3D Face Recognition For Cows

Deepak Yeleshetty[1], Luuk Spreeuwers[2], Yan Li[3]

Abstract: This paper presents a method to recognize cows using their 3D face point clouds. Face is chosen because of the rigid structure of the skull compared to other parts. The 3D face point clouds are acquired using a newly designed dual 3D camera setup. After registering the 3D faces to a specific pose, the cow's ID is determined by running Iterative Closest Point (ICP) method on the probe against all the point clouds in the gallery. The root mean square error (RMSE) between the ICP correspondences is used to identify the cows. The smaller the RMSE, the more likely that the cow is from the same class. In a closed set of 32 cows with 5 point clouds per cow in the gallery, the ICP recognition demonstrates an almost perfect identification rate of 99.53%.

Keywords: Cows, Biometrics, Visual identification, 3D face recognition, Pointcloud registration, Iterative Closest Point, Realsense cameras.

1 Introduction

Biometric identification is an efficient and a reliable method because it uses the unique natural discriminating features of each subject without the need of an external identification document or an attached device. This paper aims to design, implement, test and qualify a system that identifies cows using their 3D face point clouds. The project was carried out in the Product Development Group of the Dutch agri-tech company - Lely Industries N.V. Despite existing electrical cow identification methods, computer vision is opted due to its reliability, cost effectiveness and non-invasive property. Twisted Infrared (IR) tag around the neck sometimes results in failure of identification as the tag faces the cow's body. The IR sensor's batteries are non-replaceable, making the current system expensive. The tag around the neck also causes discomfort to the cow. Cameras are cost-effective and visual biometric identification methods like face recognition are non-invasive. The anatomy of the cow shows us that the skull is rigid and symmetric [JC07], which gives enough reason to pursue the cows' 3D faces for identification. This paper attempts to answer if we can uniquely identify cows based on their 3D face shapes. A new dual 3D camera setup is designed to capture the face of the cow. The 3D face is registered to a specific pose by finding the region of interest and correcting the rotation angles using the vertical symmetry of the cow's face [Sp11]. For recognition, the probe point cloud is compared with all the point clouds in the gallery using ICP [BM92] and inlier RMSE, a metric from the python library Open3D [ZPK18] is used to identify the cow. The identification rate for a herd of 32 cows is 99.53%, which proves that the 3D face shape can be used to identify cows.

[1] University of Twente, Data Management and Biometrics Group, Enschede, NL, yeleshetty.deepak@gmail.com
[2] University of Twente, Data Management and Biometrics Group, Enschede, NL, l.j.spreeuwers@utwente.nl
[3] Lely Technologies, Maassluis, NL, yli@lely.com

The paper is organized as follows: Related work on visual cow identification, human face recognition and point cloud registration methods is explained in section 2. The methodology of the proposed system is explained in section 3. The results are presented and analyzed in section 4. The paper is concluded with an insight on future scope in section 5.

2 Related Work

The authors of [Be19] demonstrated a Deep Learning method to identify cows based on multiple perspectives of their 2D face images. The accuracy for a closed set of 561 images from 52 cows was observed to be 89%. Their paper explains the shortcomings in terms of 2D landmark annotation, owing to the shape of the cow's face and states that multiple views yield better identification results.

ICP based recognition systems have been explored for 3D human faces [Ma05], however, as mentioned in [Sp11], ICP takes several seconds to register and recognize. The author in [Sp11] describes a fast and accurate 3D face registration and recognition method with a rank-1 identification rate of 99%. For 3D face registration, the region of interest (ROI) is estimated by fitting a cylinder. The vertical symmetry plane is obtained by finding the rotation around y and z axes. The angle between the nose bridge and the vertical axis is maintained at $\frac{\pi}{6}$ rad. Recognition is done by estimating the likelihood ratio of the probe's PCA-LDA features after comparing with those in the gallery. This method overcomes the time complexity of ICP and speeds up 3D face recognition for humans. As opposed to humans, cows lack the luxury of publicly available face database. Additionally, 2D face registration for cows is challenging as the face creates self occlusion for even a minute change of pose.

