Relatively-Paired Space Analysis

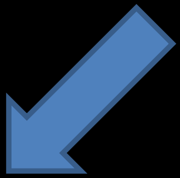
Zhanghui Kuang Kwan-Yee K. Wong
Department of Computer Science
The University of Hong Kong

Multi-modality representation

Modality photo

Multi-modality representation

Modality photo



Modality sketch

Multi-modality representation

Modality photo

Modality sketch

Modality fingerprint

Multi-modality representation

Modality photo

Modality sketch

Modality fingerprint

Modality sound wave

Cross-modality pattern recognition

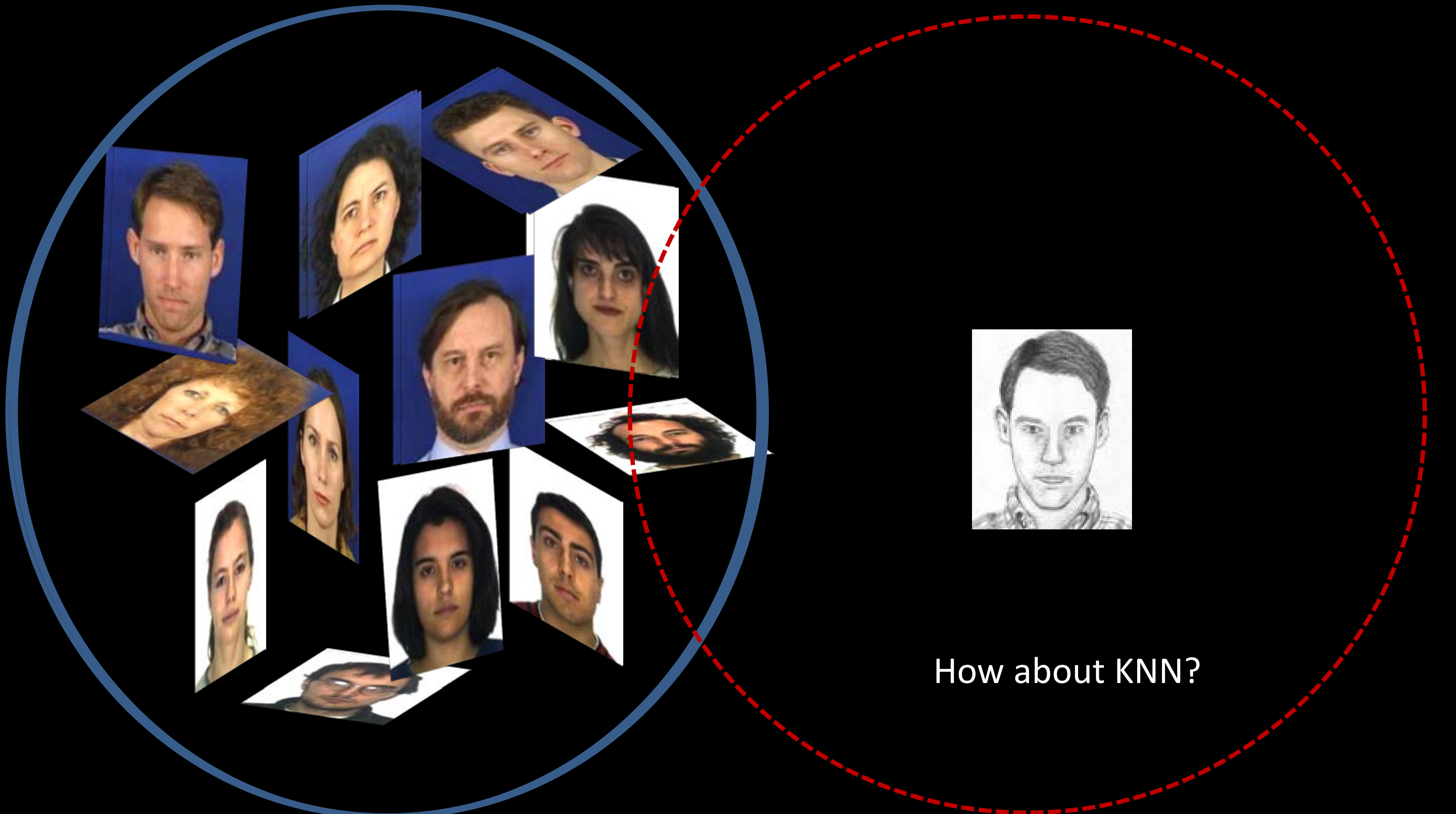
Gallery dataset with labels

Cross-modality pattern recognition

Gallery dataset with labels

Sketch query

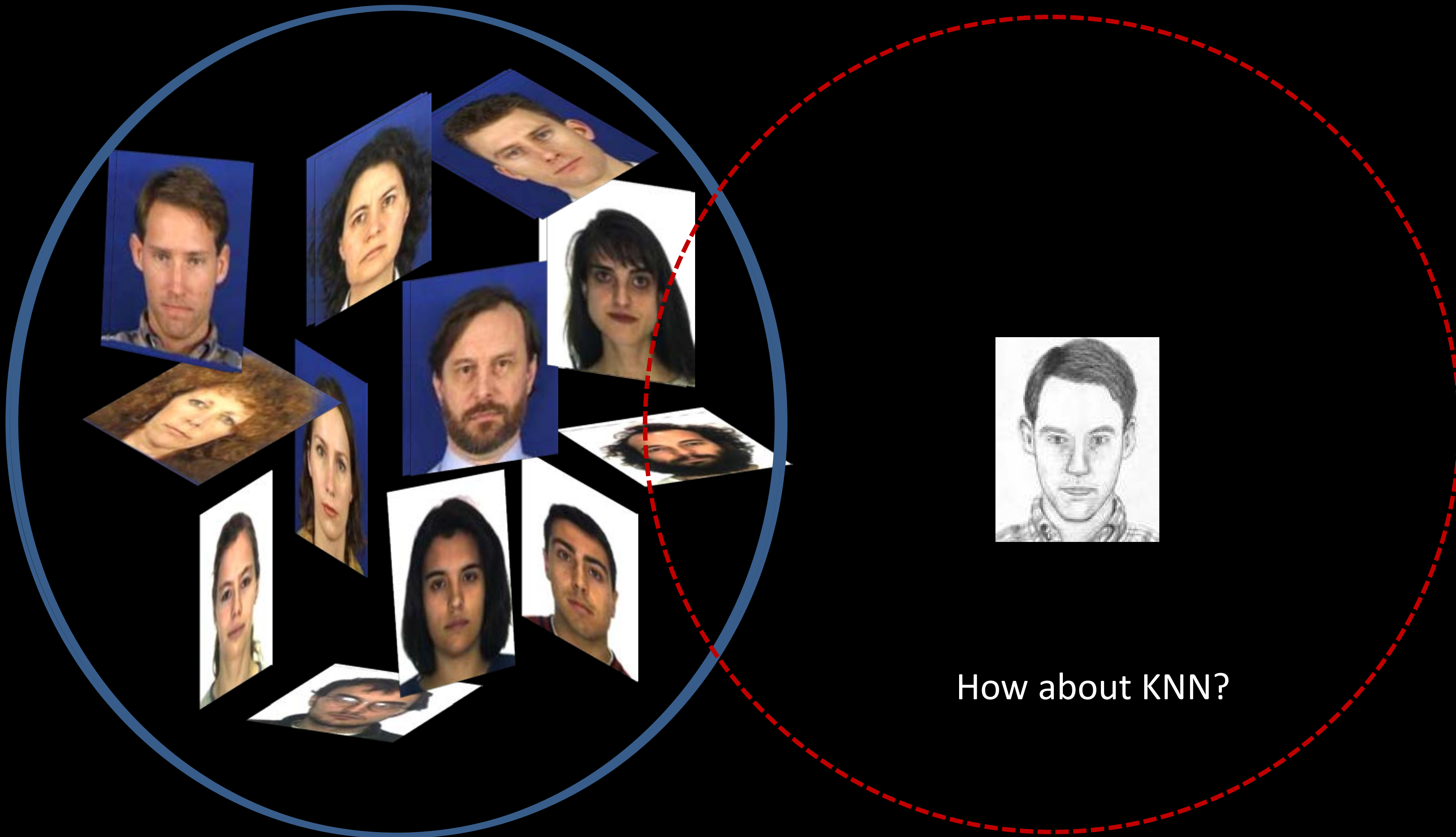
Cross-modality pattern recognition



Gallery dataset with labels

Sketch query

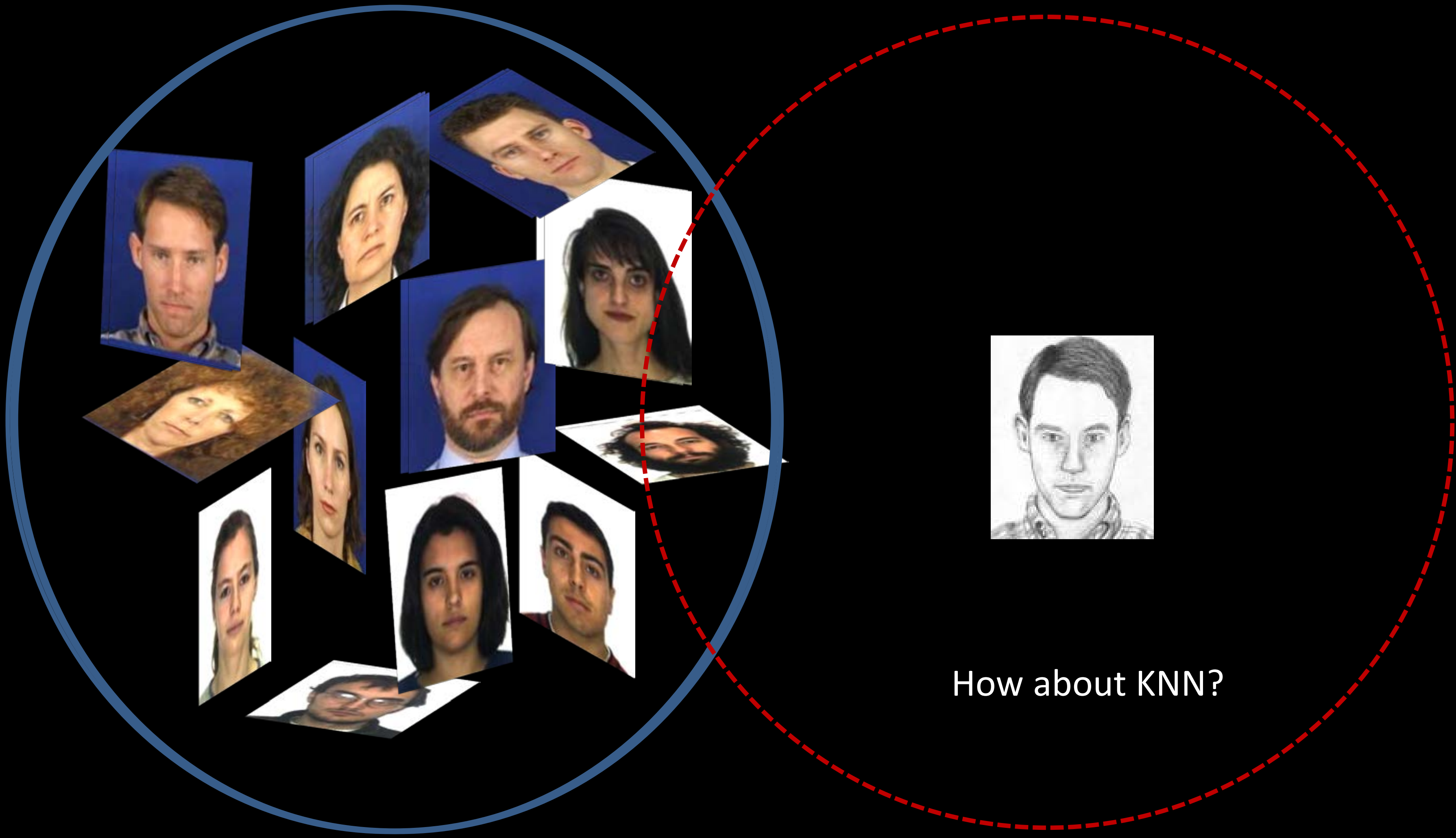
Cross-modality pattern recognition



Gallery dataset with labels

Sketch query

