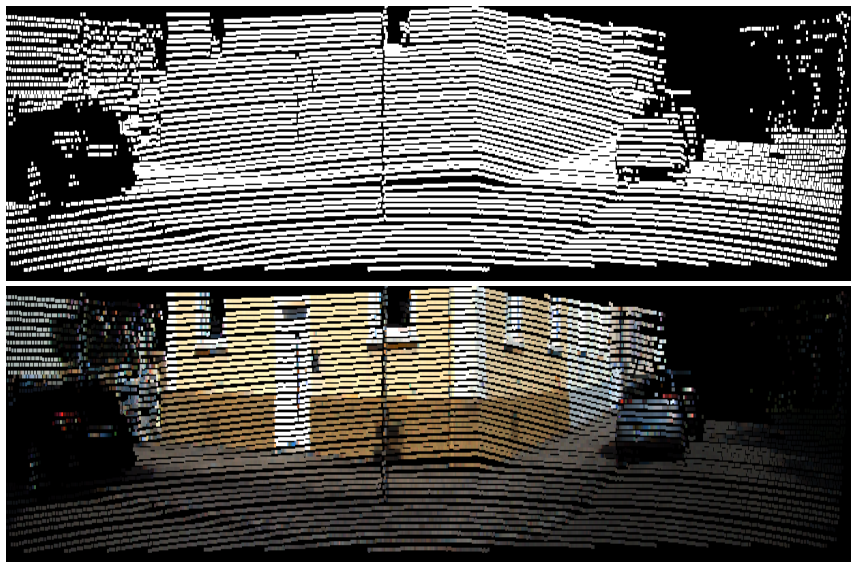
c-M2DP: A FAST POINT CLOUD DESCRIPTOR WITH COLOR INFORMATION TO PERFORM LOOP CLOSURE DETECTION

Leonardo Perdomo, Diego Pittol
Mathias Mantelli, Renan Maffei
Mariana Kolberg and Edson Prestes

August, 2019

MOTIVATION

- 3D spatial data can provide more descriptive scenes than only 2D images;
- Color can increase 3D descriptiveness[1], but is insufficiently investigated for loop closure detection;
- Availability in recent benchmark platforms[2].



Colored point cloud generated using LIDAR and camera.

[1] Tombari, Salti & Stefano (2011), Feng, Liu & Liao (2015), Logoglu, Kalkan & Temizel (2016)

[2] Pandey, McBride & Eustice (2011), Geiger, Lenz & Urtasun (2012)

PROPOSAL

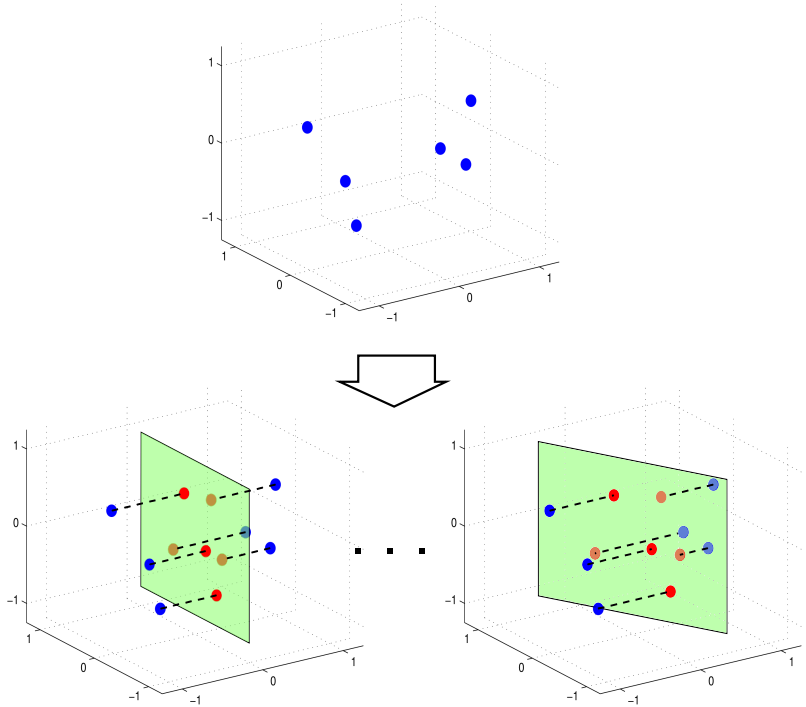
- Color M2DP (c-M2DP), a global descriptor comprising of color and shape data computed from the point cloud:
  - Takes advantage of M2DP's structure, extending it to compute color signatures from multiple 2D projections;
- An improved loop closure detection, using the c-M2DP descriptor on point cloud sequences generated through camera-LIDAR fusion, or stereo depth estimation.

