Efficient Region Tracking with Parametric Models of Geometry and Illumination

Gregory D. Hager, Peter N. Belhumeur

Presented by: Jongwoo Lim, Oct 15 2002

The tracking problem

We want to keep track of the interesting object or region in a video sequence.

Intrinsic difficulties

- movements in 3D
- illumination changes
- shape changes (deformation)
- occlusions

Tracking a moving object

Image constancy assumption

$$I(\mathbf{x}, t_0) = I(\mathbf{f}(\mathbf{x}; \boldsymbol{\mu}_t), t) \qquad \text{for all } \mathbf{x} \in R$$

the current image of the object is a ‘warped’ version of the initital ‘reference’ image.

Goal

estimate the warping parameter $\boldsymbol{\mu}$ in each frame

Tracking... : Objective function

Objective function

$$O(\boldsymbol{\mu}) = \sum_{\mathbf{x}\in R} \{I(\mathbf{f}(\mathbf{x};\boldsymbol{\mu}),t) - I(\mathbf{x},t_0)\}^2$$
$$= ||\,\mathbf{I}(\boldsymbol{\mu},t) - \mathbf{I}(\mathbf{0},t_0)\,||^2$$

notaions :

$$\mathbf{I}(\boldsymbol{\mu},t) \triangleq \begin{bmatrix} I(\mathbf{f}(\mathbf{x}_1,\boldsymbol{\mu}),t) \\ I(\mathbf{f}(\mathbf{x}_2,\boldsymbol{\mu}),t) \\ \cdots \\ I(\mathbf{f}(\mathbf{x}_N,\boldsymbol{\mu}),t) \end{bmatrix}$$

$$\mathbf{I}_{\mu_i}(\boldsymbol{\mu},t) = \frac{\partial}{\partial \mu_i}\mathbf{I}(\boldsymbol{\mu},t)$$
$$\mathbf{I}_t(\boldsymbol{\mu},t) = \frac{\partial}{\partial t}\mathbf{I}(\boldsymbol{\mu},t)$$

Tracking... : Parameter estimation

Given $\boldsymbol{\mu}_t$ for time t, the parameter at $(t+\tau)$ can be written as

$$\boldsymbol{\mu}_{t+\tau} \;=\; \boldsymbol{\mu}_t + \delta\boldsymbol{\mu}$$

Rewrite the objective function on $\delta\boldsymbol{\mu}$

$$O(\delta\boldsymbol{\mu}) \;=\; \| \,\mathbf{I}(\boldsymbol{\mu}_t + \delta\boldsymbol{\mu}, t+\tau) \;-\; \mathbf{I}(\mathbf{0}, t_0)\, \|^2$$

Tracking... : Approximation

Talyor expansion of $\mathbf{I}$ about $\boldsymbol{\mu}$ and t
when $\mathrm{M}(\boldsymbol{\mu}, t)$ is Jacobian matrix of $\mathbf{I}$ w.r.t $\boldsymbol{\mu}$

$$\mathbf{I}(\boldsymbol{\mu} + \delta\boldsymbol{\mu}, t + \tau) \;\simeq\; \mathbf{I}(\boldsymbol{\mu}, t) + \mathbf{M}(\boldsymbol{\mu}, t)\,\delta\boldsymbol{\mu} + \tau\,\mathbf{I}_t(\boldsymbol{\mu}, t)$$

and $\tau\,\mathbf{I}_t(\boldsymbol{\mu}, t) \;\simeq\; \mathbf{I}(\boldsymbol{\mu}, t + \tau) - I(\boldsymbol{\mu}, t)$

$$\begin{aligned} O(\delta\boldsymbol{\mu}) &\simeq \;\|\, \mathbf{I}(\boldsymbol{\mu}, t) + \mathbf{M}\,\delta\boldsymbol{\mu} + \tau\,\mathbf{I}_t(\boldsymbol{\mu}, t) - \mathbf{I}(\mathbf{0}, t_0)\,\|^2 \\ &\simeq \;\|\, \mathbf{M}\,\delta\boldsymbol{\mu} + \mathbf{I}_t(\boldsymbol{\mu}, t + \tau) - \mathbf{I}(\mathbf{0}, t_0)\,\|^2 \end{aligned}$$

