

Adaptive and Constrained Algorithms for Inverse Compositional Active Appearance Model Fitting

— CVPR 2008 Paper Supplemental Material —

George Papandreou and Petros Maragos
School of Electrical and Computer Engineering
National Technical University of Athens, Greece

<http://cvsp.cs.ntua.gr>

1. Including Priors into Flexible Warp-based Inverse Compositional Algorithms

In this note we expand the discussion of Section 4.1 of the main paper on computing the inverse-compositional to additive parameter update $(4+n) \times (4+n)$ Jacobian matrix $J_{\tilde{\mathbf{p}}}$. Full details are given for the particularly interesting case of the thin-plate spline warp [2].

Our starting point is the relationship $\mathbf{W}(\mathbf{x}; \tilde{\mathbf{p}} + J_{\tilde{\mathbf{p}}}d\tilde{\mathbf{p}}) \approx \mathbf{W}(\mathbf{W}(\mathbf{x}; -d\tilde{\mathbf{p}}); \tilde{\mathbf{p}})$, which holds for all points \mathbf{x} in the image plane to first order in $d\tilde{\mathbf{p}}$ [1,4]. Differentiation w.r.t. $d\tilde{\mathbf{p}}$ yields

$$\underbrace{\frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}}}_{2 \times (4+n)} \bigg|_{(\mathbf{x}; \tilde{\mathbf{p}})} \underbrace{J_{\tilde{\mathbf{p}}}}_{(4+n) \times (4+n)} \approx - \underbrace{\frac{\partial \mathbf{W}}{\partial \mathbf{x}}}_{2 \times 2} \bigg|_{(\mathbf{x}; \tilde{\mathbf{p}})} \underbrace{\frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}}}_{2 \times (4+n)} \bigg|_{(\mathbf{x}; \tilde{\mathbf{p}}=0)}. \quad (1)$$

Equation (1) gives $2 \times (4+n)$ constraints per image point \mathbf{x} .

Since the warp \mathbf{W} is uniquely determined by positions of the shape landmarks, it suffices to apply Eq. (1) L times, once for the spatial position \mathbf{x}_l , $l = 1, \dots, L$ of each landmark on the mean shape \mathbf{s}_0 . Putting together the L resulting terms in a single block matrix equation yields

$$\underbrace{\begin{bmatrix} \frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}} \big|_{(\mathbf{x}_1; \tilde{\mathbf{p}})} \\ \vdots \\ \frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}} \big|_{(\mathbf{x}_L; \tilde{\mathbf{p}})} \end{bmatrix}}_{2L \times (4+n)} \underbrace{J_{\tilde{\mathbf{p}}}}_{(4+n) \times (4+n)} \approx - \underbrace{\begin{bmatrix} \frac{\partial \mathbf{W}}{\partial \mathbf{x}} \big|_{(\mathbf{x}_1; \tilde{\mathbf{p}})} & \frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}} \big|_{(\mathbf{x}_1; \tilde{\mathbf{p}}=0)} \\ \vdots \\ \frac{\partial \mathbf{W}}{\partial \mathbf{x}} \big|_{(\mathbf{x}_L; \tilde{\mathbf{p}})} & \frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}} \big|_{(\mathbf{x}_L; \tilde{\mathbf{p}}=0)} \end{bmatrix}}_{2L \times (4+n)}. \quad (2)$$

Denoting as $\frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}} \big|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})}$ the $(2L) \times (4+n)$ stacked matrix of derivatives on the left-hand-side and as $\frac{\partial \mathbf{W}}{\partial \mathbf{x}} \big|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})} \odot \frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}} \big|_{(\mathbf{x}_{1:L}; \mathbf{0})}$ the stacked block-by-block matrix product on the right-hand-side of the previous equation, we can write it more compactly as

$$\underbrace{\frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}} \big|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})}}_{2L \times (4+n)} \underbrace{J_{\tilde{\mathbf{p}}}}_{(4+n) \times (4+n)} \approx - \underbrace{\frac{\partial \mathbf{W}}{\partial \mathbf{x}} \big|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})} \odot \frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}} \big|_{(\mathbf{x}_{1:L}; \mathbf{0})}}_{2L \times (4+n)}. \quad (3)$$

Solving this with the method of least squares yields the Jacobian estimate

$$J_{\tilde{\mathbf{p}}} = - \left(\frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}} \bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})}^T \frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}} \bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})} \right)^{-1} \left(\frac{\partial \mathbf{W}}{\partial \mathbf{x}} \bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})} \odot \frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}} \bigg|_{(\mathbf{x}_{1:L}; \mathbf{0})} \right), \quad (4)$$

which is Eq. (22) of our main paper.