Two cameras are used in this project to overcome self occlusion. Point cloud registration is the process of estimating the rigid body transformation matrix that aligns the perspectives from both cameras, giving us a complete view of the subject. ICP [BM92] estimates the transformation between two point clouds (source to target) by minimizing the distances between correspondences, given an initial transformation. The transformation matrix is iteratively updated to minimize the point to point distances over the correspondence set. Let $C = \{(p,q)\}$ be the correspondence set with correspondence pairs $p \in P$ and $q \in Q$, where P and Q are the target and the source point clouds respectively. The two main ICP result metrics described in the library Open3D [ZPK18] are called Fitness (F) and Inlier Root Mean Square Error (I_{RMSE}).

$$F = \frac{N_c}{N_p} \qquad\qquad I_{RMSE} = \frac{1}{N_c} \sum_{(p,q)\in C} d_{p-q}$$

N_c is the number of correspondences, N_p is the number of points in the target point cloud and d_{p-q} is the mean squared distance between the correspondences. Fitness describes the overlapping area between the two point clouds. Inlier RMSE is the average of the mean square point to point distances of the correspondences(*Inliers*). A good registration results in a high fitness value (in the range [0,1]) and a low inlier RMSE value.

This paper aims to demonstrate the 3D face registration method explained in [Sp11], on cows. The recognition method will be based on point-to-point ICP [BM92].

3 Methodology

The proposed system's methodology can be divided into three steps: Data Acquisition, 3D Face Registration and ICP Based Recognition These steps can be seen in figure 1. The camera setup is designed to capture the 3D recordings of the cow's face. From the recordings, the required frames are captured and the point clouds are extracted. L-R registration method is performed on the extracted point cloud pairs to obtain 3D faces. The 3D faces are de-noised and transformed to a common pose as described in section 3.2 [Sp11]. Point to point ICP [BM92] is performed on every probe point cloud against all the point clouds in the gallery. The resulting inlier RMSE score is used to identify cows.

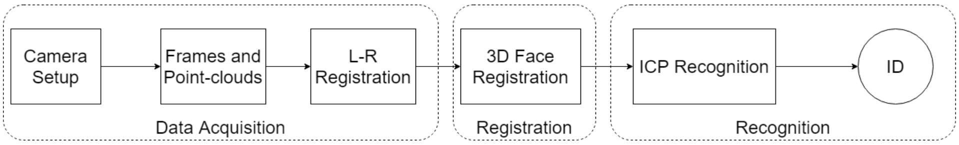


Fig. 1: Project Pipeline

3.1 Data Acquisition

A new dual 3D camera [4] setup is designed to acquire the 3D face of the cow. Both cameras are designed to face forward with no tilt because tilting the cameras would increase the field of view, which brings other cows in the frame and affects the further steps. Since the approximate ear to ear width of the cow was about 35 cm, the cameras were placed 70 cm apart. An illustration of the setup can be seen in figure 2, where C_L and C_R represent the left and right cameras respectively. F_W, F_D and S_B denote the approximate face width, approximate distance of the face from the setup and the fixed setup baseline respectively. Figure 3 shows the setup used in the farm.

Table 1 shows the specification of the camera setup and the camera itself. Due to the auto exposure setting in the Realsense camera, it was observed that the 3D points were very poorly estimated for cows with white fur or a surface that reflects light. So, only black or dark skinned cows are used in this project for identification. From a one-day data acquisition session, 1442 point clouds from 32 cows were collected and are used in this project. Each cow has 10 to 75 point clouds. Five point clouds per cow are stored in the gallery and the remaining are used as probe. The cows were treated gently without any discomfort throughout the data acquisition process.

