ALGORITHM 2: Pre-fixed Region Selection

Input: Source image I^{src} and source text T^{src}

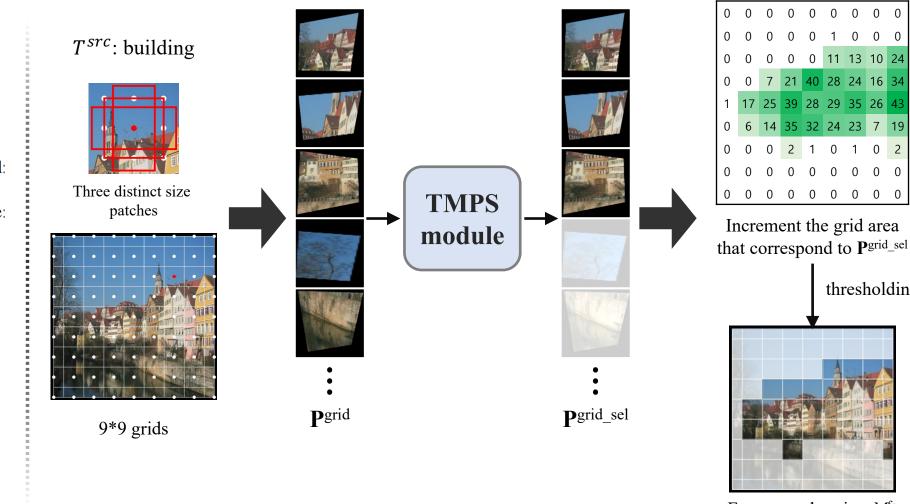
Output: A binary mask containing objects corresponding to the source text M^{fg}

Parameter: τ : Threshold for # of selections.

- 1: Divide I^{src} into L uniform square grids.
- 2: Generate a set of three distinct-sized patches per grid: $\mathbf{P}^{\mathrm{grid}} = \{P_1^{\mathrm{grid}}, \dots, P_{3L}^{\mathrm{grid}}\}.$
- 3: Obtain selected patches P^{grid_sel} using TMPS module: $\mathbf{P}^{\text{grid_sel}} = TMPS(\mathbf{P}^{\text{grid}}, T^{\text{src}}).$
- 4: Initialize a voting matrix $V \in \mathbb{R}^{H \times W}$ with all elements set to zero.
- 5: for each pixel in the selected patches $\mathbf{P}^{\text{grid_sel}}$ do
- Increment the corresponding element in V.
- 7: end for
- 8: Determine the pre-fixed foreground region M^{fg} :

$$M^{\mathrm{fg}}(i,j) = \begin{cases} 1, & \text{if } V(i,j) \ge \tau \\ 0, & \text{otherwise} \end{cases}$$

9: return M^{fg}



Foreground region M^{fg}

thresholding

0 0 0 1 0 0 0