We move forward and show how the matrices involved in Eq. (4) can be computed. Regarding the $(2L) \times (4 + n)$ matrix $\frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}}\bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})}$, we need compute the $\frac{\partial \mathbf{W}}{\partial \mathbf{p}}\bigg|_{(\mathbf{x}; \mathbf{p})}$ Jacobian. Applying the chain rule on $\mathbf{W}(\mathbf{x}, \tilde{\mathbf{p}}) = \mathbf{S}_t(\mathbf{W}(\mathbf{x}, \mathbf{p}))$ and considering separately the similarity \mathbf{t} and deformation \mathbf{p} parameters gives

$$\frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}}\bigg|_{(\mathbf{x}; \tilde{\mathbf{p}})} = \left[\frac{\partial \mathbf{S}}{\partial \mathbf{t}}\bigg|_{(\mathbf{W}(\mathbf{x}, \mathbf{p}); \mathbf{t})} \quad \frac{\partial \mathbf{S}}{\partial \mathbf{x}}\bigg|_{(\mathbf{W}(\mathbf{x}, \mathbf{p}); \mathbf{t})} \cdot \frac{\partial \mathbf{W}}{\partial \mathbf{p}}\bigg|_{(\mathbf{x}; \mathbf{p})} \right] \quad (5)$$

Taking advantage of the fact that we only need to evaluate the quantities above on the landmark positions \mathbf{x}_l , it is easy to show (*c.f.* [4, Sec. 4.1.2]) that

$$\frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}}\bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})} = \left[[\mathbf{s}_{\mathbf{p}} \quad \mathbf{s}_{\mathbf{p}}^+ \quad \mathbf{1}_x \quad \mathbf{1}_x^+] \quad (1 + t_1) [\mathbf{s}_1 \quad \dots \quad \mathbf{s}_n] + t_2 [\mathbf{s}_1^+ \quad \dots \quad \mathbf{s}_n^+] \right], \quad (6)$$

where $\mathbf{s}_{\mathbf{p}} = \mathbf{s}_0 + \sum_{i=1}^n p_i \mathbf{s}_i$ is the deformed shape, given the parameters \mathbf{p} , \mathbf{s}^+ denotes the shape \mathbf{s} rotated counter-clockwise by 90° and $\mathbf{1}_x = [1 \ 0 \ \dots \ 1 \ 0]^T$ is the shape with 1's in the x -coordinate and 0's in the y -coordinate.

Regarding the $2L \times 2$ matrix $\frac{\partial \mathbf{W}}{\partial \mathbf{x}}\bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})}$, we need compute the Jacobian $\frac{\partial \mathbf{W}}{\partial \mathbf{x}}\bigg|_{(\mathbf{x}; \mathbf{p})}$. Application of the chain rule on $\mathbf{W}(\mathbf{x}, \tilde{\mathbf{p}}) = \mathbf{S}_t(\mathbf{W}(\mathbf{x}, \mathbf{p}))$ gives

$$\frac{\partial \mathbf{W}}{\partial \mathbf{x}}\bigg|_{(\mathbf{x}; \tilde{\mathbf{p}})} = \frac{\partial \mathbf{S}}{\partial \mathbf{x}}\bigg|_{(\mathbf{W}(\mathbf{x}, \mathbf{p}); \mathbf{t})} \frac{\partial \mathbf{W}}{\partial \mathbf{x}}\bigg|_{(\mathbf{x}; \mathbf{p})} = \begin{bmatrix} 1 + t_1 & -t_2 \\ t_2 & 1 + t_1 \end{bmatrix} \frac{\partial \mathbf{W}}{\partial \mathbf{x}}\bigg|_{(\mathbf{x}; \mathbf{p})}. \quad (7)$$

