

# Match Move 2

## Inserting CG Objects into Live Video

School of Computing

## Agenda

- Know your camera
- Track 2D features
- Recover 3D info
- Recover camera motion
- Insert CG objects

CS5245: Vision & Graphics for Special Effects School of Computing

## Where Matchmoving fits in

Excerpted from Matchmoving, by Tim Dobbert

CS:

## Geometry of 2 views

- $p_1, p_2$  are corresponding points.
- A set of corresponding points over time is called a track.

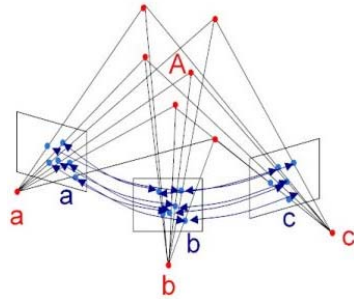
- Camera centre has moved from  $C_1$  to  $C_2$ , via rotation  $R$  and translation  $t$
- 3D scene point  $P$  projects to  $p_1$  in image 1, and  $p_2$  in image 2
- Relationship  $p_2^T F p_1 = 0$
- $F$ : fundamental matrix
- Using multiple scene points, we can solve for  $C_1, C_2, P, F$
- Sequence of  $C_1, C_2, \dots$  is camera motion

Ref: Computer Vision, A Modern Approach by Forsyth & Ponce

CS5245: Vision & Graphics for Special Effects School of Computing

## Structure from Motion

- Structure from Motion is a computer vision technique for recovering 3D motion from a video sequence.
  - 3D coordinates (shape or structure) of scene points
  - Camera parameters (which may change in each frame)
    - Rotation
    - Translation
    - Focal length



CS5245: Vision & Graphics for Special Effects **School of Computing**

## 2D Tracking

- Successful SFM depends on tracking 2D points
  - i.e. establishing point correspondences across video sequence
  - Tracking can be done manually or automatically.

CS5245: Vision & Graphics for Special Effects **School of Computing**



**School of Computing**

## 2D Tracking

- |   |   |
|---|---|
| <ul style="list-style-type: none"> <li>• Manual</li> <li>• Pros:                             <ul style="list-style-type: none"> <li>○ Robust (tolerate changes in illum, repeated structures)</li> </ul> </li> <li>• Cons:                             <ul style="list-style-type: none"> <li>○ Tedious</li> <li>○ May not be accurate</li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>• Automatic</li> <li>• Pros:                             <ul style="list-style-type: none"> <li>○ Easy, fast</li> <li>○ Subpixel accuracy</li> </ul> </li> <li>• Cons:                             <ul style="list-style-type: none"> <li>○ Not robust (fooled by illum variation, repeated structures)</li> </ul> </li> </ul> |
|---|---|

CS5245: Vision & Graphics for Special Effects **School of Computing**

## Tracking tips

- Track stationary objects
- Track true 3D points
- Sample 3D space
- Maintain minimum number of tracks
- Avoid
  - Large change in illum, motion, pose, size
  - Repeated structures
  - Rich texture
  - Featureless regions

CS5245: Vision & Graphics for Special Effects **School of Computing**

## Track stationary objects

- SFM won't work if both camera & object are moving
  - So track background (stationary) objects, e.g. buildings, lamp posts, stationary vehicles
  - Avoid trees, people, animals, clouds, shadows, water



CS5245: Vision & Gr

## Track true 3D points

- Avoid shadows, boundary between objects

Overlap between buildings



CS5245: Vision &

## Sample 3D space

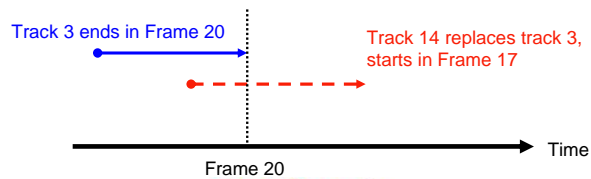
- Tracks should be placed at various heights, depths and widths to get the 3D layout of the scene.



