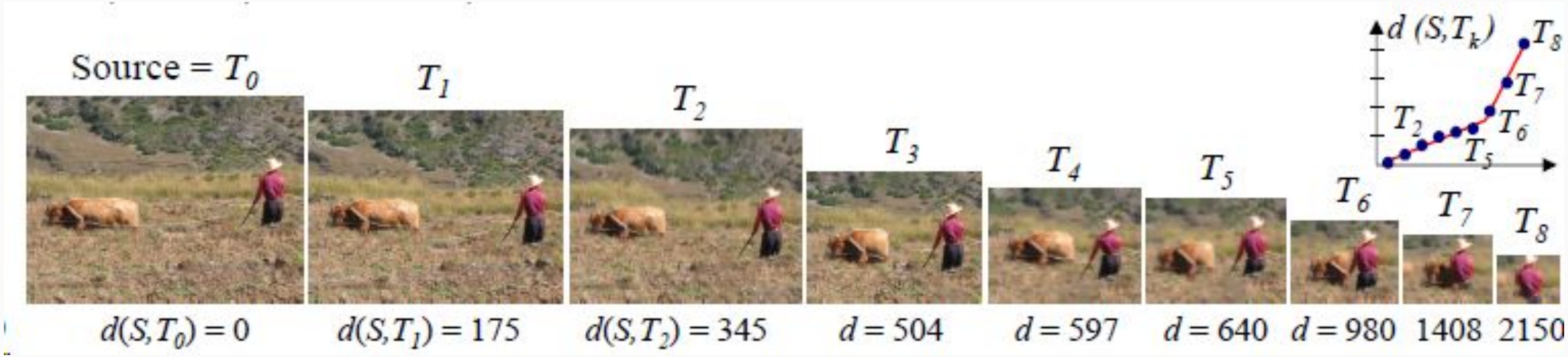
Summarizing Visual Data Using Bidirectional Similarity

By
Hema
Sailaja

Visual Summary

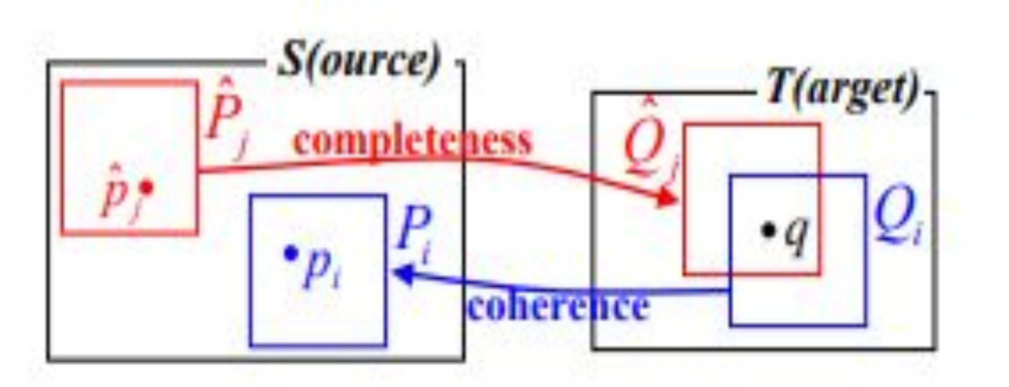
Goal:

Produce a smaller image that summarizes the content of the larger image

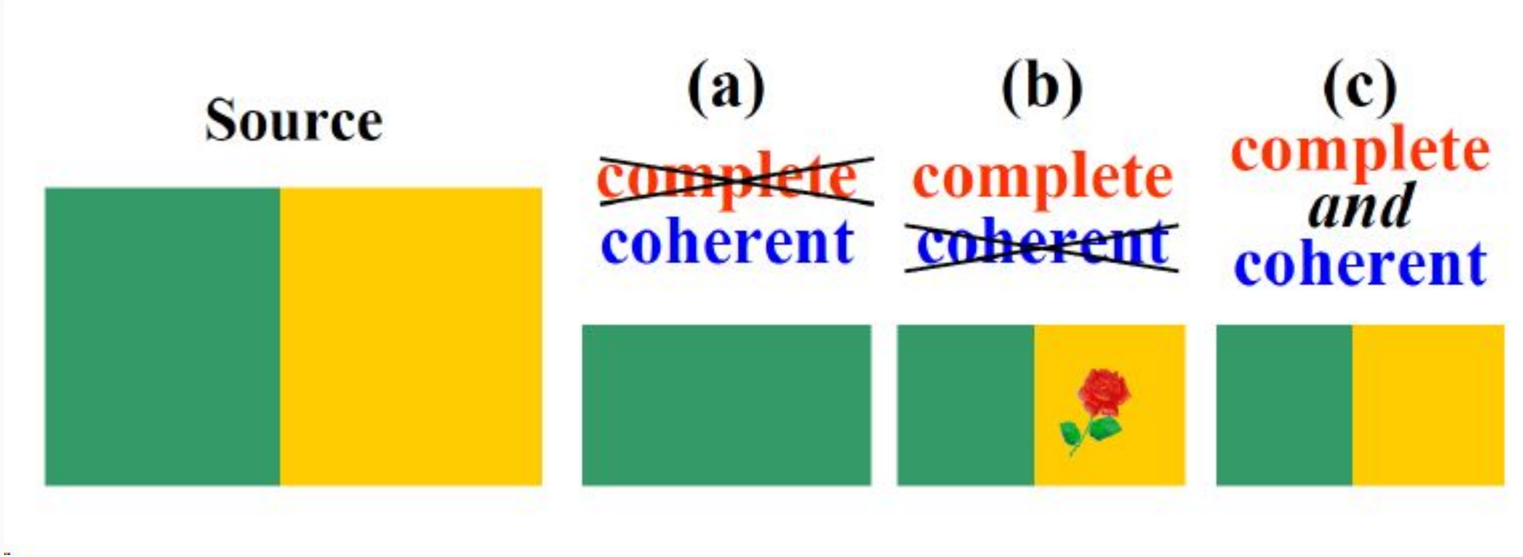


Comparing two images

- Completeness
- Coherence



Similarity Distance



Bidirectional similarity measure

- Error or dissimilarity measure

$$d(S,T) = \overbrace{\frac{1}{N_S} \sum_{P \subset S} \min_{Q \subset T} D(P,Q)}^{d_{\text{complete}}(S,T)} + \overbrace{\frac{1}{N_T} \sum_{Q \subset T} \min_{P \subset S} D(Q,P)}^{d_{\text{cohere}}(S,T)}$$

Where,
S and T are the source and Target images, P and Q are the patches of fixed size of S and T.
D is the Sum of Squared Difference between the patches.

The Summarization (Retargeting) Algorithm

The Iterative Update rule: contribution of a pixel to the coherence measure

- Let q be a pixel of T, q lies inside m neighboring patches $Q_1, Q_2, \ldots Q_m$
- These patches are matched to $P_1, P_2, \ldots P_m$ in source image S
- The positions corresponding to q in $P_1, P_2, \ldots P_m$ are $p_1, p_2, \ldots p_m$

Hence, the contribution is

$$\frac{1}{N_T} \sum_{i=1}^{m} \|S(p_i) - T(q)\|^2$$

The Iterative Update rule: contribution of a pixel to the completeness measure

- Let q be a pixel of T,
- q lies inside n neighboring patches $\hat{Q}_1, \hat{Q}_2, \cdots, \hat{Q}_n$ that are the nearest patch to some patches of S $\hat{P}_1, \hat{P}_2, \cdots \hat{P}_n$
- The positions corresponding to q in $\hat{P}_1, \hat{P}_2, \cdots, \hat{P}_m$ are $\hat{p}_1, \cdots, \hat{p}_m$

Hence, the contribution is

$$\frac{1}{N_S} \sum_{i=1}^{n} \|S(\hat{p}_i) - T(q)\|^2$$

Color Update

The best T(q) should minimise

$$\frac{1}{N_S}\sum_{j=1}^{n}(S(\hat{p}_j) - T(q))^2 + \frac{1}{N_T}\sum_{i=1}^{m}(S(p_i) - T(q))^2$$

