Saliency-Based Detection of Identity Documents Captured by Smartphones

Minh Ôn Vũ Ngọc, Jonathan Fabrizio, Thierry Géraud
EPITA Research and Development Laboratory (LRDE), France
firstname.lastname@lrde.epita.fr

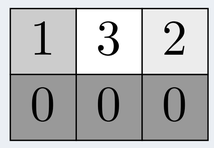
At a Glance

Problem statement:
- Automatic segmentation of identity documents in smartphone photos.
- Many challenges: unknown background, lighting condition, acquisition problems ...

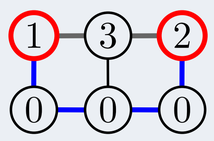
Why our approach is interesting:
- New different approach.
- Fast, efficient but state-of-the-art results.

Main contributions:
- An extension to color images of the Dahu distance.
- A framework to detect identity documents in photos or videos.
- Comparison of different saliency-based methods.

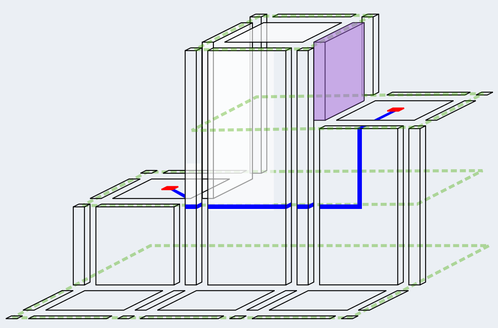
Image representations for computing barrier distances



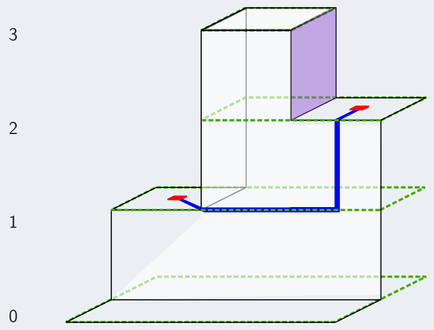
(a) An image $\boldsymbol{u}$.



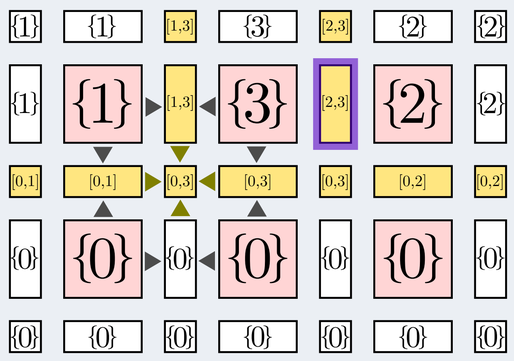
(b) $\boldsymbol{u}$ as a graph.



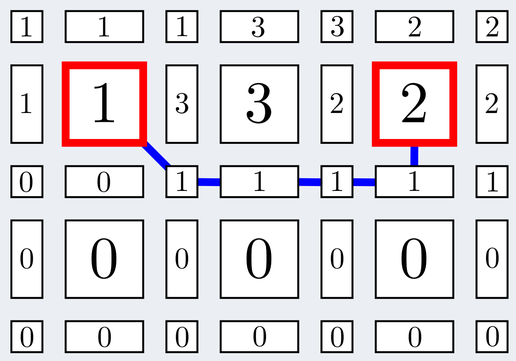
(c) 3D version of $\widetilde{\boldsymbol{u}}$ given in (e).



(d) $\boldsymbol{u}$ as a surface.



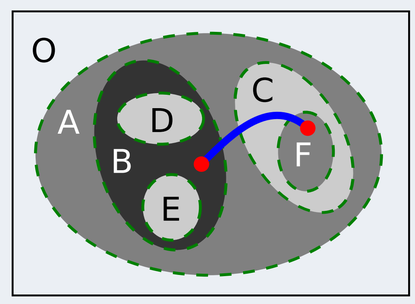
(e) Interval-valued image $\widetilde{\boldsymbol{u}}$.



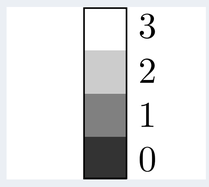
(f) A minimal path in a $\overline{\boldsymbol{u}} \hookleftarrow \widetilde{\boldsymbol{u}}$.

