WHERE TO PLACE: A REAL-TIME VISUAL SALIENCY BASED LABEL PLACEMENT FOR AUGMENTED REALITY APPLICATIONS

Neel Rakholia (Stanford University), Srinidhi Hegde (TCS Research)
Ramya Hebbalaguppe (TCS Research)

Outline

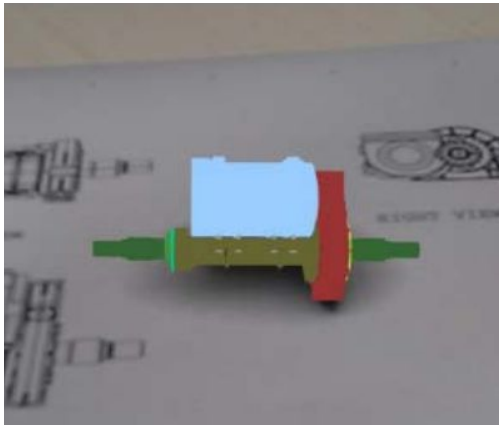
1. Introduction to label placement in AR and its importance
2. Previous works
3. Our proposed Method: Where2Place
4. Evaluation
5. Live demo of single label placement on phone
6. Live demo of multi-label placement in cluttered scene
7. Conclusion

Outline

1. **Introduction to label placement in AR and its importance**
2. Previous works
3. Our proposed Method: Where2Place
4. Evaluation
5. Live demo of single label placement on phone
6. Live demo of multi-label placement in cluttered scene
7. Conclusion

Introduction to label placement in AR and its importance

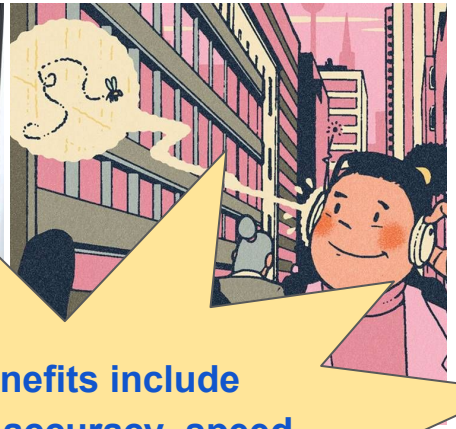


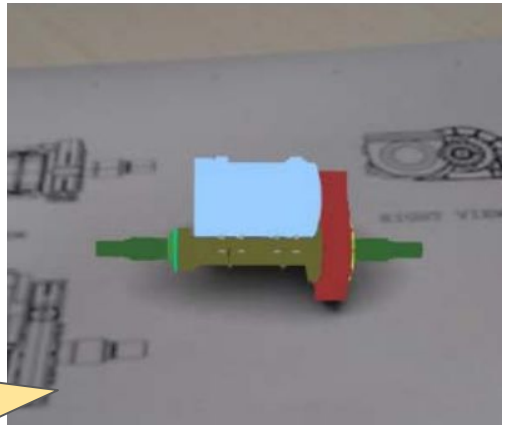


Label Placement in Augmented Reality

- Augmented Reality is a metaphor for situated Computing.
- Augmentations can take the form of Text/label, image, sound, and Video.
- Textual overlays/labels add contextual information in Augmented Reality (AR) applications.

Images from internet





Label Placement in Augmented Reality

Key benefits include improved accuracy, speed and lowering one's cognitive load

situated

- Aug tations n take the form of Text/label, imag , sound, and Video.
- Textual overlays/labels add contextual information in Augmented Reality (AR) applications.

Images from internet

Why label placement is important ?

Reproduced from Grasset et al. -“Image-driven view management for augmented reality browsers,” in Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)

Objectives for label placement

Label should not occlude the salient regions

Label should be closed to object of interest

Labels should not overlap each other

Labels should work in real-time and temporally coherent on device

Outline

1. Introduction to label placement in AR and its importance
2. **Previous works**
3. Our proposed Method: Where2Place
4. Evaluation
5. Live demo of single label placement on phone
6. Live demo of multi-label placement in cluttered scene
7. Conclusion

Related Works

- Related studies of optimal placement of textual labels based on – **Geometry based layout** and **Image-based layouts** for rendering labels, aesthetic rule, and adaptive overlays.

- These approaches work on images - not suited for real-time camera streams. These are **computationally heavy** and **lack real-time performance**.

- Owing to occlusions, dim light scenarios, scene variations in the live field of view, overlays have their own challenges.

