Local-convexity reinforcement for scene reconstruction from sparse point clouds

Maxime Lhuillier

Institut Pascal, UMR 6602 CNRS/UCA/SIGMA, Clermont-Ferrand

IC3D, 11th december 2019, Brussels, Belgium

Introduction

Motivations

3D reconstruction of outdoor environments using consumer 360 camera

Potential applications: content creation for VR, scene modeling

Advantages: weak experimental constraints, avoiding costly 3D scanners

360 camera examples

Source: 360rumors.com

Introduction

Reminders about automatic scene modeling from images (still images & vidéo)

- There are usually two main steps: sparse and dense reconstructions
- Sparse: estimate surface from 3D reconstruction of sparse features in images
- Dense: refine the surface using all image pixels
- Sparse alone can also be useful in several contexts: large scale environments, limited hardware resource, if applications do not need high levels of details
- Most sparse methods compute a 3D Delaunay triangulation whose tetrahedra are labeled freespace or matter (the surface is between freespace-matter)

Paper contribution

- We propose corrections based on local-convexity to improve this labeling
- They can be used as preprocessing or postprocessing of most sparse methods

Local-convexity (LC)

Definition

Let M be a set of points in $\mathbb{R}^k$

M is locally-convex at $x \in M$ if x has a neighborhood $N \subset \mathbb{R}^k$ such that $N \cap M$ is convex

Reminder: X is convex if line segment $yz \subseteq X$ for all points $y \in X$ and $z \in X$

In the original definition [Karshon07], N is an Euclidean (spherical) ball centered at x

In our context

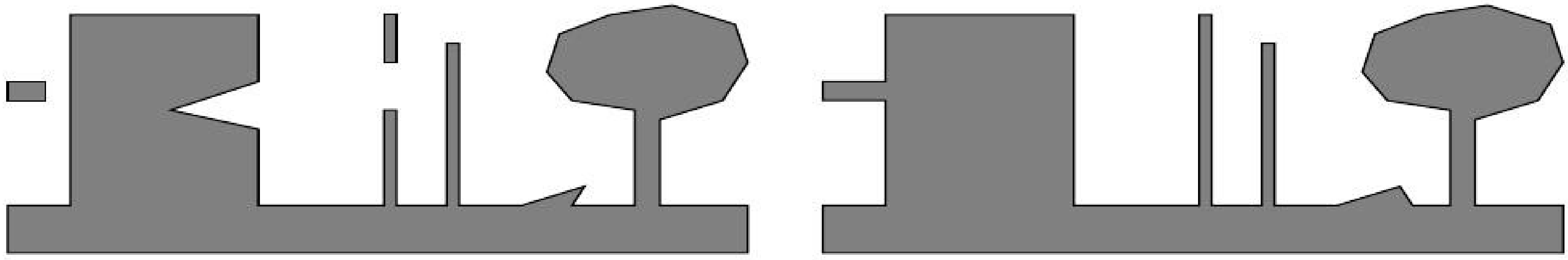
M is the matter in $\mathbb{R}^3$, ie the union of the matter tetrahedra in the 3D Delaunay triangulation T

The most notable errors in M are falsely-labeled freespace tetrahedra due to bad or lacking points (eg a tetrahedron is freespace if it is crossed by a line-of-sight between a point and a camera)

In short, a freespace tetrahedron is relabeled matter by our corrections if all its vertices are in matter tetrahedra and if the tetrahedron is “small enough”

N can be anisotropic due to the prevalence of vertical structures in usual scenes

Local-convexity for a simple example



What we have before correction (left)

The matter M is gray and the freespace is white

A balcony and a post are disconnected, a building has a spurious concavity

What we would like after correction (right)

Fill the building concavity and the gaps of the balcony and the post (eg the gap of the post is filled by choosing points y and z in different components of the post to connect it)

Fill neither the gap between the two posts nor the gap between the ground and the tree foliage (both locality and anisotropy are useful to do this)

Previous work

LC and its variants : convexity, weak-convexity, star-shapeness

LC was ignored for surface reconstruction from images, thus we consider other problems

1) LC is used to segment surface regions into the ground surface and objects [Moosmann09]

2) Point clouds are segmented into weakly convex regions [Asafi13]

3) A star-shaped objet is segmented in an image if the starshape center is given [Veksler08]

4) Foreground and background are segmented in an image by enforcing convexity [Gorelick14]

LC alternatives for shape completion

Previous work uses LC alternatives but needs strong assumptions to estimate primitives

