DIETARY ASSESSMENT AND OBESITY AVOIDANCE SYSTEM BASED ON VISION: A REVIEW

Salwa Khalid Abdulateef[1,2], Massudi Mahmuddin[1], Nor Hazlyna Harun[1], and Yazan Aljeroudi[3]

[1]*School of Computing, College of Art and Sciences, Universiti Utara Malaysia, Malaysia*
[2]*College of Computer and Mathematics Sciences , Tikrit University, Iraq*
[3]*International Islamic University, Malaysia*
khalid.salwa@yahoo.com[1,2], ady@uum.edu.my[1], *hazlyna@uum.edu.my*[1], *yazan.aljeroudi@gmail.com*[3]

ABSTRACT. Using technology for food objects recognition and estimation of its calories is very useful to spread food culture and awareness among people in the age of obesity due to the bad habits of food consumption and wide range of inappropriate food products. Image based sensing of such system is very promising with the large expanding of camera embedded portable devices such as smartphones, PC tablets, and laptops. In the past decade, researchers have been working on developing a reliable image based system for food recognition and calories estimation. Different approaches have tackled the system from different aspects. This paper reviews the state of the art of this interesting application, and presents its experimental results. Future work of research is presented in order to guide new researchers toward potential tracks to create more maturity and reliability to this application.

Keywords: Pattern recognition, image segmentation, classification, statistical estimation

INTRODUCTION

The nature of the current style of life and lacking of time of preparing food with proper amount of calories and fat, all of those factors are creating an increasing concern of using technology to serve in this matter. Balancing food quantitative content of calories, fat, sugar, vitamins, minerals and other micronutrients is an efficient way to avoid obesity and to maintain body health. Obesity is defined by the World Health Organization (WHO) as abnormal or excessive fat accumulation that may risk human being health in all different ages (Lockwood, 2013). Many researchers in health sectors associate obesity with diabetes, heart disease, cancer, high blood pressure and high cholesterol (Finkelstein, Trogdon, Cohen, & Dietz, 2009). Moreover, obesity is very risky of elderly people life due to difficulty of moving and danger of falling and causing injuries. Statistically, in 2014, more than 1.9 billion adults, 18 years and older, were overweight. Over 600 million of these were obese. Besides, obesity is expanding in new generation; 42 million children under the age of 5 were overweight or obese in 2013 (WHO, 2015).

Historically, manual approaches have been used for the purpose of food dietary assessment or calories estimation. Examples are: food records, 24 hours dietary recall, and food frequency. The most common drawback of these manual methods is the problem of underreporting due to the complexity of recording food and lacking of needed knowledge in ordinary people to take care of this matter. One promising technology for food intake advisory system is vision-based techniques. Actually, vision-based is very promising to be provided to individual level due to the accurate camera, the high computational power, and low cost of this technology that is provided in almost every smartphone. Besides, this technology is non-invasive which makes it very practical for all ages, and cases of health conditions (Wilson & Sarin, 2007).

The remaining of the paper is organized as follows: in Section 2, a framework for the system is presented. All the phases are introduced with the state of the art methods of tackling them. Section 3, states the conclusion and the future direction of research in this application.

BUILDING FRAMEWORK FOR AN IMAGE-BASED DIETARY ASSESSMENT

The application of image-based dietary assessment is still recent topic. Researchers have tackled this application from different perspectives. Thus, we see it is very appealing to build taxonomy to present the general image of the state of the art. In our taxonomy, the system of image-based dietary assessment is presented as phases. Each phase has different approaches, and the overall aggregation of the phases is reflected in the performance of the system.

In the Figure 1, a diagram shows the general phases of the system. As it can be seen, the system is arranged into 5 main phases. Each phase performance is based significantly on the outcome of the previous phase. In the next subsection, present briefly the different aspect of phase. Also, it should be noted that the current taxonomy and presentation is the first one in the literature for an image-based dietary assessment.

