Robust computer vision system for marbling meat segmentation

Gabriel Fillipe Centini Campos*, José Luis Seixas Jr.+, Ana Paula A. C. Barbon++, Alan Salvany Felinto*, Ana Maria Bridi++, Sylvio Barbon Junior.*

* Department of Computer Science, Londrina State University, Londrina, 86057-970, Brazil
+ Department of Data Science and Engineering, Eötvös Loránd University, Faculty of Informatics, Budapest, 1117, Hungary
++ Department of Zootechnology, Londrina State University, Londrina, 86057-970, Brazil

Received 3rd of May 2015; accepted 5th March 2020

Abstract

In this study, we developed a robust automatic computer vision system for marbling meat segmentation. Our approach can segment intramuscular fat from meat samples using images acquired with different quality devices in an illumination varying environment, where there was external ambient light and artificial light; thus, professionals can apply this method without specialized knowledge in terms of image treatment or equipment, as well as without disruption to normal procedures, thereby obtaining a robust solution. The proposed approach for marbling segmentation is based on data clustering and dynamic thresholding. Experiments were performed using two datasets that comprised 82 images of 41 longissimus dorsi muscles acquired by different sampling devices. The experimental results showed that the computer vision system performed well with over 98% accuracy and a low number of false positives, regardless of the acquisition device employed.

Key Words: Automation, Computer Vision, Image Segmentation, K-means, Marbling meat

1 Introduction

Meat quality evaluations comprise several processes, which are based on robust techniques that require sophisticated, complex, and costly equipment. Thus, these analyses take a long time to perform, as well as requiring skilled professionals and chemical products that degrade the samples [26]. Furthermore, many of these analyses are subjective (visual and sensory) and they are based only on an evaluation by a qualified professional, such as marbling analysis [30, 27, 24].

Marbling is caused by the intramuscular distribution of fat and it is one of the main attributes of meat that determines consumer choice when purchasing because it influences the palatability, texture, and tenderness of meat [18, 4]. Furthermore, marbling level influences consumers choice, since a high marbling level indicates a superior meat quality [5]. In general, a specialist determines the marbling degree based on a visual assessment...
of the longissimus dorsi (LD) muscle, which is supported by standardized images that are labeled according to numerical scales related to the amount of fat between the muscle fibers, excepting those from sample fat coverage. A higher degree is associated with a greater amount of fat. This methodology is highly dependent on professional subjectivity and it may be questionable [2], such as color evaluation. Thus, human inspections of meat marbling are variable, laborious, tedious, and unsuitable for routine large-scale assessments [29], which has motivated the development of solutions based on computer vision systems.

Some studies have demonstrated the difficulty of segmentation to solve the marbling problem because the acquisition systems require high levels of control. For example, Girolami et al. [8] performed image acquisition in a wooden box with dedicated and controlled lighting, where the camera required support and positioning at an appropriate distance. In Ziadi et al. [32], an NIR device was used in specific acquisition scenarios to obtain overall marbling measurements, where the evaluation was performed with a complex experimental device.

The laboratory where the images were acquired in the present study had ordinary lighting with a series of fluorescent lamps but no dimmer switches or diffusers, and no other lighting was added. Thus, the proposed system is sufficiently robust to work in illumination varying environments that are influenced by surrounding light, as described in Mahadeo et al. [19]. Our motivation was to overcome conventional recognition limitations, which demand image acquisition in controlled environments. In Jung and Scharcanski [16], a robust method was also used to reduce the segmentation cost and to avoid intensive post-processing.

An important feature of our proposed method is that images are acquired with the least possible control, thereby allowing professionals to apply this method without specialist knowledge, sample treatments, or equipment, as well as without disrupting normal procedures to yield a robust solution.

Ulrici et al. [28] highlighted the ability of digital cameras to provide satisfactory detailed color descriptions of objects during food quality analysis because the light sensors yield values in the Red, Green and Blue (RGB) channels; thus, this color model was used in our method.

