storage, we usually use a monotone function to reduce its value range such as \( \sqrt{|F(u,v)|} \) or \( \log|F(u,v)| \).

In Fig. 4(a), this is the energy spectrum of the original image, where the yellow-orange color indicates the major energy of symbol “+” and this area is easily contaminated by the background noise (i.e. blue and green areas). Fig. 4(b) shows the outcome of a low-pass filter, where only the central area of the symbol is kept but the edges of symbol “+” are mixed with the background. In Fig. 4(c), based on mean shift, the filtering result shows that the central area of the symbol is outstanding and the edges are also kept well. The results show that the traditional low-pass filter has good performance on image smoothing but affects the edge details.

Fig. 5. (a) is the output of 27 iterations before it reaches convergence with mean shift. Fig. 5. (b) is the result of involving 7 iterations to reach convergence with the entropy based mean shift. In Fig. 6, the entropy value gradually decreases before the convergence. We calculate the entropy value at each iteration. The mean shift iteration stops at a given thresholding. Fig. 7 shows the result of using a mean shift based method. Fig. 7(a) illustrates the energy distribution of the original image with many spikes (noise) on the background. Fig. 7(b) is the outcome of mean shift, where the background contains much less spikes than Fig. 7(a) and this attributes to the strong filtering capability of mean shift.

The performance of mean shift filtering can be measured using the Mean Square Error (MSE) defined as
\[ \text{MSE} = \sqrt{\frac{1}{BG} \sum_{x=0}^{B-1} \sum_{y=0}^{G-1} [(I(x,y) - I(x,y))^2]} \]  

Where, \( I(x,y) \) is denoted as the filtered output image. \( I(x,y) \) is denoted as the original input image. Table 1 shows the performance of the three different filtering approaches.

Table 1. The filtering performance of three filtering algorithms

<table>
<thead>
<tr>
<th>Performance</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison</td>
<td>( \alpha = 0.05 )</td>
</tr>
<tr>
<td>Median filter</td>
<td>28.8642</td>
</tr>
<tr>
<td>Butterworth filter</td>
<td>28.4517</td>
</tr>
<tr>
<td>Proposed filter</td>
<td>26.9695</td>
</tr>
</tbody>
</table>

4.3 Analysis of the image gradient transform

Gradient image is more important than an image itself. An color image has three color channels. In this paper, the
Di Zenzo’s method [26] is used to calculate gradient $\nabla I$, directly in the RGB space. The experiment proves that calculating gradient with color vectors can achieve better performance than gray gradient images.

Fig. 8 shows the color gradient without filtering, and Fig. 8(c) illustrates the color gradient with a mean shift filter.

### 4.4 Comparison of segmentation performance

In this section, comparison results of the segmentation performance are shown for our method, Vincent’ method [13], Bentsson’ method [16] and Gaoli’ method [13]. Fig. 9 shows the segmentation results of cluttered pellet image. Fig. 9(a) shows the result of the Vincent’ method, Fig. 9(b) is the outcome of the Bentsson’ method, Fig. 9(c) is the result of the Gaoli’ method, and Fig. 9(d) is the result of our method. The number of regions obtained by each approach is shown in Table 2. We clearly see that the Vincent and Bentsson’s algorithm suffers from over-segmentation. The Vincent’s method is sensitive to inhomogeneous background and cluttered targets. The Bentsson’ method has better performance than the Vincent’s method due to the use of the H-minima marker approach, although it still needs further improvements. The Gaoli’s method produces false minimum basin regions with low-pass filtering, but it has poor performance on edge segmentation because of the low-pass filter that only removes weak edges. The proposed method greatly reduces over-segmentation and obtains much better segmentation results than all the other methods. Meanwhile, it locates edges accurately in the inhomogeneous background.

![Fig. 8 Gradient transform of an image](image)

**Fig. 8 Gradient transform of an image**

<table>
<thead>
<tr>
<th>Method</th>
<th>Number</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vincent’ algorithm</td>
<td>1977</td>
<td>1.75</td>
</tr>
<tr>
<td>Bentsson’ algorithm</td>
<td>536</td>
<td>2.04</td>
</tr>
<tr>
<td>Gaoli’ algorithm</td>
<td>173</td>
<td>1.98</td>
</tr>
<tr>
<td>The proposed method</td>
<td>70</td>
<td>1.28</td>
</tr>
</tbody>
</table>

**Table 2 region number and executing time(s)**

![Fig. 10 Fake-color marked image](image)

**Fig. 10 Fake-color marked image**

### 5 Conclusion and future work

In this paper, a robust segmentation algorithm for cluttered grain images is proposed using an entropy based mean shift filter and modified Marked Watershed Transform. We used the entropy information to improve the mean shift convergence speed. Meanwhile, the H-minimal depth threshold is adaptively obtained to overcome the threshold selection. Experimental results show that the proposed approach can obtain better segmentation results than the classical Watershed algorithms with lower computational consumption. Moreover, it has better performance in terms of anti-noise and edge detection abilities than the other state of the art techniques.

There are several research directions to extend this work...
that we are currently considering. The first is to make use of the sparse representation not only for image compression, but also for feature extraction. Additionally, the mean shift filter can be developed using other gradient descent techniques to speed up the convergence.

References

Kun Zhang received the B.Sc. degree in automation technology form Jiangsu University, China, in 2005. The M.Sc. degree in Control engineering from the University of Zhejiang, Zhejiang, China, in 2010, and now he is a doctoral student in Shanghai university. China. His research interest covers machine vision, fuzzy control, and networked advanced control.

Minrui Fei (corresponding author) received his B.Sc. degree from Shanghai University in 1982. M.Sc. degree from Shanghai University in 1992, and Ph.D. degree from Shanghai University in 1997. His research interests include distributed and fieldbus control, systems, networked advanced control and system implementation, key technology and applications in multi-fieldbus conversion and performance evaluation, as well as the application of virtual reality and digital simulation in industry. He has been a Professor and Doctoral Supervisor at Shanghai
Huiyu Zhou received the B.Sc. degree in radio technology from Huangzhong University of Science and Technology, Huangzhong, China, in 1990, the M.Sc. degree in biomedical engineering from the University of Dundee, Dundee, U.K., in 2002 and the Ph.D. degree in computer vision from the Herriot-Watt University, Edinburgh, Scotland in 2006. Currently, he is a Lecturer in the School of Electronics, Electrical Engineering and Computer Science Queen’s University Belfast, Belfast, U.K. His research interests include computer vision, human motion analysis, intelligent systems, and human computer interface. He has published widely in international journals and conferences. E-mail: h.zhou@qub.ac.uk