REFERENCE FRAME

- Rotation and shift invariance in 3D space;
- Point cloud $\boldsymbol{P}$ centroid is computed and used as the reference frame origin;
- PCA is performed on $\boldsymbol{P}$, with both the 1st and 2nd principal components defined as the x-axis and y-axis, respectively.

MULTIPLE 2D PROJECTIONS

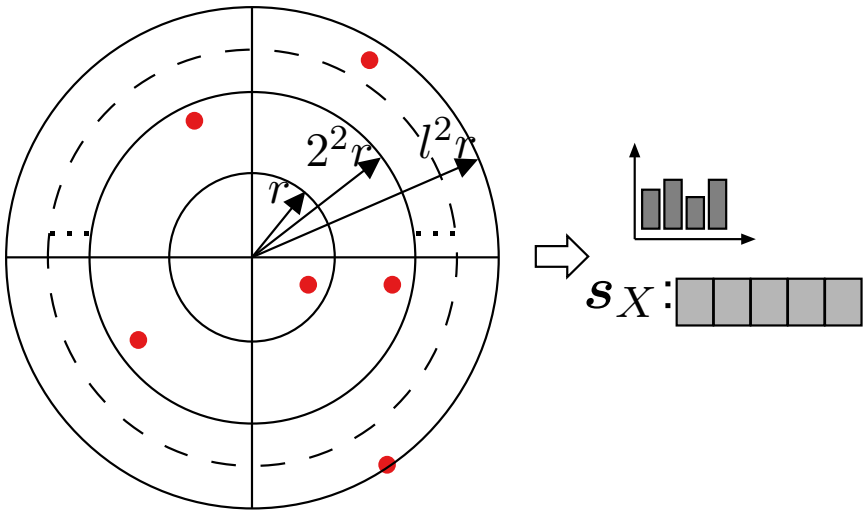
- A 2D plane X is defined with normal vector $\boldsymbol{m}$ and $[\theta, \phi]$ parameters, which are used to generate distinct 2D planes;
- $\boldsymbol{P}$ is projected onto each 2D plane, in order to compute shape and color signatures from each 2D projection.



Projecting a point cloud $\boldsymbol{P}$ on multiple 2D planes.

SHAPE SIGNATURE

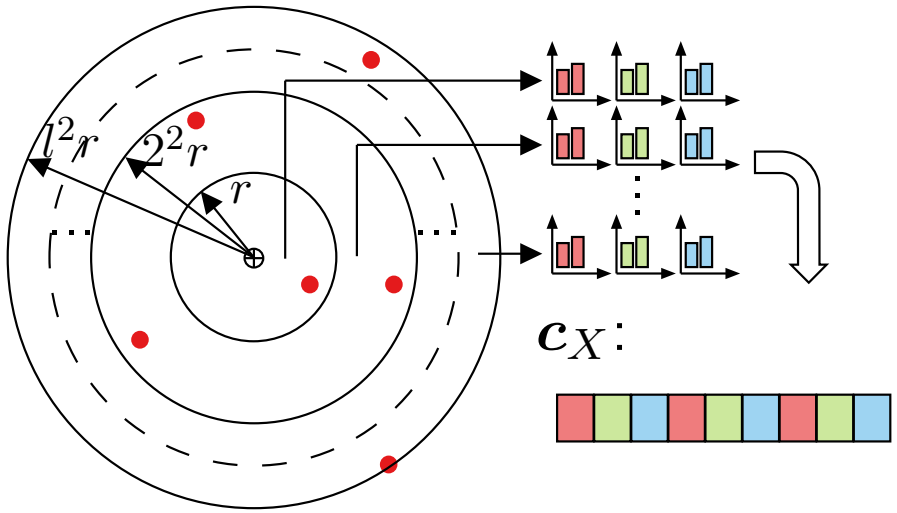
- Each plane is split into l concentric circles with varying radii $[r, 2^2 r, \ldots, l^2 r]$;
- Each concentric circle is divided in t shape bins, indexed by the x-axis;
- Shape signature $\boldsymbol{s}_X$ is computed by counting the points within each bin.



