Master in
Computer Vision
Barcelona

Module: Video Analysis

Lecture 2: Video segmentation

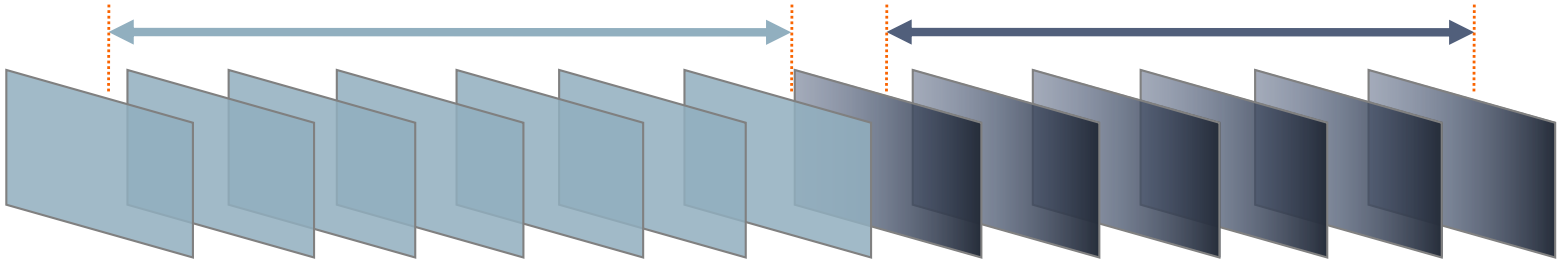
Lecturer: Montse Pardàs

Outline

- Introduction to video segmentation
- Shot segmentation
- Moving object segmentation
    - Introduction
    - Still background estimation
    - Variable background estimation: Single Gaussian
    - Variable background estimation: Multiple Gaussians
    - Other change detection techniques
  - Shadow detection
  - Connected component analysis and tracking
- Region segmentation
    - Spatial segmentation and tracking
    - Spatio-temporal segmentation

Introduction to video segmentation

- **Different meanings**:
  - **Shot detection**: The segmentation is seen as a 1D (temporal) problem.
  - A **temporal segmentation** is performed, dividing the video into segments that should be related to the different shots in the sequence.



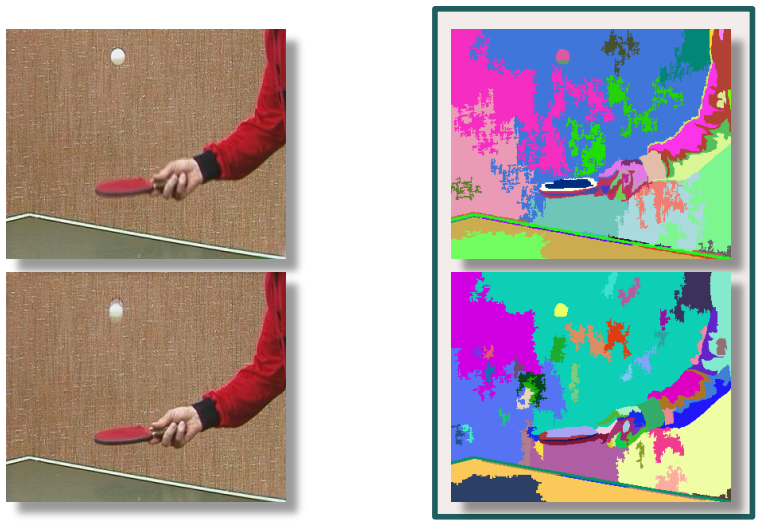
Introduction to video segmentation

- **Different meanings**:

  Moving object segmentation: Detect and segment moving objects from a video sequence of a fixed camera
  - Background – static scene
  - Foreground – moving objects

