MaskBot: Real-time Robotic Projection Mapping with Head Motion Tracking

Miguel Altamirano Cabrera
Skolkovo Institute of Science and Technology (Skoltech)
Moscow, Russia
miguel.altamirano@skoltech.ru

Igor Usachev
Skolkovo Institute of Science and Technology (Skoltech)
Moscow, Russia
igor.usachev@skoltech.ru

Juan Heredia
Skolkovo Institute of Science and Technology (Skoltech)
Moscow, Russia
juan.heredia@skoltech.ru

Jonathan Tirado
Skolkovo Institute of Science and Technology (Skoltech)
Moscow, Russia
jonathan.tirado@skoltech.ru

Aleksey Fedoseev
Skolkovo Institute of Science and Technology (Skoltech)
Moscow, Russia
aleksey.fedoseev@skoltech.ru

Dzmitry Tsetserukou
Skolkovo Institute of Science and Technology (Skoltech)
Moscow, Russia
d.tsetserukou@skoltech.ru

Figure 1: MaskBot system: a) The module with the projector and webcam located at the 6DoF UR3 robot’s end-effector. b) Face detection and projection of 68 points. c) The projection mapping on the user’s face.

ABSTRACT

The projection mapping systems on the human face are limited by the process latency and the users’ movement. The area of the projection is restricted by the position of the projectors and cameras. We are introducing MaskBot, a real-time projection mapping system guided by a 6 Degrees of Freedom (DoF) collaborative robot. The collaborative robot locates the projector and camera in front of the user’s face to increase the projection area and reduce the system’s latency. A webcam is used to detect the face orientation and to measure the robot-user distance. Based on this information we modify the projection size and orientation. MaskBot projects different images on the user’s face, such as face modifications, make-up, and logos. In contrast to the existing methods, the presented system is the first that introduces a robotic projection mapping. One of the prospective applications is to acquire a dataset of adversarial images to challenge face detection DNN systems, such as Face ID.

CCS CONCEPTS

- Hardware → Displays and imagers; Emerging interfaces;
- Computer systems organization → Robotics.

KEYWORDS

Robotics; Projection Mapping; Motion Capture; Tracking System

1 INTRODUCTION

Projection mapping has been used in many fields to solve some important object visualization problems in Augmented Reality. Inami et al. [Inami et al. 2000] introduced projection on objects covered by retroreflective material to allow the users to handle objects of arbitrary shape. The users look at bright images projected on the surface of the object covered with retroreflective material. This approach recreates the shape of the observed object and creates the illusion to see a different one. This method can be applied to
generate many illusions and to change the visual shape of objects. However, to create different projections on the human face, the use of retroreflective material on the user’s face will cover the area of interest to project.

The technology introduced by Matrosov et al. generates a projection on the floor as a screen to interact with a drone by foot gestures [Matrosov et al. 2016]. The novelty of this work is based on its intuitive interface composed by a RGB-D camera and a projector mounted over a flying robot. The proposed technology recognizes the user’s foot gestures and generate a corresponding projection. The LightAir is not suitable for the dataset collection as it is a challenge to achieve smooth drone hovering. CollMot Entertainment [Ackerman 2020] developed a system that uses drones to generate an enormous screen in the sky, and then laser projectors draw on that screen. The low-latency requirements for face projection mapping are much higher than those required for non-deformable objects. Bermano et al. [Bermano et al. 2017] proposed a system to considerably reduce the latency in the face projection mapping. In this project, the facial expressions are accurately detected and can be predicted with Kalman filters. However, the equipment used by them is too sophisticated, and the face’s position is not possible to be changed. Bokaris et al. introduced a system that applies a more straightforward approach. This system makes a real-time facial projection mapping to create the effect of make-up on the face of the users [Bokaris et al. 2019]. However, projection mapping is limited to the area of the projector.