Color Update:

$$T(q) = \frac{\frac{1}{N_S}\sum_{j=1}^{n} S(\hat{p}_j) + \frac{1}{N_T}\sum_{i=1}^{m} S(p_i)}{\frac{n}{N_S} + \frac{m}{N_T}}$$

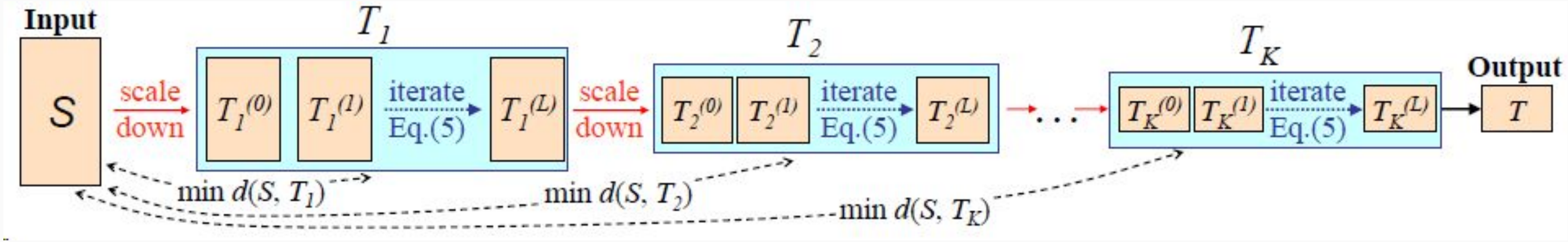
Iterative Update rule

Given a source signal S, we want to reconstruct a target signal T that optimizes the similarity measure