Tracking... : Parameter update

If we define the error vector $\mathbf{e}(t+\tau)$ as

$$\mathbf{e}(t+\tau) \;=\; \mathbf{I}(\boldsymbol{\mu}_t, t+\tau) - \mathbf{I}(\mathbf{0}, t_0)$$

and solve $\nabla O \;=\; 0$ using , then

$$\mathrm{M}^t\mathrm{M}\,\delta\boldsymbol{\mu} \;=\; -\mathrm{M}^t\,\mathbf{e}(t+\tau)$$

$$\boldsymbol{\mu}_{t+\tau} \;=\; \boldsymbol{\mu}_t \;-\; \left(\mathrm{M}^t\mathrm{M}\right)^{-1}\mathrm{M}^t\,\mathbf{e}(t+\tau)$$

Tracking... : Efficiency

To calculate $\delta\boldsymbol{\mu}$, M is necesary, but,

$$\mathrm{M}(\mu, t) = \begin{bmatrix} \boxed{\nabla_{\mathbf{f}}\, I(\mathbf{f}(\mathbf{x}_1;\boldsymbol{\mu}), t)^t} \; \mathbf{f}_{\mu}(\mathbf{x}_1;\boldsymbol{\mu}) \\ \nabla_{\mathbf{f}}\, I(\mathbf{f}(\mathbf{x}_2;\boldsymbol{\mu}), t)^t \; \mathbf{f}_{\mu}(\mathbf{x}_2;\boldsymbol{\mu}) \\ \vdots \\ \nabla_{\mathbf{f}}\, I(\mathbf{f}(\mathbf{x}_N;\boldsymbol{\mu}), t)^t \; \mathbf{f}_{\mu}(\mathbf{x}_N;\boldsymbol{\mu}) \end{bmatrix}$$

M includes time-varing quantities.

Tracking... : Making time-independent

If μ is exact, from Image Constancy assumption,

$$\nabla_{\mathbf{x}} I(\mathbf{x}, t_0) = \mathbf{f}_{\mathbf{x}}(\mathbf{x};\boldsymbol{\mu})^t \boxed{\nabla_{\mathbf{f}} I(\mathbf{f}(\mathbf{x};\boldsymbol{\mu}), t)}$$

M does not depend on t (at least directly).

$$\mathbf{M}(\mu) = \begin{bmatrix} \boxed{\nabla_{\mathbf{x}} I(\mathbf{x}_1, t_0)^t \, \mathbf{f}_{\mathbf{x}}(\mathbf{x}_1;\boldsymbol{\mu})^{-1}} \, \mathbf{f}_{\mu}(\mathbf{x}_1;\boldsymbol{\mu}) \\ \nabla_{\mathbf{x}} I(\mathbf{x}_2, t_0)^t \, \mathbf{f}_{\mathbf{x}}(\mathbf{x}_2;\boldsymbol{\mu})^{-1} \, \mathbf{f}_{\mu}(\mathbf{x}_2;\boldsymbol{\mu}) \\ \vdots \\ \nabla_{\mathbf{x}} I(\mathbf{x}_N, t_0)^t \, \mathbf{f}_{\mathbf{x}}(\mathbf{x}_N;\boldsymbol{\mu})^{-1} \, \mathbf{f}_{\mu}(\mathbf{x}_N;\boldsymbol{\mu}) \end{bmatrix}$$

Tracking... : Making time-independent

If μ is exact, from Image Constancy assumption,

$$\nabla_{\mathbf{x}} I(\mathbf{x}, t_0) \;=\; \mathbf{f}_{\mathbf{x}}(\mathbf{x};\boldsymbol{\mu})^t \boxed{\nabla_{\mathbf{f}} I(\mathbf{f}(\mathbf{x};\boldsymbol{\mu}), t)}$$

M does not depend on t (at least directly).

$$\mathbf{M}(\mu) \;=\; \begin{bmatrix} \nabla_{\mathbf{x}}\, I(\mathbf{x}_1, t_0)^t \boxed{\mathbf{f}_{\mathbf{x}}(\mathbf{x}_1;\boldsymbol{\mu})^{-1}\, \mathbf{f}_{\mu}(\mathbf{x}_1;\boldsymbol{\mu})} \\ \nabla_{\mathbf{x}}\, I(\mathbf{x}_2, t_0)^t\, \mathbf{f}_{\mathbf{x}}(\mathbf{x}_2;\boldsymbol{\mu})^{-1}\, \mathbf{f}_{\mu}(\mathbf{x}_2;\boldsymbol{\mu}) \\ \vdots \\ \nabla_{\mathbf{x}}\, I(\mathbf{x}_N, t_0)^t\, \mathbf{f}_{\mathbf{x}}(\mathbf{x}_N;\boldsymbol{\mu})^{-1}\, \mathbf{f}_{\mu}(\mathbf{x}_N;\boldsymbol{\mu}) \end{bmatrix}$$

