ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

IVRG

Jin B., Yildirim G., Lau C., Shaji A., Ortiz-Segovia M., Süsstrunk S.

MODELING FACE IMPORTANCE FOR SALIENCY DETECTION

Abstract

We present an algorithm to model the importance of human faces in visual saliency tasks. The perceived importance of faces is determined as a function of the face size and the number of faces through crowd sourcing experiments. We create a new visual saliency database that includes face and non-face images, and collect ground truth also through crowd sourcing. Evaluation on this database shows that our face and object saliency method outperforms eight state-of-the-art saliency algorithms.

Why face saliency?

In consumer type images, portraits and people images are predominant. Current saliency algorithms [4-8] assume faces to be equally important. However, the importance of faces varies in different images:

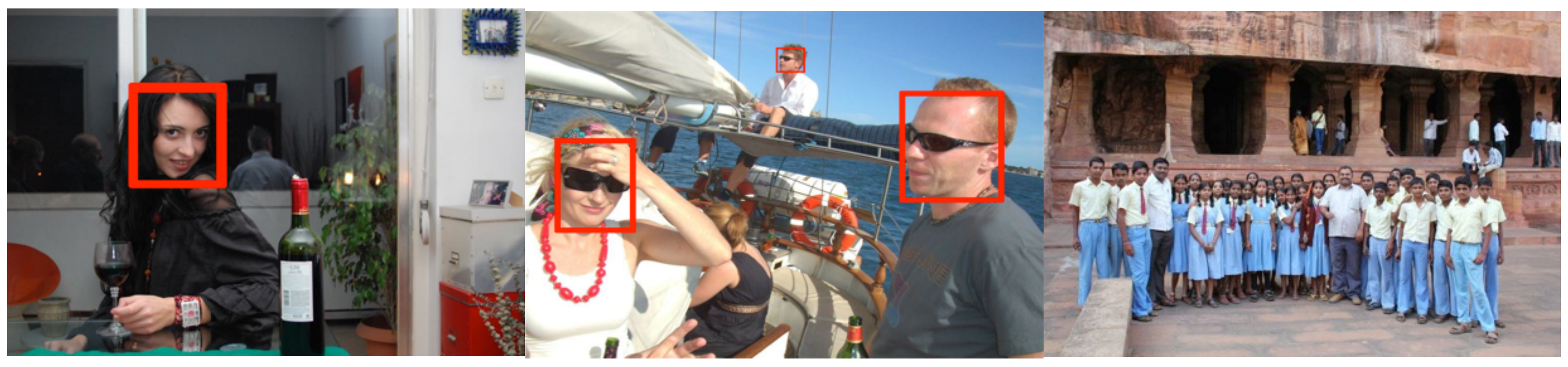

Fig. 1. In the left image, the face is large enough to be recognized by the observer and thus is salient. In the middle image there are multiple faces which are not equally important. In the right image there are numerous faces, resulting in individual faces not being considered important.

Face importance model

We conduct crowd-sourcing experiments on a dataset of 152 face images. People are asked to label the important regions of the image. We obtain our face importance model by analyzing the distribution over the face size and the number of faces:

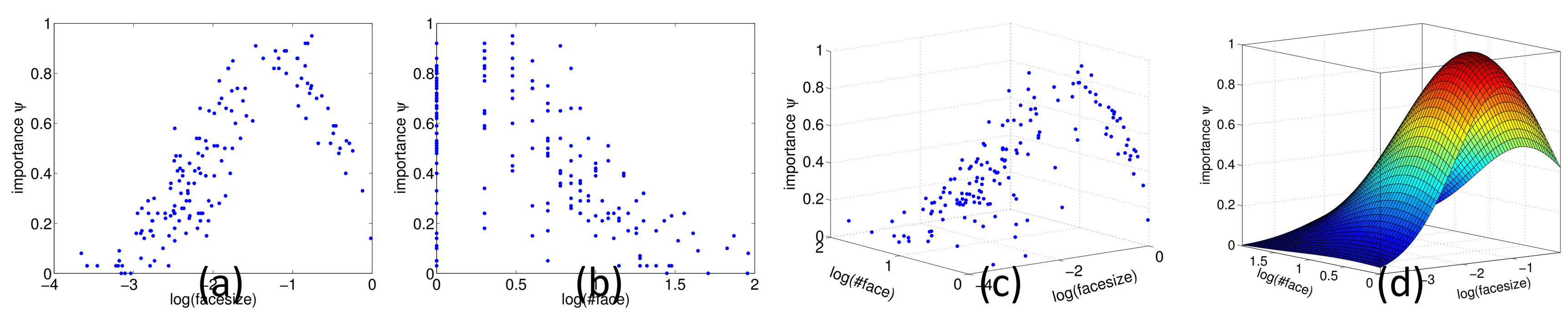

Fig. 2. (a) The distribution of face importance over face size. (b) The distribution of face importance over the number of faces. (c) The joint distribution of face importance over face size and the number of faces. (d) Our 2D Gaussian face-importance model.

