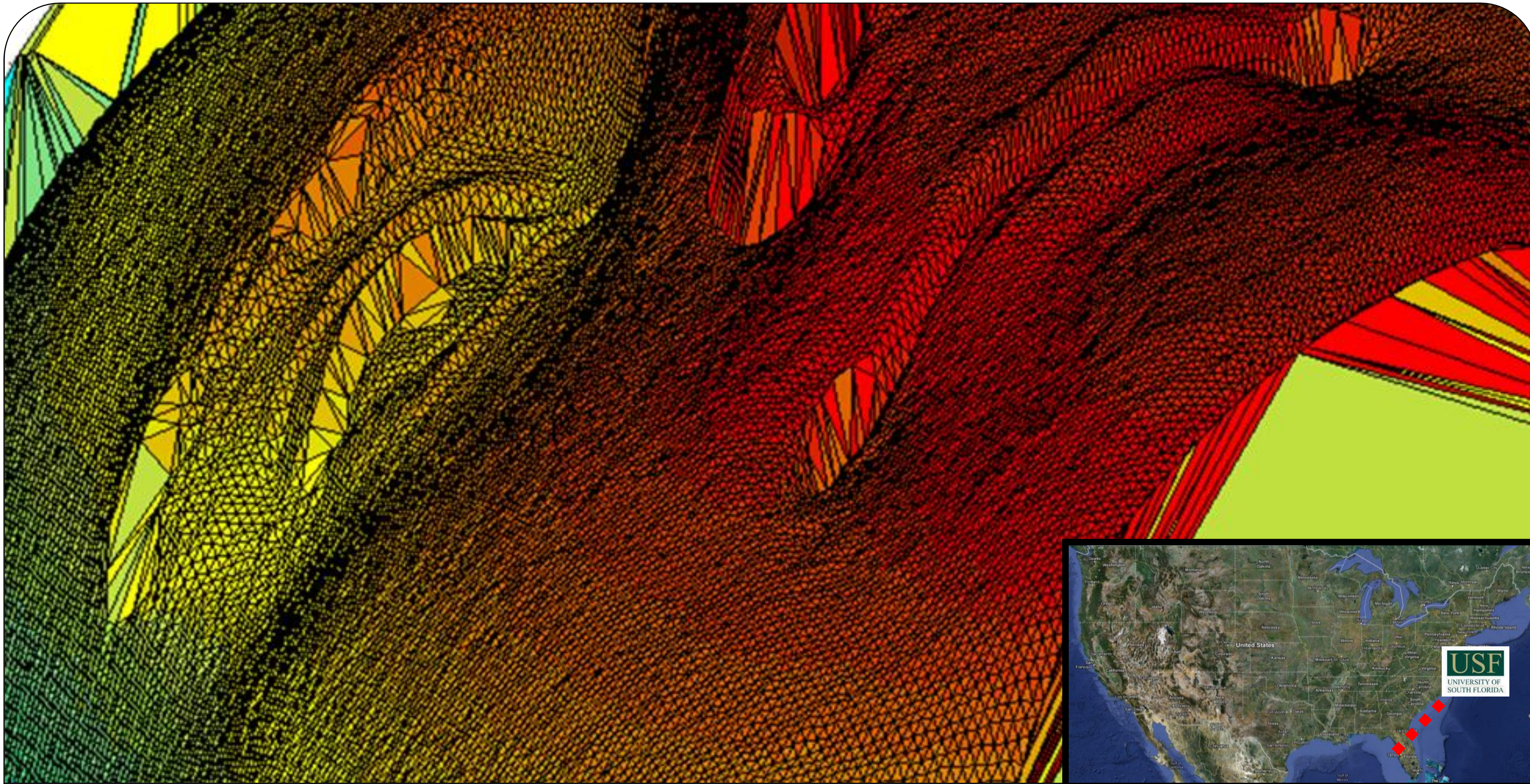


Method for Calculating View-Invariant 3D Optical strain

Matthew Shreve, Sergiy Feflilatyev, Nestor Bonilla, Gerardo Hernandez, Dmitry Goldgof, Sudeep Sarkar

Computer Vision and Pattern Recognition Group

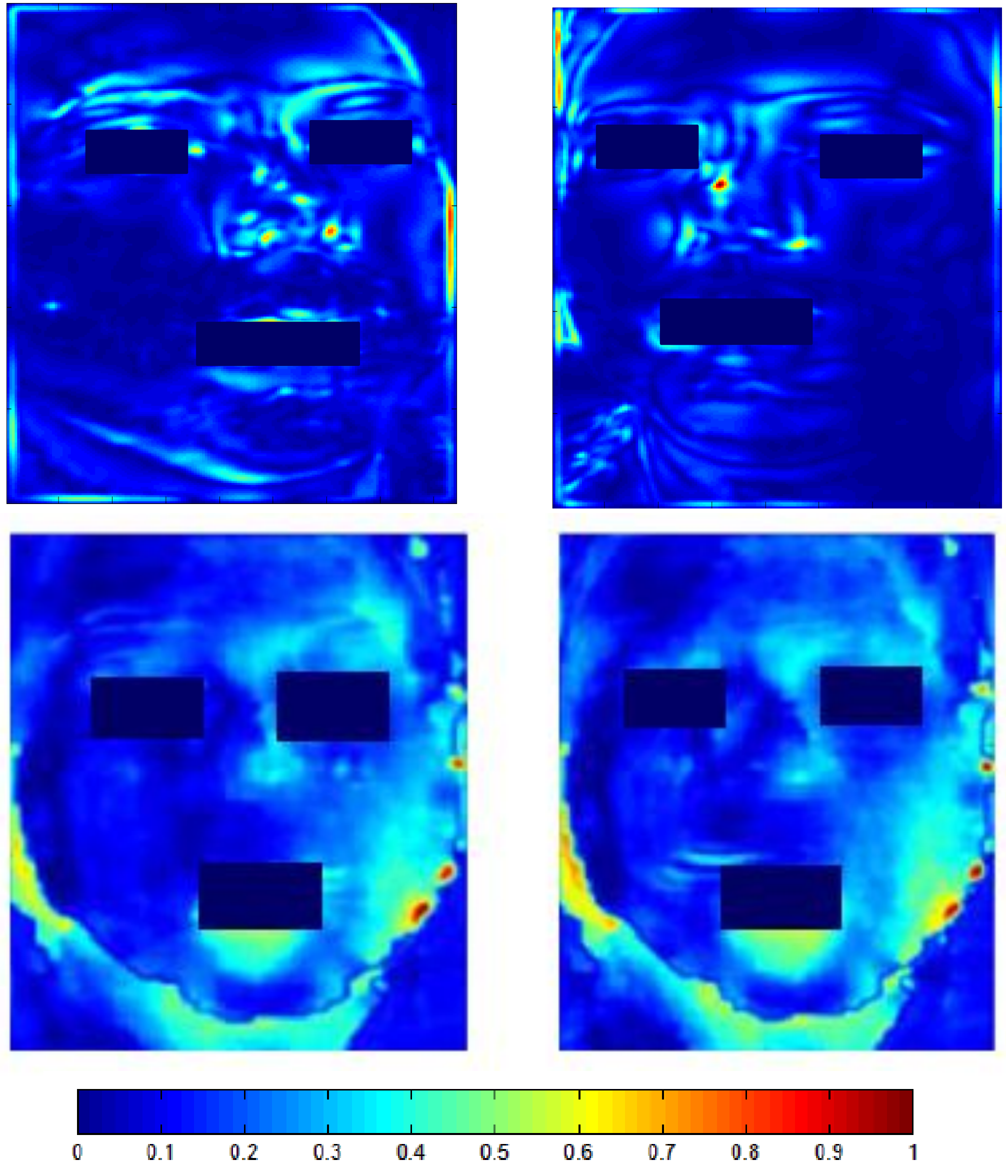
WDIA 2012

Contribution

- We have worked on several applications using 2-D optical strain, however 2-D optical strain is not invariant to view.
- Therefore we attempt to aide this problem by projecting 2-D motion on to a 3-D surface.
- In brief, the correspondence issue is solved using 2-D displacements, which are then updated using rough 3-D estimations.