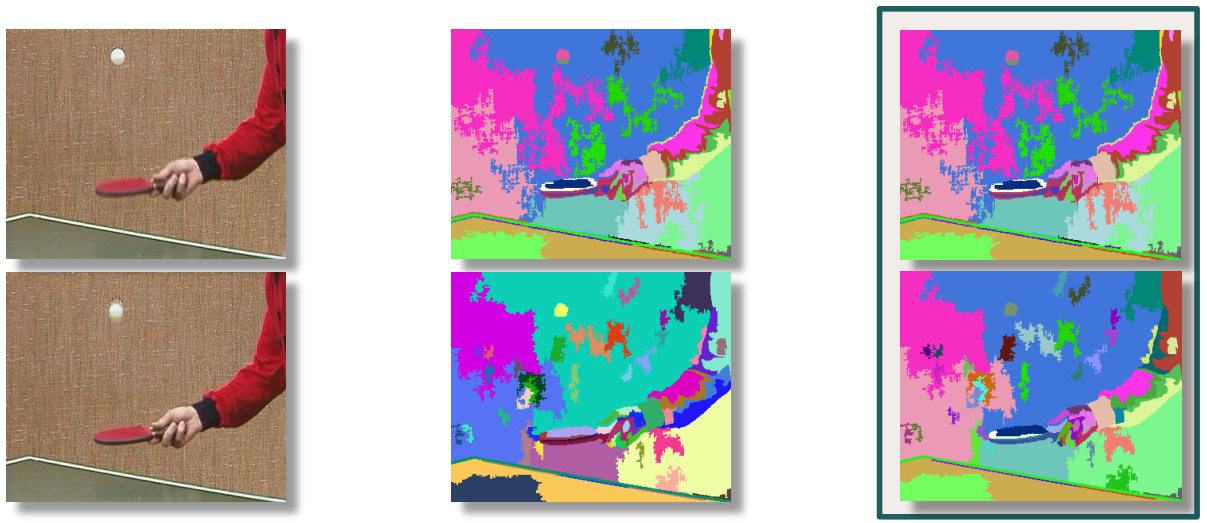
Introduction to video segmentation

- **Different meanings**:
  - **2D Region segmentation**: The sequence is understood as a set of isolated images. **Spatial segmentation** is performed on each separated image (2D signal) in the sequence. Regions have to be related to the objects in the scene.
    - It is the same problem as the **image segmentation** problem.
    - Object labels in consecutive images **may not be coherent**.



Introduction to video segmentation

- **Different meanings**:
  - **3D region segmentation**: The sequence is understood as a set of temporally related images. **Spatial/Temporal segmentation** is performed on the sequence (3D signal). Regions have to be related to the objects in the scene and their temporal evolution.
    - Direct extension of the **image segmentation** problem to the 3D case.
    - Object labels in consecutive images **should be coherent**.



Introduction to video segmentation

- The division in three steps proposed for image segmentation can be adopted:

- **Remove useless or annoying information**
- **Preserve shape information**

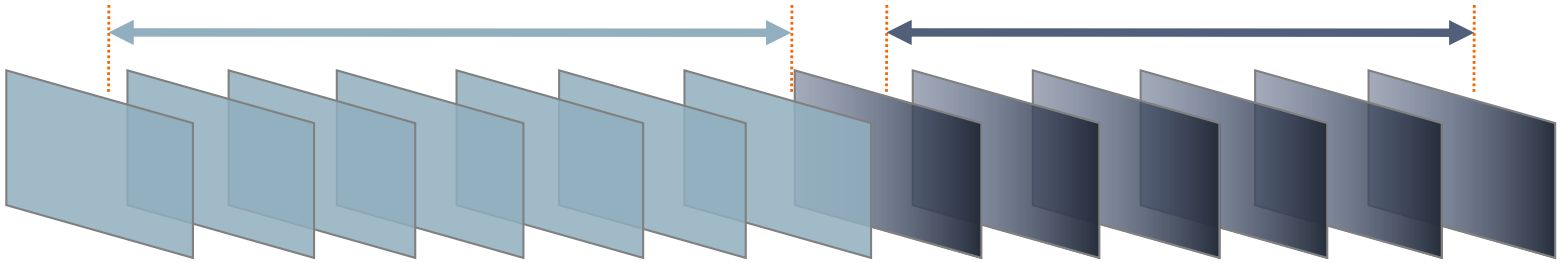
- **Feature space**:
  Gray level/Color
  Texture
  Depth
  Motion
  Frame difference
  DFD
  Histogram
- Elements belonging to the same regions should correspond to **homogeneous features**

- **Decision space**:
  1D: Temporal
  2D: Spatial
  3D: Spatial/temporal
- **Partition definition**:
  - ✓ Transition
  - ✓ Homogeneity

Outline

- Shot segmentation
- Moving object segmentation
- Region segmentation

Shot segmentation (I)

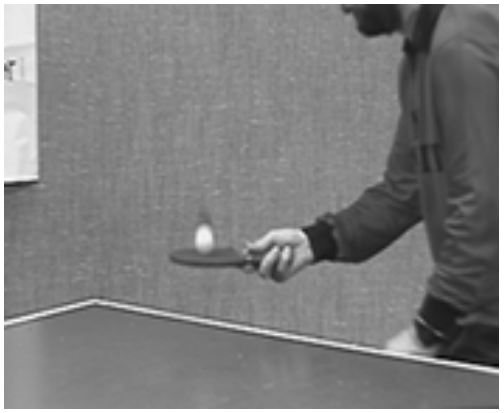


- A **video shot** is a sequence of frames captured by one camera in a single continuous action in time and space.