In this study, our objective was to develop an automatic technique for the segmentation of intramuscular fat in the LD muscle, which could yield high accuracy even in low marbling samples using a robust computer vision system in uncontrolled acquisition conditions, where the camera used could vary in its quality and capacity. We developed a method to determine the quality of meat via the automated visual inspection of meat marbling, which we tested in different real scenarios that affected the evaluation of the degree of marbling.

2 Meat marbling segmentation approaches

The first step in image information analysis is segmentation. Jaglan et al. [14] described several different techniques for segmentation, such as histogram thresholding, feature space, region-based approaches, edge detection approaches, fuzzy approaches, and neural network approaches.

Our approach is an alternative to the automatic segmentation method proposed by Jackman et al. [10], where very good results were obtained at low computational costs and the method is highly reproducible. Furthermore, this study showed main problems and difficulties that can be addressed. Thus, we developed an automatic method that can be used in uncontrolled acquisition conditions regardless of the device quality in order to facilitate the improvement of low-fat breeds based on low quality acquisition devices.

2.1 Previous solution

The first segmentation step is background subtraction because the background is not the object of interest and it is not used in the analysis. This first step was described previously. In Jackman et al. [10], the segmentation method is based on clustering and thresholding, where they used a trial and error approach to determine the threshold value and to remove the blue background. This method could be improved by developing a non-empirical approach.

Next step aims to remove fat at muscles’ edges, which is not part of the marbling. K-means algorithm was used for this purpose in Jackman et al. [10] because it is a powerful algorithm with relatively low complexity,
i.e., \(O(NKd)\), where \(N\) is the input set, \(K\) is the number of clusters, and \(d\) is the problem dimension [31]. In our case, the problem dimension comprises the RGB channel values. It was also reported in Jackman et al. [10] that K-means avoid some false positives pixels if compared with another clustering algorithms.

The application of the K-means algorithm causes two problems. First, it can generate empty clusters and thus lead to runtime errors. Second, when it is applied to images acquired in an illumination varying environment, the muscle darkness caused by light effects might cause K-means to assign dark red pixels as the background. Therefore, the clustering results may include a high number of false positives, as shown in Figure 1a.

Specular reflection is another problem that has been reported in several studies [13, 12, 10], which also generates false positives. Many light spots are observed in real scenarios where the pixels have high blue channel values rather than red and green channel values, as shown in Figure 1a. Figure 1b was generated using our proposed approach, which is discussed in the next section.

![Figure 1](image1.png)

Figure 1: Detecting the problems caused by K-means in terms of dark red muscle and blue light spots

### 2.2 Proposed approach

The proposed approach for marbling segmentation is divided into two parts. The first part is presented in Figure 2 and the second part is shown in Figure 3.

![Figure 2](image2.png)

Figure 2: First part of the proposed segmentation approach

After reading the image, the red and blue channels are swapped to perform a hue channel thresholding in the Hue, Saturation and Intensity (HSI) color space to remove background. This channel swap is performed to facilitate the blue background removal, using just a single threshold point in the H channel. After the background has been detected and segmented, the color channel is reverted and the image is converted back to the Red, Green and Blue (RGB) color space. Then, a three class K-means clustering is applied to estimate marbling, muscle, and background pixels. As a first step to avoid false positive pixels, a region-growing approach is applied to the K-means muscle mask, removing objects smaller than 2.5% of the image’s size (as used in
Based on the mean brightness of all samples, muscle region brightness is changed to the population mean brightness, which helps to determine a threshold value. After the DoubleGreen step [10], thresholding yields a new muscle region based on values determined using Otsu’s method. The first marbling mask is obtained by filling holes in the muscle region. At this stage, the marbling mask still has a lot of very small false positive regions, requiring a second small region removal step (less than $1.5 \times 10^{-3}$% percent of the image size). Finally, any remaining false positive regions in the marbling mask are removed via pixel classification (dark red muscle and blue light spot regions).