Computation of the deformation field Jacobian $\frac{\partial \mathbf{W}}{\partial \mathbf{x}}\bigg|_{(\mathbf{x}; \mathbf{p})}$ depends on the warp family under consideration. For the often used *thin-plate spline* warp [2], we can write the warp function $\mathbf{W}(\mathbf{x}, \mathbf{p})$ in the form of a generalized linear model (*c.f.* [3, App. F])

$$\mathbf{W}(\mathbf{x}, \mathbf{p}) = \underbrace{W(\mathbf{p})}_{2 \times (L+3)} \underbrace{k(\mathbf{x})}_{(L+3) \times 1}, \quad (8)$$

where the vector $k(\mathbf{x})$ is given by

$$k(\mathbf{x}) = [U(|\mathbf{x} - \mathbf{x}_1|) \quad \dots \quad U(|\mathbf{x} - \mathbf{x}_L|) \quad 1 \quad x \quad y]^T, \quad (9)$$

$U(r) = r^2 \ln r^2$ is the spline kernel, and $W(\mathbf{p})$ is determined by requiring that the warp maps s_0 to $s_{\mathbf{p}}$. The final result is

$$\frac{\partial \mathbf{W}}{\partial \mathbf{x}}\bigg|_{(\mathbf{x}; \mathbf{p})} = W(\mathbf{p}) \frac{dk(\mathbf{x})}{d\mathbf{x}}\bigg|_{\mathbf{x}} = W(\mathbf{p}) \begin{bmatrix} 2(1 + \ln r_{\mathbf{x},1}^2)(\mathbf{x} - \mathbf{x}_1)^T \\ \vdots \\ 2(1 + \ln r_{\mathbf{x},L}^2)(\mathbf{x} - \mathbf{x}_L)^T \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad (10)$$

where $r_{\mathbf{x},l} = \|\mathbf{x} - \mathbf{x}_l\|_2$. We need to evaluate $\frac{\partial \mathbf{W}}{\partial \mathbf{x}}\bigg|_{(\mathbf{x}; \mathbf{p})}$ for each landmark point \mathbf{x}_l . Since the term $k(\mathbf{x})$ does not depend on the shape parameter \mathbf{p} , $\frac{dk(\mathbf{x})}{d\mathbf{x}}\bigg|_{\mathbf{x}}$ can be pre-computed and be subsequently used in every AAM iteration.

The cost of computing the Jacobian matrix $J_{\tilde{\mathbf{p}}}$ can be analyzed as follows: (a) Computing the $(2L) \times (4 + n)$ matrix $\frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}}\bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})}$ is $\mathcal{O}(nL)$. (b) Computing the $2L \times 2$ matrix $\frac{\partial \mathbf{W}}{\partial \mathbf{x}}\bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})}$ is $\mathcal{O}(L)$. (c) Forming the stacked block-by-block matrix product $\frac{\partial \mathbf{W}}{\partial \mathbf{x}}\bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})} \odot \frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}}\bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})}$ is $\mathcal{O}(nL)$. (d) Forming the $(4 + n) \times (4 + n)$ least-squares system matrix $\frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}}\bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})}^T \frac{\partial \mathbf{W}}{\partial \tilde{\mathbf{p}}}\bigg|_{(\mathbf{x}_{1:L}; \tilde{\mathbf{p}})}$ is $\mathcal{O}(n^2L)$. (e) Inverting the same system matrix is $\mathcal{O}(n^3)$. Overall, the last two operations dominate the cost of the procedure, whose overall complexity is thus $\mathcal{O}(n^2L + n^3)$.

References

- [1] S. Baker, R. Gross, and I. Matthews. Lucas-Kanade 20 years on: A unifying framework - Part 4. Technical Report CMU-RI-TR-04-14, Robotics Institute, CMU, 2004.
- [2] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Tr. on PAMI*, 11(6):567–585, 1989.
- [3] T. F. Cootes and C. J. Taylor. Statistical models of appearance for computer vision. http://www.isbe.man.ac.uk/~bim/Models/app_models.pdf, 2004.
- [4] I. Matthews and S. Baker. Active appearance models revisited. *Int. J. of Comp. Vision*, 60(2):135–164, 2004.