To combine the point clouds from both cameras, L-R Registration method is followed (L-R indicates Left - Right cameras). L-R registration is divided into two steps: a feature-based

[4] Intel Realsense D435

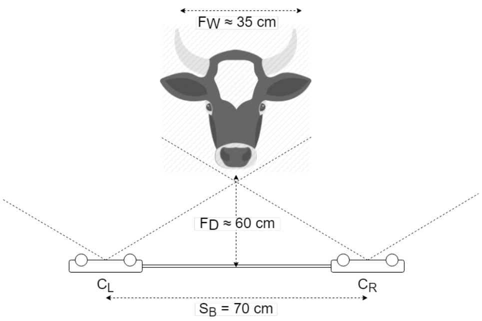


Fig. 2: Data Acquisition - Illustration

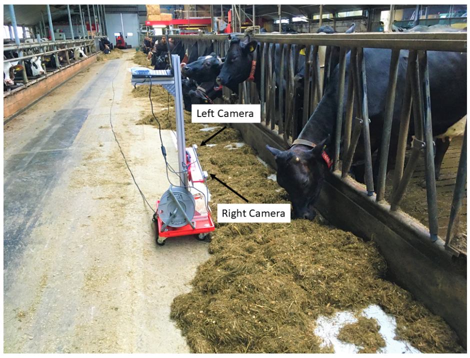


Fig. 3: Hardware set up

Description	Value
Approx. dist. cow to camera	60 cm
Setup baseline	70 cm
Cow face width	35 cm
Diagonal Field of view (per camera)	$95° \pm 3°$
Resolution	848×480 px
Frames per second	15 fps

Tab. 1: Camera Setup Specification

Description	Value
Voxel size for downsampling	0.02 m
Downsampling search radius	0.04 m
FPFH features search radius	0.1 m
RANSAC distance threshold	0.03 m

Tab. 2: Point Cloud Registration Parameters

Global Registration method that estimates a coarse transformation matrix and a *Local Registration* method that refines the transformation.

Fast Point Feature Histograms (FPFH) features are estimated on the down-sampled point cloud [RBB09]. A coarse transformation is obtained from the FPFH correspondences between the left (source) and the right (target) point clouds using RANdom SAmple Consensus (RANSAC) [FB81]. The RANSAC model is set to converge when the distance between majority of the correspondences reaches a global minimum. This coarse transformation matrix is fed to the ICP algorithm as an initial transformation estimate and yields a fine transformation matrix between the left and right cameras. This resulted in a visually convincing L-R Registration. Table 2 shows the different parameters used in L-R registration method. On an Intel i7 6-core 2.20 GHz CPU, it takes roughly 2 seconds to complete L-R Registration for one pair of point clouds.

3.2 3D Face Registration [Sp11]

3D face registration involves de-noising and transforming the L-R registered 3D face point cloud to a specific pose. The chosen pose is the front view of the cow, with the nose bridge area parallel to the image plane, which results in an ideal perspective that shows the vertical symmetry of the cow's face. A slightly modified version of the face registration method explained in [Sp11] is implemented in this section.

To estimate the ROI of the point cloud, the surface normals of the point cloud are calculated and a cylinder is fit using RANSAC. The open source C++ library PCL[RC11] is used to fit a cylinder and extract the ROI using the defined parameters (Table 3) on the point cloud. The input and output of the ROI estimation is illustrated in figure 4. As opposed to

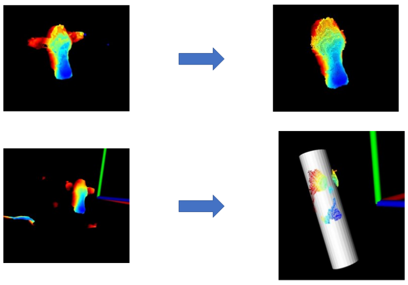


Fig. 4: Estimating ROI using PCL Cylinder Fitting

Description	Value
Radius Interval	$[0.15, 0.20]$ m
Inlier Distance Threshold (from Axis)	$radius \pm 0.05$ m
Max. RANSAC Iterations	1000

