Similarity ~~and Dissimilarity~~ Measures

Problem statement

Observed images

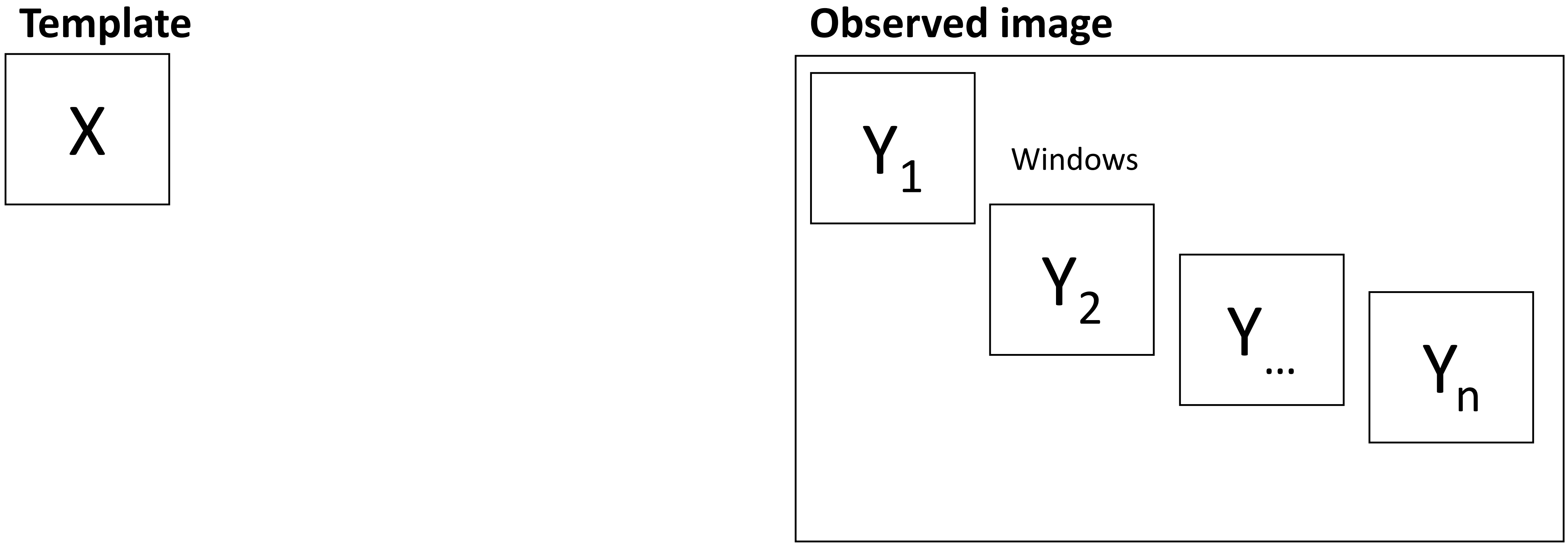
X

Saved images

Y_1 Y_2 $Y_{\dots}$ Y_n

Try to find the most similar to image X from saved images!

Problem statement (2)



Try to find the best matching window in an observed image!

More general statement

- X and Y are sets (doesn’t have to be images)
- Find a measure/metric on how similar X and Y are.

Images

We are often using 1D coordinates instead of 2D

X

$$\begin{matrix} x_1 & x_2 & x_3 & x_4 \\ x_5 & x_6 & x_7 & x_8 \\ x_9 & x_{10} & x_{11} & x_{12} \\ x_{13} & x_{14} & x_{15} & x_{16} \end{matrix}$$

Y

$$\begin{matrix} y_1 & y_2 & y_3 & y_4 \\ y_5 & y_6 & y_7 & y_8 \\ y_9 & y_{10} & y_{11} & y_{12} \\ y_{13} & y_{14} & y_{15} & y_{16} \end{matrix}$$

Similarity measures

- Pearson correlation coefficient
- Tanimoto measure
- Stochastic sign change
- Deterministic sign change
- Minimum ratio
- Spearman’s rho
- Kendall’s tau
- Greates deviation
- Ordinal measure
- Correlation ratio
- Energy of joint probability distribution
- Material similarity
- Shannon mutual information
- Rényi mutual information
- Tsallis mutual information
- F-Information measures

Pearson correlation coefficient

- $r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} * \sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$, $\quad r = \frac{\sum xdiff * ydiff}{xdiff\ distance * ydiff\ distance}$
- $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i,$ The moment of x and y (average)
- $r = \frac{1}{n}\sum\left(\frac{x_i - \bar{x}}{\sigma_x}\right)\left(\frac{y_i - \bar{y}}{\sigma_y}\right)$
- r = 1 -> perfect positive correlation
- r = -1 -> perfect negative correlation

Pearson correlation coefficient (2)

- Can be extended to work with different rotation [1] and scale [2,3]

[1] De Castro, E., Morandi, C.: Registration of translated and rotated images using finite Fourier transforms. IEEE Trans. Pattern Anal. Mach. Intell. **9**(5), 700–703 (1987)

[2] Chen, Q.-S.: Matched filtering techniques. In: Le Moigne, J., Netanyahu, N.S., Eastman, R.D. (eds.) Image Registration for Remote Sensing, pp. 112–130. Cambridge University Press, Cambridge (2011)

[3] Reddy, B.S., Chatterji, B.: An FFT-based technique for translation, rotation and scale invariant image registration. IEEE Trans. Image Process. **5**(8), 1266–1271 (1996)

Pearson correlation coefficient (optimizations for template matching)

- *Require data is 2D image*
  - $C = F^{-1}[F(X) * F^{*}(Y)]$
- Phase correlation:
  - $C_p = F^{-1}\left[\frac{F \circ G^{*}}{|F \circ G^{*}|}\right]$

- Convolution/correlation corresponds to multiplication in frequency space!
- To be explained
- $\circ$ - The Hadamard product

Pearson correlation coefficient (Phase correlation)

- From signal processing, we know… (probably not):
  - A shift in image space corresponds a phase multiplication in frequency space
  - Shift: $f(x - \Delta x) \rightarrow$ Multiply: $F(x)e^{\phi}$. ϕ is some phase stuff
- $g_b(x, y) = g_a\big((x - \Delta x) mod\ M, (y\ - \Delta y) mod\ N\big)$
- $G_b(u, v) = G_a e^{-2\pi i\left(\frac{u\Delta x}{M} + \frac{v\Delta y}{N}\right)}$

[4]: Wikipedia | Phase correlation | https://en.wikipedia.org/wiki/Phase_correlation

Pearson correlation coefficient (Phase correlation derivation)

- $R(u,v) = \frac{G_a \circ G_b^*}{|G_a \circ G_b^*|}$
- $= \frac{G_a \circ G_a^* e^{-2\pi i\left(\frac{u\Delta x}{M}+\frac{v\Delta y}{N}\right)}}{\left|G_a \circ G_a^* e^{-2\pi i\left(\frac{u\Delta x}{M}+\frac{v\Delta y}{N}\right)}\right|}$

Doesn't change length

- $= \frac{G_a \circ G_a^* e^{-2\pi i\left(\frac{u\Delta x}{M}+\frac{v\Delta y}{N}\right)}}{|G_a \circ G_a^*|}$
- $= \text{stuff} \times e^{-2\pi i\left(\frac{u\Delta x}{M}+\frac{v\Delta y}{N}\right)}$
- Stuff has no phase, so transforming back gives only offset!

[4]: Wikipedia | Phase correlation | https://en.wikipedia.org/wiki/Phase_correlation

Pearson correlation coefficient (Orientation correlation)

- Use gradients to avoid problems with absolute level differences.
- $X \rightarrow U_d(x, y) = sign\left(\frac{\partial X(x,y)}{\partial x} + j\,\frac{\partial X(x,y)}{\partial y}\right)$
- $Y \ \rightarrow V_d(x, y) = sign\left(\frac{\partial Y(x,y)}{\partial x} + j\,\frac{\partial Y(x,y)}{\partial y}\right)$
- Then apply fourier stuff on images.

Pearson correlation coefficient (Takeaway)

- Relatively simple operator and simple to understand.
- Can be extended to work with different scale and rotation.
- Consider using frequency space for large images. (Fourier transform)
- Consider using phase correlation for certain images.
- Consider using orientation correlation if images have significant level differences.