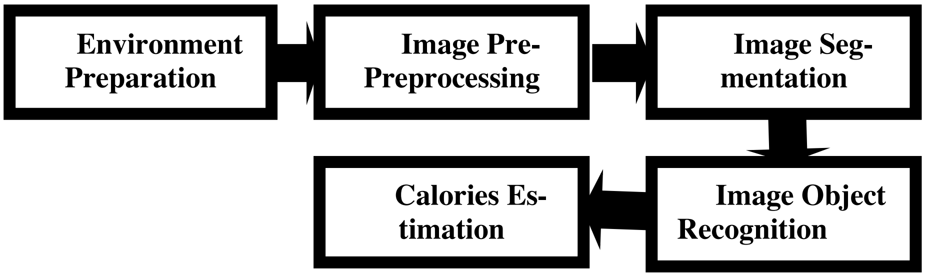

Figure 1. General phases of the system

Environment Preparation

Pouladzadeh, Shirmohammadi, & Al-Maghrabi (2014) have required the thumb of the user to exist in the scene of the environment in order to be used as a reference. Others have proposed known fiducial object in the scene (Zhu et al. 2008). Another approach was used by Martin, Kaya, and Gunturk, (2009) was based on calibration card that was used as reference card in the pictures to account for the viewpoint and distance of the camera. Also, it is possible to use dimension of a plate or a cup, and prepared tablecloth such as chessboard type.

Image Pre-processing

Different classical approaches of image processing has been applied by some researchers to remove noise (Villalobos, Almaghrabi, Pouladzadeh, & Shirmohammadi, 2012). Some researhcers have relied on morphological to remove noise (Patel, Jain, & Joshi, 2012). Blurring removal techniques have been used by (Neelamegam, Abirami, Priya, & Valantina, 2013; Xu, He, Khanna, Boushey, & Delp, 2013). Some researchers have used an interactive approach to detect blurring by asking the user to take the image again (Ahmad, Khanna, Kerr, Boushey, & Delp, 2014). In our perspective, we see that blurring removal can be performed automatically with good performance if an advantage is taken of the inertial sensing that is available most smartphones. Thus adding more autonomy to the system.

Other appraches have been developed to to remove shadows, and light reflection that arises (Sun & Du, 2004). One difficulty of food image is specular highlight, which occurs on food plate because of illumination nature of the environment and reflection nature of the plate. Some researchers have developed specular highlight removal on food images. He, Khanna, Boushey, and Delp (2012b) used an assisting grey scale image generated from the original image, and an independent component analysis for removing specular highlight from food image.

Image Segmentation

This part is probably the most difficult and challenging part of image-based dietary assessment. Different approaches have been developed in the image segmentation field in general, and in the food image segmentation application in special. Some researchers have performed comparative study between three state of the art methods of image segmentation in the food image application study (He, Khanna, Boushey, & Delp, 2013). Three methods of segmentation on food images: active contour, normalized cuts, and local variation. Their study showed that local variation was more consistent in terms of performance with changing the configuration parameters comparing with the other two methods. Also, active contour has a performance similar to local variation as the number of initial contour increases, with a drawback of overhead of computational complexity. From this study, it can be seen that initial information provided from prior knowledge such as the number of segments are an essential part of the functionality of the three methods. Typically, in segmentation methods there is a concern about the computational time of the execution. This study does not show any measure of the computational complexity of the three methods.

Another concern in image segmentation that has been tackled is the dealing with the aspect of under or over segmentation. In He, Xu, Khanna, Boushey, and Delp (2013) an iterated food segmentation method. The iterated food segmentation has used feedback based on the classifier confidence score in order to prevent under or over segmentation.

Some researchers have presented image segmentation as an optimization problem. He, Khanna, Boushey, & Delp (2012a) used snakes or active contours for food images segmentation. The basic idea is to optimize an energy function for a selected area in the image by the contour. The global minimum of the energy function is reached when the contour is aligned with the edges of the food objects. They improved the basic method by starting from equally distributed circles as initial contours. However, this approach is limited to evenly colored food objects. In other words, the energy function will not reach its minimum value at the borders of the food objects when there are internal edges in the objects surface.

