**A.S.I.S.T**

Automatic Semantically Invariant Scene Transformation

Or Litany, Tal Remez, Daniel Freedman, Lior Shapira, Alex Bronstein, Ran Gal

---

**OBJECTIVE**

- **scene**
- **model dataset**
- **transformed scene**

**Task:** Given a database of objects (from a known set of classes):
- Detect instances of objects within the scene
- Replace these instances with similar objects (exemplars)
- Replacement should respect: semantics (class), geometric resemblance

---

**FORMULATION**

\[
E\left(\{w_{cp}\}, \{T_{e}\}, \{v_{e}\}\right) = \sum_{i=1}^{6} \lambda_{i} E_{i}\left(\{w_{cp}\}, \{T_{e}\}, \{v_{e}\}\right)
\]

- **Point weight per exemplar**
- **Exemplar rigid transformation**
- **Exemplar vote**

- Semantic class estimation
- Geometric data Term
- Spatial smoothness
- Soft segmentation
- Non-collision Term
- Weight-vote coupling
- Constraints

---

**ALGORITHM**

**Input:** Point cloud, set of exemplars

**Output:** Subset of exemplars to insert, and their pose

**Initialization:**
- Random-Forest to get per class probability
- Mean-Shift to set initial exemplar positions

**Iterative Minimization - repeat:**
- Alternate between registration (ICP) and segmentation
- Fix \(\{w_{cp}\}, \{T_{e}\}\) and update votes \(\{v_{e}\}\)
- Increase collision penalty \(\lambda_{6}\)

---

**PIPELINE ILLUSTRATION**

![Pipeline Illustration](pipeline.png)

**RESULTS**

- **Our dataset (Kinect sensor)**
- **Comparison with Li et al.**
- **Our dataset (Tnago sensor)**
- **Comparison with Nan et al.**

---

**QUANTITATIVE EVALUATION**

- Automatically generated synthetic scenes
- Non-overlapping objects of class: bed, chair, sofa, table, and toilet.
- Criteria: semantic IoU and geometric IoU
- Notice the improvement with additional iterations (\(N_{out}\))

**Measure**

<table>
<thead>
<tr>
<th>Measure</th>
<th>(\text{Class})</th>
<th>bed</th>
<th>chair</th>
<th>sofa</th>
<th>table</th>
<th>toilet</th>
<th>geo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proc. ((N_{out} = 5))</td>
<td>0.98</td>
<td>0.96</td>
<td>1.00</td>
<td>0.91</td>
<td>0.97</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Proc. ((N_{out} = 1))</td>
<td>0.97</td>
<td>0.93</td>
<td>0.99</td>
<td>0.91</td>
<td>0.97</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Rec ((N_{out} = 5))</td>
<td>0.95</td>
<td>0.94</td>
<td>0.96</td>
<td>0.98</td>
<td>0.98</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>Rec ((N_{out} = 1))</td>
<td>0.83</td>
<td>0.88</td>
<td>0.97</td>
<td>0.78</td>
<td>0.88</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>(F_{1}) ((N_{out} = 5))</td>
<td>0.96</td>
<td>0.95</td>
<td>0.98</td>
<td>0.94</td>
<td>0.99</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>(F_{1}) ((N_{out} = 1))</td>
<td>0.91</td>
<td>0.91</td>
<td>0.95</td>
<td>0.84</td>
<td>0.92</td>
<td>0.94</td>
<td></td>
</tr>
</tbody>
</table>

---

**CONCLUSIONS**

- A unified approach for segmentation & replacement
- Simple mechanisms: Random Forest + ICP + QuadProgs
- Easily adjustable to new classifiers
- Main limitation is the restriction to rigid transformations.

---

**CONTACT INFORMATION:**

![QR Code](qr_code.png)