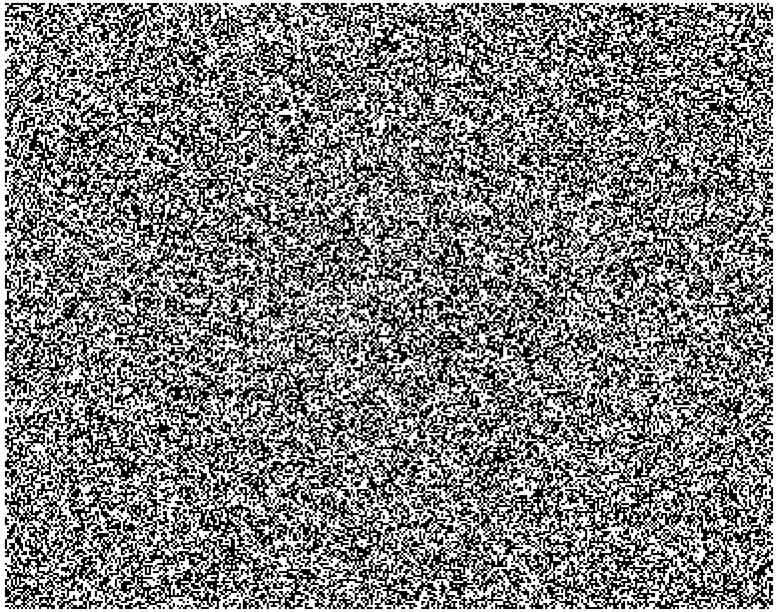
Tanimoto Measure

- $S_T = \frac{X^t Y}{\|X\|^2 + \|Y\|^2 - X^t Y} = \frac{X^t Y}{\|X - Y\| + X^t Y}$
- Gives the same result as Pearson correlation coefficient
- $\|X - Y\|$ corresponds to $\sigma_x \sigma_y$ in the Pearsons coefficient.
- Division by $X^t Y$ corresponds to normalization regarding the mean in the Pearson coefficient.
- No need to calculate σ_x and σ_y.

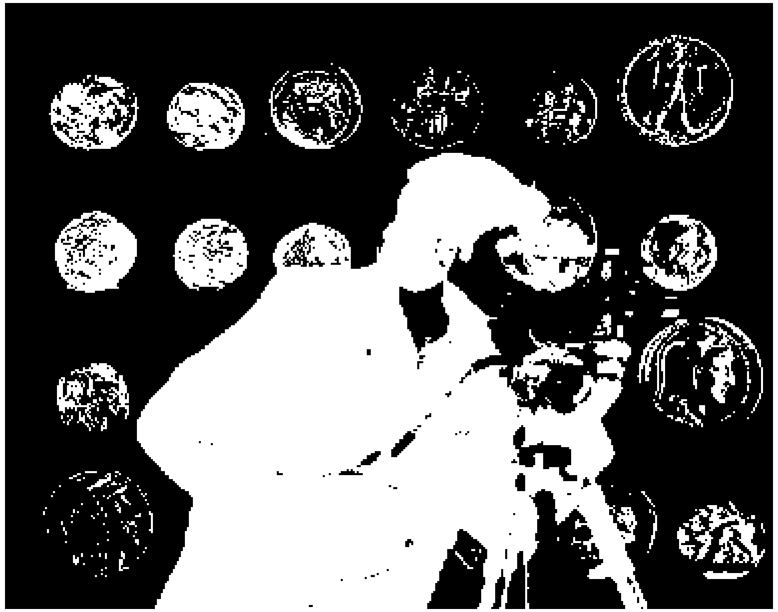
Stochastic Sign Change

- Calculate difference image
  - $D = \{x_i - y_i \mid i = 1, \dots, n\}$
- Count number of sign changes and $x_i = y_i$
- If images are equal, only the noise is visible in D, which changes often.

Stochastic Sign Change



Stochastic Sign Change

Deterministic sign change

- Add «noise» to X «manually» and deterministic:
  - $z_i = x_i + q(-1)^{-1}$
- Calculate difference image:
  - $D = \{z_i - y_i \mid i = 1, \dots, n\}$
- You can use training to optimize q.
- Is supposed to give better results than stockastic sign change [5]

[5] Venot, A., Devaux, J.Y., Herbin, M., Lebruchec, J.F., Dubertret, L., Raulo, Y., Roucayrol, J.C.: An automated system for the registration and comparison of photographic images in medicine. IEEE Trans. Med. Imaging **7**(4), 298–303 (1988)

Minimum ratio

- $r_i = \min\left\{\frac{y_i}{x_i}, \frac{x_i}{y_i}\right\}$
- $m_r = \frac{1}{n}\sum_{i=1}^{n} r_i$
- This is a metric going from 0 to 1
- $m_r = 1 \Rightarrow \text{Identical images}$
- Sensitive to noise. Insensitive to intensity differences.
- Paper proves that this is in fact a metric.

Metrics

- Limited range: $S(X, Y) \leq S_0$ for some S_0 (for example 1).
- Reflexitivity: $S(X, Y) = S_0 \Leftrightarrow X = Y$
- Symmetry: $S(X, Y) = S(Y, X)$
- Triangle Inequality: $S(X, Y)S(Y, Z) \leq [S(X, Y) + S(Y, Z)]S(X, Z)$

Pixel rank

- $R(x_i)$ is the rank of pixel.
- Given all unique pixel values, the rank is the index in a list if the pixel values were sorted.
- Unique pixel values required: Solve ties between pixels by smoothing with a small filter.
  - Gaussian blur with $\sigma \approx 1$

Rank example

Image

3.4	43	45	32
23	231	1	41
125	12	214	124
15	21	51	24

Rank image

2	10	11	9
6	16	1	8
14	3	15	13
4	5	12	7

Spearman's Rho

- Assuming unique ranks.
- $\rho = 1 - \frac{6 \sum_{i=1}^{n} [R(x_i) - R(y_i)]^2}{n(n^2 - 1)}$
- Equivalent of calculating pearson correlation coefficient with unique rankings.

Spearman’s Rho (Some results)

- Compared to pearsons correlation coefficient
  - Better for impulse noise.
  - Slightly better for intensity differences.
  - Because pixels must be sorted, it is computationally heavier than previous methods.

[6] Ayinde, O., Yang, Y.-H.: Face recognition approach based on rank correlation of gaborfiltered images. Pattern Recognit. **35**, 1275–1289 (2002)
[7] Muselet, D., Trémeau, A.: Rank correlation as illumination invariant descriptor for color object recognition. In: Proc. 15th Int’l Conf. Image Processing, pp. 157–160 (2008)

Kendall's Tau

- Concordance: $sign(x_j - x_i) = sign(y_j - y_i), j \neq i$
- Discordance: $sign(x_j - x_i) = -sign(y_j - y_i), j \neq i$
- $N_c -$ Number of pairs of concordances for all valid pairs of i and j.
- $N_d -$ Number of pairs of discordances for all valid pairs of i and j.
- Kendalls's tau: $\tau = \frac{N_c - N_d}{n(n-1)/2}$

Kendall’s Tau (2)

- Very computationally heavy
- Same discriminative power as spearman’s rho.
- More «strict» than Pearson correlation coefficient

Greatest deviation

- Assuming image with unique intensities giving a rank image.
- $d_i = \sum_{j=1}^{i} I\left[R(x_i) \leq i < R(y_j)\right]$
- $D_i = \sum_{j=1}^{i} I[n + 1 \ - R(x_i) > R(y_i)]$
- $R_g = \frac{\max\limits_i(D_i) - \max\limits_i(d_i)}{\frac{n}{2}}$
- R_g varies between -1 and 1.
- Less sensitive to impulse noise than r
- Computationally heavy.

Ordinal measure

- $D_i = \sum_{j=1}^{i} I[n + 1 \ - R(x_i) > R(y_i)]$ (as previously calculated)
- $R_o = \frac{\max_i(D_i)}{\frac{n}{2}}$
- Half the computational expense as Greatest deviation (which is still quite a lot!)

Correlation ratio

- Similar images -> Y is a single valued function of X.
- How much does pixels in Y deviate from the same pixels in X given a pixel intensity in X?
  - $\sigma_i = \sqrt{\frac{1}{n_i}\sum_{x_i}(Y[x_i] - m_i)^2}$
  - $m_i = \frac{1}{n_i}\sum_{x_i}(Y[x_i])$ (or some other mean).
  - We go over all pixels in X with intensity i, and calculate the deviation in Y on the corresponding pixel.
- In an image, we can calculate the mean scatter:
  - $\sigma_m = \frac{1}{256}\sum_{i=0}^{255}\sigma_i$

Correlation ratio

- Calculate the variance of sigmas:
  - $D^2 = \left\{\frac{1}{n}\sum_{i=0}^{255} n_i \sigma_i^2\right\}$
  - $n = \sum_{i=0}^{255} n_i$
- Calculate a similarity measure from variance from D
  - $\eta_{yx} = \sqrt{1 - D^2}$
- Works well for intensity differences in X and Y.
- Relatively expensive operator, but still linear.