- **Feature extraction and decision**
  - Frame difference (FD)
  - Frame histogram comparison
  - Displaced frame difference (DFD)

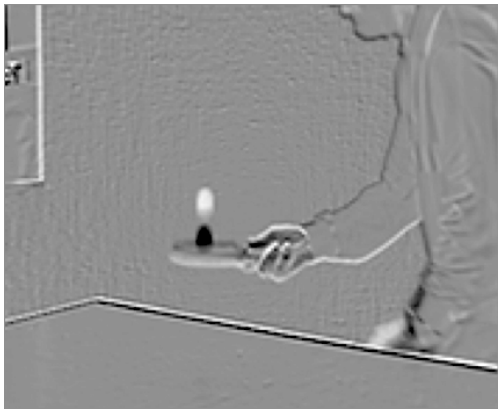
Temporal gradients

Shot segmentation (I)

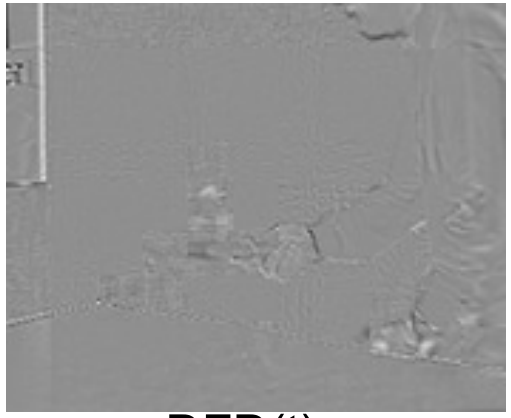


I(t-1)

I(t)



FD(t)

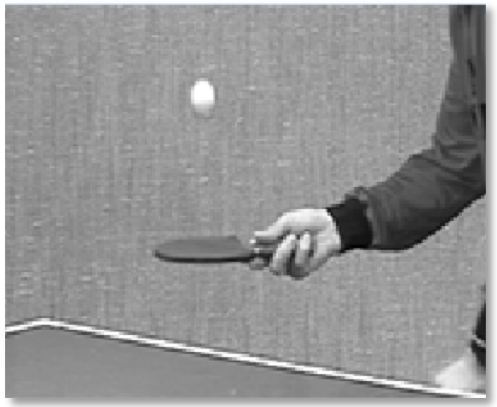


DFD(t)

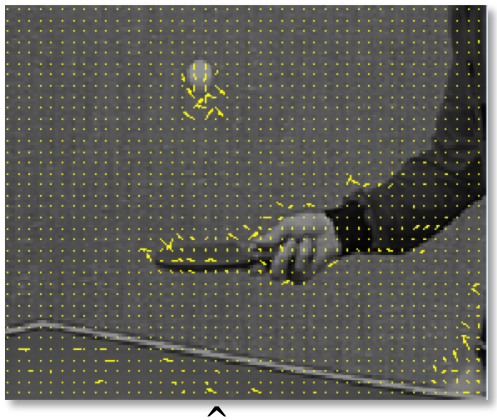
2D Motion analysis

- In practice, the optical flow has to be estimated and may differ from the ideal one.
- If $\hat{\vec{D}}(\vec{r})$ denotes the estimated optical flow, the reference image to which the optical flow is applied is called the motion compensated image.

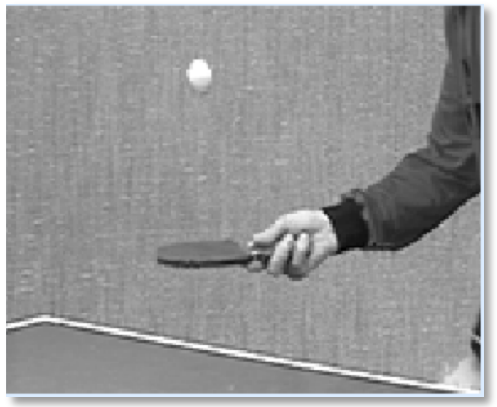
$$I^{MC}(\vec{r},t) = I(\vec{r} - \hat{\vec{D}}(\vec{r}), t - \Delta t)$$



