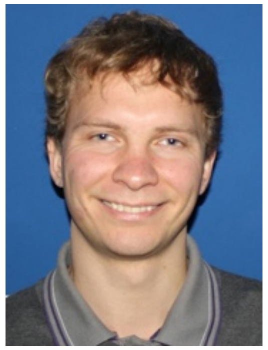
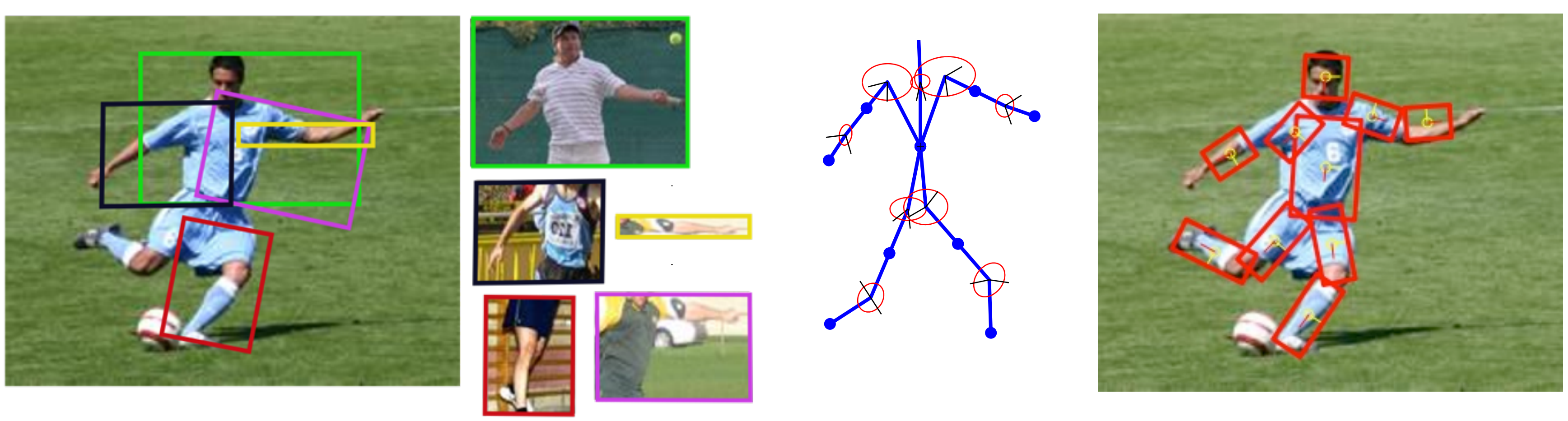
Poselet Conditioned Pictorial Structures




Leonid Pishchulin[1] Micha Andriluka[1] Peter Gehler[2] Bernt Schiele[1]

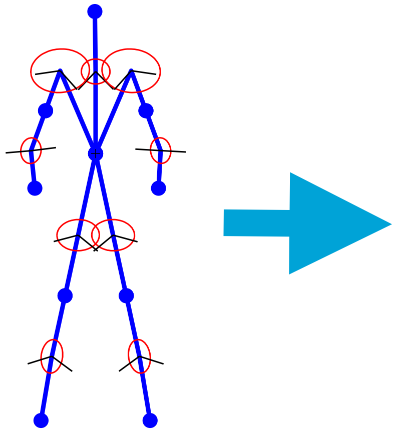
[1]Max Planck Institute for Informatics [2]Max Planck Institute for Intelligent Systems

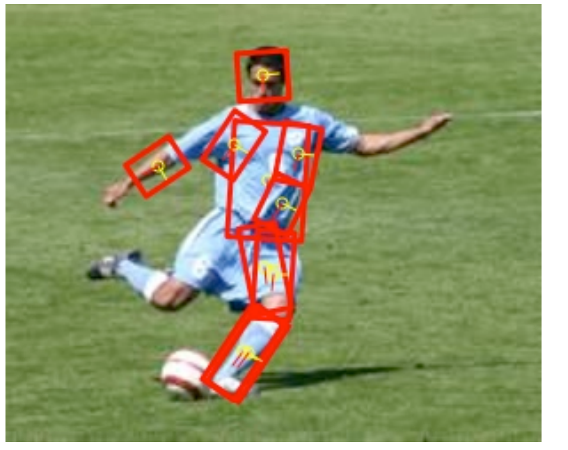


max planck institut
informatik

Motivation

State of the art: tree-structured models



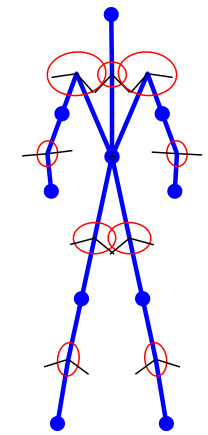


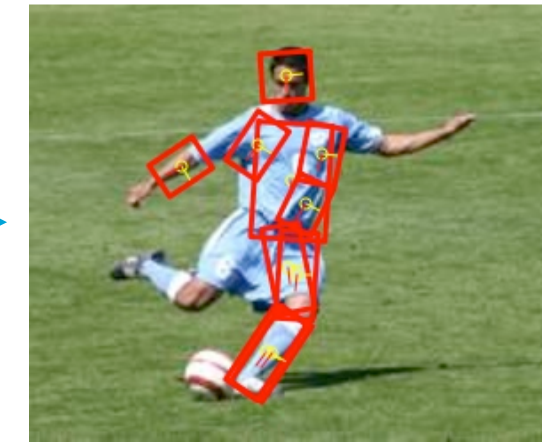
- ***generic*** kinematic tree
- capture ***adjacent*** part dependencies ***only***

Motivation

MAX-PLANCK-GESELLSCHAFT

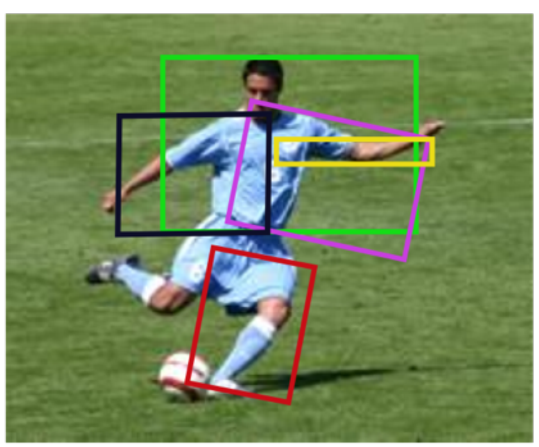
State of the art: tree-structured models

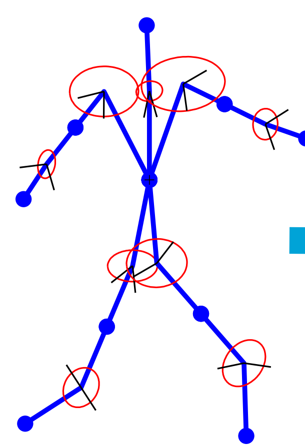


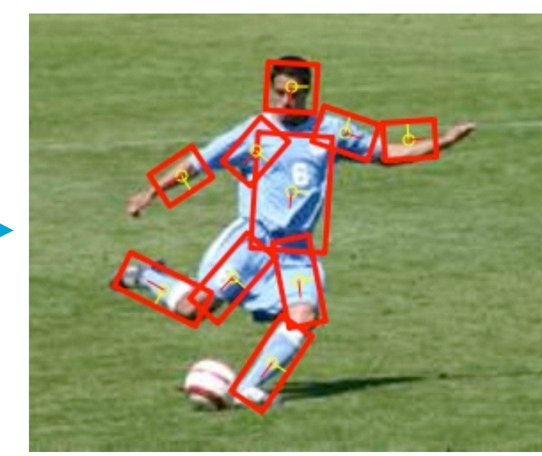


- ***generic*** kinematic tree
- capture ***adjacent*** part dependencies ***only***

This work





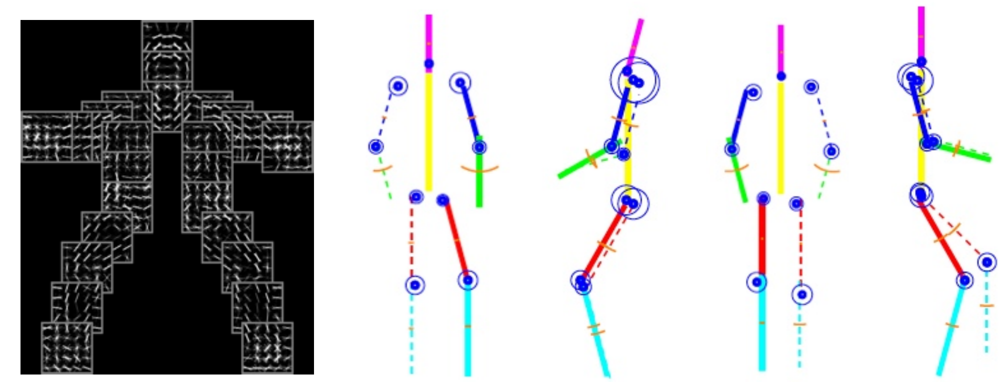


✓ ***poselet conditioned*** kinematic tree

✓ ***poselets*** capture ***non-adjacent*** part dependencies

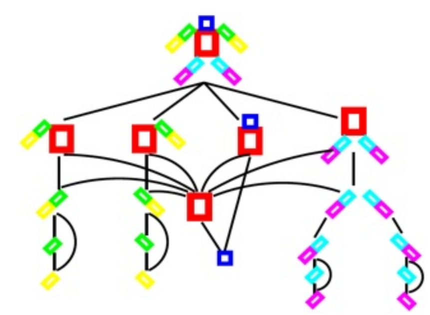
Related work

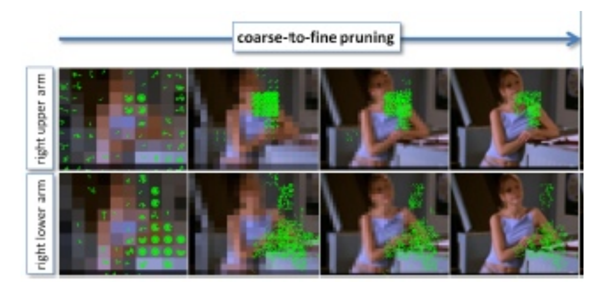
- Mixtures-of-trees models
  - ✓ efficient and exact inference
  - \- capture adjacent part dependencies only



[Johnson&Everingham, BMVC’10]
[Yang&Ramanan, CVPR’11]
[Eichner&Ferrari, ACCV’12] ...