Cross-modality pattern recognition



Gallery dataset with labels

Sketch query

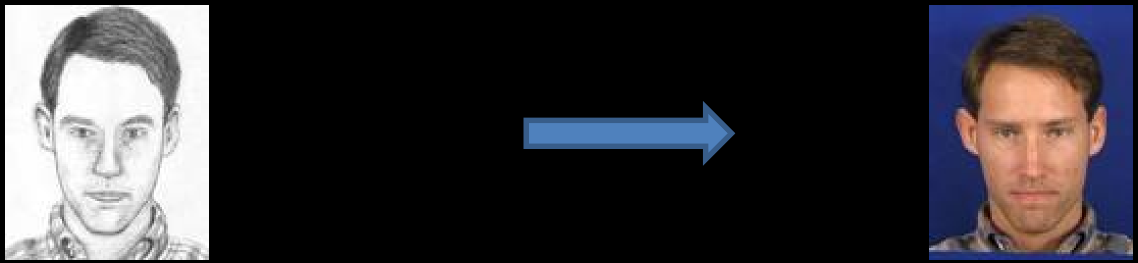
Related Approaches

- Transform one modality into another

[Wang and Tang PAMI 2010, Zhou *et al.* CVPR 2012， Blanz et al. CVPR 2005]

Related Approaches

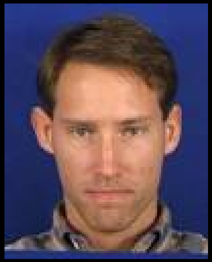
- Transform one modality into another



[Wang and Tang PAMI 2010, Zhou *et al.* CVPR 2012， Blanz et al. CVPR 2005]

Related Approaches

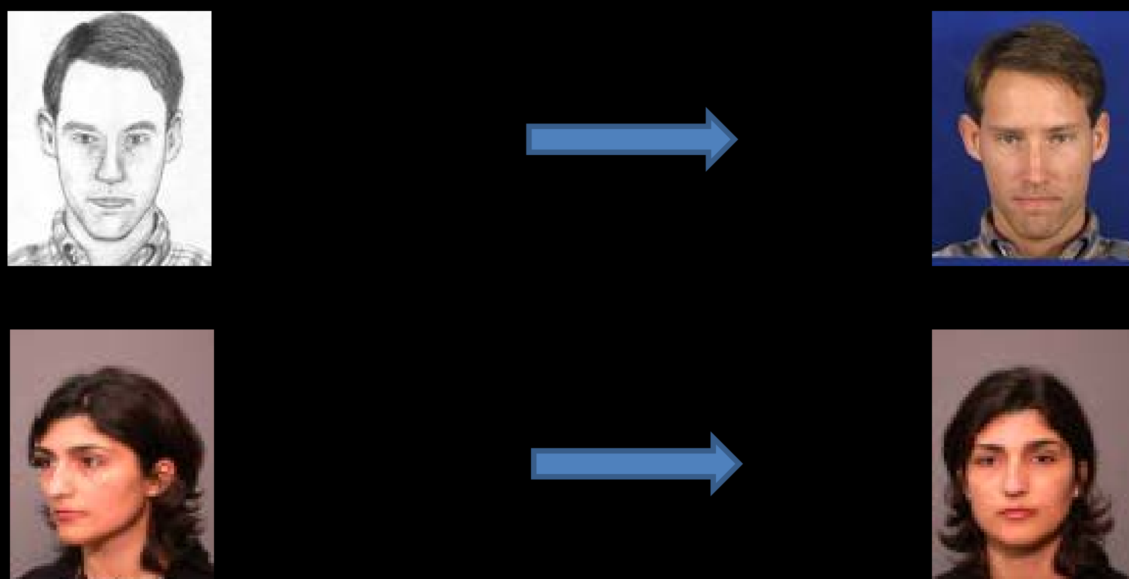
- Transform one modality into another



[Wang and Tang PAMI 2010, Zhou *et al.* CVPR 2012， Blanz et al. CVPR 2005]

Related Approaches

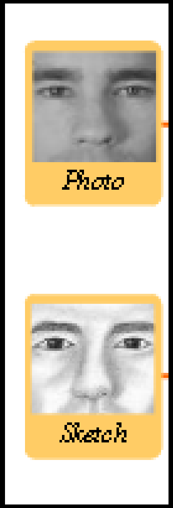
- Transform one modality into another



[Wang and Tang PAMI 2010, Zhou *et al.* CVPR 2012， Blanz et al. CVPR 2005]

Related Approaches

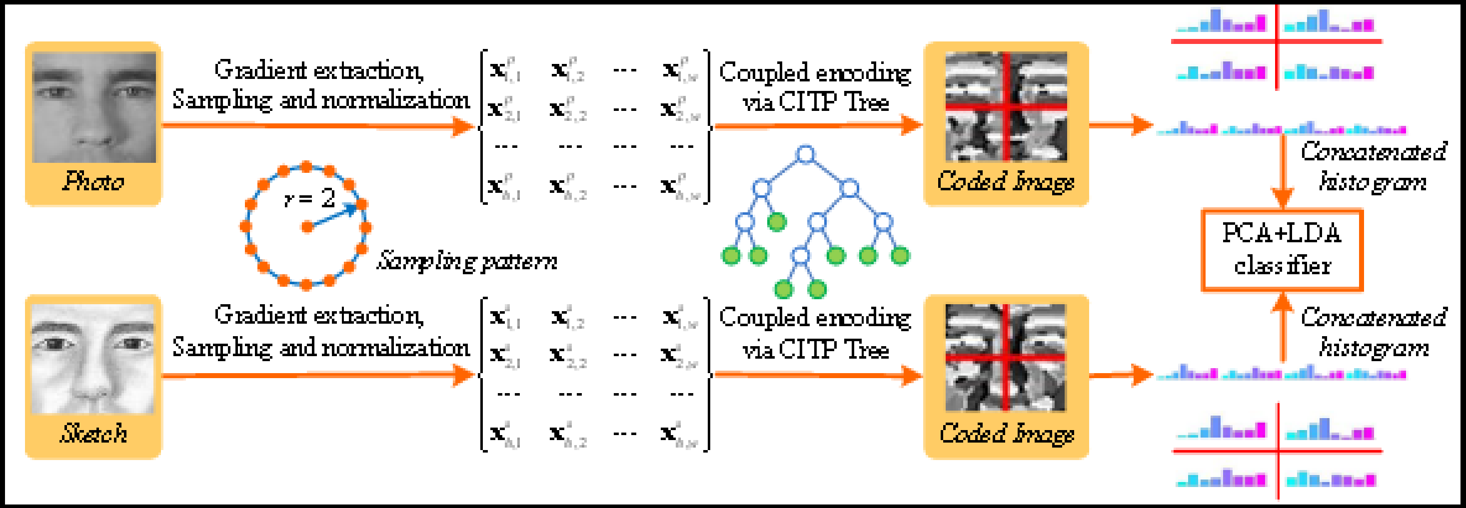
- Extract modality-invariant features



[Zhang *et al.* CVPR 2011]