We use a histogram-based approach instead of static and empirical threshold value to remove background. The proposed threshold value is calculated dynamically as a pixel value at the interval between the highest histogram value and the next peak. The empirical thresholding method reported previously (fixed at a value found empirically) is substituted for a value obtained using Otsu’s method [21].

Problems where dark red muscle is assigned as background and specular reflection are solved in final marbling process stage, where the differences between muscle, brightness and marbling pixels are more evident. Dark red muscle and fat classification is achieved by analyzing RGB values using the following proposed equation.

\[
\text{Pixel} = \begin{cases} 
\text{Muscle}, & \text{if } R > G + B \\
\text{Fat}, & \text{if } R \leq G + B 
\end{cases} \quad \text{(Dark muscle classification)}
\]

Brightness and fat are distinguished by the following proposed equation.

\[
\text{Pixel} = \begin{cases} 
\text{Brightness}, & \text{if } B^2 > R.G \\
\text{Fat}, & \text{if } B^2 \leq R.G 
\end{cases} \quad \text{(Brightness classification)}
\]

An example of marbling image with some dark red muscle and blue light spot pixels is shown in Figure 1a, as the image after their removal is shown in Figure 1b.

## 3 Materials and Methods

The overall experiment is described in Figure 4, which shows acquisition, sample preparation, evaluation, and results.

### 3.1 Samples

Forty-one LD samples were removed between 12th and 13th ribs of each Nelore breed animal. Animals were fed on pasture, thus they received basic forage feeding. They were slaughtered at a federally inspected abattoir and image acquisition was performed at 24 hours postmortem. Pasture-fed cattle from this breed are more robust and agitated and they have a lower capacity to deposit intramuscular fat [7].
Figure 4: Experimental model

Image sampling was performed at the Food Analysis Laboratory (LANA) at State University of Londrina. Samples were boneless with no fat layer extraction and they were cut at a thickness of 2 cm. After cutting the steaks, they were analyzed subjectively by experts using traditional method for marbling evaluation with a photographic standard and each sample was assigned a score. Thereafter, samples were acquired for computer vision analysis at the laboratory in an illumination varying environment, which was illuminated by ambient daylight and cool white fluorescent artificial lighting.

3.2 Image acquisition

LANA environment had no special illumination. The acquisition process did not use any controlling structures, light calibration devices, or reflection homogenization. Acquired images were divided into two datasets in order to investigate different types of devices and their effects on the model performance.

Images in Dataset 1, which were acquired with a high quality device, and Dataset 2, which were acquired with a low quality device, were obtained in an illumination varying environment with blue paper as background and a paper color checker to facilitate further size and color evaluations, as shown in Figure 5a. Both cameras were CMOS image sensors, which were selected due to their low costs, low power consumption, requirement for a single power supply, and small size that enabled system integration [29].

Dataset 1 was acquired using a digital single-lens reflex camera and a tripod that supported the device at 37 cm above the sample. The camera was configured with automatic settings. The image sensor comprised 16.2 megapixels and there was a high quality lens, which was optimally engineered to gather more light.

Dataset 2 was created with a smartphone which has a 5.0 megapixel camera, and was also configured with automatic settings. Image acquisition distance was about 37 cm and it was stabilized manually with no tripod.

After acquisition, all of Dataset 1 images were adjusted to a $1550 \times 1700$ window and Dataset 2 images to a $1000 \times 1100$ window, before exporting them in both .bmp and .png formats for different purposes.
3.3 Performance measures

Performance comparisons were not made based on final marbling definition to avoid traditional processes subjectivity, but instead, we used a model that delimited fat area in image as precisely as possible. Fat area is the main interest in marbling analysis, i.e., results are concerned with the area which is considered to be fat instead of pattern levels analyzed by professionals.

Thus, output of the algorithm was fat segment detected relative to fat amount in gold standard images rather than the sample marbling level. Thus, images were marked to show what should be considered marbling fat and the output after segmentation was compared with these marks.

