A Deformation and Lighting Insensitive Metric for Face Recognition Based on Dense Correspondences

Anne Jorstad[1] **David Jacobs[1]** **Alain Trouvé[2]**

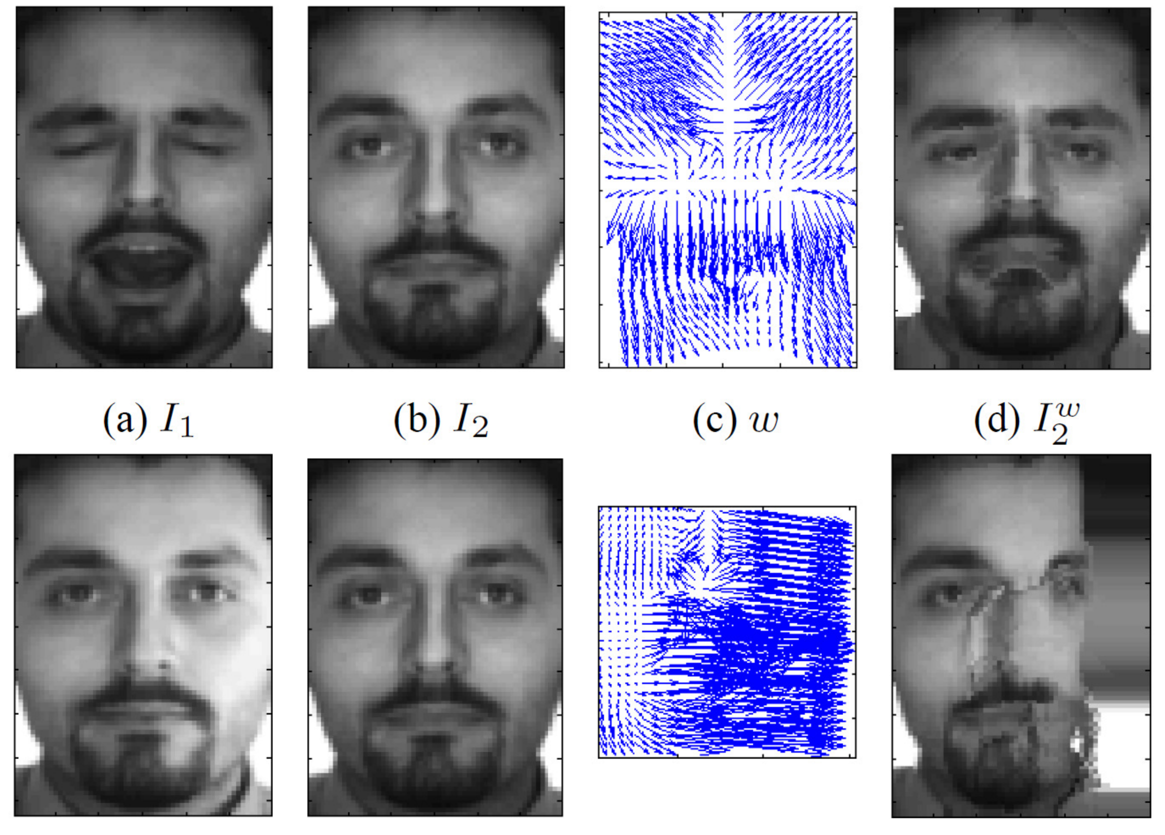
[1] University of Maryland Institute for Advanced Computer Studies

[2] Centre de Mathématiques et de Leurs Applications, Ecole Normale Supérieure de Cachan, France

Optical Flow for Face Recognition

Goal: Compute the lowest cost deformation and intensity changes relating two face images, and use these for recognition.

Poor results are achieved when traditional optical flow is used:



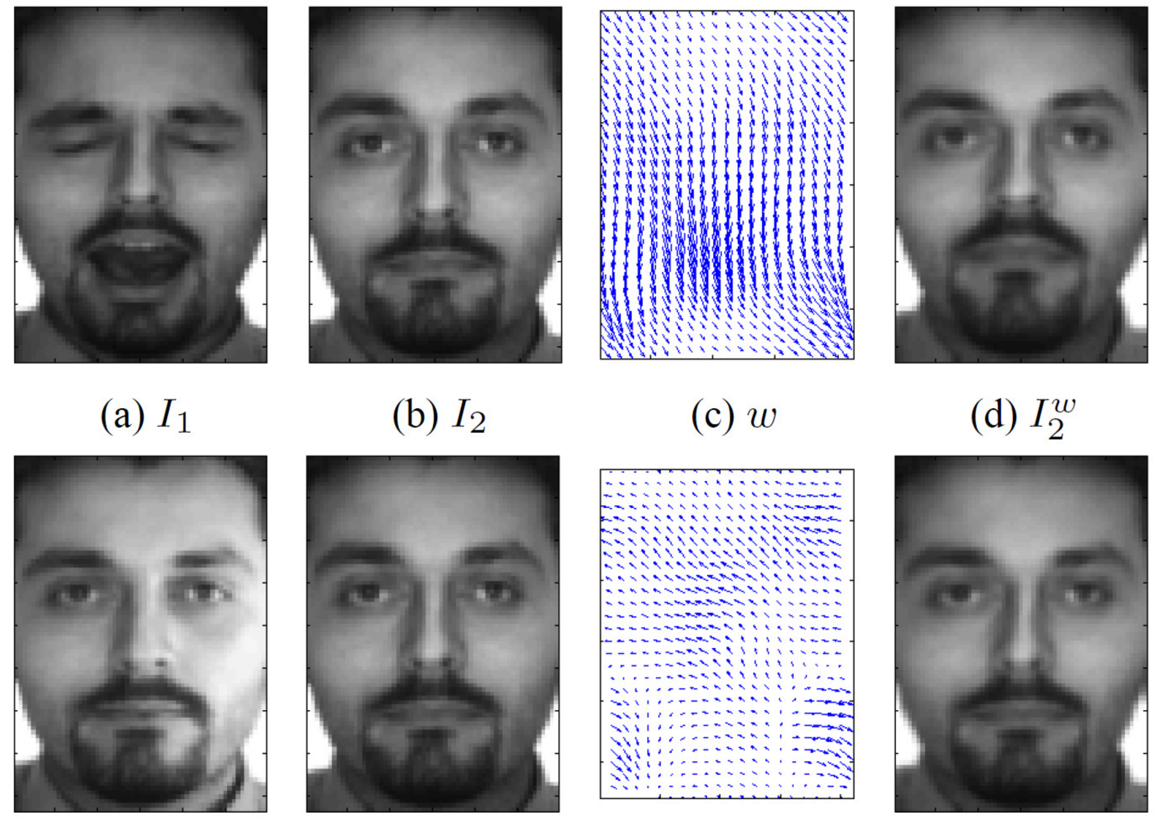
(a) I_1 (b) I_2 (c) w (d) I_2^w

(e) I_1 (f) I_2 (g) w (h) I_2^w

w : *the flow from* I_1 *to* I_2

I_2^w : I_2 *warped backwards along* w *to attempt to match* I_1

The flow calculated with our proposed method is more stable:



(a) I_1 (b) I_2 (c) w (d) I_2^w

(e) I_1 (f) I_2 (g) w (h) I_2^w

References:

[1] J. W. Neuberger. *Sobolev Gradients and Differential Equations, 2nd Edition*. Springer, 2010.

[2] A. Martinez and R. Benavente. The AR Face Database. *CVC Technical Report #24*, 1998.

This research was funded by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), through the Army Research Laboratory (ARL). All statements of fact, opinion or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of IARPA, the ODNI, or the U.S. Government.

The New Deformation and Lighting Insensitive (DLI) Metric

Goal: Define a metric on face image flows that is insensitive to changes in expression and lighting.

$$E_{\text{DLI}}(w) = (1-\lambda)E_b(w) + \lambda E_r(w)$$

Photometric term :

$$E_b(w) = \frac{1}{2}\sum_{ij} \frac{\|\nabla(I_2^w - I_1)\|^2}{\|\nabla I_1\|^2 + \epsilon^2} = \frac{1}{2}\sum_{ij}(E^x_{b_{ij}})^2 + (E^y_{b_{ij}})^2$$

- Invariant to multiplication by a scalar and addition by a constant.
- Insensitive to changes caused by the effects of lighting variation in 3D scenes (ex: changing the location of a light can magnify or weaken the gradient at the edge of a polyhedron, as the two sides forming the edge are exposed to the light differently).