Computing the shape signature $\boldsymbol{s}_X$.

COLOR SIGNATURE

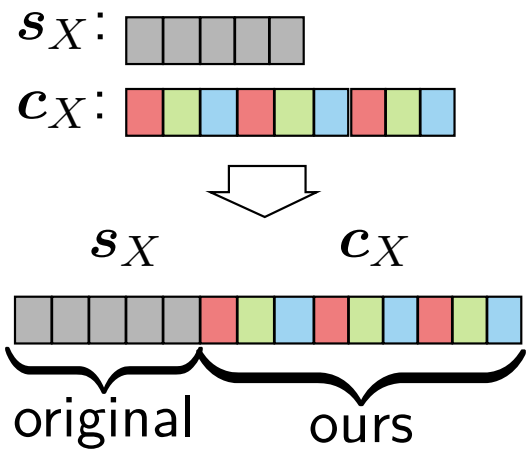
- For every concentric circle, histograms of the color channels, which are divided in j bins, are computed;
- These histograms are concatenated them into a single color signature vector $\boldsymbol{c}_X$;



Computing the color signature $\boldsymbol{c}_X$.

SIGNATURE VECTOR

- Both $\boldsymbol{s}_X$ and $\boldsymbol{c}_X$ are normalized and concatenated into a single signature vector;
- The signature matrix $\boldsymbol{A}$ is augmented by a row with the concatenated vector.



Concatenated shape and color signatures.

DIMENSIONALITY REDUCTION

- For every 2D projection, both shape and color signatures are generated, concatenated and included into $\boldsymbol{A}$;
- SVD of $\boldsymbol{A}$ is computed, with the resulting 1st left and right singular vectors being concatenated and used as the final descriptor.

M2DP AND C-M2DP PARAMETERS

- M2DP parameters (shared with c-M2DP) were the same from original work;
- Color bins for each color channel were set $t = j$;
- Color channels were fixed $h = 3$, i.e. RGB, HSV and CIELab;