Barrier τ of a path π: $\tau_u(\pi) = \max_{\pi_i \in \pi} u(\pi_i) - \min_{\pi_i \in \pi} u(\pi_i)$.
MB distance: $d_u^{\text{MB}}(x, x') = \min_{\pi \in \Pi(x, x')} \tau_u(\pi)$,
Dahu distance: $d_u^{\text{DAHU}}(x, x') = \min_{\overline{u} \hookleftarrow \widetilde{u}} d_{\overline{u}}^{\text{MB}}(h_x, h_{x'})$,
where h_x: 2D element of the complex corresponding to x

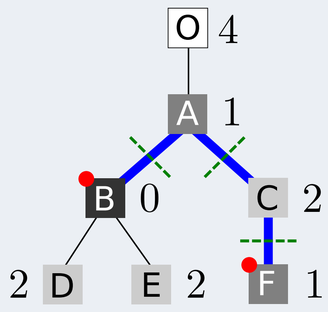
The Dahu distance on the tree of shapes



(g) Image $\boldsymbol{u}$.



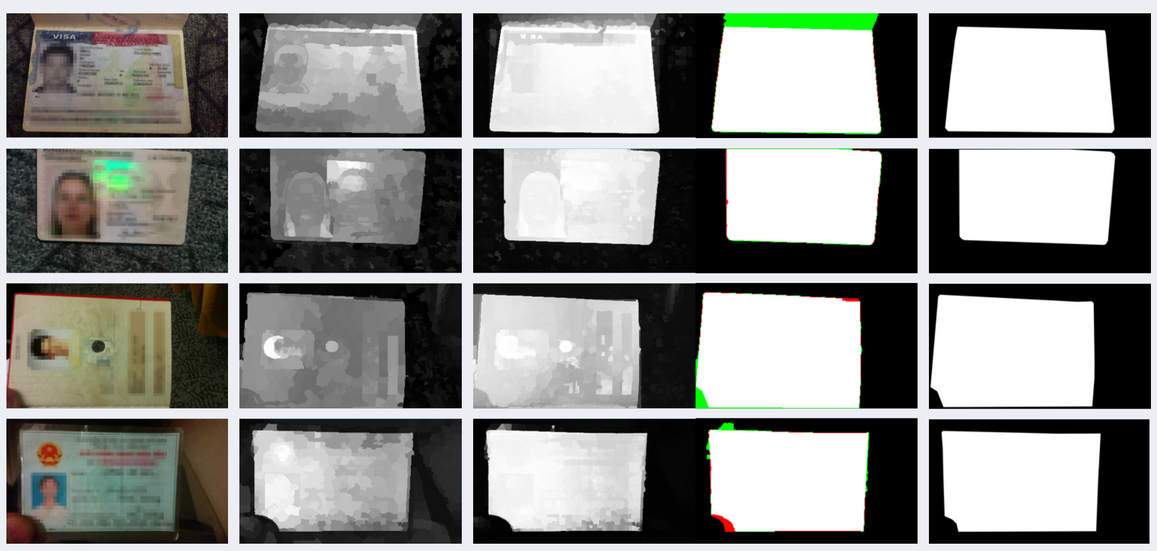
(h) Gray scale.



(i) Tree $\mathfrak{S}(\boldsymbol{u})$.

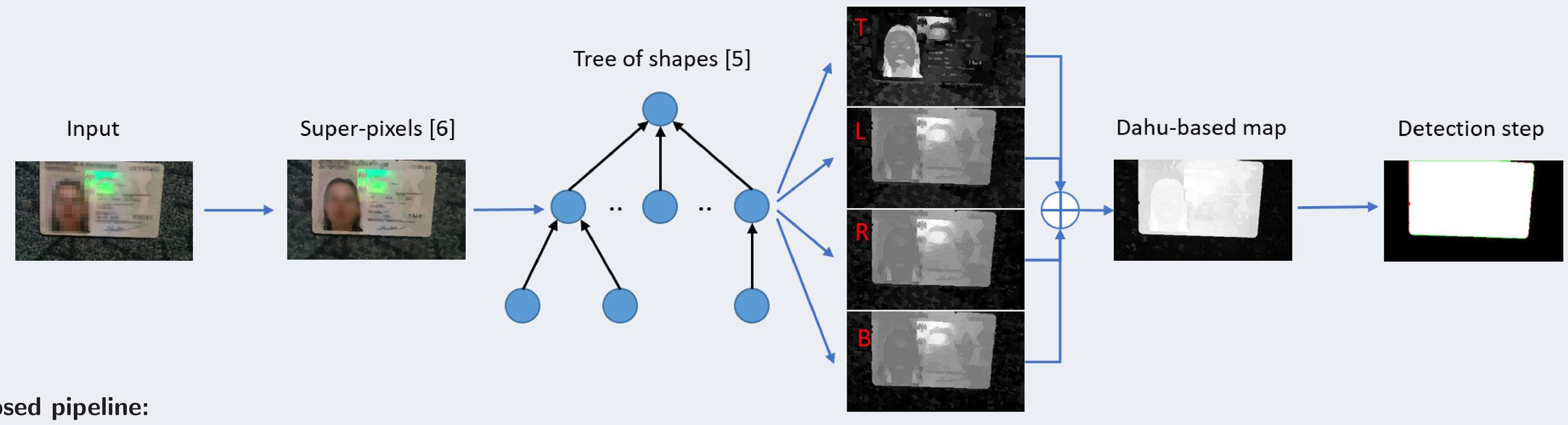
Gray Image u: $d_u^{\text{DAHU}}(x, x') = d_{\mathfrak{S}(u)}^{\text{MB}}(t_x, t_{x'})$, t_x: node containing x [1].
Color Image u: $d_{\text{u}}^{\text{DAHU}}(x, x') = \sum_{i \in \{R,G,B\}} \tau_{\text{u}}^{(i)}(\dot{\pi}(t_x, t_{x'}))$
with $\tau_{\text{u}}^{(i)}(\dot{\pi}) = \max_{t \in \dot{\pi}} \mu_{\text{u}}^{(i)}(t) - \min_{t \in \dot{\pi}} \mu_{\text{u}}^{(i)}(t)$, $\mu_{\text{u}}^{(i)}(t)$: color of node t.

Some qualitative results



(j) Input (k) Dahu-based... (l) ...normalized (m) Detection (n) Ground truth

Proposed solution



Proposed pipeline:

Step 1: Simplify image into super-pixels;
Step 2: Compute a tree of shapes from the graph of super-pixels;
Step 3: Produce a saliency map by using the Dahu distance on the tree of shapes;
Step 4: Apply detection step.

Quantitative results

Method	MAE	$\overline{F_\beta}$
GS [2]	0.328	0.573
MR [3]	0.299	0.642
SO [4]	0.265	**0.7461**
Dahu	**0.178**	**0.7465**

Table: MAE and $\overline{F_\beta}$

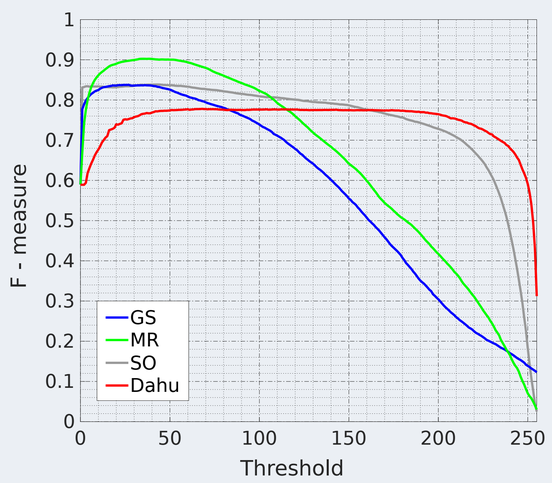


Figure: F_β w.r.t. saliency map thresholding.

Selected Bibliography

[1] T. Géraud, Y. Xu, E. Carlinet, and N. Boutry, "Introducing the Dahu pseudo-distance.," in *Proc. of ISMM*, ser. LNCS, vol. 10225, 2017, pp.55-67.

[2] Y. Wei, F. Wen, W. Zhu and J. Sun, "Geodesic saliency using background priors," in *Proc. of ECCV*, pp. 29-42, 2012.

[3] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in *Proc. of CVPR*, 2013, pp. 3166-3173.

[4] W. Zhu, S. Liang, Y. Wei, and J. Sun, "Saliency optimization from robust background detection," *Proc. of ICPR*, 2014, pp. 2814-2821.

[5] E. Carlinet, and T. Géraud, "MToS: A tree of shapes for multivariate images," *IEEE Transactions on Image Processing*, vol. 24, no.12, pp.5330-5342, 2015.

[6] R. Achanta et al., "SLIC superpixels compared to state-of-theart superpixel methods," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp.2274-2282, 2012.