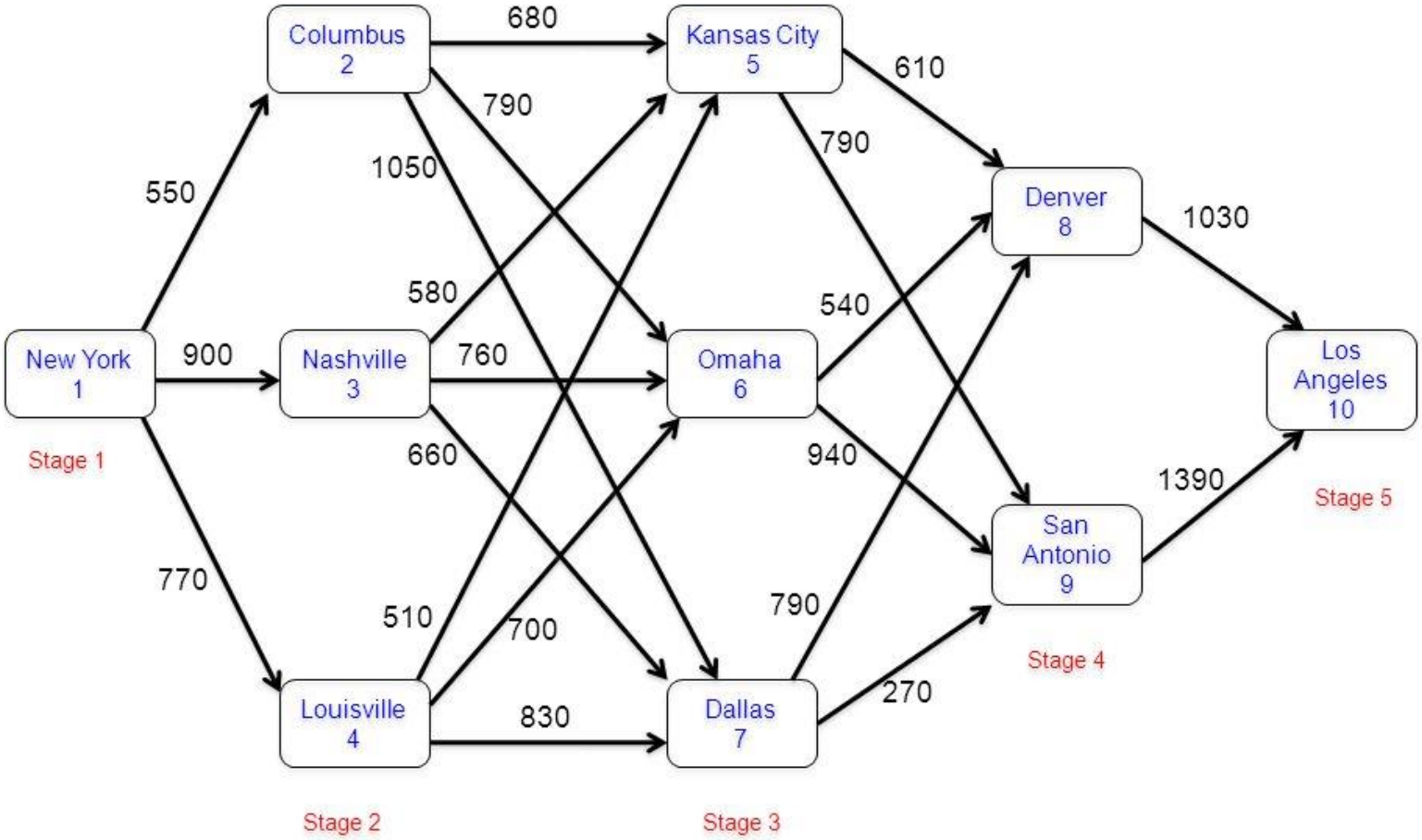