Tab. 3: Cylinder fitting parameters

humans, cows have a longer and relatively flatter nose bridge. So, a plane P_α with normal N_α is fitted on the ROI point cloud using RANSAC in PCL. This plane always fits on the nose bridge with a very minor tilt. The x-y plane is called P_{xy} with normal N_z. The angle γ between P_α and P_{xy} is calculated using their normals and the ROI is rotated around the x-axis by this angle. A plane $P_{n\alpha}$ is fit on the rotated ROI point cloud and it is translated along the positive z- axis to a distance $d_z = D - 0.1$ where D is the distance between the planes $P_{n\alpha}$ and P_{xy}. This will translate the point cloud approximately 10 cm from the x-y plane. To estimate the rotation angles along the y and z axes (θ and ϕ), we use the vertical

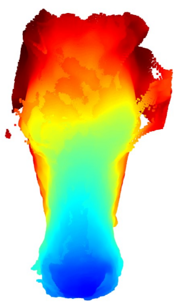


Fig. 5: Completely registered point cloud

Description	Value
θ interval	$[-\frac{\pi}{9}, \frac{\pi}{9}]$ rad.
θ step size	$\frac{\pi}{90}$ rad.
ϕ interval	$[-\frac{\pi}{4}, \frac{\pi}{4}]$ rad.
ϕ step size	$\frac{\pi}{50}$ rad.
Sliding interval	$[-\frac{3}{4}width, \frac{3}{4}width]$
Sliding step	0.05 m
$diff_z$ threshold	0.01

Tab. 4: Symmetric Orientation Parameters

symmetry of the cow's face and implement parts of the Symmetry Plane section in [Sp11]. A low resolution range image is constructed for multiple steps of rotation around the y axis (θ), by projecting the point cloud on the x-y plane with 5x5 mm grids. The value of each pixel in this range image is equal to the average depth (z coordinate) of points projected to the corresponding grid. The image is rotated in-plane in multiple steps which is the same as rotating along z-axis (ϕ). For every step in ϕ, The range image is mirrored and is slid horizontally in $[-\frac{3}{4}w, \frac{3}{4}w]$ with a step size of $d = 5$ mm, where w is the width of the range image. For every step d, the pixel-wise difference ($diff_z$) between the image and its mirror is computed. The pixel is said to contribute to the symmetry if the $diff_z$ value lies below the threshold (0.01). The θ and ϕ step corresponding to the maximum number of contributing pixels are the required angles to *straighten* the cow's face. The result is called a completely registered point cloud (figure 5). The parameters used to obtain the symmetric orientations is summarized in the table 4. It was observed that some gallery point clouds are incorrectly registered but the source of these irregularities is not investigated in this project.

3.3 ICP Based Recognition

ICP based recognition method is identical to the two-step L-R registration method. It is a computationally expensive and time consuming process as each of the probe point cloud is compared with all 160 gallery point clouds (32 cows with 5 point clouds each). Figure 6 shows an overview of the ICP based recognition method. ICP on each probe generates 160 Fitness and Inlier RMSE scores. The inlier RMSE scores are grouped for each cow in gallery and the average scores per cow is computed, which results in a reduced set R_s of 32 scores. The gallery ID corresponding to the minimum inlier RMSE score of R_s is the predicted ID. Out of 1282 probes from 32 cows, 1276 probes are correctly predicted, yielding an identification rate of 99.532%. Recognizing each cow takes about 300 seconds on an Intel i7 6-core 2.20 GHz CPU. With further improvements in the data acquisition process and implementation of a version of [Sp11], the recognition process could be much faster.

