# Probabilistic Graphical Models

## Lectures 11

Conditional Random Fields

# Remember: Generative Models

- System variables $X_1, X_2, \dots, X_N$
- Generative Model: Joint distribution $p(x_1, x_2, \dots, x_N)$
- If you have the joint distribution, you have everything

- Prediction:
    - Having $p(x,y,z) = Pr(X=x, Y=y, Z=z)$, predict $x,y,z$

# Remember: Generative Models

- System variables $X_1, X_2, \dots, X_N$
- Generative Model: Joint distribution $p(x_1, x_2, \dots, x_N)$
- If you have the joint distribution, you have everything

- Prediction:
  - Having $p(x,y,z) = Pr(X=x, Y=y, Z=z)$, predict x,y,z
  - Find the most likely configuration of system variables

# Remember: Generative Models

1928
K. N. Toosi University of Technology

- System variables $X_1, X_2, \dots , X_N$
- Generative Model: Joint distribution $p(x_1, x_2, \dots , x_N)$
- If you have the joint distribution, you have everything

- Prediction:
  - Having p(x,y,z) = Pr(X=x, Y=y, Z=z), predict x,y,z
  - Find the most likely configuration of system variables

$$x^*, y^*, z^* = \arg\max_{x,y,z} p(x, y, z)$$

# Remember: Generative Models

- System variables $X_1, X_2, \dots , X_N$
- Generative Model: Joint distribution $p(x_1, x_2, \dots , x_N)$
- If you have the joint distribution, you have everything

- Prediction:
    - Having $p(x,y,z) = Pr(X=x, Y=y, Z=z)$
    - If we know $Z = z_0$, predict x,y

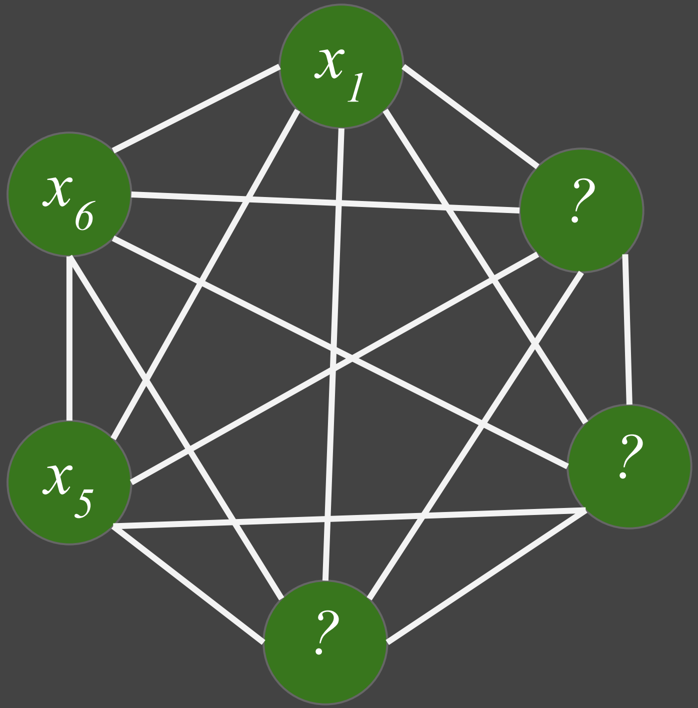
# Remember: Generative Models

- System variables $X_1, X_2, \dots, X_N$
- Generative Model: Joint distribution $p(x_1, x_2, \dots, x_N)$
- If you have the joint distribution, you have everything

- Prediction:
    - Having p(x,y,z) = Pr(X=x, Y=y, Z=z)
    - If we know $Z = z_0$, predict x,y