Related Approaches

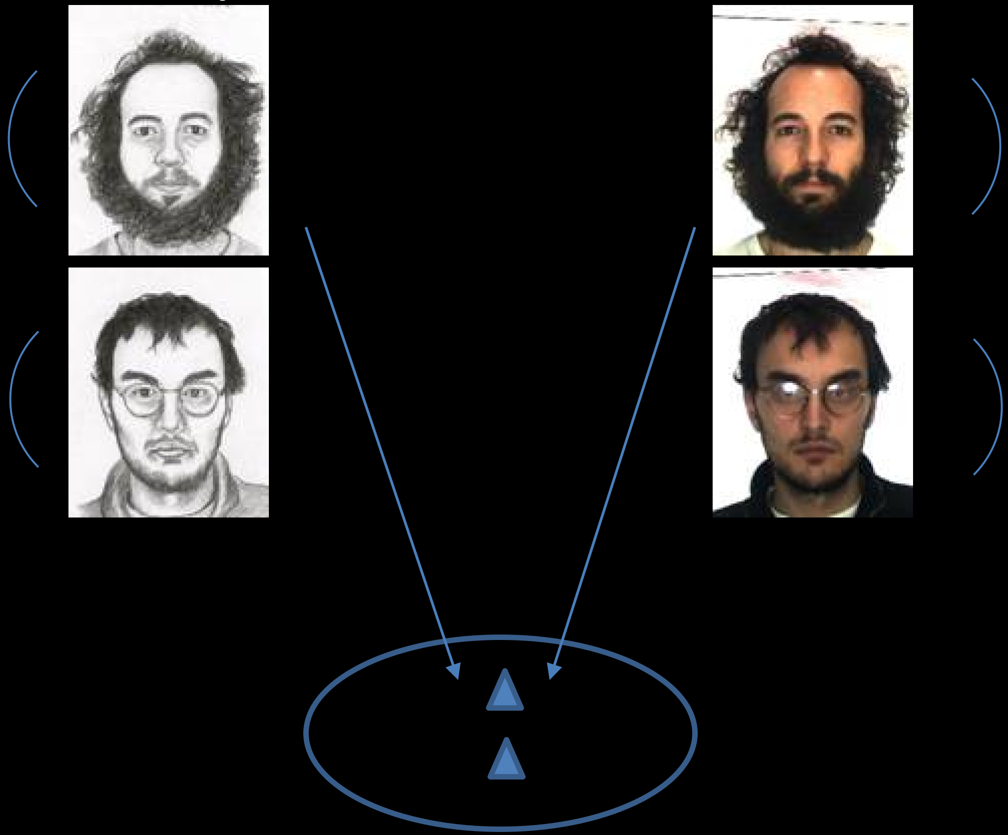
- Extract modality-invariant features



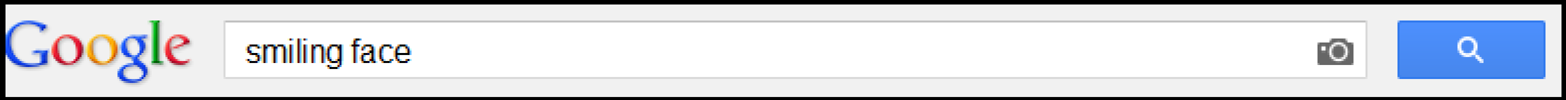
[Zhang *et al.* CVPR 2011]

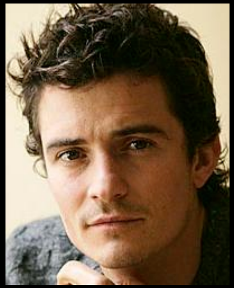
Related Approaches

- Find a latent common space between different modalities using absolutely-paired observations (CCA, DCCA, PLS)

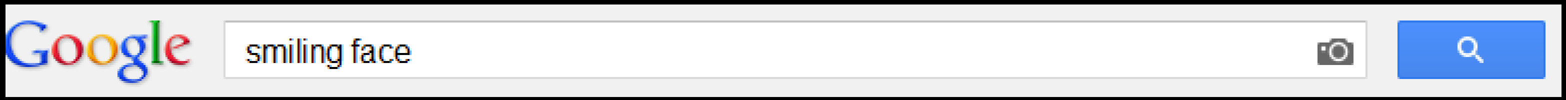


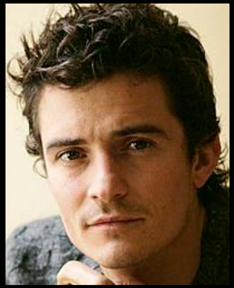
Relatively-Paired Scenario

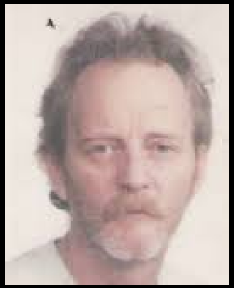




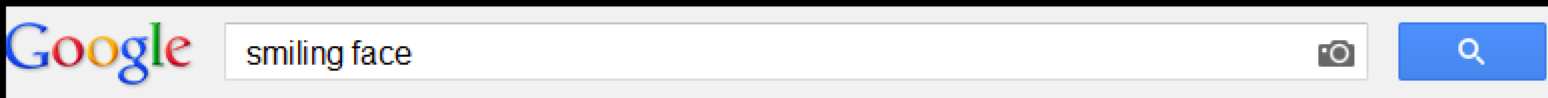
Relatively-Paired Scenario

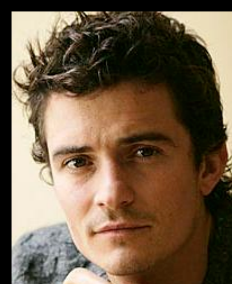




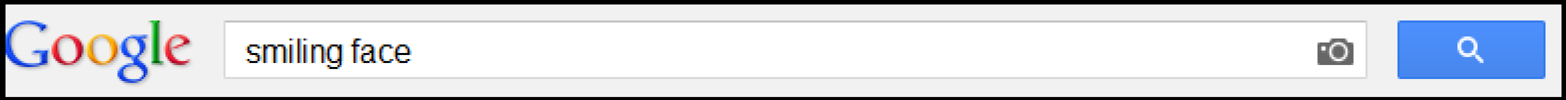


Relatively-Paired Scenario

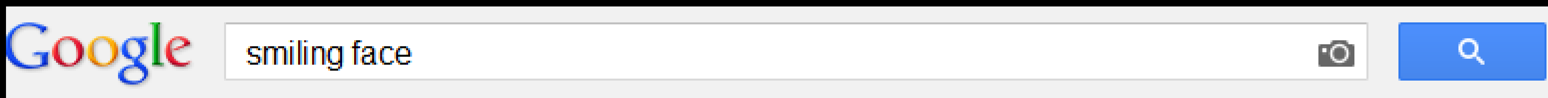




Relatively-Paired Scenario



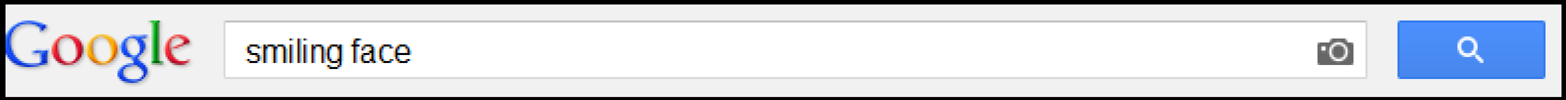
Relatively-Paired Scenario



(

, Smiling face)

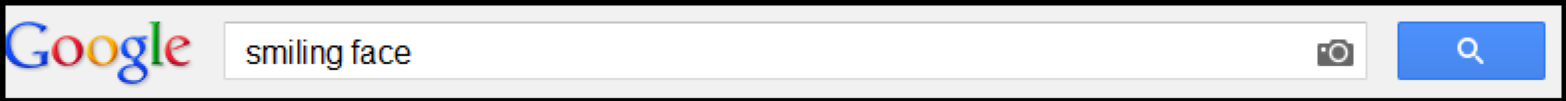
Relatively-Paired Scenario



( , Smiling face)

