On the Effectiveness of Visible Watermarks Supplemental

1 Image/Watermark Decomp.

We derive the solution to the optimization problem in Eq. 11 in the continuous domain, i.e., replacing the sum by an integral. In this case, the optimal solution must satisfy the *Euler-Lagrange* equations, given by

$$\frac{\partial L}{\partial I^k(\mathbf{p})} - \frac{\partial}{\partial x}\frac{\partial L}{\partial(I_x^k(\mathbf{p}))} - \frac{\partial}{\partial y}\frac{\partial L}{\partial(I_y^k(\mathbf{p}))} \tag{1}$$

$$\frac{\partial L}{\partial W^k(\mathbf{p})} - \frac{\partial}{\partial x}\frac{\partial L}{\partial(W_x^k(\mathbf{p}))} - \frac{\partial}{\partial y}\frac{\partial L}{\partial(W_y^k(\mathbf{p}))} \tag{2}$$

for all pixel locations $\mathbf{p}$, where L is the integrand in Eq. 11. That is,

$$\begin{aligned} L = \quad & L_{\text{data}}(I_k, W_k, \alpha) + \lambda_I L_{\text{reg}}(\nabla I_k) + \lambda_w L_{\text{reg}}(\nabla W_k) + \\ & \lambda_\alpha L_{\text{reg}}(\nabla\alpha) + \beta L_f(\nabla(\alpha W_k)) + \gamma L_{\text{aux}}(W, W_k), \end{aligned} \tag{3}$$

where L_x is the corresponding integrand to term E_x in the paper (i.e., the expression inside the sum). Keeping only the relevant terms in L, Eq. (1-2) are given by:

$$\frac{\partial L_{\text{data}}}{\partial I^k(\mathbf{p})} - \lambda_I \frac{\partial}{\partial x}\frac{\partial L_{\text{reg}}(\nabla I_k)}{\partial(I_x^k(\mathbf{p}))} - \lambda_I \frac{\partial}{\partial y}\frac{\partial L_{\text{reg}}(\nabla I_k)}{\partial(I_y^k(\mathbf{p}))} \tag{4}$$

$$\frac{\partial(L_{\text{data}} + \gamma L_{\text{aux}})}{\partial W^k(\mathbf{p})} - \frac{\partial}{\partial x}\frac{\partial(\beta L_f + \lambda_w L_{\text{reg}})}{\partial(W_x^k(\mathbf{p}))} - \frac{\partial}{\partial y}\frac{\partial(\beta L_f + \lambda_w L_{\text{reg}})}{\partial(W_y^k(\mathbf{p}))} \tag{5}$$

Let $\boldsymbol{\alpha} = \text{diag}(\alpha)$, $\bar{\boldsymbol{\alpha}} = \text{diag}(1-\alpha)$ be diagonal matrices, where α and $1-\alpha$ are the diagonals, respectively. We define the following notations:

$$\boldsymbol{\Psi'_{\text{data}}} = \text{diag}(\Psi'((\alpha W^k + (1-\alpha)I_k - J_k)^2))$$

$$\boldsymbol{\Psi'_{\mathbf{w}}} = \text{diag}(\Psi'((|\alpha_x|W_x^k + |\alpha_y|W_y^k)^2))$$

$$\boldsymbol{\Psi'_{\mathbf{I}}} = \text{diag}(\Psi'((|\alpha_x|I_x^k + |\alpha_y|I_y^k)^2))$$

$$\boldsymbol{\Psi'_{f}} = \text{diag}(\Psi'(\|\nabla(\alpha W^k) - \nabla W_m)\|^2)$$

$$\boldsymbol{\Psi'_{\text{aux}}} = \text{diag}(\Psi'((W^k - W)^2))$$

$$\mathbf{\Psi'rI} = \mathrm{diag}(\Psi'((|\alpha_x|I_x^2 + |\alpha_y|I_y^2))$$

$$\mathbf{\Psi'rw} = \mathrm{diag}(\Psi'((|\alpha_x|W_x^2 + |\alpha_y|W_y^2))$$

With these notations in hand, (1-2) can be explicitly written as

$$\left[\begin{array}{c|c} \boldsymbol{\alpha}^2\mathbf{\Psi'_{data}} + \lambda_w \boldsymbol{L_w} + \beta \boldsymbol{A_f} & \boldsymbol{\alpha}\bar{\boldsymbol{\alpha}}\mathbf{\Psi'_{data}} \\ \hline \boldsymbol{\alpha}\bar{\boldsymbol{\alpha}}\mathbf{\Psi'_{data}} & \bar{\boldsymbol{\alpha}}^2\mathbf{\Psi'_{data}} + \lambda_I \boldsymbol{L_I} \end{array}\right] \left[\begin{array}{c} W^k \\ I_k \end{array}\right] = \left[\begin{array}{c} b_w \\ b_I \end{array}\right] \tag{6}$$