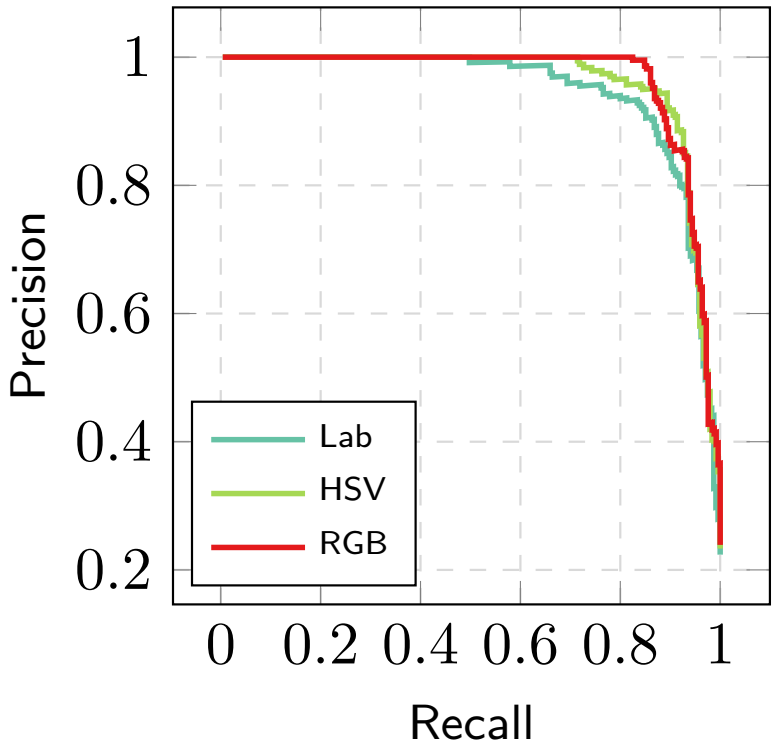
M2DP and c-M2DP Parameters

Parameter	**M2DP**	**c-M2DP**
Azim. angles (b)	4	4
Elev. angles (q)	16	16
Conc. circles (l)	8	8
Shape bins (t)	16	16
Color bins (j)	-	16
Vector length	192	576

c-M2DP COLOR SPACE

- c-M2DP color space was chosen after evaluating it using RGB, HSV and CIELab.

Descriptor	**Recall at Precision** 100%
RGB	82.5%
HSV	71.4%
CIELab	49.8%



KITTI 06 LIDAR-camera.

TIME EFFICIENCY

- In camera-LIDAR sequences:
  - c-M2DP computing time is only 23.2% higher than M2DP;
  - c-M2DP is 22.6% faster to compute than CSHOT.

Average times computing and matching a descriptor.

Descriptor	**Computing (s)**	**Matching (s)**
M2DP	$\mathbf{0.0674} \pm 0.0041$	$\mathbf{0.0043} \pm 0.0004$
c-M2DP	$\mathbf{0.0830} \pm 0.0052$	$\mathbf{0.0051} \pm 0.0006$
CSHOT	0.1072 ± 0.0168	0.0059 ± 0.0005

TIME EFFICIENCY

- In stereo sequences:
  - Overall increase in the average times computing the descriptors;
  - c-M2DP computing time is only 18.8% higher than M2DP;
  - CSHOT computational burden, with an average time 315.9% higher than c-M2DP.