- Loopy models
  - ✓ non-adjacent part dependencies
  - \- approximate inference





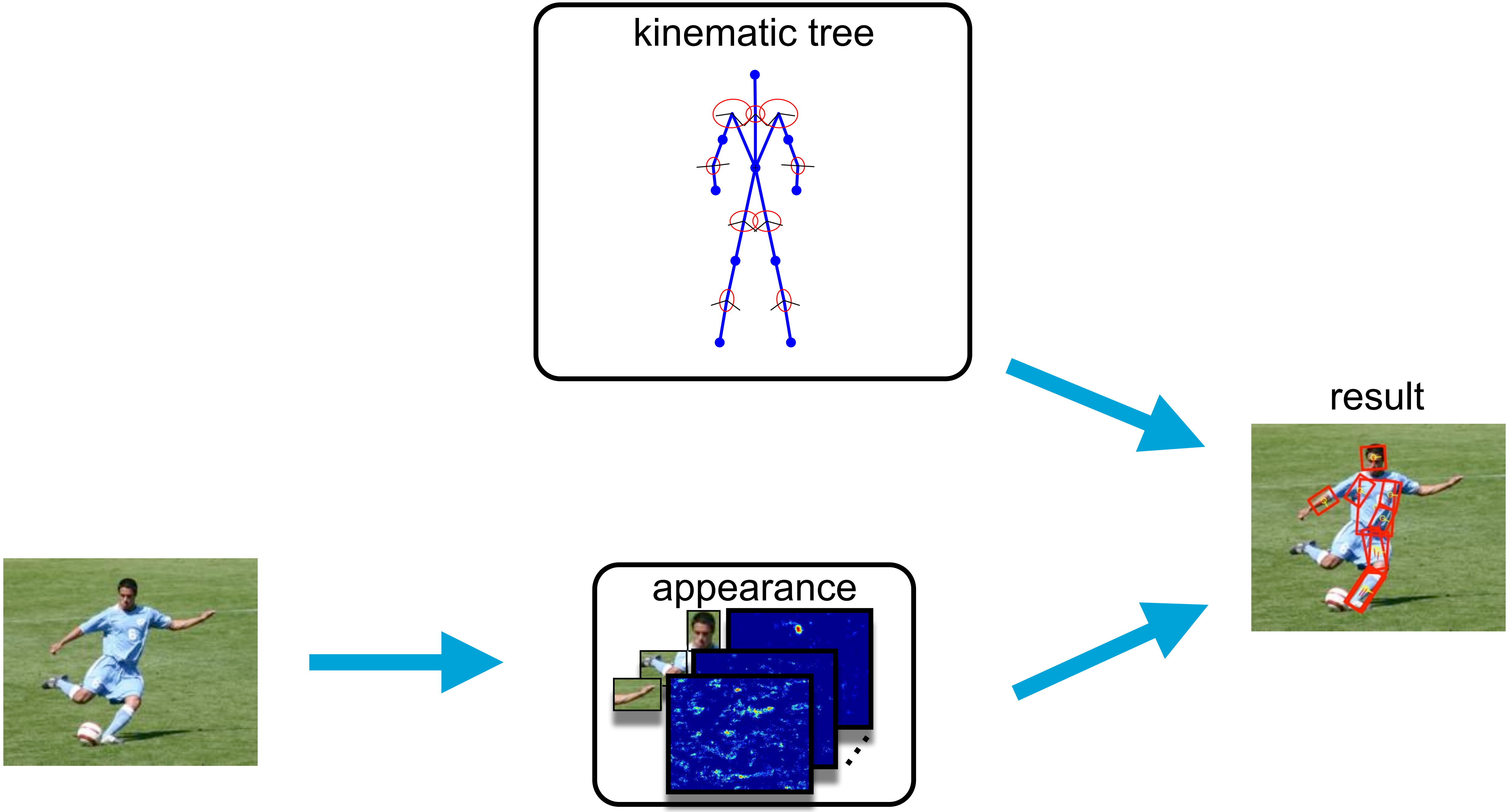
[Sapp et al., ECCV’10]
[Tran&Forsyth, ECCV’10]
[Wang et al., CVPR’11] ...

- Holistic methods
  - ✓ non-adjacent part dependencies
  - \- much training data required

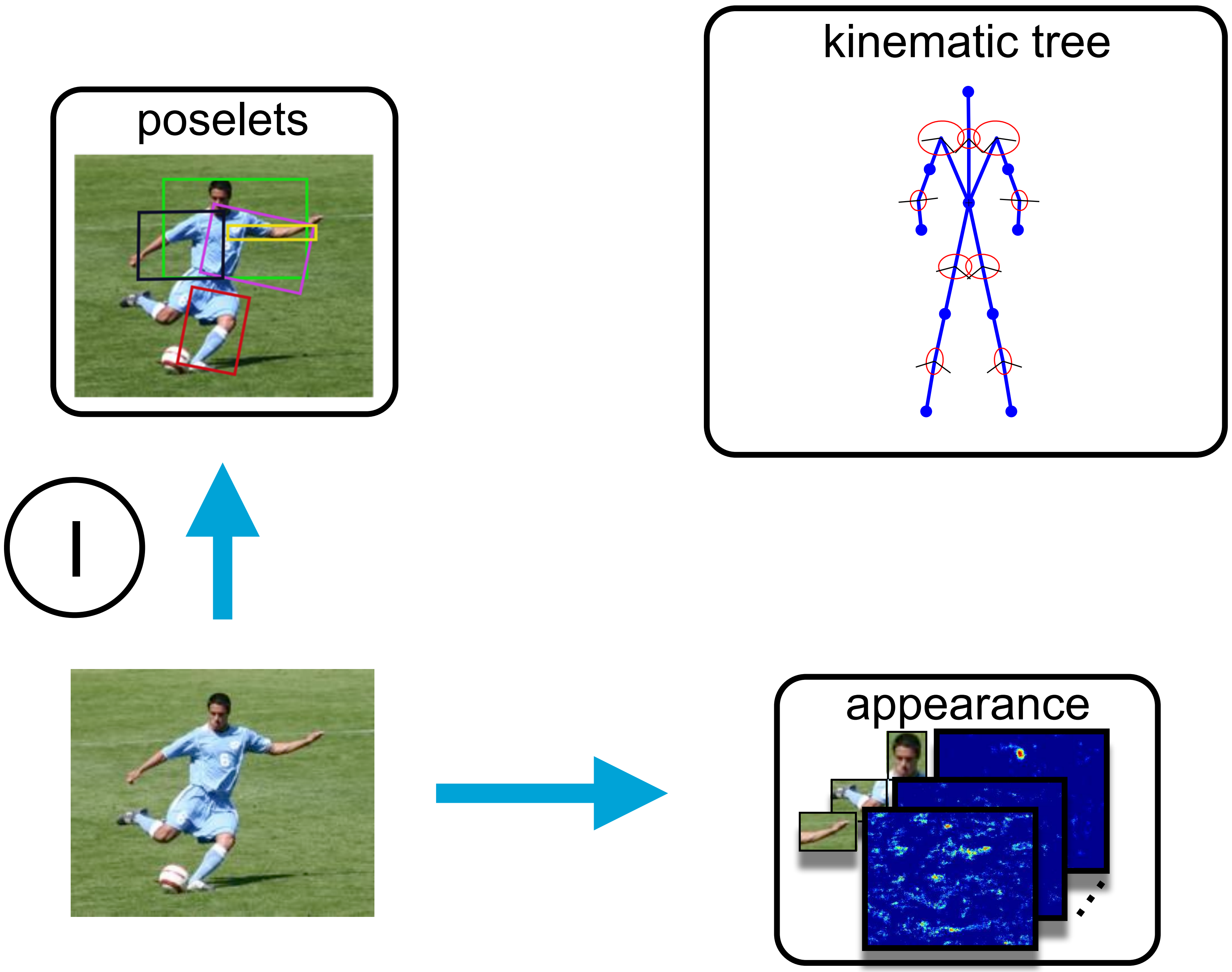
[Agarwal&Triggs, PAMI’02]
[Urtasun&Darrell, ICCV’09]
[Ionescu et al., ICCV’11] ...