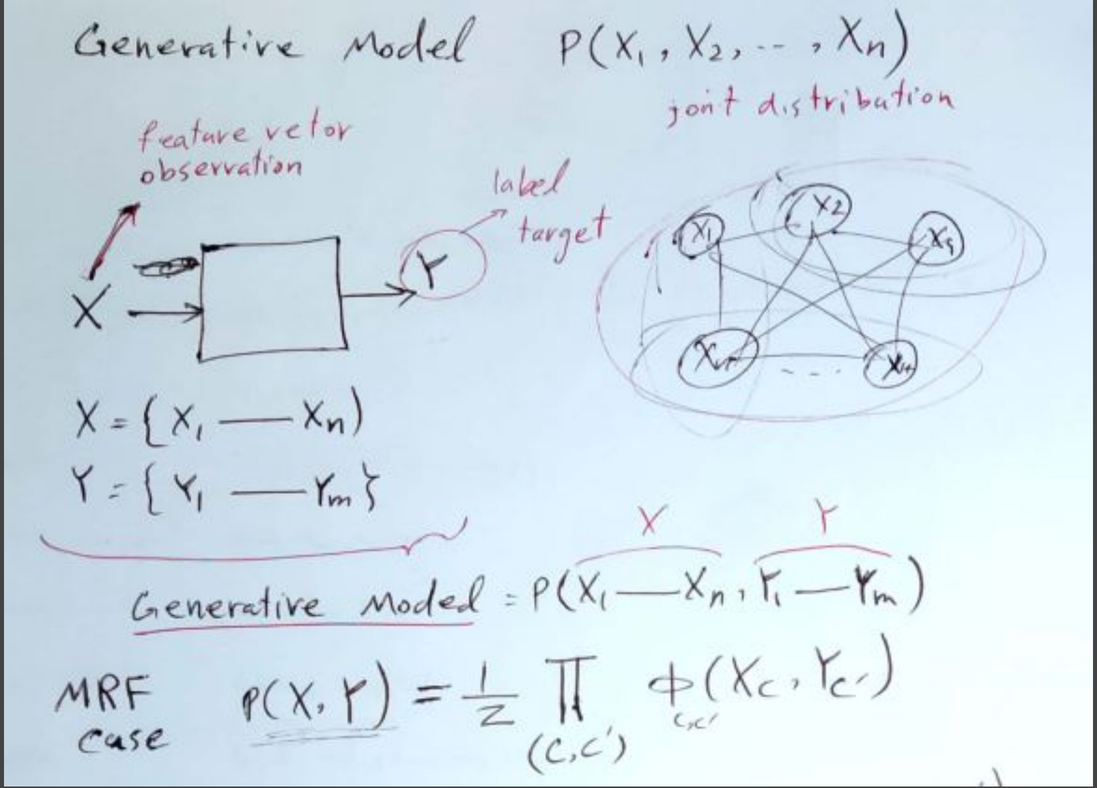
$$x, y = \arg\max_{x,y} p(x, y, z_0)$$

# Generative Models



# Generative Models

K. N. Toosi
University of Technology



Generative Model $P(X_1, X_2, \dots, X_n)$

joint distribution

feature vetor
observation

label
target

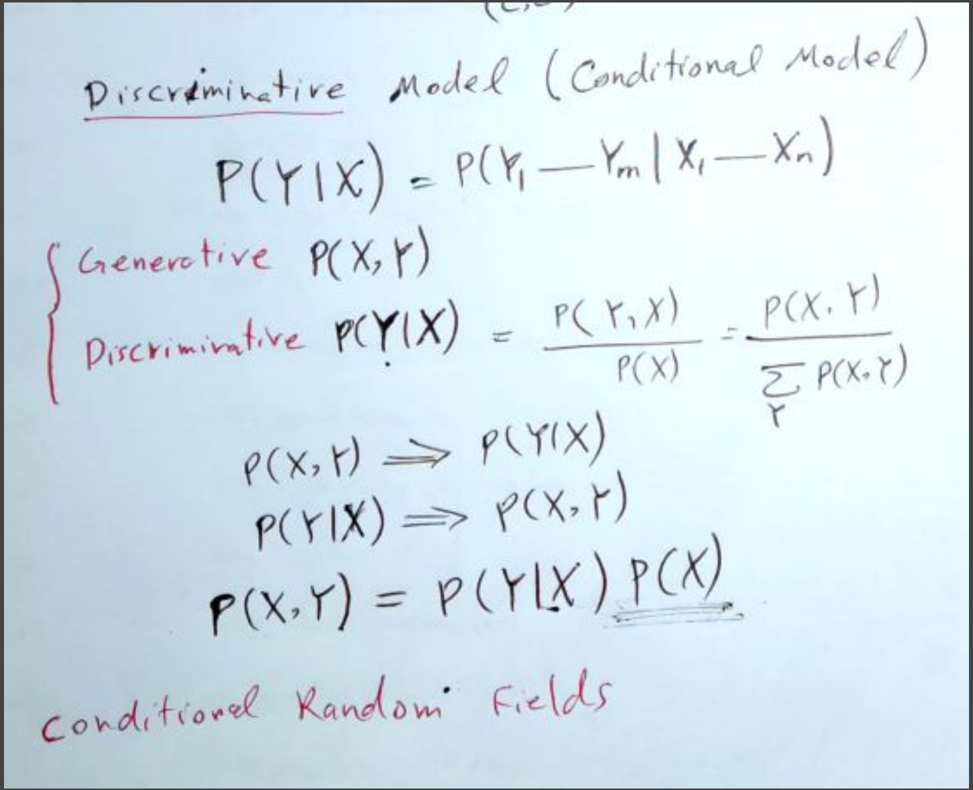
$X = \{X_1 —— X_n)$

$Y = \{Y_1 —— Y_m\}$

Generative Model $= P(\overbrace{X_1 —— X_n}^{X}, \overbrace{Y_1 —— Y_m}^{Y})$

MRF
case $P(X, Y) = \frac{1}{Z} \prod_{(c,c')} \phi_{c,c'}(X_c, Y_{c'})$

# Generative vs Discriminative Models



Discriminative Model (Conditional Model)

$$P(Y|X) = P(Y_1 — Y_m | X_1 — X_n)$$

Generative $P(X,Y)$

Discriminative $P(Y|X) = \frac{P(Y,X)}{P(X)} = \frac{P(X,Y)}{\sum_{Y} P(X,Y)}$

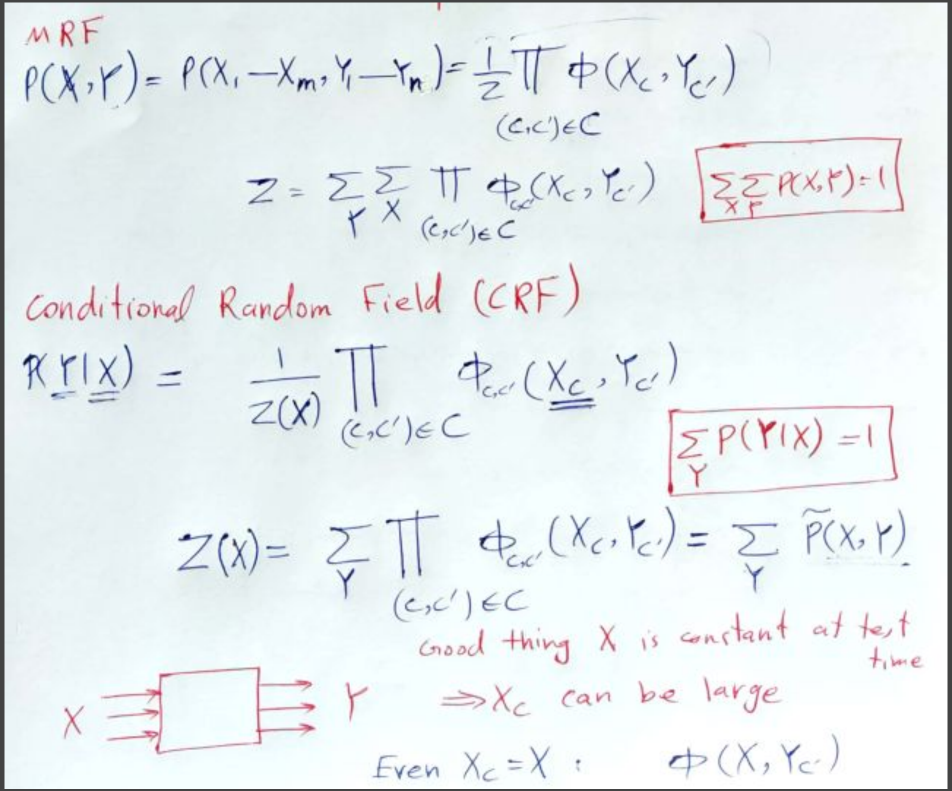
$$P(X,Y) \Longrightarrow P(Y|X)$$

$$P(Y|X) \Longrightarrow P(X,Y)$$

$$P(X,Y) = P(Y|X)\,\underline{\underline{P(X)}}$$

Conditional Random Fields

# Conditional Random Fields



MRF

$$P(X,Y) = P(X_1 — X_m, Y_1 — Y_n) = \frac{1}{Z} \prod_{(c,c')\in C} \phi(X_c, Y_{c'})$$

$$Z = \sum_{Y} \sum_{X} \prod_{(c,c')\in C} \phi_{cc'}(X_c, Y_{c'}) \qquad \boxed{\sum_{X}\sum_{Y} P(X,Y) = 1}$$

