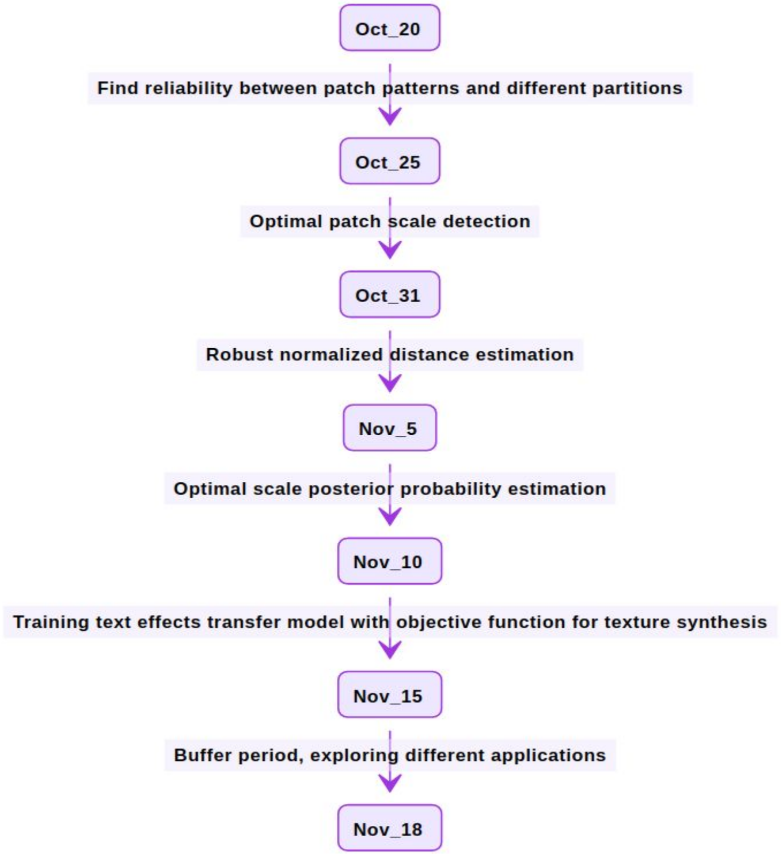
DIP project
Mid Evaluation

255 Shades of Gray

Team members :
Nihar Potturu (2018102039)
Abhiram Bondada (2018102036)
Tushar Patra (2018102041)
Samartha SM (2018101094)

Brief Timeline



Optimal Patch Scale Detection

The algorithm detects the optimal patch scale scal(q) to depict texture patterns around pixel q.

A fixed patch of size mxm is used to obtain the optimal patch scale.

The source image S is downsampled with scale rate of $1/s^{l-1}$ where s is downsampling factor and l is scale to get downsampled source S_l. Now $Q_l(q)$ is the patch centered at q/s^{l-1} in S_l. Similarly, S'_l and $Q'_l(q)$ are downsampled stylized source S' and patch centered at q/s^{l-1} in S'_l.

$\hat{q}$ is computed using the objective function as follows,

$$\hat{q} = \arg\min \|Q_\ell(q) - Q_\ell(\hat{q})\|^2 + \|Q'_\ell(q) - Q'_\ell(\hat{q})\|^2$$

Filter criterion at scale l is, $\zeta_\ell(q, \hat{q}) = \left(\sigma_\ell + \sqrt{d_\ell(q, \hat{q})} > \omega\right)$

where, $\sigma_l = \sqrt{Var(Q'_\ell(q))}/2$

If the above criterion is true, then it means the the variance is larger than required and hence we need a finer scale. Hence, we iterate patch scale from max patch scale L to 1 to find a patch scale l that gives the above criterion as false.