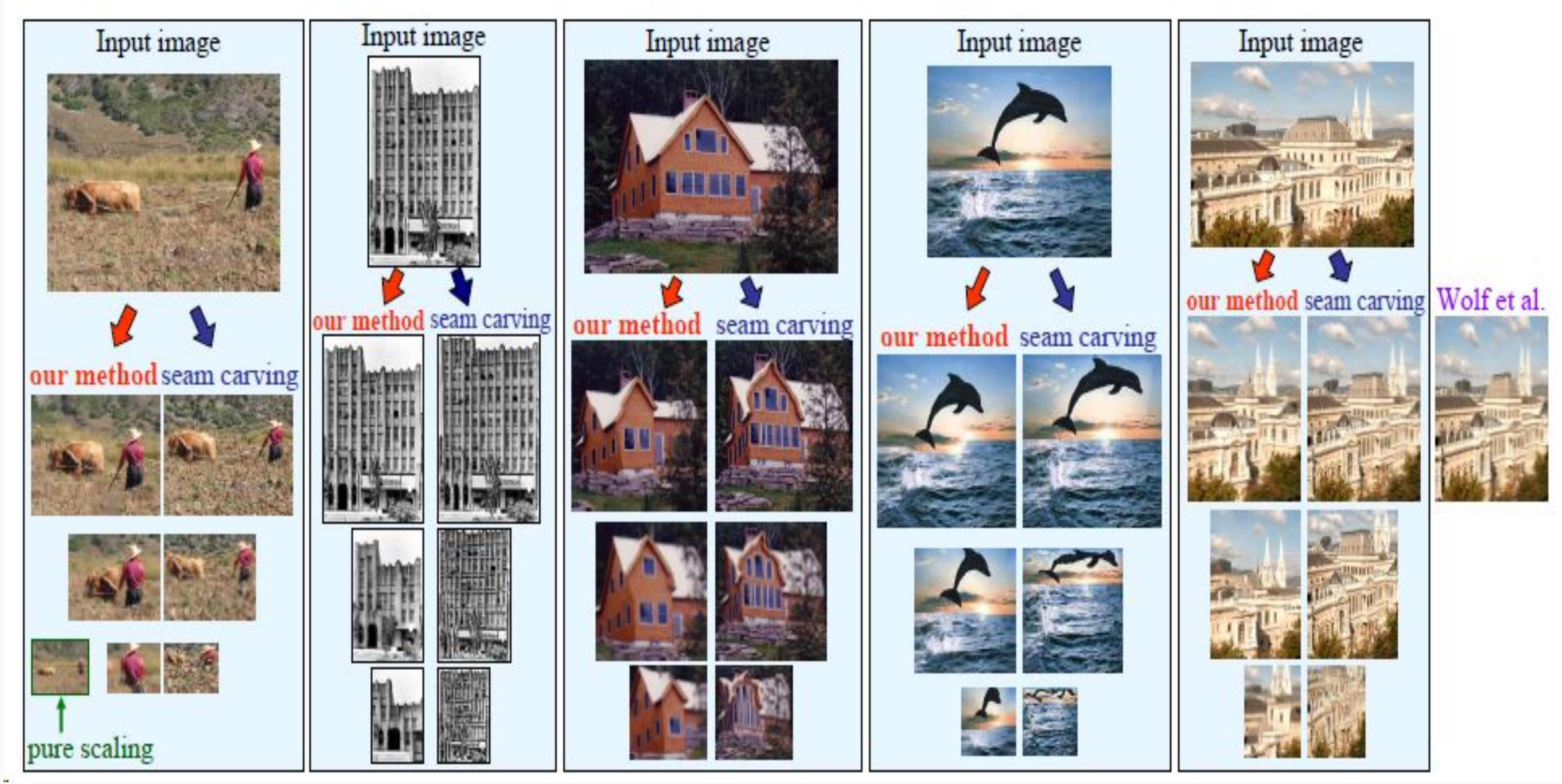
$$T_{output} = \arg\min_{T} d(S, T).$$

Gradual Resizing

- When the target has a very different size from the source: what is a good initial guess?
- Iterative process: downsample the image and apply the reconstruction