Pictorial Structures [Felzenszwalb&Huttenlocher, IJCV'05]

MAX-PLANCK-GESELLSCHAFT



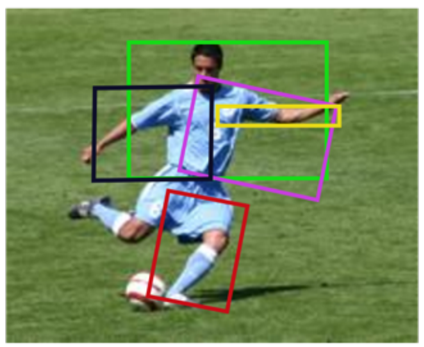
Poselet Conditioned Pictorial Structures



Poselet Conditioned Pictorial Structures

II

poselets

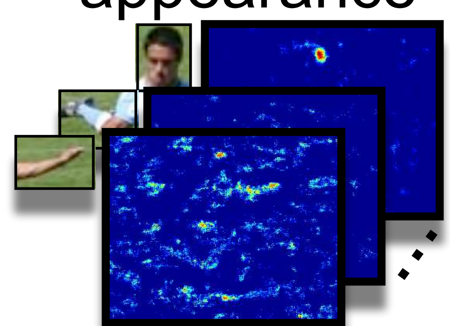


pairwise prediction

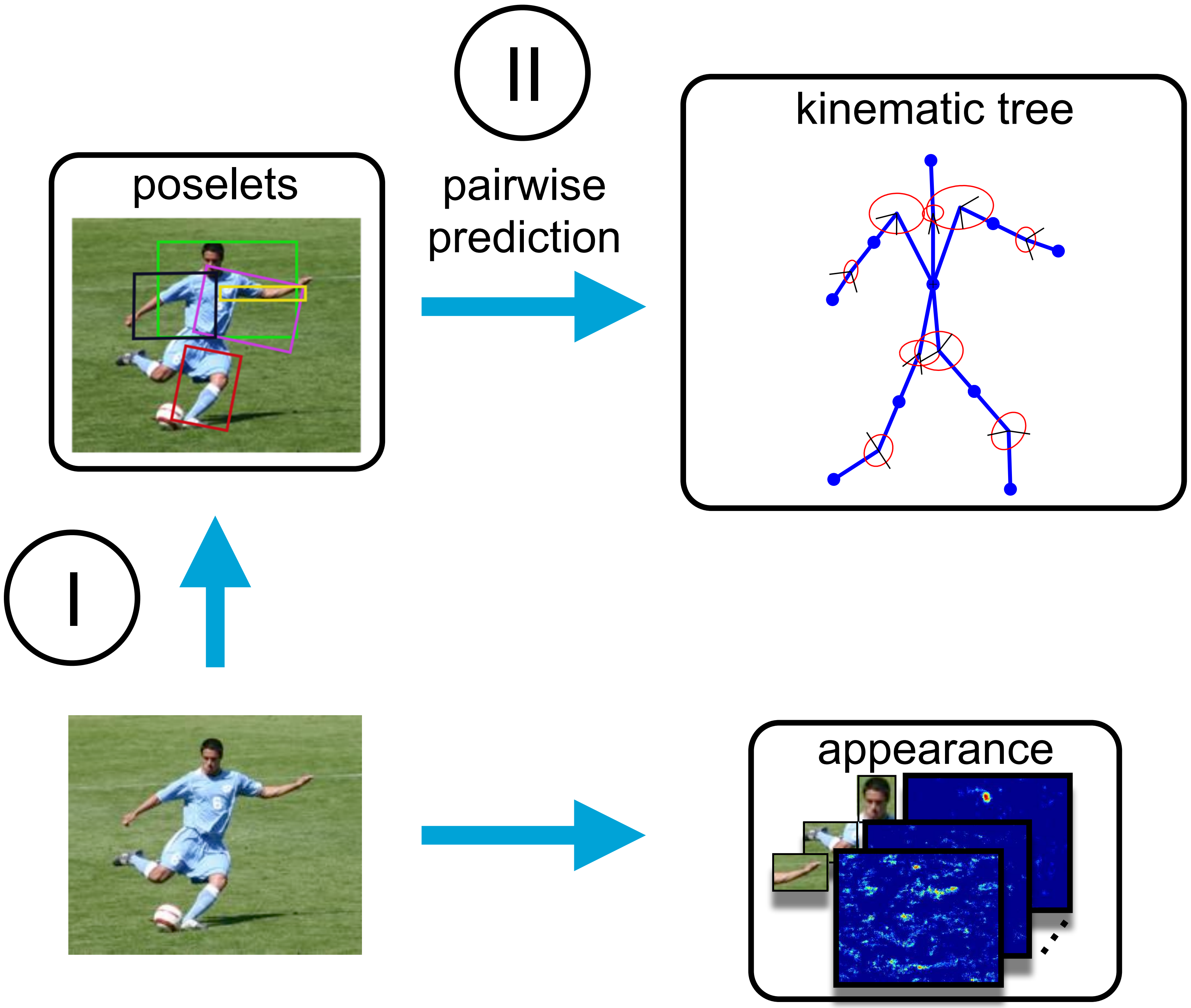
kinematic tree

I

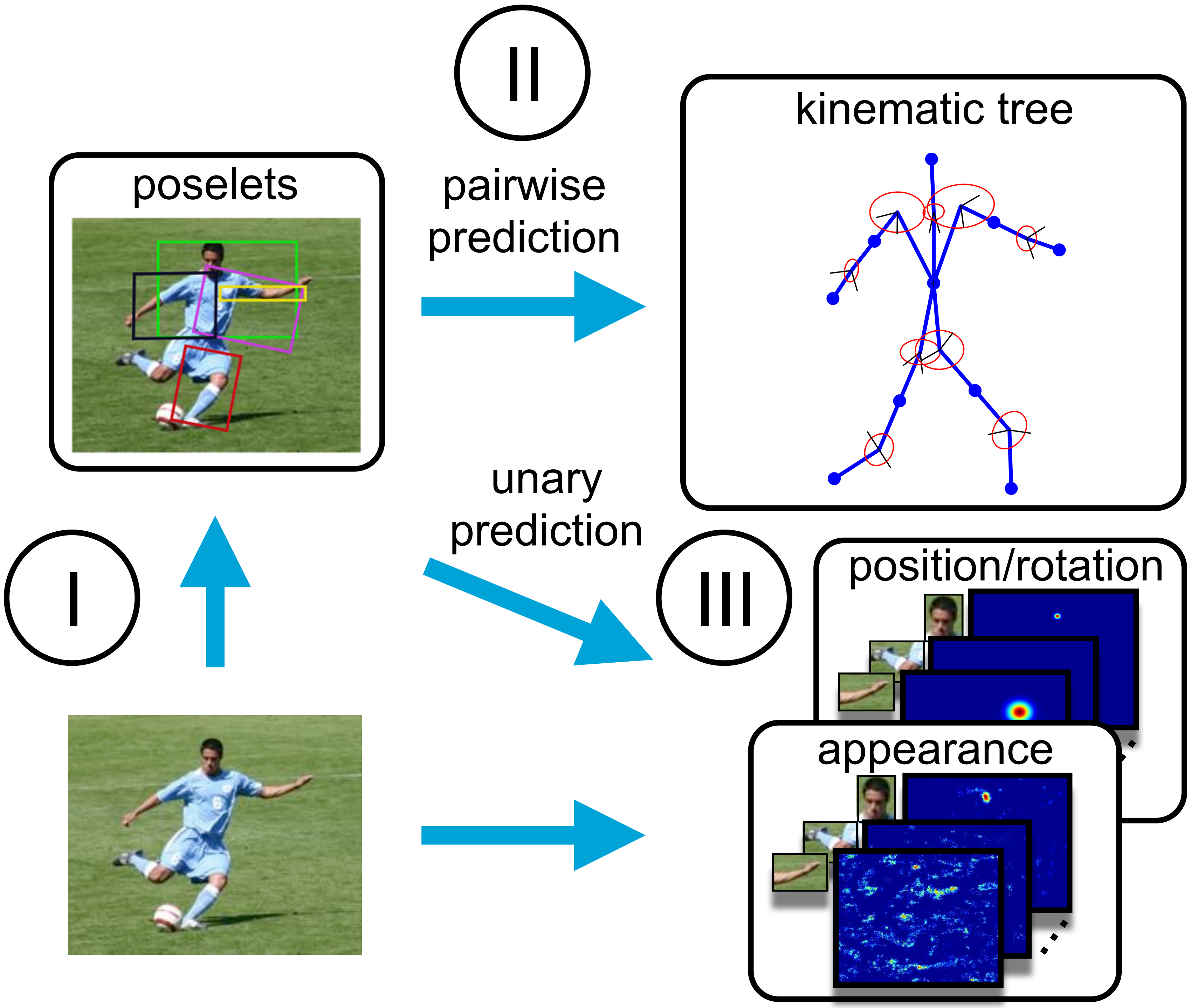
appearance



