Leow Wee Kheng

CS4243 Computer Vision and Pattern Recognition

Background Removal

Here's an image…

- We often just want the eagle

Background Removal

- Related to tracking and segmentation
  - Tracking

    Tracks location of moving object in video.
  - Segmentation

    Separate object and background in single image.
  - Background removal

    Separate object and background given > 1 image.

Background Removal

- Two general approaches:
  - With known background, also called clean plate.
  - Without known background.

With Clean Plate

- Clean plate: background only image

- Subtract clean plate P from image I

$$D(x,y) = \left| I(x,y) - P(x,y) \right|$$

absolute difference

- Colour image has 3 components
  - R: red, G: green, B: blue
  - So, get 3 sets of differences

$$D_R(x,y) = \left| I_R(x,y) - P_R(x,y) \right|$$

$$D_G(x,y) = \left| I_G(x,y) - P_G(x,y) \right|$$

$$D_B(x,y) = \left| I_B(x,y) - P_B(x,y) \right|$$

- Combine 3 sets of differences into 1 set

$$D(x,y) = \alpha_R D_R(x,y) + \alpha_G D_G(x,y) + \alpha_B D_B(x,y)$$

  - α_R, α_G, α_B are constant weights.
  - Usually, $\alpha_R + \alpha_G + \alpha_B = 1$.
  - In the case of equal weights, $\alpha_R = \alpha_G = \alpha_B = 1/3$.

absolute difference

- Finally, fill in foreground object colour

$$F(x, y) = \begin{cases} I(x, y) & \text{if } D(x, y) > \Gamma \\ B & \text{otherwise} \end{cases}$$

  - Γ is threshold.
  - If $D(x, y) > \Gamma$, pixel at (x, y) is foreground pixel.
  - B is constant background colour, e.g., black.

absolute difference

Notice

- Some parts of the eagle’s tail are missing. Why?