Graph cut segmentation was used by Pouladzadeh, Shirmohammadi, and Yassine (2014). This approach is also formulated as an optimization problem of energy (cut) at the object boundary. There is real concern about the computational complexity of this method. Besides, it is applicability on mixed food objects.

Stick Growing and Merging (SGM) approach was used by Sun and Du (2004) for segmenting complex images of many types of foods including pizza, apple, pork and potato. In this approach, the food image itself is segmented by new region growing-and-merging. Pizza is partitioned into cheese, ham, tomato sauce. Before segment made preprocessing to remove shadows, and light reflection. Sobel was applied to detect the edge of the objects. The objects with clear edges usually obtain smooth boundaries, but the objects with inside inhomogeneous character have less smooth boundaries and larger number of sub-regions. The method is fast if the data volume is reduced. The limitation is minimal stick length and dealing with the images that have less texture.

Pyramidal mean-shift filtering, a region growing algorithm and region merging has been used for image segmentation by (Anthimopoulos, Dehais, Diem, & Mougiakakou, 2013), this method was evaluated only on carbohydrate food type with plate and background the same color.

Image Object Recognition

This part can be seen from two perspectives: the type of features that are extracted, and the type of classifier that is involved.

In terms of features, most researchers have extracted colors, circles and scale-invariant feature transform (SIFT) features from each image (Kitamura, Yamasaki, & Aizawa, 2008 ; Morikawa, Sugiyama, & Aizawa, 2012 ; Aizawa, Maruyama, Li, & de Silva, 2013), others have extracted color histograms and Discrete Cosine Transform (DCT) coefficients and SIFT features (Aizawa, de Silva, Ogawa, & Sato, 2010 ; Kong & Tan, 2012), or SIFT on the on the Hue Saturation Value (HSV) color space (Anthimopoulos, Gianola, Scarnato, Diem, & Mougiakkou, 2014), others have extracted color histogram, color correlograms and Speeded Up Robust Features (SURF) features (Miyazaki, de Silva, & Aizawa, 2011; Yang, Chen, Pomerleau, & Sukthankar, 2010), shape features have been extracted and used by Shroff, Smailagic and Siewiorek (2008), Pairwise local features have been extracted and used by (Yang et al., 2010).

In Villalobos et al. (2012) and Pouladzadeh, Shirmohammadi, and Arici (2013) have extracted texture, size, color and shape features. In our opinion, it is not simple to decide which type of features is best for providing meaningful information for classifier about the food content or type in the image. This type of problem should be addressed from quantitative perspective, in other words, there should be a comparative study of the influence of feature selection on the performance of the classifier.

The feature selection has an impact on the accuracy. Bosch, Zhu, Khanna, Boushey, and Delp (2011) have investigated the real effect of features selection on the accuracy of the classification of food objects. The finding was combining global and local features extracted from food segments give more discriminative information about food objects types. Other more specific type of features that have been used Color features: Scalable Color Descriptor (SCD), Color Structure Descriptor (CSD), Dominant Color Descriptor (DCD), and Color Layout Descriptor (CLD) are extracted, while for texture features Gradient Orientation Spatial-Dependence Matrix (GOSDM), Entropy Based Categorization and Fractal Dimension Estimation (EFD), and Gabor Based Decomposition and Fractal Dimension Estimation (GFD) (He et al., 2013).

In terms of classifier type, Support Vector Machine (SVM) was the most used type of classifier among all different types of classifiers (Pouladzadeh et al. 2013; Kitamura, Yamasaki, & Aizawa 2008; Yang, Chen, Pomerleau, & Sukthankar 2010; Maruyama, de Silva, Yamasaki, & Aizawa 2010; Anthimopoulos et al. 2014; Pouladzadeh et al. 2014).