$I(\vec{r},t-\Delta t)$



$\hat{\vec{D}}(\vec{r})$

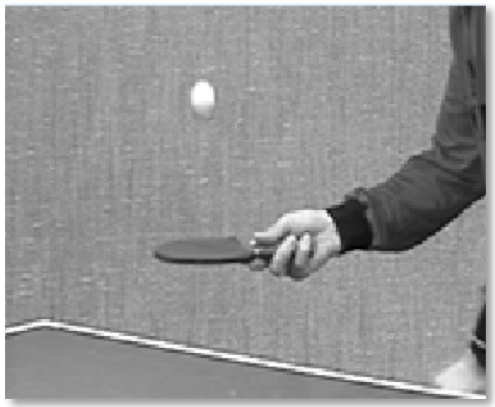


$I^{MC}(\vec{r},t)$

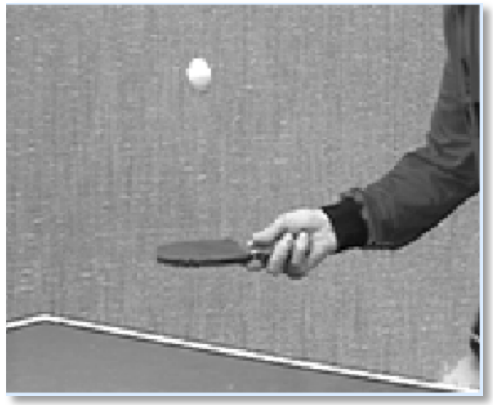
2D Motion analysis

The difference between the motion compensated frame and the real current image is called the **Displaced Frame Difference** (DFD).

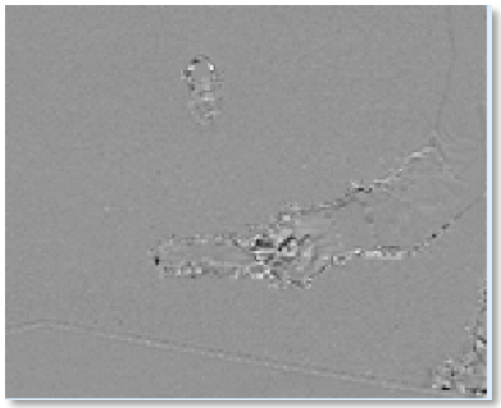
$$DFD(\vec{r},\hat{\vec{D}}(\vec{r}))=I(\vec{r},t)-I(\vec{r}-\hat{\vec{D}}(\vec{r}),t-\Delta t)$$