Tracking... : Factoring M

If we can factor $\mathrm{f}_{\mathbf{x}}(\mathbf{x};\boldsymbol{\mu})^{-1}\,\mathrm{f}_{\mu}(\mathbf{x};\boldsymbol{\mu}) \;=\; \Gamma(\mathbf{x})\,\Sigma(\boldsymbol{\mu})$,
we have

$$\mathrm{M}(\mu) \;=\; \begin{bmatrix} \nabla_{\mathbf{x}}\, I(\mathbf{x}_1, t_0)^t\, \Gamma(\mathbf{x}_1) \\ \nabla_{\mathbf{x}}\, I(\mathbf{x}_2, t_0)^t\, \Gamma(\mathbf{x}_2) \\ \vdots \\ \nabla_{\mathbf{x}}\, I(\mathbf{x}_N, t_0)^t\, \Gamma(\mathbf{x}_N) \end{bmatrix} \Sigma(\boldsymbol{\mu}) \;=\; \mathrm{M}_0 \Sigma(\boldsymbol{\mu})$$

and M_0 can be calculated in off-line.

Tracking... : Algorithm

Offline

- Define the target region
- Acquire and store the reference template
- Compute $\mathbf{M}_0$ and $\Lambda = \mathbf{M}_0^t \mathbf{M}_0$

Online

- Compute $\mathbf{e}(t+\tau)$ using most recent $\boldsymbol{\mu}$
- Solve the system $\Sigma^t \Lambda \Sigma \delta\boldsymbol{\mu} = -\Sigma^t \mathbf{M}_0^t \mathbf{e}(t+\tau)$
- Update $\boldsymbol{\mu} \leftarrow \boldsymbol{\mu} + \delta\boldsymbol{\mu}$