Algorithm Optimal Patch Scale Detection

Input: Image S, S', parameters L, s, ω

Output: Optimal scale $\text{scal}(q)$ for each pixel q

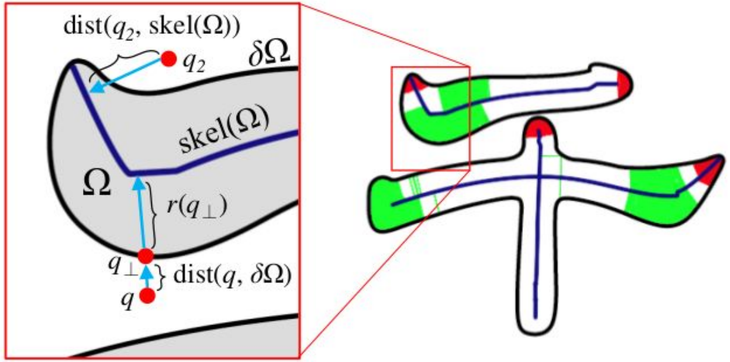
Initialize $R = \{q | q \in S\}$ and $\text{scal}(q) = 1, \forall q \in R$
for $\ell = L, ..., 2$ **do**
    for all $p \in R$ **do**
        Compute $\hat{q} = \arg\min_{\hat{q}} d_\ell(q, \hat{q})$
        if $\zeta_\ell(q, \hat{q})$ is false **then**
            $\text{scal}(q) = \ell$
            $R = R \setminus \{q\}$
        end if
    end for
end for

Robust Normalized Distance Estimation

We know that, there is good correlation between distance of pixel from the skeleton and the pattern to of the text.

So we find the distance transform from the skeleton to classify the pixel into different patch levels. But there are two problems with this :

1. Method is not invariant to text radius
2. Some sharp edges can cause skeleton to be touching the boundary instead of staying in the interiors (as shown in the image - red region)



So, the paper solves the problem in the as discussed in next slide

We are trying to normalize the distance to make it invariant with the width. Simply dividing the distance by the text width is unreliable because the inaccuracy of the obtained $skel(\Omega)$ leads to errors both in the numerator and denominator as well.

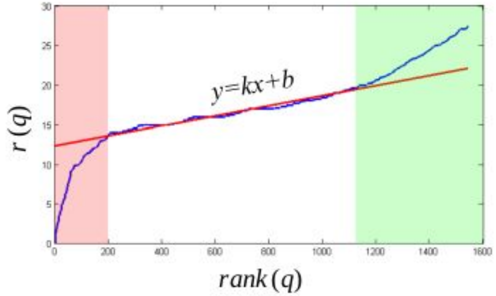
Ω-> Text Region

$\delta\Omega$ -> Contour/Boundary

$|\delta\Omega|$ -> Contour/Boundary

$r(q)$ -> Text width / Shortest distance from point on contour to skeleton ($q \in contour$)

$\tilde{r}(q)$ -> Corrected text width (obtained by removing outliers using regression coefficients) ($q \in contour$)



(d) Linear relation of $r(q)$ and $rank(q)$

We can observe that the outliers assemble at small values.