Conditional Random Field (CRF)

$$P(\underline{Y}|\underline{\underline{X}}) = \frac{1}{Z(X)} \prod_{(c,c')\in C} \phi_{c,c'}(\underline{\underline{X_c}}, Y_{c'})$$

$$\boxed{\sum_{Y} P(Y|X) = 1}$$

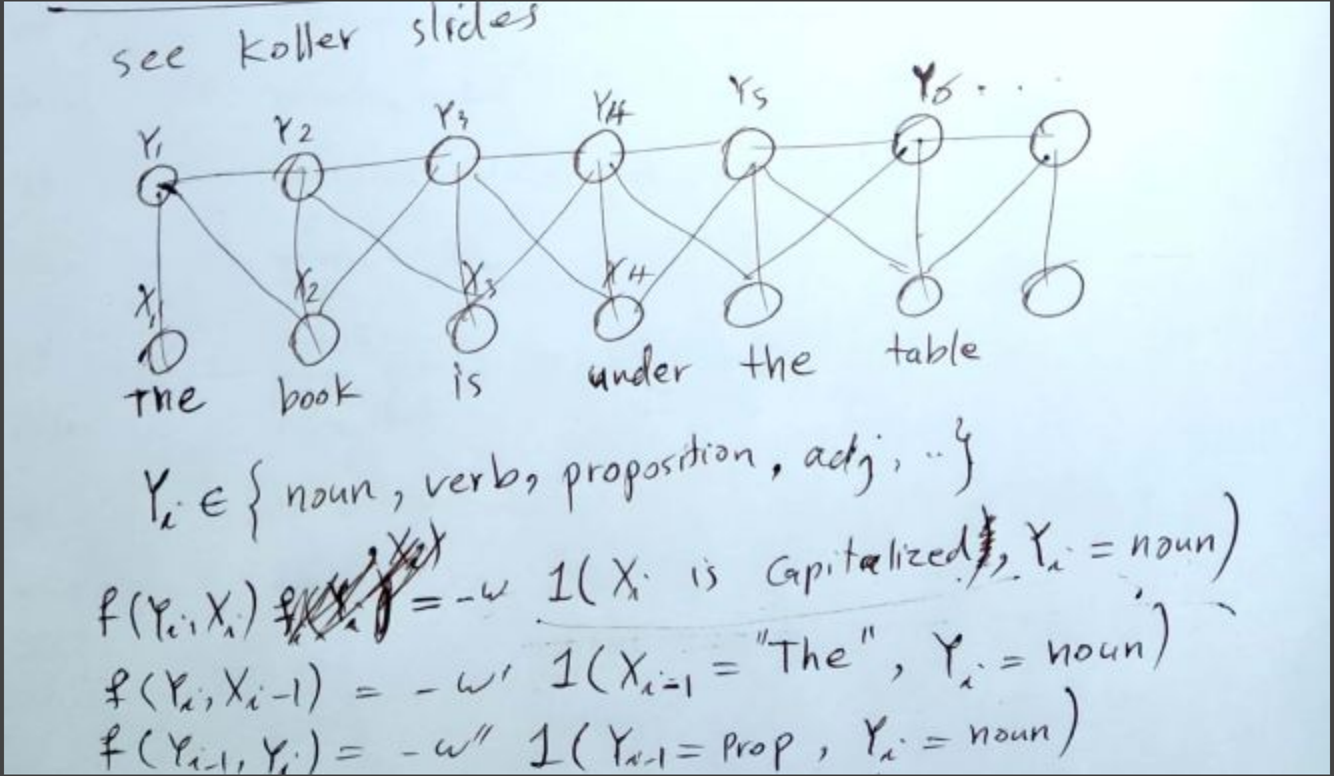
$$Z(X) = \sum_{Y} \prod_{(c,c')\in C} \phi_{c,c'}(X_c, Y_{c'}) = \sum_{Y} \tilde{P}(X,Y)$$

Good thing X is constant at test time

X → → Y $\Rightarrow X_c$ can be large

Even $X_c = X$ : $\phi(X, Y_{c'})$

# Example



see koller slides

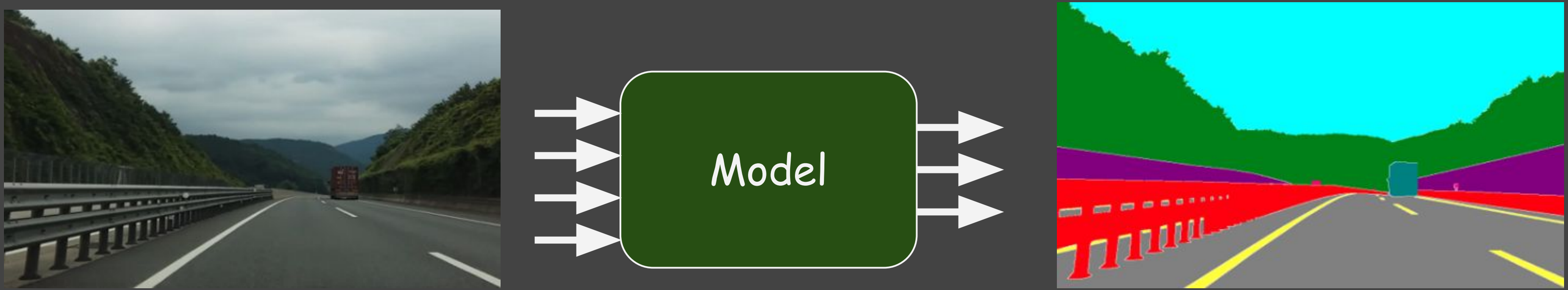
$Y_i \in \{ noun, verb, proposition, adj, .. \}$