( , Smiling face)

Relatively-Paired Scenario



( , Smiling face)

V

( , Smiling face)

V

( , Smiling face)

Exploit distance

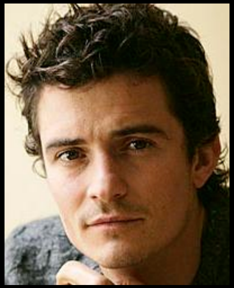
Exploit distance

Distance ( , Smiling face)

Exploit distance

Distance ( , Smiling face)

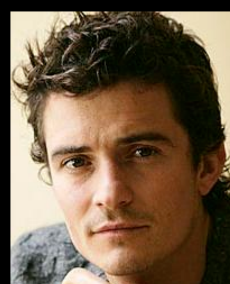
^

Distance ( , Smiling face)

Exploit distance

Distance ( , Smiling face)

^

Distance ( , Smiling face)

^

Distance ( , Smiling face)

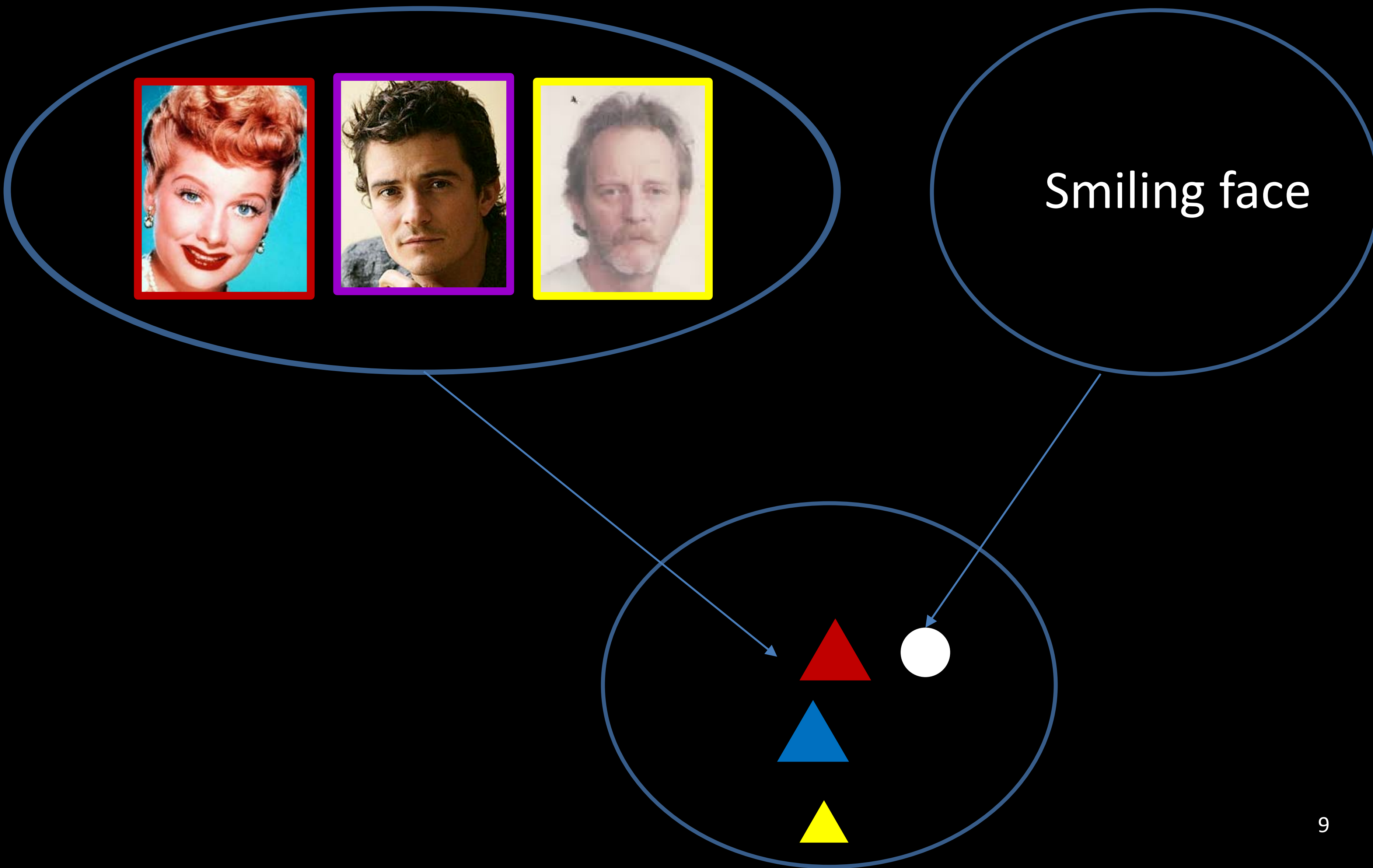
Exploit distance

Distance ( , Smiling face)

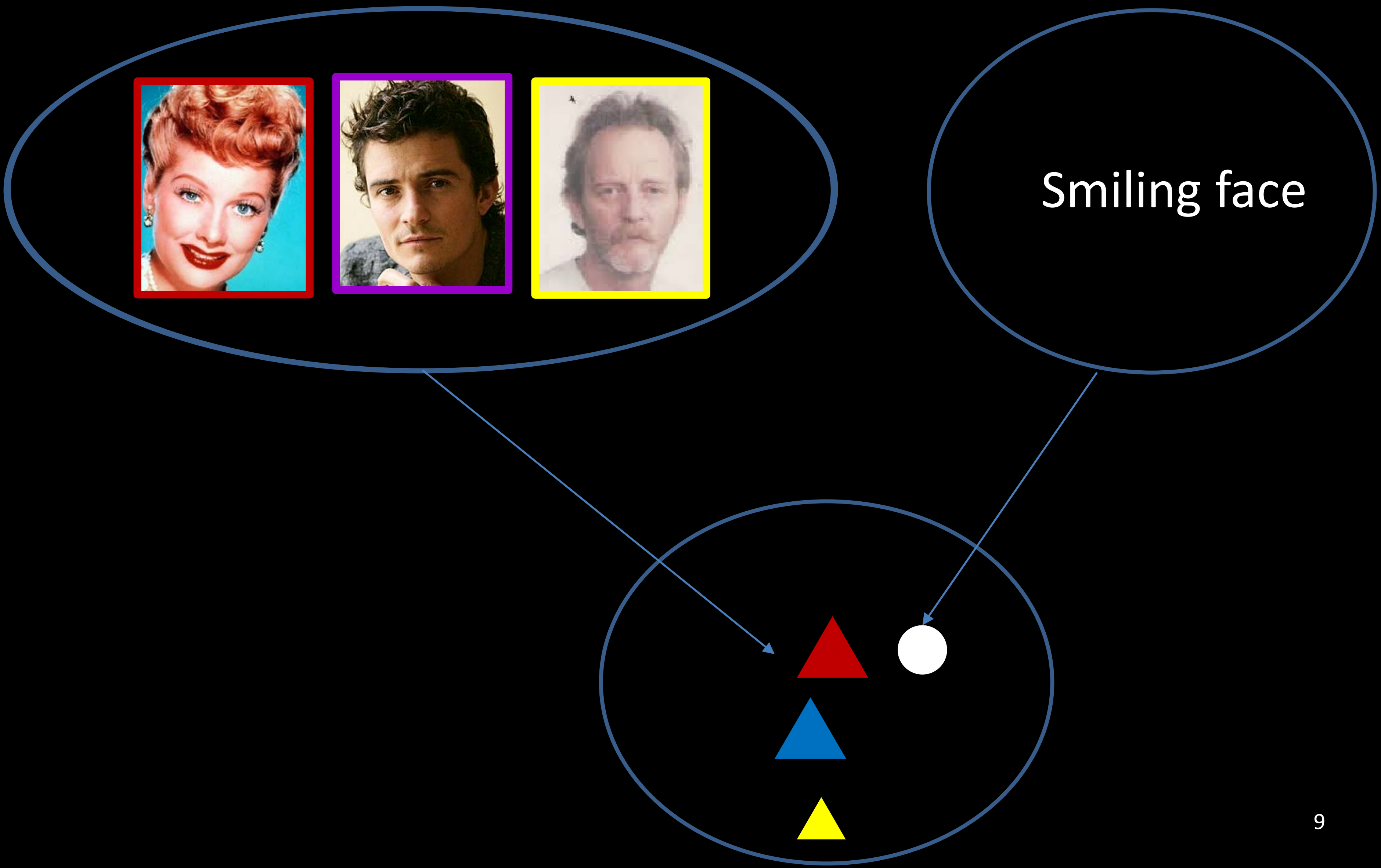
How to define distance?