Joint probability distribution

- Measure a pixel in X with intensity X_i
- Measure a pixel in Y at the same index with intensity Y_i. (Can also be measured with an offset).
- Increment pixel (X_i, Y_i) in JD image.
- Obtain final JPD by dividing JD by n. (Divide by the number of pixles in the image to get the probability)

Joint probability distribution (Example)

X

$$\begin{matrix} \boxed{1} & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 \\ 6 & 5 & 2 & 1 \\ 7 & 7 & 1 & 1 \end{matrix}$$

Y

$$\begin{matrix} \boxed{2} & 3 & 2 & 4 \\ 1 & 1 & 2 & 5 \\ 6 & 5 & 4 & 3 \\ 5 & 7 & 2 & 1 \end{matrix}$$

JD

Intensities in Y as index

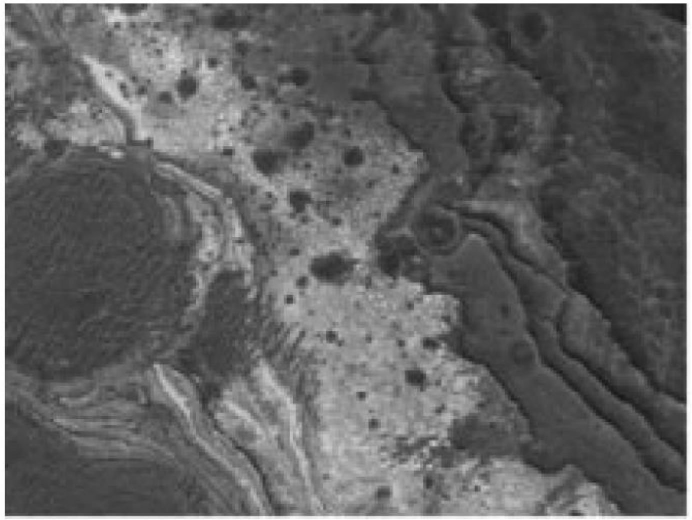
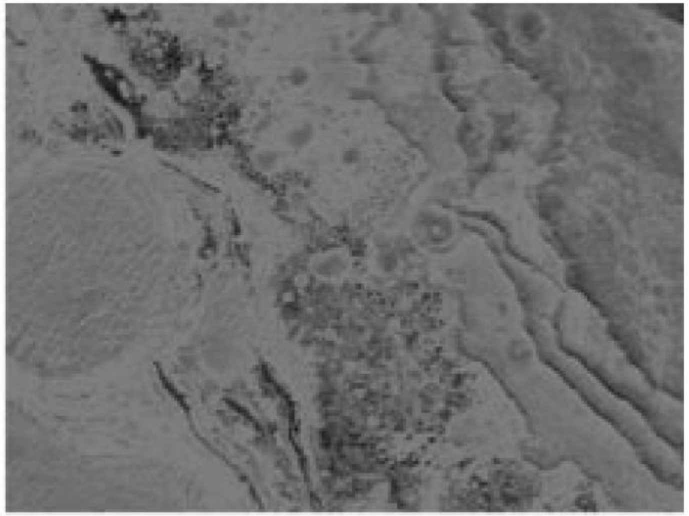
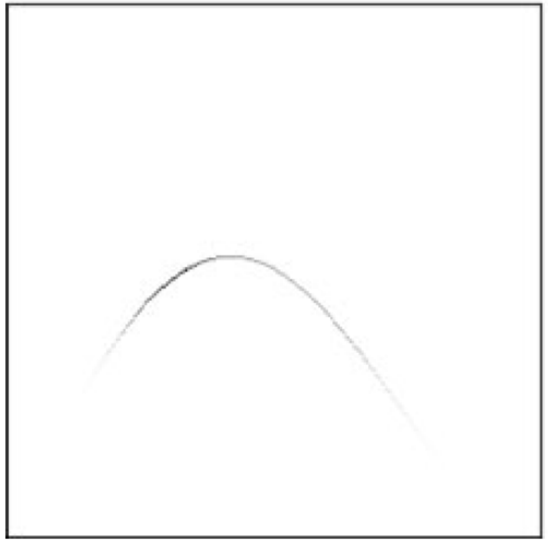
Intensities in X as index

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{0} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

Energy of Joint Probability Distribution

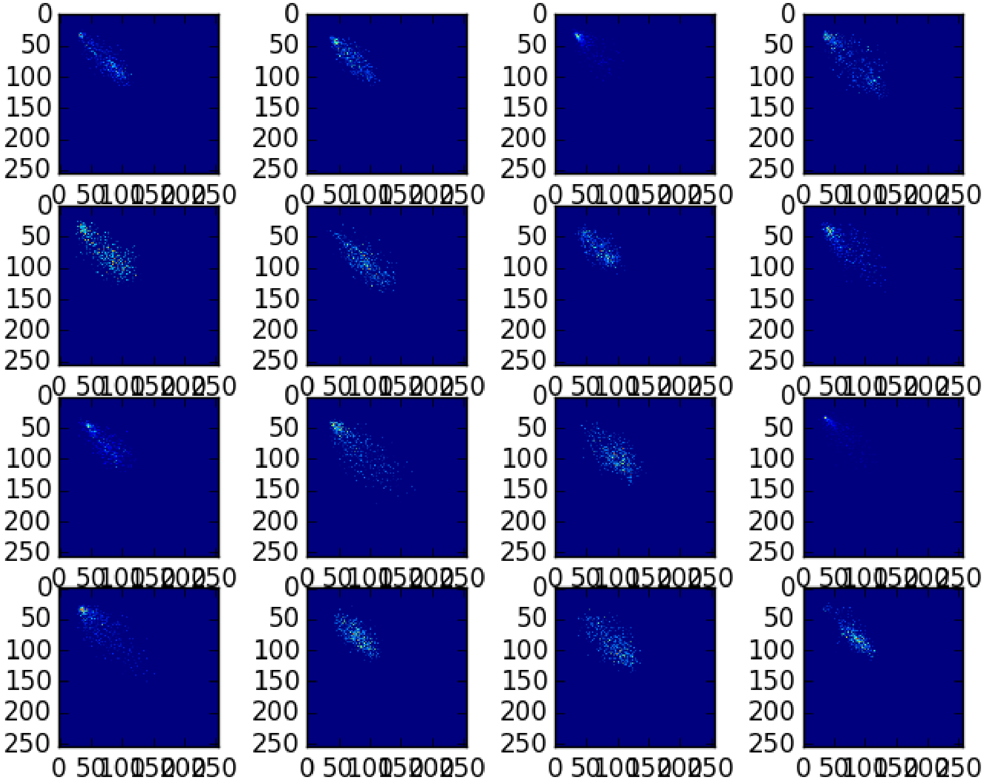
- Calculate JDP of image.
- Calculate energy
  - $E = \sum_{i=0}^{255} \sum_{j=0}^{255} p_{ij}^2$
  - p_{ij} is the pixel in the JPD with index (i, j)
- Can handle intensity transformations and translations.
- A relatively cheap operation.
- Not a metric ☹

Energy of Joint Probability Distribution (Example from paper)

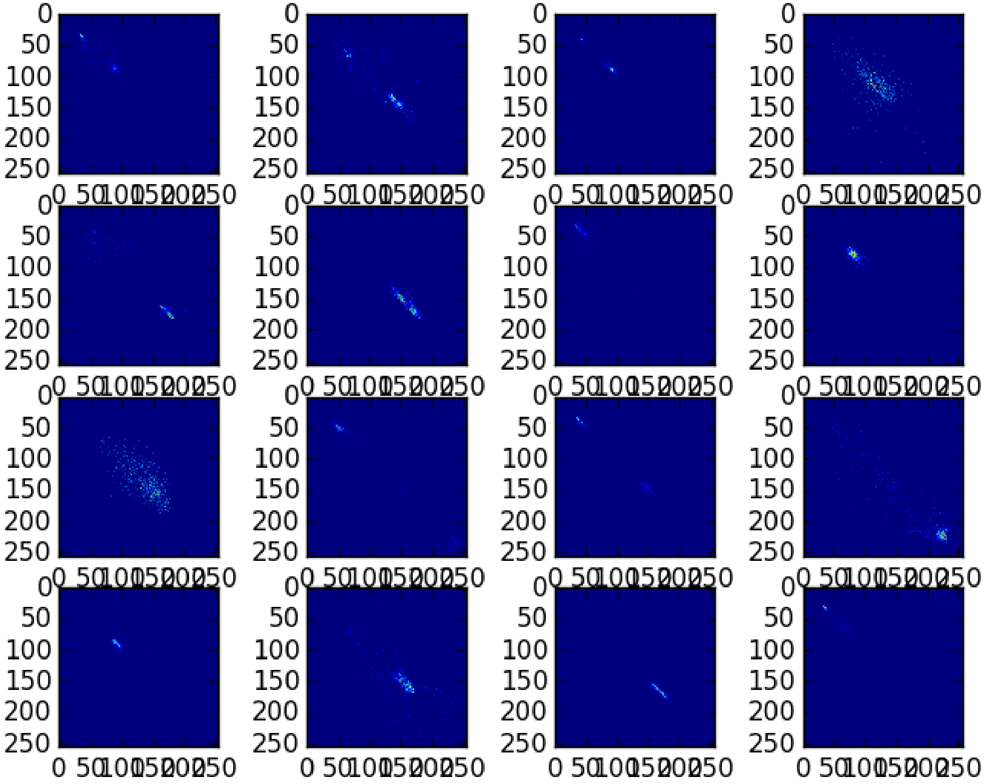
[8] Goshtasby, A.A | **Chapter 2**
Similarity and Dissimilarity Measures F | Book: Image Registration Principles, tools and methods | SPRINGER

Energy of Joint Probability Distribution (My example)



Some image matched against images of trees in airplane photos

Energy of Joint Probability Distribution (My example)



Some image matched against images of not trees in airplane photos

Material similarity

- Problem with Energy of Joint Probability Distribution:
  - Noise gives a «spread out» distribution in the JDP
- Solution:
  - Use material similarity.