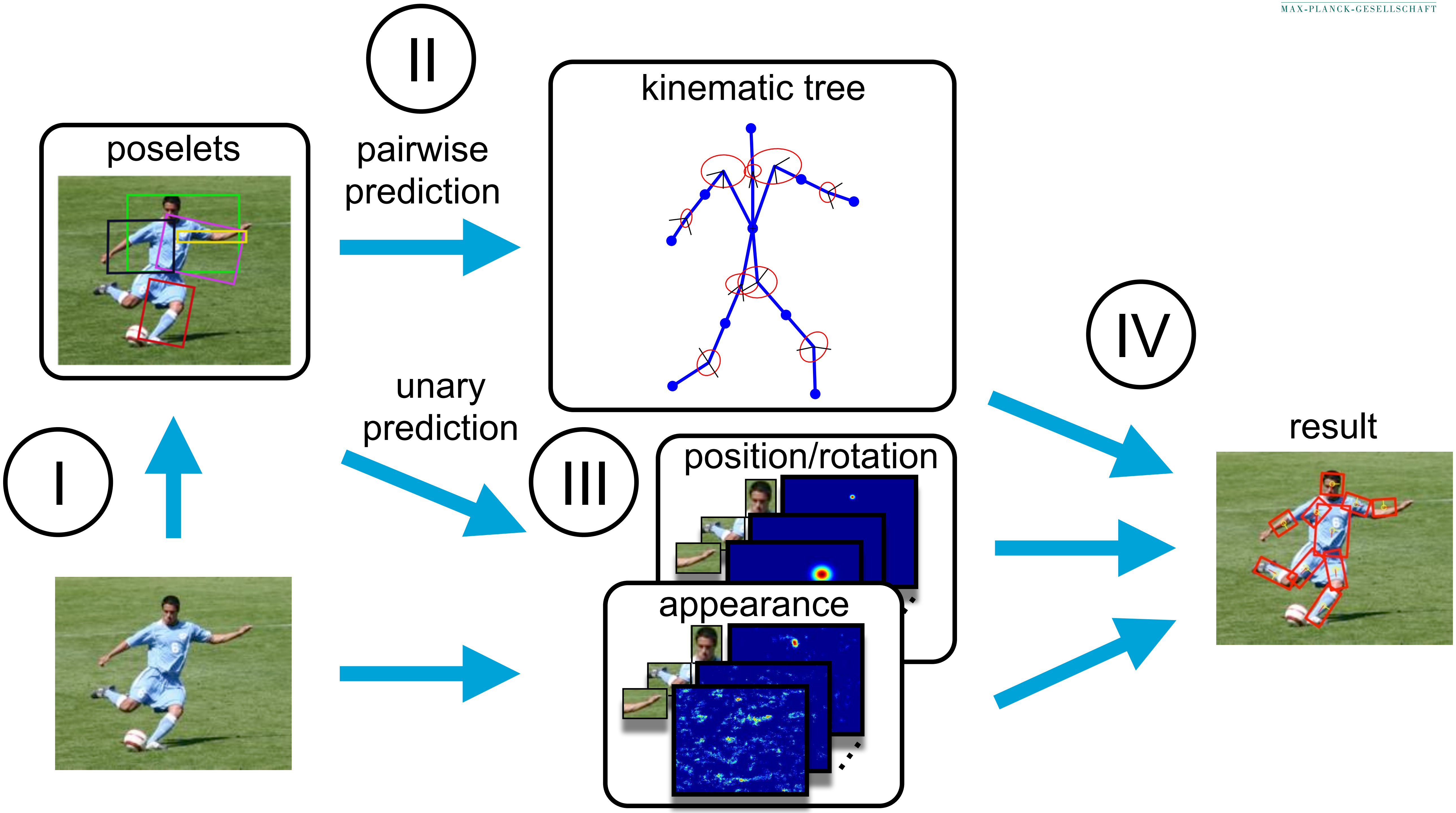
Poselet Conditioned Pictorial Structures



Poselet Conditioned Pictorial Structures



Poselet Conditioned Pictorial Structures

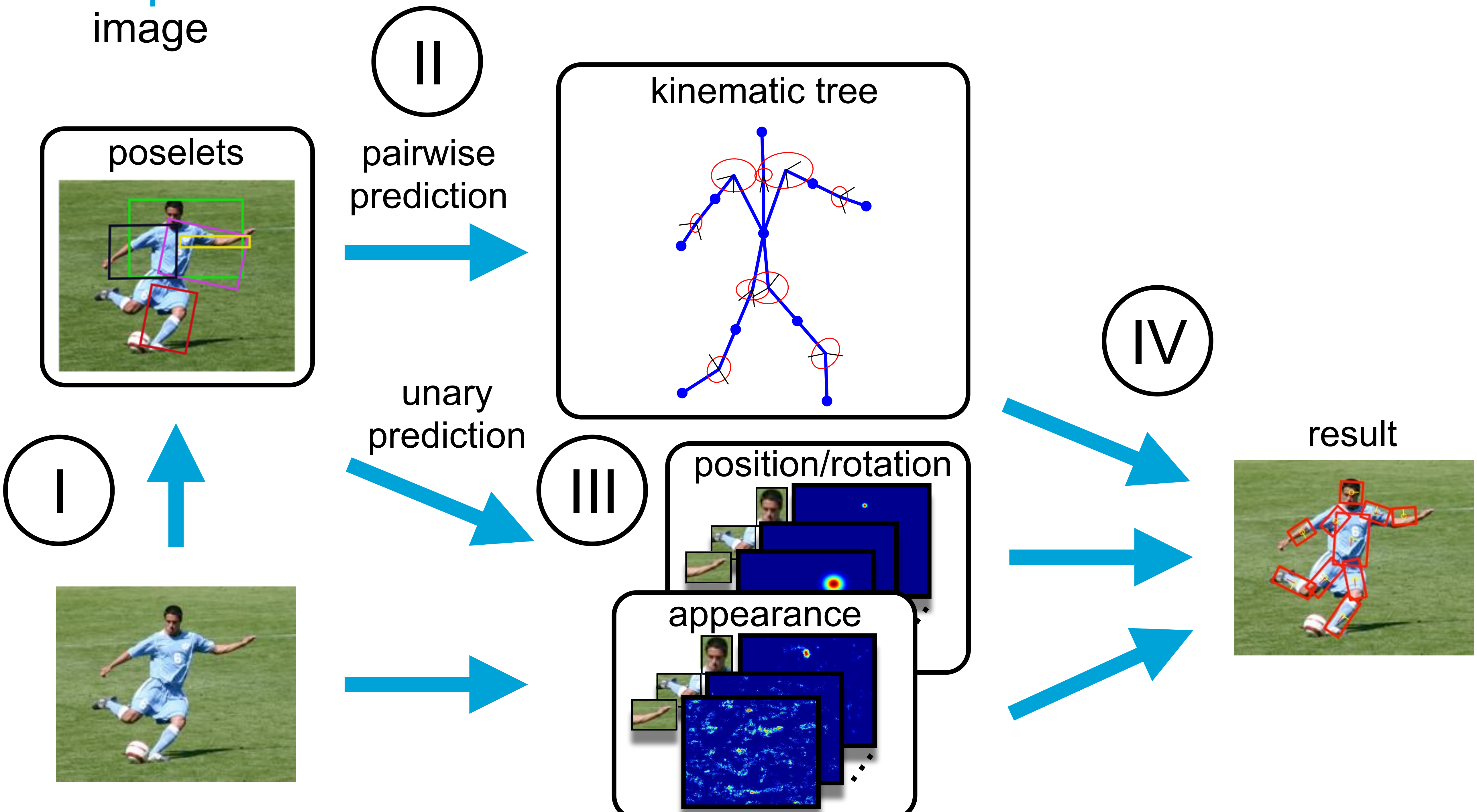


Poselet Conditioned Pictorial Structures

part

$$E(L; D) = \sum_{m=1}^{M} E^u(l_m; D) + \sum_{n \sim m} E^p(l_n, l_m; \quad)$$

image



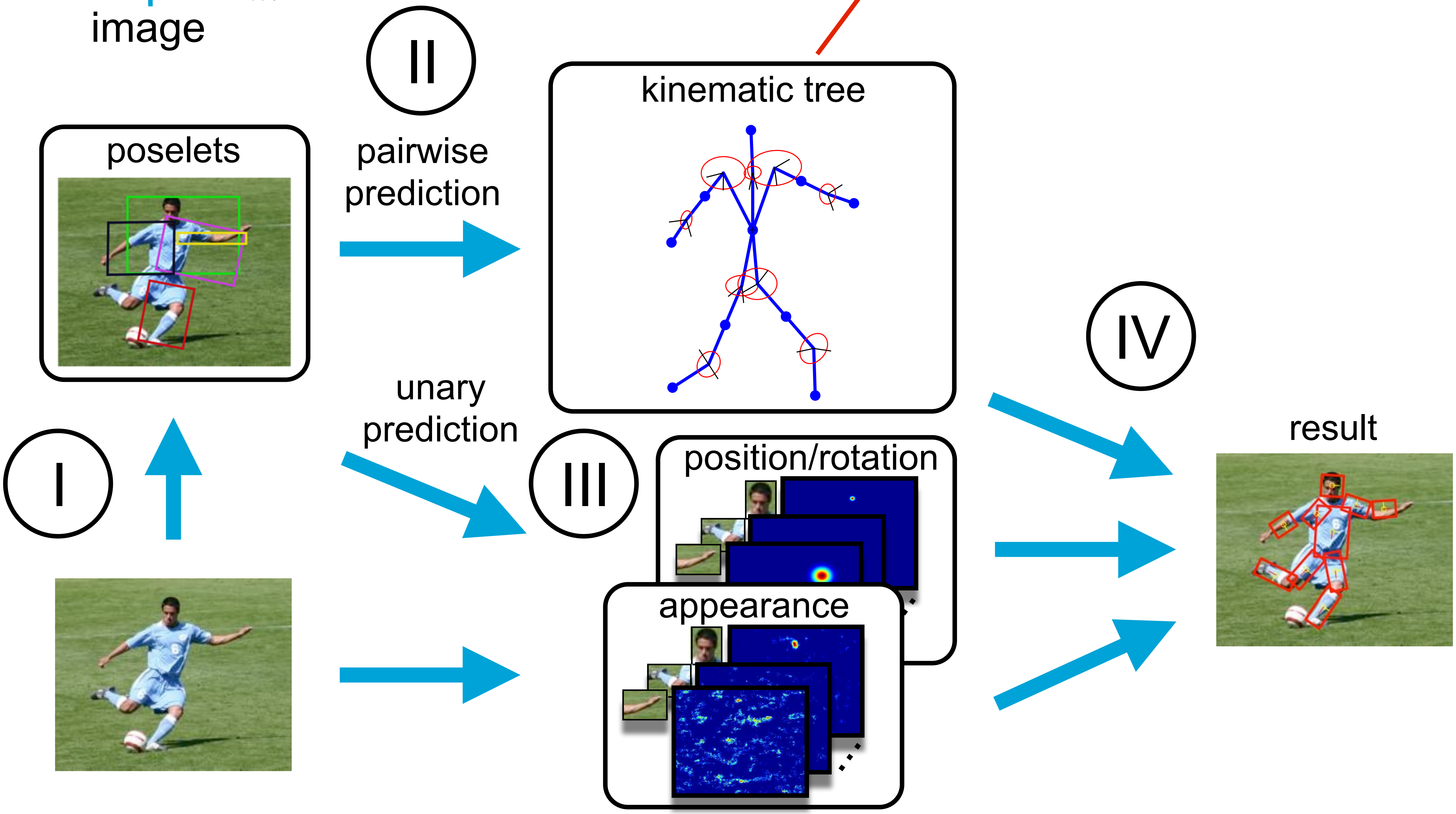
Poselet Conditioned Pictorial Structures