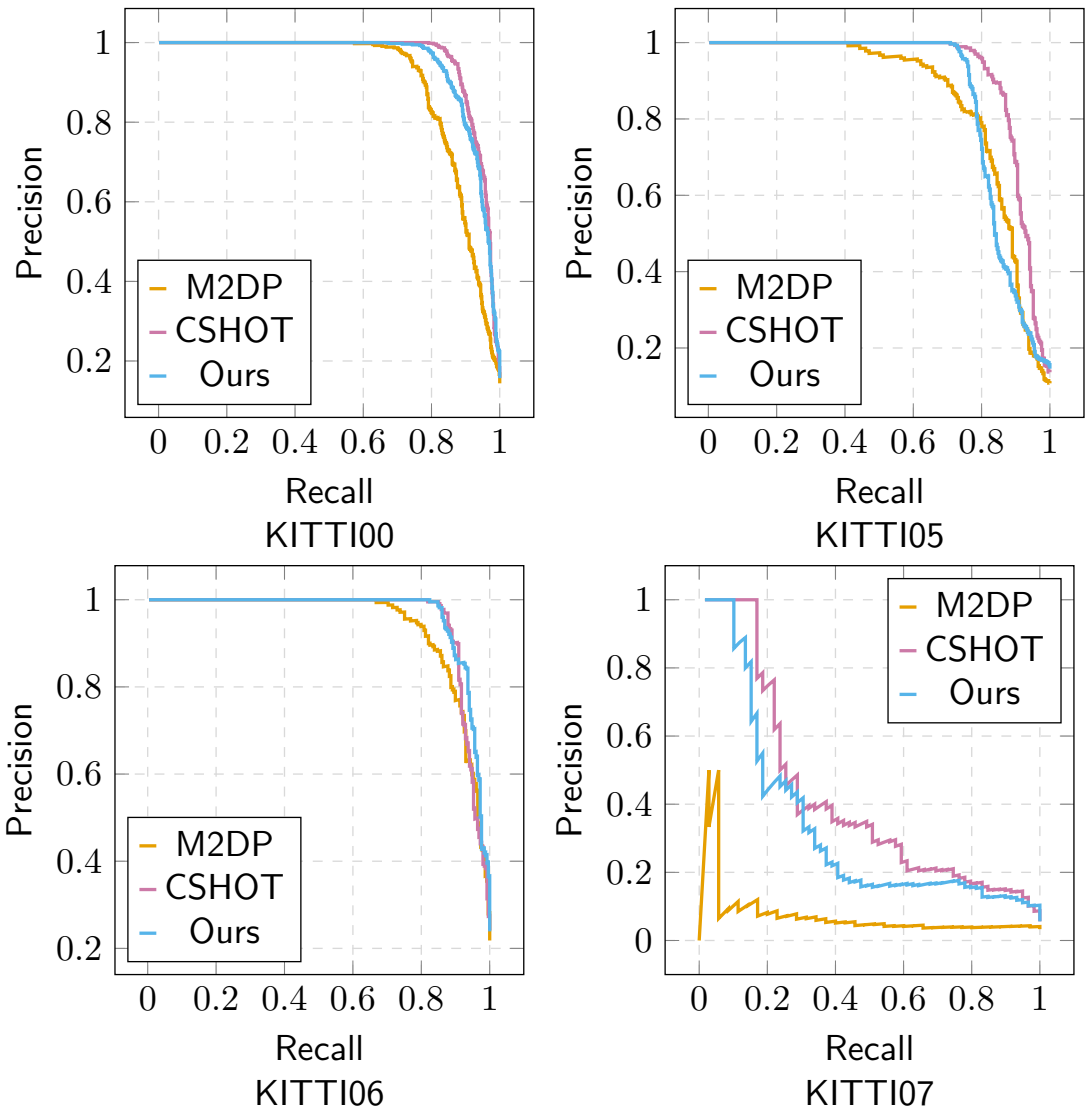
Descriptor	**Computing (s)**	**Matching (s)**
M2DP	$\mathbf{0.3584} \pm 0.0816$	$\mathbf{0.0044} \pm 0.0008$
c-M2DP (Ours)	$\mathbf{0.4259} \pm 0.0956$	$\mathbf{0.0054} \pm 0.0006$
CSHOT	1.7711 ± 1.0159	0.0061 ± 0.0005

PRECISION-RECALL CAMERA-LIDAR

Recall at Precision 100% on KITTI Camera-LIDAR Point Clouds

Sequence	M2DP	c-M2DP (Ours)	CSHOT
KITTI00	0.574303	0.673295	$\mathbf{0.791549}$
KITTI05	0.408935	$\mathbf{0.708861}$	0.708108
KITTI06	0.668122	$\mathbf{0.824701}$	0.818898
KITTI07	0	0.101695	$\mathbf{0.169492}$