2-D
Surprise

3-D
Surprise

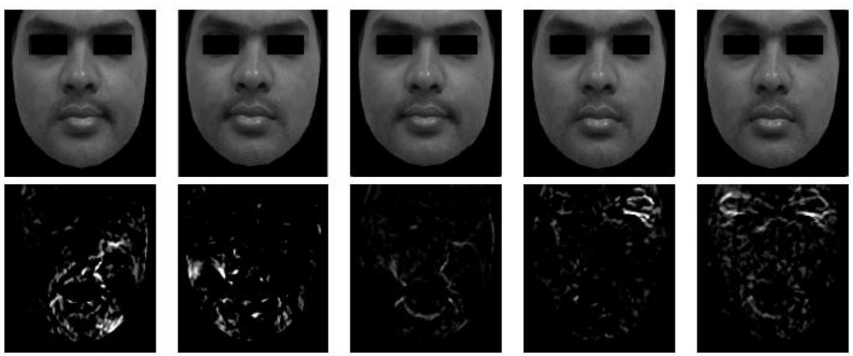


Background

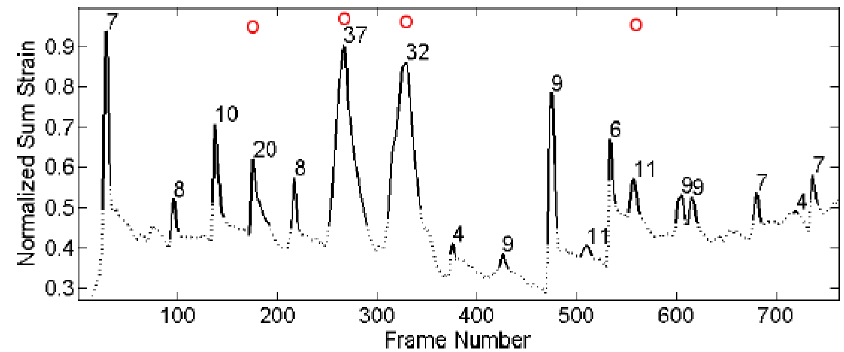
- Optical strain describes the change, or variation, of the motion vectors in a local neighborhood.
- Optical Strain maps (as applied to facial motion analysis) describe a biomechanical property of facial skin tissue.
- Derived from the non-rigid motion which occurs on face during facial expressions

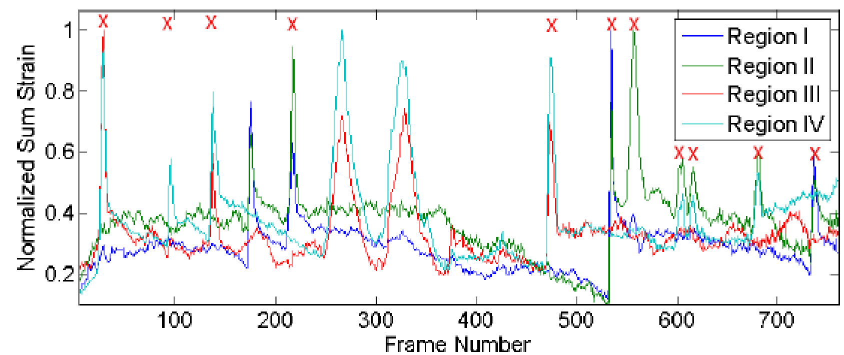
Example Application: Expression Spotting

- Given an input sequence, find the frame boundaries of when expression occur.
- Method is able to identify both macro-expressions (>1/3 second) and micro-expressions (<1/3 second).



(a) (b) (c) (d) (e)





Optical Strain

Given optical flow $\mathbf{u} = [u, v]^T$

Then we can define the finite strain tensor

$$\varepsilon = \frac{1}{2}[\nabla \mathbf{u} + (\nabla \mathbf{u})^T]$$

Which can be expanded to

$$\varepsilon = \begin{bmatrix} \varepsilon_{xx} = \frac{\partial u}{\partial x} & \varepsilon_{yx} = \frac{1}{2}\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) & \varepsilon_{zx} = \frac{1}{2}\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right) \\ \varepsilon_{xy} = \frac{1}{2}\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) & \varepsilon_{yy} = \frac{\partial v}{\partial y} & \varepsilon_{zy} = \frac{1}{2}\left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z}\right) \\ \varepsilon_{xz} = \frac{1}{2}\left(\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z}\right) & \varepsilon_{yz} = \frac{1}{2}\left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z}\right) & \varepsilon_{zz} = \frac{\partial w}{\partial z} \end{bmatrix}$$

And so the strain magnitude is defined as:

$$\varepsilon_m = \sqrt{\varepsilon_{xx}^2 + \varepsilon_{yy}^2 + \varepsilon_{zz} + \varepsilon_{xy}^2 + \varepsilon_{yx}^2 + \varepsilon_{zx}^2 + \varepsilon_{yz}^2}$$

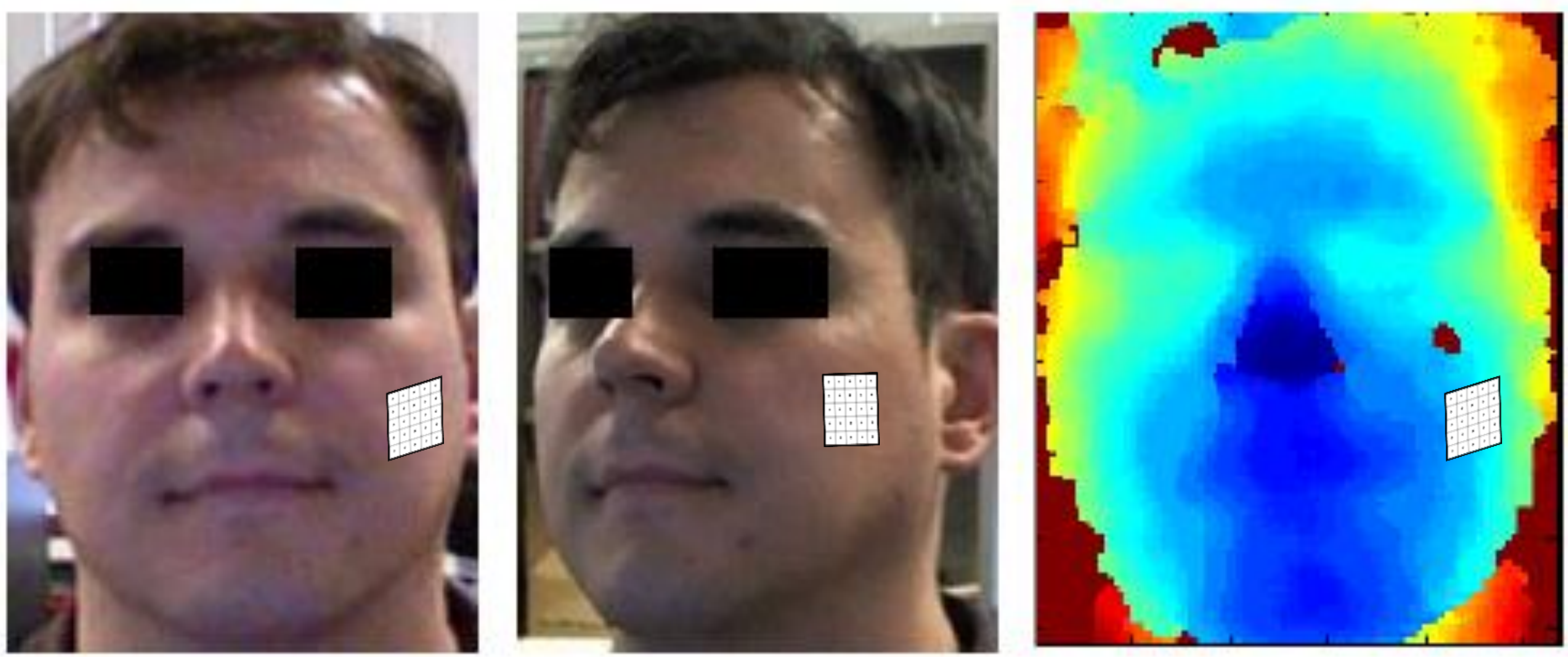
Which can be normalized to 0-255 for visualization purposes:

Optical Strain

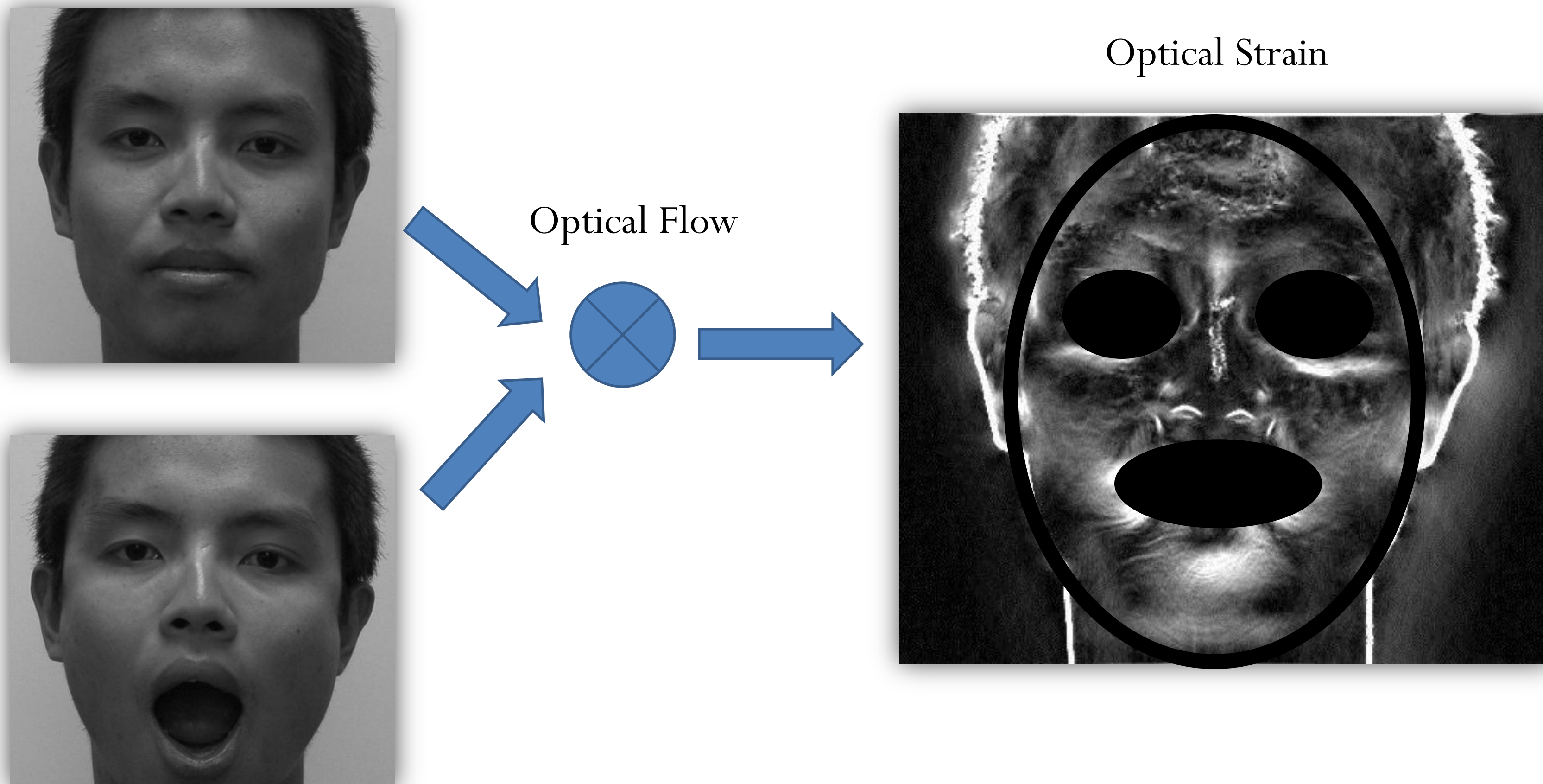
- Use change in 2 depth maps to approximate $\varepsilon_{zz} = \frac{\partial w}{\partial z}$
- Can then use optical flow for u,v approximation, over a planar surface defined by the neighborhood defined by the rectangle (x-r,y-r,x+r,y+r) over w'

