An intelligent machine vision-based smartphone app for beef quality Evaluation

Soleiman Hosseinpour ∗ , Ali Hakimi Ilkhchi, Mortaza Aghbashlo

Nanobioelectronics Lab., Department of Mechanical Engineering of Agricultural Machinery, Faculty of Agricultural Engineering and Technology, College of Agriculture and

Natural Resources, University of Tehran, Karaj, Iran

Keywords: Android app，Beef tenderness，Image invariants，Machine vision，Smartphone

Beef tenderness is the most important attribute correlated with beef quality, consumer satisfaction, and purchasing decisions. Nowadays, a rapid, non-invasive, and non-destructive evaluation and prediction of beef tenderness and quality from fresh product attributes is desired in industries, laboratories, and markets dealing with beef handling, processing, analyzing, and buy and sell. In this study, a new machine vision-based smartphone app was developed and verified for the first time in order to predict beef tenderness from fresh beef image captured under uncontrolled conditions. In order to eliminate the effects of uncontrolled imaging conditions, an illumination-, rotation-, scale-, and translation-invariant image processing algorithm was developed so that acommon user can easily capture the image of the beef sample with more degree of freedom in terms of luminance, rotation, scale, and translation with no worries about the accuracy of the results. The obtained preprocessed image textural features were well correlated with instrumental data obtained using Warner-Bratzler shear force measurement through artificial neural network technique. The developed android app was installed on a LG G4 H815 smartphone and its performance was assessed using thirty unseen beef samples. The prob-ability of occurrence of 2-D correlation coefficients obtained from the analyses of all the beef samples subjected to the image processing algorithm showed the average probability of 0.92, which strongly supported the ro-bustness of the developed algorithm. The best obtained neural network model could predict the tenderness values with mean absolute percentage error (MAPE) of 3.28% and coefficient of determination (R 2 ) of 0.97. The app promisingly predicted the beef tenderness values of the unseen samples with mean squared error (MSE) of 3.34, MAPE of 3.74%, and R 2 of 0.99. Accordingly, the developed app can be a low-cost and user-friendly tool for

predicting beef tenderness and quality from its real-world image.

## Introduction

Beef eating quality and palatability attributes are mostly characterized by tenderness (texture), juiciness, and flavor (Aaslyng, 2002).These sensory properties are important for customer satisfaction and making purchasing decisions (Platter et al., 2005). In the case of

quality, tenderness is one of the most important attributes regarding sensory experience and eating satisfaction. According to Platter et al. (2005), consumers are willing to pay more for tender beef products. Tenderness is evaluated by objective and subjective methods like

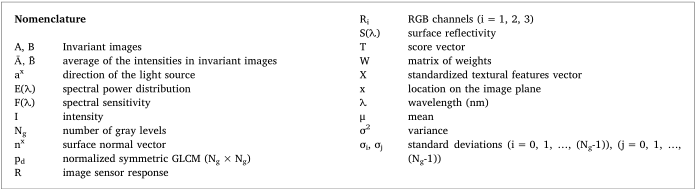
Warner Bratzler shear force (WBSF) and consumer panel, respectively (Destefanis et al., 2008), but, the majority of these methods are invasive and destructive assessment approaches and need sample preparation. Several relatively new methods like visible, hyperspectral imaging

(ElMasry et al., 2012), spectroscopy (Bowling et al., 2009), ultrasonic technology (Tait, 2016), and X-ray computed tomography (Prieto et al., 2010) have been developed and used for beef tenderness prediction;however, most of these methods are not cost-effective and need so-

phisticated instruments.

In meat industry, a rapid, non-invasive, and non-destructive evaluation and prediction of meat quality from fresh product attributes is desired. Machine (computer) vision technology is a powerful and widely used tool for food quality inspection because of being reliable,

robust, cost-effective, non-invasive, and non-destructive (Hosseinpour et al., 2014). Therefore, numerous researches have been devoted in the literature where image processing technique has been used to predict the beef quality (Li et al., 2001; Jackman et al., 2010). Beef image features like color, marbling, and texture can excellently reflect the beef quality (Xiong et al., 2014). Notably, surface image texture associated with connective tissue contents and fiber bundle size in beef mussel is an important indicator of beef tenderness (Swatland, 2006; Jabri et al.,2010).