$f(Y_i, X_i) = -w \; 1(X_i \text{ is Capitalized}, Y_i = noun)$

$f(Y_i, X_i - 1) = -w' \; 1(X_{i-1} = \text{"The"}, Y_i = noun)$

$f(Y_{i-1}, Y_i) = -w'' \; 1(Y_{i-1} = Prop, Y_i = noun)$

# Example - Image Segmentation

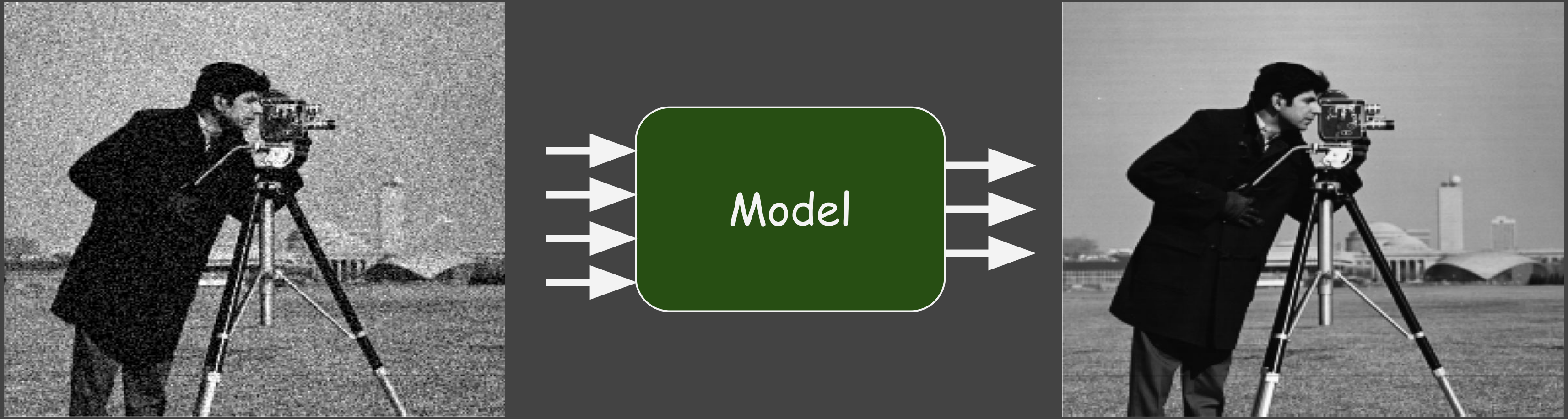


Model

https://sites.google.com/site/highwaydrivingdataset

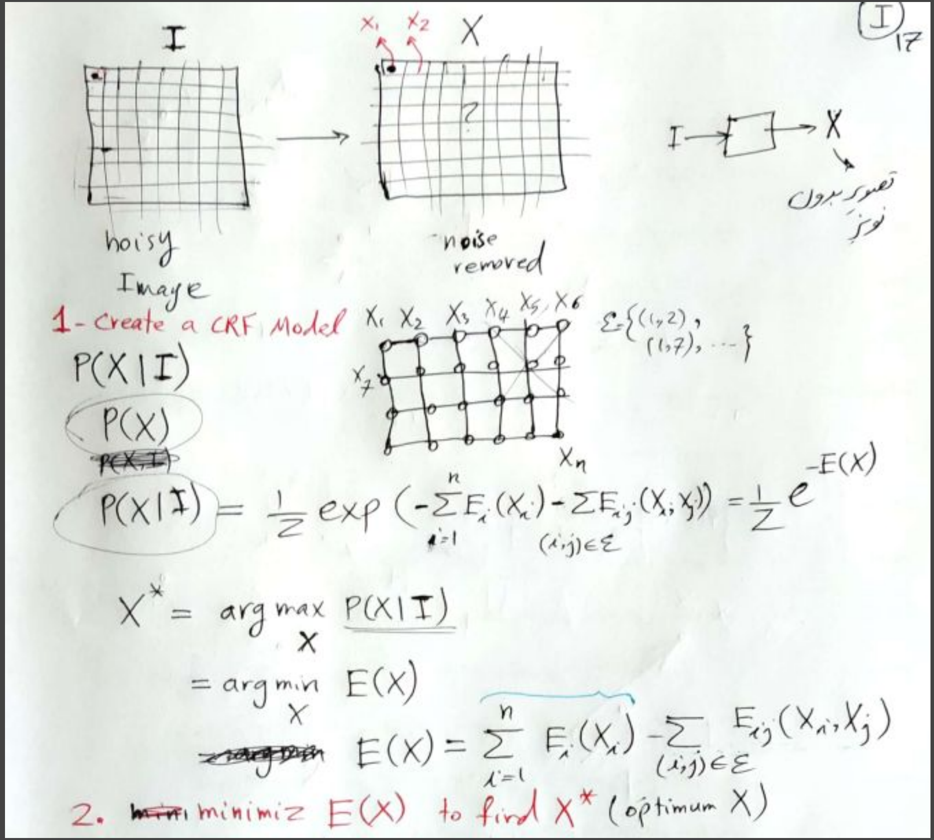
# Example - Image Denoising

K. N. Toosi
University of Technology



Model

# Example: Noise removal



I

noisy Image

$x_1$ $x_2$ X

?

noise removed

$I \rightarrow \square \rightarrow X$

تصویر بدون نویز

ⓘ 17

1- Create a CRF Model

$P(X|I)$

$P(X)$

~~$P(X,I)$~~

$X_1$ $X_2$ $X_3$ $X_4$ $X_5$ $X_6$

$X_7$

$X_n$

$\mathcal{E} = \{(1,2), (1,7), \cdots\}$

$$P(X|I) = \frac{1}{Z} \exp\left(-\sum_{i=1}^{n} E_i(X_i) - \sum_{(i,j)\in\mathcal{E}} E_{ij}(X_i,X_j)\right) = \frac{1}{Z} e^{-E(X)}$$

