MATCHING MOTION DESCRIPTORS FOR SHORT TERM ACTION RECOGNITION

Robert B. Fisher

School of Informatics

University of Edinburgh

MATCHING DESCRIPTORS I

Start with single $\boxed{?}$ matching

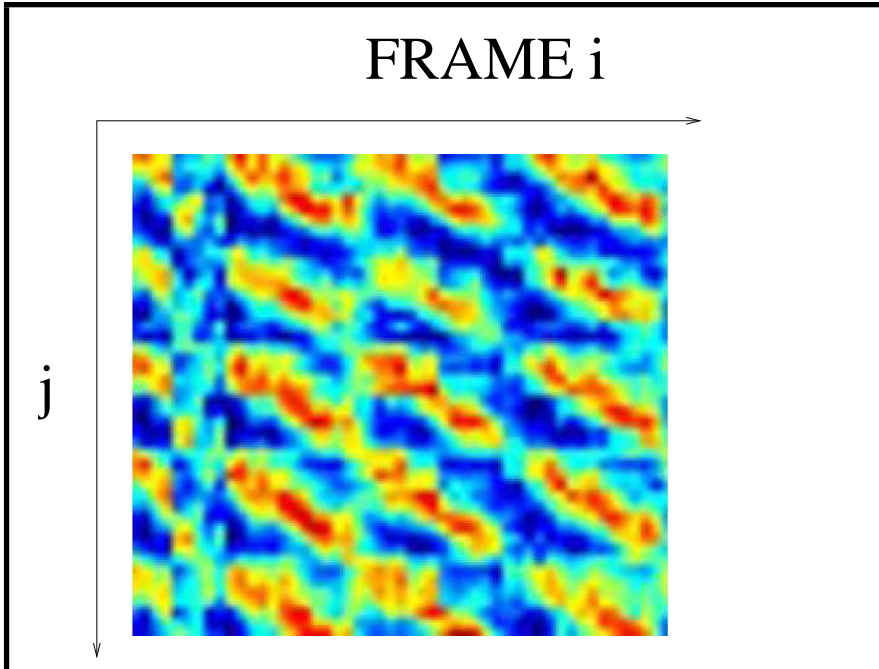$$m(i,j) = \sum_{c=1}^{4} \sum_{x,y \in I} a_c^i(x,y) b_c^j(x,y)$$

Where

Frame i of seq. a, frame j of seq. b

$c = 1,2,3,4$ optical flow components

$(x,y) =$ pixel positions
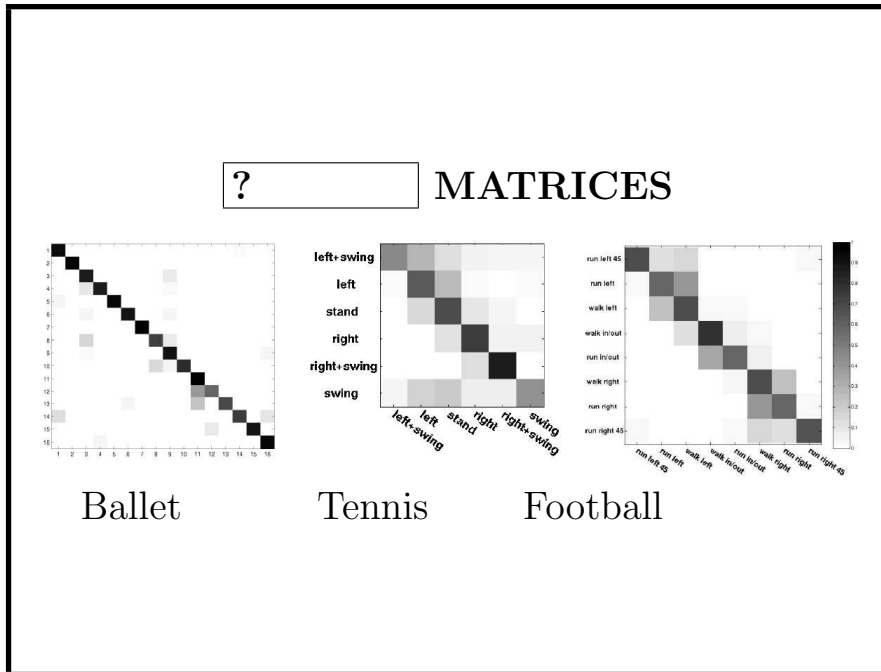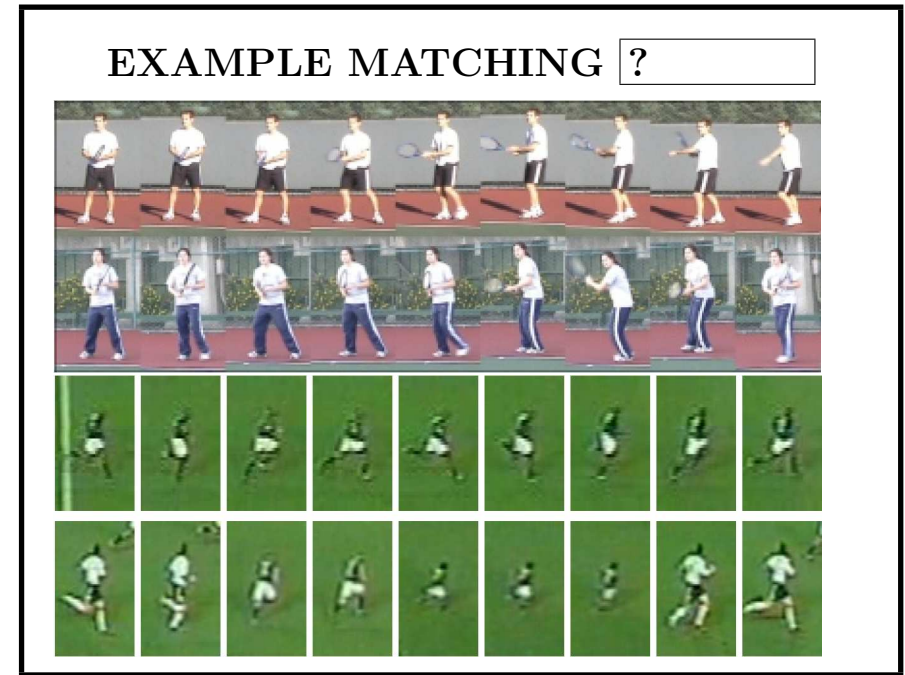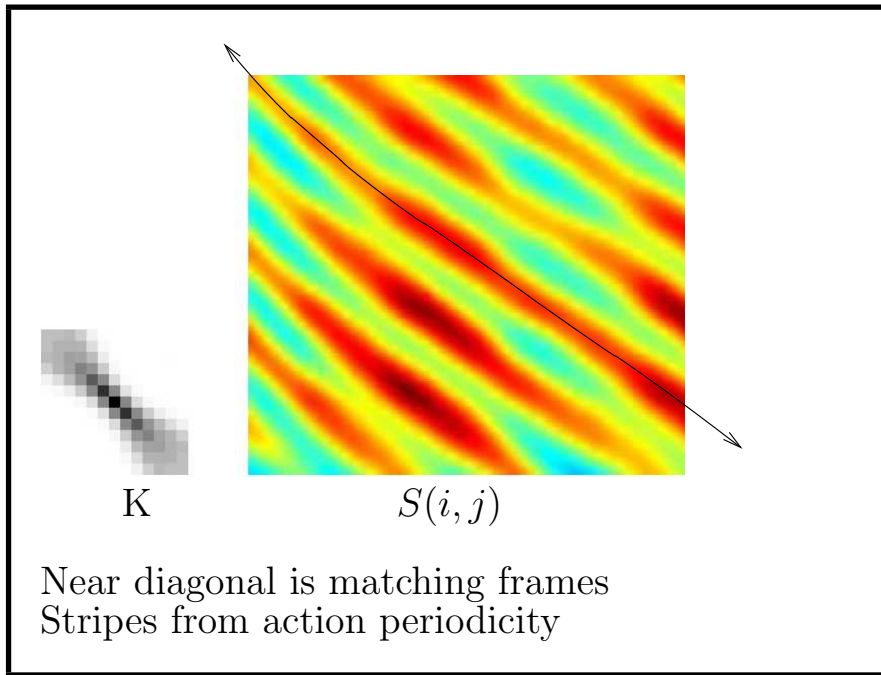
FRAME i

j

$\boxed{?}$ DESCRIPTORS II

Problem: match score from a single frame a bit noisy

So use time window of $T = 50$ frames

$$S(i,j) = \sum_{r=-T/2}^{r=+T/2} \sum_{s=-T/2}^{s=+T/2} K(r,s) m(i+r, j+s)$$

Weighted sum of nearby in time frames (convolution)

K $S(i,j)$

Near diagonal is matching frames
Stripes from action periodicity

EXAMPLE MATCHING ?

? MATRICES

Ballet   Tennis   Football

WHAT WE HAVE LEARNED

1. Short term action recognition technique

2. Based on stabilized optical flow of local medium sized windows

3. Encodes ? structure better

4. But: still viewpoint and scale dependent