Technical University of Denmark

DTU

An explorative study on pork loin recognition

Larsen, Anders Boesen Lindbo; Hviid, Marchen Sonja; Larsen, Rasmus; Dahl, Anders Lindbjerg

[Link back to DTU Orbit](#)

DTU Library
Technical Information Center of Denmark

An explorative study on pork loin recognition

Anders Boesen Lindbo Larsen[1], Marchen Sonja Hviid[2],
Rasmus Larsen[1], and Anders Lindbjerg Dahl[1]

[1] Technical University of Denmark
{abll,rlar,abda}@dtu.dk
[2] Danish Meat Research Institute
mahd@teknologisk.dk

Abstract. Bag-of-words (BoW) image description has shown good performance for a large variety of image recognition scenarios. We investigate approaches to alleviating a standard BoW image description pipeline representations for the specific task of recognizing pork loins. Specifically, we extend the BoW description to include depth maps, perform non-rigid image registration to align the images, and apply PCA dimensionality reduction on the BoW descriptors. Our results show that the combination of image registration and PCA yields a more distinctive recognition.

1 Introduction

The goal of our work is to recognize pork loins in order to track them. The motivation behind the project is to facilitate meat traceability in slaughterhouses. In recent years, traceability has become an increasingly important aspect of the meat industry. For consumers, meat safety and quality is a persistent concern strengthened by reoccurring food recalls and scandals as well as increased animal welfare awareness [1].

Currently, meat tracking in slaughterhouses is made possible using RFID tags on carrier devices. However, these carrier devices allow only tracking at batch-granularity as they carry multiple meat cuts. It is not possible to attach RFID tags to individual meat cuts because the risk of losing an RFID tag into the product is too high. In comparison, a robust visual recognition method would be able to accommodate the tracking problem in a non-intrusive manner.

In this work we explore image recognition methods for enabling meat traceability in slaughterhouse environments. We have constructed a baseline method using the popular BoW approach. Compared to standard visual recognition challenges, our dataset is characterized by low inter- and intra-variability of the objects and by trivial background segmentation. We try to exploit these limitations and propose extensions to the baseline recognition algorithm.

2 Dataset

The dataset for our experiment is constructed using 211 pork loins. The photographing setup (see Figure 1a) is the same for both photo sessions. We use a

2 Larsen, Hviid, Larsen, Dahl

Microsoft Kinect camera that captures a depth map along with a standard RGB image of the loin. Examples of both images are shown in Figure 1b. Next to the camera a fluorescent tube is mounted spreading light at a wide angle. A selection

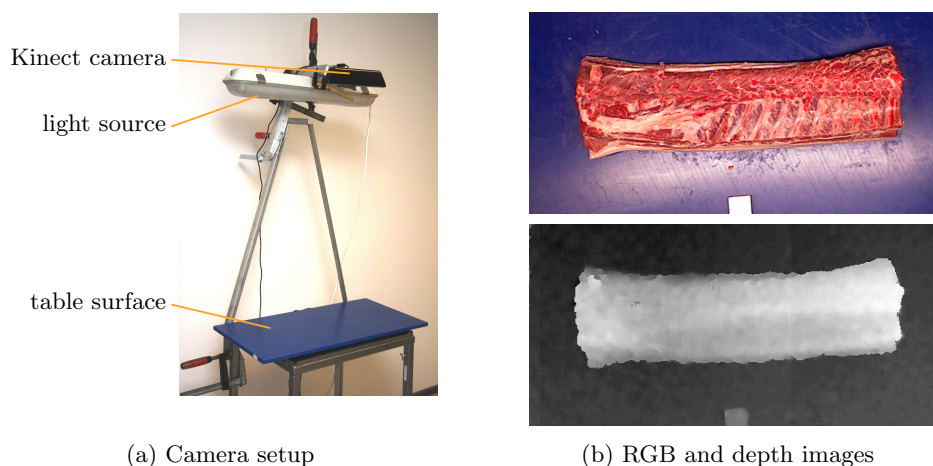(a) Camera setup   (b) RGB and depth images

Fig. 1: Experiment setup and dataset example.

of the loins undergo different perturbation scenarios in an attempt to simulate a slaughterhouse treatment. The perturbations are:

Rough treatment 19 loins are knocked hard onto a table before the second photo session.