$$\begin{aligned} \boldsymbol{L_I} &= \boldsymbol{D_x}^T \boldsymbol{c_x} \mathbf{\Psi}'_{\mathbf{rI}} \boldsymbol{D_x} + \boldsymbol{D_y}^T \boldsymbol{c_y} \mathbf{\Psi}'_{\mathbf{rI}} \boldsymbol{D_y} \\ \boldsymbol{L_w} &= \boldsymbol{D_x}^T \boldsymbol{c_x} \mathbf{\Psi}'_{\mathbf{rw}} \boldsymbol{D_x} + \boldsymbol{D_y}^T \boldsymbol{c_y} \mathbf{\Psi}'_{\mathbf{rw}} \boldsymbol{D_y} \\ \boldsymbol{A_f} &= \boldsymbol{\alpha}^T (\underbrace{\boldsymbol{D_x}^T \mathbf{\Psi'_{\boldsymbol{f}}} \boldsymbol{D_x} + \boldsymbol{D_y}^T \mathbf{\Psi'_{\boldsymbol{f}}} \boldsymbol{D_y}}_{\boldsymbol{L_f}}) \boldsymbol{\alpha} + \gamma \mathbf{\Psi'_{aux}} \end{aligned}$$

and $\boldsymbol{D_x}, \boldsymbol{D_y}$ denote the horizontal and vertical derivatives operators. The vectors b_w, b_I are given by

$$\begin{aligned} b_w &= \boldsymbol{\alpha}^T \mathbf{\Psi'_{data}} J^k + \beta \boldsymbol{L_f} W_m + \gamma \mathbf{\Psi'_{aux}} W \\ b_I &= \bar{\boldsymbol{\alpha}}^T \mathbf{\Psi'_{data}} J^k. \end{aligned}$$

The weighting matrices $\boldsymbol{c_x}, \boldsymbol{c_y}$ are given by $\boldsymbol{c_x} = diag(|\alpha_x|)$ and $\boldsymbol{c_y} = diag(|\alpha_y|)$.

As mentioned in Sec. 3.2, we solve Eq. 6 using iterative reweighed least square, i.e., iterating between computing the non linear terms $\mathbf{\Psi'}$ based on the current estimate, and updating the solution of I_k and W_k.

II. Matte Update: The EL equation for Eq. 12 is given by

$$\left(\sum_k \mathbf{\Psi'_k} + \lambda_\alpha \boldsymbol{L_\alpha} + \beta \tilde{\boldsymbol{A}}_{\boldsymbol{f}} \right) \alpha = \sum_k \mathbf{A_k}(J - I_k) + \beta \boldsymbol{W}^T \boldsymbol{L_f} W_m, \tag{7}$$

where $\boldsymbol{L_f}$ as defined above, $\boldsymbol{W} = \mathrm{diag}(W)$ and

$$\begin{aligned} \mathbf{\Psi'_k} &= \mathrm{diag}\left(\Psi'\left((\alpha W + (1-\alpha)I^k - J^k)^2 (W - I_k)\right)\right) \\ \boldsymbol{L_\alpha} &= \boldsymbol{D_x}^T \mathbf{\Psi'_{\boldsymbol{\alpha}}} \boldsymbol{D_x} + \boldsymbol{D_y}^T \mathbf{\Psi'_{\boldsymbol{\alpha}}} \boldsymbol{D_y} \\ \tilde{\boldsymbol{A}}_{\boldsymbol{f}} &= \boldsymbol{W}^T \boldsymbol{L_f} \boldsymbol{W} \end{aligned}$$

As before, $\mathbf{\Psi}'_\alpha = \mathrm{diag}(\Psi'(\|\nabla\alpha\|^2))$.

2 Blend Factor Estimation

We assume a small per-image deviation from a global blend factor c, i.e., the opacity of the k^{th} image watermark is $c_k \cdot c$. We solve for a per-image r c_k by minimizing the following objective

$$\Psi((c_k c\alpha W + (1 - c_k c\alpha) I_k - J_k)^2) + \lambda_c (c_k - 1)^2, \tag{8}$$

where λ_c is the weight of the regularization term (which controls the amount of deviation from the global opacity), and Ψ is a robust function. Minimizing

Input: A collection of watermarked images $\{J_k\}$.
(optional) In case of random watermark position – bounding box around watermarked region in a single image J_i
Output: Watermark W, alpha matte α, watermark free collection $\{I_k\}$

1. Compute initial matted watermark & detect all watermarks (Sec. 3.1)
2. Initialize α using single-image matting
3. Estimate global (average) blend factor c
4. **for** $t = 1$ **to** T **do**
   - **for** $k = 1$ **to** K **do**
     - **I. Image–Watermark Decomposition**: Solve for I_k and W_k, keeping α and W fixed.
     - **II. Opacity Estimation (Optional):** Solve for small per-image variation in opacity c_k.
     - **III. Flow Estimation (Optional):** Solve for small per-image watermark perturbation ω_k
   - **end**
   - **IV. Watermark Update**: Solve for W keeping $\{I_k, W_k, c_k, \omega_k\}$, and α fixed.
   - **IIV. Matte Update**: Solve for α keeping $\{I_k, W_k\, c_k, \omega_k\}$, and W fixed.