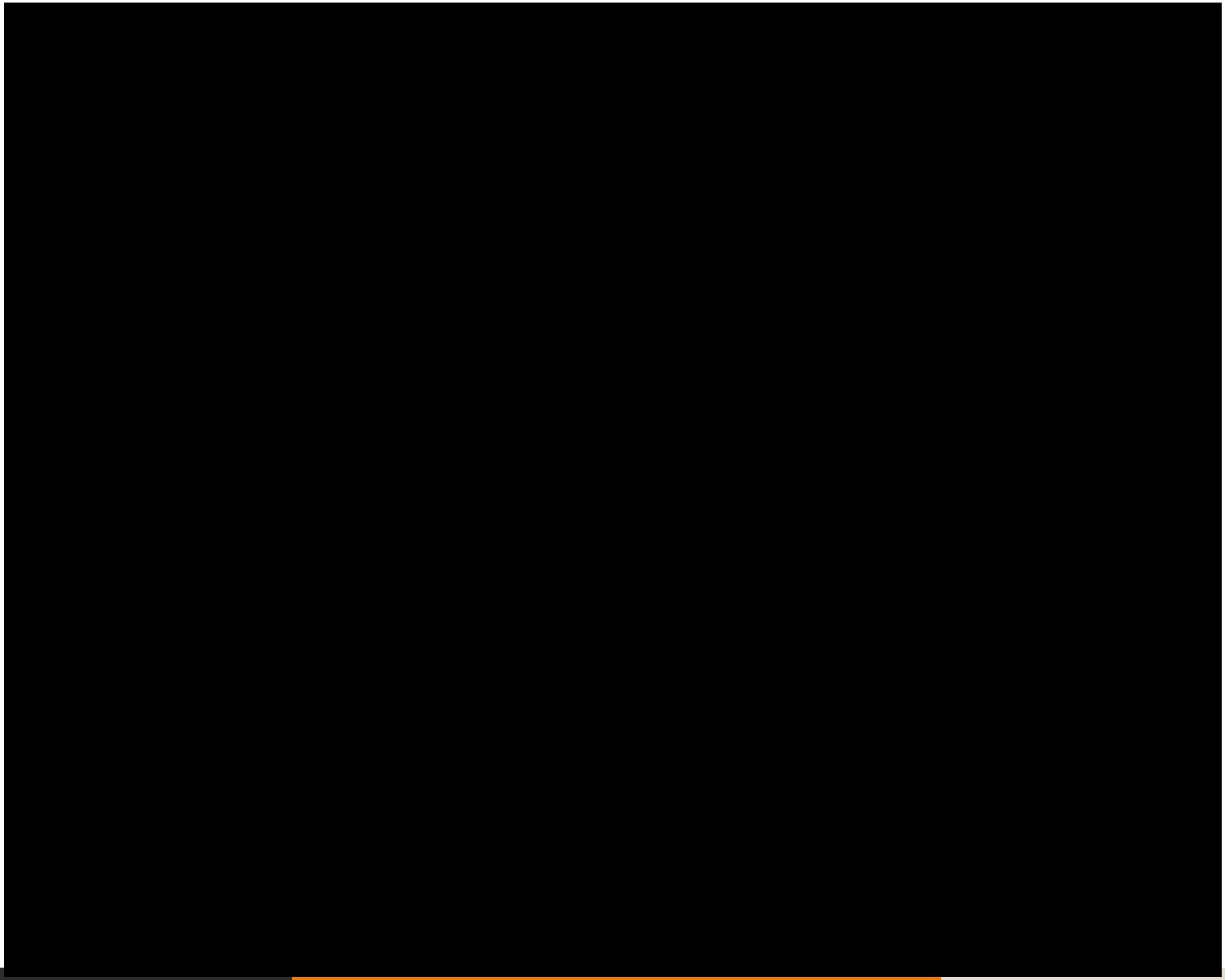
Dynamic Clean Plate

- Stationary camera
  - Stationary background.
  - Need only one image as clean plate.
- Moving camera
  - Moving background.
  - Need a video clean plate.
  - With motion-controlled camera, controlled lighting
    - Shoot clean plate video.
    - Shoot target video with same camera motion.
  - Remove background with corresponding clean plate.

clean plate

scene video

background removed



Without Clean Plate

- Background removal without clean plate is more difficult.
- Possible if moving objects do not occupy the same position all the time.
- 3 cases
  - Stationary camera, fixed lighting.
  - Stationary camera, varying lighting.
  - Moving camera.

Stationary Camera, Fixed Lighting

- Consider these video frames:

  - Moving object occupies a small area.
  - Moving object does not occupy the same position.
  - What if we average the video frames?

Averaging

- Mean of video frame

$$M(x, y) = \frac{1}{n} \sum_i I_i(x, y)$$

  - i : frame number
  - n : number of frames

Notes:

- The above direct formula can lead to overflow error.
- Refer to colour.pdf for a better formula.

Case 1: average over whole video

- Averaging gives mostly background colours.
- Some faint foreground colours remain.

Case 2: average over first 3 seconds

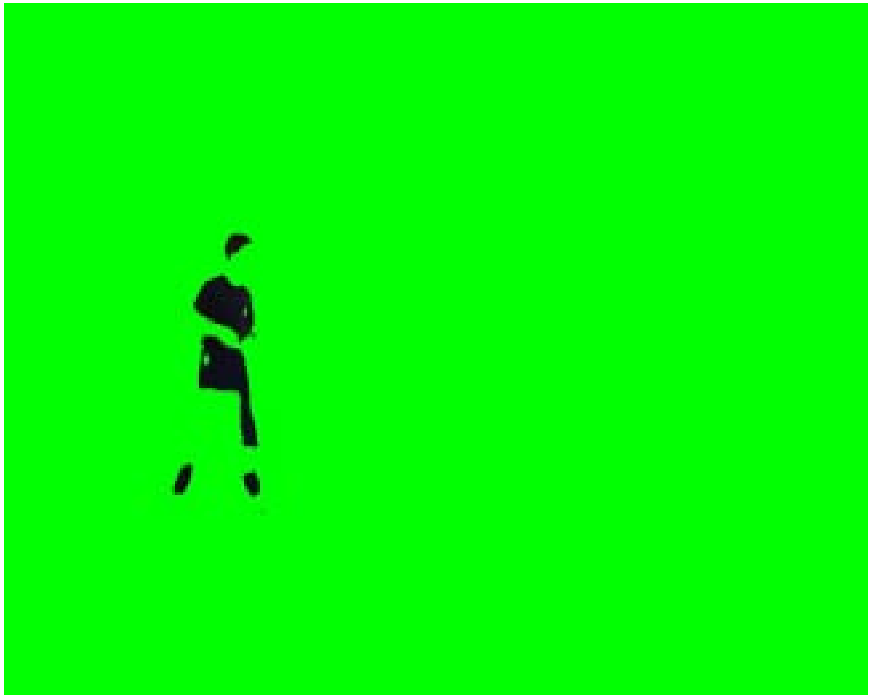
- Foreground colours are more localised in one region.
- Foreground colours are stronger.