Regularization term :

$$E_r(w) = \frac{1}{2}\langle K^{-1}w,\ w\rangle_G = \frac{1}{2}\sum_{ij}(E^x_{r_{ij}})^2 + (E^y_{r_{ij}})^2$$

- K : a symmetric positive definite matrix (eg a gaussian)
- G : a generalized inner product defined on MxNx2 structures (dimensions of the flow)

Optimization

Use a modified gradient descent algorithm to minimize $E_{\text{DLI}}(w)$.

Sobolev Gradient[1]: $\nabla_K E = K\nabla E$

- Smoother, results in superior rates of convergence.

Optimize over a dual variable α:

$$w_n = K\alpha_n \qquad \textit{(this is a convolution)}$$

$$\alpha_{n+1} = \alpha_n - \Delta t \cdot \nabla E(w_n)$$

Learning for Improved Results

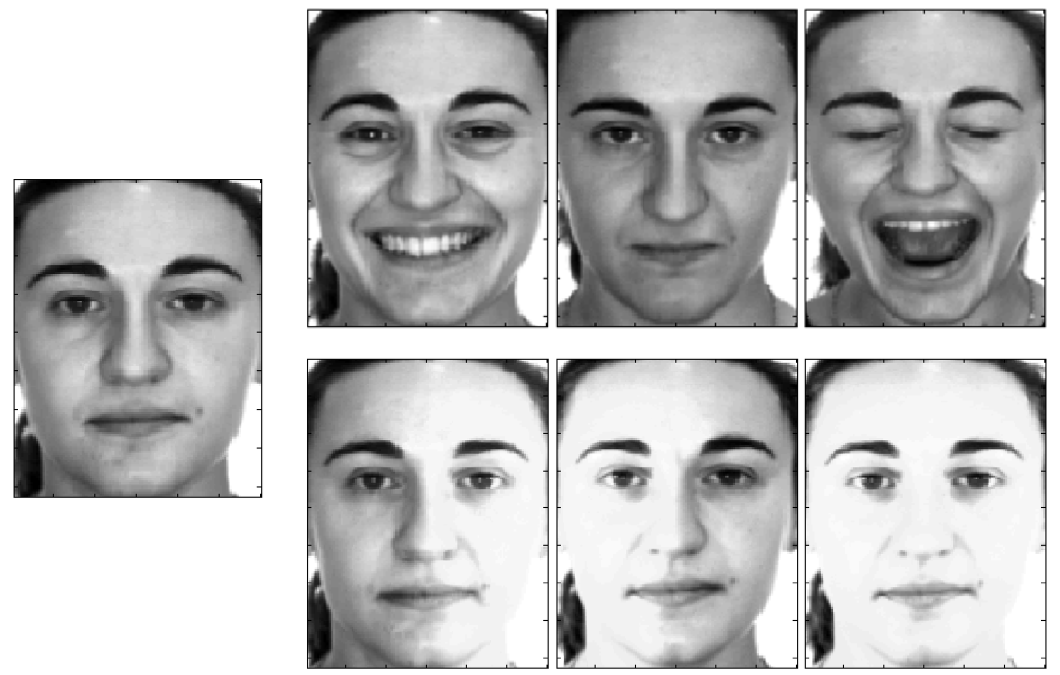
Given pixel matching costs for known image pairs,

- Use Maximum Likelihood estimation to learn typical Gaussian distributions through 4D cost vectors $\vec{E}_{ij} = [E^x_{b_{ij}}\ E^y_{b_{ij}}\ E^x_{r_{ij}}\ E^y_{r_{ij}}]$.
- Assume pixel independence.
- Learn separate models for same-person and different-person image pairs, to calculate probability that two images are from the same person and also from different people.
- The final similarity measure: $S(I_1, I_2) = \dfrac{P_{\text{same}}(\vec{E}(w))}{P_{\text{diff}}(\vec{E}(w))}$

Experiments

Identity of unknown neutral face determined by gallery image resulting in lowest matching cost.

Images from the AR Face Database[2] with variations in expression and lighting.



Variation	*Accuracy*	*Variation*	*Accuracy*
Smile	97.6%	**Left light**	98.8%
Frown	91.6%	**Right light**	99.6%
Scream	79.6%	**Both lights**	98.4%

Cost Function	*Expression*	*Lighting*	*Overall*
Direct	82.0%	96.0%	89.0%
After Learning	89.6%	**98.9%**	**94.3%**
Smile gallery After Learning	86.8%	91.2%	89.0%
Borders removed After Learning	85.1%	96.4%	90.7%