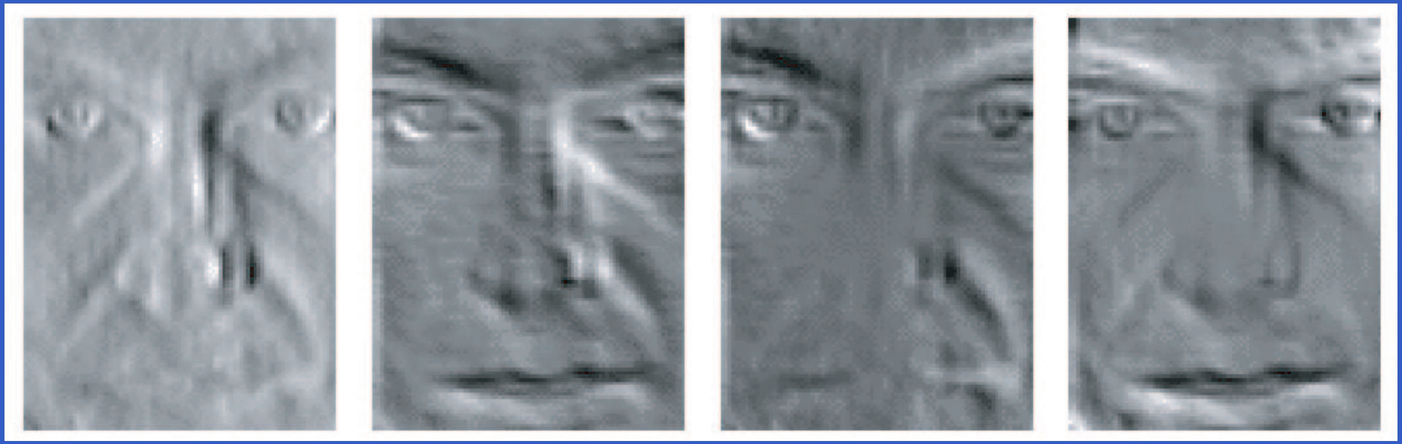
Tracking... : A few examples

Model	$\boldsymbol{\mu}$, $\mathbf{f}(\mathbf{x};\boldsymbol{\mu})$	$\Gamma(\mathbf{x})$	$\Sigma(\boldsymbol{\mu})$
Pure Trans. [Lucas,Kanade]	$\mathbf{u} = (u, v)$ $\mathbf{x} + \mathbf{u}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
Rigid Motion + Scale	$(\mathbf{u}, \theta, s)$ $s\mathbf{R}(\theta)\mathbf{x} + \mathbf{u}$	$\begin{bmatrix} 1 & 0 & -y & x \\ 0 & 1 & x & y \end{bmatrix}$	$\begin{bmatrix} \frac{1}{s}\mathbf{R}(-\theta) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & \frac{1}{s} \end{bmatrix}$
Affine [Shi,Tomasi]	$\mathbf{u}, \mathbf{A} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ $\mathbf{A}\mathbf{x} + \mathbf{u}$	$\begin{bmatrix} 1 & 0 & x & 0 & y & 0 \\ 0 & 1 & 0 & x & 0 & y \end{bmatrix}$	$\begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}^{-1} \end{bmatrix}$
Sample Nonlinear	(u, v, a) $\mathbf{x} + \begin{bmatrix} u \\ v + \frac{1}{2}ax^2 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & x & \frac{x^2}{2} \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -a & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Tracking... : Motion template

$$\mathbf{M} = [\, \mathbf{I}_{\mu_1} \,|\, \mathbf{I}_{\mu_2} \,|\, \cdots \,|\, \mathbf{I}_{\mu_n} \,]$$

Each column can be thought as a ‘motion template’ which represents the changes in brightness induced by the motion.



x-translation y-translation rotation scale

Tracking... : Aperture problem

We have assumed that $\mathrm{M}^t\mathrm{M}$ has full rank.

If $\mathrm{M}^t\mathrm{M}$ is rank deficient, there exists a motion γ

$$\left(\nabla_{\mathbf{f}} I^t \, \mathbf{f}_\mu \, |_{\mathbf{x}=\mathbf{x}_i} \right) \gamma \; = \; 0, \qquad 1 \leq i \leq N$$

This motion corresponds to the displacement of every pixel in the region is orthogonal to the local image gradient.

Illumination insensitive tracking

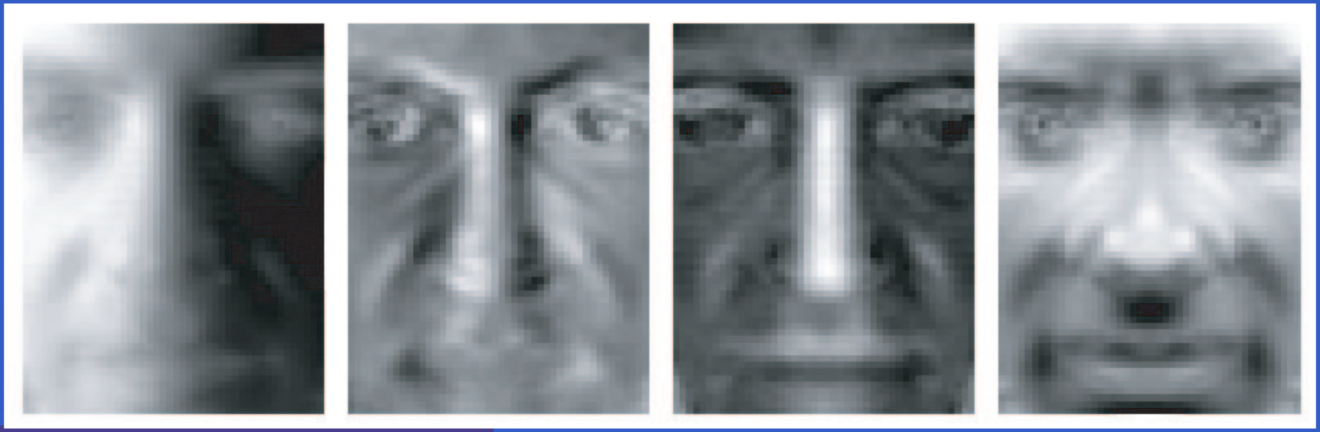
A few ‘basis’ images can be used to account for large changes in illumination.