![Acquired image](image1.png) ![Gold standard for the acquired image](image2.png)

Figure 5: Gold standard example

Figure 5b shows an example with marks included, as well as the original image used as reference for marking the image in Figure 5a. Marks were made under supervision of a professional who was also involved in acquisition process.

Marked images shown in Figure 5b were used as a gold standard for comparison with images obtained using algorithms described in Akbari and Fei [1], where each pixel received labels shown in Table 1.

<table>
<thead>
<tr>
<th>True</th>
<th>False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Fat identified</td>
</tr>
<tr>
<td>Negative</td>
<td>Muscle identified</td>
</tr>
<tr>
<td></td>
<td>Muscle and Brightness identified as fat</td>
</tr>
<tr>
<td></td>
<td>Fat identified as muscle</td>
</tr>
</tbody>
</table>

Table 1: Pixel labels

Reference marks were compared with output obtained to evaluate the performance of the proposed computer vision system. The accuracy of the proposed approach was quantified by comparing all images with gold standards. Images were also analyzed using previously described approach [10] in same manner, which successfully achieved automatic meat marbling segmentation using fuzzy-c-means instead of our faster K-means algorithm. Moreover, they proposed a partial solution to brightness problem by removing small flecks and by applying a DoubleGreen contrast enhancer, which was useful to find marbling pixels that were not assigned by K-means. Labels were used for comparison and to perform statistical tests, as described by Lav et al. [17].
4 Results and Discussion

4.1 Performance of the Proposed Approach

The system output was compared with gold standards and true positives rates are shown in Figure 6. Each boxed sequence in Figure 6 shows fat identified correctly by proposed approach and previously described method, which obtained maximum results of 80.72% and 85.96%, respectively. Segmentation results for Dataset 1 are represented by light boxes and other boxes represent true positive results for Dataset 2. Average difference in true positives for Dataset 1 was about 7.04%.

In uncontrolled acquisition conditions using a high-quality device, true positive rates were similar with both approaches. For Dataset 2, as shown in Figure 6, superior results were obtained using the proposed approach when processing images acquired by a low-quality acquisition device. Proposed approach obtained an average true positive rate of 22.04% compared with 10.04% using previously described method.

Poor quality camera and the lack of environmental control when obtaining Dataset 2 yielded images with many interpolated pixels and the lack of contrast also made it difficult for automatic methods to determine threshold values for clustering, thereby leading to this difference in true positive rates.

Moreira et al. [20] defined clustering as the division of data into groups that comprise similar data, or data with a group of features that can be distinguished. To generalize data so that they can be represented as groups, it is necessary to define thresholding values for features, which determine data that belong to a class or group.

Technique described by Otsu was used to determine threshold values in present study [21]. Figure 7 shows threshold values obtained by Otsu’s method for Dataset 1 and Dataset 2, as well as static and empirical values used in previously described study.

![Figure 6: Comparison of the true positives rates with the proposed approach and the previously described method [10]](image)

Processing accuracy, including background subtraction and muscle identification, for proposed approach and previously described method were about 98.61% and 98.40%, respectively, for Dataset 1. For Dataset 2, accuracy rates were 98.28% and 98.08%, respectively, due to lower amounts of pixels that represented fat in the overall image, this comparison is available in Figure 8.
In order to verify the performance of both approaches using images acquired in uncontrolled conditions, we evaluated false positive rates, which are shown in Figure 9. In Figure 9, both lines indicate proportional growth in false positives, but proposed approach produced results lower than those obtained by previously described method. Thus, our approach facilitates better recognition when handling images with dark muscle and brightness.

Figure 10 shows the precision when processing Datasets 1 and 2. In both experiments, results obtained by proposed approach were better than those produced by previously described method. Average result with proposed approach were 91.35% with no outliers, since previously described method yielded less than 85.35% with outliers.