Subtract background from video frame

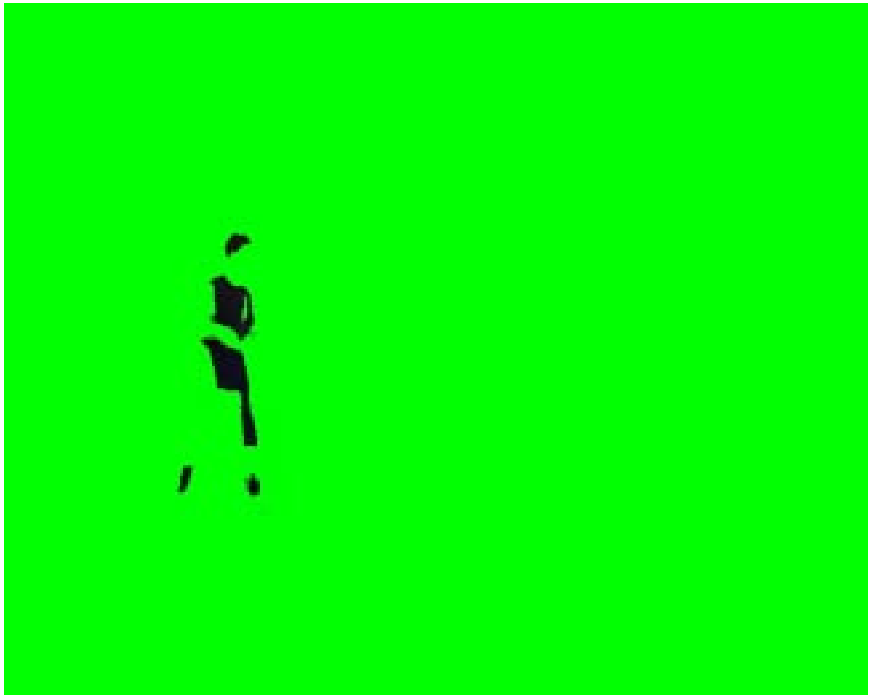
Case 1

Case 2

Copy foreground colours to foreground pixels



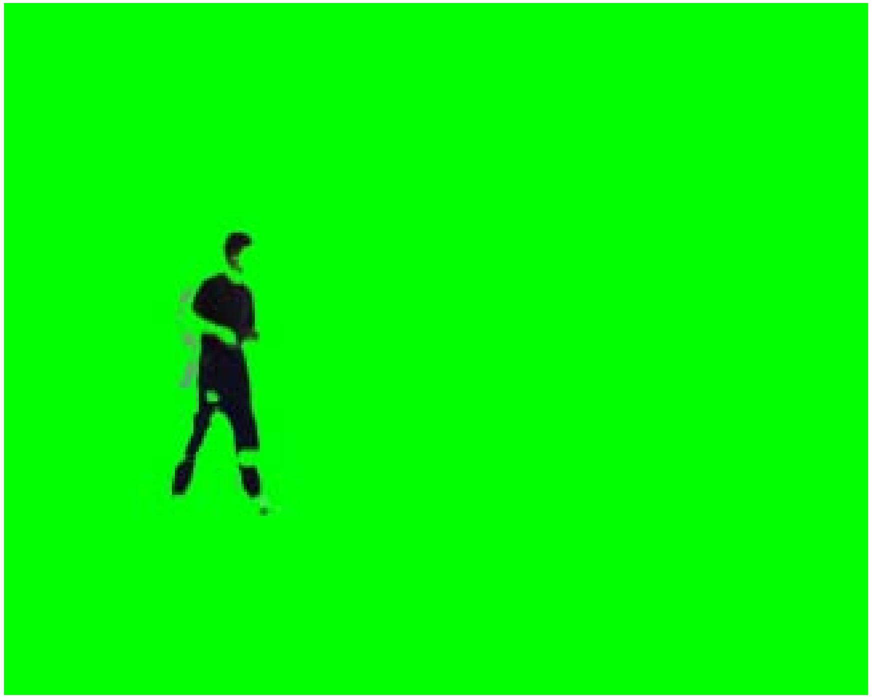
Case 1



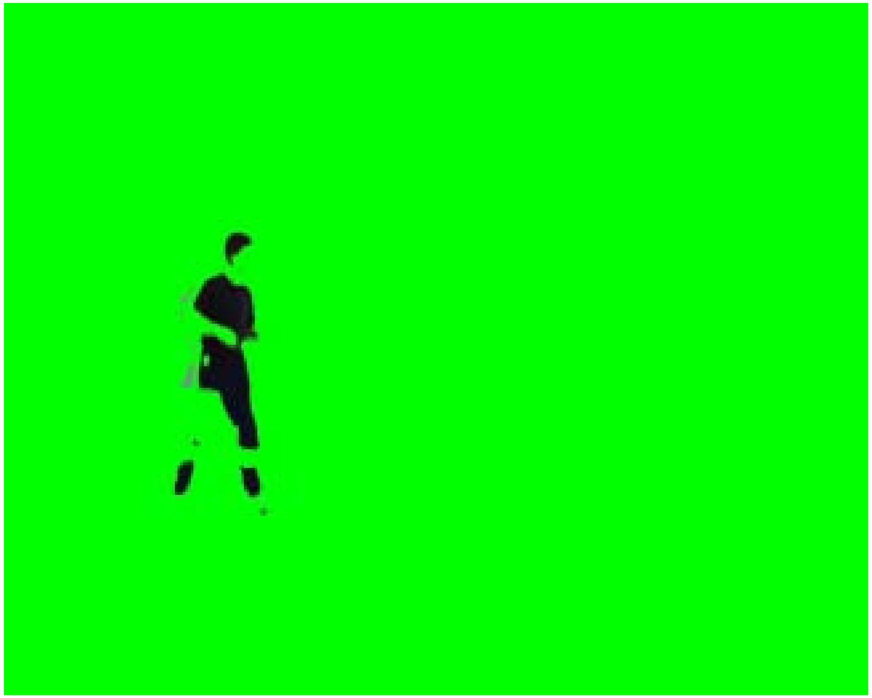
Case 2

- Background colours are removed: true rejection.
- Some foreground colours are missing: false rejection.

Use lower thresholds



Case 1



Case 2

- More foreground colours are found: true acceptance.
- Background colours are also found: false acceptance.