MAX-PLANCK-GESELLSCHAFT

part

$$E(L;D) = \sum_{m=1}^{M} E^u(l_m;D) + \sum_{n\sim m} E^p(l_n,l_m;D)$$

image

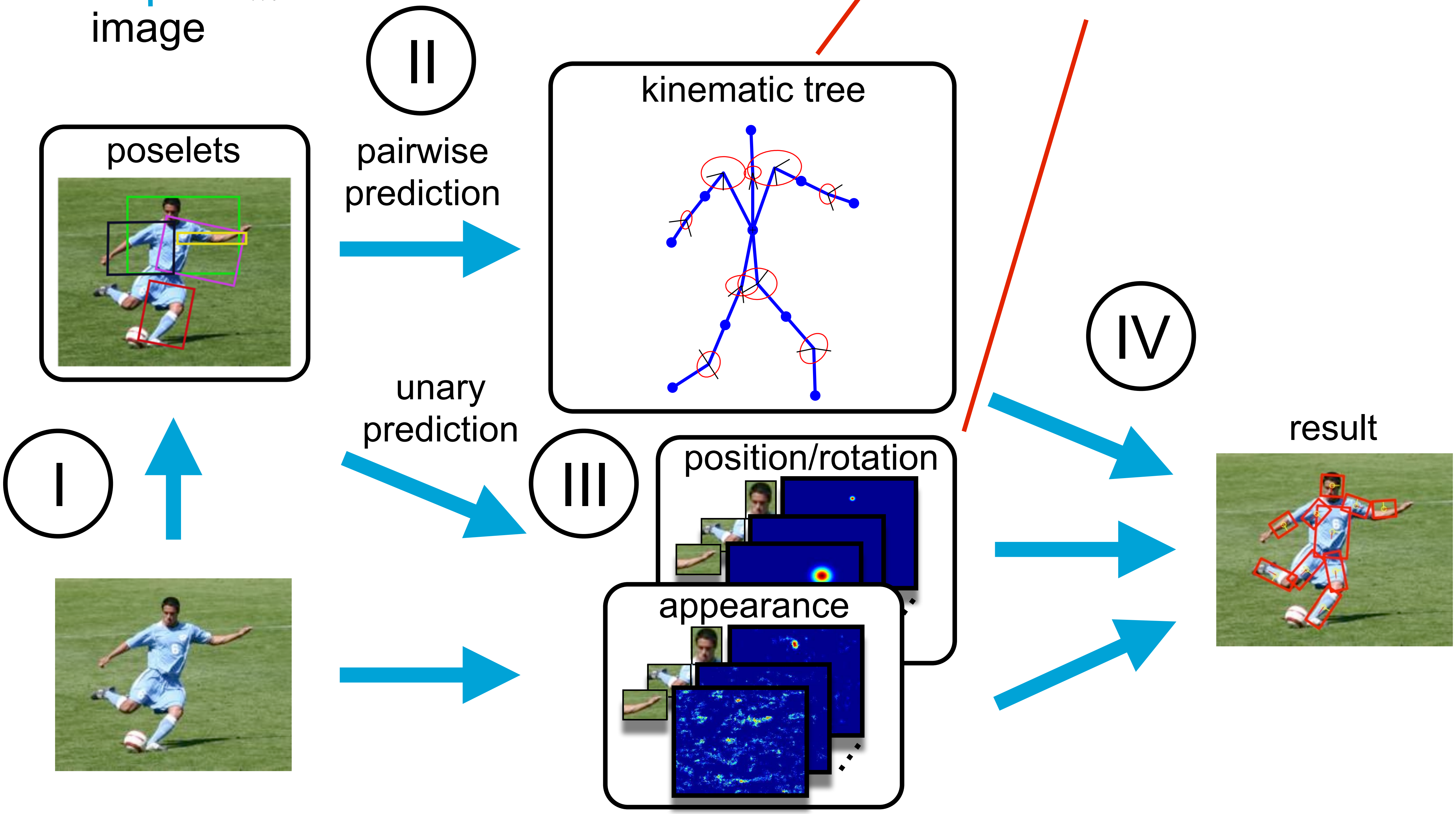


Poselet Conditioned Pictorial Structures

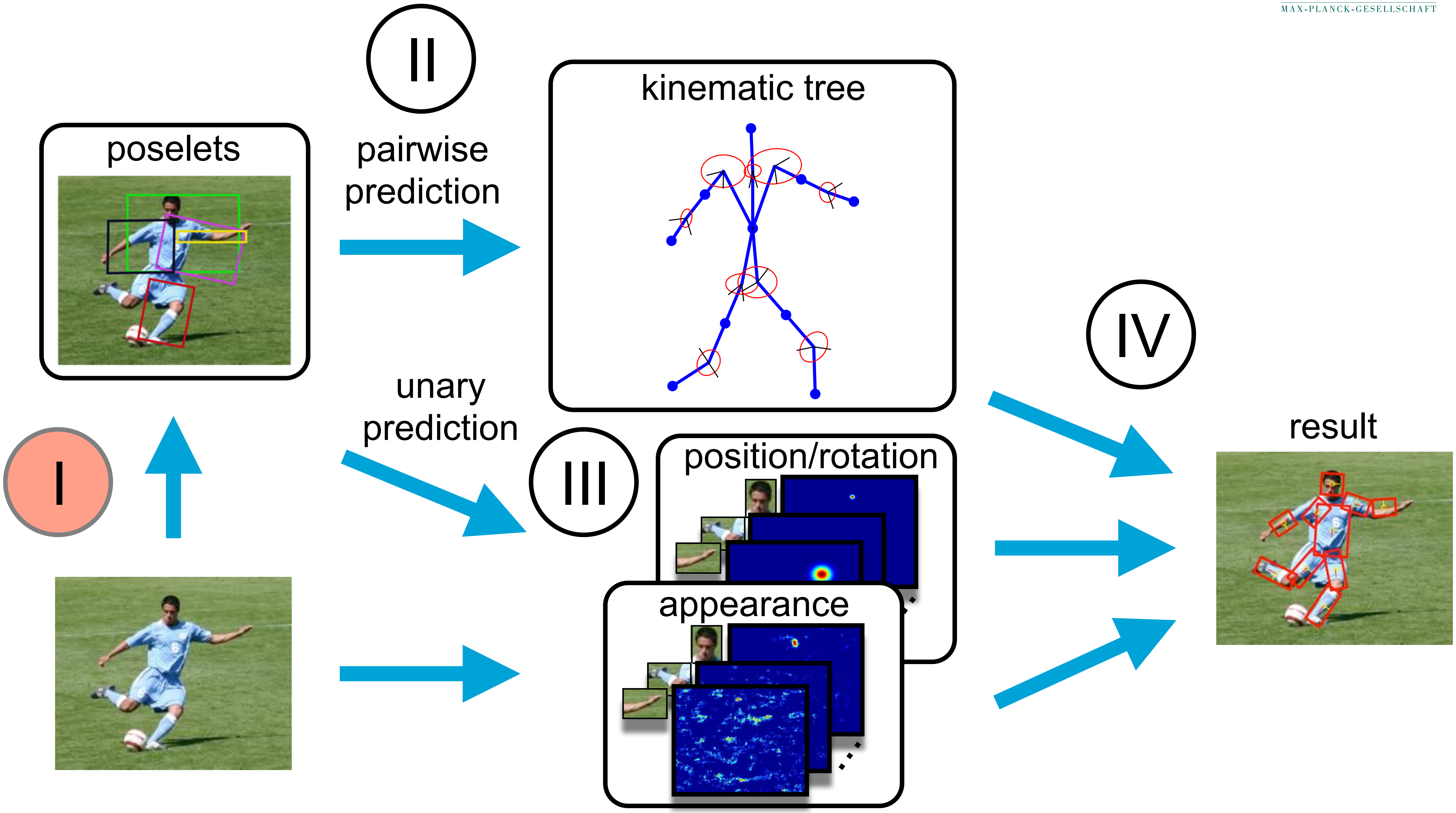
part

$$E(L;D) = \sum_{m=1}^{M} E^u(l_m;D) + \sum_{n\sim m} E^p(l_n,l_m;D) + w_p \sum_{m=1}^{M} E^{u,poselet}(l_m;D)$$

