Independent Components Analysis

1.discriminative model and generative model

MLP:

$$p(\boldsymbol{t}|w)=\prod_{n=1}^{N} y_n^{t_n}\{1-y_n\}^{1-t_n}$$

RBM:

$$P(\boldsymbol{x},\boldsymbol{h})=\frac{e^{Energy}(\boldsymbol{x},\boldsymbol{h})}{\boldsymbol{Z}}$$

2.Principal component analysis(PCA)

Goal: using orthogonal transformation to convert the high-dimensional data to low-dimensional linearly uncorrelated variables.

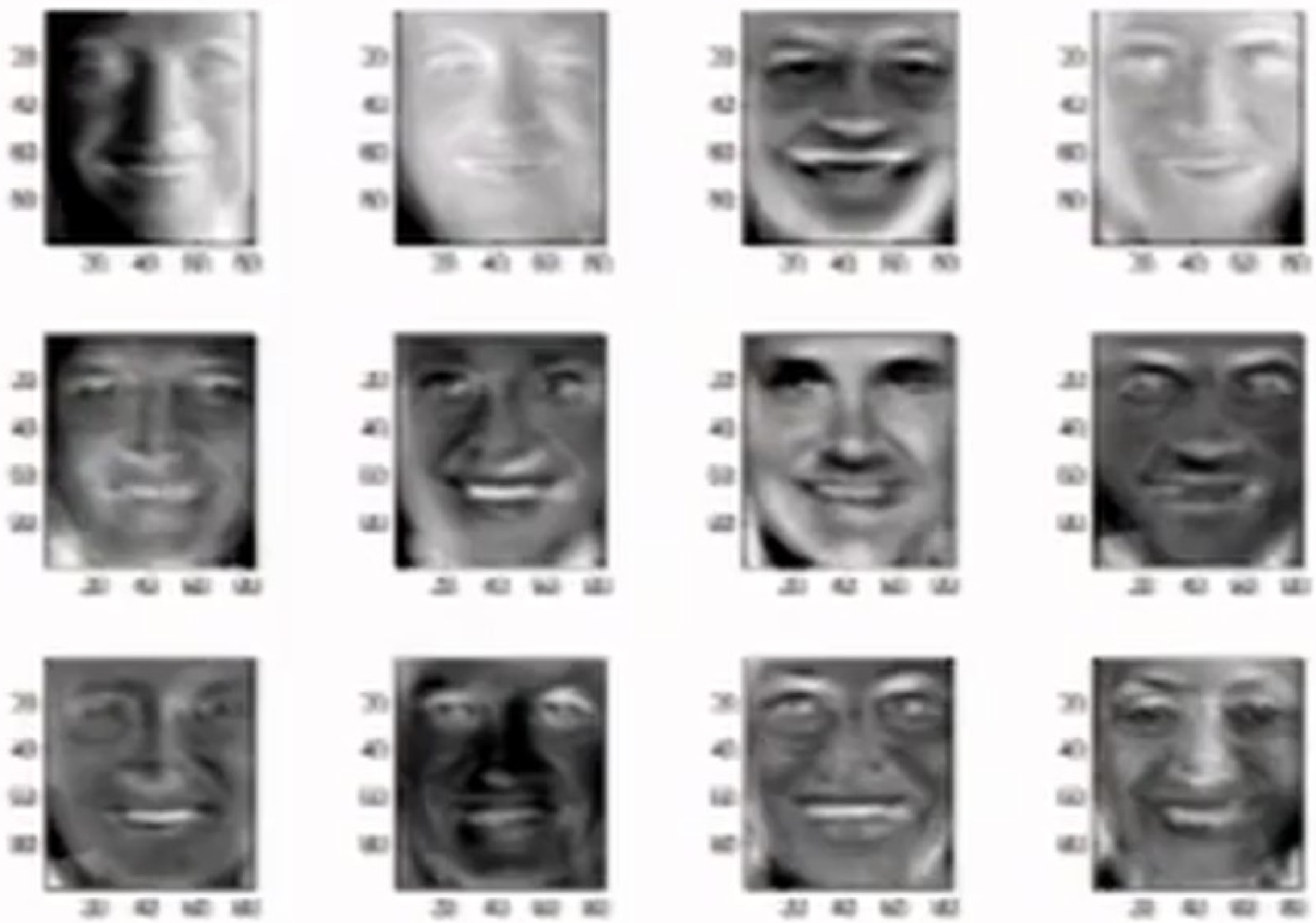
$$I(x,y)=\sum_{i=1}^{n} A_i(x,y)s_i$$

I(x,y) is the image pixels of one patch. Assume n equals to the number of pixels.

x,y means the pixel position.

Each I(x,y) is the linear weighted sum of features.