Another example

Averaging video frames

Case 1: over whole video

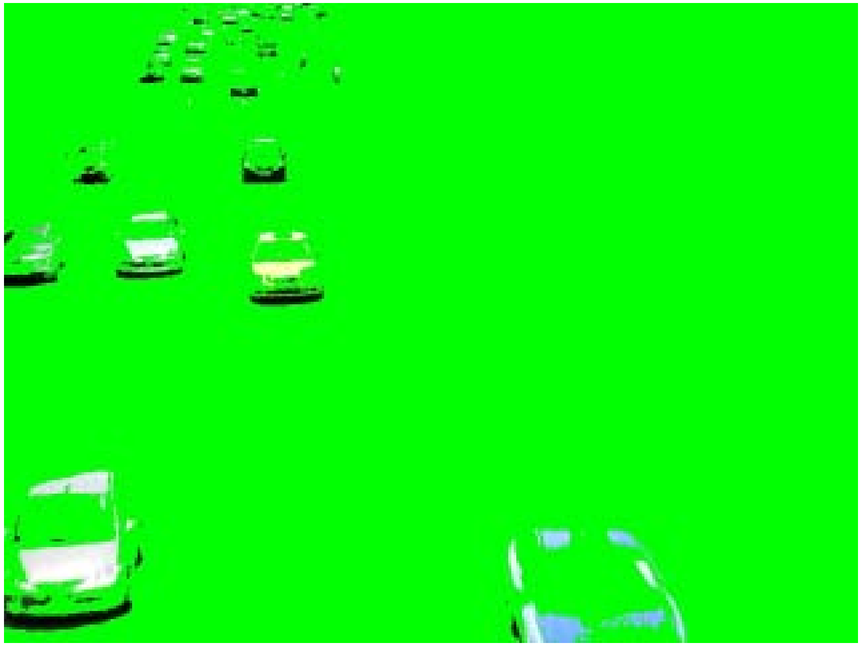
Case 2: over first 3 seconds

Subtract background from video frame

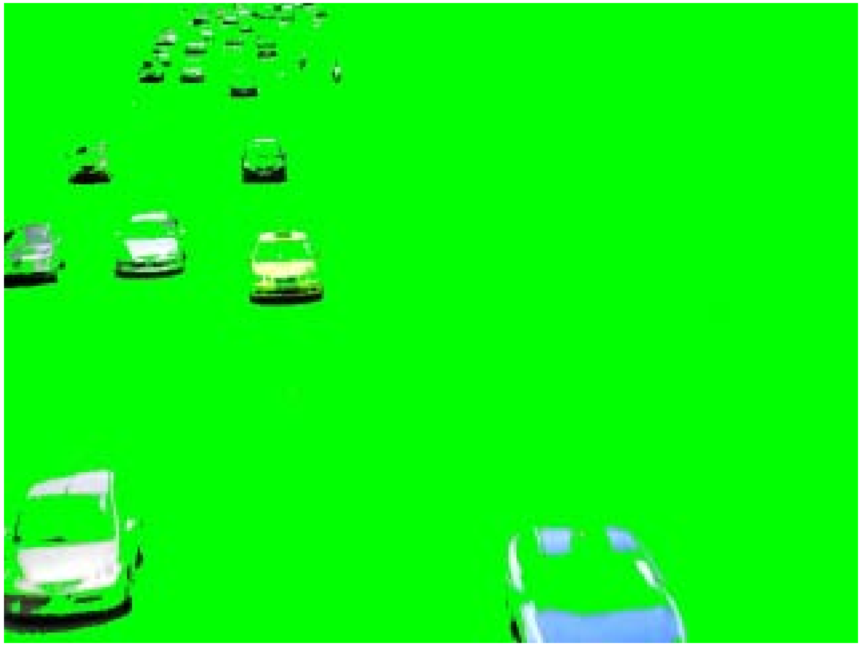
Case 1

Case 2