Material similarity

- Construct two JPD's.
  - One visiting every k'th pixel in the image, starting from 0.
  - One visiting every k'th pixel in the image, starting from k/2.
  - We are kind of calculating JPD's of subsampled images from different subsamples (a strange subsampling!?)
- Two similar images will have similar JPD's of this kind
- Material similarity: $S_m = \sum_{i=0}^{255} \frac{\min\{p_{ij_1}, q_{ij_2}\}}{|j_1 - j_2| + d}$
- P is JDP constructed starting from 0.
- Q is JDP constructed starting from k/2
- p_{ij_2}, q_{ij_2} pixel at P and Q at positions (i, j_2) in P and (i, j_2) in Q
- j_1, j_2 is the row number in P and Q which contains the peak of index i

Material similarity

X, Y

1	2	4	3
7	6	5	4
2	2	2	3
6	5	4	1

JDP

P

X, Y

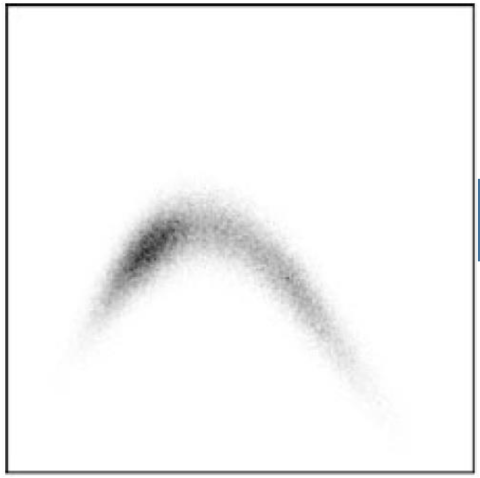
1	2	4	3
7	6	5	4
2	2	2	3
6	5	4	1

JDP

Q

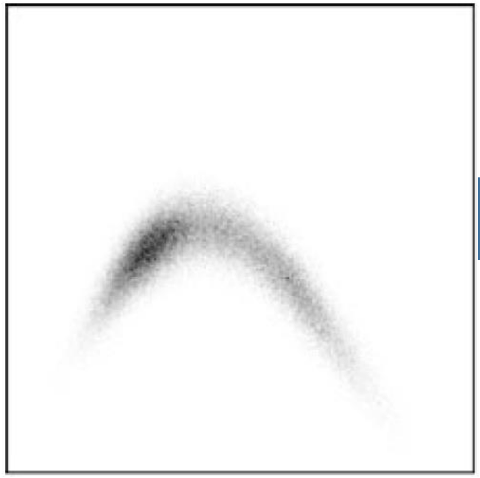
Material similarity

P



Smooth

Q



Smooth

[8] Goshtasby, A.A | **Chapter 2**
Similarity and Dissimilarity Measures F | Book: Image Registration Principles, tools and methods | SPRINGER

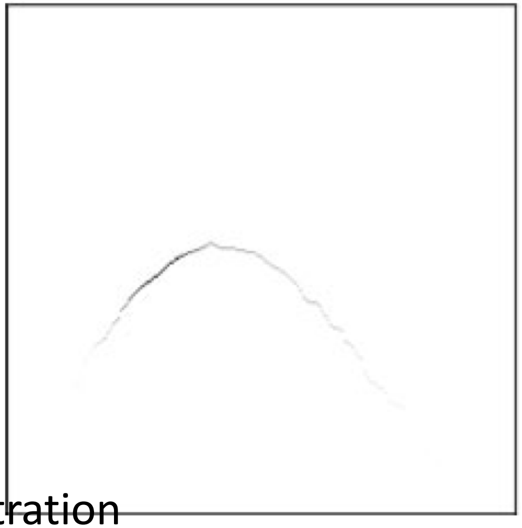
Material similarity

P

Peaks

Q

Peaks

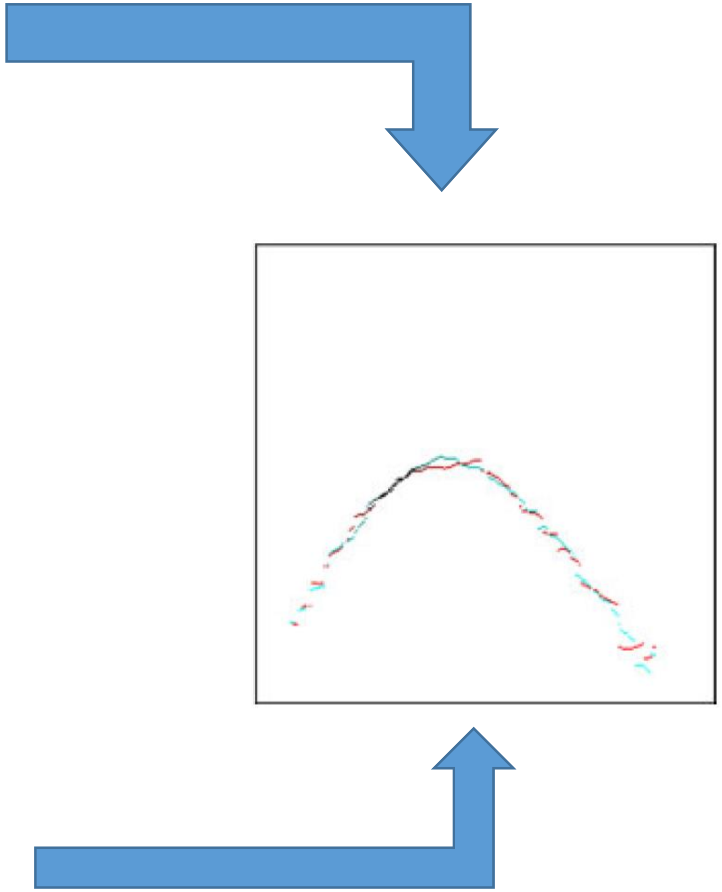


[8] Goshtasby, A.A | **Chapter 2**
Similarity and Dissimilarity Measures F | Book: Image Registration Principles, tools and methods | SPRINGER

Material similarity

P

Q



Note: There is no joining in the actual algorithm. This last slide is solely for demonstration purposes, to show how similar images might look in this space.

[8] Goshtasby, A.A | **Chapter 2**
Similarity and Dissimilarity Measures F | Book: Image Registration Principles, tools and methods | SPRINGER

Material similarity (Summary)

- Better for noise than «Energy of Joint Probability Distribution»
- Relatively cheap computation

Shannon mutual information

- History: Register multimodal images.
- Mutial information: Measure the disparity of the JPD of X and Y.
- $S_{MI} = \sum_{i=0}^{255} \sum_{j=0}^{255} p_{ij} \log_2 \frac{p_{ij}}{p_i p_j} = E_i + E_j - E_{ij}$
  - $E_i = -\sum_{j=0}^{255} p_j \log_2 p_j$
  - $E_j = -\sum_{i=0}^{255} p_i \log_2 p_i$
  - $E_{ij} = -\sum_{i=0}^{255} p_{ij} \log_2 p_{ij}$
  - $p_i = \sum_{j=0}^{255} p_{ij}$ Probability of intensity i in image X
  - $p_j = \sum_{i=0}^{255} p_{ij}$ Probability of intensity j in image Y

Shannon mutual information

- Using entropy measures. (disparity)
- Sensitive to noise
- Relatively low computational cost
- A very popular similarity measure
- Lots of variations exist
  - High-order mutual information: Calculate JDP with offsets in j and i.
  - Variations using gradients, point coordinates, phase etc.
  - Used for multiresolution, monomodal, dynamic images, etc.

Rényi mutual information

- Another measure of entropy: Rényi entropy
  - $E_\alpha = \frac{1}{1-\alpha}\log_2\left(\sum_{i=0}^{255} p_i^\alpha\right)$, This is Rényi entropy over probability distribution P
  - $\log_2(p_{max}) \leq E_\alpha \leq \log_2(256), p_{max} = \max_{p_i \epsilon P}(p_i)$
- Rényi mutial information:
  - $R_\alpha = \frac{E_\alpha^i + E_\alpha^j}{E_\alpha^{ij}}$
  - E_α^i is the Rényi entropy of the distribution $p_i = \sum_{j=0}^{255} p_{ij}$ for $i = 1, \dots, 255$
  - E_α^j is the Rényi entropy of the distribution $p_i = \sum_{i=0}^{255} p_{ij}$ for $\text{j} = 0, \dots, 255$
  - E_α^{ij} is the Rényi entropy of the distribution $\{p_{ij}: i, j = 0, \dots, 255\}$

Rényi mutual information (2)

- Choise of α magnifies higher values in JDP which reduces the effect of outliers
  - It’s therefore better for impulse noise
- Slightly heavier to compute than “Shannon mutual information”

Tsallis mutual information

- Yet another entropy mutual information!
- Tsallis entropy:
  - $S_q = \frac{1}{q-1}\left(1 - \sum_{i=0}^{255} \sum_{j=0}^{255} p_{ij}^q\right)$
- Tsallis mutual information:
  - $R_q = S_q^i + S_q^j + (1-q)S_q^i S_q^j - S_q$
  - $S_q^i = \frac{1}{q-1}\sum_{j=0}^{255} p_{ij}\left(1 - p_{ij}^{q-1}\right)$
  - $S_q^j = \frac{1}{q-1}\sum_{i=0}^{255} p_{ij}\left(1 - p_{ij}^{q-1}\right)$

Tsallis mutual information

- $q > 1 \ \rightarrow$ Outliers are less important. No logarithmic function.
- Less sensitive to noise than Rényi
- Computation is relatively cheap (the same as Rényi)
- Value of q depends on the application.

