Morphological Gender Recognition by a Social Robot and Privacy Concerns

[Late Breaking Reports]

Arnaud Ramey
RoboticsLab, Univ. Carlos III of Madrid
Leganés, Spain
arnaud.ramey@m4x.org

Miguel A. Salichs
RoboticsLab, Univ. Carlos III of Madrid
Leganés, Spain
salichs@ing.uc3m.es

ABSTRACT

An intuitive and robust user recognition system is at the key of a natural interaction between a social robot and its users. The gender of a new user can be guessed without explicitly asking it of her, which can then be used to personalize the interaction flow. In this LBR, a novel algorithm is used to estimate the gender of a person based on its morphological shape. More specifically, the vertical outline of the breast of the user is used to estimate his or her gender, based on similar shapes seen during training.

On early benchmarks with databases that represent well the diversity of human body shapes, the accuracy rate is close to 90% and outperforms a state-of-the-art algorithm. Our algorithm provides a fast and seamless estimation flow and needs limited computation resources, which tailor it for HRI. Its usefulness has been proved by integrating it in a social robot. However, its use raises concerns among the users about their privacy, which will lead to further study.

Categories and Subject Descriptors

I.2.10 [**Vision and Scene Understanding**]: Modeling and recovery of physical attributes; H.3.4 [**Systems and Software**]: User profiles and alert services; I.2.9 [**Robotics**]: Sensors

General Terms

Algorithms, Experimentation, Human Factors

1. INTRODUCTION

To obtain a natural interaction with the human user, it is important that the robot gathers information about her. The queries can be explicit, such as questions asked to her or an interactive display, but the range of applications widens if some features can be inferred with no user action required. These features are called *soft biometrics*. One of the earliest multi-modal user tracking systems [3] used the user height and the color of the skin to obtain a robust tracking of the users around a multimedia kiosk. Clothes color histograms are also used in [6] for logging in a computer and continuously checking the user identity.

Considering the shape of the breast of a person is a natural way of estimating his or her gender [5]. However, this method turned out to perform poorly in real time and with challenging body shapes. Furthermore, as further underlined in [1], the users are concerned about the privacy of this detection and might consider it intrusive.

The work presented in this paper aims at determining the gender of the user in a similar way, but overcoming the limitations of the previous work.

2. DESCRIPTION OF THE SYSTEM

The input needed for our system is a 3D (colorless) point cloud of the user. This can be obtained through a range of input devices, such as a structured light depth sensor such as Microsoft Kinect, or a stereo vision device. The output is a real-time estimation of the gender of the visible users.

User detection and mask computation. The point cloud given by the device needs to be cut down to keep only the points of the user. Our system sports two sources for this user mask. A device-specific solution is to use the one directly given by the device SDK (Kinect NITE SDK) The other way is to obtain thanks to face detection ([8]) a seed pixel belonging to the user; and to detect the edges in the depth mask thanks to a Canny filter; a propagative floodfill in the edge image from that seed generates the final user mask. In both cases, the point cloud is obtained by reprojecting to world coordinates the pixels of the depth map that also belong to the user mask.

Breast feature computation. Laws' algorithm ([5]), which is the most recent and similar algorithm, is based on extracting several horizontal slices of the point cloud of the user for a span of height values that corresponds to the breast location in a body. The contour of each slice is matched against the pattern of a breast shape using non-linear regression. Metrics on the pattern that fits best one of the slices are then used to evaluate the gender of the user. However, female users with a moderate breast and over-weight males generate similar horizontal slices.

Our algorithm is instead based on computing the vertical contour of the torso of the user. The pipeline is illustrated in Figure 1 The orientation of the user is determined by fitting an ellipse to the shadow generated by the projection of the point cloud onto the ground plane. The projection plane is chosen to be vertical, going through the center of

HRI'14, March 3–6, 2014, Bielefeld, Germany.
ACM 978-1-4503-2658-2/14/03.
http://dx.doi.org/10.1145/2559636.2563714 .

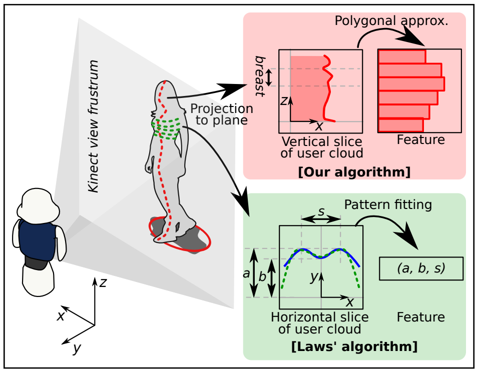

Figure 1: *Algorithms pipelines for both [5] and our algorithm. The black shadow is the projection mentioned in section 2, and the ellipse the best fit to that shadow. The projection planes for both methods are indicated in dashed lines.*

mass of the user and parallel to the short axis of the ellipse. The points of the cloud with a height corresponding to the breast, according to morphological rules, and that are close to that plane are projected onto it. The outline of the user thus generated is approximated by a fixed length polygonal line, which consists in the final feature.