Other researchers have used Bayesian probabilistic for classification (Kong & Tan, 2012 ; (Aizawa et al., 2013 ; Maruyama et al., 2010), others have used Radial Basis Function (RBF) kernel for classification (Anthimopoulos et al., 2013), Some approaches have used feed forward Neural Network (NN). Back propagation learning algorithm has been used for training NN (Shroff et al., 2008; Savakar, 2012).

It is important to notice that the accuracy of the classification that was reported in previous works is not reliable indicator to compare between the classifiers due to different factors.

Some of them are related to the simplicity of the scenarios that is conducted for validation in some approaches and because of providing different assumptions for the experiment. But most importantly, some of the methods were not really food objects classifications as food categorizations such as developed the web application system named Food log (Maruyama et al., 2010).

Calories Estimation

The final stage in any food dietary assessment system is calories estimation. As stated earlier, this stage accuracy is based on the previous stages. Some researchers have used simple look-up table approach, which means each food objects is corresponding to a pre-determined amount of food calories (Wu & Yang, 2009), while other researchers have incorporated more algorithmic work to calculate precisely the amount of calories based on different variables such as food size, color, and shape (Pouladzadeh et al., 2014). Some approaches has used linear regression to produce the estimation (Miyazaki et al., 2011)

Interactivity aspect in image based food dietary assessment

From a deep review of the state of the art methods it can be noticed that there are many challenging factors. Some of them are the difficult nature of this application in terms of non-regular shape, color, or definition of specific food objects as well as the environment‘s variability impact on the performance of the system. Therefore, and for the sake of meeting a minimum level of quality for the user a lot of researchers have relied on providing the system with an interactive scheme.

Morikawa et al., (2012) have provided user interaction for improvement of the accuracy of segmentation based on touch point to provide manual segmentation. Another interactive way was used by Kawano and Yanai (2013) where the user has to enter a bounding box around the food object.

In our perspective, this interactivity was most used in the phase of segmentation. The reason is that this phase is the most challenging part in the whole process. Besides, accurate segmentation results imply good shaped and partitioned food objects, and as a result, this implies accurate classification result.

CONCLUSION AND DISCUSSION

Based on the previous literature review, it can be clear that all approaches follow a systematic framework that consists of 5 phases: environment preparation, image pre-processing, image segmentation, object recognition, and calories estimation. The essence of the procedure of image based food assessment system resides in the part of object classification. Actually, there is no total agreement on the most efficient type of features or classifier to be used in the object recognition part. But good and accurate object classification requires accurate object segmentation. Therefore, the most essential and challenging aspect of this application is object segmentation and classification. Unfortunately, the non-well predefined shape and characteristic of food objects causes add hurdles that researchers are trying to overcome. Some researchers have tried to overcome the difficulty of object segmentation by adding an interactive approach to the scheme, while others have just avoided this challenge by simplifying the testing scenarios to very distinctive and well-defined food objects scenario. Our recommendations are to use interactivity in very light way in order to sustain a minimum level of autonomy in this the application. In our perspective, reinforcement and adaptive learning can be incorporated to improve the performance.

REFERENCES

Ahmad, Z., Khanna, N., Kerr, D. A., Boushey, C. J., & Delp, E. J. (2014). A mobile phone user interface for image-based dietary assessment. In *Proceedings of the IS&T/SPIE Conference on Mobile Devices and Multimedia: Enabling Technologies, Algorithms, and Applicationstrons*. doi:10.1117/12.2041334

Aizawa, K., de Silva, G. C., Ogawa, M., & Sato, Y. (2010). Food log by snapping and processing images. In *Virtual Systems and Multimedia (VSMM), 2010 16th International Conference on* (pp. 71–74). IEEE. doi:10.1109/VSMM.2010.5665963.