F-Information measure (I_α)

- $I_\alpha = \frac{1}{\alpha(\alpha-1)}\left(\sum_{i=0}^{255}\sum_{j=0}^{255}\frac{p_{ij}^{\alpha}}{\left(p_i p_j\right)^{\alpha-1}} - 1\right), \alpha \neq 0, \alpha \neq 1$
- Approaches Shannon information as α approaches 1

F-Information measure (M_α)

- $M_\alpha = \sum_{i=0}^{255} \sum_{j=0}^{255} \left| p_{ij}^{\alpha} - \left(p_i p_j\right)^{\alpha} \right|^{\frac{1}{\alpha}}, 0 < \alpha \leq 1$

F-Information measure (χ^{α})

- $\chi^{\alpha} = \sum_{i=0}^{255} \sum_{j=0}^{255} \frac{\left|p_{ij} - p_i p_j\right|^{\alpha}}{\left(p_i p_j\right)^{\alpha-1}}, \alpha > 1$

F-Information measure some notes

- Relatively cheap computations
- Choose an α that best matches the application

Operator summary

- r - Pearson correlation coefficient
- S_T - Tanimoto Measure
- D_s - Stochastic Sign Change
- D_d – Deterministic Sign Change
- m_r – Minimum ratio
- ρ – Spearman’s Rho
- τ – Kendall’s Tau
- R_g- Greates deviation
- R_o - Ordinal measure
- η_{yx} - Correlation Ratio
- E – Energy of joint probability distribution
- S_m - Material similarity
- S_{MI} - Shannon mutual information
- R_α - Rényi mutual information
- R_q - Tsallis mutual information
- I_α , M_α , χ^α - F-information

Computation summary

- r – Cheap, $O(n)$
- S_T - Cheap, $O(n)$
- D_s – Cheap, $O(n)$
- D_d – Cheap, $O(n)$
- m_r – Cheap, $O(n)$
- ρ – Quite cheap, $O(n \log_2 n)$
- τ – Expensive, $O(n^2)$
- R_g - Expensive, $\mathrm{n}\log_2\mathrm{n} + O(n^2)$
- R_o - Expensive, $\mathrm{n}\log_2\mathrm{n} + O(n^2)$
- η_{yx} - Quite cheap, $256 \cdot O(n)$
- E – Cheap, $256^2 + O(n)$
- S_m - Cheap, $256^2 + O(n)$
- S_{MI} - Cheap, $256^2 + O(n)$
- R_α - Cheap, $256^2 + O(n)$
- R_q - Cheap, $256^2 + O(n)$
- I_α , M_α , χ^α - Cheap, $256^2 + O(n)$

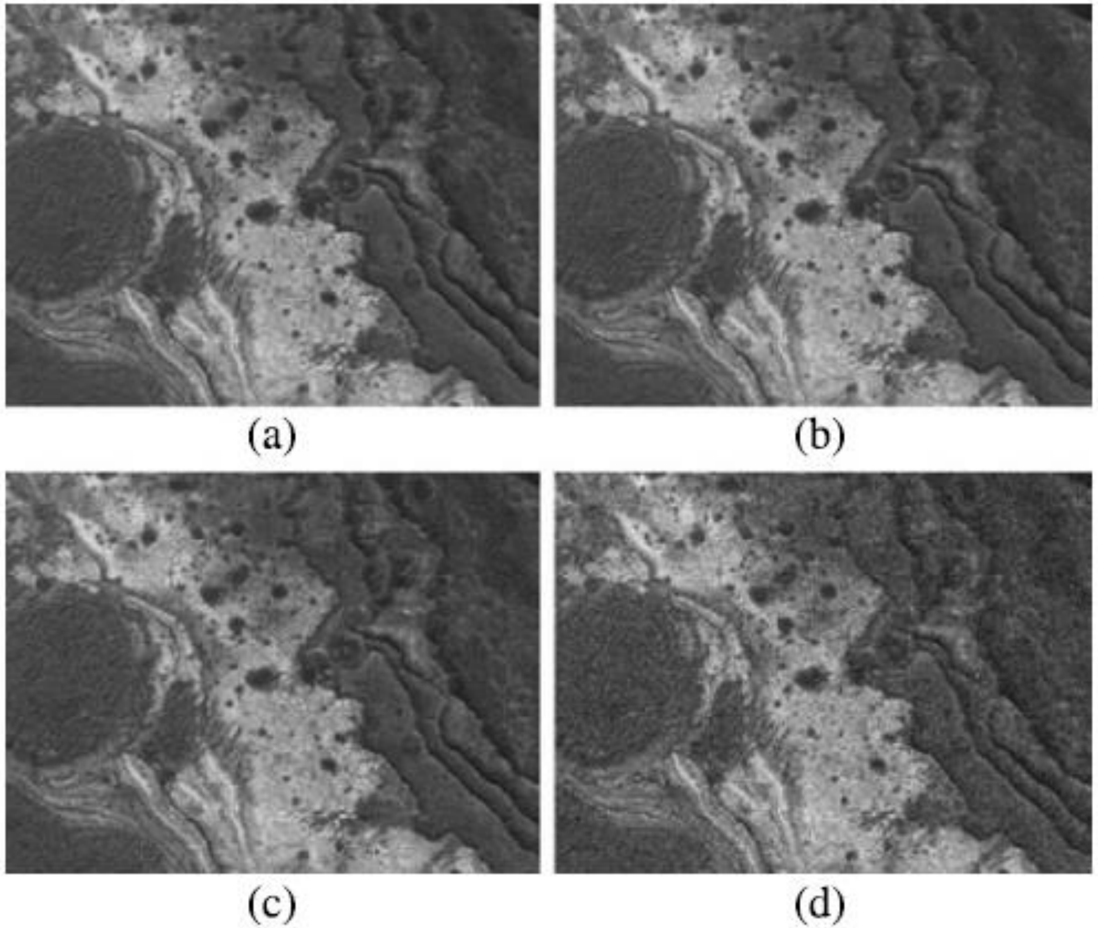
How the different measures relate

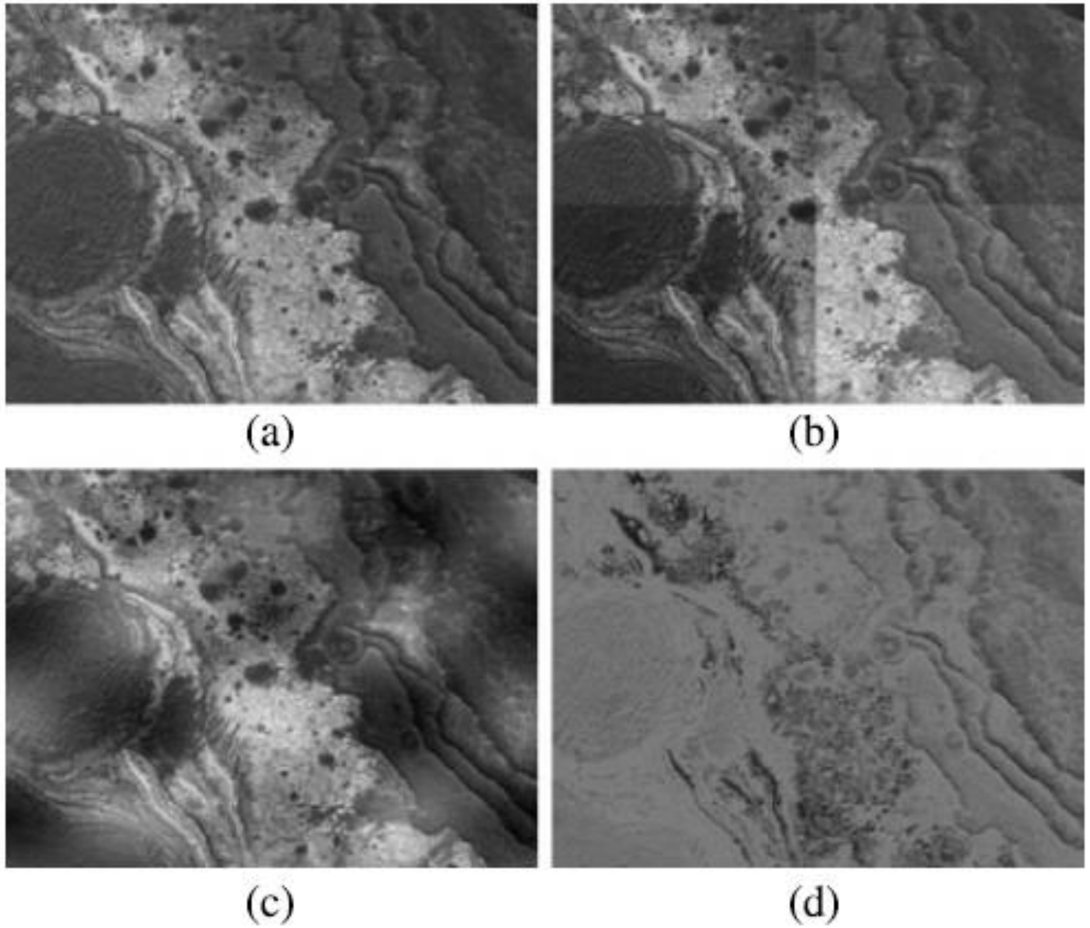
- r and S_T gives the same results. Standard deviation instead of squared euclidian distance and normalizing by their means instead of the inner product.
- $\rho =$ calculate r over Ranks of X and Y
- $bivariate(X, Y)$ is normally distributed $\rightarrow r = \sin\left(\frac{\pi\tau}{2}\right)$
- X and Y are independent $\rightarrow$ $\frac{\rho}{\tau} \rightarrow \frac{3}{2}\ as\ n \rightarrow \infty$. They have the same discrimination power. However, both ρ and τ vary between -1 and 1.
- Linear intensity transformation from X to Y $\rightarrow (\eta^2 - r^2) = 0$