Visual Summary



Applications

Image Montage

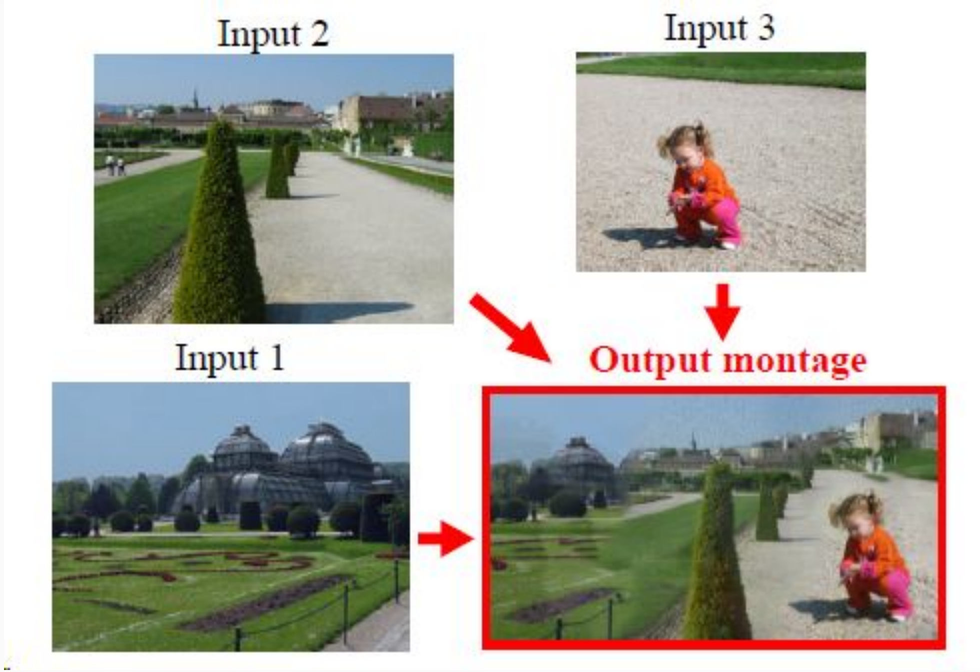