CS5245: Vision & Graphics for Special Effects **School of Computing**

## Minimum # tracks

- Matchmove software typically need 8 to 12 tracks.
- If a track leaves the field of view (or is occluded), then need to introduce a new track to replace it.
  - New track should be introduced BEFORE a track leaves.
  - Not doing so may result in sudden camera jumps.



CS5245: Vision & Graphics for Special Effects **School of Computing**

## Things to avoid

- Large change in illumination, motion, pose, size
  - Tracking algorithms typically assume illumination, motion, pose and size change very little between frames.
- Illum: cloud/sunlight movement in outdoor scenes
- Motion: large translation
  - Solution: track using image pyramid
- Pose: rotation in depth, possibly leading to self-occlusion
- Size: large translation in depth (or zoom)

CS5245: Vision & Graphics for Special Effects **School of Computing**

- Illum change



- Rotation in depth



CS5245: Vision & Graphics for Special Effects **School of Computing**

## Keeping changes small

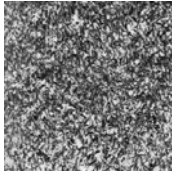
- One general way to minimize large changes in the scene is to increase frame rate.
  - So smaller change between frames
- Most video cams operate at 30 fps.
- Specialty cameras can hit 480 fps @ 4 Mpixels!
  - e.g. Phantom V10
  - Useful for slow motion effect



CS5245: Vision & Graphics for Special Effects **School of Computing**

## Things to avoid

- Repeated structures
- Rich texture
- Featureless region
- Edges (prefer corners)



CS52

ing

## Lucas-Kanade Tracker

[http://www.ri.cmu.edu/projects/project\\_515.html](http://www.ri.cmu.edu/projects/project_515.html)

Consider two images  $I$  and  $J$ :



The object moves from  $\mathbf{x} = (x, y)^T$  to  $\mathbf{x} + \mathbf{d} = (x, y)^T + (\xi, \eta)^T$ .  $\mathbf{d}$  is the displacement. So,

$$J(\mathbf{x} + \mathbf{d}) = I(\mathbf{x}) \quad (15)$$

or

$$J(\mathbf{x}) = I(\mathbf{x} - \mathbf{d}) \quad (16)$$

Due to the presence of noise,  $J(\mathbf{x})$  cannot be exactly equal to  $I(\mathbf{x} - \mathbf{d})$ . So, we compute the difference between them:

$$e = I(\mathbf{x} - \mathbf{d}) - J(\mathbf{x}) \quad (17)$$

CS5245: v

## Lucas-Kanade Tracker

The total difference is the weighted sum-squared difference

$$E = \sum_{\mathbf{x} \in W} w(\mathbf{x}) [I(\mathbf{x} - \mathbf{d}) - J(\mathbf{x})]^2 \quad (18)$$

A good match is found if  $E$  is small. So, we want to find a displacement  $\mathbf{d}$  that minimizes  $E$ .

- This algorithm readily available, e.g. Intel OpenCV
  - <http://www.intel.com/technology/computing/opencv/>

CS5245: Vision & Graphics for Special Effects **School of Computing**

## Making your own features

- Sometimes featureless regions cannot be avoided.
  - e.g. bluescreen
- You can create your own features, then digitally remove them later.
  - Masking tape
    - (Magic Glove)
  - Ping pong, tennis balls
    - (Jurassic Park)



CS5245: Vision & Graphics for Special Effects **School of Computing**

## Tracking Tips

- Practically, tracking is a guided process.
- You mark feature points in selected frames (called keyframes), then let software track in between these frames.
- You may need to manually correct tracks that slip, or introduce new tracks before old ones leave.