image

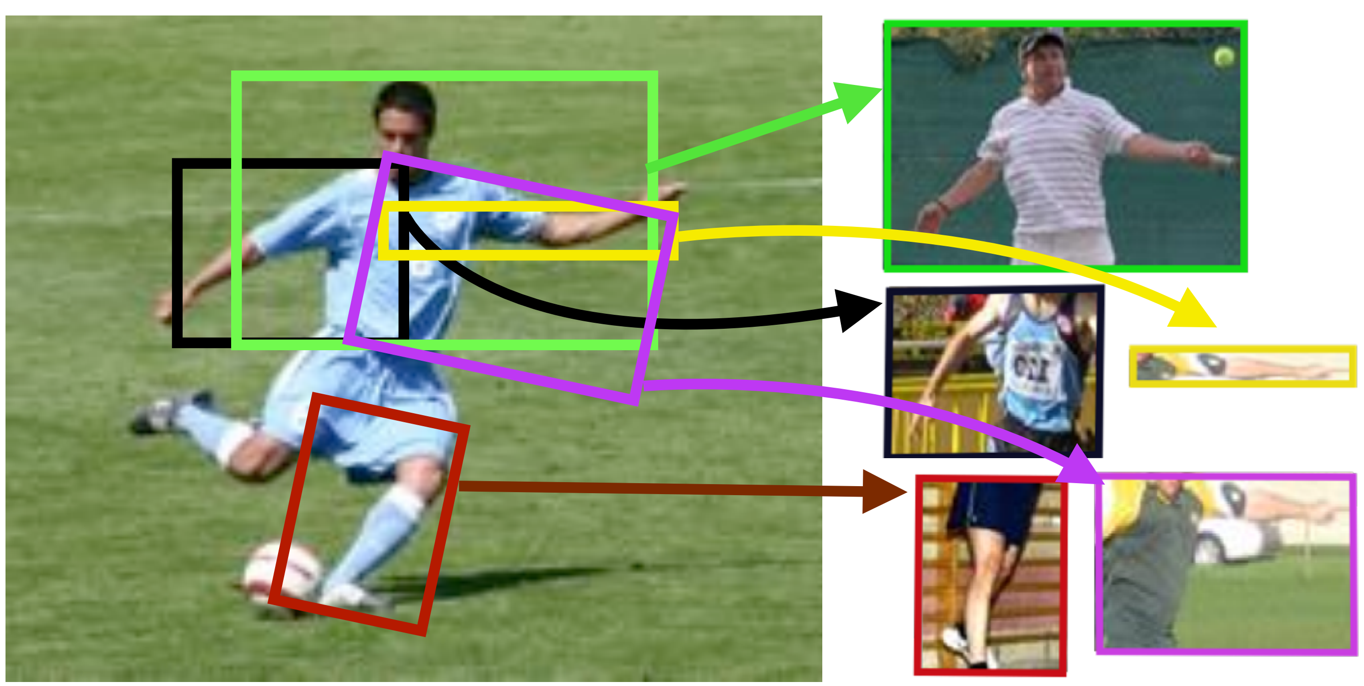


Poselet Conditioned Pictorial Structures



Poselets

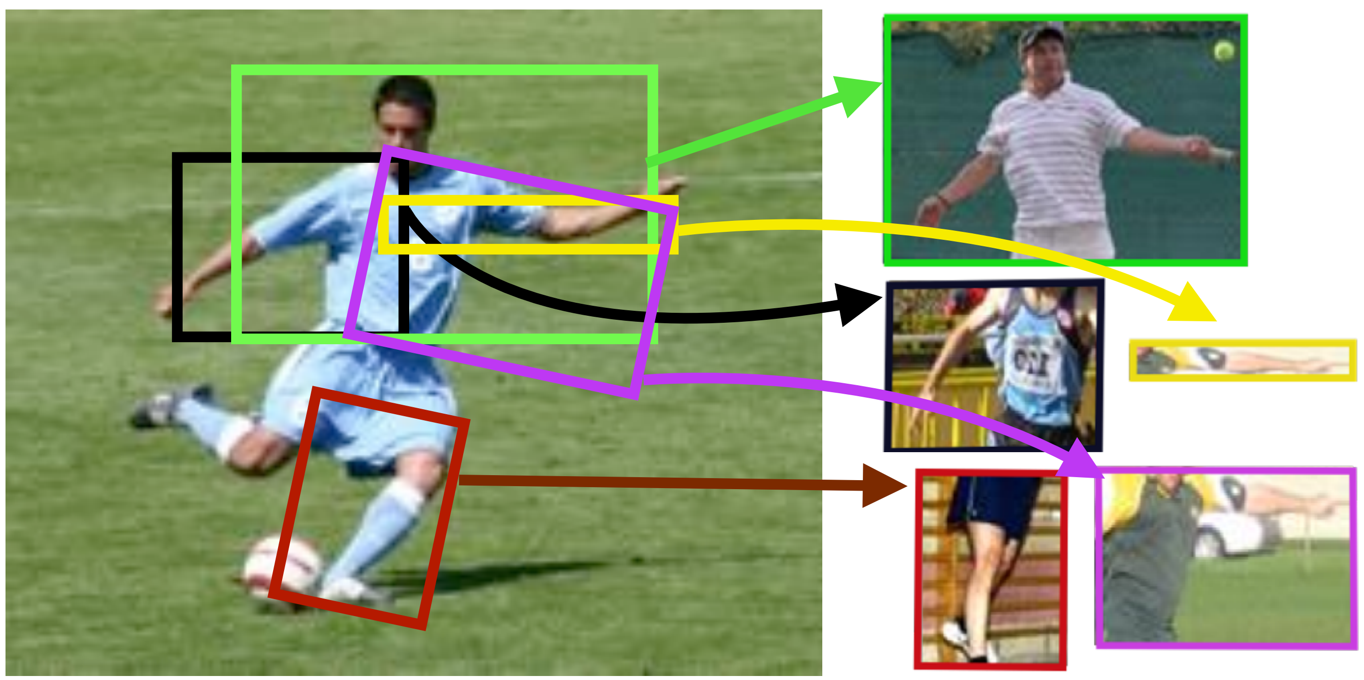
- Detect joint part configurations [Bourdev et al., ICCV'09]
  ➡ Capture ***non-adjacent part dependencies***