We are introducing MaskBot, a novel system that increases the projection area avoiding any distortion caused by the user head mis-alignment. The technology behind this project is a Computer Vision (CV) algorithm and a 6 DoF collaborative robot, which controls the orientation of the projector and the camera regardless of the user’s head posture. The position of the human head is tracked with a webcam. When human moves the head, the robotic projector always pursues the human face and emits light towards the user’s skin. The collaborative robot safety protocols (collision detection, force, and speed limitation) guarantee secure human-robot interaction.

2 MASKBOT CONFIGURATION

MaskBot provides a real-time projection mapping on moving face and on other objects. The main components of our system are a 6 DoF collaborative robot UR3 from Universal Robots, a Logitech Webcam C930e, and an ultra-compact mobile projector Dell M115HD. The webcam and projector are attached to the robot’s end-effector, as indicated in Fig. 1. The image captured by the camera is processed with a Python script. The face detection algorithm provides face center coordinates. These coordinates are recalculated into the distance between the face and the camera. Next, the robot moves towards the human face. In parallel, the projection algorithm lays an image on the particular area of the face. Thus, MaskBot provides an extra workspace due to the robot movements. Besides, the system is able to change the projection plane and to produce projection on other objects.

3 FACE PROJECTION MAPPING

We are using face landmark detection by the model in dlib trained by Sagonas et al. [Sagonas et al. 2016], where 68 landmarks are detected on the user’s face. These points are used to locate different objects on the face and to fit on the user’s face.

Figure 2: Projection of the beard, hand, glasses on the face.

The picture projected on the user’s face is created using the OpenCV library according to the calculated angle for the orientation. The distance from the user’s face to the MaskBot system is calculated by the relation between the face width detected by CV and the real face width.

During the demonstration, the users stand in front of the system. In the first instance, a camera-projector calibration is going to be carried. A series of masks will be projected on the face.

In future work, the user will be capable to change the mask by hand gesture. A second collaborative robot will move a tablet, showing the face of the user with the projected image (robotic mirror). The visitors will be able to draw their own patterns on the user’s face. A camera will capture the image of the human face modified by the projected image, and a dataset of adversarial images will be collected. We can also envision a novel communication system where humans communicate with no need to align the face towards cameras, as it is the case with Skype and Zoom. The technology can also be applied in Industry 4.0 when robots communicate the messages through projectors, e.g., they can illuminate QR-code, instructions, and content of the parcel to the partner robot and workers.

REFERENCES

Evan Ackerman. 2020. Drone Swarm Generates Smoke Screen for Huge Laser Displays - IEEE Spectrum. https://spectrum.ieee.org/automaton/robotics/drones/drone-swarm-generates-smoke-screen-for-huge-laser-displays

Amit H. Bermano, Markus Billeter, Daisuke Iwai, and Anselm Grundhofer. 2017. Makeup Lamps: Live Augmentation of Human Faces via Projection. Computer Graphics Forum 36, 2 (2017), 311–323. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13128

Panagiotis-Alexandros Bokaris, Benjamin Askenazi, and Michael Haddad. 2019. Light Me Up: An Augmented-Reality Projection System. In SIGGRAPH Asia 2019 XR (SA ’19). Association for Computing Machinery, New York, NY, USA, 21–22.

M. Inami, N. Kawakami, D. Sekiguchi, Y. Tanyagida, T. Maeda, and S. Tachi. 2000. Visuohaptic display using head-mounted projector. In Proceedings IEEE Virtual Reality 2000 (Cat. No.00CB37048). 233–240.

Mikhail Matrosov, Olga Volkova, and Dzmitry Tsetserukou. 2016. LightAir: A Novel System for Tangible Communication with Quadcopters Using Foot Gestures and Projected Image. In ACM SIGGRAPH 2016 Emerging Technologies (SIGGRAPH ’16). Association for Computing Machinery, New York, NY, USA, Article Article 16, 2 pages.

Christos Sagonas, Epameinondas Antonakos, and Georgios Tzimiropoulos. 2016. 300 Faces In-The-Wild Challenge: database and results. Image and Vision Computing 47 (Mar 2016), 3–18.