Copy foreground colours to foreground pixels



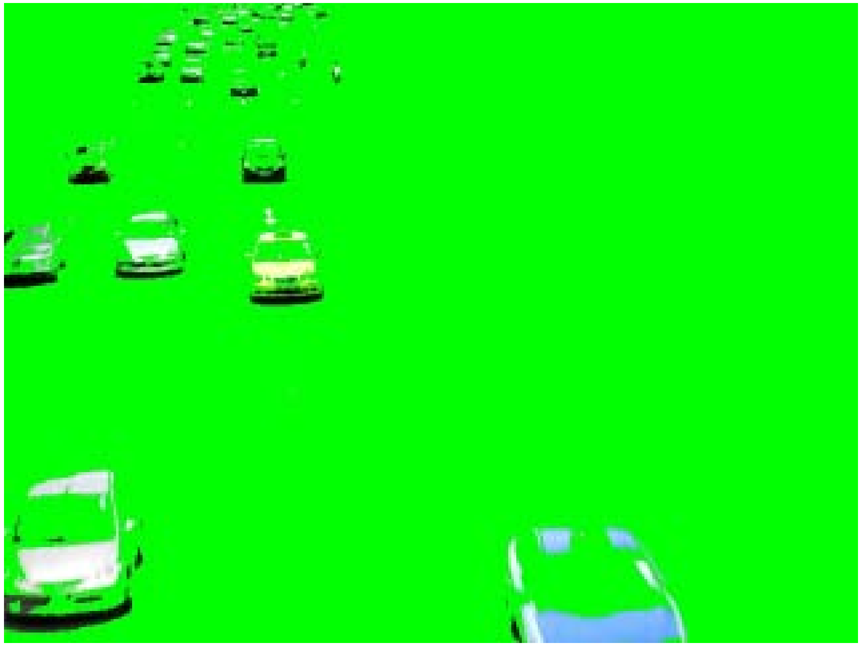
Case 1



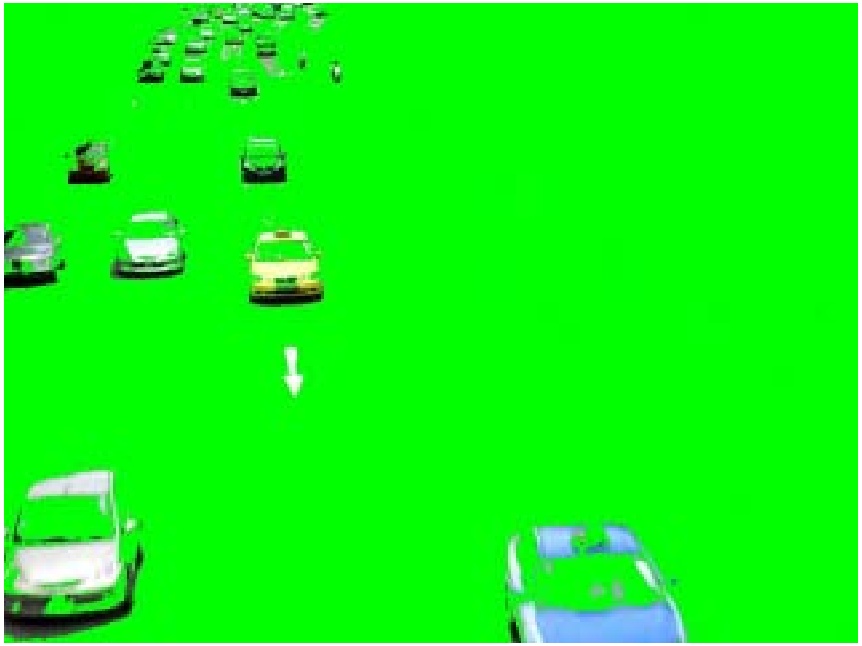
Case 2

- Background colours are removed: true rejection.
- Some foreground colours are missing: false rejection.

