Visage—a Miniaturized Face Detection Sensor

P. Nussbaum, M. Fretz, R. Krähenbühl, E. Türetken, E. Azarkhish, P. Jokic

Bringing face detection into a single low-cost and low-power component opens up new perspectives in products requiring safety or people awareness, as it can control simple functions such as access control, standby operation etc. Here a system-on-chip supporting an accelerated machine-learning algorithm with state-of-the-art optics and packaging have been devised to meet this challenge.

This component should be low-cost and contain a highly integrated system with an optic, imager and the System-on-Chip (SoC) that should not take up more than ½ cm^3. In such a system the packaging will encapsulate the compact lens, that must nevertheless provide a crisp image with a small track length, the SoC with the dedicated face detection algorithms.

Packaging

A global concept for the encapsulation of all the components has been sketched in Figure 2:

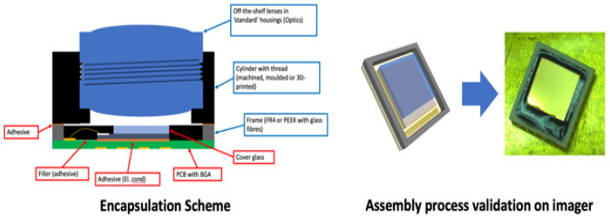

Figure 2: Encapsulation approach (left) and first tests on imager only (right).

The assembly process has been validated on a simpler encapsulation (imager only) the next step is to include the optics. The final system encapsulation includes the Visage SoC.

Optics

A commercial off-the-shelf (COTS) lens meeting the requirements of Visage has been procured, to provide a basis for comparison and to speed-up early prototyping. This version is more compact with a track length of 6 mm, see Figure 1, but is likely to give more blur.

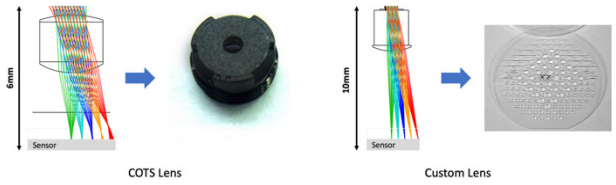

Figure 1: COTS lens (left) and CSEM lens (right).

A custom version has been designed and assembled for release in time with initial project planning. This version has a longer track length, *i.e.*, 10 mm, but is designed to deliver crisper images than the COTS lens (resulting spot size on imager of 8 um vs. 17 um) over a twofold increased focal depth range. Integration into the packaging and characterization will be completed in 2019.

Algorithms selection and training

Two algorithms capable of achieving good accuracy for face detection while fitting with the exceedingly small memory footprint (1 Mbytes) have been identified. The first is a modified AdaBoost Binary Decision Tree (BDT) algorithm that has been thoroughly optimized for Visage and the second is based on a Convolutional Neural Network (CNN) in an attempt to ultimately quantify to a binary neural network.

A data set containing more than 1 million images with uniformly distributed face sizes, locations, rotations and random distortions has been compiled to train and test the algorithms. Tests on unseen images showed 96% accuracy at less than 1% false positive rate. The operation of the detection algorithm is illustrated in Figure 3.

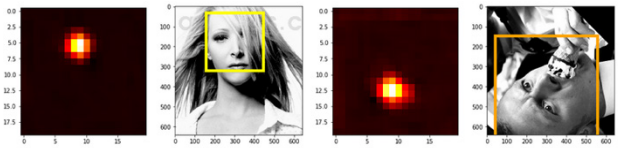

Figure 3: Probability distribution map of a face in the scene (orange blob) and the resulting face location in the source images.

System-On-Chip (SoC)

The integrated circuit supporting the machine learning accelerator has been structured around a Risc-V CPU which manages the whole system and drives the external communications (imager and detection output), see Figure 4.

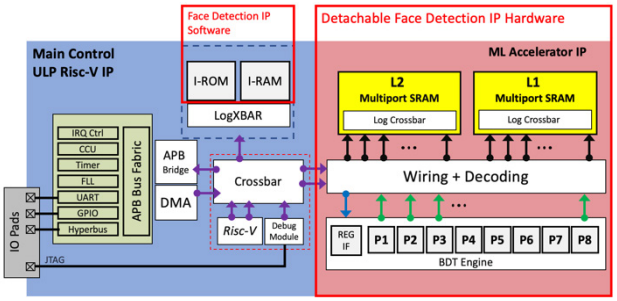

Figure 4: SoC Architecture.

The machine learning accelerator shown by the right part in the bloc diagram occupies 90% of the silicon area due to the size of the required memory. A first floorplan of the SoC can be found on the right side, showing an overall surface of 3 mm^2.

2019 Demonstrator

A first version of the demonstrator is presently being assembled. It is based on the Vision-in-Package (VIP) platform running the latest version of the BDT algorithm developed for the Visage project.

The setup implements a fully functional LED torch lamp controlled by the VIP and powered by batteries. The objective is to test the concept at the system-level and appraise the performances obtained when achieving the expected safety feature (dimming when beaming to a face).

The final component, which will be comprised of the packaged custom optics, SoC will be released in Q2 of 2020.