Human-smartphone interactions are growing day-by-day and affecting many aspects of our lives. Smartphones are the forerunner generations of cell phones equipped with powerful processors, built-in sensors, larger storage, wireless communications, and standard open

source software, enabling them to sense and understand their environment. Accordingly, smartphones have created new research op portunities to design and implement android apps in many domains such as healthcare (Higgins, J.P., 2016; Jacobs et al., 2017), medical and biotechnology (Zhang et al., 2016; Liao et al., 2016), agriculture(Han et al., 2016; Vesali et al., 2015), food industry (Yu et al., 2015;Masawat et al., 2105), medicine (Bueno et al., 2016; Cho et al., 2015),environmental monitoring, biosecurity and bioterrorism (Hutchison et al., 2015), tourism (Law et al., 2018), and internet of things (Fitzgerald et al., 2018).

In recent years, there has been a growing interest among researchers to develop image processing applications using smartphone's camera (Chung et al., 2018; Shrivastava et al., 2018). This could be attributed to the fact that the cost of smartphones is low, while their processing capabilities are enhanced thanks to the advanced processors, high-resolution cameras, and memory storage devices. Even though smartphone-based vision has created new applications in image processing and artificial intelligence, there are still many challenges to be addressed before the commercialization of such android apps. Various influential parameters like illumination, viewpoint (rotation, scaling, and translation), camera parameters (aperture, shutter speed, white balance, focus, and ISO), and hardware limitations and diversity can profoundly affect the performance, accuracy, and complexity of the smartphone-based vision systems. In order to address these issues, various approaches have been developed and employed like imaging under controlled conditions (Choodum et al., 2013) and using advanced image processing algorithms (Casanova et al., 2013). However, it is difficult or even impossible to control the imaging environment in smartphone applications. Invariant features extraction is still a key challenge in pattern recognition and image processing. Accordingly, several image processing algorithms have been developed to extract

invariant features from images captured under uncontrolled conditions in order to improve the performance of vision systems (Legaz et al., 2018; Zhu et al., 2017), but the developed algorithms have rarely been used in smartphone-based image processing (Casanova et al., 2013).

Therefore, this study was conducted for the first time to create a new smartphone app on the basis of fresh beef image captured under non-standardized and uncontrolled conditions for estimating tenderness (WBSF value) of the fresh beef sample. To this end, an illumina-

tion-, rotation-, scale-, and translation-invariant image processing algorithm was developed for the first time to extract invariant image texture features from fresh beef images. The obtained features were then correlated with Warner Bratzler shear force (WBSF) data using a

machine learning technique namely artificial neural network (ANN) model. The algorithm was used to build a user-friendly smartphone app for real-time prediction of beef tenderness under real-world imaging conditions and afterwards, the prediction accuracy of the app was

confirmed using thirty unseen meat samples. Using this app, consumers will be able to predict the required force for chewing the beef intended for buy just at the time of purchase. Finally, the developed app has a great potential to be applied in markets by common consumers and in laboratories and meat industries by specialists for evaluating beef quality.

## Materials and methods

### Samples preparation

The muscle samples of 167 carcasses were taken from a local butcher shop. The M. longissimus dorsi muscles of the all carcasses were removed in the store. All the muscle samples were divided into two groups including 137 and 30 carcasses as the training and testing samples, respectively. To carry out experiments, 2.5-cm-thick steak samples were taken from 12th/13th rib interface over the longitudinal direction of the muscle. Therefore, the surface of the beef samples was smooth as much as possible. The fresh samples were used in experiments immediately without storage. First, the required images were taken from beef samples by smartphone. The samples were then broiled to measure the meat tenderness using WBSF standard according to the research guidelines for cookery methods as prescribed by American Meat Science Association (AMSA, 2005).

### 2.2 Image acquisition

A LG G4 H815 smartphone with high resolution CCD sensor camera (5312 × 2988 pixels, 1/2.6″ inches, f/1.8) was used to capture the fresh beef images. The images with resolution of 2976 ×2976 pixels were captured intentionally under different conditions in terms of il-

lumination, rotation, distance, and translation between the camera and the sample planes. The images were captured without any background (close up images) from the samples. The close up image does not require any preprocessing steps because the entire image is the region of interest.