Use lower thresholds



Case 1



Case 2

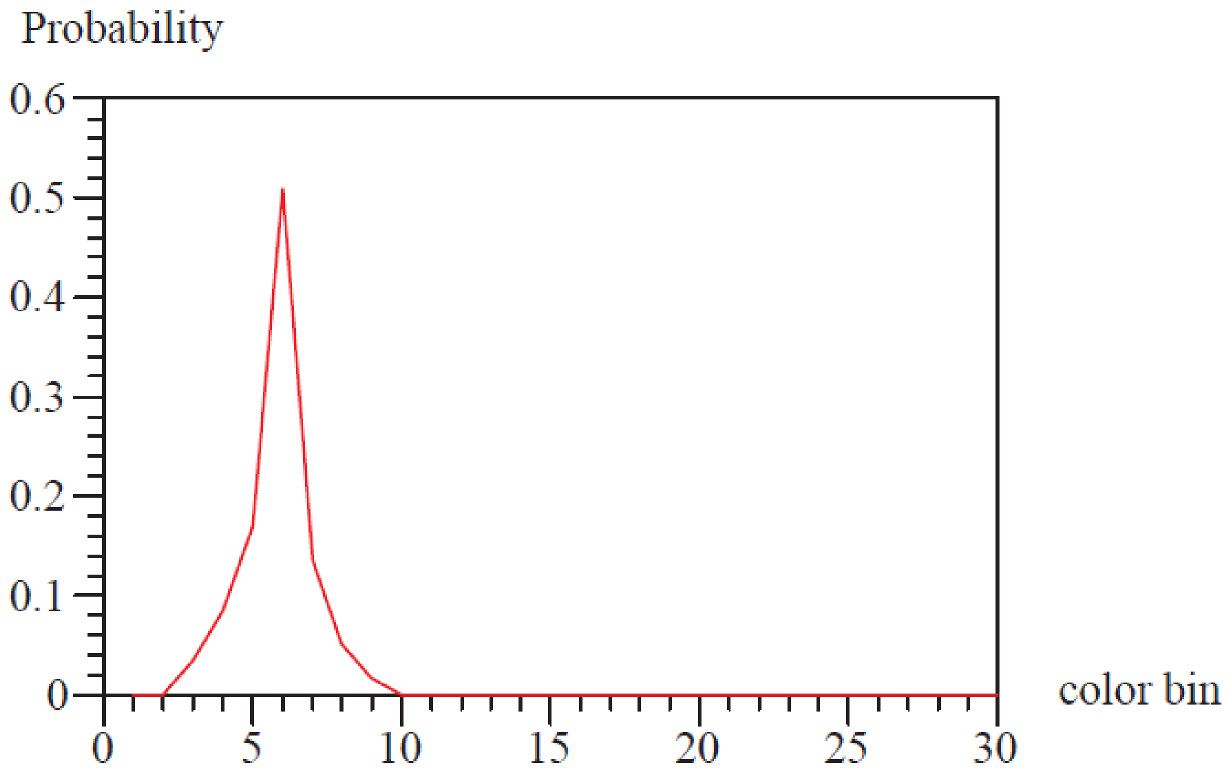
- More foreground colours are found: true acceptance.
- Background colours are also found: false acceptance.

Background Modelling

- Averaging is simple and fast but not perfect.
- Better than average: colour distribution.
  - For each pixel location,

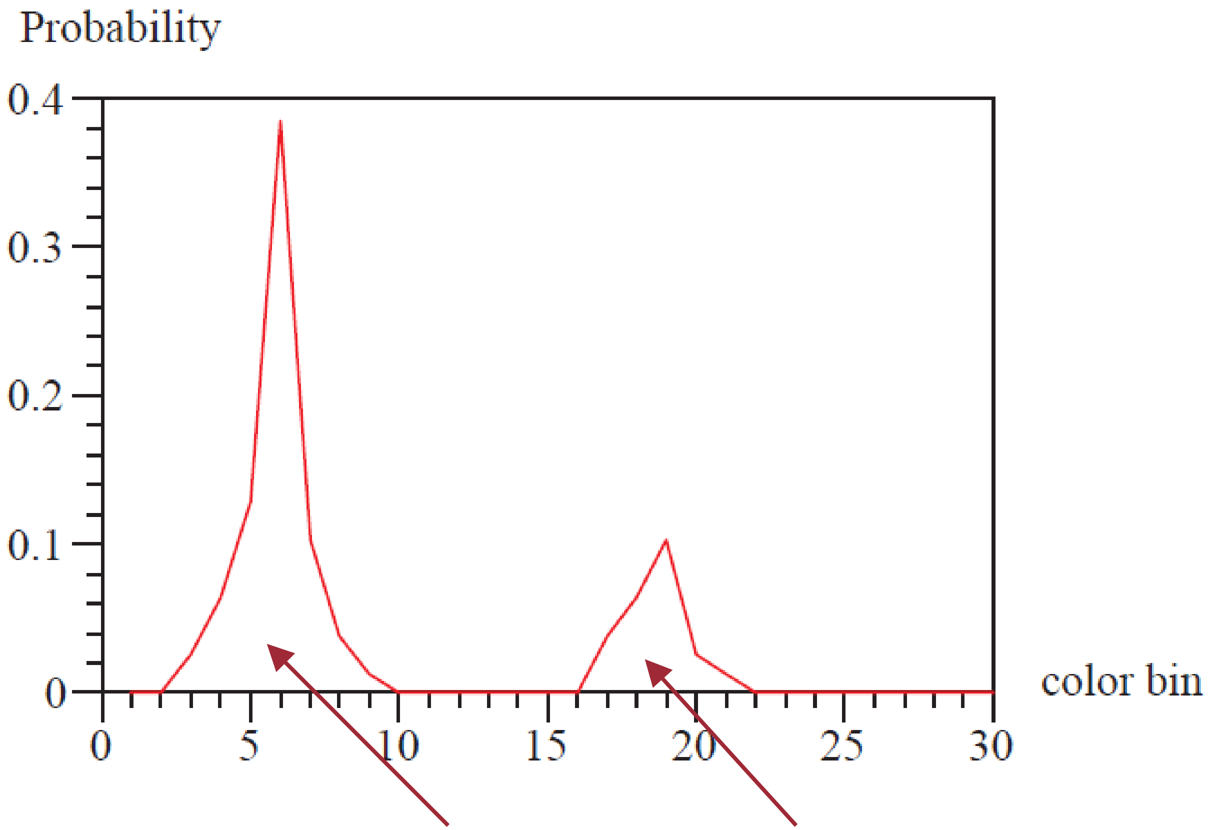
    compute distribution of colours over whole video.