2.PCA



2.PCA

$$s_i = \sum_{(x,y)} W_i(x,y) I(x,y)$$

si is the projection of I(x,y)

maximize $var(s) = E\{s^2\} - (E\{s\})^2$

same to maximize

$$\frac{1}{T}\sum_{i=1}^{T}\left(\sum_{x,y} W(x,y) I_t(x,y)\right)^2$$

constraints: $\|W\| = \sqrt{\sum_{x,y} W(x,y)^2} = 1$

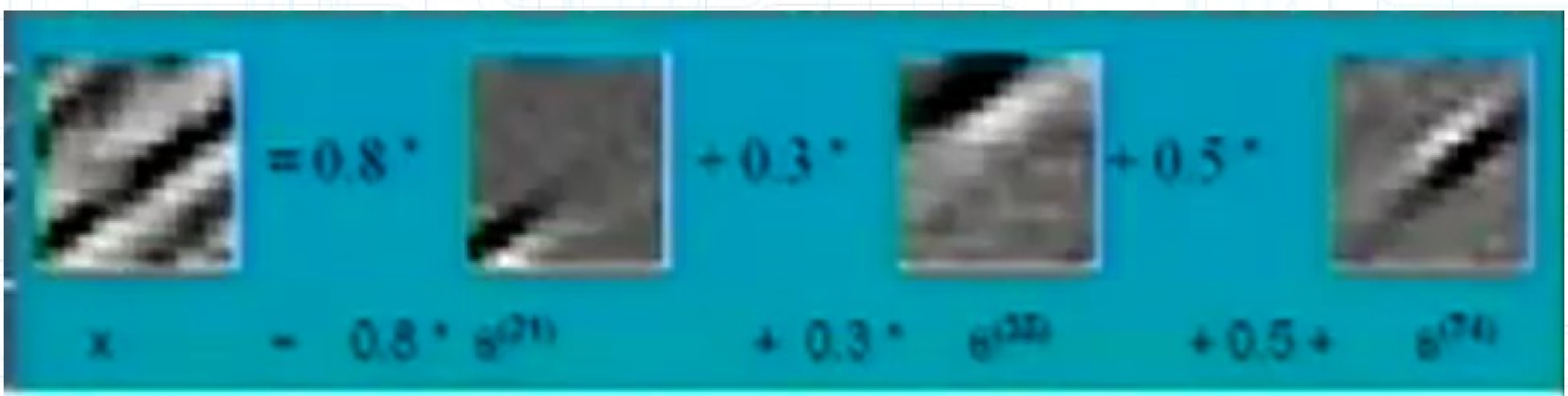
$$\sum_{x,y} W_j(x,y) W(x,y) = 0 \; forall \; j = 1, \cdots, k$$

4.Independent Components Analysis

Goal: Separating the some unknown latent and independent variables of the input.

eg. sound signal

In the image system is the **si**.



4.ICA

Three assumption:

1.si are statistically independent

2.si are non-gaussian

3.the linear system defined by Ai is invertible

4.ICA

After preprocessing, we get zi

$$z_i = \sum_{j=1}^{m} b_{ij} s_j$$

Transformation: $s_i = \sum_{j=1}^{n} v_{ij} z_j$

Assume we know the probability density functions of si:

$$p(s_{1,} \cdots, s_n) = \prod_{i=1}^{n} p_i(s_i)$$

4.ICA

Then the probability density functions of zi is:

$$p(\boldsymbol{z})=|det(\boldsymbol{V})|\prod_{i=1}^{n} p_i(\boldsymbol{v_i^T z})=|det(\boldsymbol{V})|\prod_{i=1}^{n} p_i(\sum_{j=1}^{n} v_{ij} z_j)$$

For T times of observations,likelihood is

$$L(\boldsymbol{v_1},\cdots,\boldsymbol{v_n})=\prod_{t=1}^{T} p(\boldsymbol{z_t})=\prod_{t=1}^{T}\left[|det(\boldsymbol{V})|\prod_{i=1}^{n} p_i(\boldsymbol{v_i^T z})\right]$$

Get the logarithm

$$\log L(\boldsymbol{v_1},\cdots,\boldsymbol{v_n})=T\log|det(\boldsymbol{V})|+\sum_{i=1}^{n}\sum_{t=1}^{T}\log p_i(\boldsymbol{v_i^T z})$$

5.R(reconstruction)ICA

Two drawbacks: difficult to learn overcomplete features and sensitive to whitening

ICA:

$$\underset{W}{minimize} \sum_{i=1}^{m} \sum_{j=1}^{k} g(W_j x^{(i)}) \quad \text{subject to } WW^T = I$$

RICA:

$$\underset{W}{minimize} \frac{\lambda}{m} \sum_{i=1}^{m} \| W^T W x^{(i)} - x^{(i)} \|_2^2 + \sum_{i=1}^{m} \sum_{j=1}^{k} g(W_j x^{(i)})$$

$$\underset{W}{minimize} \lambda \| W W^T - I \|_F^2 + \sum_{i=1}^{m} \sum_{j=1}^{k} g(W_j x^{(i)})$$

5.several questions

Why we need white the component?

What is the specific process to solve the principal component?

The proof of two lemma?