Our saliency method

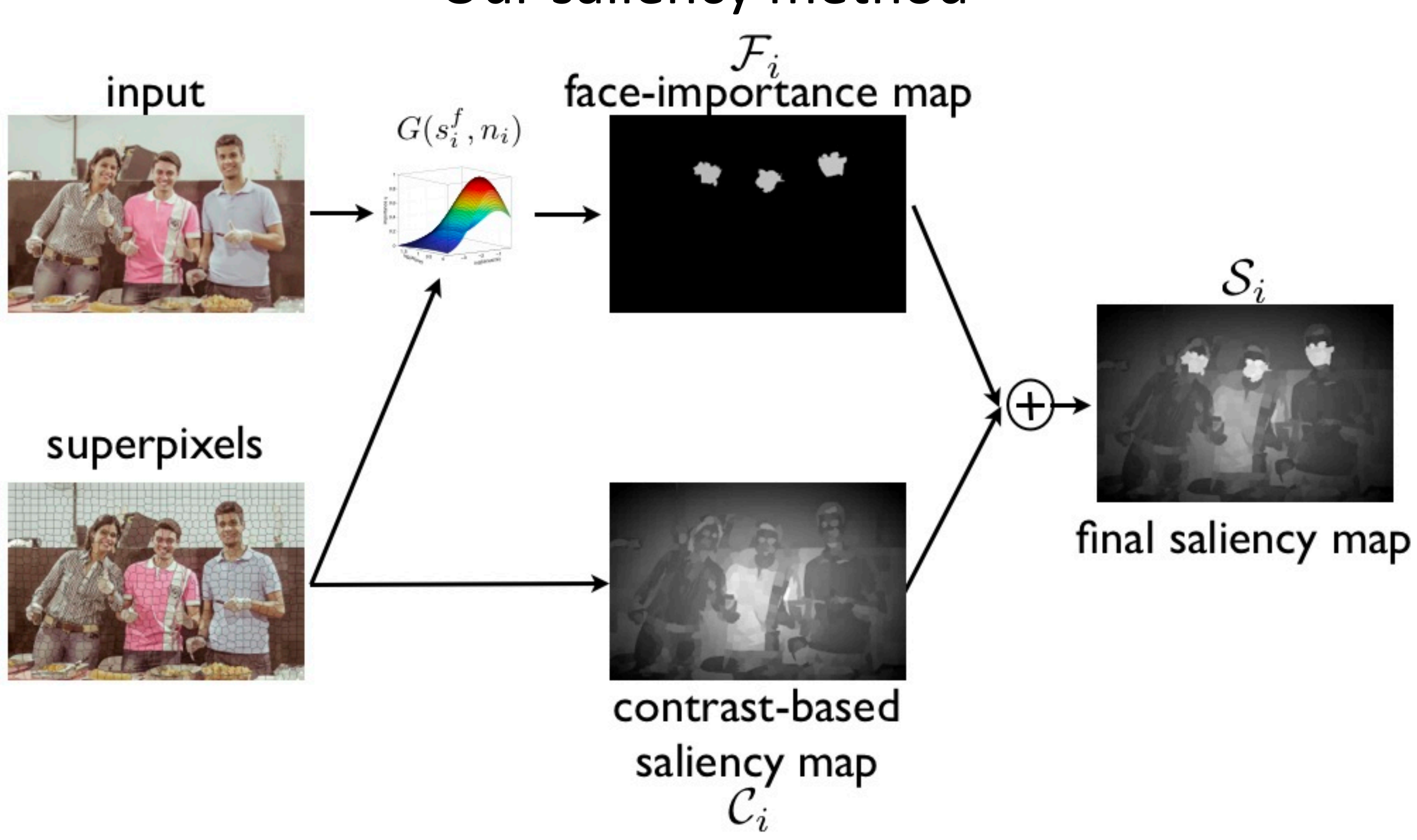

Fig. 3. The flowchart of our algorithm. We first over-segment an input image into superpixels. A contrast-based saliency map is computed considering both local and global contrast. The face-importance map is calculated using face detection results through our face-importance model. The final saliency map is the addition of the two maps.

Evaluation dataset

800 images with multi-level ground-truth, 632 of which include human faces.