- Lambertian surface : $E = a\,\mathbf{n} \cdot \mathbf{s}$
- no shadowing : three basis images required
- shadowing : more basis images for good approximation

Illumination... : Basis images

$$\mathbf{B} = [\,\mathbf{B}_1\,|\,\mathbf{B}_2\,|\,\ldots\,|\mathbf{B}_m\,]$$

- $\mathbf{B}_1 = \mathbf{I}(\mathbf{0}, t_0)$: the template image
- $\mathbf{B}_2 = (1, 1, \ldots, 1)^t$ to model global brightness
- $\mathbf{B}_{3:m}$: eigen vectors of images taken under varying illumination



Illumination... : Solution

$$\mathbf{I}(\boldsymbol{\mu}_t, t) \;=\; \mathbf{I}(\mathbf{0}, t_0) + \mathbf{B}\,\boldsymbol{\lambda}_t$$

M is factored to M_0 and Σ with approximation.

We have a closed form solution for $\delta\boldsymbol{\mu}$ when Σ is invertible, so no overhead in online computation.

$$\delta\boldsymbol{\mu} \;=\; -\Sigma^{-t}(\mathrm{M}_0^t \mathbf{N} \mathrm{M}_0)^{-1} \mathrm{M}_0^t \mathbf{N}\,\mathbf{e}(t+\tau)$$

$$\mathbf{N} \;=\; (\mathbf{1} - \mathbf{B}(\mathbf{B}^t\mathbf{B})^{-1}\mathbf{B}^t)$$

Resistant to occlusion

We treat occlusions as 'outliers' in estimation.

$$O_R(\boldsymbol{\mu}) = \sum_{\mathbf{x}\in\mathcal{R}} \rho(I(\mathbf{f}(\mathbf{x};\boldsymbol{\mu}),t) - I(\mathbf{x},t_0))$$

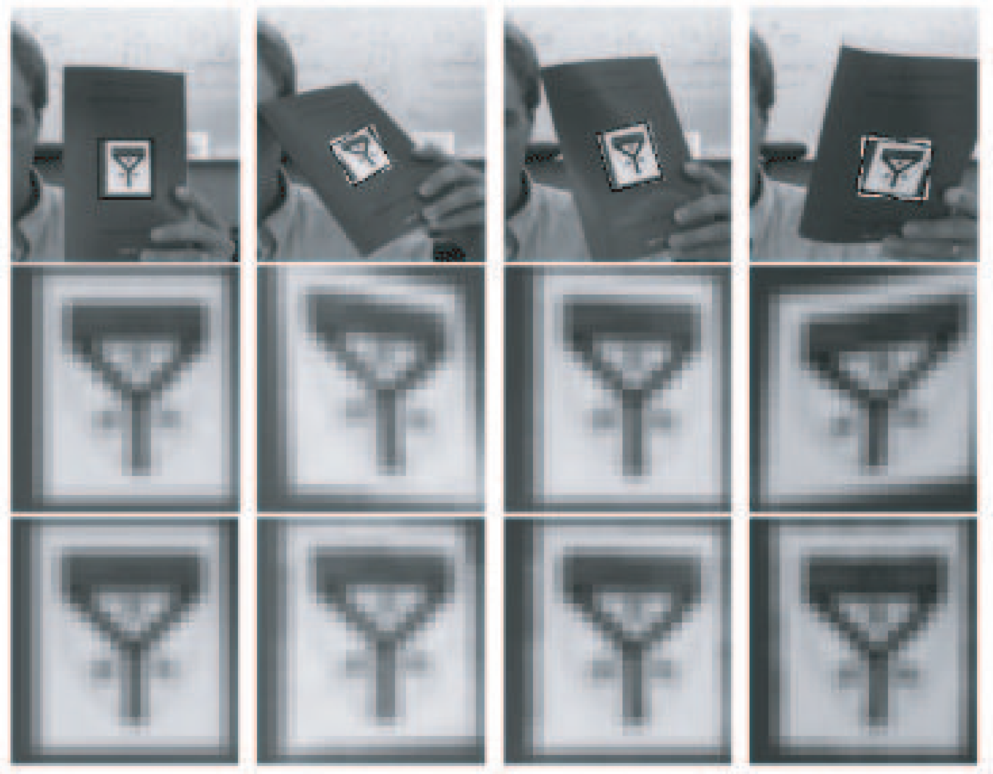
By iteratively adjusting the weight of each pixel, control the deterioration of the outliers.