CS5245: Vision & Graphics for Special Effects **School of Computing**

## Agenda

- Know your camera
- Track 2D features
- Recover 3D info
- Recover camera motion
- Insert CG objects

CS5245: Vision & Graphics for Special Effects **School of Computing**

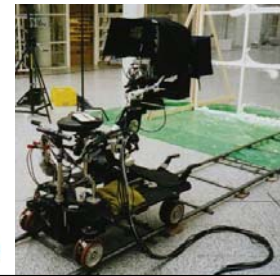
## Recover 3D info, camera param

- This is usually automatically performed by Matchmoving software, once 2D tracking is done.
  - Referred to as "calibration".
- How good is the calibration?
  - Bad calibration means CG won't fit well into live video.
  - See supplementary file: good\_calib.pdf

CS5245: Vision & Graphics for Special Effects **School of Computing**

## Types of camera moves

- Dolly/crane
  - Camera mounted on track (dolly), crane, or motion control rig
  - Usually this means translation (no rotation)

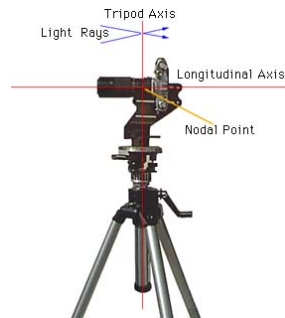


CS5245: Vision & Graphics for Special Effects **School of Computing**

## Types of camera moves

### ● Pan

- Horizontal rotation
- *Nodal pan* rotates about the optical centre → pure rotation
- Non-nodal pan will have translation as well



CS5245: Vision & Graphics for Special Effects **School of Computing**

## Types of camera moves

### ● Slight dolly

- Small translation
- Can be mistaken for zoom or rotation

### ● Lock-off

- Means “no camera motion”
- In reality there will be slight motion (few pixels), so alignment of images may be necessary.

### ● Zoom

- Change in focal length
- Some matchmove programs can't handle this well

CS5245: Vision & Graphics for Special Effects **School of Computing**

## Types of camera moves

### ● Handheld / shaky footage

- Can lead to motion blur
- Result in jerky camera motion
- May need to stabilize video sequence before matchmove

### ● Shaky footage may be *deliberate*

- For artistic effect
- So you will need to de-stabilize after matchmove / compositing

CS5245: Vision & Graphics for Special Effects **School of Computing**

## Inserting CG objects

### ● After matchmoving, you will have 3D camera motion info + 3D scene points.

### ● To insert CG objects, you typically need to:

- Re-create 3D model of scene in CG
- Match scale (set fitting)
- Re-create camera motion in CG
- Render your CG objects according to camera viewpoint
- Composite with live video

### ● Last 2 steps called *matchamation*

CS5245: Vision & Graphics for Special Effects **School of Computing**

## Re-create 3D world in CG

- Low-res proxy objects



CS5245: Vision & Graphics for Special Effects **School of Computing**

## Set fitting

- Recall that 3D info recovered only up to unknown scale.
- Match scale by measuring actual distances in scene.
  - Min. 3 orthogonal distances, but more is better
    - Length, breadth, height
- There is another ambiguity:
  - What is it?
  - How to overcome it?



CS5245: Vision & Graph


## Matchmation

- Matching CG objects onto scene
  - CG objects must “touch” real objects
- Requires lots of manual tweaking
  - Multiple passes may be necessary
  - May need to morph image/objects
  - e.g. Forrest Gump with JFK



CS5245: Vision & Graphics for Special Effects **School of Computing**

## Examples

- Making of Magic Glove 
- Michael Jackson Black or White video
  - (search YouTube)
- Feather in Forrest Gump 
- BMW/Sprint commercial  
[http://www.theorphanage.com/ocp/ocp/qt\\_popup/5/1006/video\\_small](http://www.theorphanage.com/ocp/ocp/qt_popup/5/1006/video_small)

CS5245: Vision & Graphics for Special Effects **School of Computing**