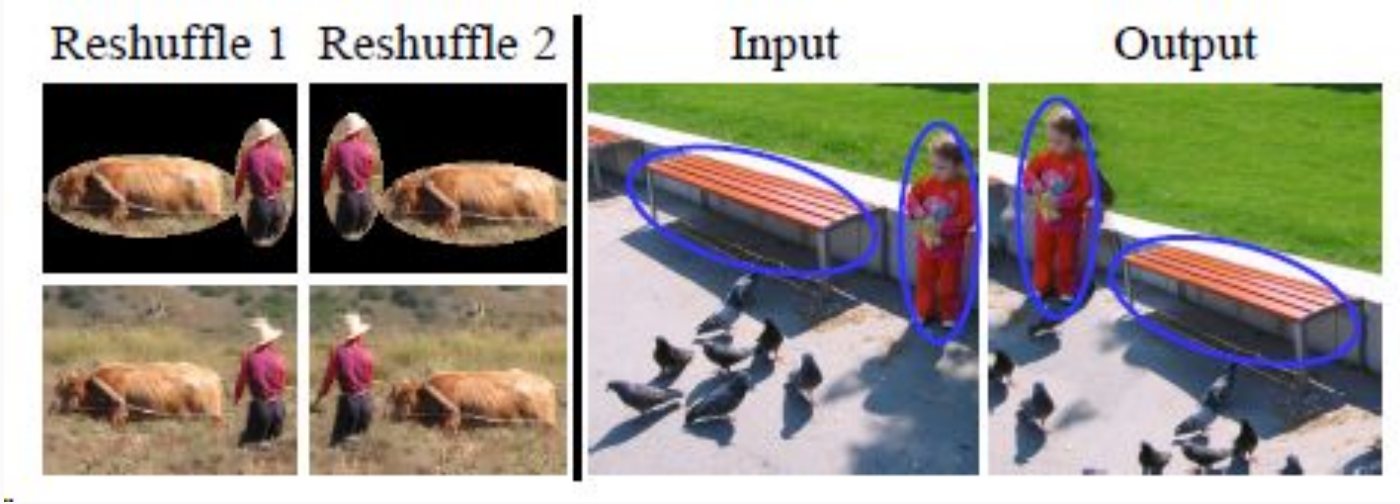
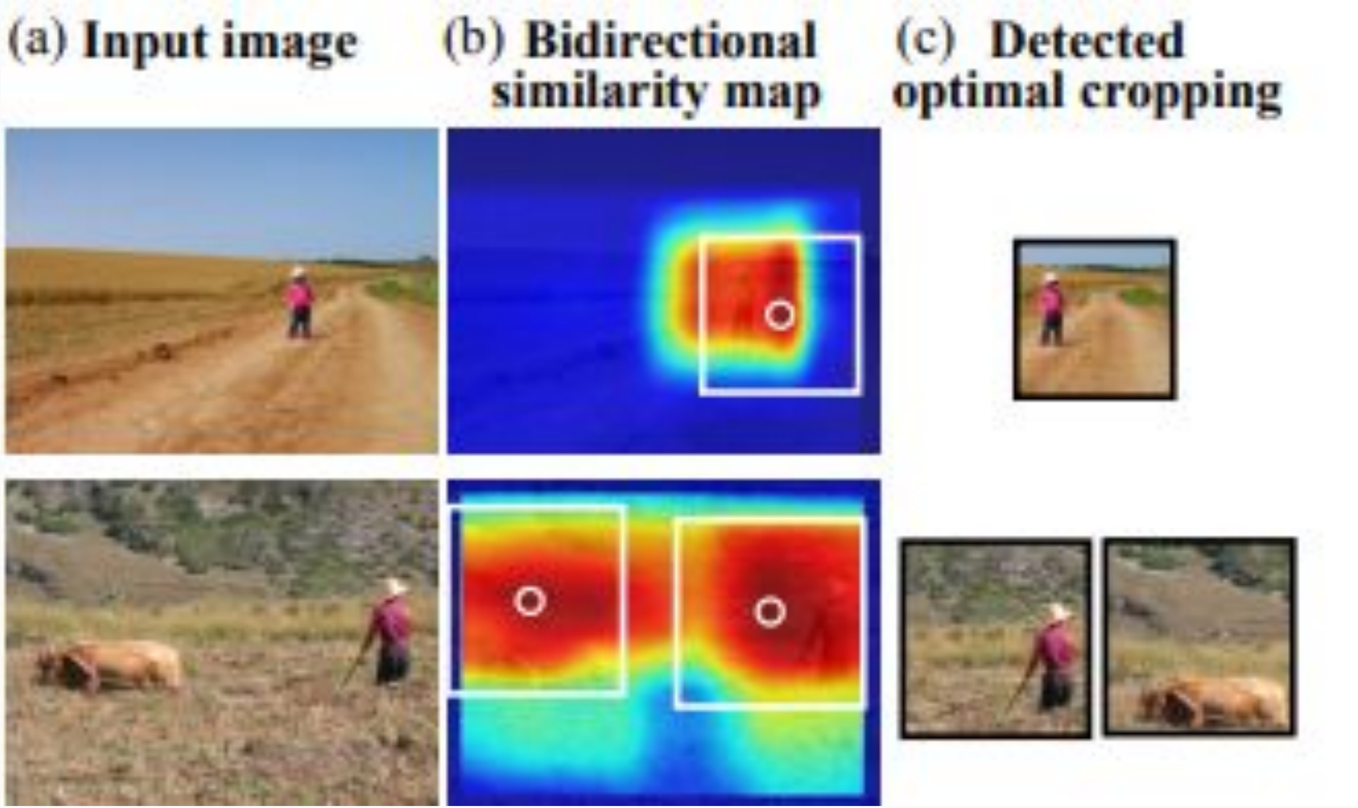