- Plane is defined using linear regression over 25 points to solve for a, b, c

$$\mathrm{P} = ai + bj + c$$



Optical Strain

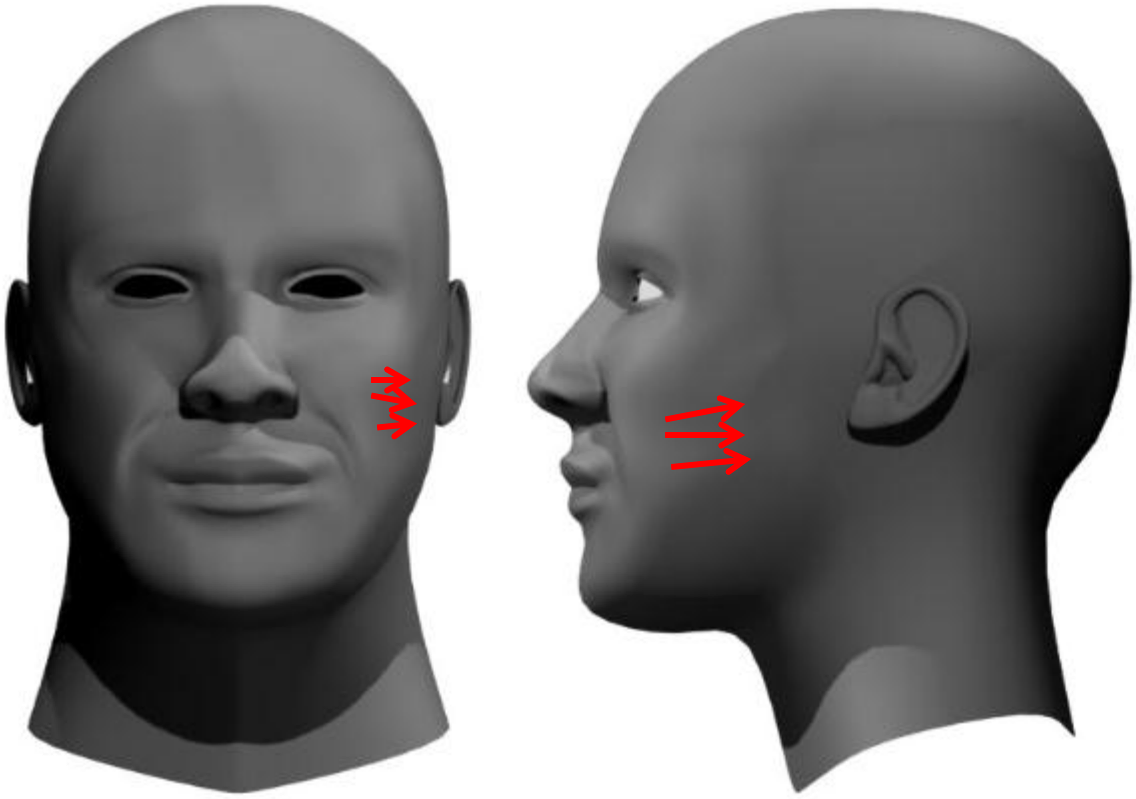


Strain Example

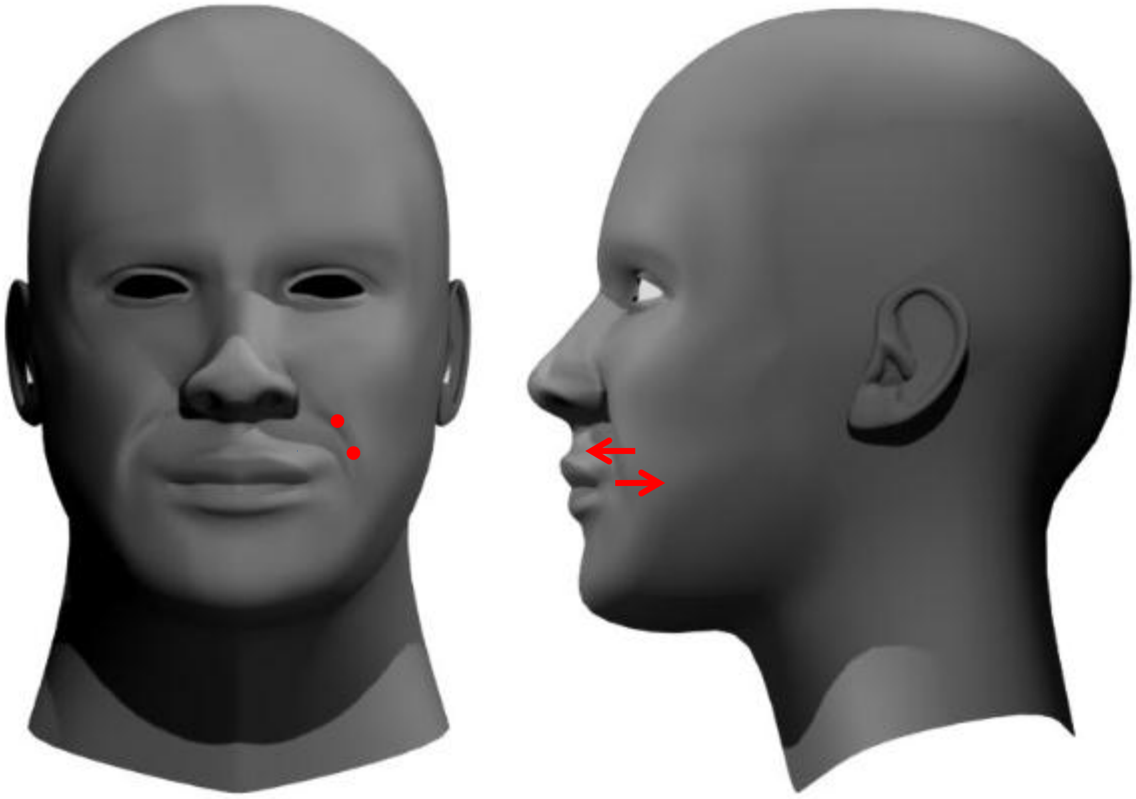
3D Optical Strain

- Intuitively, improvements should be found for:
  - Horizontal motion that occurs along the side of the face. Vectors are often projected as smaller displacement because of parallax.
    - These vectors could be reconstructed using 3D information, which would more accurately match true displacement.
  - Similarly, motion perpendicular to the camera axis lost due to 2D projection.

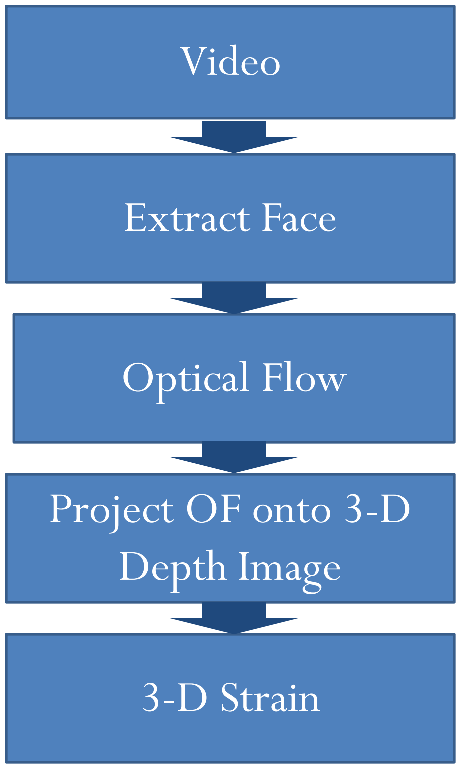
3D Optical Strain



3D Optical Strain



Method



Video of subject's face performing an expression such as smile, surprise

Currently done by manually locating both eyes, but can be automated

Optical flow is calculated between the beginning and peak of the expression

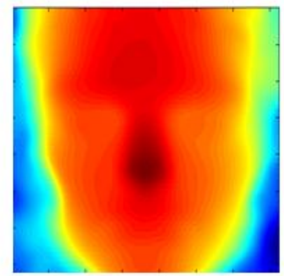
3-D Optical flow is then estimated by projecting the 2-D displacements on to the registered 3-D model.

3-D Strain is then obtained using the central difference method

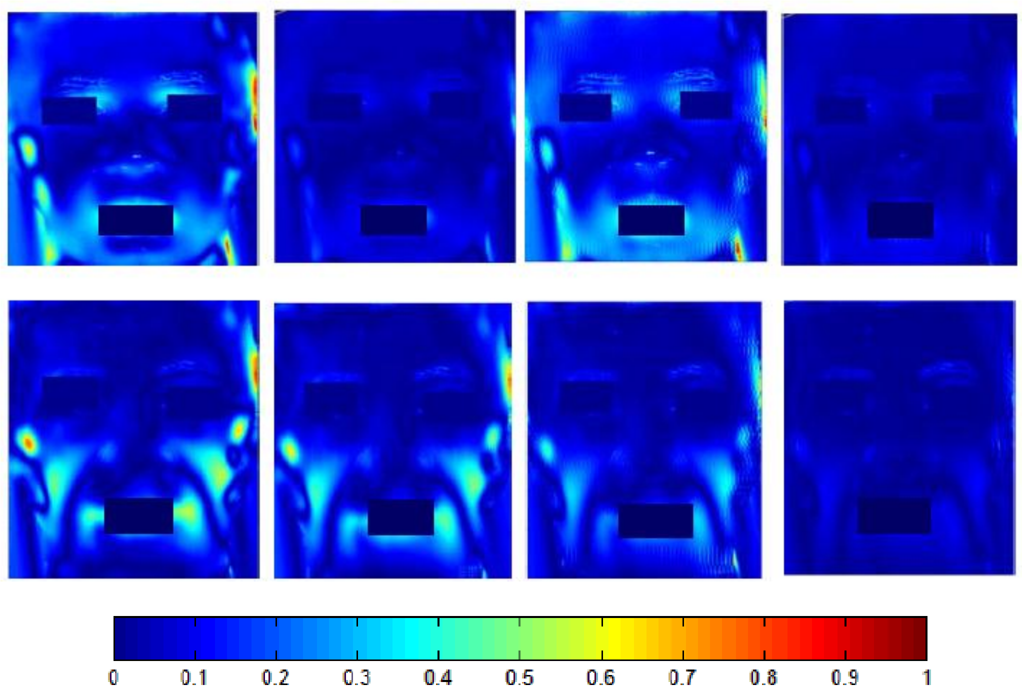
Two Experiments

- Experiment 1 – Performance at multiple depth resolutions
- Experiment 2 – View Invariance