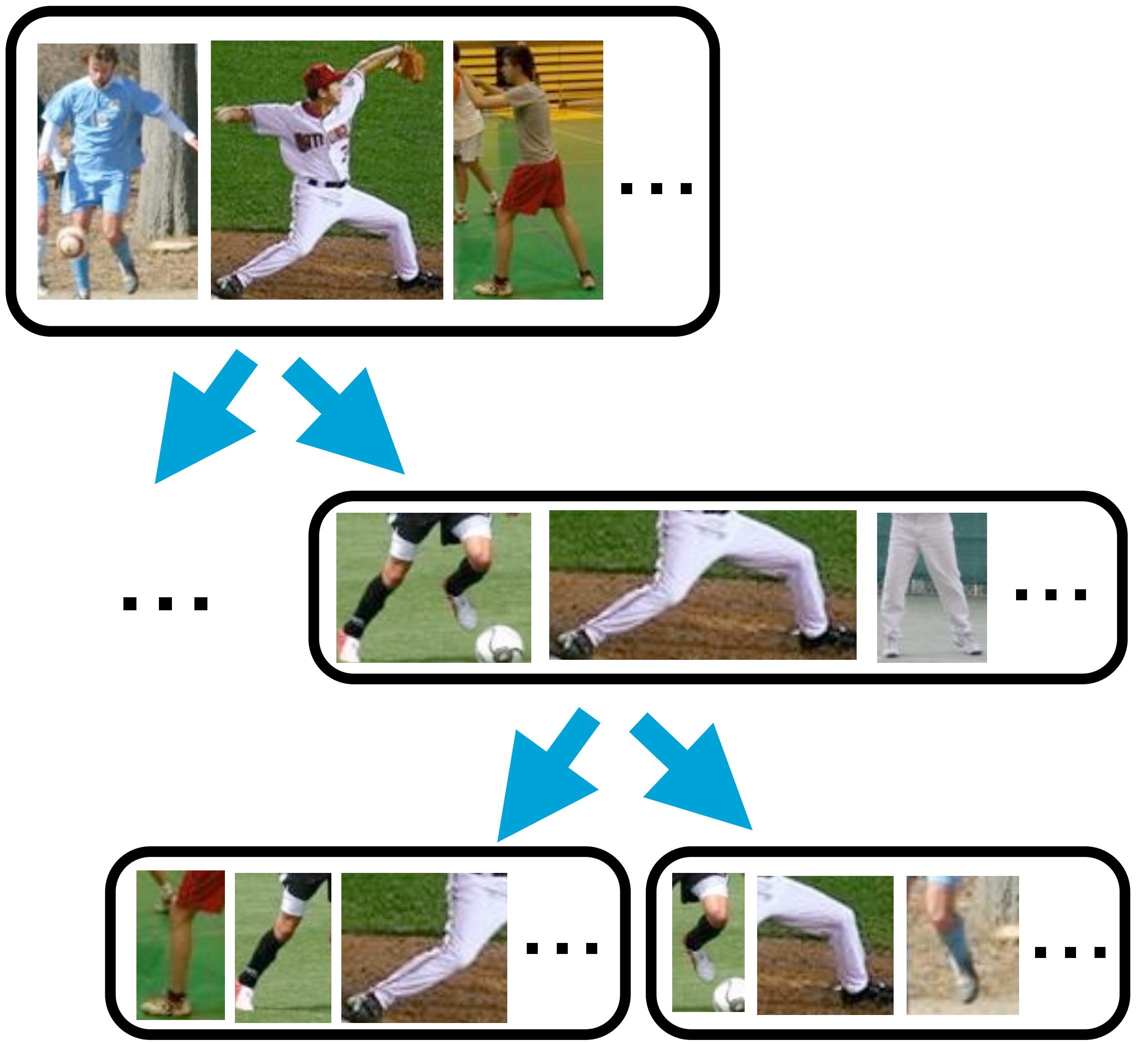


Top detections **Poselet cluster medoids**

Poselets

- Detect joint part configurations [Bourdev et al., ICCV'09]
  ➡ Capture ***non-adjacent part dependencies***
- Trained for different levels of abstraction

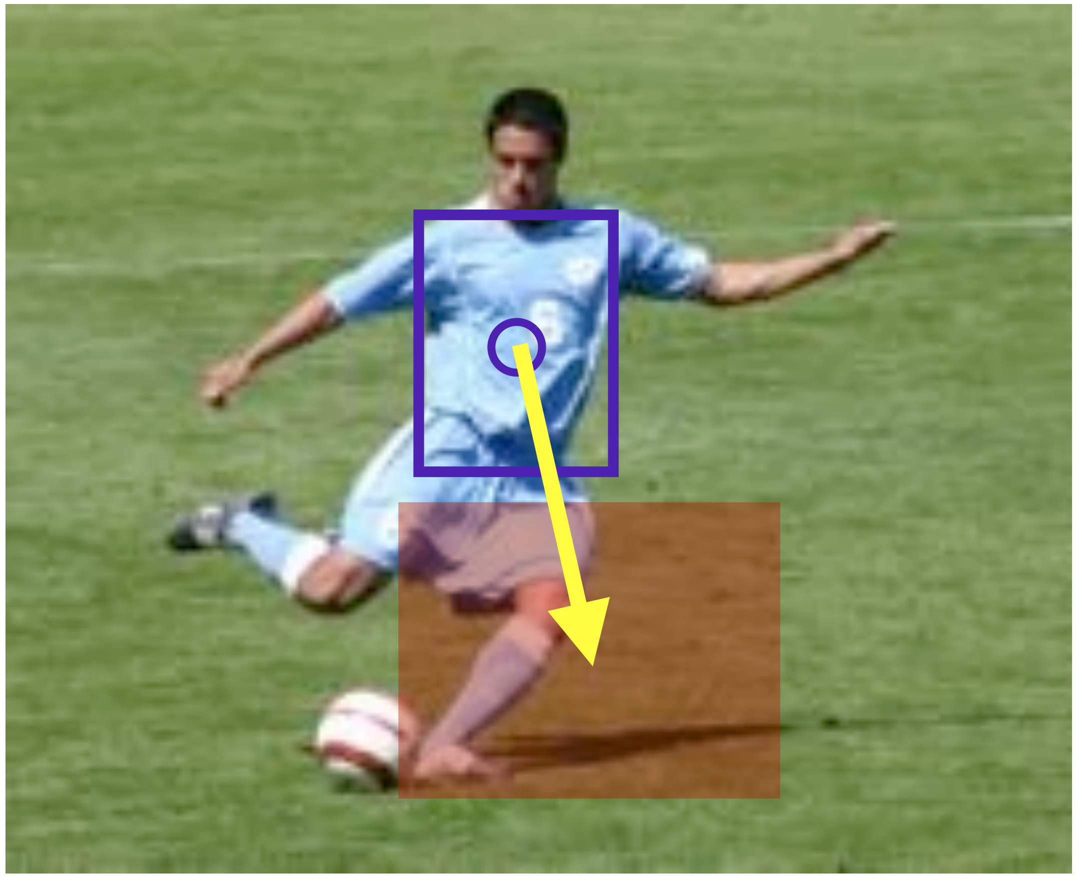




Top detections **Poselet cluster medoids**

Mid-level representation

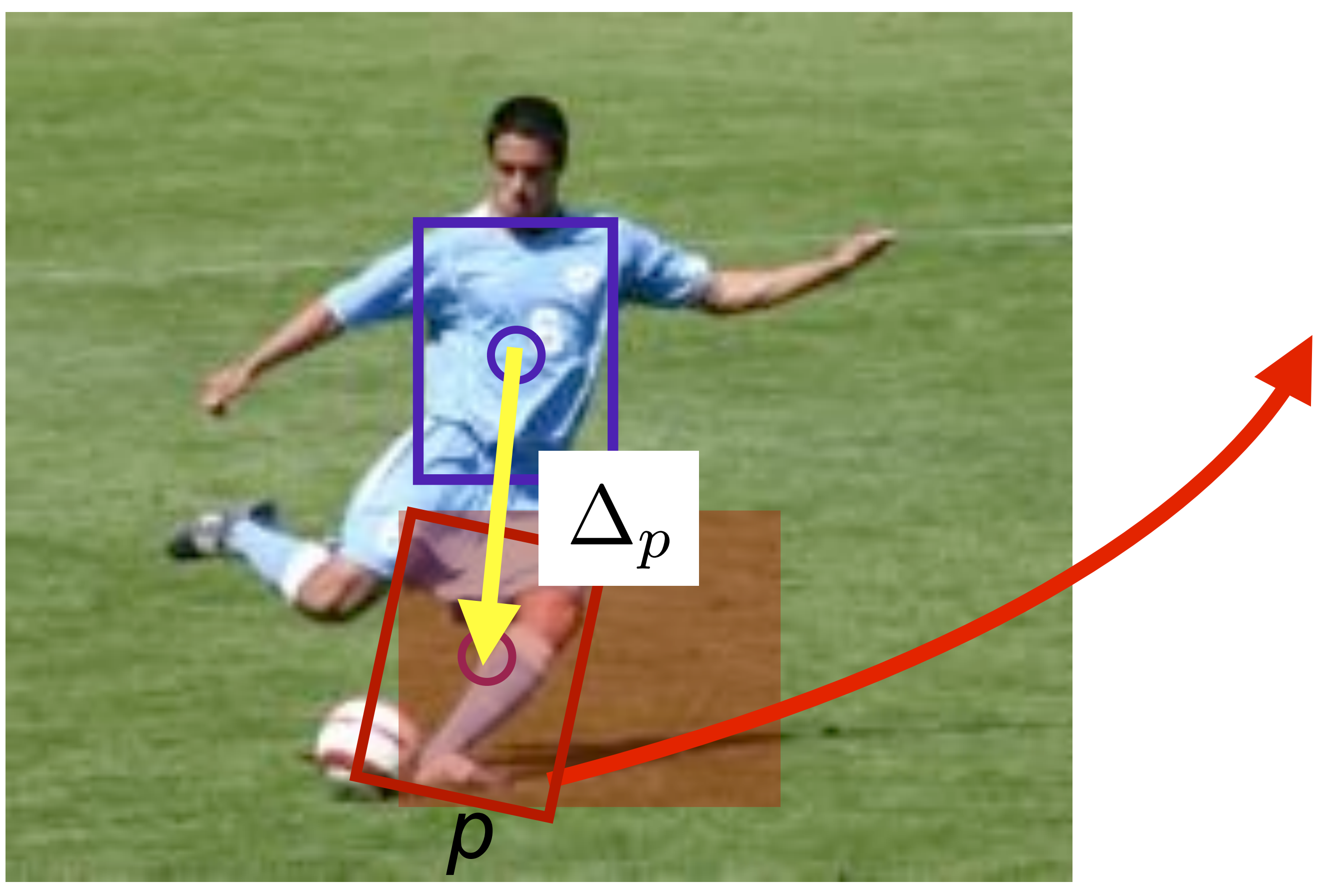
- Poselet responses vector
  - use torso detector [Pishchulin et al., CVPR'12]
  - ➡ collect top responses from local regions relative to torso



$$\begin{pmatrix} score_1 \\ \Delta_1 \\ . \\ . \\ score_p \\ \Delta_p \\ . \\ . \\ score_P \\ \Delta_P \end{pmatrix}$$

