Real-Time Salient Closed Boundary Tracking using Perceptual Grouping and Shape Priors

University of Alberta, CANADA

Xuebin Qin, Shida He, Zichen Zhang, Masood Dehghan and Martin Jagersand

BMVC 2017 LONDON

1. Overview

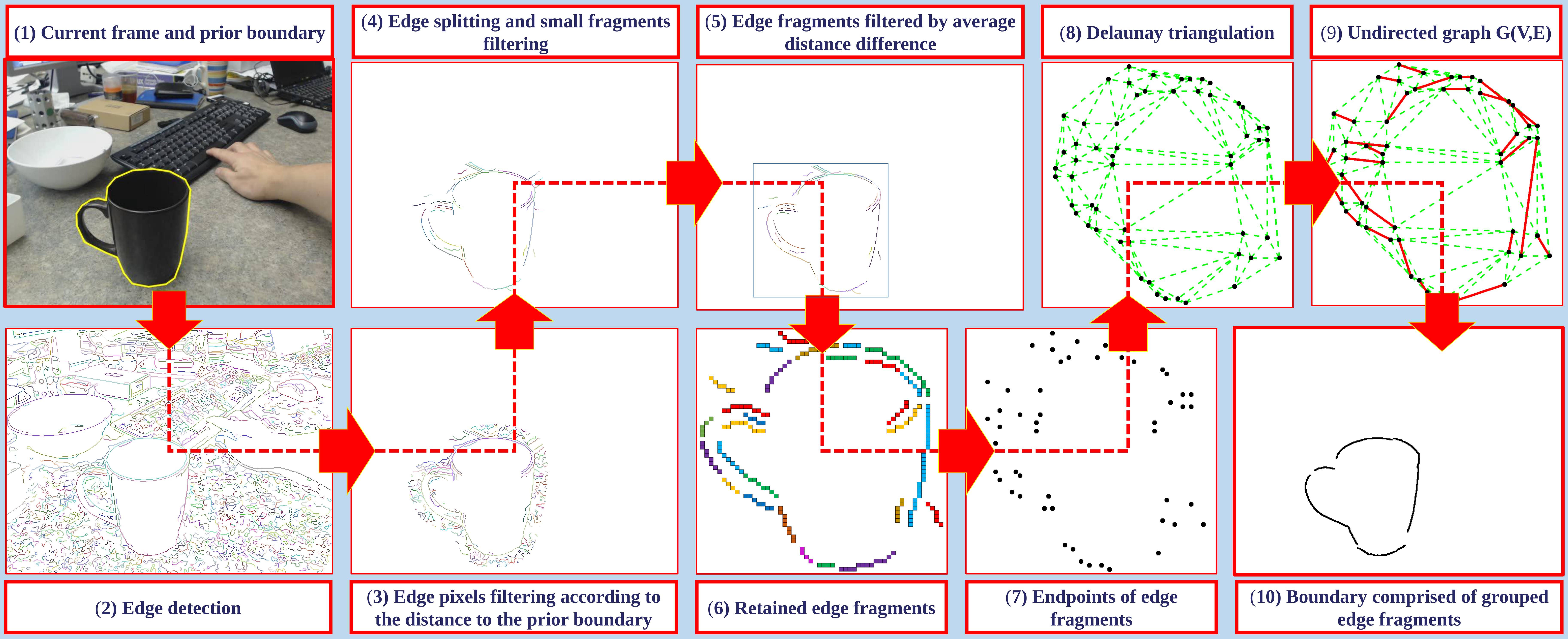
$$\Gamma(B) = \frac{|G_B| + |DD_B|}{\iint_{R(B)} dxdy}$$

This paper presents a real-time method for accurate salient closed boundary tracking via a combination of shape constraints and perceptual grouping on *edge fragments*. We encode the Gestalt law of proximity and the prior shape constraint in a novel ratio-form grouping cost ($\Gamma(B)$). The proximity and prior constraint are depicted by the relative gap length ($|G_B|$) and average distance difference ($|DD_B|$) along the to-be-tracked boundary with respect to its area ($|\iint_{R(B)} dxdy|$). The perimeter and area variations of boundary grouping are also constrained as $v(P)<e_P$ and $v(A)<e_A$.

We then search the optimal boundary from an undirected graph $G = (V,E)$, which has the minimal $\Gamma(B)$ and satisfies variation constraints ($v(P)<e_P$ and $v(A)<e_A$). We validated our tracker **(EFG)** on a public video dataset and compared its results with those of other two methods **(RRC)** and **(BDSP)**.

$$\begin{cases} v(P) = min(\frac{P_{prior}}{P_{cur}}, \frac{P_{cur}}{P_{prior}}) < e_P \\ v(A) = min(\frac{A_{prior}}{A_{cur}}, \frac{A_{cur}}{A_{prior}}) < e_A \end{cases}$$

2. Method



3. Experimental Results

3.1 Quantitative Evaluations

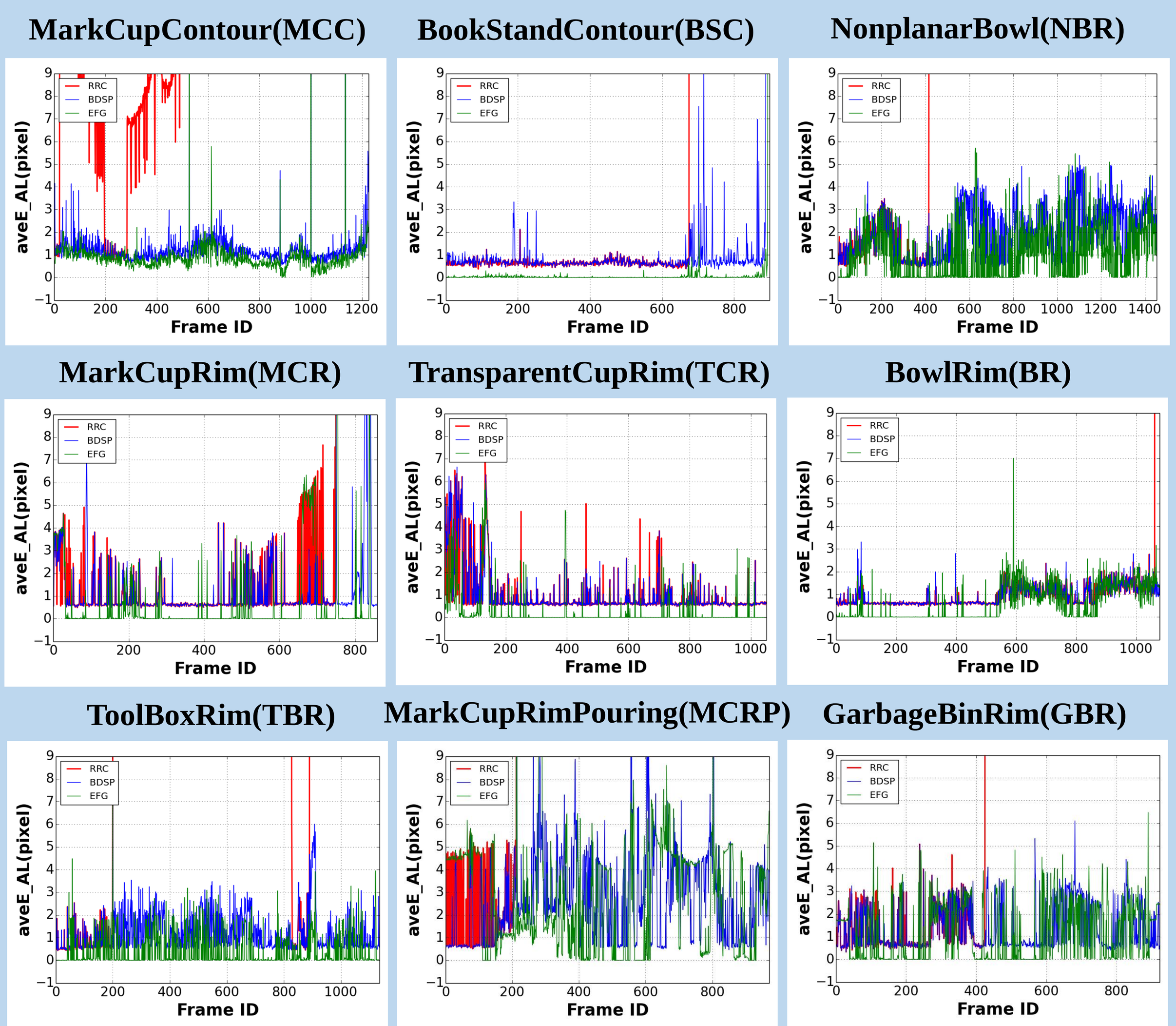


Figure 1: Average alignment error *aveE_AL* of the closed boundary tracking achieved by trackers RRC, BDSP and our EFG.

Video	**MCC**	**BSC**	**NBR**	**MCR**	**TCR**	**BR**	**TBR**	**MCRP**	**GBR**
RRC	56.09	58.06	120.48	19.29	1.01	3.37	74.09	58.18	108.28
BDSP	1.31	1.24	1.99	1.18	1.03	0.94	1.21	**2.73**	1.38
EFG	**0.99**	**0.18**	**1.08**	**0.68**	**0.26**	**0.61**	**0.46**	2.80	**1.22**

Table 1: Average *aveE_AL* (pixel) of each sequence.

Video	MCC	BSC	NBR	MCR	TCR	BR	TBR	MCRP	GBR
$aveEFs$	27	23	25	36	30	25	29	35	28
$aveGFs$	145	120	137	200	162	136	155	192	154
$T(ms)$	**28.11**	**20.37**	**21.60**	**36.39**	**26.45**	**22.31**	**26.39**	**35.08**	**30.92**

Table 2: Average tracking time costs for each video sequence: $aveEFs$ and $aveGFs$ are the average numbers of Edge Fragments and Gap Filling segments. $2 \times aveEFs$ is the number of graph vertices and $aveEFs + aveGFs$ is the number of graph edges. T is the average time cost of each frame tracking in milliseconds. It includes edge segments detection, splitting, filtering, graph construction and optimization. The frame loading time is not included because this process can be implemented in parallel and will not influence the tracking speed significantly.