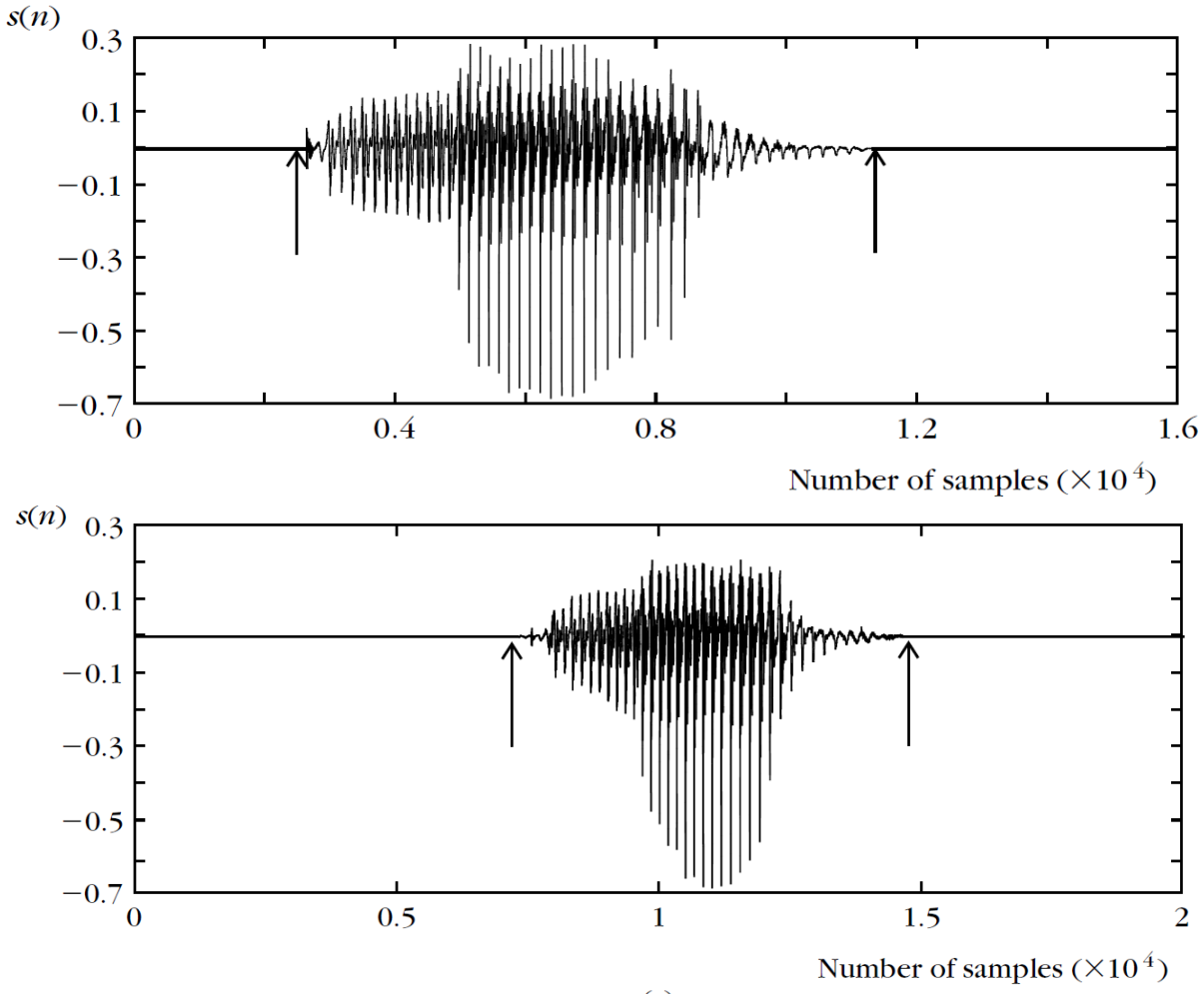
Shortest Path Problem



DYNAMIC TIME WARPING IN SPEECH RECOGNITION

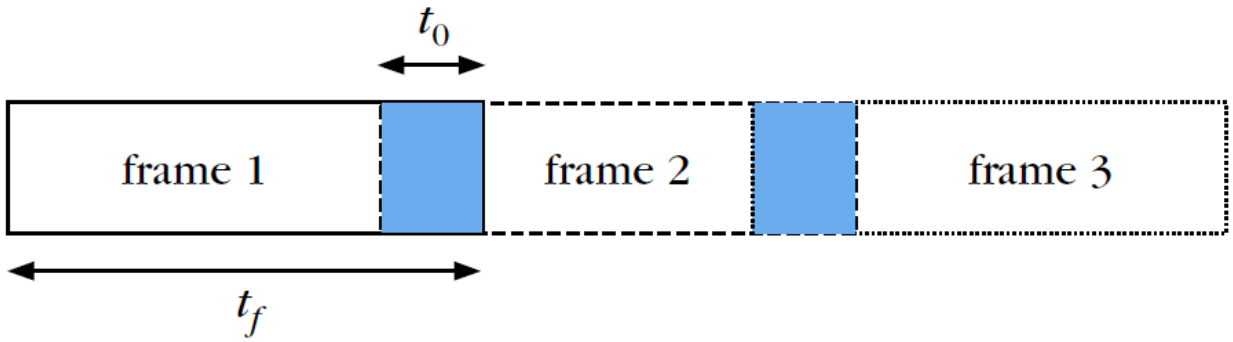
Isolated Word Recognition

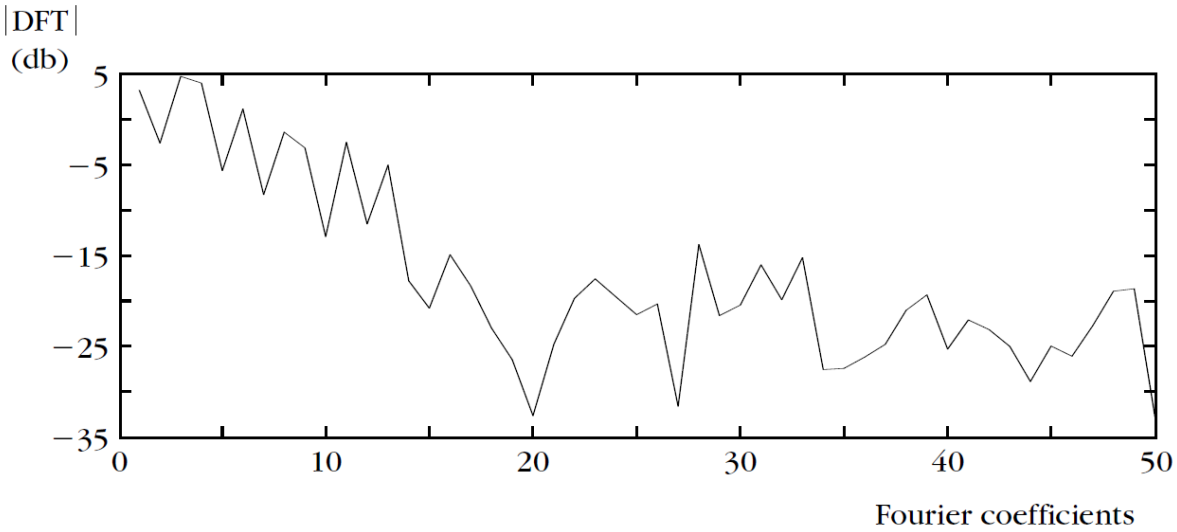
- Two “love”s



Feature Extraction

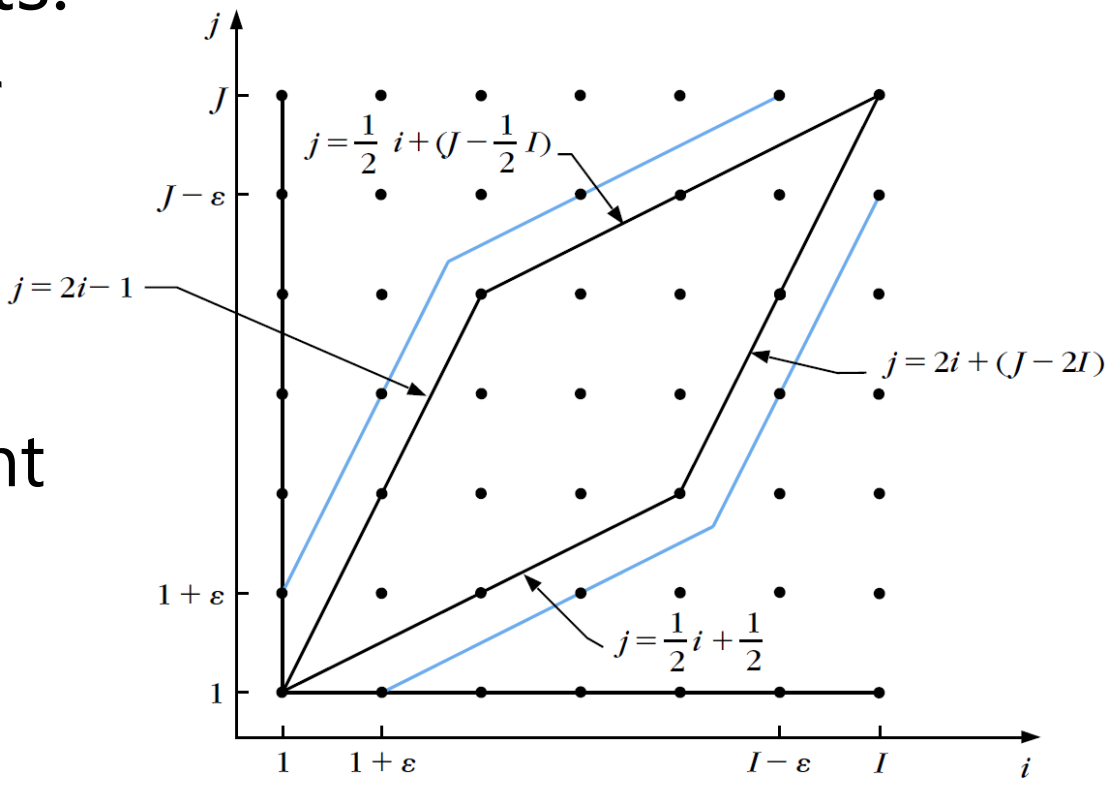
- 512-DFT with 100-sample overlap
- Use the first 50 coefficients only





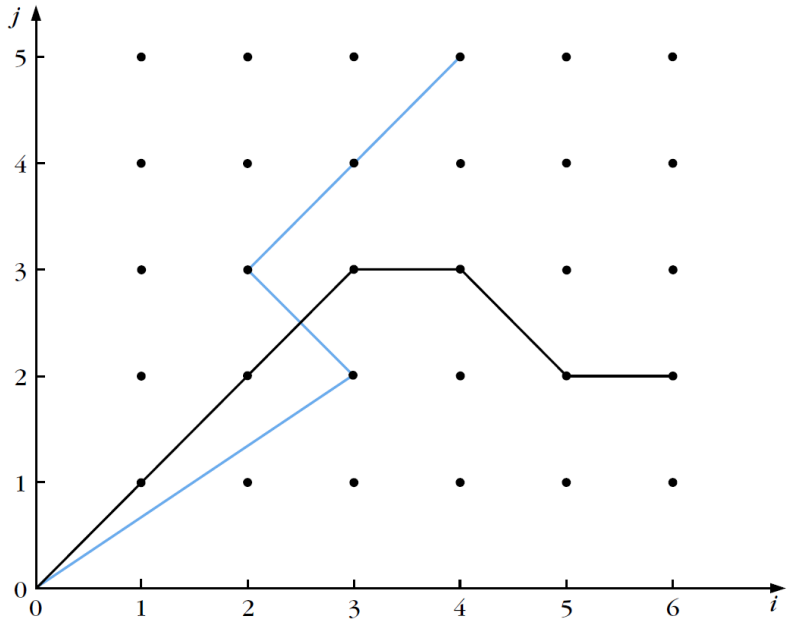
Constraints

- End-point constraints:
  - Start at $(1,1)$ and end at (I,J)
- Global constraints:
  - Stretching factor $\frac{1}{2} < s < 2$
  - Blue curves with relaxed end-point constraints



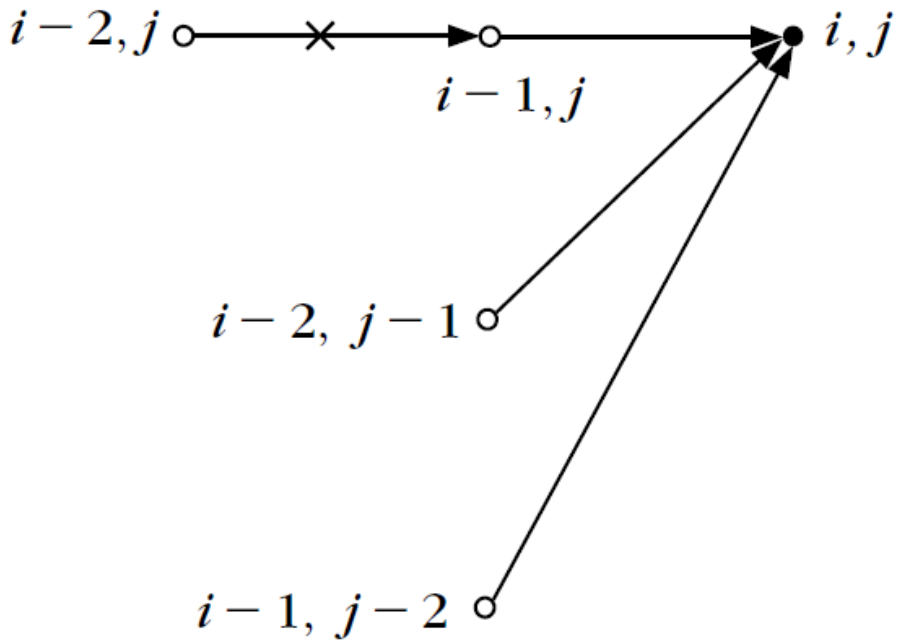
Constraints

- Monotonicity constraints
  - $i_{k-1} \leq i_k$ and $j_{k-1} \leq j_k$



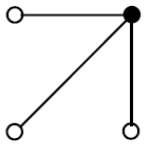
Constraints

- Itakura local constraints
  - Two successive horizontal transitions are not allowed

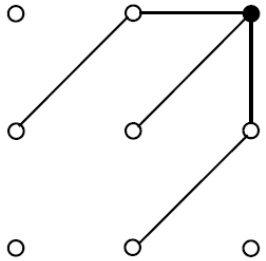


Constraints

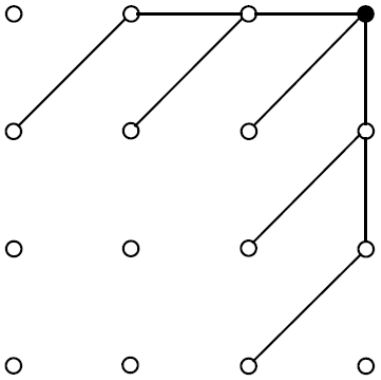
- Sakoe and Chiba local constraints



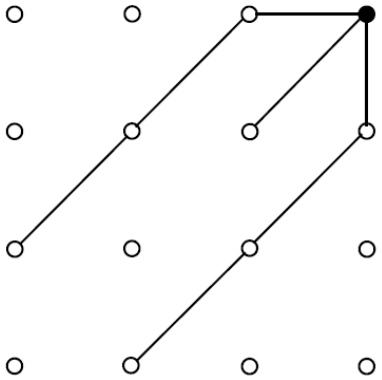
(a)