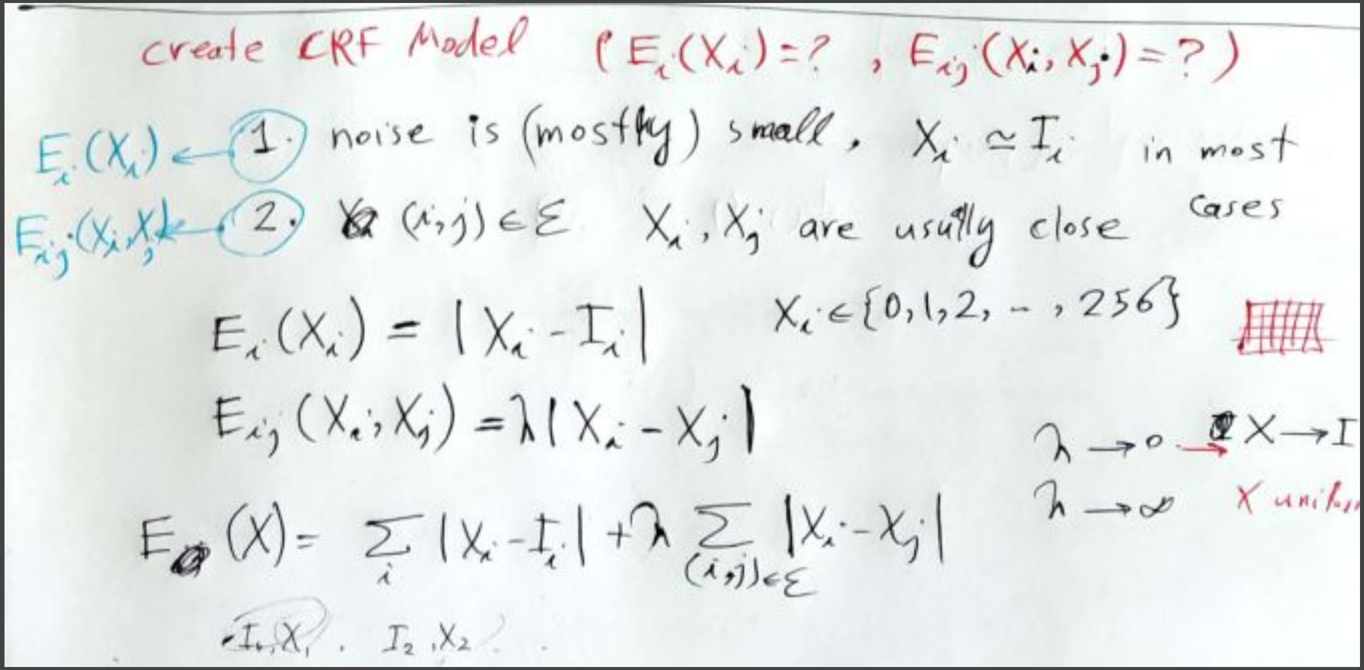
$$X^* = \underset{X}{\arg\max}\ P(X|I)$$

$$= \underset{X}{\arg\min}\ E(X)$$

~~argmin~~ $$E(X) = \sum_{i=1}^{n} E_i(X_i) - \sum_{(i,j)\in\mathcal{E}} E_{ij}(X_i,X_j)$$

2. ~~mini~~ minimiz $E(X)$ to find $X^*$ (optimum X)

# Example: Noise removal



create CRF Model ( $E_i(X_i) = ?$ , $E_{ij}(X_i, X_j) = ?$ )

$E_i(X_i) \leftarrow$ ① noise is (mostly) small, $X_i \simeq I_i$ in most cases

$E_{ij}(X_i, X_j) \leftarrow$ ② $(i,j) \in \mathcal{E}$ $X_i, X_j$ are usully close

$$E_i(X_i) = |X_i - I_i| \qquad X_i \in \{0, 1, 2, \dots, 256\}$$

$$E_{ij}(X_i, X_j) = \lambda |X_i - X_j|$$

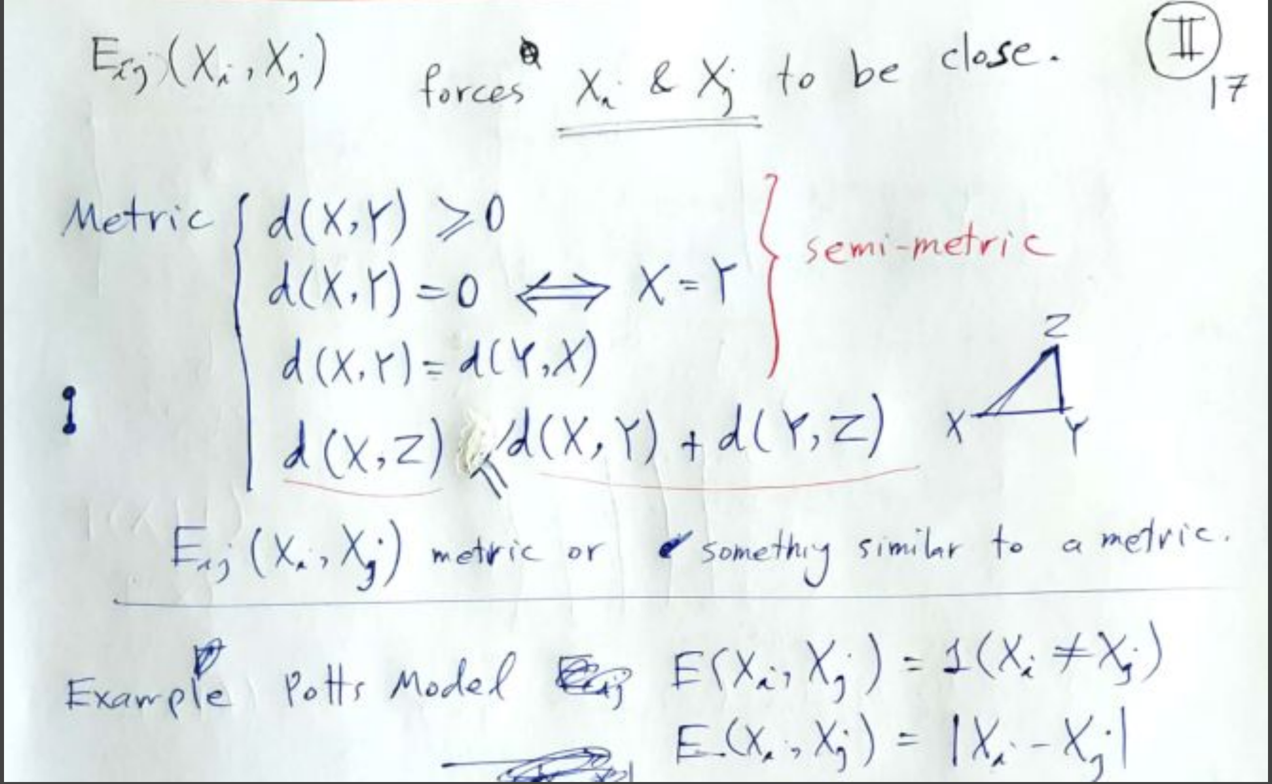
$\lambda \to 0 \to X \to I$

$\lambda \to \infty$ X unif

$$E(X) = \sum_i |X_i - I_i| + \lambda \sum_{(i,j) \in \mathcal{E}} |X_i - X_j|$$

$I_1, X_1$ . $I_2, X_2$ . .