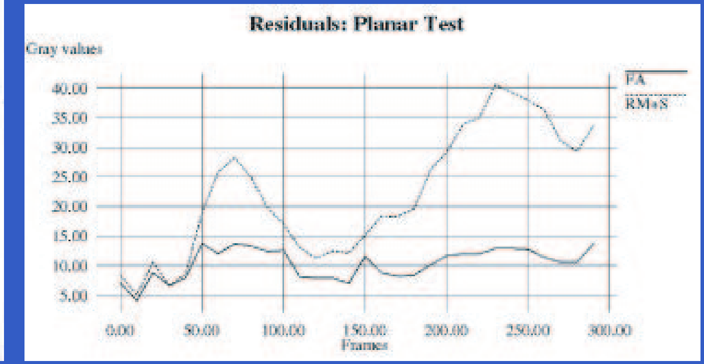
$$\mathbf{r}^i = \mathbf{e}(t+\tau) - \mathbf{M}(\boldsymbol{\mu})\delta\boldsymbol{\mu}^i$$

$$\Sigma^t \Lambda \Sigma\, \delta\boldsymbol{\mu}^{i+1} = \Sigma^t \mathbf{M}_0^t\, \mathbf{W}^i \mathbf{r}^i$$

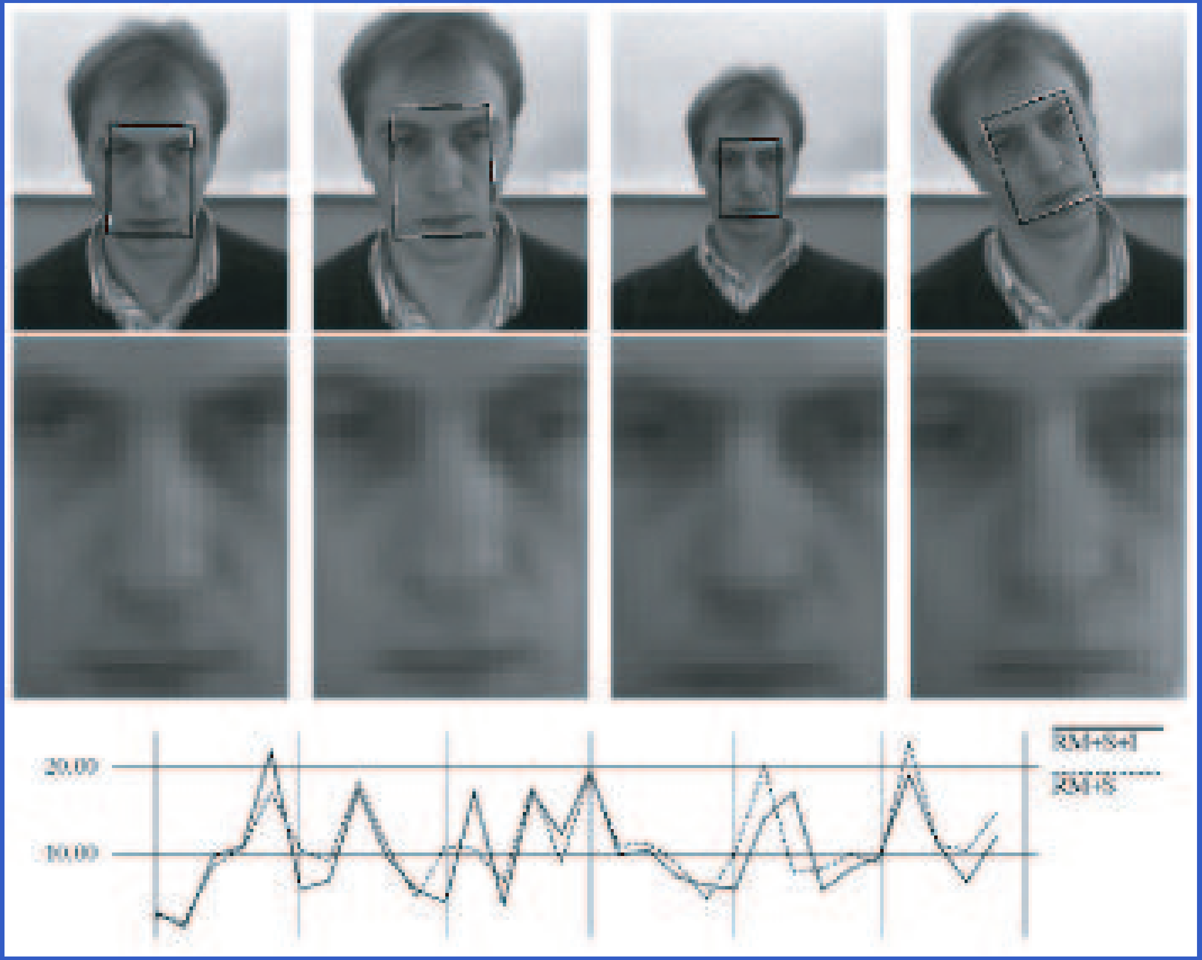
Results : Planar

RM+S vs Affine model

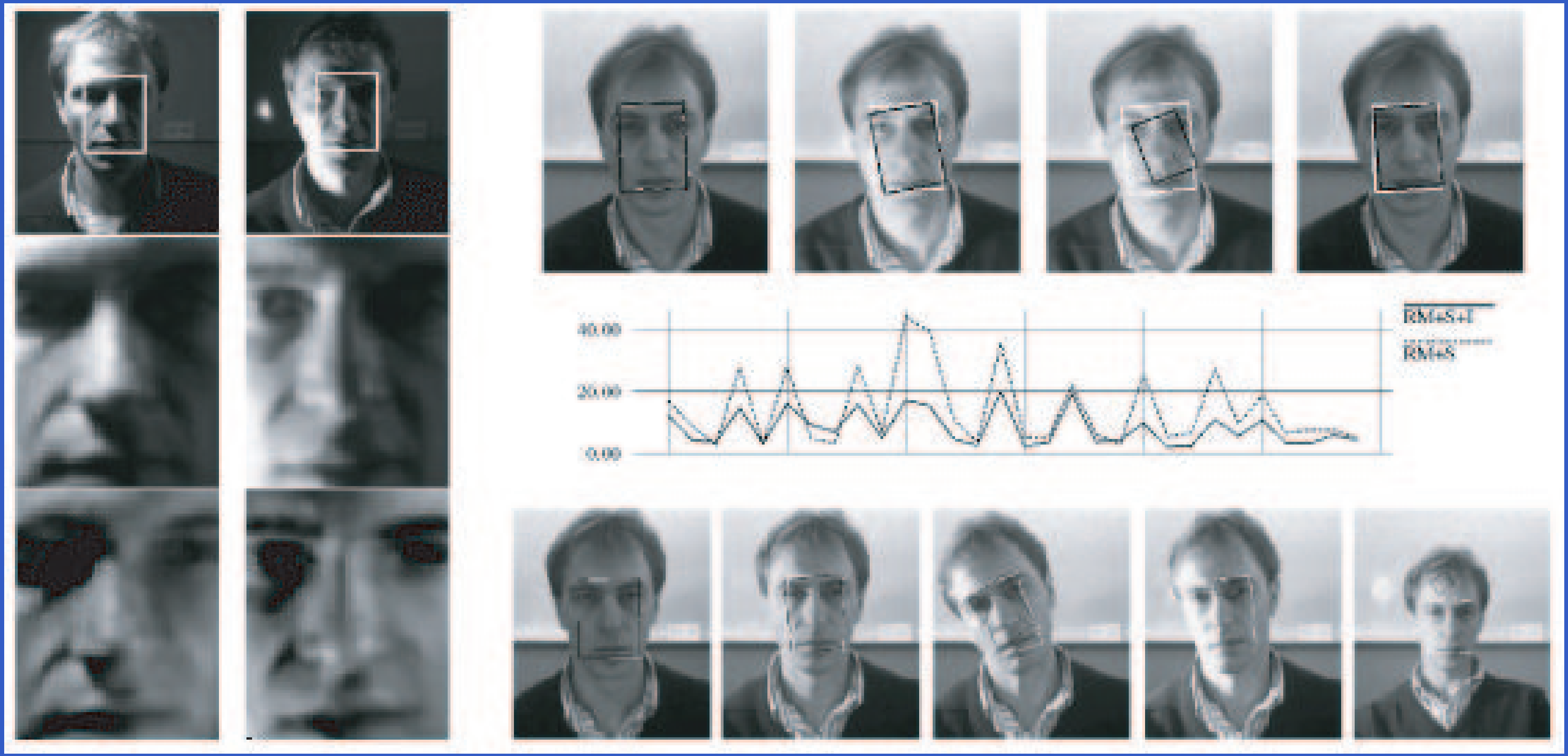