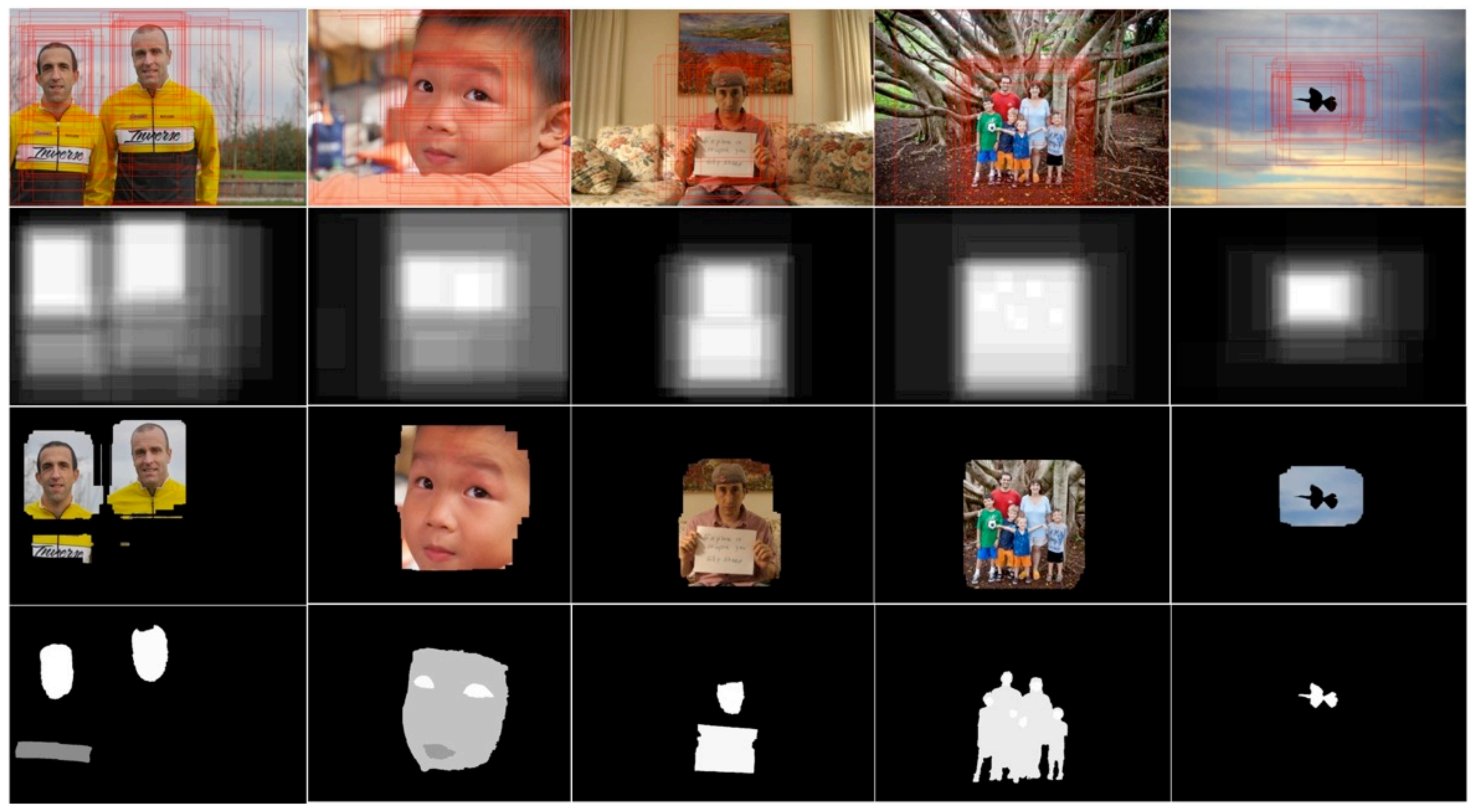

Fig. 4. The first row: the combination of all the labels from crowd-sourcing; the second row: the corresponding weight map; the third row: the foreground area using Ostu's threshold; the last row: our multi-level ground truth.

Evaluation

Comparison with 8 state-of-the-art saliency algorithms: AMC[1], CH[2], GBMR[3], SMVJ[4], LR[5], Judd[6], Borji[7], SC[8].