1) Planes are estimated from a dense point cloud, then a simplified surface is generated using priors: the prevalence of vertical structures and orthogonal intersections [Chauve10]

2) A 3D space partition is built assuming that the surface is only horizontal or vertical [Oesau14]

3) A sweep surface (which generalizes ruled surface) is estimated for architectural scenes [Wu12]

Method 1: explicit neighborhoods

There are several ways to apply LC to the matter tetrahedra of the 3D Delaunay triangulation T

Choose a norm ||.|| to define neighborhoods by its unit ball

Its unit ball is a cylinder (vertical axis, size similar to the expected level of detail of the shape)

LC of M: if $a \in M$ and $b \in M$ and $\|a-b\| \leqslant 1$, then the line segment ab should be included in M

Principle of method 1

1) Check whether every vertex v_i of T is in a matter tetrahedron (boolean $b(v_i)$ is true if yes)

2) Check whether every edge $v_i v_j$ of T is in matter using LC of M (boolean $b(v_i v_j)$ is true if yes)

3) Relabel matter every freespace tetrahedron of T if its 4 vertices and 6 edges have true booleans

Two steps for computing the vertex booleans b(v)

Step 1 is a simple initialization: let b(v)=true iff there is a matter tetrahedron in T with vertex v

Step 2: search a tetrahedron $v_i v_j v_k v_l$ using T vertices in the neighborhood of v, reset b(v)=true if the 4+6 booleans are true and $v \in v_i v_j v_k v_l$ (ie v is in the convex hull of v_i, v_j, v_k and v_l)

Methods 2&3: implicit neighborhoods

Another kind of methods enumerates a lot of sets of freespace tetrahedra included in T, then all tetrahedra in a set S are relabeled matter if there are enough matter tetrahedra that surround S

Choice of sets

They should be neither too large nor too numerous for computation

The tetrahedra in S share a common vertex and S is strongly connected

S is strongly connected if its adjacency graph D is connected (D has a vertex for each tetrahedron in S, D has an edge for each pair of tetrahedra in S that have a common triangle face)

Isotropic surrounding measure c(S)

Let ∂S be the boundary of S (ie the set of every triangle that is a face of a single tetrahedron in S)

Let $\partial_m S$ be the triangles in ∂S that are also faces of matter tetrahedra in T/S

Let c(S) be the area of $\partial_m S$ divided by the area of ∂S (thus $0 \leqslant c(S) \leqslant 1$)

If c(S) is close to 1, S is well surrounded by matter tetrahedra. In this case, most tetrahedra in S are in the convex hull of matter points (in $\partial_m S$), thus S should be relabeled matter by LC of M

Method 2: remove surface peaks

Here every tried set S is a strongly connected component of the freespace tetrahedra sharing a vertex. In contrast to method 1, method 2 does not limit the geometric size of the relabeled tetrahedra.

Reminders

The surface is the set of triangles separating matter tetrahedra and freespace tetrahedra (of T)

A peak is a vertex such that the ring of its incident surface triangles defines a small solid angle

Principle of method 2

For each vertex v of the surface:

1) Compute the set(s) S sharing v by a graph traversal in the adjacency graph of the T tetrahedra

2) Compute c(S) and the solid angle w(S) of S with apex v (note: $w(S)=\sum_{vabc \in S} SolidAngle(v, vabc)$)

3) Relabel matter S if w(S) is less than a threshold and c(S) is greater than another threshold

(Note: method 2 generalizes the Peak Removal operation in [Lhuillier13] for non-manifold surfaces)

Method 3: use anisotropy & freespace confidence

In contrast to method 2, the sets tried by method 3 take into account the scene anisotropy (i.e. the prevalence of vertical structures) and the freespace confidence of the tetrahedra.

Freespace confidence

Every tetrahedron has a freespace confidence: the number of lines-of-sight that cross it.
The tetrahedra with the smallest freespace confidences should be relabeled matter in priority.