Distance ( , Smiling face)

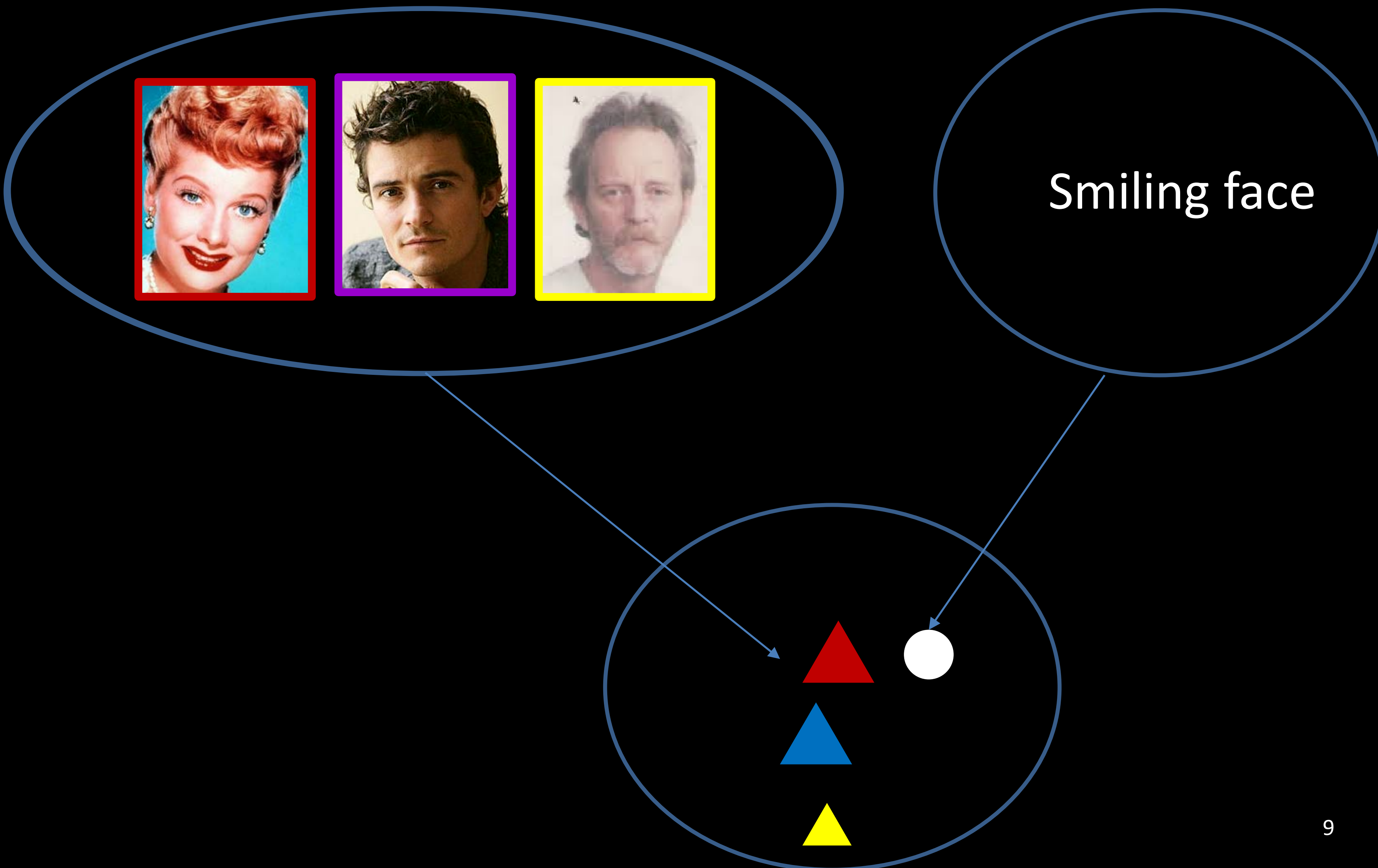
Relatively-Paired Space Analysis



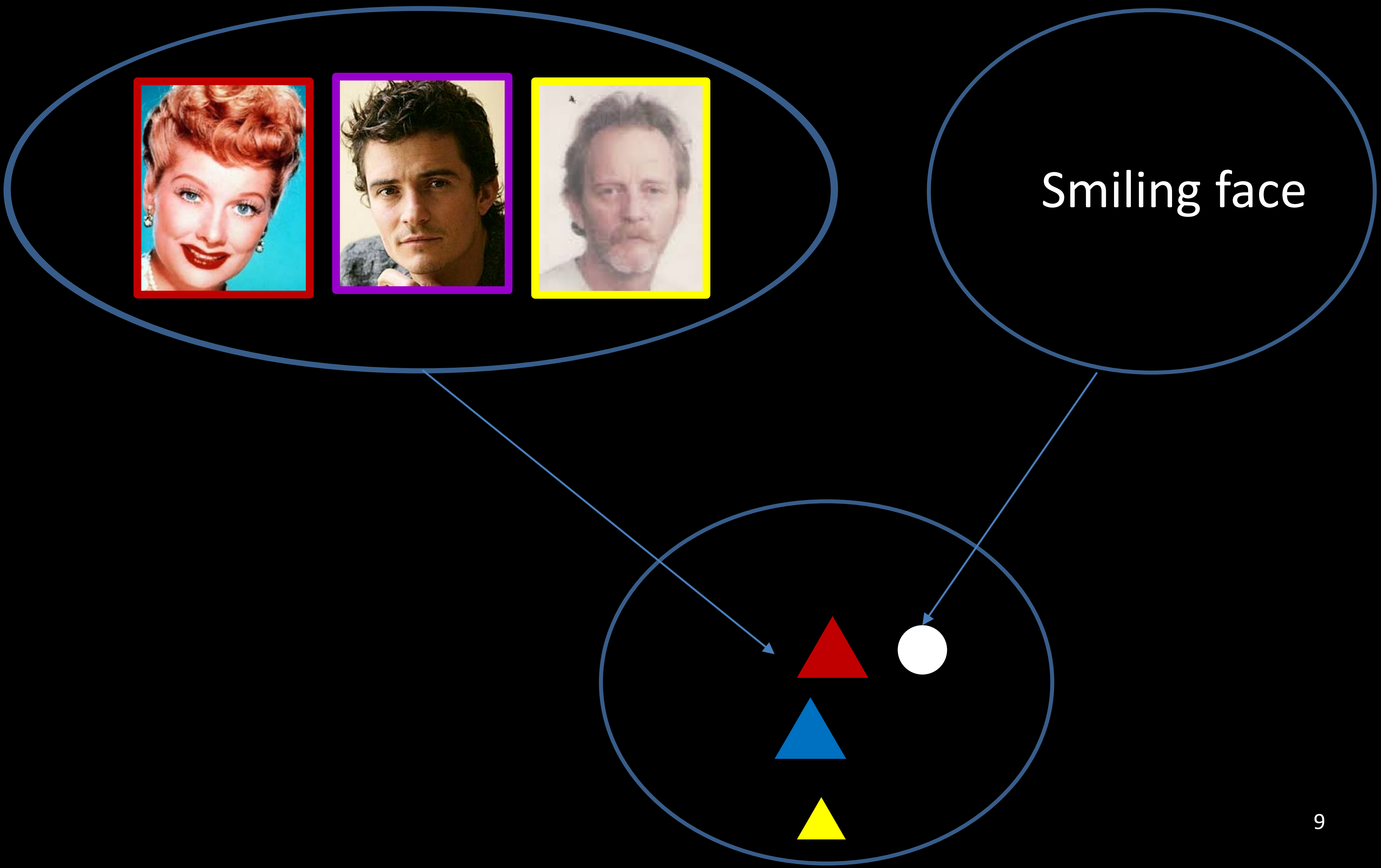
Relatively-Paired Space Analysis



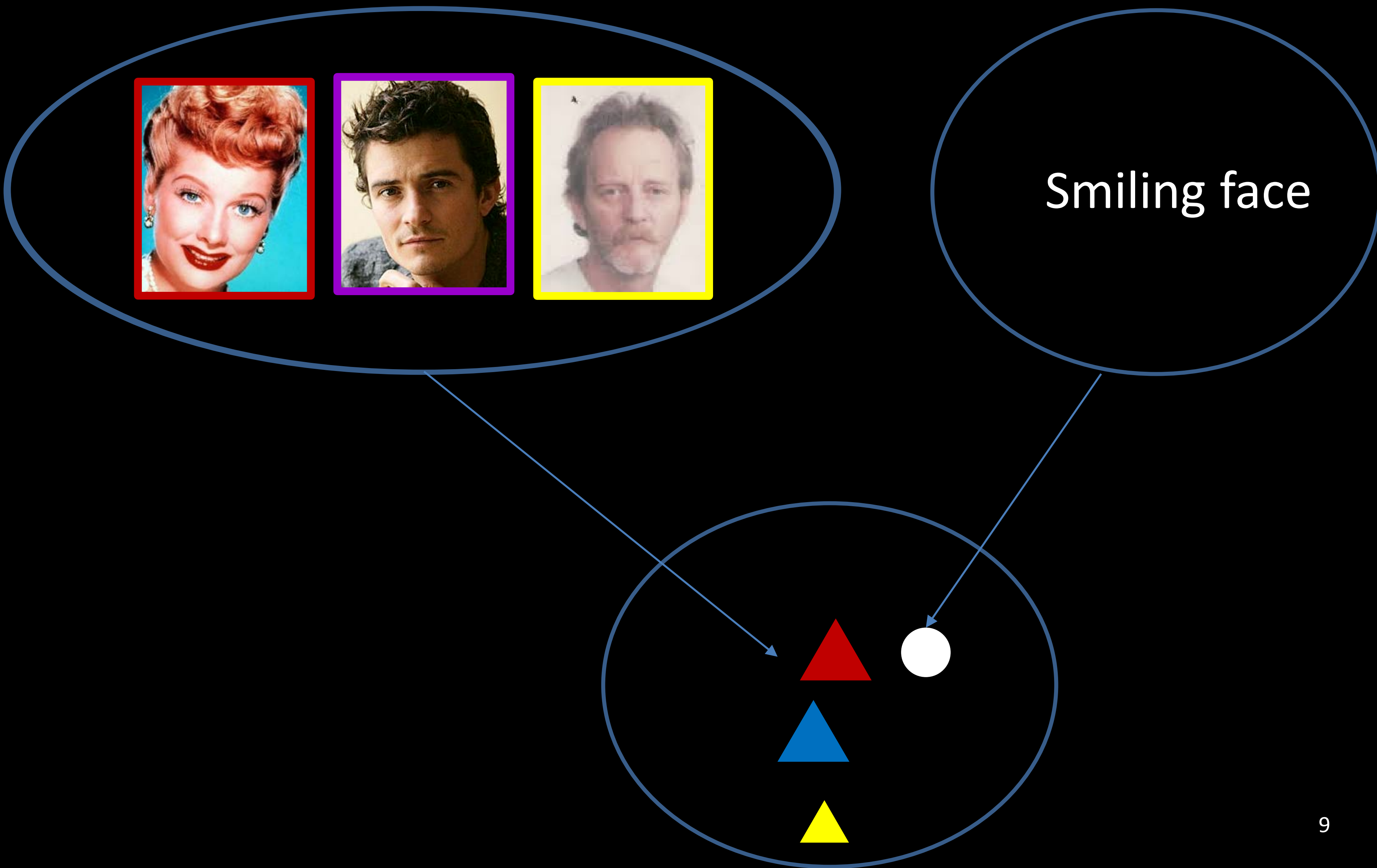
Relatively-Paired Space Analysis



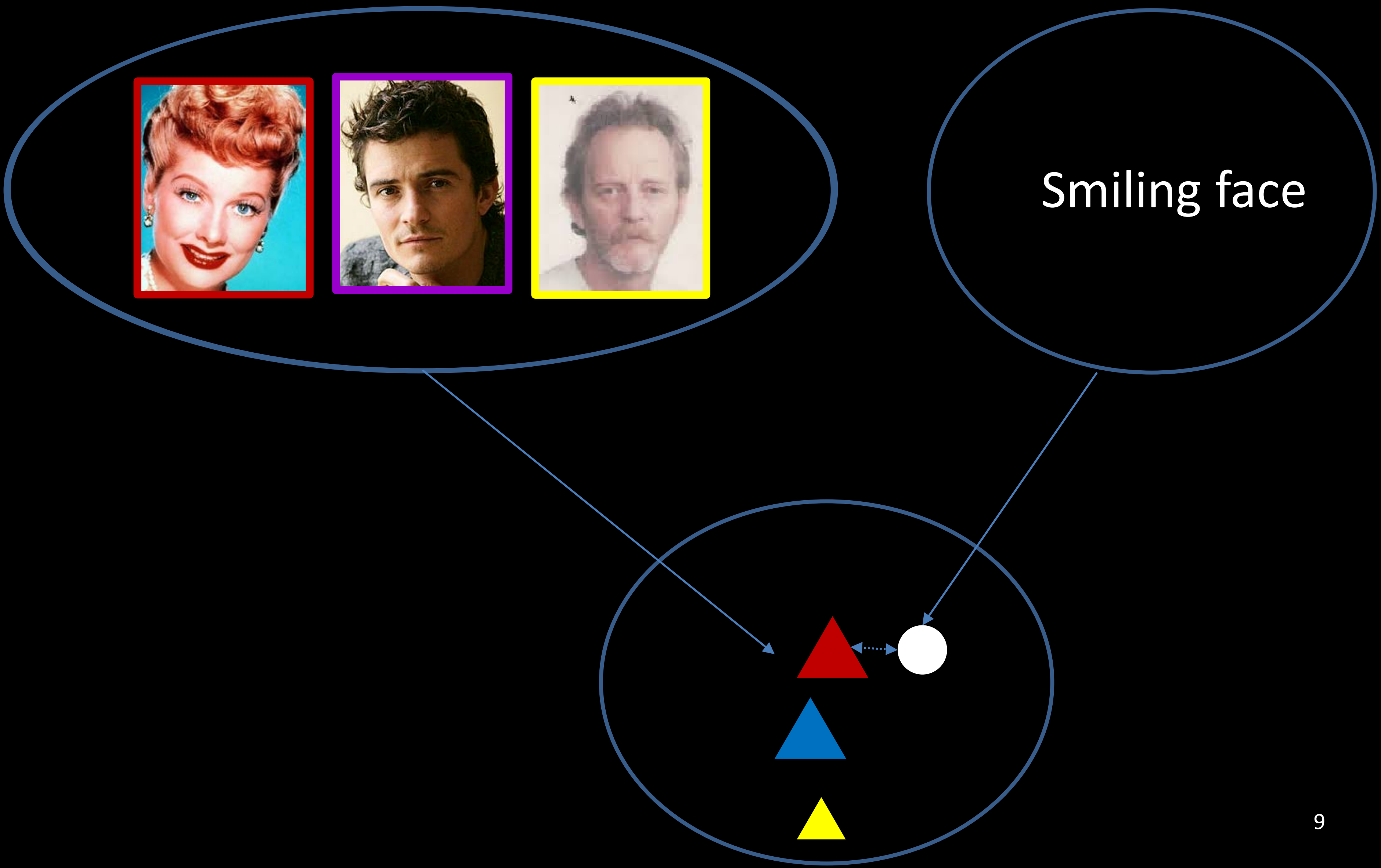
Relatively-Paired Space Analysis



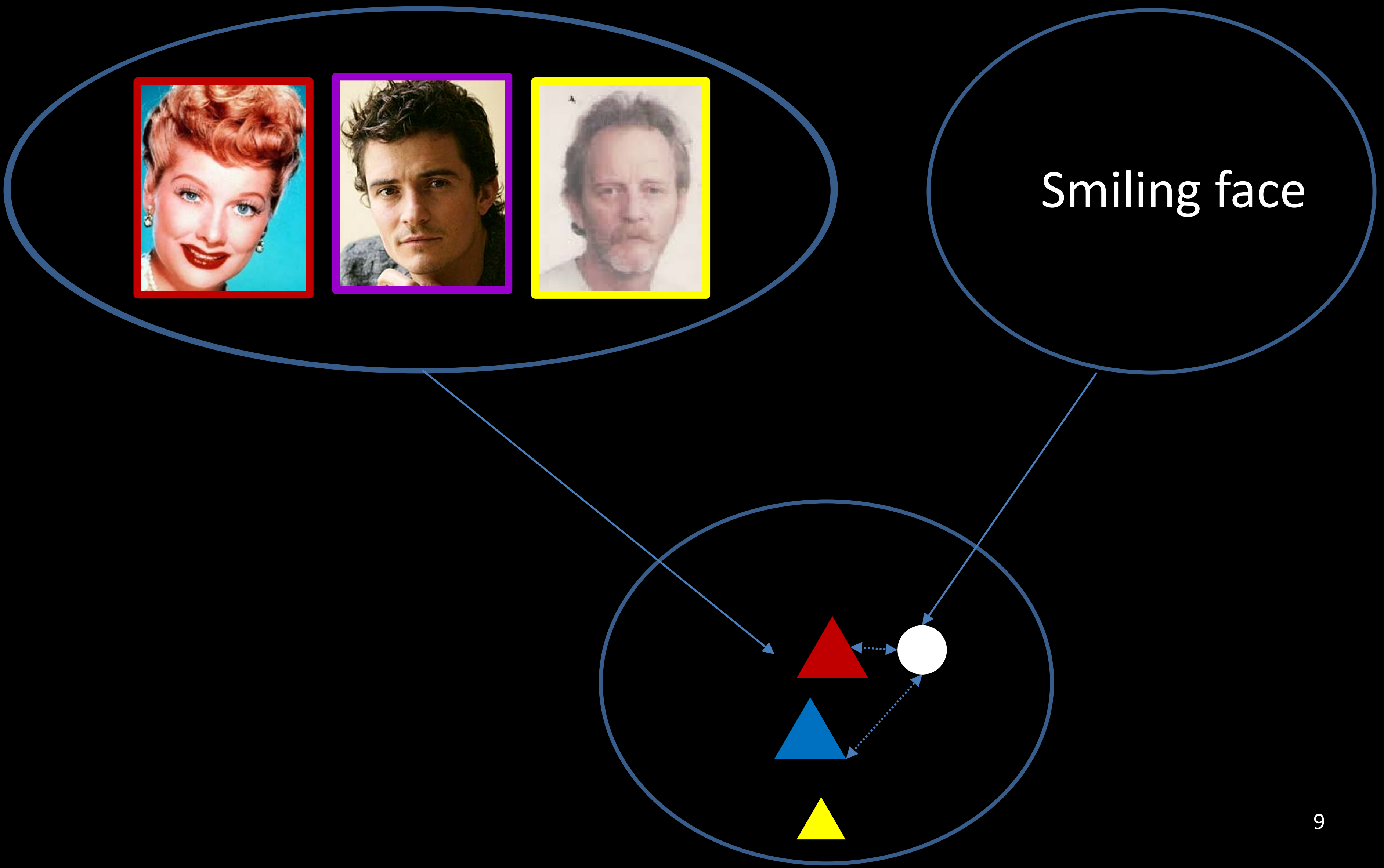
Relatively-Paired Space Analysis



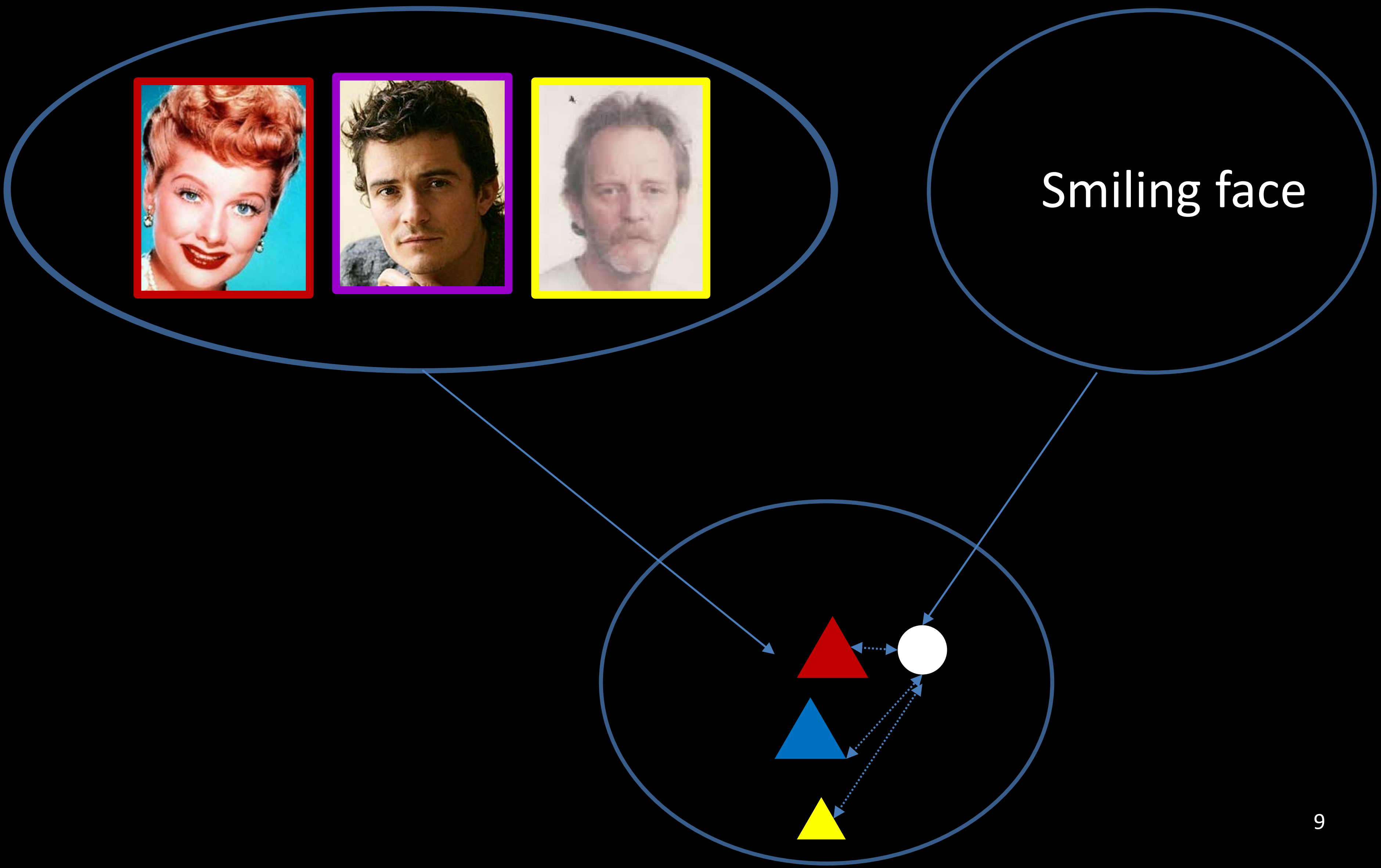
Relatively-Paired Space Analysis



Relatively-Paired Space Analysis



Relatively-Paired Space Analysis



Relatively-Paired Space Analysis

$$\min \| A \|_F^2 \quad + \quad \gamma \sum \ell\,(x_i, x_j, x_k)$$

Relatively-Paired Space Analysis

$$\min \boxed{\| A \|_F^2} \qquad + \qquad \gamma \sum \ell\,(x_i, x_j, x_k)$$

Relatively-Paired Space Analysis

$$\min \| A \|_F^2 \quad + \quad \gamma \sum \ell\,(x_i, x_j, x_k)$$

Relatively-Paired Space Analysis

$$\min \| A \|_F^2 \qquad + \qquad \gamma \boxed{\sum \ell\,(x_i, x_j, x_k)}$$

Relatively-Paired Space Analysis

$$\min \| A \|_F^2 \qquad + \qquad \gamma \sum \ell\,(x_i, x_j, x_k)$$

Relatively-Paired Space Analysis

$$\min \| A \|_F^2 \quad + \quad \boxed{\gamma} \sum \ell \left(x_i, x_j, x_k\right)$$