Incorrect trimming Pieces of meat and bones are cut off from 18 loins before the second photo session.

Incorrect hanging 19 loins are stored overnight by hanging them sideways on Christmas trees (storage hooks) which causes bends.

Illumination and orientation changes 37 loins are rotated between 45° and 180° around the optical axis before being photographed. This creates variations in lighting because the light falls differently on a rotated object.

3   Baseline algorithm

The basis algorithm is divided into the following 4 steps [2].

1. Segmentation The pork loin is segmented from the background using a Markov random field on the depth image.

2. Canonization The segmented pork loin images are then brought to a canonized form through histogram equalization and orientation detection followed by a rotation to a common orientation. Moreover the RGB images are converted to gray-scale because the color information is mainly in the red channel.

3. **Description** From the canonized images we perform BoW image description by extracting 8 histograms in a 2×4 grid to match the shape of a pork loin. The image features used in the BoW are DAISY descriptors [3] extracted from the gray-scale version of the RGB image.

4. **Matching** We measure the similarity of two pork loin images by calculating the distance between their histograms. For every pork loin from day 1 a match is established to the pork loin from day 2 with the smallest χ^2 distance $\chi^2(\boldsymbol{x}, \boldsymbol{y}) = \sum_{n=1}^{D} \frac{(\boldsymbol{x}(n) - \boldsymbol{y}(n))^2}{\boldsymbol{x}(n) + \boldsymbol{y}(n)}$, where D is the dimensionality of the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ and $\boldsymbol{x}(n)$ is the nth element of $\boldsymbol{x}$.

Note that because the dataset is small, we have used the entire dataset for training, validation and testing.

3.1   Performance

Using the baseline algorithm, all 211 pork loins are recognized correctly. To investigate the sensitivity of the recognition method we want to inspect loins that have been poorly matched in our experiments. We measure the quality of a match by its distinctiveness $d = \frac{d_\mathrm{i} - d_\mathrm{c}}{d_\mathrm{i} + d_\mathrm{c}}$, where d_c is the distance of the correct match and d_i is the distance of the nearest incorrect match. A large d means that the matching pork loin image pair from day 1 and 2 stand out from the rest of the loins. A small d means that there exist a mismatching loin from day 2 with an image description similar to the pork loin from day 1. In Figure 2, we illustrate the distinctiveness statistics for each perturbation scenario. We see that the baseline method is very close to yielding a few mismatches as the distinctiveness of the lowest outliers come close to 0 (a negative value means an incorrect match). However, the main part of the remaining loins is matched with a comfortable margin to the nearest incorrect match. That is, the interquartile range of the distribution of d is above 0.
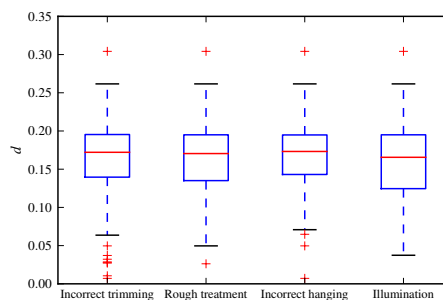

Fig. 2: Box plots showing the statistics of the match distinctiveness d of the baseline recognition method. Rectangles represent the *interquartile range* IQR $=$ Q3 $-$ Q1. The whiskers are placed at Q1 $- 1.5 \cdot$ IQR and Q3 $+ 1.5 \cdot$ IQR. The plusses denote outliers.

4   Larsen, Hviid, Larsen, Dahl

4   Extensions to the baseline algorithm

In the following, we attempt to ameliorate the performance of the recognition algorithm by proposing 3 different extensions.