Anisotropic surrounding measure c'(S)

Reminder: c(S) is a ratio between two sums of areas of triangles in the boundary of S

Obtain c'(S) by replacing every triangle area in c(S) by its projection on the horizontal plane

Principle of method 3

For each vertex v of the surface:

1) Order $\Delta_{1,} \Delta_{2,} \ldots \Delta_n$ the freespace tetrahedra sharing v by increasing freespace confidence

2) Enumerate the sets S by all strongly connected components of all subsets $\{\Delta_{1,} \Delta_{2,} \ldots \Delta_i\}, i \leq n$

3) Relabel matter S if S maximizes c'(S) and c'(S) is greater than a threshold

Notations for experiments

Combinations of our correction methods

Method 2+1: first apply method 2 (peak removal), then apply method 1 (explicit N)

Method 2+3, first apply method 2 (peak removal), then apply method 3 (implicit N)

More details in the paper (omited in the talk: final peak removal in the matter side)

Previous surface reconstruction methods and their corrected version

A manifold method in [Lhuillier18] named M

Method M(2+1): first apply 2+1 on the tetrahedron labeling obtained by ray-casting, then apply M. The definition of M(2+3) is very similar.

Graph-cut methods in [Vu12] and [Jacosek11] named G1 and G2, respectively

Method Gi(2+1): first apply Gi, then apply 2+1. The definition of Gi(2+3) is similar

Dataset

The sparse input point cloud is computed from two 2.5k videos at 30Hz.
They are taken by a Garmin Virb 360 camera moving in a city (The trajectory is 6.7km long)
It is mounted on the top of a car using a small mast.
There are 5.5M vertices in T which are reconstructed from 6.5k keyframes selected in the videos by structure-from-motion. Curves are reconstructed (more details in [Lhuillier18])

Fig. 2. One keyframe, the Garmin Virb 360 camera, and a bottom view of the vertices of T.

Only use corrections (1/3)

We compare the initial labeling by ray-casting (ie a tetrahedron is freespace iff it is crossed by a line-of-sight) and its relabeling by several methods.

Fig.3 shows labeling differences near a 5-way crossroad by projecting the freespace tetrahedra (in black) to the horizontal plane.

Our methods removes the largest freespace tetrahedra that are inconsistent with the geometry of the urban corridors.

Fig. 3. Large scale view of our corrections. From left to right: vertices of T, initial labeling by ray-casting, results of the correction methods $2 + \tilde{2}$, $2 + 1$ and $2 + 3$. Matter is white and freespace is black.

Only use corrections (2/3)

Fig.4 shows that our methods change topology: trunk components are connected, tunnels (ie holes in the matter) in a traffic sign are filled.