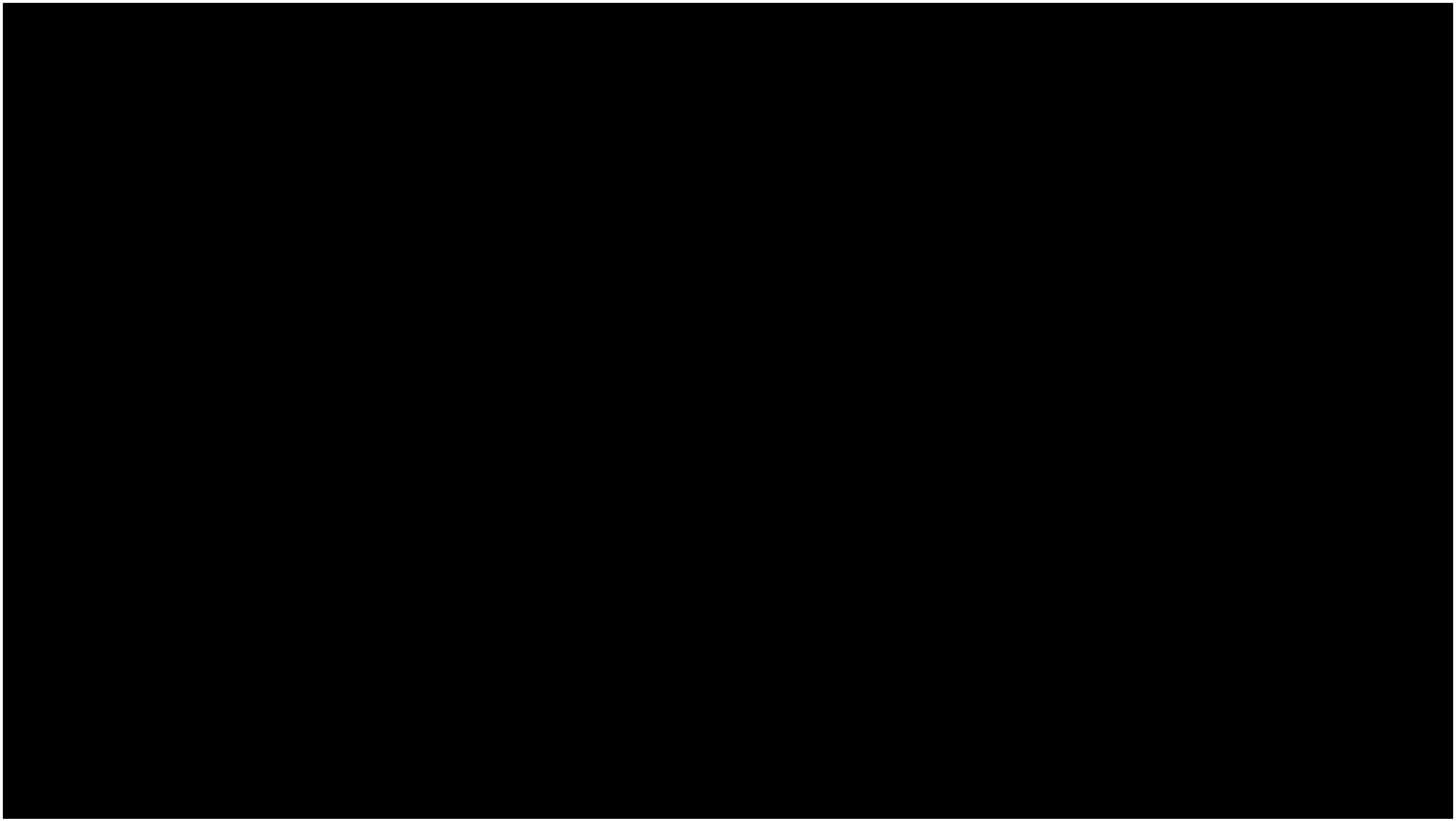
$I(\vec{r},t-\Delta t)$



$I^{MC}(\vec{r},t)$

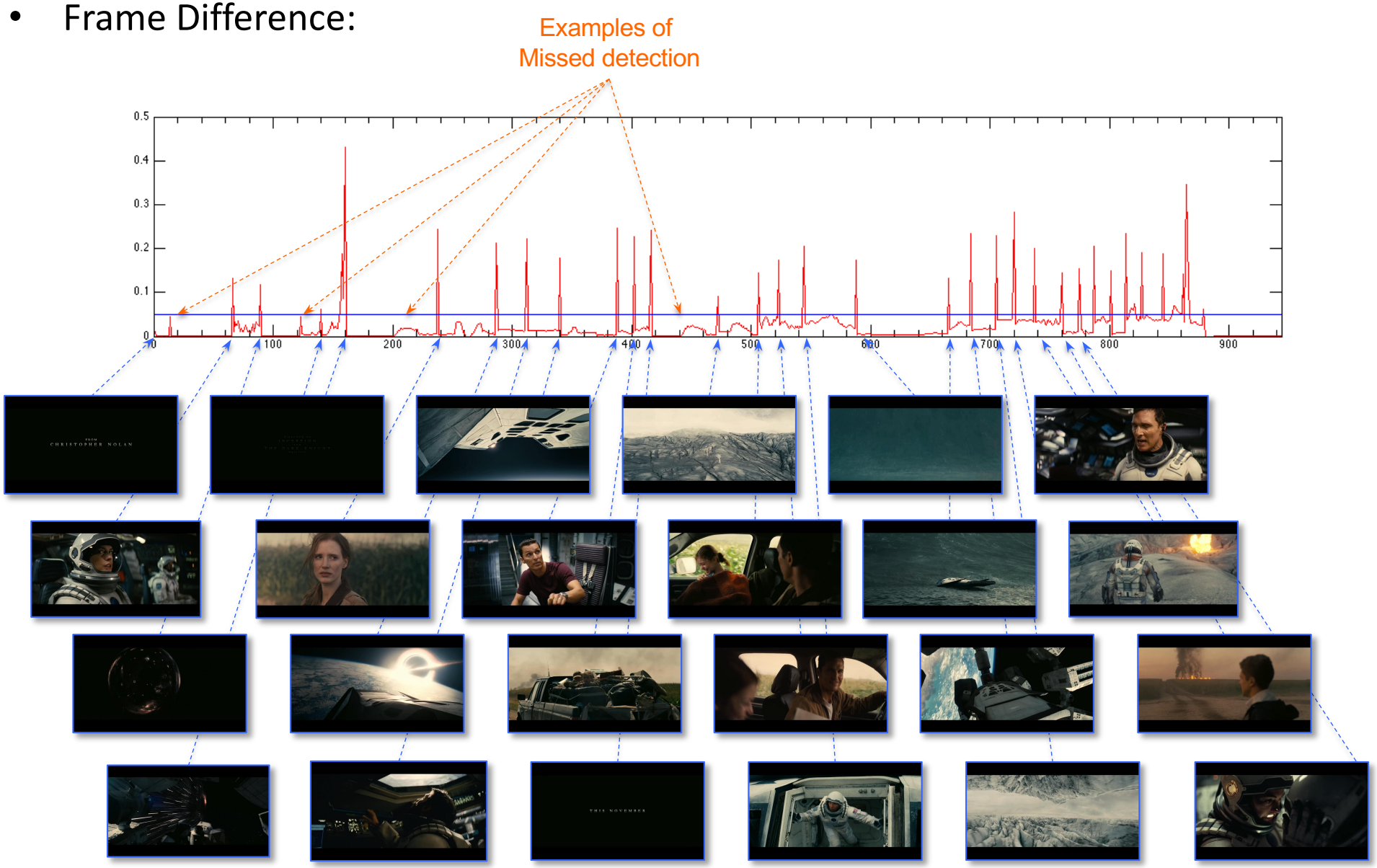


$DFD(\vec{r},\hat{\vec{D}}(\vec{r}))$

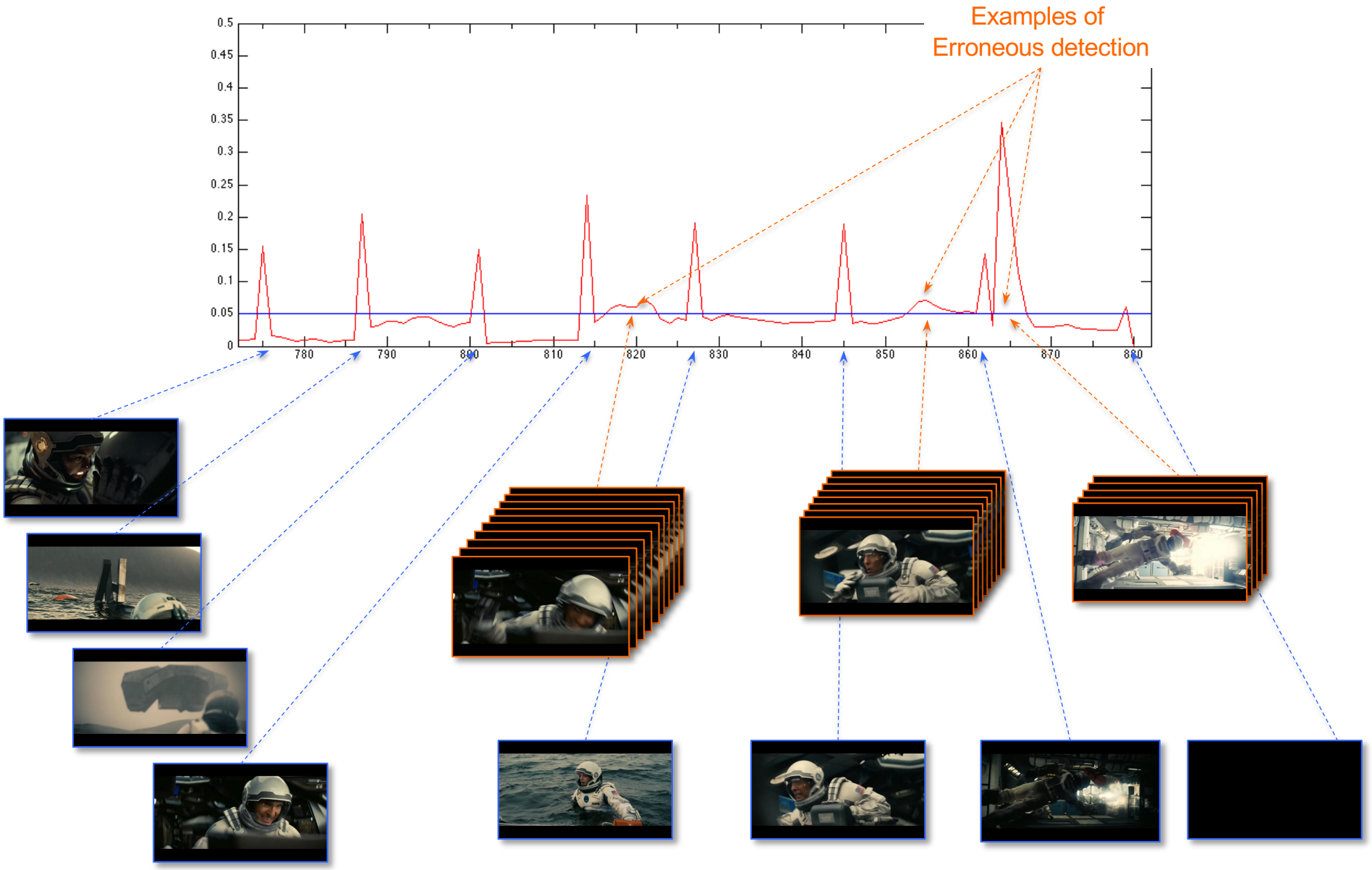


Shot segmentation (II)

- Frame Difference:

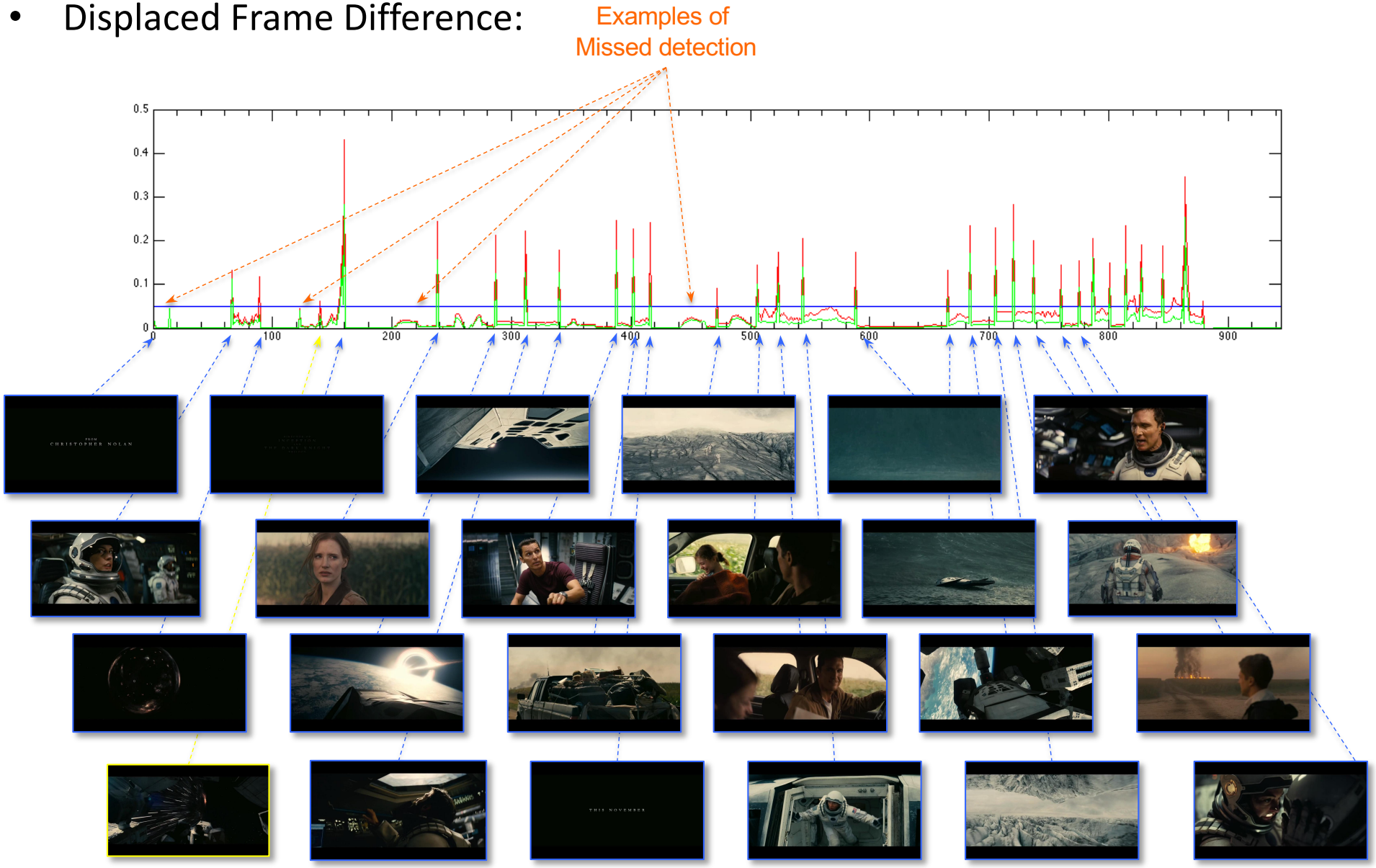


Shot segmentation (III)