Image Completion and Synthesis

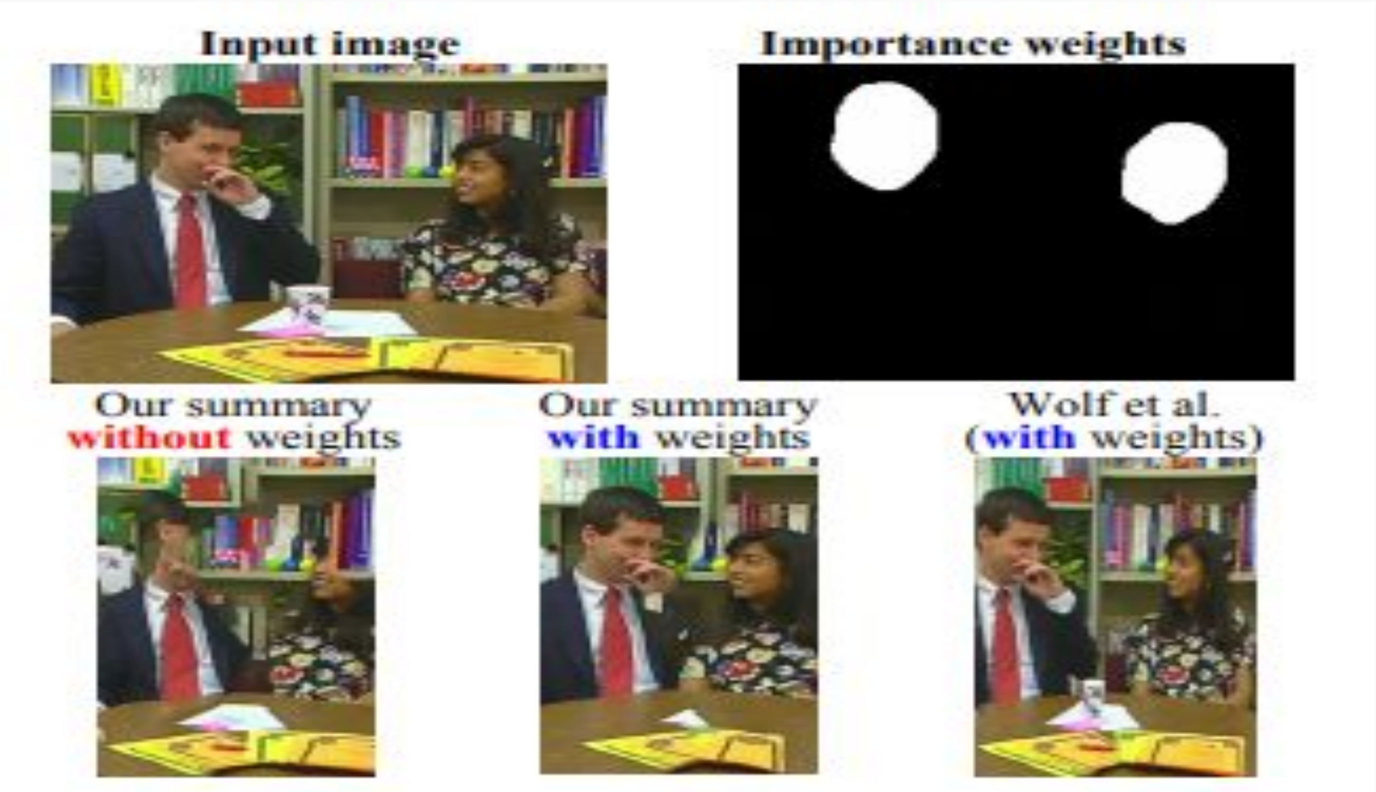
Photo Reshuffling



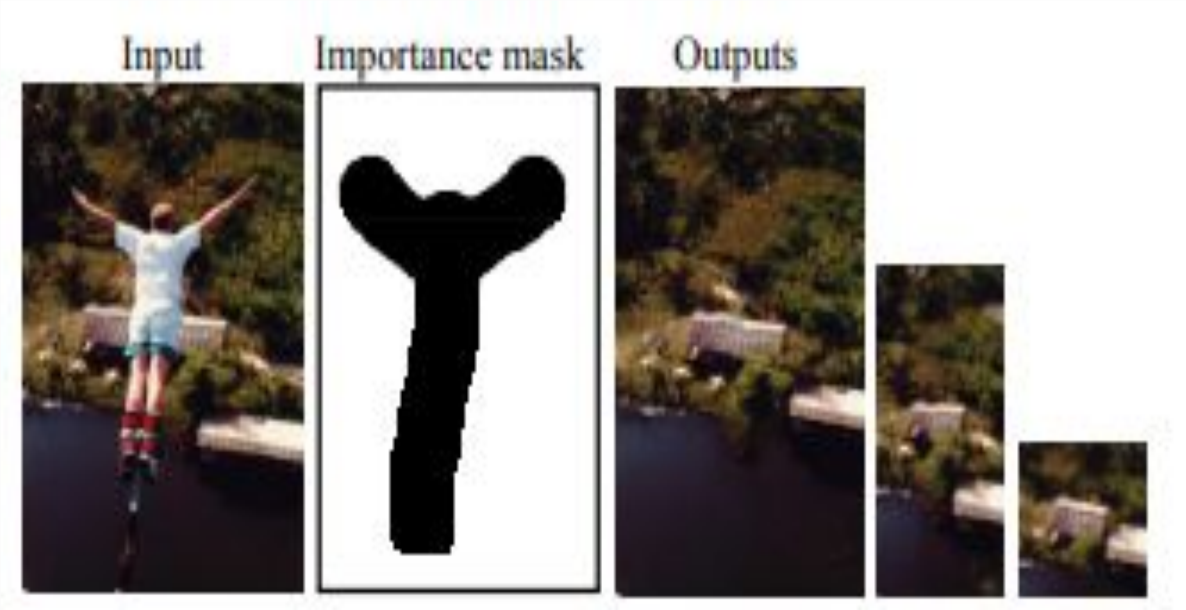
Automated Optimal Cropping



Incorporating Non Uniform Importance



Summarization with object removal constraints



Thank you