Formulas

$\tilde{r}(q) = max(dist(q, skel(\Omega)), 0.2k|\delta\Omega| + b)$

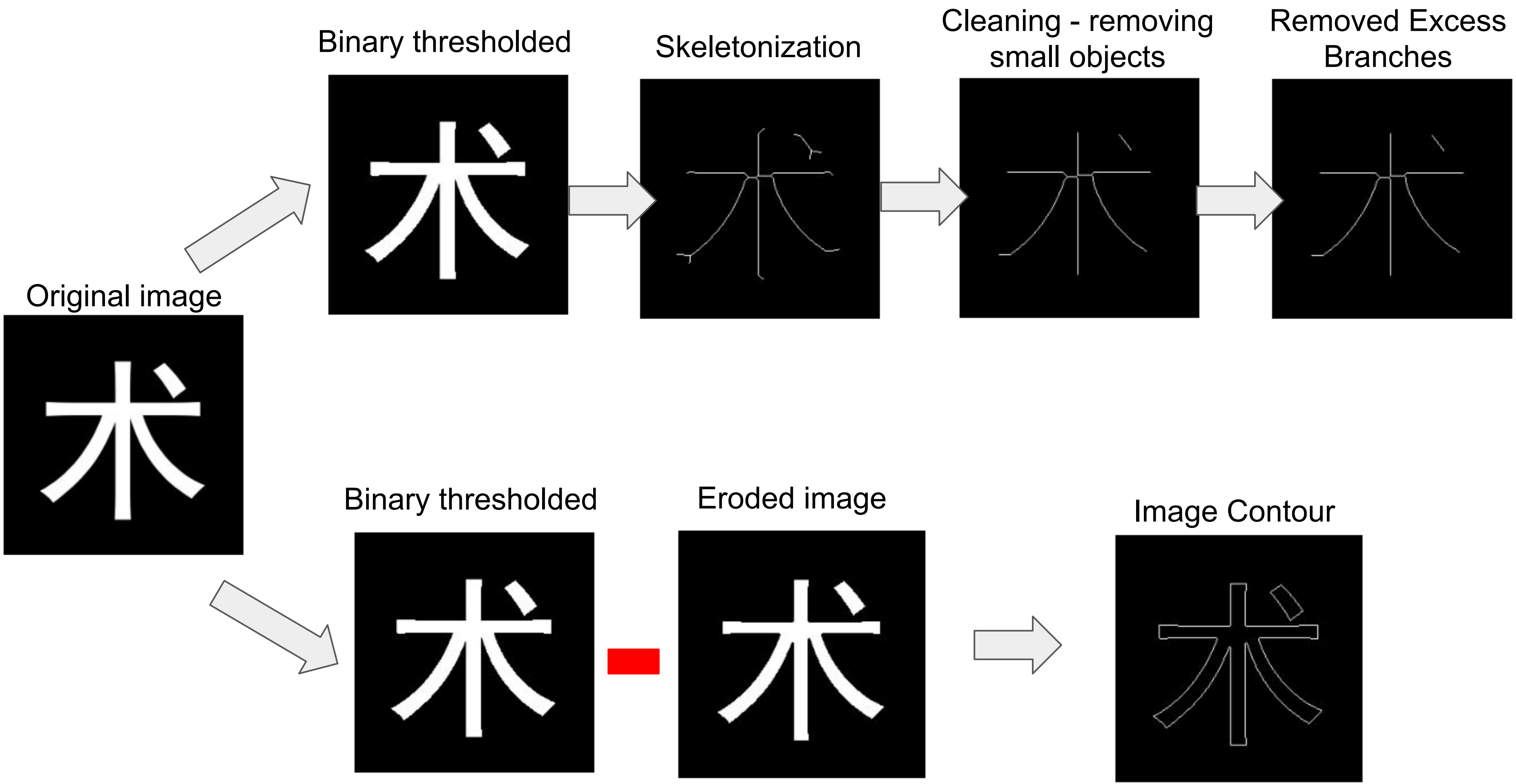
This eliminates the outliers as they assemble at small values and we assume them to be the lefmost 20% points.

$0.2k|\delta\Omega| + b \geq$ dist (q,skel(\Omega)) for q $\in$ outliers.

We calculate the normalized distance as follows :

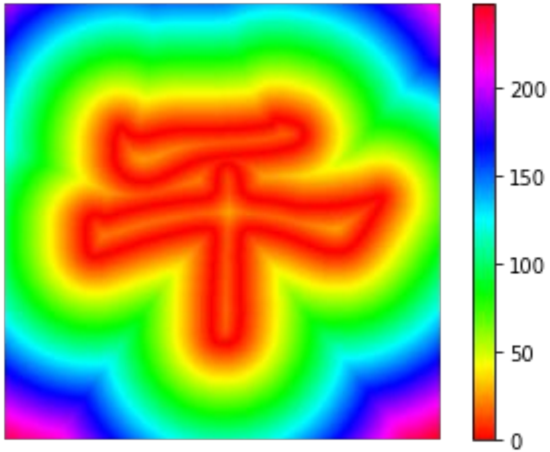
$$\tilde{dist}(q, skel(\Omega)) = \begin{cases} 1 + dist(q, skel(\Omega))/\bar{r}, & \text{if } q \notin \Omega \\ 1 - dist(q, skel(\Omega))/\tilde{r}(q_{\perp}), & \text{otherwise} \end{cases}$$