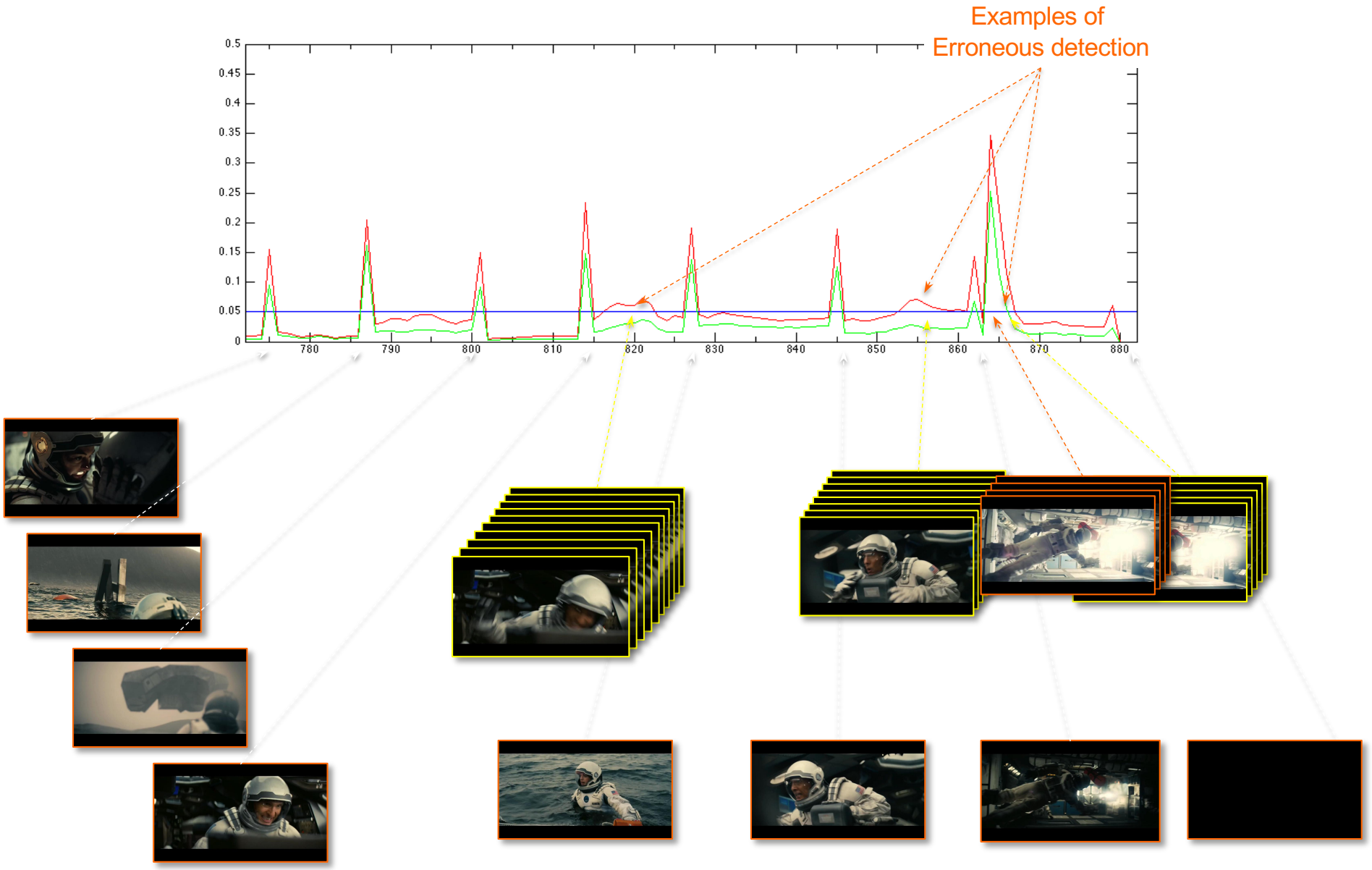


Shot segmentation (IV)

- Displaced Frame Difference:



Shot segmentation (V)



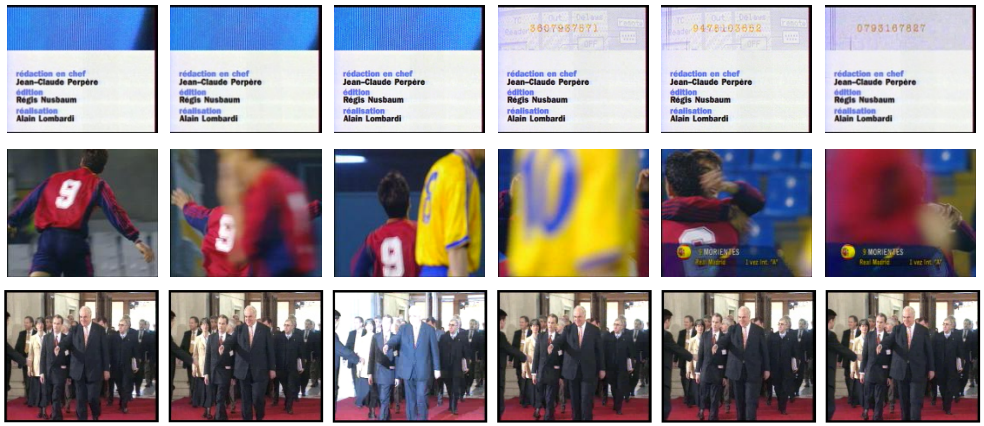
Shot segmentation (VI)

The DFD comparison allows predicting the current image with the information of the previous image leading to a more robust estimation of the shot transition. However, some transitions are difficult to handle (dissolve and fade, for instance)

- Examples of **subtle shot transitions**.

These changes correspond to shot transitions

- Examples of **internal transitions or motion**.



These changes *do not* correspond to shot transitions