Absence of outliers in precision results indicates the robustness of the proposed method. The performance of proposed approach was better with Dataset 2, where it obtained higher precision rates in uncontrolled acquisition conditions with low quality devices. Therefore, proposed approach satisfied the objective of reducing the need for a controlled environment and high quality devices, as well as addressing the major issues related to the acquisition process to yield a robust solution.

Low precision of both approaches, i.e., 40.89% and 26.47% for proposed approach and previously described method, respectively, was due to low level of marbling fat in samples used in this study and acquisition conditions (i.e., brightness and light conditions).
Table 2 shows robustness of proposed approach, which is attributable to the ability of the model to obtain results independently of meat marbling quality, device settings, and other controls. Mean standard deviation in precision with proposed method (4.47%) compared with previously described method (10.72%) demonstrates greater robustness of our method. Based on segmentation precision and meat marbling quality, it is clear that high marbling quality yields better results in terms of true positive rate and thus greater precision. Highest precision results were obtained using proposed approach.

Table 2: Mean standard deviation in the meat marbling quality

<table>
<thead>
<tr>
<th>Approach Precision</th>
<th>Marbling Meat Quality Degree</th>
<th>Mean Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low D1</td>
<td>Medium D1</td>
</tr>
<tr>
<td>Proposed</td>
<td>30.06%</td>
<td>38.90%</td>
</tr>
<tr>
<td>Literature</td>
<td>24.64%</td>
<td>38.51%</td>
</tr>
</tbody>
</table>

4.2 K-means Clustering

Clustering processes are the traditional choice for understanding and learning the characteristics of data, as described in Jain [15] and specifically for food quality in Jackman et al.[10] and [11], and Sun [23]. We used K-means algorithm to determine the similarity of the features of muscle, fat, and background based on color to allow meat marbling segmentation. However, results of this study had moderate precision compared with visual evaluations for meat with a low amount of marbling. This indicates that K-means has limitations when using only RGB color space as descriptors to segment fat in an illumination varying environment where brightness, shadows, and other acquisition conditions generate noise.
Figure 10: Precision of the previously described method [10] and the proposed approach with Datasets 1 and 2

Figure 11: Pixels color distribution using K-means
Figure 11 shows three representations of K-means partitions, i.e., muscle, fat, and background, for one sample. Based on red and blue perspective (Figure 11a), and red and green perspective (Figure 11b), we can see that there is a high level of discrimination between classes; thus, these colors are good descriptors. However, in blue and green perspective (Figure 11c), cluster disjunctions are not consistent, thereby reducing overall precision. In order to improve discrimination level between classes we can use other image descriptors, such as regions [25] [23] and differences between color channels [22]. K-means algorithm is one of the best known non-iterative clustering methods for unsupervised learning [6], but given the use of gold standard in our methodology, we could apply supervised learning modeling. A supervised learning method could improve results and different types of evaluation approaches might be used, such as discriminative paraconsistent machine [9] or traditional artificial neural networks [3, 25, 23].

5 Conclusion

Our proposed approach could segment fat in muscle from various marbled meat samples in an illumination varying environment with external ambient light and artificial light. This is a great advance compared with traditional analyses (visual and subjective), where experts must rely on patterns to assign the marbling of samples, because the analysis performed by each evaluator is prone to errors and inter-operator differences.

Current state-of-the-art for marbling meat evaluations based on computer vision includes several methods related to light, brightness, high computational costs, and high precision based on specific equipment and controlled conditions. However, we propose an approach that preserves features of similar methods but we add greater robustness in uncertain acquisition environments, thereby helping to avoid human inconsistency and subjectivity in evaluations.

Furthermore, static empirical variables removal allows the algorithm to obtain information requisite automatically from samples, thereby allowing it to become independent of environment or device to generate dynamic thresholds in a robust process, which can also utilize low cost devices in computer vision system.

6 Acknowledgments

The authors would like to thank CAPES, Fundação Araucária (Brazilian Agencies) and CNPQ (grant #420562/2018-4) for financial support. We are grateful to Grupo de Pesquisa e Análises de Carne (GPAC) for sharing meat marbling samples.

References