Experiment 1
Performance at multiple depth resolutions



Depth Map



3D Strain maps with depth sampled at 1:1, 1:2, 1:3, 1:4 ratios

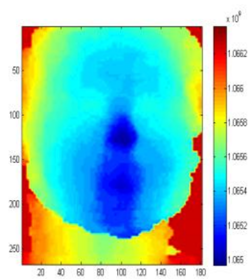
Experiment 1
Performance at multiple depth resolutions

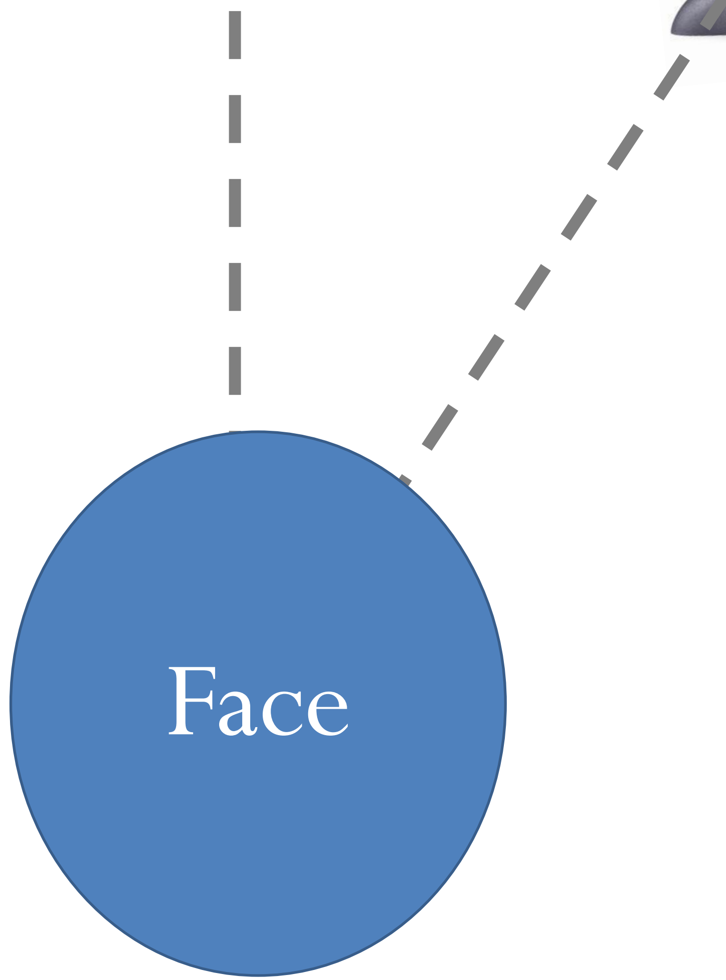
Table 1. Correlation coefficients for 40 expression (20 smile, 20 surprise) after subsampling at the given ratios and compared with a 1:1 sampling ratio.

Exp. / Ratio	**1:2**	**1:3**	**1:4**
Smile	.90±.11	.80±.12	.69±.18
Surprise	.95±.02	.89±.05	.79±.15
Both	.93±.08	.85±.10	.74±.17

KINECT

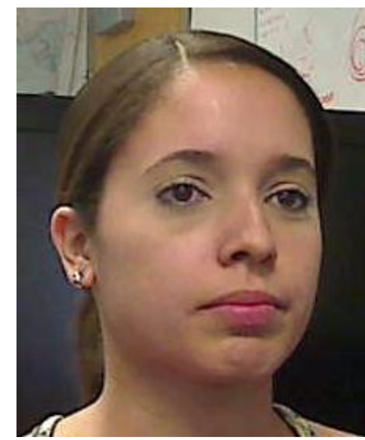
- Low resolution depth (face must be sufficiently distant from camera – 1 meter)
- Poor optics for RGB image for optical flow
- Optical Flow fails