Outline

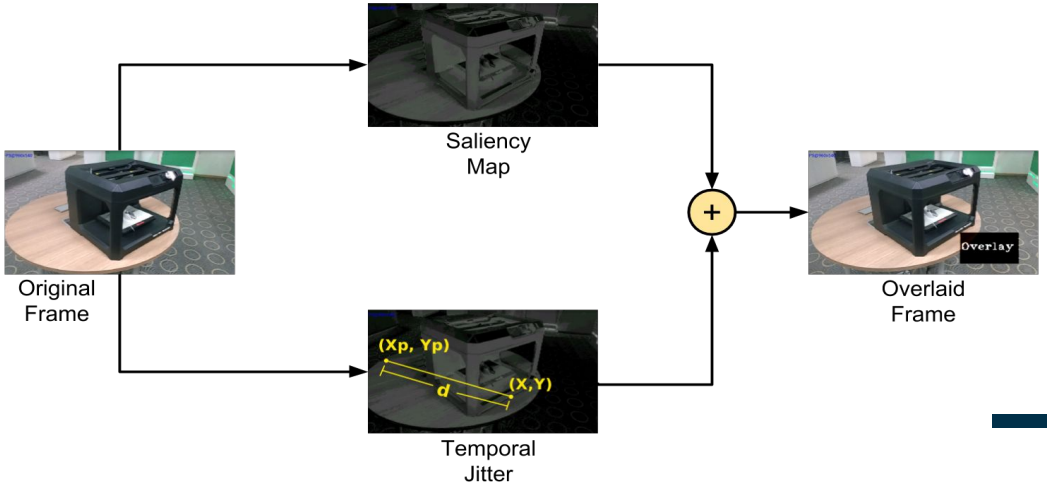
1. Introduction to label placement in AR and its importance
2. Previous works
3. **Our proposed Method: Where2Place**
4. Evaluation
5. Live Demo of single label placement on phone
6. Live demo of multi-label placement in cluttered scene
7. Conclusion

First attempt towards single label placement

1. **Label should be placed at real-time** - applications are scene summarisation, effective captioning of a scene, AR applications
2. **Label should be placed on non-salient areas -** interesting areas should not be occluded by the label
3. **Label should be temporally coherent** - facilitating in scene understanding than causing jitter

Therefore, we formulate this method by an objective function that minimizes -

- **Occlusion** with visually salient regions in scenes of interest
- **Temporal jitter** for facilitating coherence in real-time AR applications



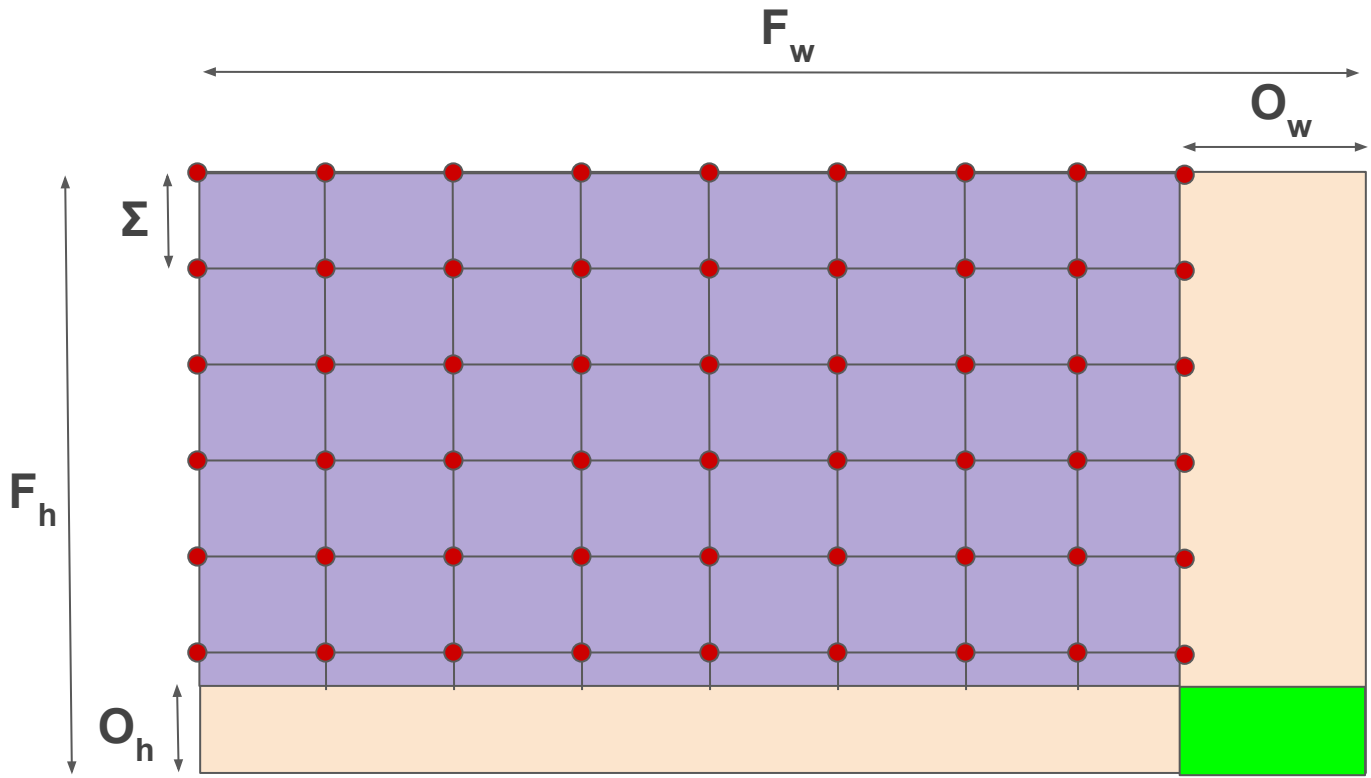
Algorithm Overview

- Input to our algorithm - video frames
- First leg - compute saliency maps[1]
- Second leg - compute Euclidean distance between previous and current overlay positions
- Combine saliency and Euclidean distance - predict overlay position