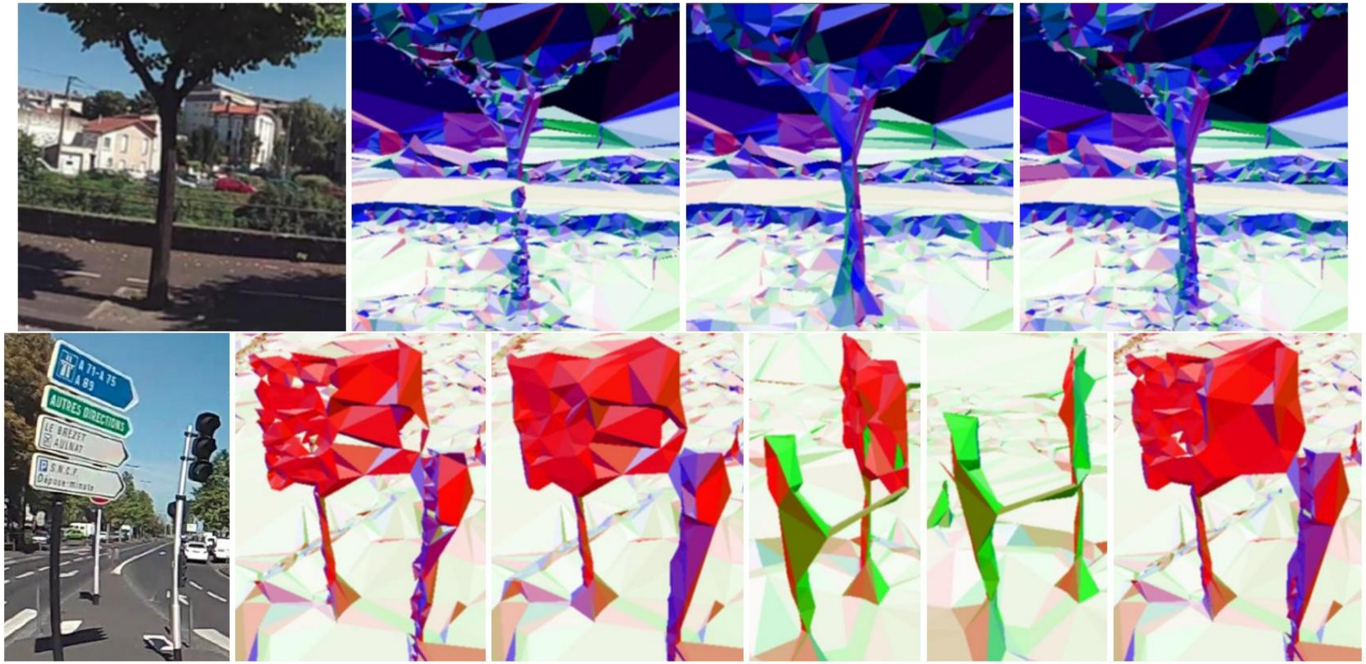


Fig. 4. Small scale views of our corrections. From left to right: input image, results of the correction methods $2+\tilde{2}$, $2+1$ and $2+3$ (best viewed in colors and by zooming in). Vertical triangles are red-green-blue and horizontal triangles are white/black.

Only use corrections (3/3)

So we numerically evaluate the topology changes.

Tab.1 shows that our corrections globally improve the topology: the number of holes and the percentage of non-manifold vertices in the surface decrease.

Method	ray-casting	$2+\tilde{2}$	2+1	2+3
comp. number (β_0)	4.8k	5.5k	1.5k	5.1k
tunnel number (β_1)	185k	105k	18k	31k
non-manifold vertex	31%	9.8%	5.0%	4.4%
freespace	57%	51%	38%	49%
computation time	102s	22s	38s	62s

Table 1. Topology changes provided by our corrections. See also the percentage of the freespace tetrahedra in T and the computation time (using one core of a standard laptop).

Use corrections with previous methods (1/3)

Now we compare the previous surface reconst. methods (M, G1, G2) and their corrected versions.

In Fig.5, 2+3 improves the surface normal of all methods in the lowest textured face of the building.

Here 2+1 has smaller effects than 2+3 since (1) Delaunay edges at the building are large and (2) the correction method 2+1 enforces a geometric size limit on the relabeled tetrahedra.

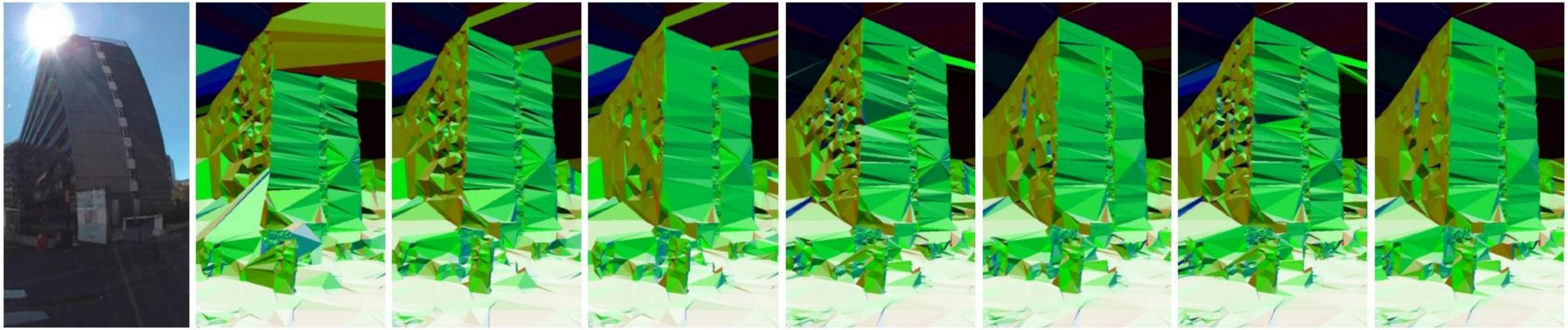


Fig. 5. Apply our corrections with previous methods to a building (best viewed in colors and by zooming in). From left to right: input image, results of methods M, M(2+1), M(2+3), G_1, $G_1(2+3)$, G_2, $G_2(2+3)$.

Use corrections with previous methods (2/3)

Fig.6 shows that our corrections can improve sharp edges if they are vertical and convex (an edge is convex if the matter side of the surface near the edge is convex)

But our corrections can oversmooth sharp edges if they are concave.