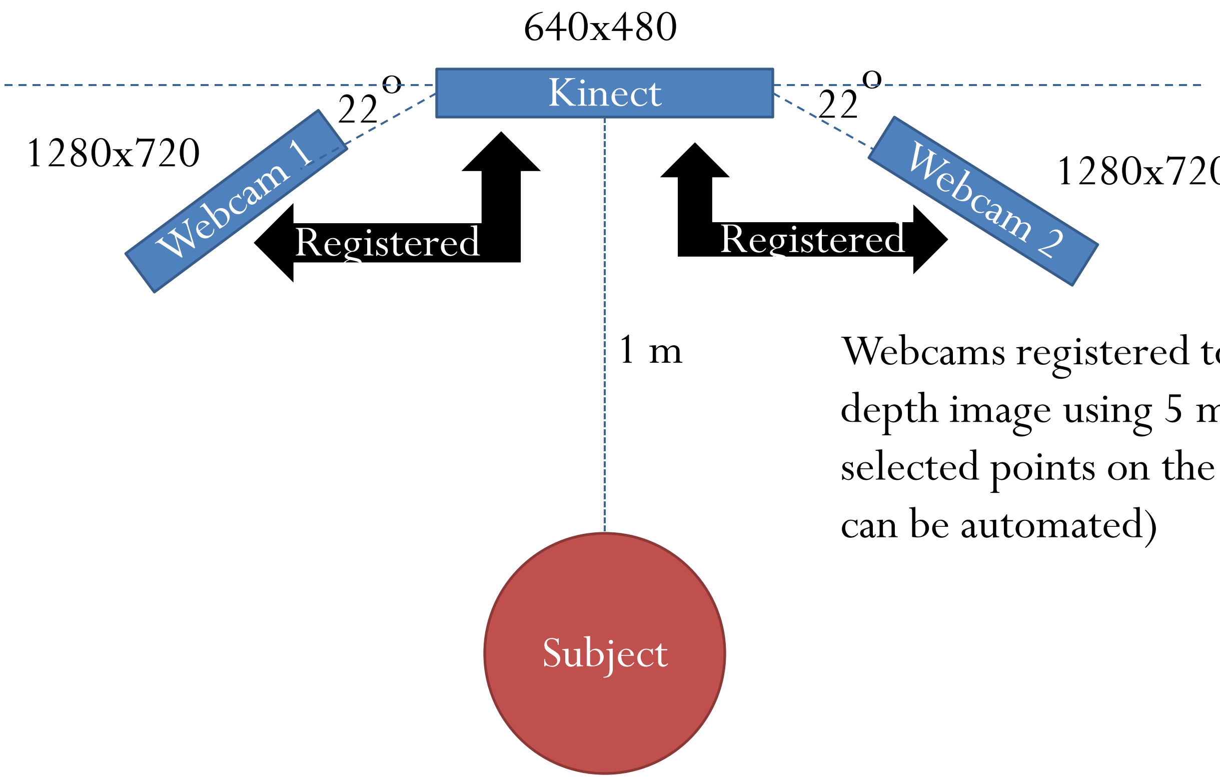


Webcam

- HD Resolution

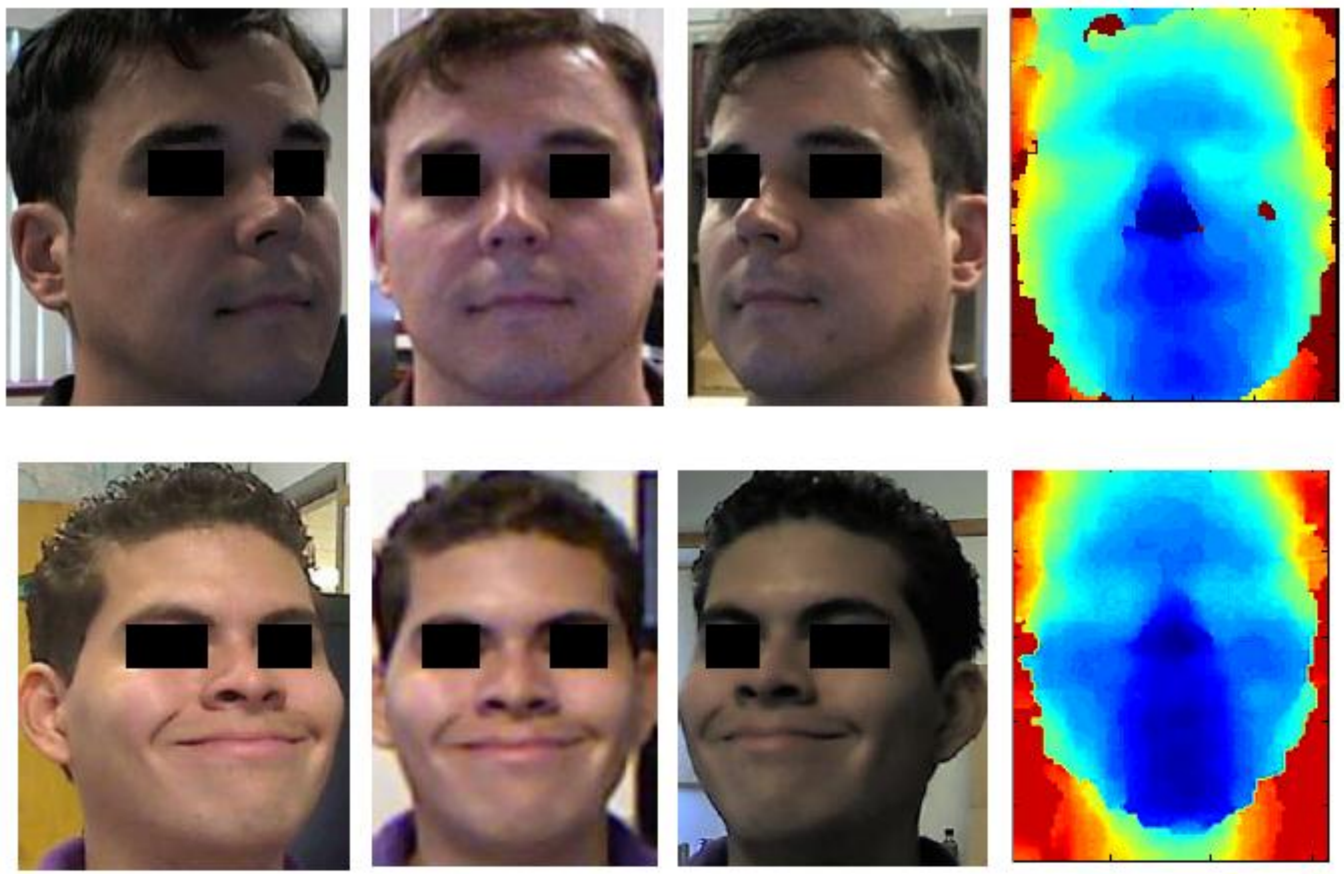


Experiment 2
View Invariance



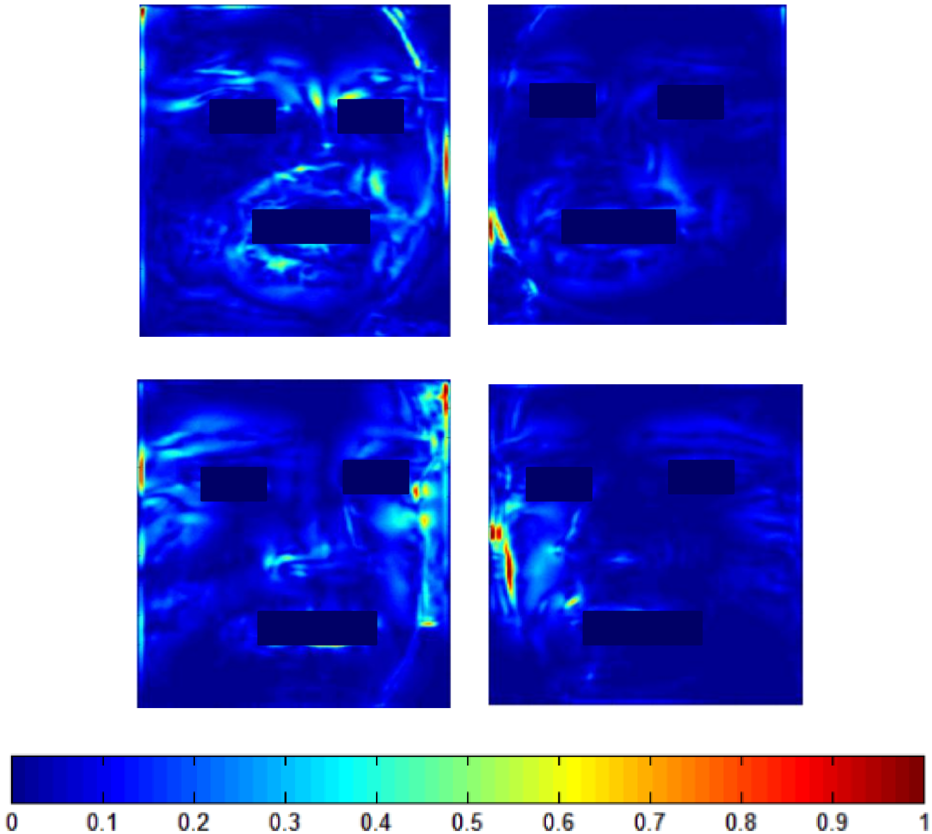
Webcams registered to Kinect depth image using 5 manually selected points on the face (this can be automated)