- For a background pixel:



  - Single cluster of colours (due to random variation).
  - Peak: most frequent colour.

- For a pixel that is background most of the time:



  - Two clusters: background, foreground.
  - Relative height: duration covered by foreground.

k-means clustering

- A method for grouping data points into clusters.
- Represent each cluster C_i by a cluster centre $\mathbf{w}_i$.
- Repeatedly distribute data points and update cluster centres.

k-means clustering

1. Choose k initial cluster centres $\mathbf{w}_1(0), \ldots, \mathbf{w}_k(0)$.
2. Repeat until convergence
   - Distribute each colour $\mathbf{x}$ to the nearest cluster $C_i(t)$

$$\mathbf{x} \in C_i(t) \quad \text{if } \|\mathbf{x} - \mathbf{w}_i\| < \|\mathbf{x} - \mathbf{w}_j\| \ \forall j \neq i$$

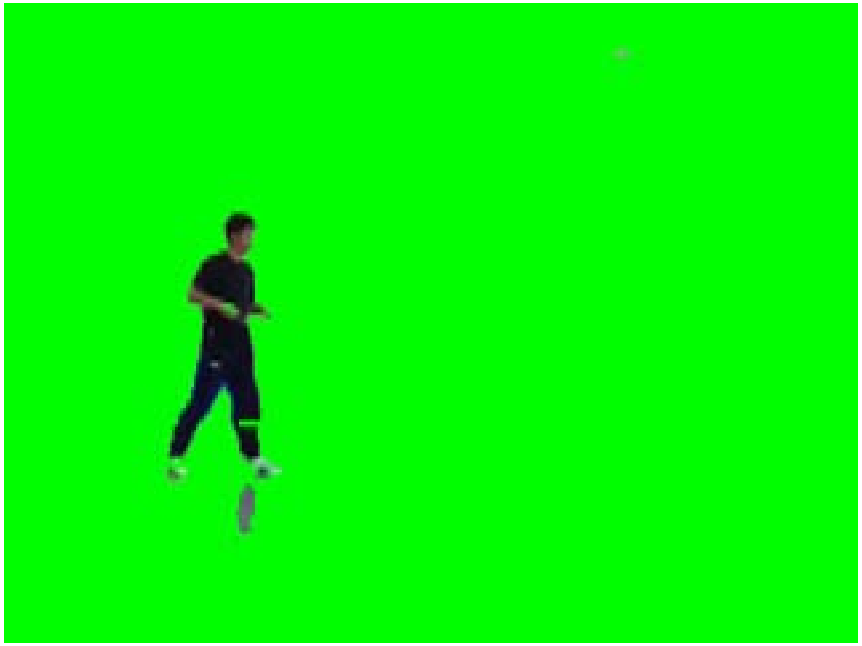
   - Update cluster centres:
     Compute mean of colours in cluster

t is iteration number

$$\mathbf{w}_i(t+1) = \frac{1}{|C_i(t)|} \sum_{\mathbf{x} \in C_i(t)} \mathbf{x}$$

- For background removal, can choose $k = 2$
  - One for foreground, one for background.
- Initial cluster centres
  - Get from foreground and background in video.
- Possible termination criteria
  - Very few colours change clusters.
  - Fixed number of iterations.
- After running clustering
  - If foreground area is small, then smaller cluster is foreground.