# Example: Noise removal



Ⅱ 17

$E_{ij}(X_i, X_j)$ forces $X_i$ & $X_j$ to be close.

Metric
$$\begin{cases} d(X,Y) \geq 0 \\ d(X,Y) = 0 \iff X = Y \\ d(X,Y) = d(Y,X) \\ d(X,Z) \leq d(X,Y) + d(Y,Z) \end{cases}$$

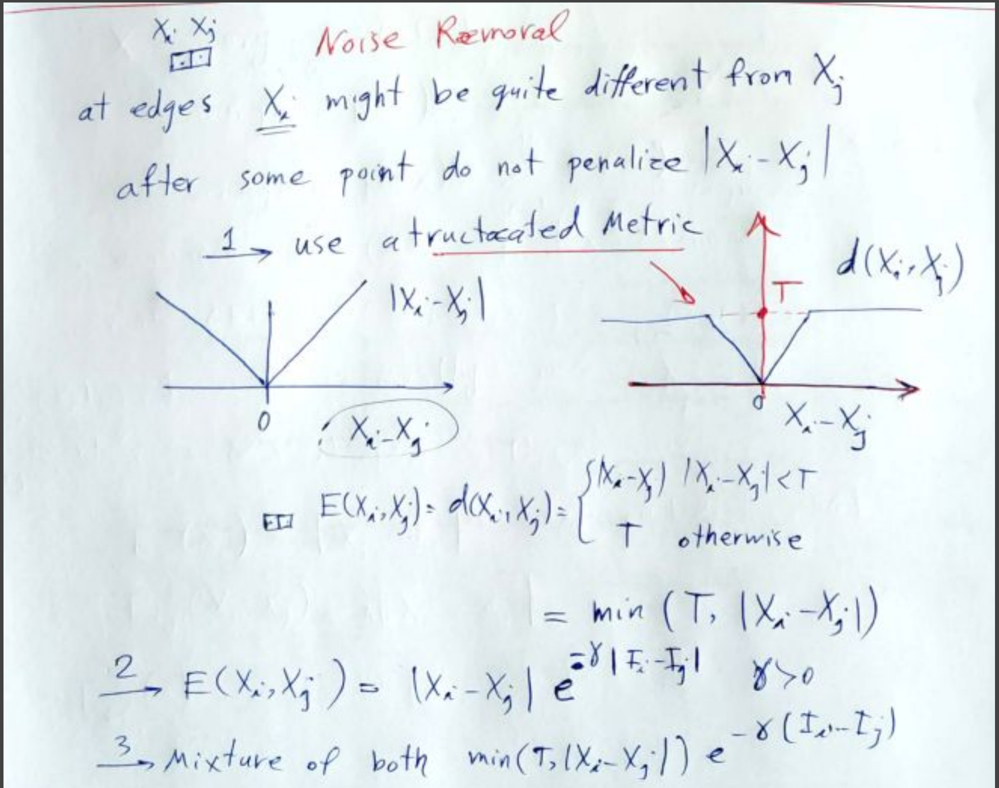
semi-metric

$E_{ij}(X_i, X_j)$ metric or something similar to a metric.

Example Potts Model ~~$E_{ij}$~~ $E(X_i, X_j) = \mathbb{1}(X_i \neq X_j)$

$E(X_i, X_j) = |X_i - X_j|$

# Example: Noise removal



$X_i$ $X_j$

## Noise Removal

at edges $X_i$ might be quite different from $X_j$

after some point do not penalize $|X_i - X_j|$

$\xrightarrow{1}$ use a truncated Metric

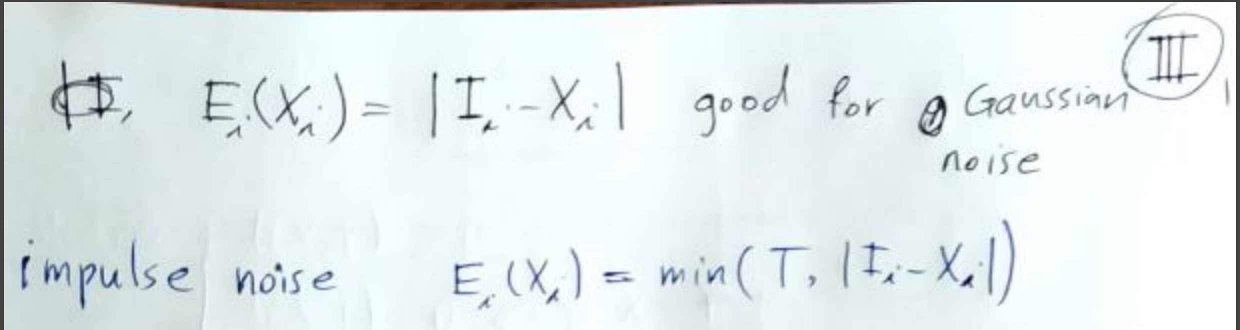
$$E(X_i, X_j) = d(X_i, X_j) = \begin{cases} (X_i - X_j) & |X_i - X_j| < T \\ T & \text{otherwise} \end{cases}$$

$$= \min(T, |X_i - X_j|)$$

$\xrightarrow{2}$ $E(X_i, X_j) = |X_i - X_j| \, e^{-\gamma |I_i - I_j|}$ $\gamma > 0$

$\xrightarrow{3}$ Mixture of both $\min(T, |X_i - X_j|) \, e^{-\gamma (I_i - I_j)}$

# Example: Noise removal



III

$\Phi$, $E_i(X_i) = |I_i - X_i|$ good for Gaussian noise

impulse noise $E_i(X_i) = \min(T, |I_i - X_i|)$

# Example - Stereo Matching

https://www.researchgate.net/publication/326505595_Semi-Dense_3D_Reconstruction_with_a_Stereo_Event_Camera