Relatively-Paired Space Analysis

$$\min \| A \|_F^2 \qquad + \qquad \gamma \sum \ell\,(x_i, x_j, x_k)$$

Relatively-Paired Space Analysis

$$\min \parallel A \parallel_F^2 \quad + \quad \gamma \sum \ell\,(x_i, x_j, x_k)$$

$$\boxed{A = W^T W, W = [W_1, \ldots, W_M]}$$

Relatively-Paired Space Analysis

$$\min \| A \|_F^2 \quad + \quad \gamma \sum \ell\,(x_i, x_j, x_k)$$

$$\boxed{A = W^T W, W = [W_1, \ldots, W_M]}$$

$$\boxed{\begin{gathered} \left(1 - Tr(AC_{ijk})\right)_+ \\ C_{ijk} = C_{ik} - C_{ij} \\ C_{ij} = (A_{\Omega_{t_i}} x_i - A_{\Omega_{t_j}} x_j)\left(A_{\Omega_{t_i}} x_i - A_{\Omega_{t_j}} x_j\right)^T \\ W_{\Omega_m} = W A_{\Omega_m} \end{gathered}}$$

Relatively-Paired Space Analysis

$$\min \| A \|_F^2 \quad + \quad \gamma \sum \ell\,(x_i, x_j, x_k)$$

$$\boxed{A = W^T W, W = [W_1, \dots, W_M]}$$

$$\left(1 - Tr(AC_{ijk})\right)_+$$

$$C_{ijk} = C_{ik} - C_{ij}$$

$$C_{ij} = (A_{\Omega_{t_i}} x_i - A_{\Omega_{t_j}} x_j)\left(A_{\Omega_{t_i}} x_i - A_{\Omega_{t_j}} x_j\right)^T$$

$$W_{\Omega_m} = W A_{\Omega_m}$$

Primal problem

$$\min \| A \|_F^2 + \gamma \sum \left(1 - Tr(AC_{ijk})\right)_+$$

Relatively-Paired Space Analysis

Dual problem