### 2.3 Features extraction

Since there exist positional and lighting variations during image acquisition with smartphones, illumination- and affine-invariant (i.e., translation-, rotation- and scale-invariant) methods should be applied to remove the effects of such uncontrolled conditions and to extract robust textural features. The invariant feature extraction algorithm proposed herein consists of three subroutines including illumination invariant, affine invariant, and textural features extraction algorithms. The overall feature extraction procedure can be divided into the following steps:

i) the captured color images were resized to 25% of the original size (resized to 744 ×744 pixels) to increase time efficiency of the algorithms,

ii) median filter (3 × 3) was employed to eliminate the environmental noises from the images as well as salt and pepper noises which can be caused by sharp and unforeseen disturbances in the image due to, for example, some defects in the CCD or in the transmission of the image. The mechanism of the median filter is to run through the image array by array, replacing each entry with the median of neighboring arrays,

iii) color images were converted from RGB space to the illuminationinvariant gray level space,

iv) the translation-, rotation-, and scale-invariant space were derived from the illumination-invariant gray level space, and

v) the textural features were extracted from the last obtained space.

#### 2.3.1. Illumination-invariant space

Illumination-invariant color space was achieved using the method proposed by Maddern et al. (2014). This invariant space is extracted from the camera response function, showing the response of an image sensor R with a spectral sensitivity function F(λ) to the light I reflected

from a scene with surface reflectivity S(λ) under illumination source with spectral power distribution E(λ) as follows:



where a^x and n^x represent the direction of the light source and the surface normal vector, respectively (Fig. 1). By modeling the spectral sensitivity function as a Dirac delta function and separating the components of Eq. (1) using logarithmic transform, the camera response function simply changes to the following equation:



where G is a^n ·x^x called the geometry factor. Maddern et al. (2014) extracted one-dimensional gray level space I from Eq. (2) as follows:



where R1 , R2 , R3 are the sensor responses in the three RGB channels. The α value is determined by having the values of the three wavelengths corresponding to the peak of sensitivities (λ 1 < λ 2 < λ 3 ) as follows:



In order to determine the α value and, consequently, illuminationinvariant gray level space, peak values of the sensor spectral responses are required. Fig. 2 shows the details of the procedure used for deriving illumination-invariant gray level space from an image in the RGB color space.

#### 2.3.2. Rotation-, scale-, and translation-invariant space

To achieve a rotation-, scale-, and translation-invariant space, Fourier-Mellin transform was applied to the obtained illumination-invariant gray level space. As illustrated in Fig. 3, Rotation-, scale-, and translation-invariant algorithm is composed of three successive trans-

forms including Fourier transform, log-polar transform, and Fourier transform. The log-polar transform is the most popular Coordinate transformation (Asselin and Arsenault, 1994), mapping rotation and scale to translation because of its topological nature. On the other hand, translation-, rotation-, and scale-invariant space can be achieved based on the shift theorem of the Fourier transform. Fourier shift theorem states that a translation in the spatial domain corresponds to a linear phase term in the frequency domain.

To investigate the robustness of the developed image processing algorithm, the image capturing procedure of each beef sample was performed in three levels of luminance (400, 700, and 1000 lx) three levels of rotation (0°, 45°, and 90°), three levels of scale (0.5, 0.8, and

1), three levels of translation (0, 25, and 50% of the image width) in each horizontal and vertical direction in the image plane, and three replications. Then, robustness of the image processing algorithm was measured via 2-D correlation coefficient (Eq. (5)) of each pair of images for each beef sample captured in different conditions in terms of illumination, rotation, scale, and translation.



where, A ¯ and B ¯ are respectively the average of the intensities in images A and B in the illumination-, rotation-, scale-, and translation-invariant space. Finally, statistical analysis was performed using MATLAB software (Release, 2016b) to find the effect of replications on the 2-D correlation coefficients.

#### 2.3.3. Texture features

Gray level co-occurrence matrix (GLCM) is the most popular and the first statistical method used for analyzing texture of food images (Zheng et al., 2006). A GLCM is a probability distribution matrix derived based on second-order statistical technique presented by Haralick and Shanmugam (1973). The widely used GLCM technique was considered in this study for extracting image textural features. Various textural features such as variance, correlation, homogeneity, energy, entropy, sum entropy, and sum variance (Eqs. (6)–(12)) and some statistical features such as mean, variance, and entropy were respectively extracted from the amplitude and real parts of the normalized GLCM of the obtained illumination- and affine-invariant space for encoding the randomness, linearity among pixels, pixel similarity, and textural uniformity of the space:

