CHAPTER 7 – TEMPLATE MATCHING

- The Goal: Given a set of reference patterns known as TEMPLATES, find the one to which an unknown pattern matches best. That is, each class is represented by a single typical pattern.

- The crucial point is to adopt an appropriate "measure" to quantify similarity or matching.

- These measures must accommodate, in an efficient way, deviations between the template and the test pattern. For example, the word beauty may have been read a beeauty or beuty, etc., due to errors.

- Typical Applications
  - Speech Recognition
  - Motion Estimation in Video Coding
  - Data Base Image Retrieval
  - Written Word Recognition
  - Bioinformatics

- Measures based on optimal path searching techniques
  - Representation: Represent the template by a sequence of measurement vectors

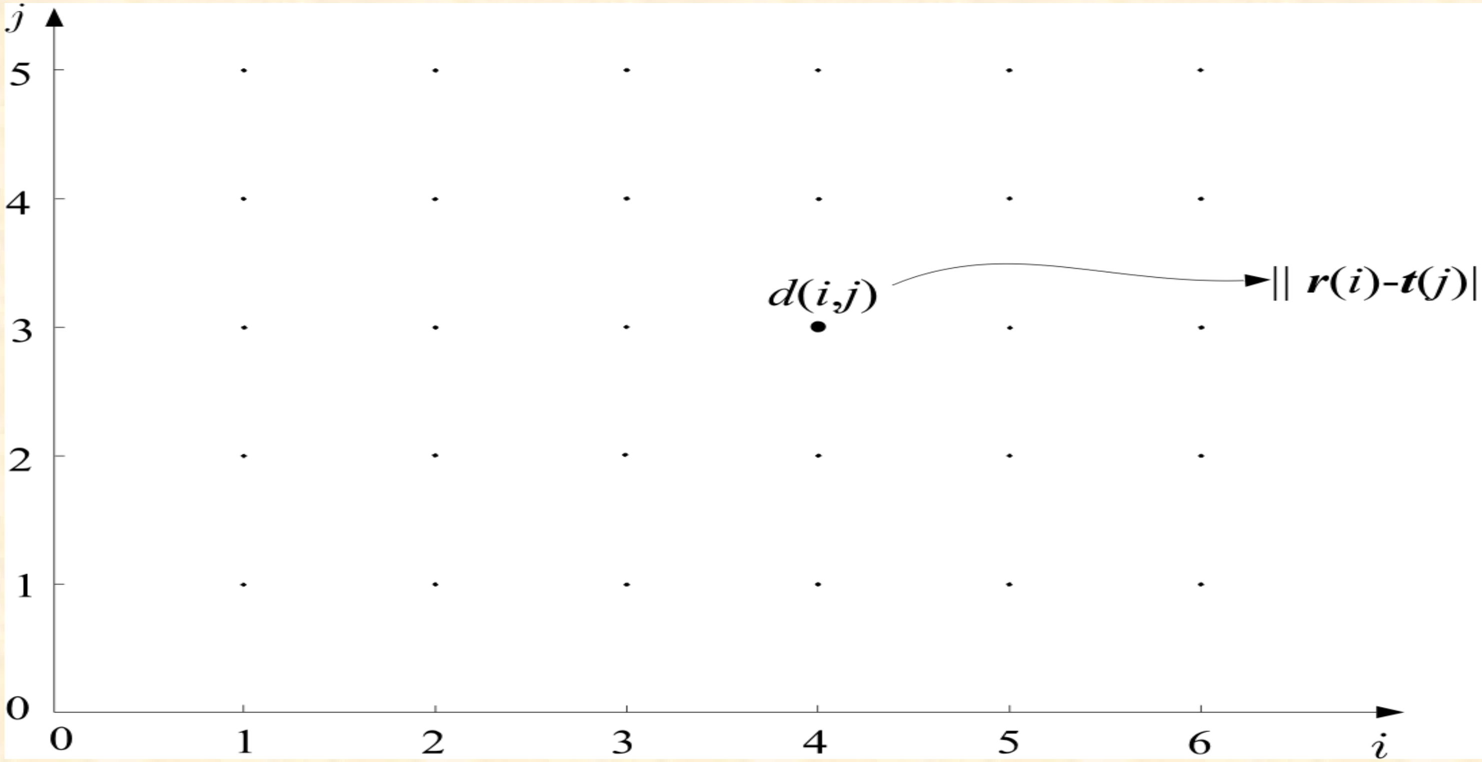
Template: $\underline{r}(1), \underline{r}(2), \ldots, \underline{r}(I)$

Test pattern: $\underline{t}(1), \underline{t}(2), \ldots, \underline{t}(J)$

- In general $I \neq J$

- Form a grid with I points (template) in the horizontal and J points (test) in the vertical axes, respectively.

- Each point (i,j) of the grid measures the distance between $\underline{r}(i)$ and $\underline{t}(j)$

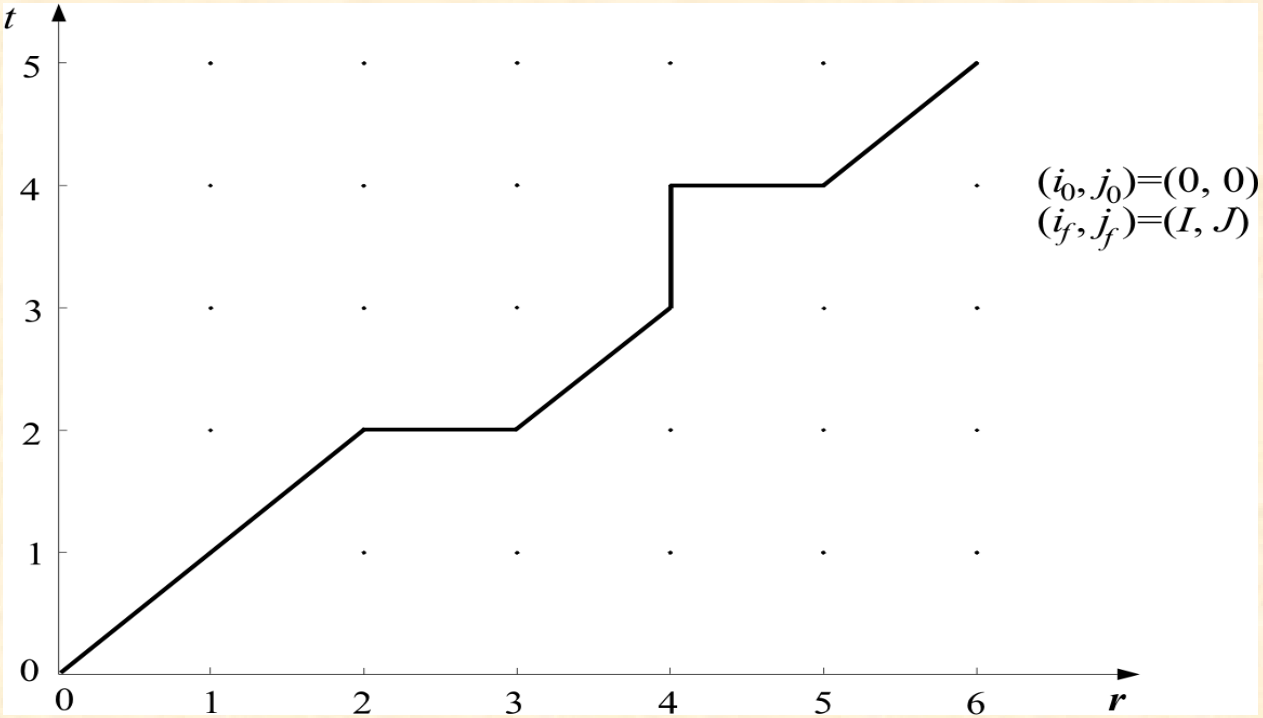


- Path: A path through the grid, from an initial node (i_0, j_0) to a final one (i_f, j_f), is an ordered set of nodes $(i_0, j_0), (i_1, j_1), (i_2, j_2) \ldots (i_k, j_k) \ldots (i_f, j_f)$.

- Each path is associated with a cost

$$D = \sum_{k=0}^{K-1} d(i_k, j_k)$$

where K is the number of nodes across the path.



- Search for the path with the optimal cost $D_{opt.}$
- The matching cost between template $\underline{r}$ and test pattern $\underline{t}$ is $D_{opt.}$

BELLMAN'S OPTIMALITLY PRINCIPLE

- Optimum path:

$$(i_0, j_0) \xrightarrow{opt} (i_f, j_f)$$

- Let (i,j) be an intermediate node, i.e.,

$$(i_0, j_0) \to \ldots \to (i, j) \to \ldots \to (i_f, j_f)$$

Then write the optimal path through (i, j)

$$(i_0, j_0) \underset{(i,j)}{\overset{opt}{\longrightarrow}} (i_f, j_f)$$

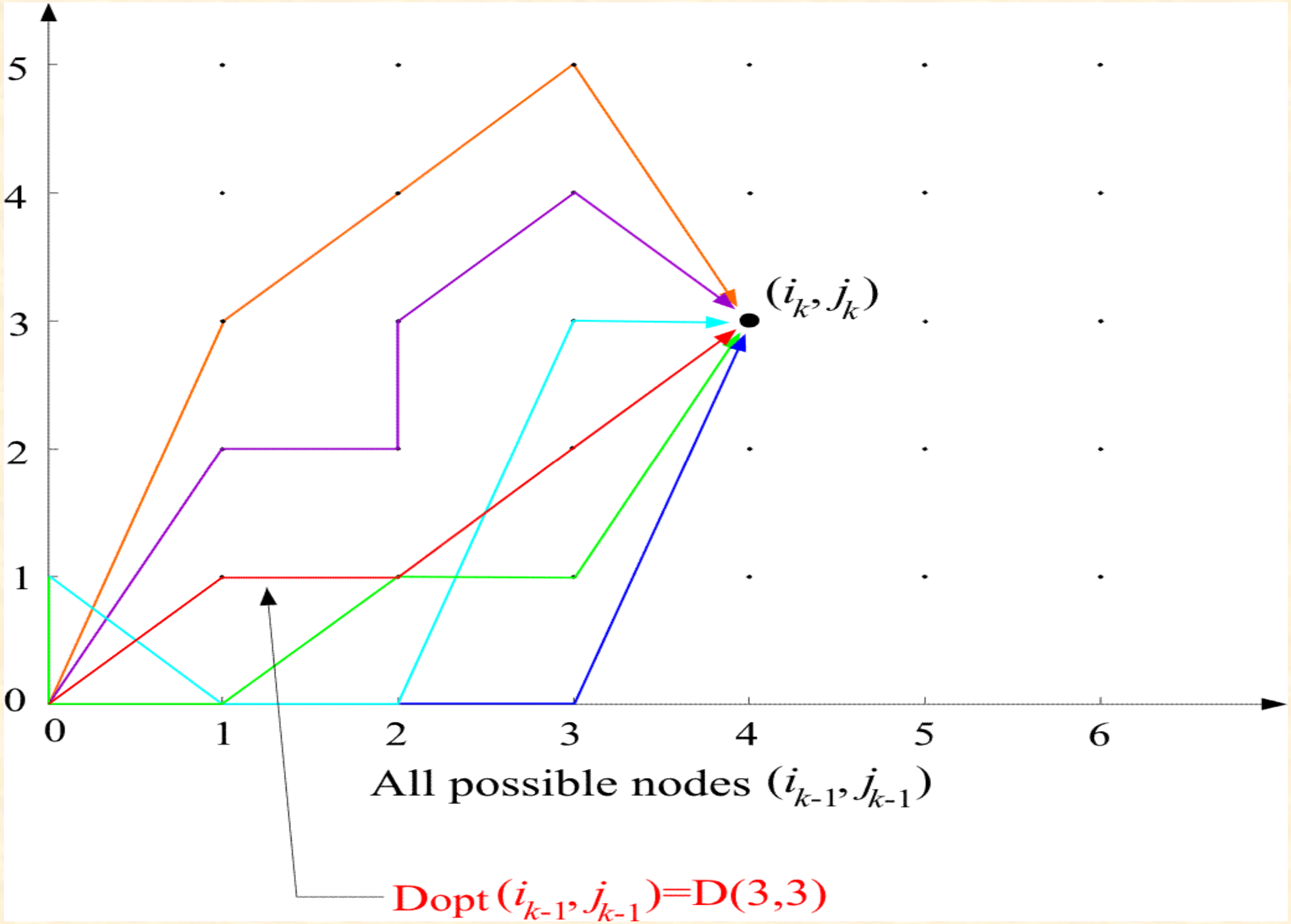
❖ Bellman's Principle:

$$(i_0, j_0) \xrightarrow{opt} (i_f, j_f) = (i_0, j_0) \xrightarrow{opt} (i, j) \oplus$$
$$(i, j) \xrightarrow{opt} (i_f, j_f)$$

❖ In words: The overall optimal path from (i_0, j_0) to (i_f, j_f) through (i,j) is the concatenation of the optimal paths from (i_0, j_0) to (i,j) and from (i,j) to (i_f, j_f).

❖ Let $D_{opt.}(i,j)$ is the optimal path to reach (i,j) from (i_0, j_0), then Bellman's principle is expressed mathematically as:

$$D_{opt}(i_k, j_k) = opt\{D_{opt}(i_{k-1}, j_{k-1}) + d(i_k, j_k)\}$$



❖ The Edit distance

- It is used for matching written words. Applications:
  - Automatic Editing
  - Text Retrieval
  - Bioinformatics

- The measure to be adopted for matching, must take into account:
  - Wrongly identified symbols
    e.g. “befuty” instead of “beauty”
  - Insertion errors, e.g., “bearuty”
  - Deletion errors, e.g., “beuty”

- Edit distance: **Minimal** total number of changes, C, insertions I and deletions R, required to change pattern A into pattern B,

$$D(A,B) = \min_{j}[C(j) + I(j) + R(j)]$$

where j runs over All possible variations of symbols, in order to convert $A \longrightarrow B$

- Allowable predecessors and costs

  - $(i-1, j-1) \rightarrow (i, j)$

$$d(i, j \,|\, i-1, j-1) = \begin{cases} 0, & \text{if } t(i) = r(j) \\ 1, & \text{if } t(i) \neq r(j) \end{cases}$$

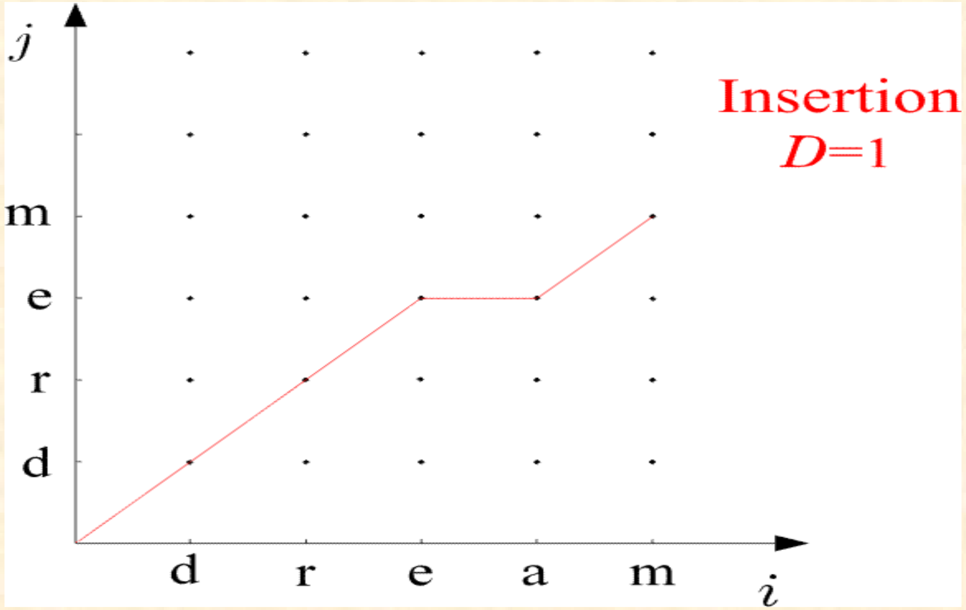
  - Horizontal

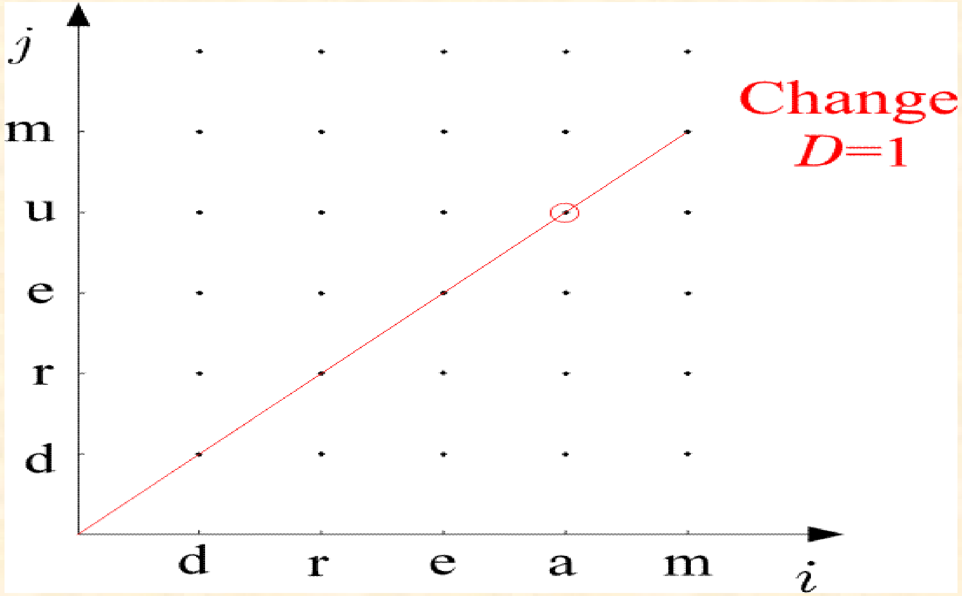
$$d(i, j \,|\, i-1, j) = 1$$

  - Vertical

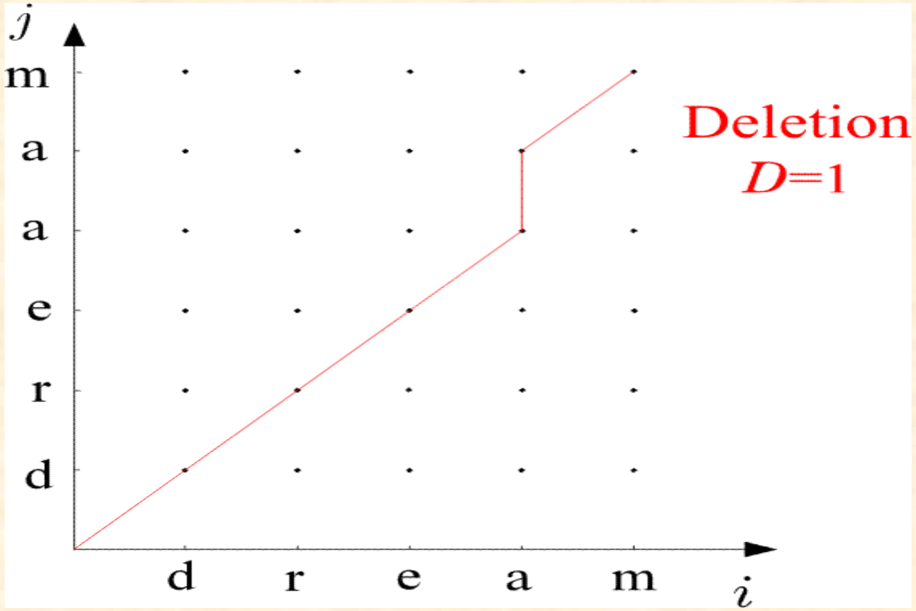
$$d(i, j \,|\, i, j-1) = 1$$

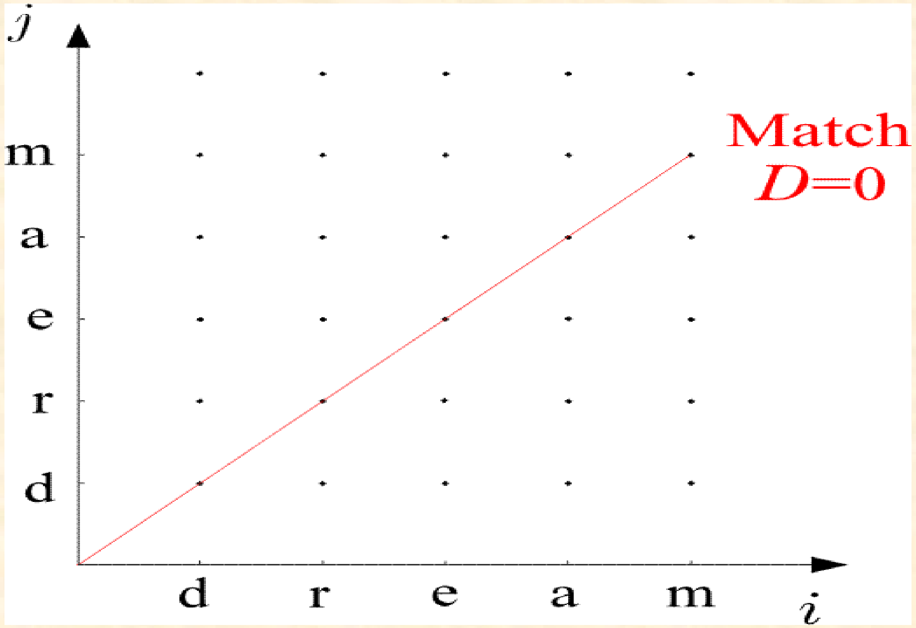
❖ Examples:





❖ Examples:



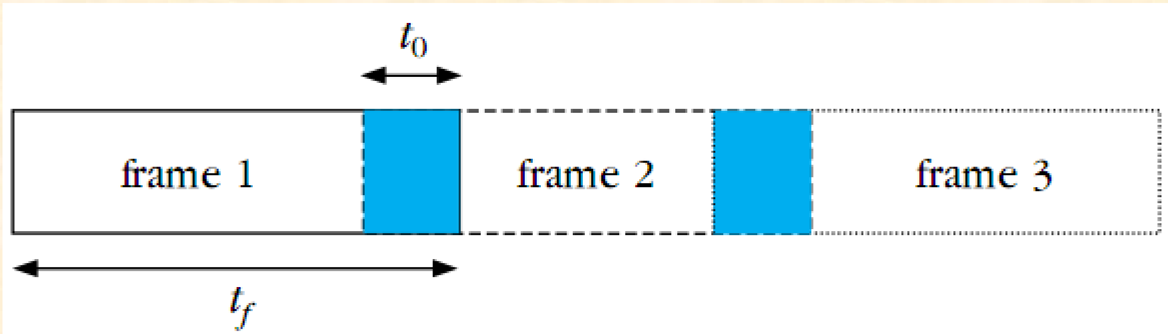


❖ Dynamic Time Warping in Speech Recognition

The isolated word recognition (IWR) will be discussed.

- The goal: Given a segment of speech corresponding to an unknown spoken word (test pattern), identify the word by comparing it against a number of known spoken words in a data base (reference patterns).

- The procedure:
  - Express the test and each of the reference patterns as sequences of feature vectors , $\underline{r}(i)$, $\underline{t}(j)$.
  - To this end, divide each of the speech segments in a number of successive frames.



- For each frame compute a feature vector. For example, the DFT coefficients and use, say, ℓ of those:

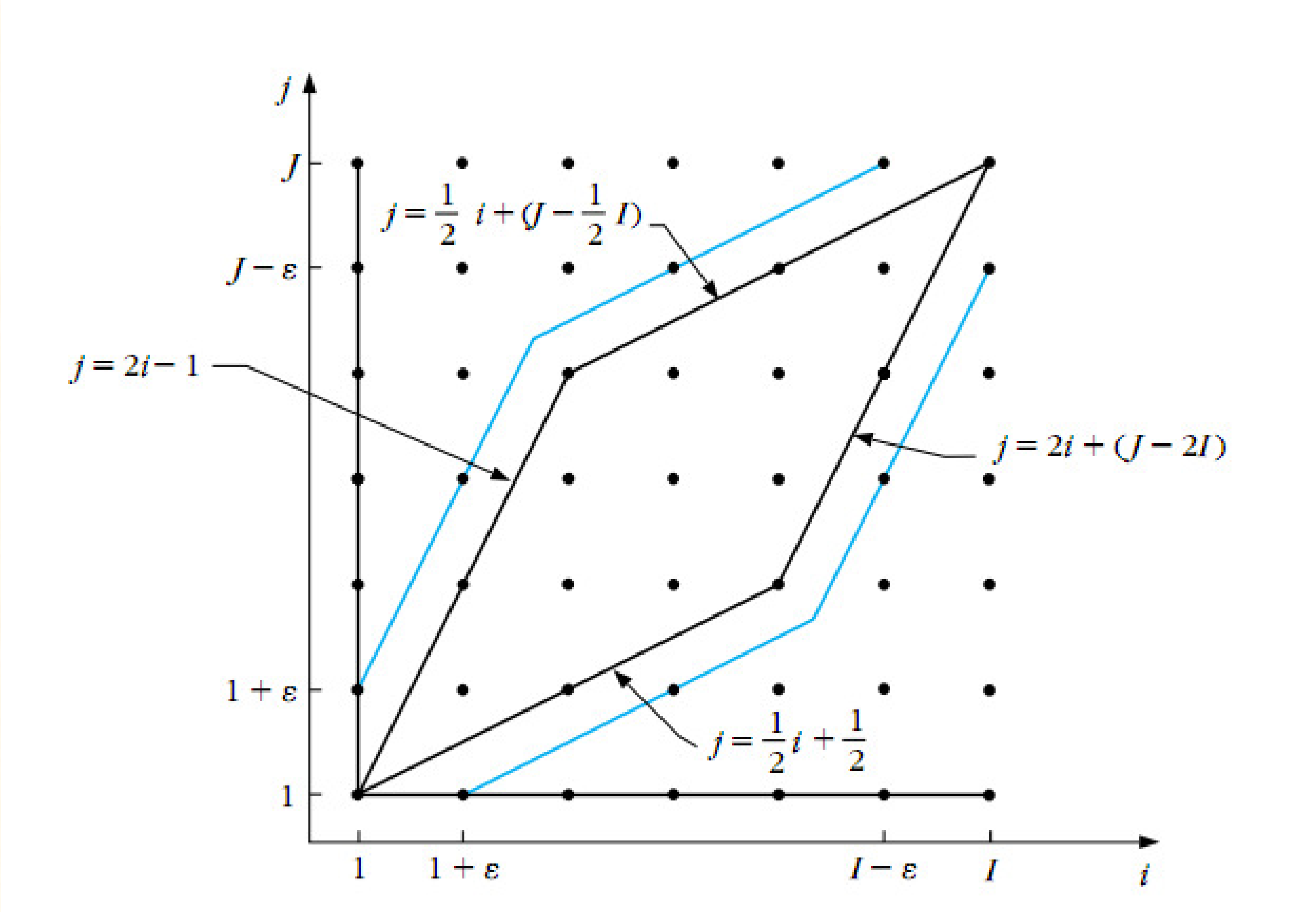
$$\underline{r}(i) = \begin{bmatrix} x_i(0) \\ x_i(1) \\ \dots \\ \dots \\ x_i(\ell - 1) \end{bmatrix}, \; i = 1, \dots, I \qquad \underline{t}(j) = \begin{bmatrix} x_j(0) \\ x_j(1) \\ \dots \\ \dots \\ x_j(\ell - 1) \end{bmatrix}, \; j = 1, \dots, J$$

- Choose a cost function associated with each node across a path, e.g., the Euclidean distance

$$\left\| \underline{r}(i_k) - \underline{t}(j_k) \right\| = d(i_k, j_k)$$

- For each reference pattern compute the optimal path and the associated cost, against the test pattern.
- Match the test pattern to the reference pattern associated with the minimum cost.

- Prior to computing the path one has to choose:
  - The global constraints: Defining the region of space within which the search for the optimal path will be performed.



- The local constraints: Defining the type of transitions allowed between the nodes of the grid.