Aizawa, K., Maruyama, Y., Li, H., & de Silva, G. (2013). Food Balance Estimation by Using Personal Dietary Tendencies in a Multimedia Food Log, *15*(8), 2176–2185. doi:10.1109/TMM.2013.2271474

Anthimopoulos, M., Dehais, J., Diem, P., & Mougiakakou, S. (2013). Segmentation and recognition of multi-food meal images for carbohydrate counting. In *2013 IEEE 13th International Conference on Bioinformatics and Bioengineering (BIBE),* 1–4. IEEE. doi:10.1109/BIBE.2013.6701608.

Anthimopoulos, M., Gianola, L., Scarnato, L., Diem, P., & Mougiakkou, S. (2014). A Food Recognition System for Diabetic Patients based on an Optimized Bag of Features Model. *IEEE Journal of Biomedical and Health Informatics*, *18*(4), 1261–1271. doi:10.1109/JBHI.2014.2308928

Bosch, M., Zhu, F., Khanna, N., Boushey, C. J., & Delp, E. J. (2011). Combining global and local features for food identification in dietary assessment. In *2011 18th IEEE International Conference on Image Processing (ICIP),* 1789–1792. IEEE. doi:10.1109/ICIP.2011.6115809

Finkelstein, E. A., Trogdon, J. G., Cohen, J. W., & Dietz, W. (2009). Annual medical spending attributable to obesity: payer-and service-specific estimates. *Health Affairs*, *28*(5), w822–w831. doi:10.1377/hlthaff.28.5.w822.

He, Y., Khanna, N., Boushey, C. J., & Delp, E. J. (2012a). Snakes assisted food image segmentation. In *2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP),* 181–185. IEEE. doi:10.1109/MMSP.2012.6343437.

He, Y., Khanna, N., Boushey, C. J., & Delp, E. J. (2012b). Specular Highlight Removal for Image-Based Dietary Assessment. In *2012 IEEE International Conference on Multimedia and Expo Workshops (ICMEW),* 424–428. IEEE. doi:10.1109/ICMEW.2012.80

He, Y., Khanna, N., Boushey, C. J., & Delp, E. J. (2013). Image segmentation for image-based dietary assessment: A comparative study. In *2013 International Symposium on Signals, Circuits and Systems (ISSCS),* 1–4. IEEE. doi:10.1109/ISSCS.2013.6651268

He, Y., Xu, C., Khanna, N., Boushey, C. J., & Delp, E. J. (2013). Food image analysis: Segmentation, identification and weight estimation. In *2013 IEEE International Conference on Multimedia and Expo (ICME),* 1–6. IEEE. doi:10.1109/ICME.2013.6607548.

Kawano, Y., & Yanai, K. (2013). Real-time mobile food recognition system. In *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW),* 1–7. IEEE. doi:10.1109/CVPRW.2013.5.

Kitamura, K., Yamasaki, T., & Aizawa, K. (2008). Food log by analyzing food images. In *Proceedings of the 16th ACM international conference on Multimedia,* 999–1000. ACM. doi:10.1145/1459359.1459548.

Kong, F., & Tan, J. (2012). DietCam: Automatic dietary assessment with mobile camera phones. *Pervasive and Mobile Computing*, *8*(1), 147–163. doi:10.1016/j.pmcj.2011.07.003

Lockwood, W. (2013). Obesity: Methods of Assessment.

Martin, C. K., Kaya, S., & Gunturk, B. K. (2009). Quantification of food intake using food image analysis. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* 6869–6872. IEEE. doi:10.1109/IEMBS.2009.5333123.

Maruyama, Y., de Silva, G. C., Yamasaki, T., & Aizawa, K. (2010). Personalization of food image analysis. In *16th International Conference on Virtual Systems and Multimedia (VSMM),* 75–78. IEEE. doi:10.1109/VSMM.2010.5665964.

Miyazaki, T., de Silva, G. C., & Aizawa, K. (2011). Image-based calorie content estimation for dietary assessment. In *IEEE International Symposium on Multimedia (ISM),* 363–368. IEEE. doi:10.1109/ISM.2011.66

Morikawa, C., Sugiyama, H., & Aizawa, K. (2012). Food region segmentation in meal images using touch points. In *Proceedings of the ACM multimedia 2012 workshop on Multimedia for cooking and eating activities*, 7–12. ACM. doi:10.1145/2390776.2390779.