Optimization problem to select optimal label position -

$$\begin{aligned} \underset{(X,Y)}{\text{minimize}} \quad & S((X,Y)) + \lambda d((X,Y);(X_p,Y_p)) \\ \text{subject to} \quad & X \leq F_w - O_w, \quad X \geq 0 \\ & Y \leq F_h - O_h, \quad Y \geq 0 \end{aligned}$$

Radhakrishna Achanta, Sheila Hemami, Francisco Estrada, and Sabine Susstrunk, “Frequency-tuned salient region detection,” in Computer vision and pattern recognition, 2009. cvpr 2009. ieee conference on. IEEE, 2009, pp. 1597–1604.



Legend:

- K : Frames to skip
- λ : Temporal Coherence (Jitter/Threshold) Parameter
- Σ :Search space sampling parameter
- O_h, O_w : Overlay height and overlay width
- X_p, Y_p : Optimal position of overlay
- X, Y : Optimal position of overlay computed in the current iteration
- SM : Computed saliency map
- P : set of sampled pixels from search space
- F_w, F_h : Width and height of video frames

Algorithm 1 PROPOSED ALGORITHM

$(X_p, Y_p) = (frame\_width/2, frame\_height/2)$
For every k^{th} frame
    $SM = AchantaSal(frame)$
    for $(x, y) \in P$
        $L = \{(a, b) | x \leq a \leq x + O_w, y \leq b \leq y + O_h\}$
        $s_{x,y} = \sum_{(a,b) \in L} SM(a, b)$
        $d_{x,y} = \lambda \times Distance((X, Y), (X_p, Y_p))$
    $s_{min} = min(s_{x,y} + d_{x,y})$
    $(X, Y) := arg\_min(s_{x,y})$
    $(X_p, Y_p) := (X, Y)$ // Use linear interpolation for overlay transition

Legend:

- K : Frames to skip
- λ : Temporal Coherence (Jitter/Threshold) Parameter
- Σ :Search space sampling parameter
- O_h, O_w : Overlay height and overlay width
- X_p, Y_p : Optimal position of overlay
- X, Y : Optimal position of overlay computed in the current iteration
- SM : Computed saliency map
- P : set of sampled pixels from search space
- F_w, F_h : Width and height of video frames

Outline

1. Introduction to label placement in AR and its importance
2. Previous works
3. Our proposed Method: Where2Place
4. **Evaluation**
5. Live Demo of single label placement on phone
6. Live demo of multi-label placement in cluttered scene
7. Conclusion

Evaluation

Objective Evaluation

Label Occlusion Over Saliency (S)

$$S = \frac{\sum_{(x,y)\in L} G(x,y)}{\| L \|}$$

Our algorithm had an average LOS score of 0.042 and it took 0.021s of time to compute the overlay position.

Subjective Evaluation

Subjective metrics	Value (0-5)
Position of Overlay	4.5
Responsiveness of Overlay box	4.7
Lack of Jitter	4.2
Color of Overlay box	3.9

Outline

1. Introduction to label placement in AR and its importance
2. Previous works
3. Our proposed Method: Where2Place
4. Evaluation
5. **Live demo of single label placement on phone**
6. Live demo of multi-label placement in cluttered scene
7. Conclusion

Where2Place: Demo 1 on phone

Where2Place: Demo 2 on phone

Outline

1. Introduction to label placement in AR and its importance
2. Previous works
3. Our proposed Method: Where2Place
4. Evaluation
5. Live demo of single label placement on phone
6. **Live demo of multi-label placement in cluttered scene**
7. Conclusion

Version 2 Demo: Multilabel placement on server

Outline

1. Introduction to label placement in AR and its importance
2. Previous works
3. Our proposed Method: Where2Place
4. Evaluation
5. Live demo of single label placement on phone
6. Live demo of multi-label placement in cluttered scene
7. **Conclusion**

Conclusions

We have presented a novel method for placement of overlays on device (resource constrained environment)

We formulated it as objective function that minimizes visual saliency around the object of interest and minimises the temporal jitter facilitating coherence in real-time AR applications.

The main focus - algorithm should run in real-time on a low-end android phone/tablet

Applications include: situational awareness for museum exploratory tasks, industrial inspection and repair operations, advertisement and media, and in tourism industry.

Questions?

Contact:

sri.hegde@tcs.com

neel.rakholia@stanford.edu

ramya.hebbalaguppe@tcs.com


Thank You