3.2 Qualitative Evaluations

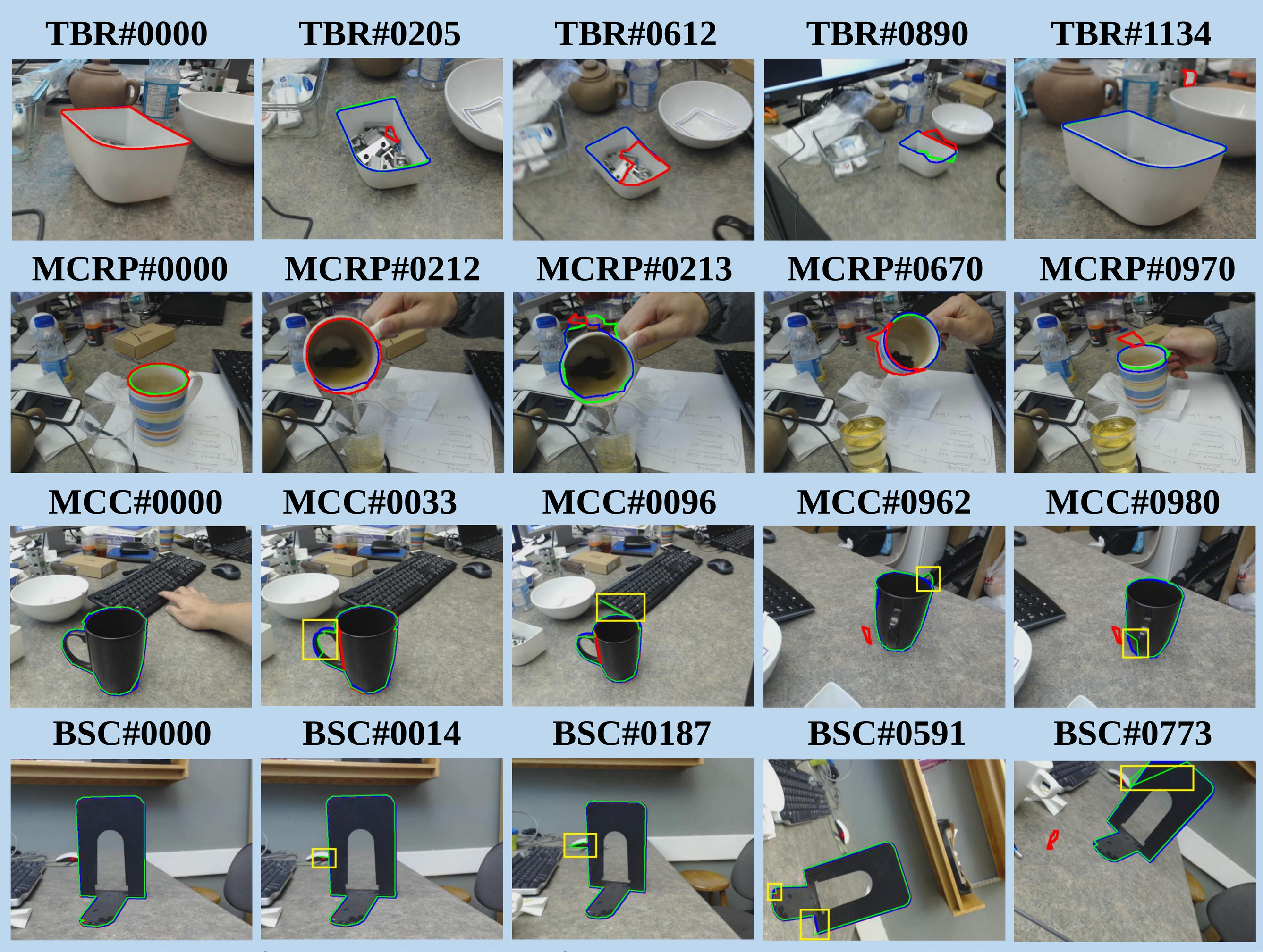


Figure 2: Failures of RRC and wiggles of BDSP: Red, green and blue boundaries are results of RRC, BDSP and EFG respectively. All of the four rows demonstrate the failure of RRC tracking. The last two rows show the wiggles (within yellow boxes) produced by BDSP.

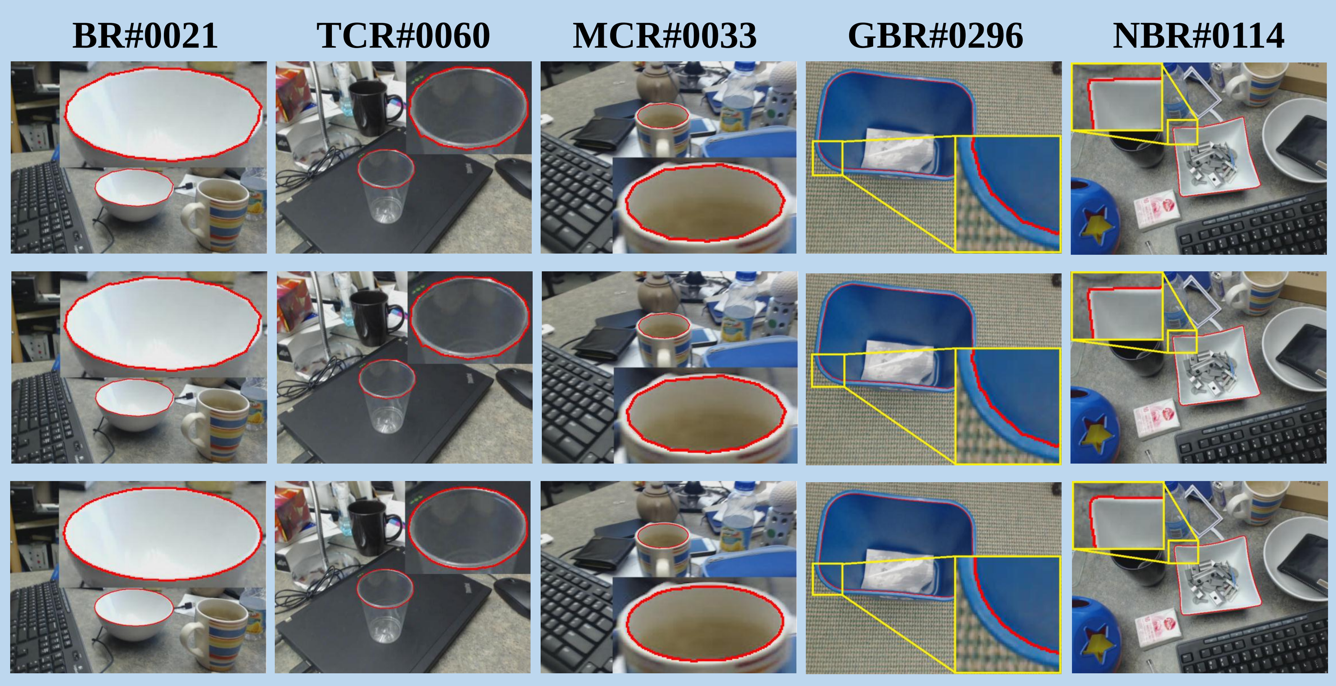


Figure 3: Comparison between line segments and edge fragments represented boundaries. From top row to bottom row are results of RRC, BDSP and our EFG respectively.

4. Conclusions

Our tracker encodes the prior shape constraint into the distance difference of deliberately split edge fragments and combines it with the boundary salient measure of relative gap length. It suppresses most of the small erroneous wiggles on the boundaries and improves the tracking accuracy. We validated our method on real-world video sequences and achieved the state-of-the-art results both qualitatively and quantitatively.

Related resources:
http://webdocs.cs.ualberta.ca/~xuebin/
Our group page:
https://webdocs.cs.ualberta.ca/~vis/