How the different measures relate (2)

- $E_\alpha \rightarrow S_{MI}\ when\ \alpha \rightarrow 1$
- $S_q \rightarrow S_{MI}\ when\ q\ \ \rightarrow 1$
- $I_\alpha \rightarrow S_{MI}\ when\ \alpha \rightarrow 1$

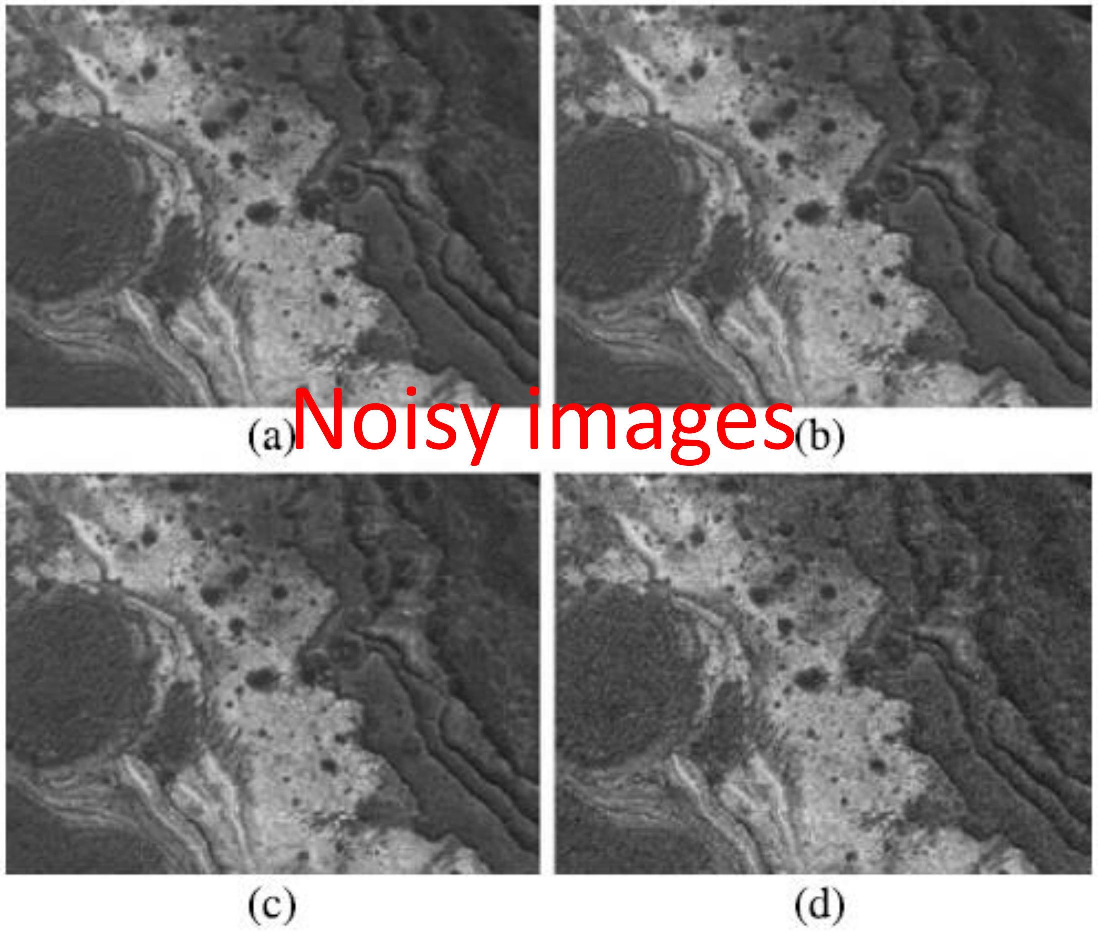
Experimental setup

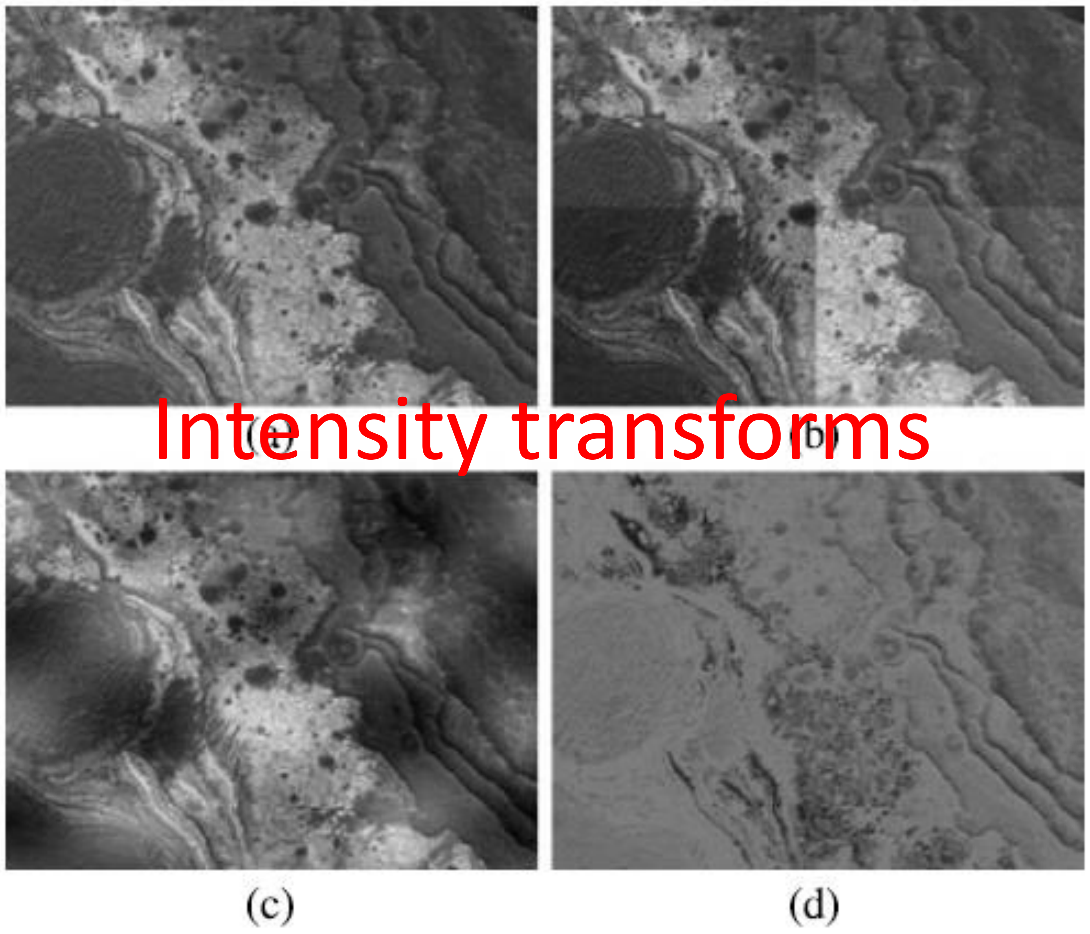




[8] Goshtasby, A.A | **Chapter 2**
Similarity and Dissimilarity Measures F | Book: Image Registration Principles, tools and methods | SPRINGER