4.1   Including depth maps

In the baseline algorithm we extract DAISY descriptors from the intensity image only. We wish to investigate if the image description can be improved by appending the BoW histograms from the depth map to the BoW histograms from the intensity images. Compared to the RGB image, the depth image provided by the Kinect camera contains visible noise, see Figure 3. Moreover, the depth image can vary significantly between two photo sessions.

In Figure 5a, the performance of this approach is shown. We see immediately that the depth information does not supplement the intensity information well as performance drops significantly. Therefore, we have not pursued further investigations in this direction.
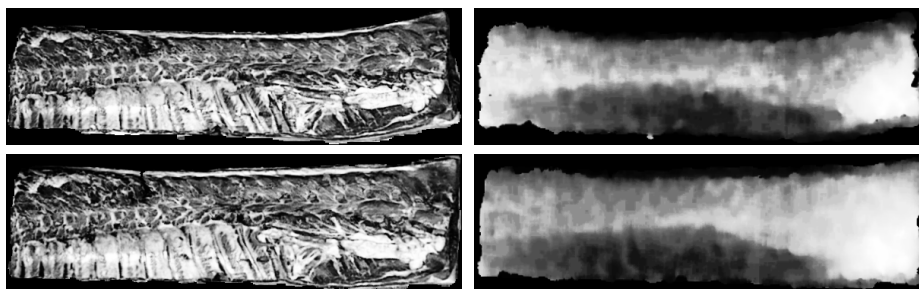

Fig. 3: Canonized images and depth maps of the same pork loin day 1 (top row) and day 2 (bottom row).

4.2   Image registration

Currently, the canonization step assumes that the pork loin is rigid such that only rotation and translation is necessary to align the images. However, in the dataset we have encountered a couple of examples where this assumption does not hold when the loin has been exposed to incorrect hanging or rough treatment. In this extension we introduce non-rigid registration of the loins to achieve invariance towards such perturbations.

Using the pork loin shape generated in the segmentation step, we detect the 4 corners of the pork loin and sample 15 and 6 points along each horizontal and vertical side of the shape respectively. From these points we perform a landmark-based registration using thin plate splines to a target shape selected

among the pork loins. An example of the image warping is shown in Figure 4 In Figure 5b, we see the matching performance using this extension. While the performance seems to improve the problematic cases in the incorrect hanging scenario, the distinctiveness of the incorrectly trimmed loins goes down yielding a single mismatch.

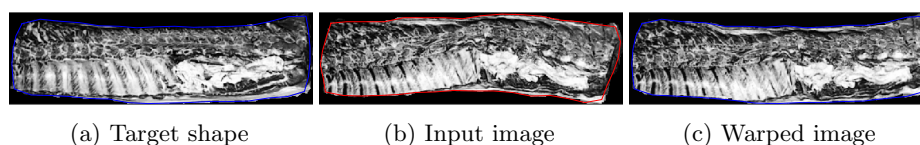(a) Target shape   (b) Input image   (c) Warped image

Fig. 4: Image registration. The blue contour is the target shape generated from the pork loin in (a). The red contour is the shape of the input pork loin.

4.3 PCA-based matching

Inspired by the eigenface approach from facial recognition, we perform a principal component analysis (PCA) from an eigenvalue decomposition of the descriptor covariance matrix. That is, we extract the 120 largest eigenvectors from the covariance matrix of the zero-meaned descriptors in the dataset. Instead of matching loins using the χ^2-distance between their descriptors, we transform the descriptors into the selected eigenvector components (the eigenfaces) and perform a matching in this space using the euclidean distance. The idea behind this approach is to obtain a more robust match caused by the spatial correlation introduced by the eigenfaces. In Figure 5c, the performance of this approach is shown. We see that the loins that have been incorrectly trimmed are more distinctive which makes sense because the eigenfaces are more robust towards local perturbations such as those caused by trimming a small region of the loin.