Experiment 2
View Invariance



Experiment 2
View Invariance

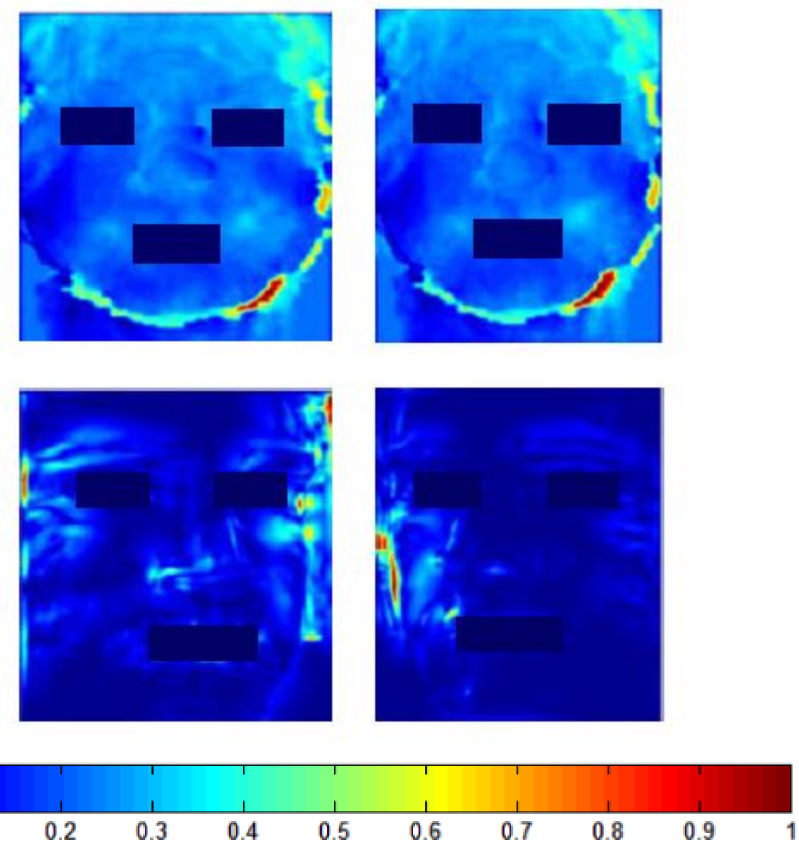
Without using 3-D



Example strain maps calculated at two views roughly 45 degrees apart, for two subjects (each row), without using 3D information. The first two pairs of columns are for the smile expression, the second pair of columns are for the surprise expression.

Experiment 2
View Invariance

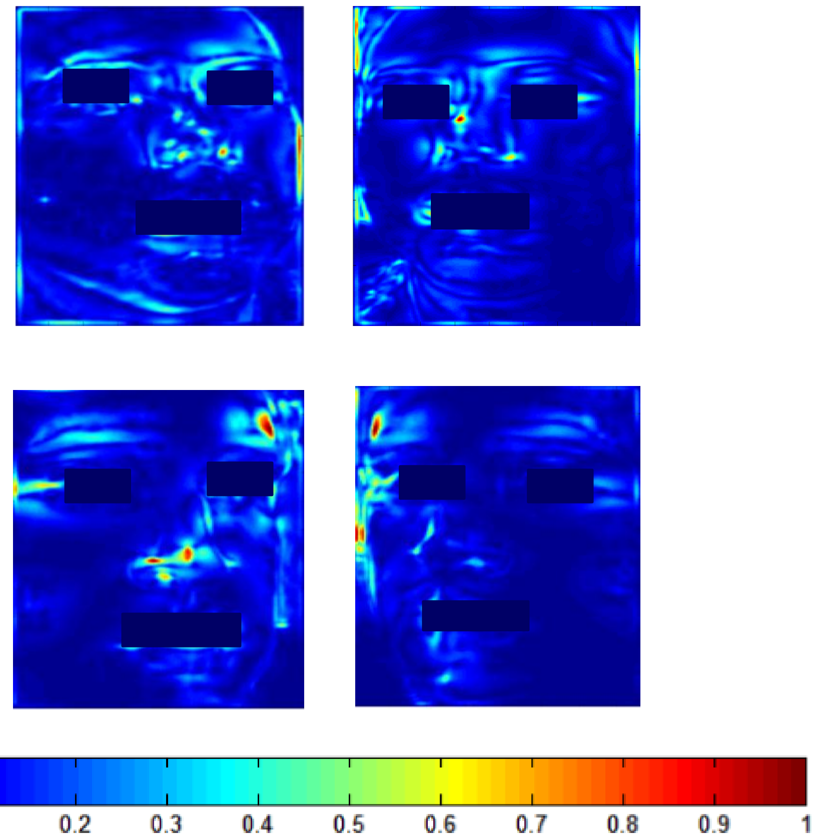
With 3-D



Example strain maps calculated at two views roughly 45 degrees apart, for two subjects (each row). The first two pairs of columns are for the smile expression, the second pair of columns are for the surprise expression.

Experiment 2
View Invariance

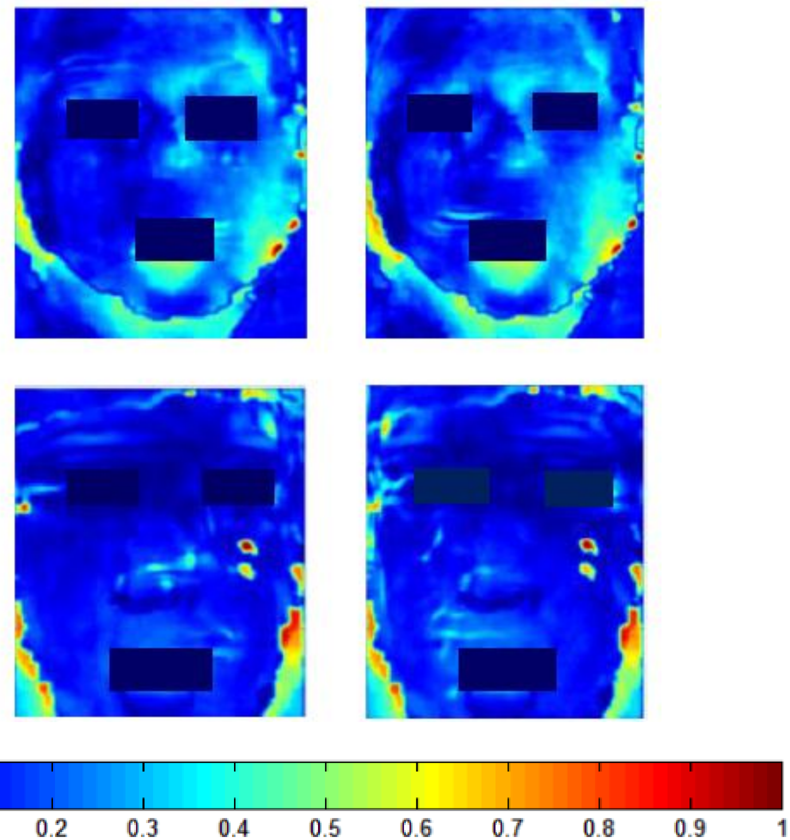
Without using 3-D



Example strain maps calculated at two views roughly 45 degrees apart, for two subjects (each row), without using 3D information. The first two pairs of columns are for the smile expression, the second pair of columns are for the surprise expression.

Experiment 2
View Invariance

With 3-D

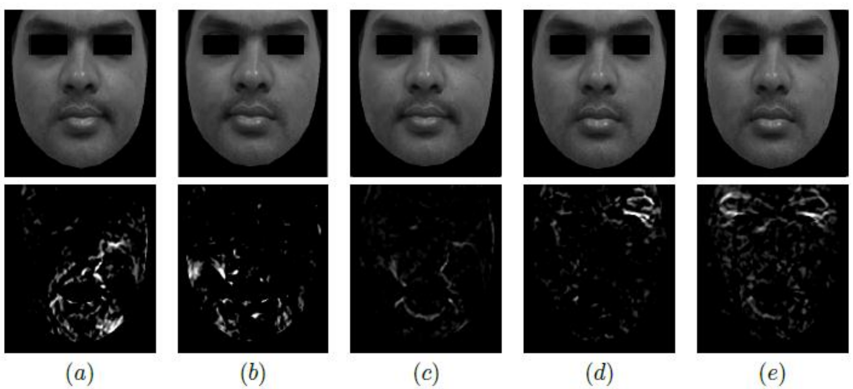


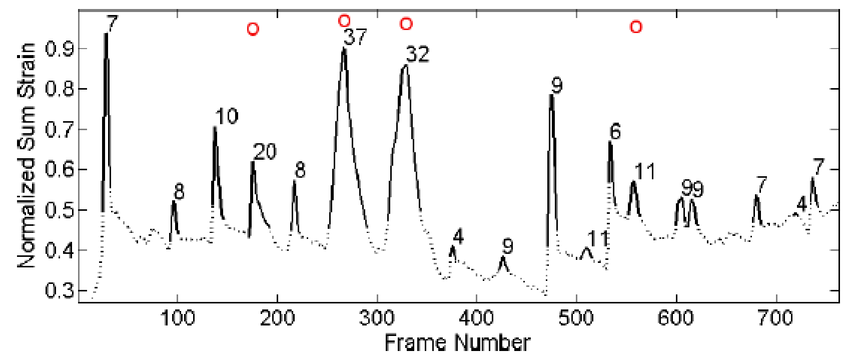
Example strain maps calculated at two views roughly 45 degrees apart, for two subjects (each row). The first two pairs of columns are for the smile expression, the second pair of columns are for the surprise expression.