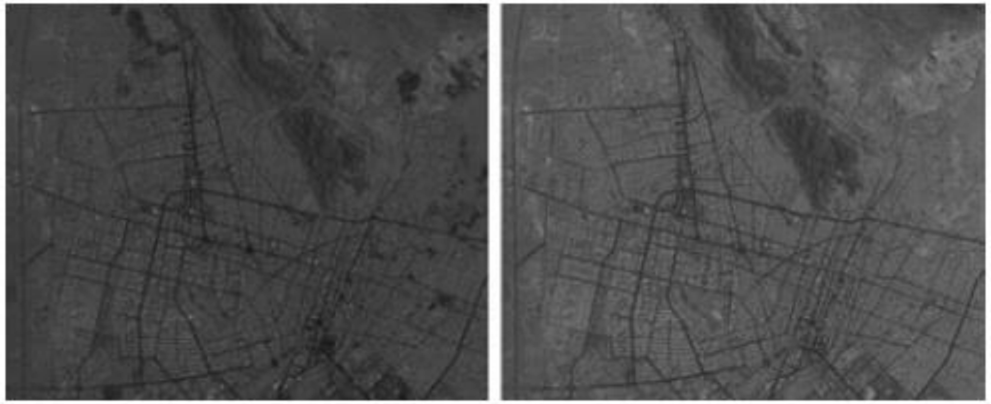
Experimental setup

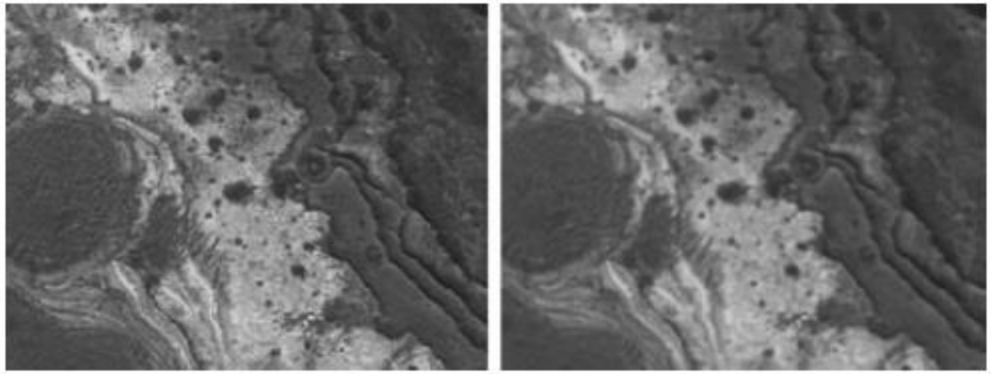




[8] Goshtasby, A.A | **Chapter 2**
Similarity and Dissimilarity Measures F | Book: Image Registration Principles, tools and methods | SPRINGER

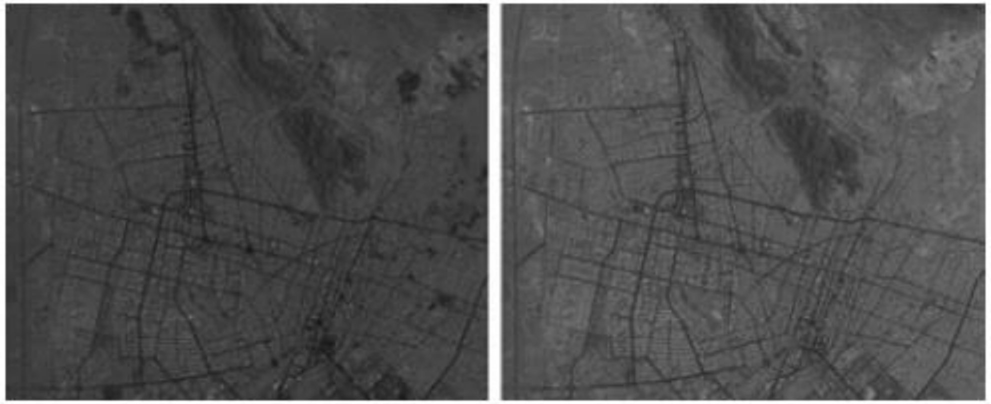
Experimental setup (2)



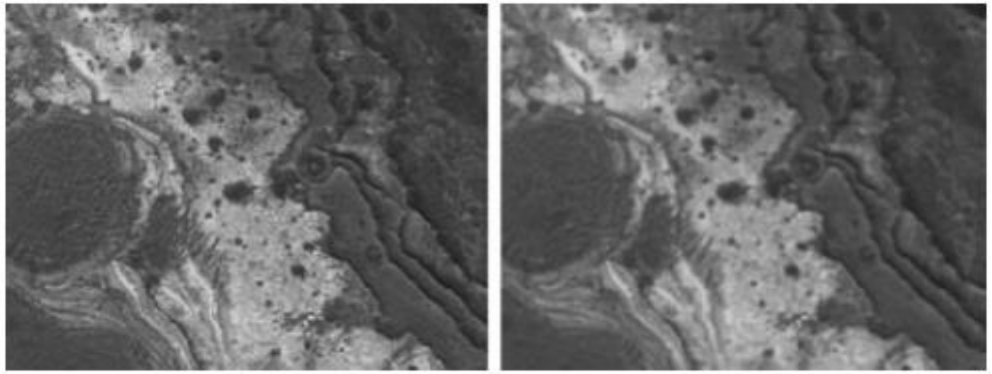


[8] Goshtasby, A.A | **Chapter 2**
Similarity and Dissimilarity Measures F | Book: Image Registration
Principles, tools and methods | SPRINGER

Experimental setup (2)



Multimodal images

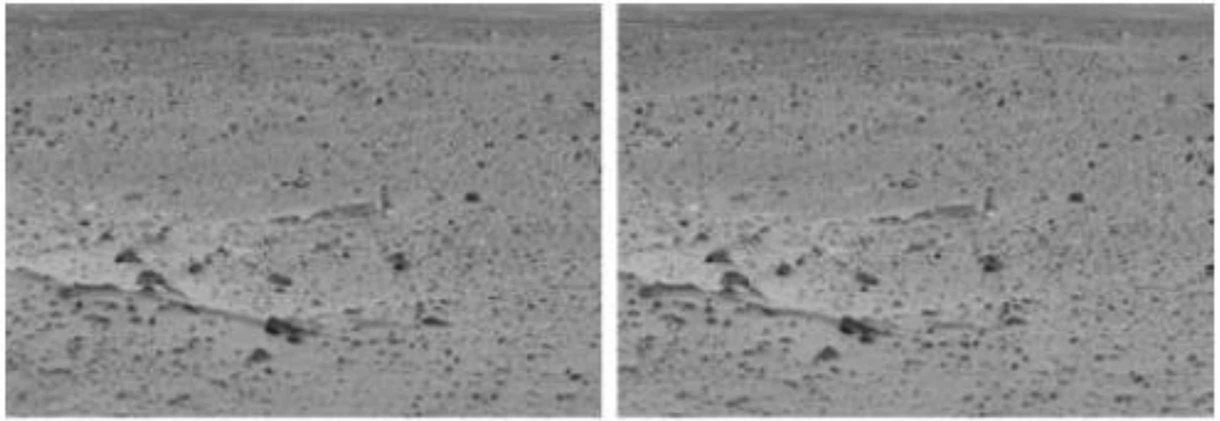


Blurred image

Different exposure

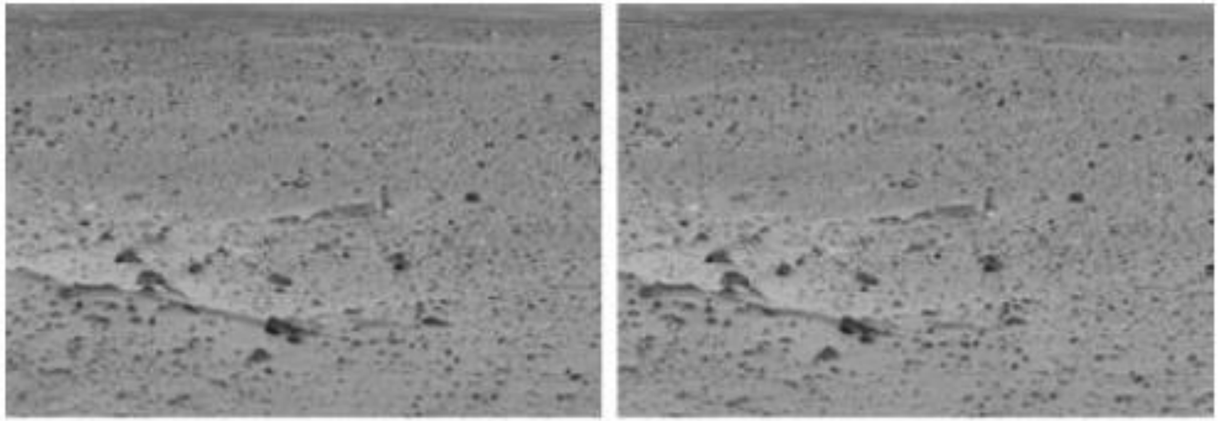
[8] Goshtasby, A.A | **Chapter 2**
Similarity and Dissimilarity Measures F | Book: Image Registration
Principles, tools and methods | SPRINGER

Experimental setup (3)



[8] Goshtasby, A.A | **Chapter 2**
Similarity and Dissimilarity Measures F | Book: Image Registration
Principles, tools and methods | SPRINGER

Experimental setup (3)



Stereo images

Stereo images

[8] Goshtasby, A.A | **Chapter 2**
Similarity and Dissimilarity Measures F | Book: Image Registration
Principles, tools and methods | SPRINGER

Experimental setup (4)

- All possible windows of size 31 in first image is searched for in the other image.
- For stereo images only a horizontal search is perfomed.
- On certain operators, weighting is used. Gaussian mask. JPDs are also weighted based on distance to center.
- RMSID used for stereo images.
- Result: How many pixels are correctly matched between images.