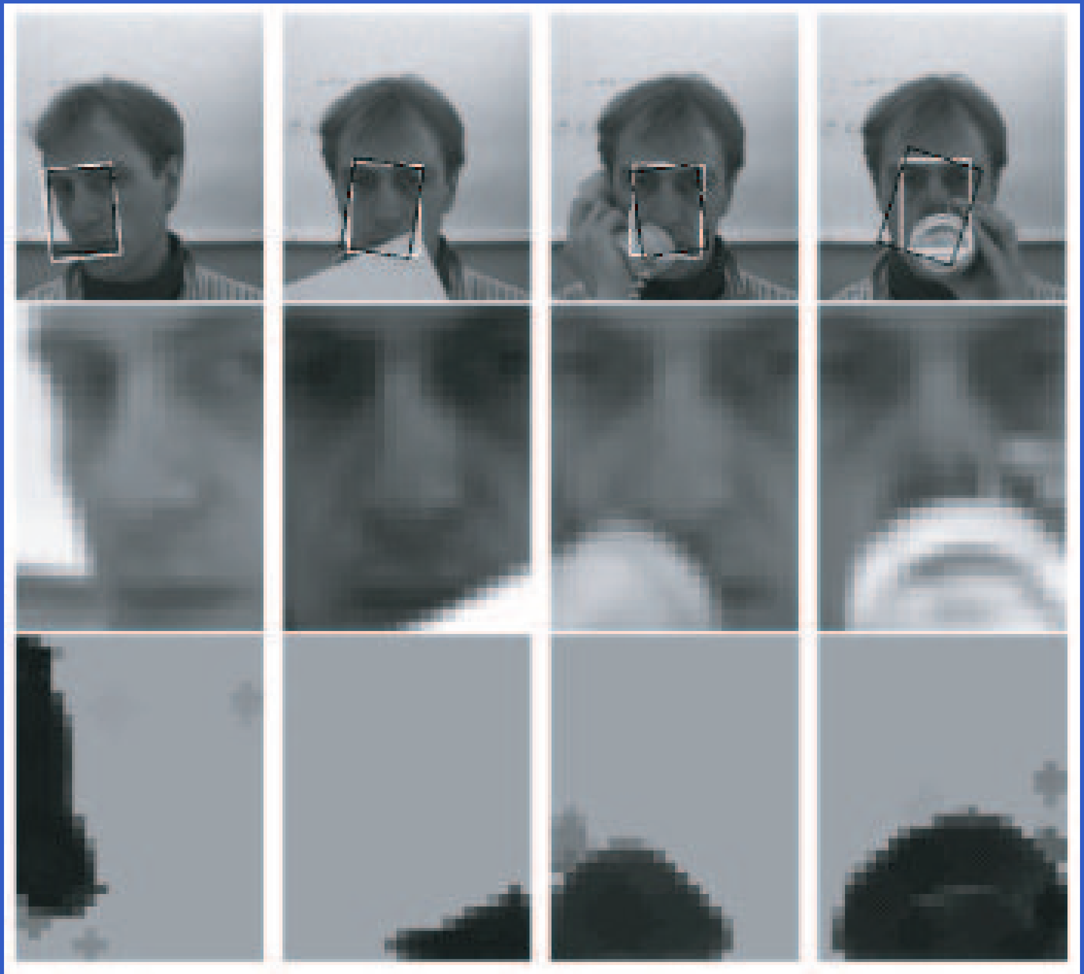
Results : Face



Results : Face + Illumination



Results : Face + Occlusion



Conclusion & Discussion

- Parameter estimation techniques under
  - geometric distortion
  - changing illumination
  - partial occlusion
- Inter-frame changes should be small
- Determine the illumination basis online
- Utilize shape information on the target

References

- Hager & Belhumeur, PAMI ’98
  Efficient Region Tracking with Parametric Models of Geometry and Illumination
- Belhumeur & Kriegman, CVPR ’96
  What Is the Set of Images of an Object Under All Possible Lighting Conditions
- Baker & Matthews, CVPR ’01
  Equivalence and Efficiency of Image Alignment Algorithms