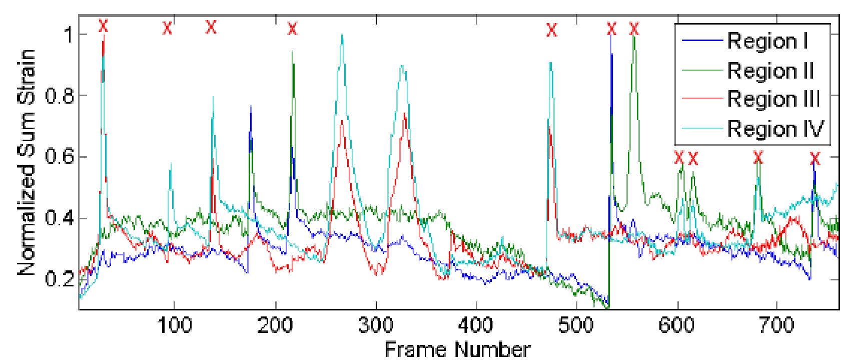
Future Work

Future Work for 3-D Optical Strain: Expression Spotting

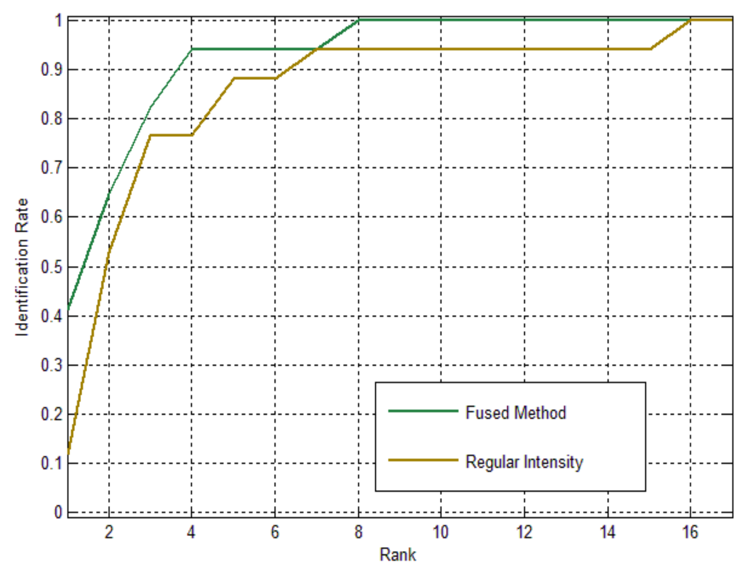
- Given an input sequence, find the frame boundaries of when expression occur.
- Method is able to identify both macro-expressions (>1/3 second) and micro-expressions (<1/3 second).







Future Work for 3-D Optical Strain: Face Identification

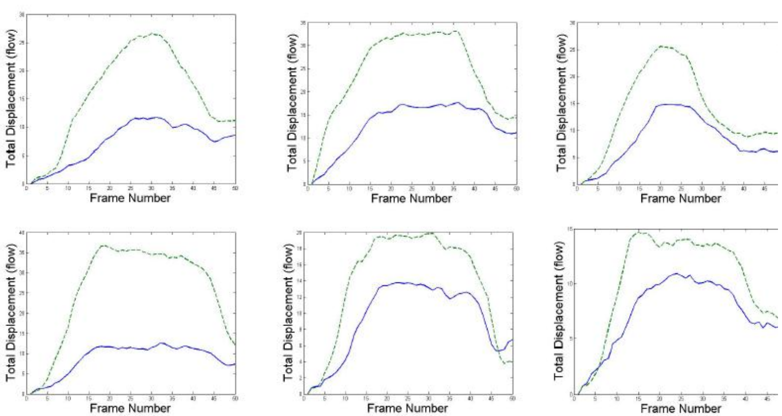


- 30% increase in rank 1 identification
- Average 20% increase in identification rate

Future Work for 3-D Optical Strain: Efficacy of Facial Reconstructive Surgery

- Developed a more rich representation of reconstructive surgery efficacy.
- Reduced the video acquisition time by as much as 5 hours for each subject.

(a) (b) (c)



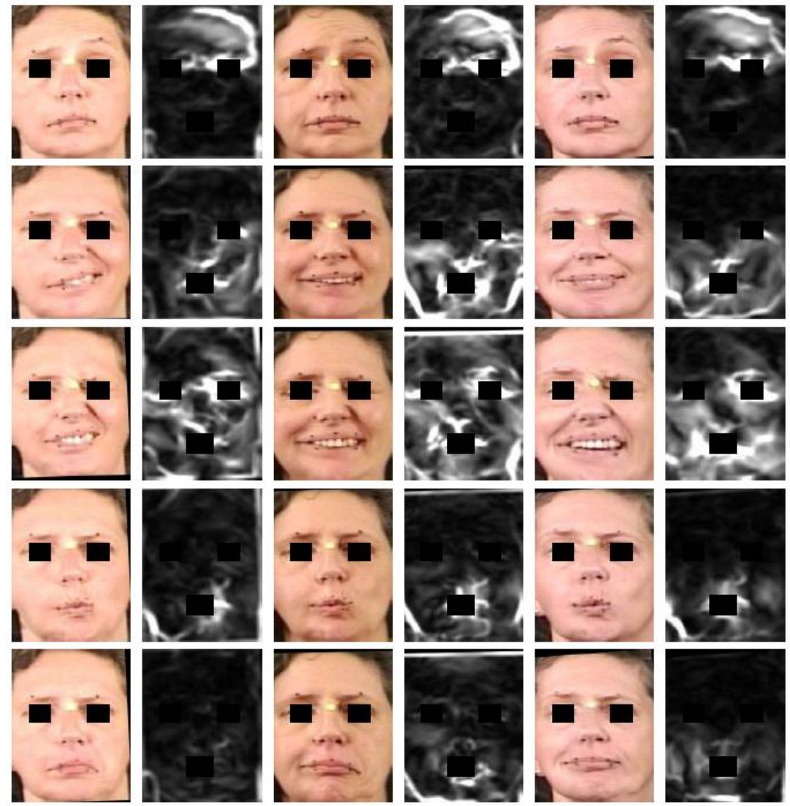


Fig. 3: Optical strain maps for five expressions, over three years. Strain maps were generated between the start and peak of each expression. Intensity values correspond to amount of deformation observed.

Conclusion

- Optical strain maps have broad significance in facial motion analysis.
- We have proposed method for calculating assisting 2-D motion analysis using a rough 3-D range sensor.
- The method has been shown to work at depth resolutions of 100x100 and 66x66 while maintaining at least 80% correlation with full (200x200) resolution.
- We have shown empirically that the strain maps from two views 45 degrees apart are highly similar.

Thanks

Questions?