Results (accuracy)

Method	Noise	More Noise	Most Noise	Lin trans	Sin trans
Pearson correlation	100.00	100.00	99.92	100.00	100.00
Tanimoto measure	100.00	100.00	99.95	100.00	100.00
Stochastic sign change	83.51	58.43	43.24	0.00	0.70
Deterministic sign change	98.50	99.05	85.81	48.20	49.45
Minimum ratio	100.00	100.00	99.61	42.29	50.41
Spearman's Rho	100.00	100.00	**99.96**	99.97	100.00
Kendall's Tau	100.00	100.00	**100.00**	100.00	100.00
Greatest deviation	99.92	99.36	91.18	97.17	94.01
Ordinal measure	99.98	99.25	90.35	94.66	87.75
Correlation ratio	100.00	100.00	99.90	100.00	99.49
Energy of JPD	100.00	82.13	16.91	100.00	87.59
Material similarity	100.00	97.82	56.06	100.00	73.11
Shannon MI	93.50	50.91	5.59	100.00	61.82
Rényi MI	98.11	54.12	5.93	100.00	73.66
Tsallis MI	100.00	83.61	17.46	100.00	90.16
I_α-information	99.85	98.06	77.72	100.00	98.92

Results (accuracy)

Method	Noise	More Noise	Most Noise	Lin trans	Sin trans
Pearson correlation	100.00	100.00	99.92	100.00	100.00
Tanimoto measure	100.00	100.00	99.95	100.00	100.00
Stochastic sign change	83.51	58.43	43.24	0.00	0.70 !
Deterministic sign change	98.50	99.05	85.81	48.20	49.45
Minimum ratio	100.00	100.00	99.61	42.29	50.41
Spearman's Rho	100.00	100.00	**99.96**	99.97	100.00
Kendall's Tau	100.00	100.00	**100.00**	100.00	100.00
Greatest deviation	99.92	99.36	91.18	97.17	94.01
Ordinal measure	99.98	99.25	90.35	94.66	87.75
Correlation ratio	100.00	100.00	99.90	100.00	99.49
Energy of JPD	100.00	82.13	16.91	100.00	87.59
Material similarity	100.00	97.82	56.06	100.00	73.11
Shannon MI	93.50	50.91	5.59	100.00	61.82
Rényi MI	98.11	54.12	5.93	100.00	73.66
Tsallis MI	100.00	83.61	17.46	100.00	90.16
I_α-information	99.85	98.06	77.72	100.00	98.92

Results (2) (accuracy)

Method	Sin trans	Mult. Mod.	Diff. Expos.	Blur	Stereo	Stereo
Pearson correlation	52.78	96.87	**98.96**	100.00	**8.44**	9.81
Tanimoto measure	52.55	96.88	95.16	100.00	**8.43**	**9.80**
Stochastic sign change	13.06	0.27	9.61	93.17	10.30	11.83
Deterministic sign change	2.25	0.00	12.33	88.24	9.00	10.01
Minimum ratio	100.0	0.10	2.81	100.00	8.60	**9.77**
Spearman's Rho	56.19	97.28	97.53	99.97	8.66	9.98
Kendall's Tau	59.44	98.64	98.23	100.00	9.04	10.08
Greatest deviation	45.39	96.16	89.15	93.62	11.66	10.92
Ordinal measure	44.05	95.24	88.71	96.07	11.31	10.91
Correlation ratio	100.00	98.27	**99.78**	100.00	10.81	10.70
Energy of JPD	100.00	98.21	79.25	85.51	12.08	11.23
Material similarity	100.00	**100.00**	**98.73**	93.84	16.52	15.46
Shannon MI	100.00	98.36	83.59	61.61	20.33	14.12
Rényi MI	100.00	98.30	79.57	67.84	17.75	12.99
Tsallis MI	100.00	98.30	84.31	89.06	10.87	10.86
I_α-information	100.00	97.59	91.18	86.71	11.14	11.59

Results (2) (accuracy)

Method	Sin trans	Mult. Mod.	Diff. Expos.	Blur	Stereo	Stereo
Pearson correlation	52.78	96.87	**98.96**	100.00	**8.44**	9.81
Tanimoto measure	52.55	96.88	95.16	100.00	**8.43**	**9.80**
Stochastic sign change	13.06	0.27	9.61	93.17	10.30	11.83
Deterministic sign change	2.25	0.00	12.33	88.24	9.00	10.01
Minimum ratio	100.0	0.10	2.81	100.00	8.60	**9.77**
Spearman's Rho	56.19	97.28	97.53	99.97	8.66	9.98
Kendall's Tau	59.44	98.64	98.23	100.00	9.04	10.08
Greatest deviation	45.39	96.16	89.15	93.62	11.66	10.92
Ordinal measure	44.05	95.24	88.71	96.07	11.31	10.91
Correlation ratio	100.00	98.27	**99.78**	100.00	10.81	10.70
Energy of JPD	100.00	98.21	79.25	85.51	12.08	11.23
Material similarity	100.00	**100.00**	**98.73**	93.84	16.52	15.46
Shannon MI	100.00	98.36	83.59	61.61	20.33	14.12
Rényi MI	100.00	98.30	79.57	67.84	17.75	12.99
Tsallis MI	100.00	98.30	84.31	89.06	10.87	10.86
I_α-information	100.00	97.59	91.18	86.71	11.14	11.59

Results (3) (Larger templates)

- Larger templates gives the following:
  - JPD based methods become slightly better (as do every measure, but most for JPD)
  - Ordinal measure also become slightly better (Spearman’s rho)
  - Sign-change, minimum ratio, correlation ratio, incremental sign distance and rank distance least affected by the increase.

Results (4) (Performance)

Small template	
Pearson correlation	5.09
Tanimoto measure	5.09
Stochastic sign change	4.29
Deterministic sign change	28.91
Minimum ratio	4.10
Spearman's Rho	107.94
Kendall's Tau	627.81
Greatest deviation	722.73
Ordinal measure	435.64
Correlation ratio	84.52
Energy of JPD	110.25
Material similarity	241.43
Shannon MI	172.97
Rényi MI	220.40
Tsallis MI	226.82
I_α-information	456.45

Large template	
Pearson correlation	12.13
Tanimoto measure	12.15
Stochastic sign change	10.78
Deterministic sign change	36.63
Minimum ratio	9.96
Spearman's Rho	320.36
Kendall's Tau	4547.6
Greatest deviation	4873.3
Ordinal measure	2718.3
Correlation ratio	224.92
Energy of JPD	118.89
Material similarity	273.27
Shannon MI	192.60
Rényi MI	232.93
Tsallis MI	230.17
I_α-information	534.25

Note: These are a lot faster on stereo images

Some notes on preprocessing

- Normalize pixels before running JPD
- Smooth images to reduce noise (or median filter etc.)
- Adaptive smoothing can be used.
- Rank transform.
  - Replace the center pixel with the number of pixels in the neighbourhood that is smaller than the center pixel.
  - Might improve accuracy on intensity transformed images.

Rank transform

0	0	1	3	2	4	5	6
1	7	5	4	5	1	2	5
4	1	2	5	5	2	1	7
5	4	8	4	6	5	1	2
0	9	4	5	1	2	6	5
4	0	4	8	2	0	5	0
q	7	9	1	0	8	1	5
7	1	8	4	7	8	5	1

0	0	1	3	2	4	5	6
1	7	5	4	5	1	2	5
4	1	2	5	5	2	1	7
5	4	8	4	6	5	1	2
0	9	4	5	1	2	6	5
4	0	4	8	2	0	5	0
q	7	9	1	0	8	1	5
7	1	8	4	7	8	5	1

…

…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
…	…	1	3	…	…	…	…
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…

How to read the paper (suggestion)

- Look at images in experimentation section an look at what best fits your application (Section 2.3.1 Experimental Setup).
- Look over the accuracy tables and performance tables on the experiment image of choice. Select those that gives acceptable results (section 2.4 Characteristics of Similarity/Dissimilarity Measures)
- If not clear, read section 2.5 (choosing a similarity/dissimilarity measure)
- Read the algorithm's section for implementation.
- Read the remaining part of what is required.

Conclusion

- There are many similarity/dissimilarity measures!
- Some are generally better than other, but it really depends on the application
  - Real time?
  - Dataset size?
  - Noise?
  - Intensity differences?
  - Modality?
  - Shifted data?
  - Performance?
  - Code size?
  - Is a metric required?