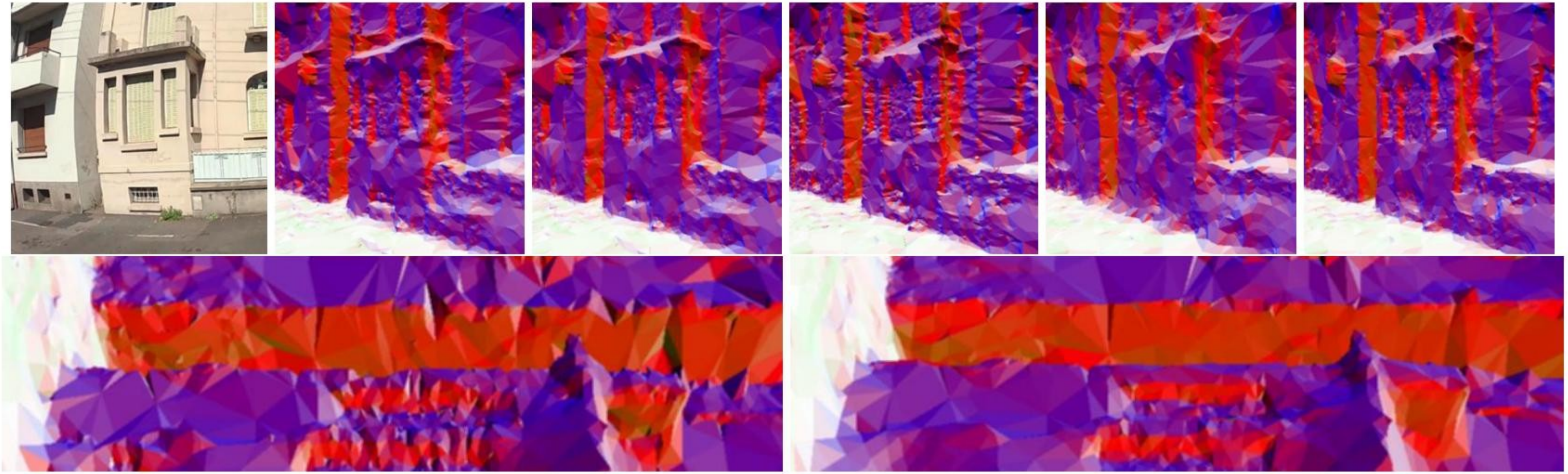


Fig. 6. Apply our corrections with previous methods to a facade (best viewed in colors and by zooming in). Top (from left to right): input image, M, M(2+3), G_1, G_1(2+1), G_1(2+3). Bottom: G_2 and G_2(2+3).

Use corrections with previous methods (3/3)

Tab.2 shows quantitative evaluations with a synthetic urban dataset (a piecewise planar scene with real texture).

The T vertices and lines-of-sight are estimated by structure-from-motion applied to synthetic videos taken by a 360 camera (whose trajectory is 621m long).

Method	G_1	$G_1(2+1)$	$G_1(2+3)$	G_2	$G_2(2+1)$	$G_2(2+3)$	M	M(2+1)	M(2+3)
80% fractile of error (cm)	28	26	26	28	26	26	24	22	22
90% fractile of error (cm)	94	84	88	94	82	88	68	58	56
triangle number	1.26M	0.86M	1.01M	1.27M	0.82M	0.98M	1.09M	0.72M	0.922M
component number (β_0)	46	28	45	27	18	29	2	3	4
tunnel number (β_1)	834	187	205	1737	247	351	32	17	18
non-manifold vertex	0.98%	1.24%	0.27%	1.58%	1.18%	0.34%	0%	0%	0%

Table 2. Quantitative improvements provided by our corrections of previous surface reconstruction methods applied to a synthetic urban dataset. There are not only accuracy improvements (error fractiles decrease) but also topology improvements (the number of spurious tunnels decreases) and mesh simplification. The true scene has $\beta_1 = 3$ tunnels and $\beta_0 = 1$ component.

Conclusion

Local-convexity is used to improve previous surface reconstruction methods based on the 3D Delaunay triangulation of the sparse input point cloud

We do not try to fit primitives but enumerate and select tetrahedra for relabeling

We reconstruct an urban scene from videos taken by a consumer 360 camera and observe improvements:

- freespace concavities are removed
- topological noise (ie the spurious tunnels in the matter) is reduced
- surface manifoldness (of graph-cut) is improved most of time
- sharp edges are improved if they are convex and vertical,
- thin structures are completed,
- both geometric error and surface complexity decrease

Limitations and future work

- Avoid oversmoothing at sharp edges if they are concave
- Vary the privileged direction along which we obtain most improvements
- Investigate other ways to reinforce local-convexity
- Last improve many steps in the complete 3D modeling process (texturing, loop closure, ...)