# Example - Stereo Matching

K. N. Toosi
University of Technology

# Example - Stereo Matching

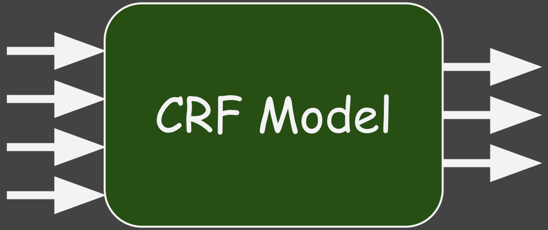
K. N. Toosi
University of Technology

# Example - Stereo Matching

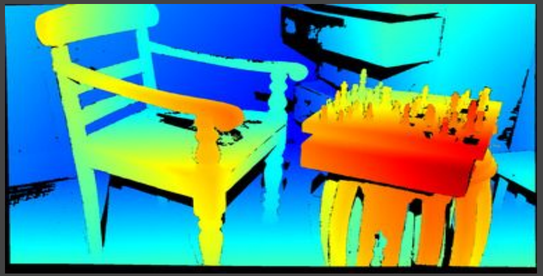
# Example - Stereo Matching

*left image*

*right image*



CRF Model

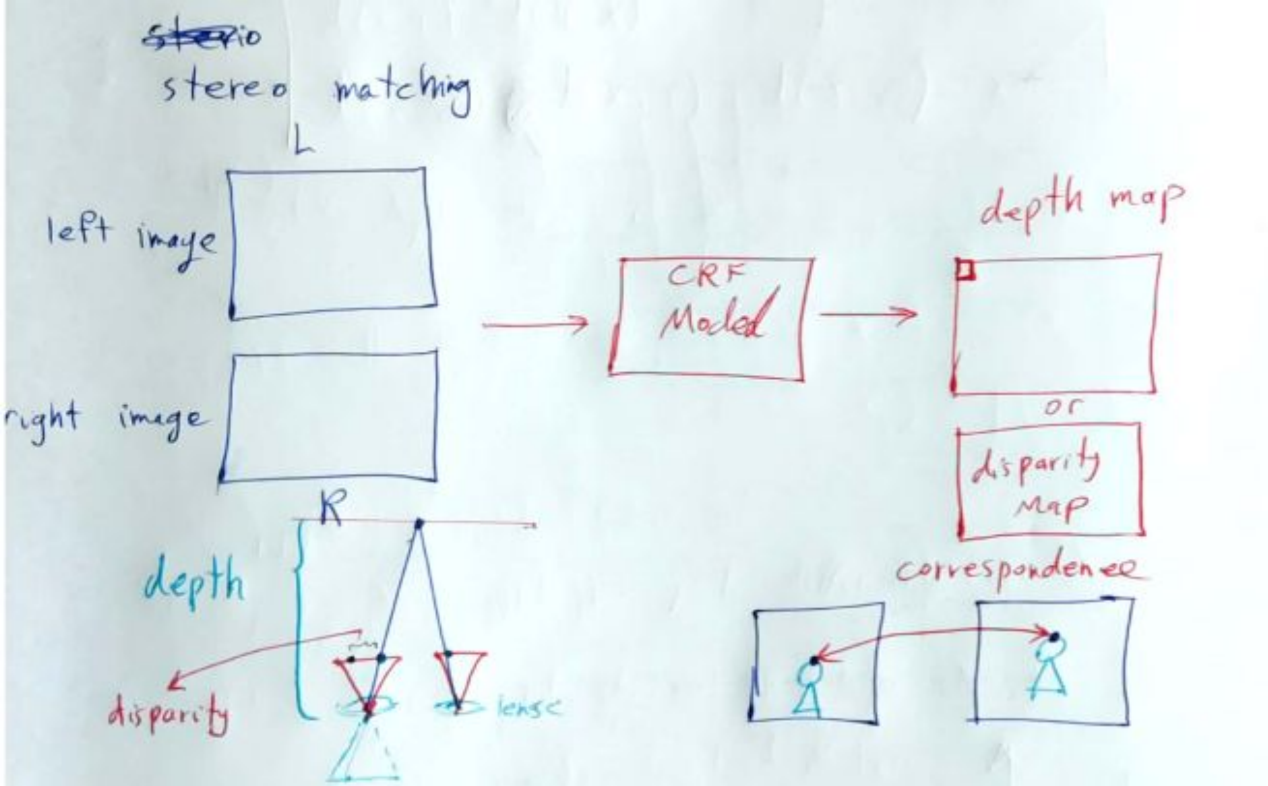


*depth map*

*OR*

*disparity map*

# Example: Stereo Matching



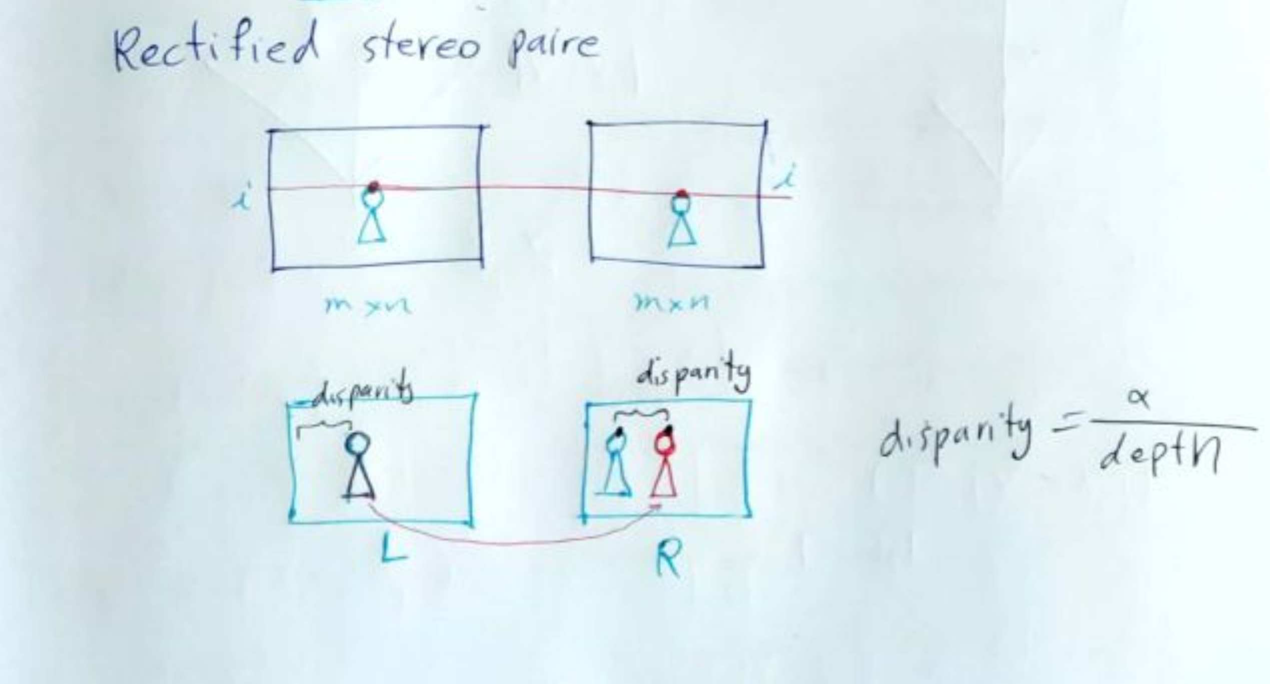
# Rectified stereo and disparity

# Rectified stereo and disparity

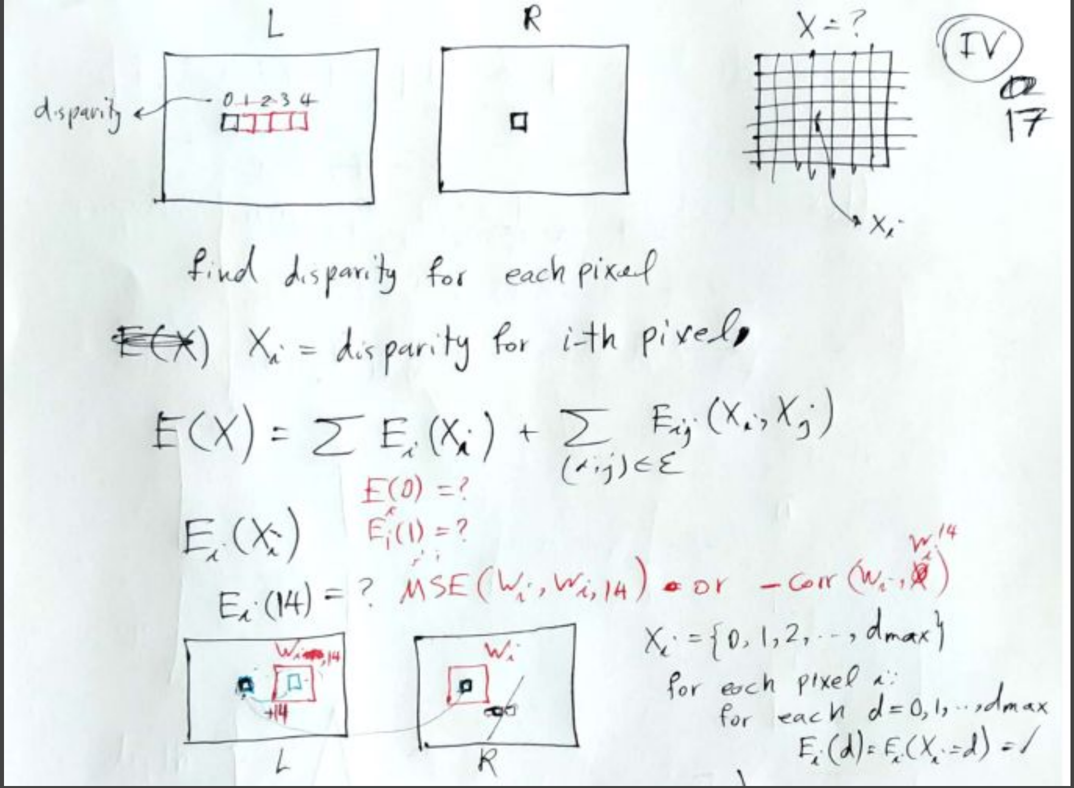
# Rectified stereo and disparity

# Rectified stereo and disparity

# Example: Stereo Matching



# Example: Stereo Matching



find disparity for each pixel

~~E(X~~) $X_i$ = disparity for i-th pixel,

$$E(X) = \sum E_i(X_i) + \sum_{(i,j)\in\mathcal{E}} E_{ij}(X_i, X_j)$$

$E_i(X_i)$

$E_i(0) = ?$

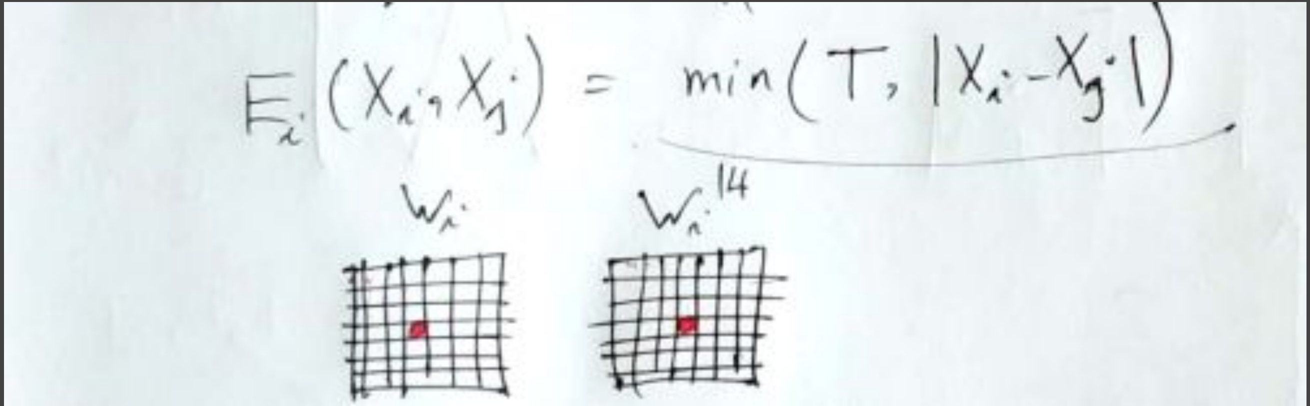
$E_i(1) = ?$

$E_i(14) = ?$ $MSE(W_i, W_{i,14})$ or $-corr(W_i, W_i^{14})$

$X_i = \{0, 1, 2, \dots, d_{max}\}$

```
for each pixel i:
    for each d = 0, 1, ..., dmax
        $E_i(d) = E_i(X_i = d) = \checkmark$
```

# Example: Stereo Matching



$$E_i(X_i, X_j) = \underbrace{\min(T, |X_i - X_j|)}$$

$W_i$ $\qquad$ $W_i^{14}$