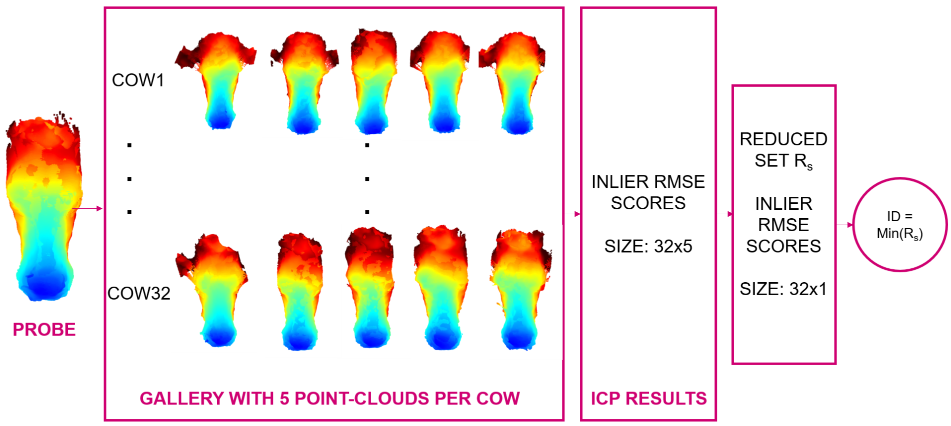
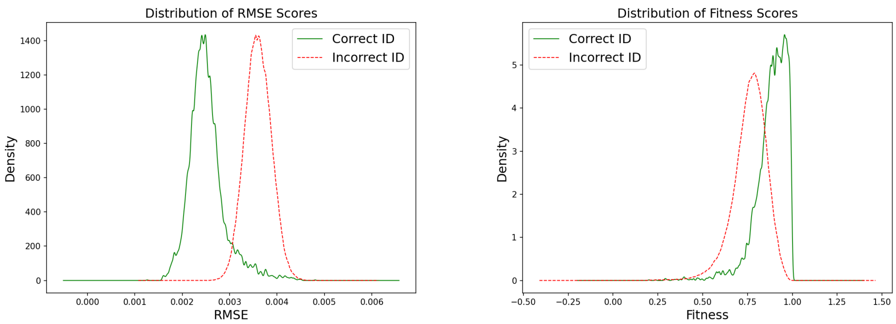


Fig. 6: ICP based recognition method

4 Results & Discussion

The metric inlier RMSE was chosen after analyzing both metrics for 1442 probes in a verification experiment. ICP was performed for all 1442 point clouds with all 160 gallery point clouds, except itself. For instance, if probes are in the gallery, ICP was performed only on 159 gallery point clouds, excluding itself. Figure 7 shows distribution plots of the



(a) Inlier RMSE (b) Fitness

Fig. 7: Distribution plots of the cows based on the chosen metrics

same vs different cows for the fitness and inlier RMSE scores. We see that fitness is not a reliable metric as the distribution shows a considerable overlap between scores 0.75 and 1.00. However, the inlier RMSE separates the same and different cows at a score threshold of approximately 0.003. Receiver Operating Characteristic (ROC) curves are plotted for both the metrics and the result is shown in figure 8. Inlier RMSE is observed to have an Equal Error Rate (EER) of about 6.5% at a threshold of 0.0032, while Fitness has an EER of 22% at a threshold of 0.836. The results show that inlier RMSE is a better metric to classify cows in this dataset.

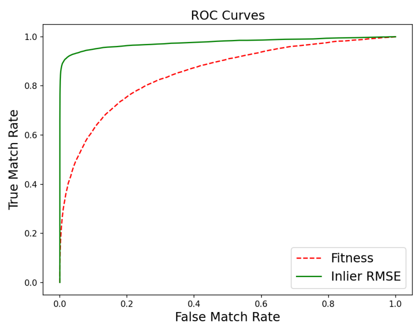


Fig. 8: ROC Curves - Fitness and Inlier RMSE

# Gallery point clouds per cow	Identification Rate (%) Min.	Max.
1	88.611	99.220
2	94.462	99.532
3	97.738	99.220
4	98.830	99.142
5	99.142	

Tab. 5: Identification Rates

In a closed set identification experiment, we see the identification rates for different number of gallery point clouds per cow. This will give us an idea of how the system could perform with respect to the amount of data in the gallery. We perform the identification experiment on 1282 probe point clouds, excluding the 160 gallery point clouds.