Background removed

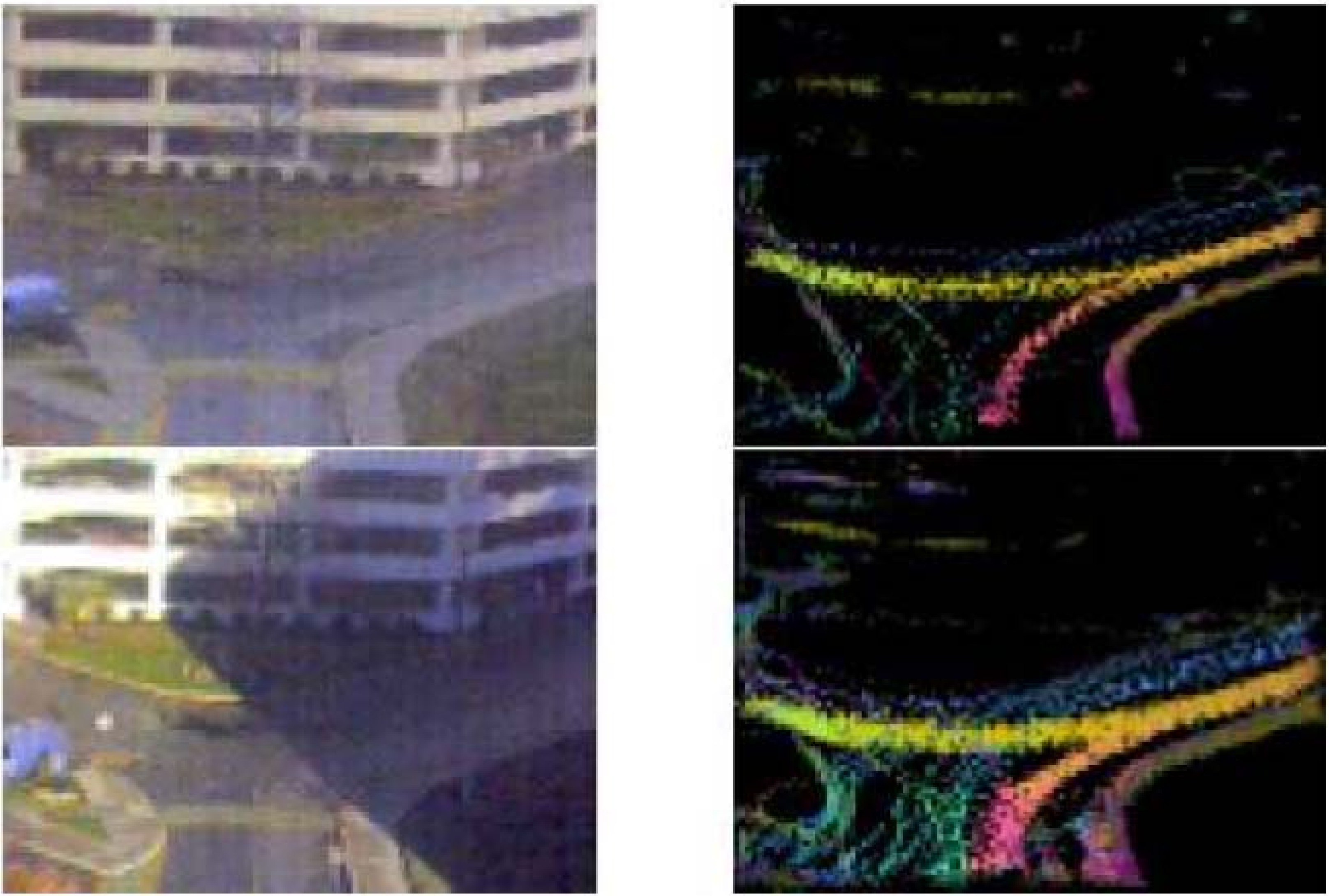


- Most background colours are removed.
- A bit of shadow remains.
- Most foreground colours are found.

Stationary Camera, Varying Lighting

- Basic ideas
  - Multiple background clusters for different lighting conditions.
  - Apply k-means clustering with $k > 2$.

Example from [Stauffer98]



Moving Camera

- Basic ideas
  - Track and recover camera motion [Bergen92].
  - Stabilise video by removing camera motion [Matsushita05].
  - Do stationary camera background removal.
  - Put back camera motion.

Summary

- With clean plate
  - Subtract clean plate from video frames.
- Without clean plate
  - Estimate background
    - Average video frame
    - Cluster pixel colours
  - Subtract estimated background from video frames.
- Moving camera
  - Stabilise video, then perform background removal.

Further Reading

- Code book method
  - OpenCV [Bradski08] chapter 9.
- Varying lighting condition
  - [Stauffer98]
- Motion estimation
  - [Bergen92]
- Video stabilization
  - [Matsushita05]

References

- G. Bradski and A. Kaebler*, Learning OpenCV*, O'Reilly, 2008.
- J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *Proc. ECCV*, pages 237–252, 1992.
- Y. Matsushita, E. Ofek, X. Tang, and H.Y. Shum. Fullframe video stabilization. In *Proc. CVPR*, volume 1, pages 50–57, 2005.
- C. Stauffer and W. E. L. Grimson. Adaptive background mixture models for real-time tracking. In *Proc. IEEE Conf. on CVPR*, 1998.