$$\min \frac{1}{2} \| Z + \sum u_{ijk} C_{ijk} \|_F - \sum u_{ijk}$$

$$s.t. Z \succcurlyeq 0, 0 \leq 0 \leq \frac{\gamma}{m}$$

Repeat

Fixing u, optimize Z
Fixing Z, optimize u

Until convergence

Relatively-Paired Space Analysis

Dual problem

$$\min \frac{1}{2} \parallel Z + \sum u_{ijk} C_{ijk} \parallel_F - \sum u_{ijk}$$

$$s.t. Z \succcurlyeq 0, 0 \leq 0 \leq \frac{\gamma}{m}$$

Repeat

Fixing u, optimize Z
Fixing Z, optimize u

Until convergence

$$\boxed{Z^* = \left(-\sum u_{ijk} C_{ijk}\right)_+}$$

Relatively-Paired Space Analysis

Dual problem

$$\min \frac{1}{2} \| Z + \sum u_{ijk} C_{ijk} \|_F - \sum u_{ijk}$$

$$s.t. Z \succcurlyeq 0, 0 \leq 0 \leq \frac{\gamma}{m}$$

Repeat

Fixing u, optimize Z
Fixing Z, optimize u

Until convergence

$$\boxed{Z^* = \left(-\sum u_{ijk} C_{ijk}\right)_+}$$

Time complexity $A = O(td^3)$

Experimental Results

- Feature fusion

Dataset: UCI Multiple Features dataset with six features, Fou, Fac, Kar, Pix, Zer and Mor
How to fuse: sum of projection of paired feature from different modalities
How to evaluate: Mean recognition rates of 1-NN in the latent common space over 10 times

Recognition rates on Multiple Features Dataset.