end

Algorithm 1: Our automatic multi-image watermark removal algorithm

this equation w.r.t. c_k, and keeping the rest of the unknowns (W, α, I_k, c) fixed, leads to

$$c_k = \Big(\lambda_c - \sum(\Psi'_k)(I_k - J_k)\alpha(W - I_k)\Big) \Big/ \Big(\lambda_c + \sum(\Psi'_k\alpha^2(W - I_k)^2)\Big), \quad (9)$$

where $\Psi'_k = \Psi'(c_k c\alpha W + (1 - c_k c\alpha)I_k - J_k)^2$. This estimation is integrated into our multi-image matting and reconstruction algorithm as additional (optional) step (see Alg. 1).

3 The Effect of Number of Images:

We tested how the number of images effects on our performance. In particular, we evaluated the impact of two factors: (i) #images used to estimate the initial matted watermark (Sec. 3.1), (ii) #images used in the multi-image matting step (Sec. 3.2). We denote these two factors by N_{init}, and $N_{matting}$, respectively.

The computed PSNR and DSSIM errors for running our algorithm with different values of N_{init}, and $N_{matting}$, on the *CVPR17* dataset, are presented in Fig. 1(a-b). An example of our reconstructions for the minimal ($N_{init} = 10, N_{matting} = 5$), and maximal ($N_{init} = 300, N_{matting} = 70$) combinations are shown in Fig. 1(c-d), respectively. As expected, the results improves as more

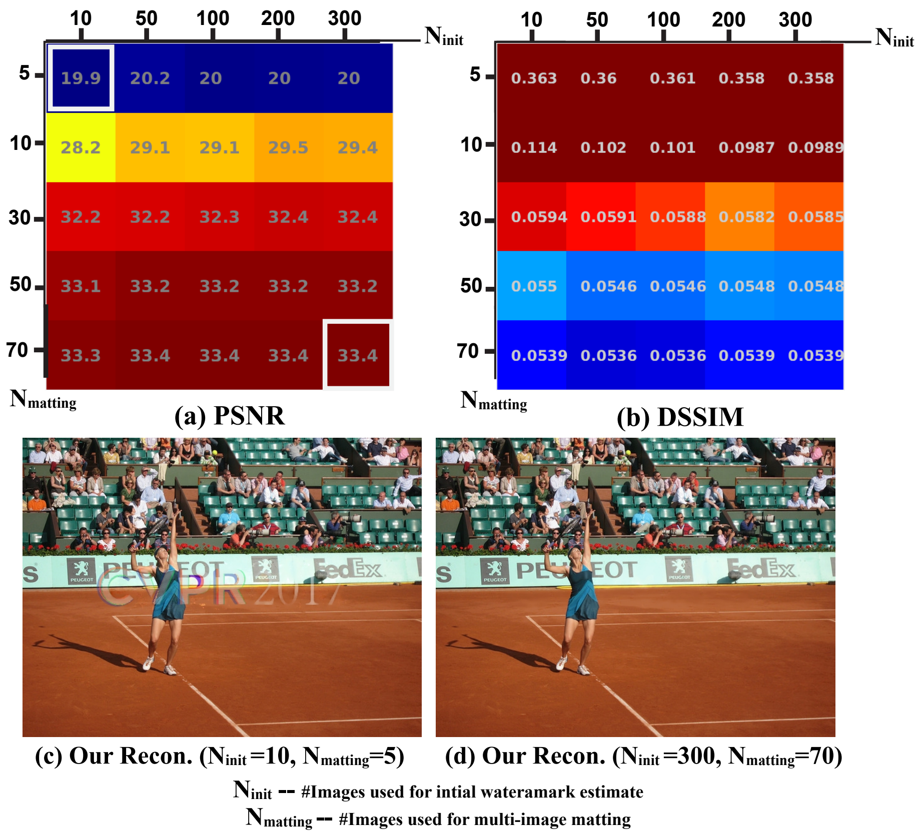

Figure 1: The effect of number of images. (a-b) Error matrix measuring PSNR and DSSIM between the ground truth and our results, respectively; the number of images used for initial matted-watermark estimation (N_{init}) is changing along the columns; the number of images used for the multi-image matting step ($N_{matting}$) is changing along the rows. (c-d) An example of our result corresponding to locations (1,1) and (5,5) in the error matrix, respectively.

images are used. However, with $N_{init} = 300$ the errors are already visually unnoticeable. Furthermore, this evaluation shows that the accuracy of the initial matted-watermark has much higher impact on the quality of the results than the number of images used for the multi-image matting step. That is, with a good initialization of the watermark in hand, it is enough to have an order of tens images for decomposition step.