Neelamegam, P., Abirami, S., Vishnu Priya, K., & Rubalya Valantina, S. (2013). Analysis of rice granules using image processing and neural network. In *IEEE Conference on Information & Communication Technologies (ICT),* 879–884. IEEE. doi:10.1109/CICT.2013.6558219

Patel, H. N., Jain, R. K., & Joshi, M. V. (2012). Automatic Segmentation and Yield Measurement of Fruit using Shape Analysis. *International Journal of Computer Applications*, *45*(7), 19–24. doi:10.5120/6792-9119.

Pouladzadeh, P., Shirmohammadi, S., & Al-Maghrabi, R. (2014). Measuring calorie and nutrition from food image. In *IEEE, Transactions on Instrumentation and Measurement,* (p. 10). IEEE. doi:10.1109/TIM.2014.2303533

Pouladzadeh, P., Shirmohammadi, S., & Arici, T. (2013). Intelligent SVM based food intake measurement system. In *IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA),* 87–92. IEEE. doi:10.1109/CIVEMSA.2013.6617401.

Pouladzadeh, P., Shirmohammadi, S., & Yassine, A. (2014). Using Graph Cut Segmentation for Food Calorie Measurement. In *IEEE International Symposium on Medical Measurements and Applications Proceedings (MeMeA),* 1–6. doi:10.1109/MeMeA.2014.6860137

Savakar, D. (2012). Identification and Classification of Bulk Fruits Image using Artificial Neural networks. *International Journal of Engineering and Innovative Technology*, *1*(3), 36–40.

Shroff, G., Smailagic, A., & Siewiorek, D. P. (2008). Wearable context-aware food recognition for calorie monitoring. In *12th IEEE International Symposium on Wearable Computers,* 119–120. IEEE. doi:10.1109/ISWC.2008.4911602.

Sun, D.-W., & Du, C.-J. (2004). Segmentation of complex food images by stick growing and merging algorithm. *Journal of Food Engineering*, *61*(1), 17–26. doi:10.1016/S0260-8774(03)00184-5

Villalobos, G., Almaghrabi, R., Pouladzadeh, P., & Shirmohammadi, S. (2012). An image procesing approach for calorie intake measurement. In *2012 IEEE International Symposium on Medical Measurements and Applications Proceedings (MeMeA),* 1–5. IEEE. doi:10.1109/MeMeA.2012.6226636.

WHO. (2015). *Obesity and Overweight*. Retrieved from http://www.who.int/mediacentre/factsheets/fs311/en/index.html

Wilson, A. D., & Sarin, R. (2007). BlueTable: connecting wireless mobile devices on interactive surfaces using vision-based handshaking. In *Proceedings of Graphics interface 2007,* 119–125. ACM.

Wu, W., & Yang, J. (2009). Fast food recognition from videos of eating for calorie estimation. In. *IEEE International Conference on Multimedia and Expo,* 1210–1213. IEEE. doi:10.1109/ICME.2009.5202718.

Xu, C., He, Y., Khanna, N., Boushey, C. J., & Delp, E. J. (2013). Model-based food volume estimation using 3D pose. In *20th IEEE International Conference on Image Processing (ICIP),* 2534–2538. IEEE. doi:10.1109/ICIP.2013.6738522.

Yang, S., Chen, M., Pomerleau, D., & Sukthankar, R. (2010). Food recognition using statistics of pairwise local features. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2249–2256. IEEE. doi:10.1109/CVPR.2010.5539907.

Zhu, F., Mariappan, A., Boushey, C. J., Kerr, D., Lutes, K. D., Ebert, D. S., & Delp, E. J. (2008). Technology-assisted dietary assessment. *Proceedings of the IS&T/SPIE Conference on Computational Imaging VI, Vol #6814*, 10. doi:10.1117/12.778616