Gender determination. A training set with pictures of users and manually labelled genders is used to generate training features and labels. A two-class SVM classifier ([2]) is trained on this data to learn the non-linear separation between male and female shapes. Once trained, the prediction output of the SVM for a test shape determines the gender of the user.

3. RESULTS AND CONCLUSIONS

The robot Maggie is equipped with a structured light sensing device (Kinect), that has already been used for gesture recognition [7]. Our algorithm is implemented in C++ and is based on the ROS architecture [1]. It supplies in real-time an estimation of the gender of the users around Maggie.

Benchmarking of the algorithm. To benchmark our algorithm against Laws' ([5]), we used the database [4]. It contains video sequences of 55 users, both female and male, walking on a stage with varying light conditions. These users represent well the possible range of body shapes: including slender females and overweight males. Both our and Laws' detectors were trained on 1100 images (20 per video), and trained on 550 (10 per video).

The results are gathered in Table 1. Note that an algorithm of gender detection based on the body shape cannot intrinsically reach a 100% accuracy, as there is no strict categorization between female and male shapes. Overweight male subjects for instance appear more curvy than some women. This limitation can be overcome by coupling this gender recognition algorithm with others, based for instance on the face of the users or their height. Yet, our algorithm correctly estimates the gender of the user in almost nine times out of ten. Laws originally tested its algorithm on a medical scan data database with full-figured, standing straight subjects, and logically performs poorly on real-life user samples.

Privacy concerns. The gender recognition algorithm runs in the background within the robot and supplies in real time

[1]http://wiki.ros.org

	Laws	Our
Accuracy rate	63%	89%
Average computation time (ms)	83	11

Table 1: *Benchmark results*

estimations of the gender of the users around it, even though the user was never seen before.

The first experiments with real users have given interesting results. On the one hand, users acknowledge the usefulness of detecting their gender for a personalized interaction. On the other hand, once told the robot considers seamlessly the shape of their torso to guess their gender, some users have shown concerns about their privacy and how measuring their chest shape is an intrusive method, thus meeting the results of [1]. However, once shown the output of the algorithm, some of these users have reported to have worse expectations than reality: the sensor and the algorithm are not as intrusive as infra-red cameras. These first conclusions now give way to deeper research using a proper experimental framework and a bigger set of users.

Conclusions. the proposed algorithm turns out to be a useful metric as it helps determining the gender of new users. It gives a generally liable estimation of it, and its lightweight nature makes it perfectly suitable for social robots.

Future works will study more in details the a-priori and a-posteriori privacy concerns of the users. We will also determine how using other techniques such as face-based gender recognition improves the accuracy of the gender estimation.

Acknowledgments

The authors gratefully acknowledge the funds provided by the Spanish MICINN (Ministry of Science and Innovation) through the project "Aplicaciones de los robots sociales", DPI2011-26980 from the Spanish Ministry of Economy and Competitiveness. Moreover, the research leading to these results has received funding from the "RoboCity2030-II-CM" project (S2009/DPI-1559), funded by "Programas de Actividades I+D en la Comunidad de Madrid" and cofunded by "Structural Funds of the EU".

4. REFERENCES

[1] Y. Cai, J. Laws, and N. Bauernfeind. Design Privacy with Analogia Graph. *IAAI*, 2010.

[2] C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 1995.

[3] T. Darrell, G. Gordon, M. Harville, and J. Woodfill. Integrated person tracking using stereo, color, and pattern detection. *International Journal of Computer Vision*, 2000.

[4] L. Igual, A. Lapedriza, and R. Borràs. Robust Gait-Based Gender Classification using Depth Cameras. In *Eurasip Journal On Image And Video Processing*, 2013.

[5] J. Laws and Y. Cai. *A privacy algorithm for 3d human body scans.* 2006.

[6] K. Niinuma, U. Park, and A. K. Jain. Soft Biometric Traits for Continuous User Authentication. *IEEE Transactions on Information Forensics and Security*, 5(4):771–780, Dec. 2010.

[7] A. Ramey, V. Gonzalez-Pacheco, M. A. Salichs, and V. González-Pacheco. Integration of a Low-Cost RGB-D Sensor in a Social Robot for Gesture Recognition. In *Proceedings of the 6th International Conference on HumanRobot Interaction*, HRI '11, pages 229–230, New York, NY, USA, 2011. ACM.

[8] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, volume 1, pages 511–518, Los Alamitos, CA, USA, Apr. 2001. IEEE Comput. Soc.