Here , $q_{\perp} \in \delta\Omega$ is the nearest pixel to q along $\delta\Omega$ and $\bar{r} = 0.5k|\delta\Omega| + b$ is the mean text width

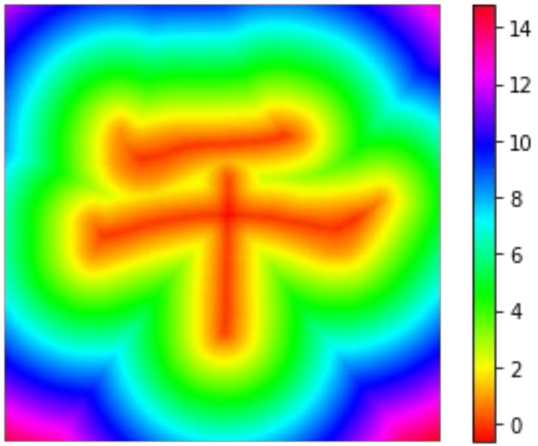

Original image
Binary thresholded
Skeletonization
Cleaning - removing small objects
Removed Excess Branches
Binary thresholded
Eroded image
Image Contour

Distance transforms - comparison

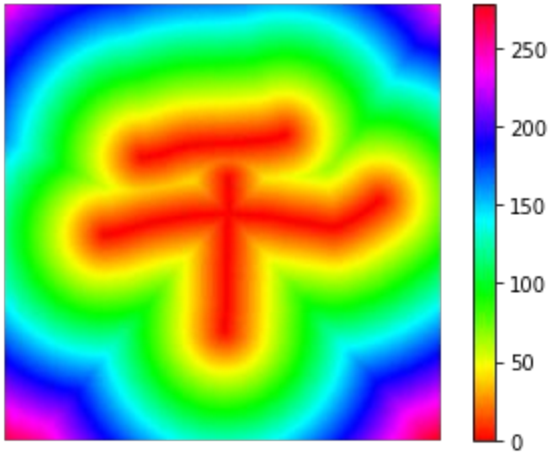
Distance transform on perimeter



Modified method Distance transform using normalisation and removal of outliers (as discussed in the paper)



Distance transform on skeleton



Optimal Scale Posterior Probability Estimation

We utilize the property of high correlation between patch patterns and their spatial distributions to derive posterior probability of optimal patch scale.

We quantify all distances into 100 bins and hence, bin(q) represents which bin q belongs to.

We construct 2D histogram hist where hist(l, x) represents number of pixels which have optimal scale as l and belongs to bin x.

$$hist(\ell, x) = \sum_q \psi(\text{scal}(q) = \ell \wedge \text{bin}(q) = x)$$

where, ψ(.) is 1 when the argument is true, else false. We get joint probability P(l, x) after normalizing the histogram.

$$\mathcal{P}(\ell, x) = hist(\ell, x) / \sum_{\ell, x} hist(\ell, x)$$

From this joint probability, we can find the posterior probability P(l|bin(q)) which represents probability of optimal scale being l when the patch has distance corresponding to bin(q).

$$\mathcal{P}(\ell|\text{bin}(p)) = \mathcal{P}(\ell, \text{bin}(p)) / \sum_{\ell} \mathcal{P}(\ell, \text{bin}(p))$$

Here, it is assumed that target images share the same posterior probability with the source image.