

## Evaluation by Single Feature Image Retrieval

We recently found a bug in the implemented evaluation code for Section 7.2. The figure below is the correct evaluation. **Note that this is only a bug fix, and there is no algorithm change.**

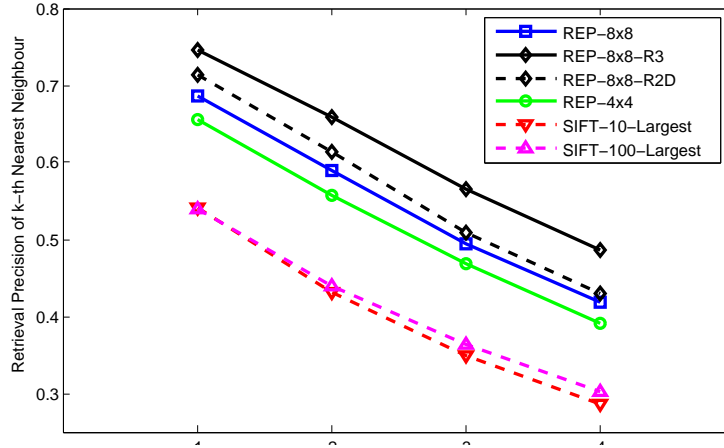


Figure 1: (Replaces Fig.7 in the submitted paper.) Evaluation by image retrieval. REP refers to our repeating elements. 8x8 and 4x4 refers to the grid size for feature descriptor. R3 refers to the elements that repeat at least 3 times. R2D refers to the features that belong some 2D repetition grids.

**Bug detail:** The sizes of some repeating elements were mistakenly set to a large incorrect value in the description stage. Orientation histogram of such a feature averages over a much larger region(may include outliers). However, they can still give decent retrieval results because the histograms from all repeating elements are similar.

After correcting the bug, all the descriptors are corresponding to the correct element regions as proposed in the paper. The figure in this document provides the correct evaluation.

**An important fact** about the retrieval failure cases in this experiment is that many are due the similar structures on different buildings, especially many buildings that have very similar window structures.