(b)



(c)



(d)

Cost

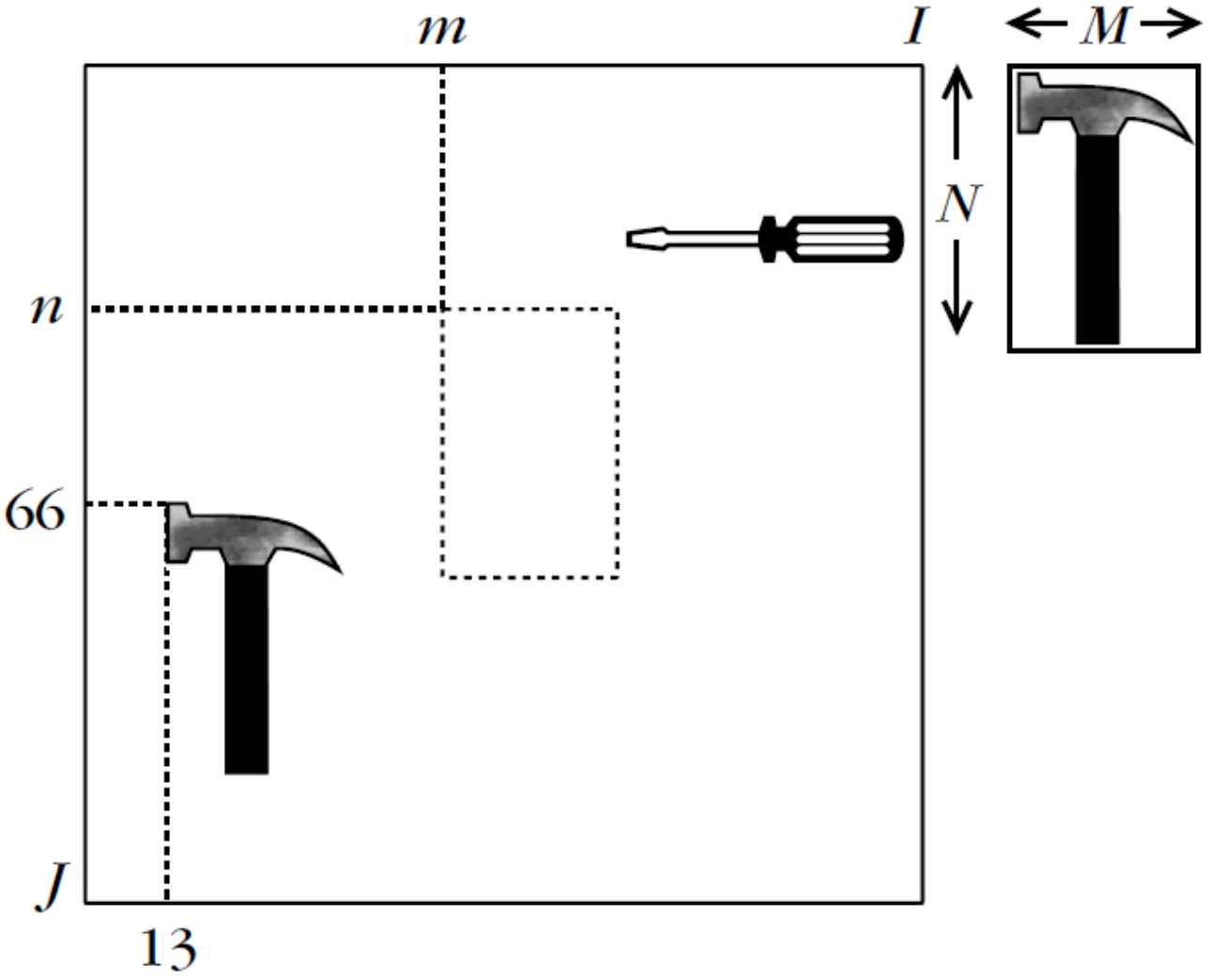
- A choice

$$d(i_k, j_k | i_{k-1}, j_{k-1}) = \|\boldsymbol{r}(i_k) - \boldsymbol{t}(j_k)\|$$

- It is then a shortest path problem

MOTION ESTIMATION AND COMPENSATION IN VIDEO CODING

Block Matching



Matching Costs

- SSD

$$D_{\mathrm{SSD}} = \sum_{i=1}^{M}\sum_{j=1}^{N} |t(i,j) - r(i,j)|^2$$

- Cross-correlation

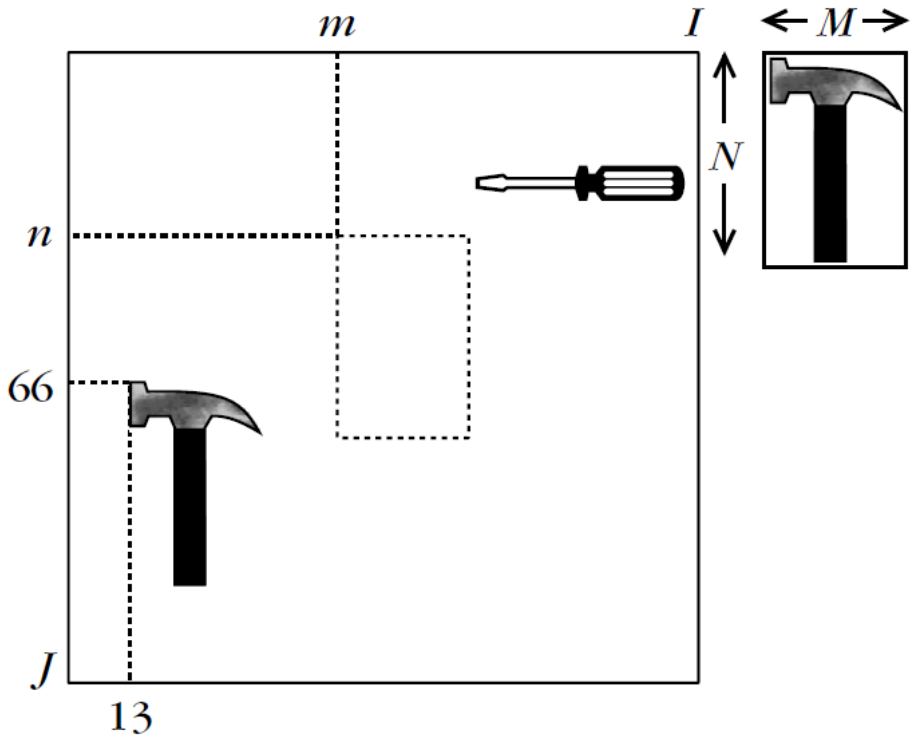
$$D_{\mathrm{CC}} = \sum_{i=1}^{M}\sum_{j=1}^{N} t(i,j)r(i,j)$$

- Cross-correlation coefficient

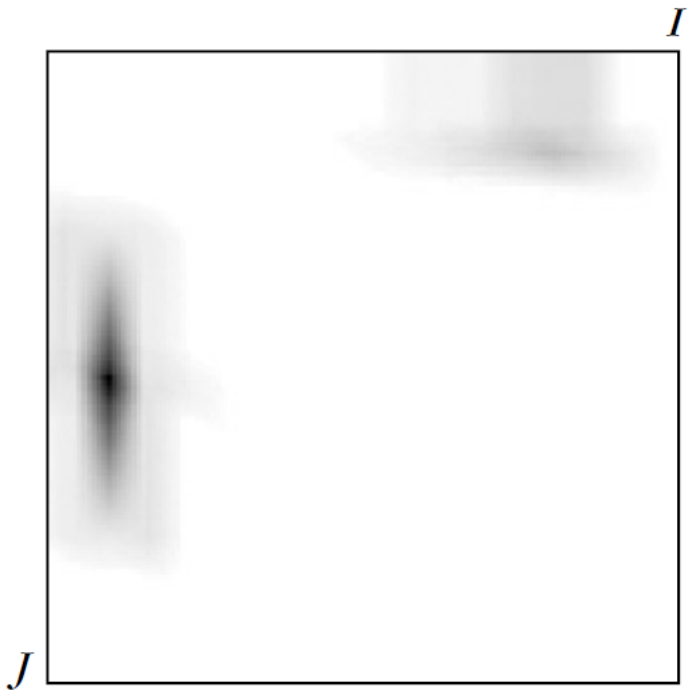
$$D_{\mathrm{NCC}} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} t(i,j)r(i,j)}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} t^2(i,j) \times \sum_{i=1}^{M}\sum_{j=1}^{N} r^2(i,j)}}$$

Matching Costs

- Correlation (the darker, the higher correlation)



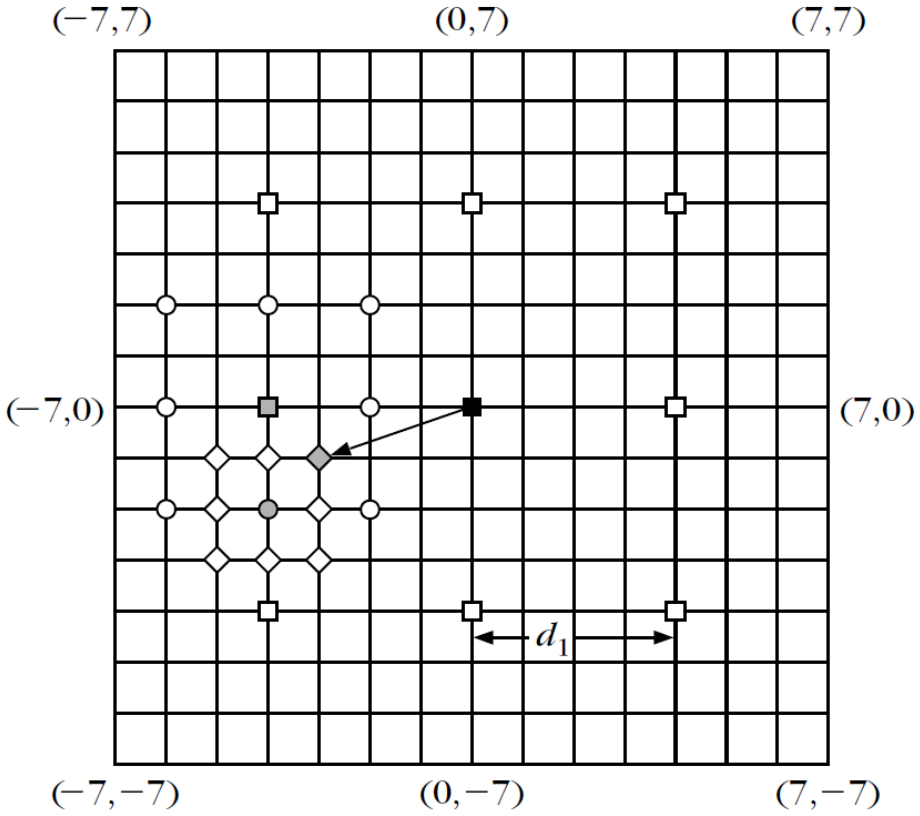
(a)



(b)

Reducing Complexity

- Logarithmic search (three-step search)



- Early termination of cost computation