PRECISION-RECALL CAMERA-LIDAR

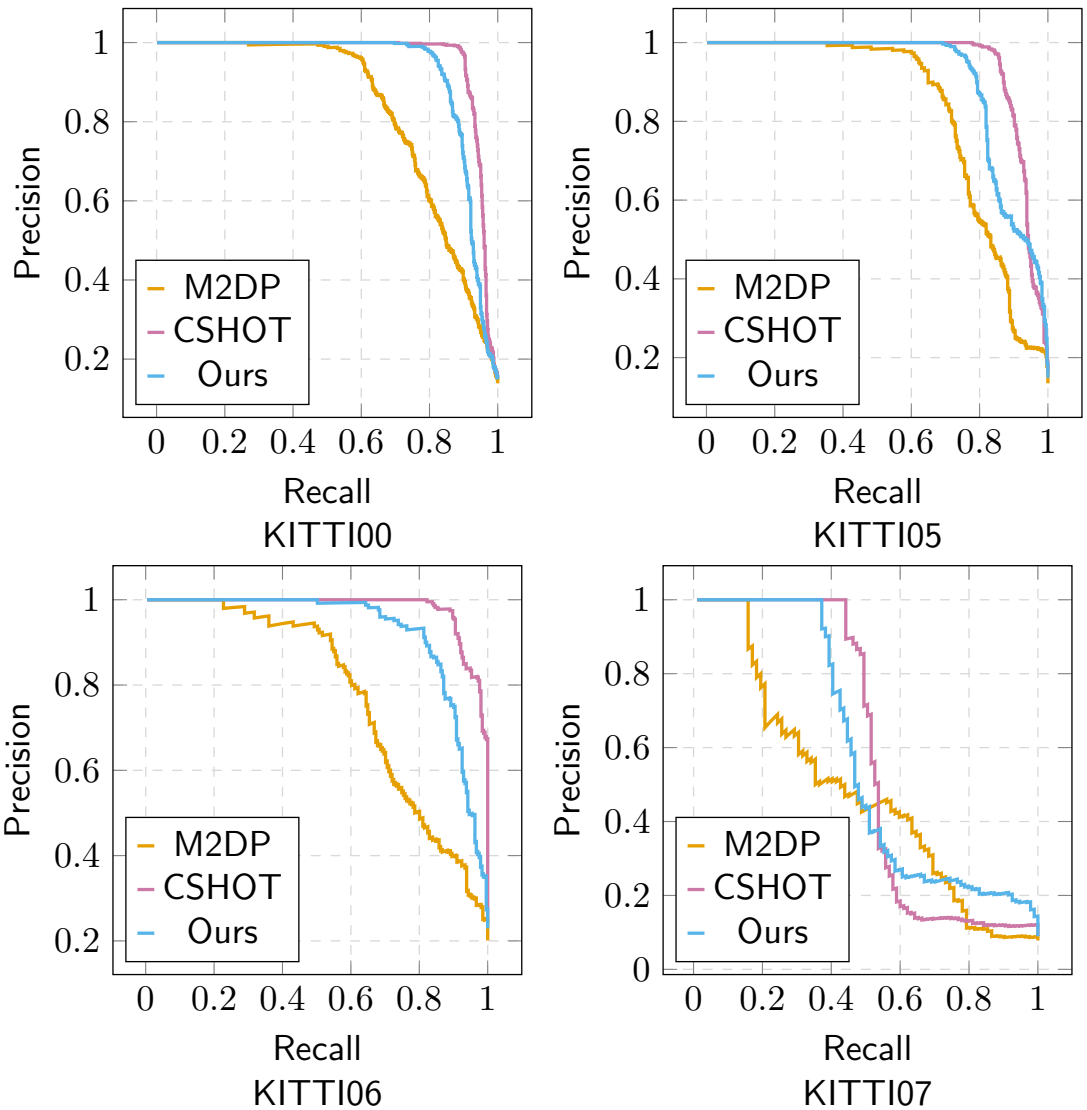


PRECISION-RECALL STEREO

Recall at Precision 100% on KITTI Stereo Point Clouds

Sequence	M2DP	c-M2DP (Ours)	CSHOT
KITTI00	0.269663	0.697466	$\mathbf{0.709402}$
KITTI05	0.353425	0.692308	$\mathbf{0.778539}$
KITTI06	0.227488	0.502075	$\mathbf{0.822835}$
KITTI07	0.158537	0.372340	$\mathbf{0.442105}$

PRECISION-RECALL STEREO



CONCLUSION

- Overall accuracy improvement of the c-M2DP descriptor over the original M2DP;
- In camera-LIDAR sequences, c-M2DP is faster to compute and shows competitive results against CSHOT;
- As expected in stereo sequences, CSHOT shows higher accuracy at the cost of being several times slower than c-M2DP.

THANKS!

Leonardo Perdomo, Diego Pittol
Mathias Mantelli, Renan Maffei
Mariana Kolberg and Edson Prestes
Instituto de Informática — UFRGS