Table 5 shows the identification rates for different number of point clouds in gallery. Minimum and maximum identification rates refer to the extremes, where the decision is made based on the worst and the best case inlier RMSE scores respectively. In the case of 1 gallery point cloud per cow, we select the lowest (best case) and the highest (worst case) inlier RMSE scores per cow. The former yields an identification rate of 99.220% and the latter yields 88.611%. Similarly, in the cases of 2, 3 and 4 gallery point clouds per cow, we choose the average of the lowest(best cases) or highest (worst cases) 2,3 and 4 inlier RMSE scores. The average of all 5 inlier RMSE values showed an identification rate of 99.142%. The trend shows us that as we keep adding more gallery point clouds per cow we get lower RMSE scores, whose contribution is clearly reflected in the Minimum Identification Rate field.

5 Conclusion & Future Scope

The objective of this project was to investigate and prove the concept of identifying cows using their face shapes in order to improve cost efficiency and the cow's comfort. The methodology involves slightly modified existing 3D face registration and recognition methods. After acquiring face point clouds from the proposed dual 3D camera setup and registering them, ICP based recognition yields near perfect identification rate of 99.532%. The results prove that we can distinguish cows based on their face shapes and opens up further possibilities in implementing a more robust registration method, speeding up the recognition process and investigating the performance on a larger scale. While the identification rate is expected to be in a similar range, computation time will increase linearly with herd size because ICP should be performed for more cows. Most medium-sized Dutch farms have over 40 cows and a real time implementation of this system requires it to be *at least* 15 times faster (20 s per cow).

To improve the speed, implementing a faster and more accurate 3D face recognition method as explained in [Sp15] for cows on a bigger dataset would be an interesting experiment. Collecting data over a longer period of time from different types of cows will show if facial variations (natural or due to sickness) will affect the system's performance. A bigger dataset will enable further research on 3D cow face recognition using conventional and Deep Learning methods. If vision based systems out-perform the traditional electrical ones, cows will be free from IR tags around the neck.

Acknowledgement

This research was fully supported by Lely Industries N.V., located at Maassluis, The Netherlands. We would like to express our gratitude to Mr. Patrick Segeren, the Head of Smart Components Team at Lely for initiating the project and for arranging all necessary resources for our research. We extend our thanks to the wonderful cows from Lely's test farms for their patient cooperation.

References

[Be19] Bergamini, L.; Porrello, A.; Dondona, A. C.; Del Negro, E.; Mattioli, M.; D'alterio, N.; Calderara, S.: Multi-views Embedding for Cattle Re-identification. CoRR, abs / 1902.04886, 2019.

[BM92] Besl, P. J.; McKay, N. D.: A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2):239–256, 1992.

[FB81] Fischler, M. A.; Bolles, R.C.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Commun. ACM, 24(6):381–395, June 1981.

[JC07] Jackson, P.G.; Cockcroft, P.D.: Clinical Examination of the Head and Neck. In: Clinical Examination of Farm Animals. John Wiley & Sons, Ltd, pp. 29–50, 2007.

[Ma05] Maurer, T.; Guigonis, D.; Maslov, I.; Pesenti, B.; Tsaregorodtsev, A.; West, D.; Medioni, G.: Performance of Geometrix ActiveIDˆ TM 3D Face Recognition Engine on the FRGC Data. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops. IEEE, pp. 154–154, 2005.

[RBB09] Rusu, R. B.; Blodow, N.; Beetz, M.: Fast Point Feature Histograms (FPFH) for 3D registration. In: 2009 IEEE International Conference on Robotics and Automation. pp. 3212–3217, 2009.

[RC11] Rusu, R. B.; Cousins, S.: 3D is here: Point Cloud Library (PCL). In: 2011 IEEE International Conference on Robotics and Automation. pp. 1–4, 2011.

[Sp11] Spreeuwers, L.: Fast and Accurate 3D Face Recognition. International Journal of Computer Vision, 93(3):389–414, Jul 2011.

[Sp15] Spreeuwers, L.: Breaking the 99% barrier: optimisation of three-dimensional face recognition. IET Biometrics, 4(3):169–178, 2015.

[ZPK18] Zhou, Q.Y.; Park, J.; Koltun, V.: Open3D: A Modern Library for 3D Data Processing. arXiv:1801.09847, 2018.