Finally, we try to combine the PCA-based matching with the image registration and show the result in Figure 5d. This approach looks promising as the eigenfaces are more robust towards the incorrectly trimmed loins that were problematic when performing image registration. Conversely, we suspect that the image registration helps the PCA-based matching because the registration causes a better image alignment which is required for a meaningful PCA.

5 Conclusion

While not all our proposed extensions to the recognition pipeline have shown good results across all perturbation scenarios, we have shown that the constrained nature of our dataset can be exploited to achieve better recognition. Notably, we have achieved invariance towards non-rigid deformations without losing distinctiveness in our image description. This allows for a new range of more flexible
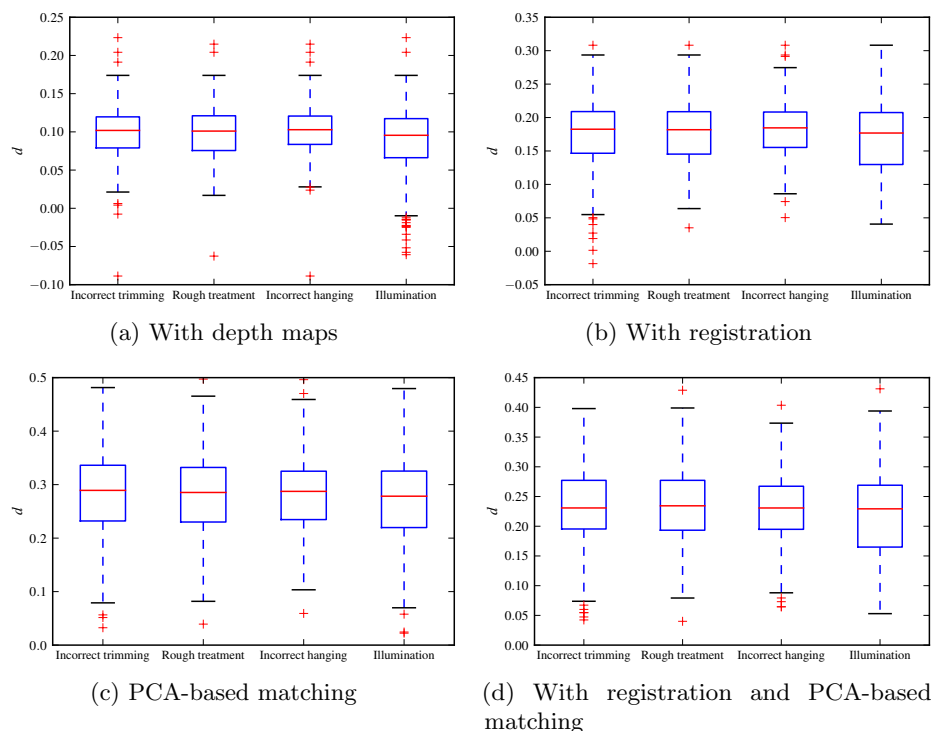
6   Larsen, Hviid, Larsen, Dahl

(a) With depth maps

(b) With registration

(c) PCA-based matching

(d) With registration and PCA-based matching

Fig. 5: Statistics of the match distinctiveness d for our proposed extensions. Note that the y-axis scale on the plots is not comparable between the plots because the descriptors exists in different spaces.

meat products to be recognized. Finally, we should remark that our experiments are carried out on a small dataset which does not allow for a proper statistical analysis of the results. On a brighter note, this study has identified new challenges that would be relevant to investigate in future experiments.

References

1. Trienekens, J., Zuurbier, P.: Quality and safety standards in the food industry, developments and challenges. International Journal of Production Economics **113**(1) (2008) 107 – 122 Research and Applications in E-Commerce and Third-Party Logistics Management.
2. Larsen, A.B.L., Hviid, M.S., Jørgensen, M.E., Larsen, R., Dahl, A.L.: Vision-based method for tracking meat cuts in slaughterhouses. Meat Science (forthcoming)
3. Tola, E., Lepetit, V., Fua, P.: Daisy: An efficient dense descriptor applied to wide-baseline stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence **32**(5) (2010) 815–830