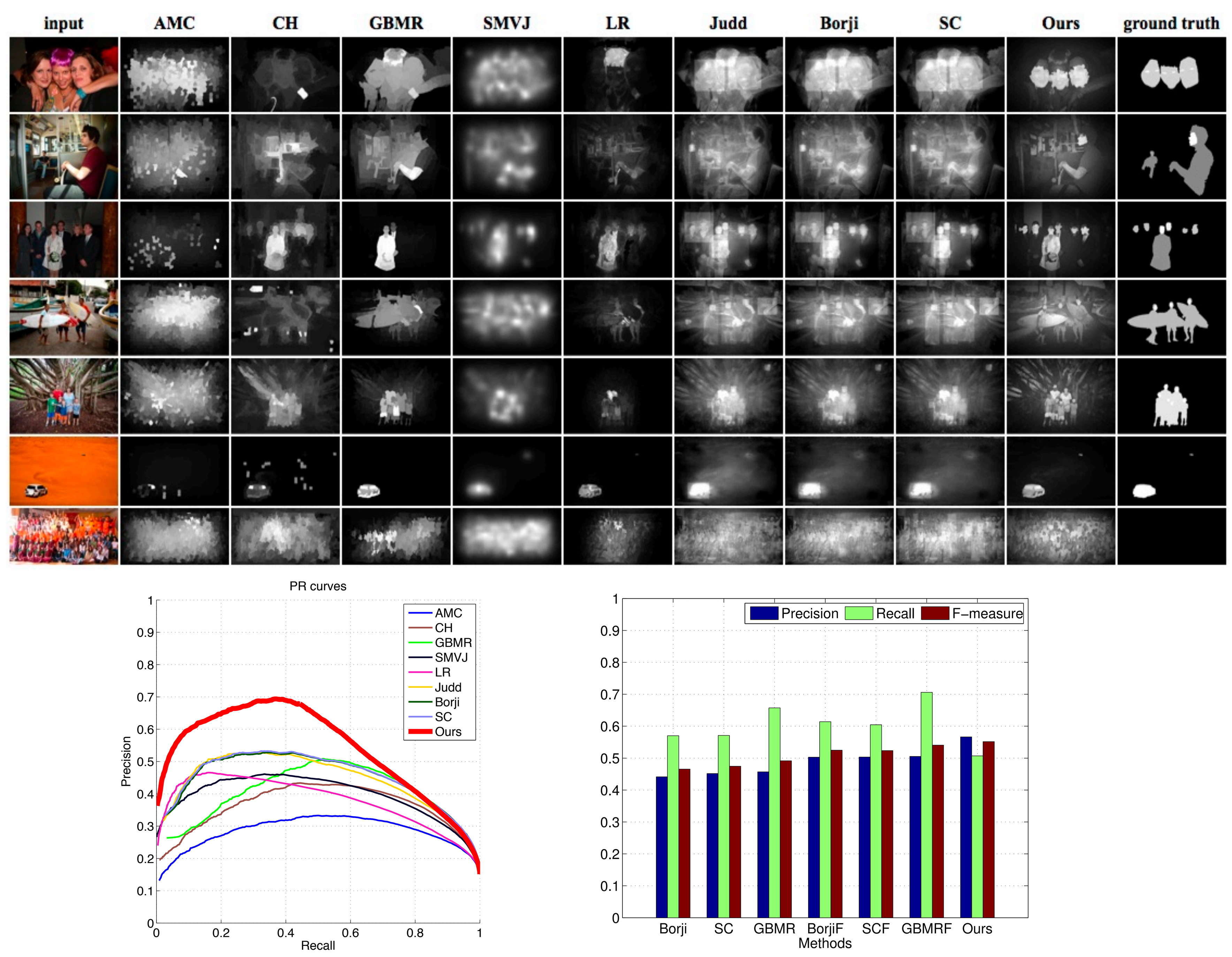

Fig. 5. Qualitative and quantitative comparison with eight state-of-the-art algorithms.

Reference

[1] Jiang, B., Zhang, L., Lu, H., Yang, M.H., Yang, C.: Saliency detection via absorbing markov chain. ICCV 2013.

[2] Li, X., Li, Y., Shen, C., Dick, A., van den Hengel, A.: Contextual hypergraph modelling for salient object detection. ICCV 2013.

[3] Yang, C., Zhang, L., Lu, H., Ruan, X., Yang, M.H.: Saliency detection via graph-based manifold ranking. CVPR 2013.

[4] Cerf, M., Harel, J., Einhäuser, W., Koch, C.: Predicting human gaze using low-level saliency combined with face detection. NIPS 2007.

[5] Shen, X., Wu, Y.: A unified approach to salient object detection via low rank matrix recovery. CVPR 2012.

[6] Judd, T., Ehinger, K., Durand, F., Torralba, A.: Learning to predict where humans look. ICCV 2009.

[7] Borji, A.: Boosting bottom-up and top-down visual features for saliency estimation. CVPR 2012.

[8] Karthikeyan, S., Jagadeesh, V., Manjunath, B.S.: Learning top down scene context for visual attention modelling in natural images. ICIP 2013.