Pairs		DCCA	bgCCA	bgDCCA	bsCCA	bsDCCA	PLS	RCE	RPSA
Fac	Fou	0.89	0.86	0.89	0.84	0.88	0.94	0.95	**0.97**
Fac	Kar	0.98	0.95	0.98	0.93	0.98	0.94	**0.98**	**0.98**
Fac	Pix	0.97	0.86	0.97	0.86	0.97	0.94	0.95	**0.98**
Fac	Zer	0.88	0.86	0.88	0.84	0.87	0.96	**0.97**	**0.97**
Fac	Mor	0.82	0.75	0.82	0.74	0.81	0.88	0.88	**0.92**
Fou	Kar	0.90	0.90	0.90	0.88	0.89	**0.97**	0.96	**0.97**
Fou	Pix	0.89	0.77	0.89	0.74	0.87	**0.98**	0.95	**0.98**
Fou	Zer	0.83	0.82	0.83	0.80	0.81	0.81	**0.85**	**0.85**
Fou	Mor	0.77	0.75	0.77	0.74	0.76	0.44	**0.80**	0.74
Kar	Pix	0.95	0.94	0.95	0.93	0.94	**0.98**	0.96	0.97
Kar	Zer	0.88	0.90	0.88	0.89	0.86	0.83	**0.96**	0.94
Kar	Mor	0.80	0.77	0.80	0.76	0.79	0.62	**0.86**	0.84
Pix	Zer	0.87	0.83	0.87	0.80	0.86	0.84	0.94	**0.95**
Pix	Mor	0.79	0.73	0.79	0.71	0.77	0.71	0.84	**0.88**
Zer	Mor	0.75	0.72	0.75	0.70	0.74	0.72	**0.77**	0.72

Experimental Results

- Feature fusion

Dataset: UCI Multiple Features dataset with six features, Fou, Fac, Kar, Pix, Zer and Mor
How to fuse: sum of projection of paired feature from different modalities
How to evaluate: Mean recognition rates of 1-NN in the latent common space over 10 times

Recognition rates on Multiple Features Dataset.

Pairs		DCCA	bgCCA	bgDCCA	bsCCA	bsDCCA	PLS	RCE	RPSA
Fac	Fou	0.89	0.86	0.89	0.84	0.88	0.94	0.95	**0.97**
Fac	Kar	0.98	0.95	0.98	0.93	0.98	0.94	**0.98**	**0.98**
Fac	Pix	0.97	0.86	0.97	0.86	0.97	0.94	0.95	**0.98**
Fac	Zer	0.88	0.86	0.88	0.84	0.87	0.96	**0.97**	**0.97**
Fac	Mor	0.82	0.75	0.82	0.74	0.81	0.88	0.88	**0.92**
Fou	Kar	0.90	0.90	0.90	0.88	0.89	**0.97**	0.96	**0.97**
Fou	Pix	0.89	0.77	0.89	0.74	0.87	**0.98**	0.95	**0.98**
Fou	Zer	0.83	0.82	0.83	0.80	0.81	0.81	**0.85**	**0.85**
Fou	Mor	0.77	0.75	0.77	0.74	0.76	0.44	**0.80**	0.74
Kar	Pix	0.95	0.94	0.95	0.93	0.94	**0.98**	0.96	0.97
Kar	Zer	0.88	0.90	0.88	0.89	0.86	0.83	**0.96**	0.94
Kar	Mor	0.80	0.77	0.80	0.76	0.79	0.62	**0.86**	0.84
Pix	Zer	0.87	0.83	0.87	0.80	0.86	0.84	0.94	**0.95**
Pix	Mor	0.79	0.73	0.79	0.71	0.77	0.71	0.84	**0.88**
Zer	Mor	0.75	0.72	0.75	0.70	0.74	0.72	**0.77**	0.72

Experimental Results

- Cross-pose face recognition

Dataset: CMUPIE with 68 subjects and 13 poses
How to evaluate: Mean recognition rates of 1-NN in the latent common space

Mean recognition rates for frontal faces (c27) gallery.

Gallery	Probe	Method	Accuracy	Method	Accuracy
c27	c05/37/25/22/29/11/14/34	PGFR	0.86	RPSA	**0.94**
c27	c05/22	TFA	**0.95**	RPSA	0.93
c27	c05/29/37/11/07/09	LLR	0.95	RPSA	**1.00**
c27	c05/29/37/11/07/09	ELF	0.90	RPSA	**1.00**
c27	c05/29/37/11/07/09/02/14/22/34/25/31	PLS	0.94	RPSA	**0.95**

[Liu and Chen CVPR 2005, Prince *et al.* PAMI 2008， Chai et al. TIP 2005, Gross et al. PAMI 2004, Sharma and Jacobs CVPR 2011]

Experimental Results

- Cross-pose face recognition

Dataset: CMUPIE with 68 subjects and 13 poses

How to evaluate: Mean recognition rates of 1-NN in the latent common space

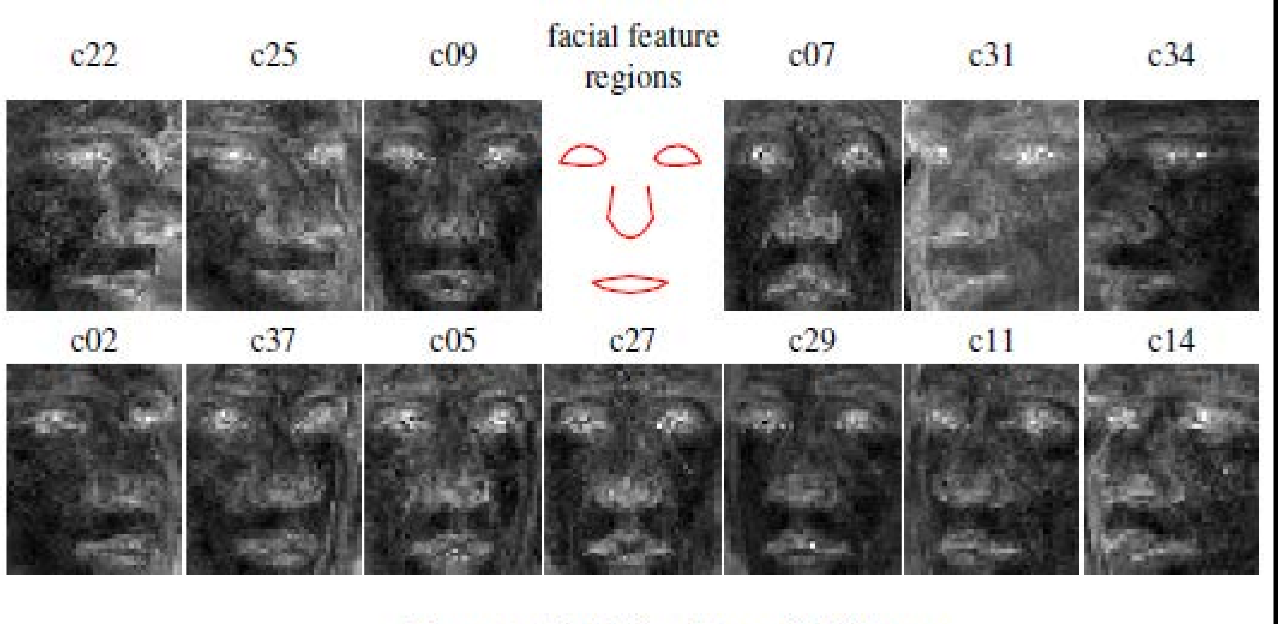
Mean recognition rates for frontal faces (c27) gallery.

Gallery	Probe	Method	Accuracy	Method	Accuracy
c27	c05/37/25/22/29/11/14/34	PGFR	0.86	RPSA	**0.94**
c27	c05/22	TFA	**0.95**	RPSA	0.93
c27	c05/29/37/11/07/09	LLR	0.95	RPSA	**1.00**
c27	c05/29/37/11/07/09	ELF	0.90	RPSA	**1.00**
c27	c05/29/37/11/07/09/02/14/22/34/25/31	PLS	0.94	RPSA	**0.95**

[Liu and Chen CVPR 2005, Prince *et al.* PAMI 2008， Chai et al. TIP 2005, Gross et al. PAMI 2004, Sharma and Jacobs CVPR 2011]

Experimental Results

- Visualization



Important pixels in photos with 13 poses

Conclusion

Conclusion

- RPSA learns a latent space according to relatively-paired observations

Conclusion

- RPSA learns a latent space according to relatively-paired observations
- RPSA is a general framework used in any pattern recognition across modality

Conclusion

- RPSA learns a latent space according to relatively-paired observations
- RPSA is a general framework used in any pattern recognition across modality
- RPSA achieves state-of-art results

Thanks for your attention.

Questions?

The code is available online
(http://i.cs.hku.hk/~zhkuang/Software.html)