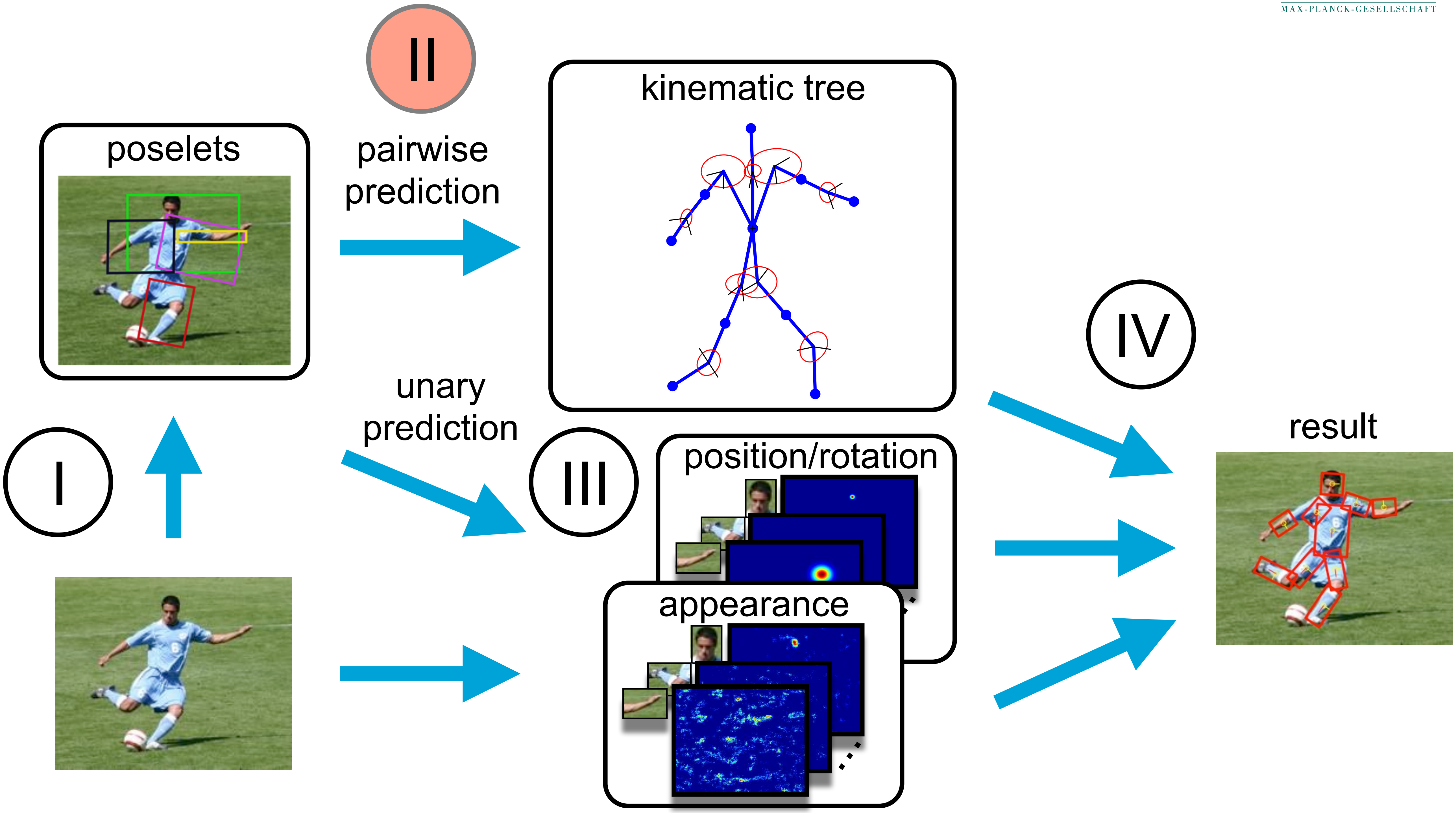
Mid-level representation

- Poselet responses vector
  - use torso detector [Pishchulin et al., CVPR'12]
  - ➡ collect top responses from local regions relative to torso



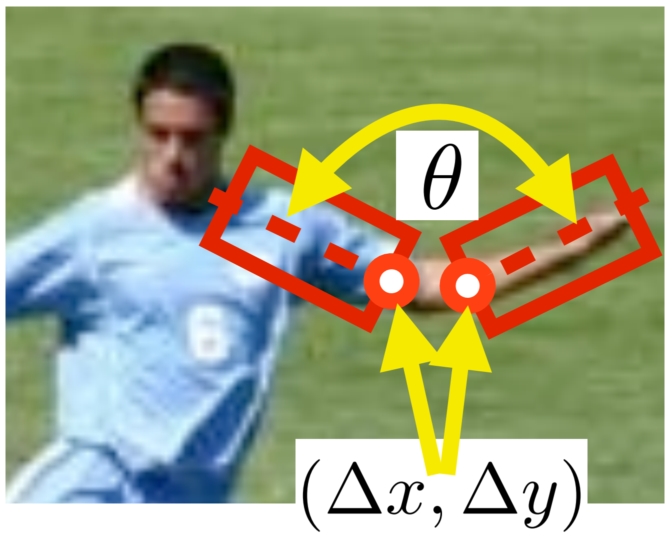
$$\begin{pmatrix} score_1 \\ \Delta_1 \\ \cdot \\ \cdot \\ score_p \\ \Delta_p \\ \cdot \\ \cdot \\ score_P \\ \Delta_P \end{pmatrix}$$

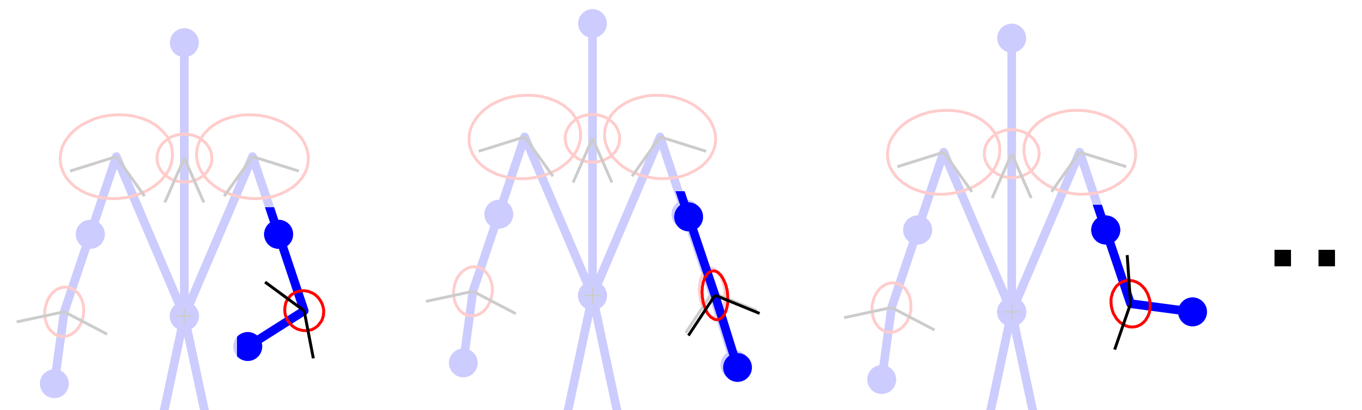
Poselet Conditioned Pictorial Structures



Learning pairwise parameters

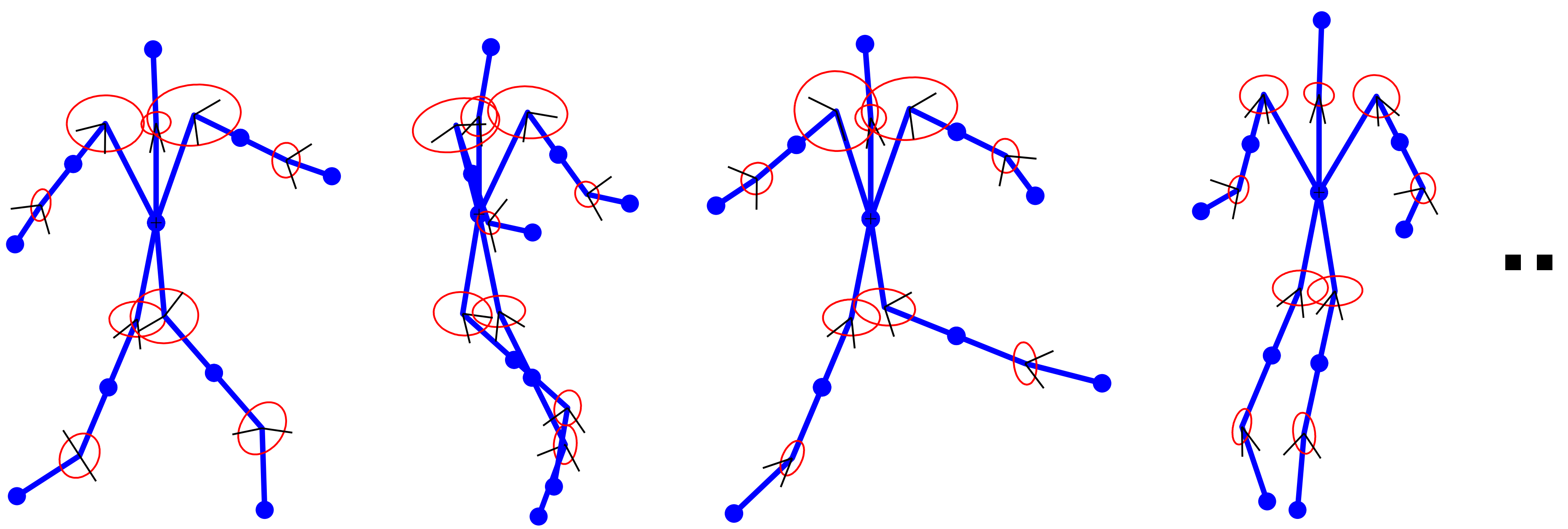
- Pairwise: relative position and rotation





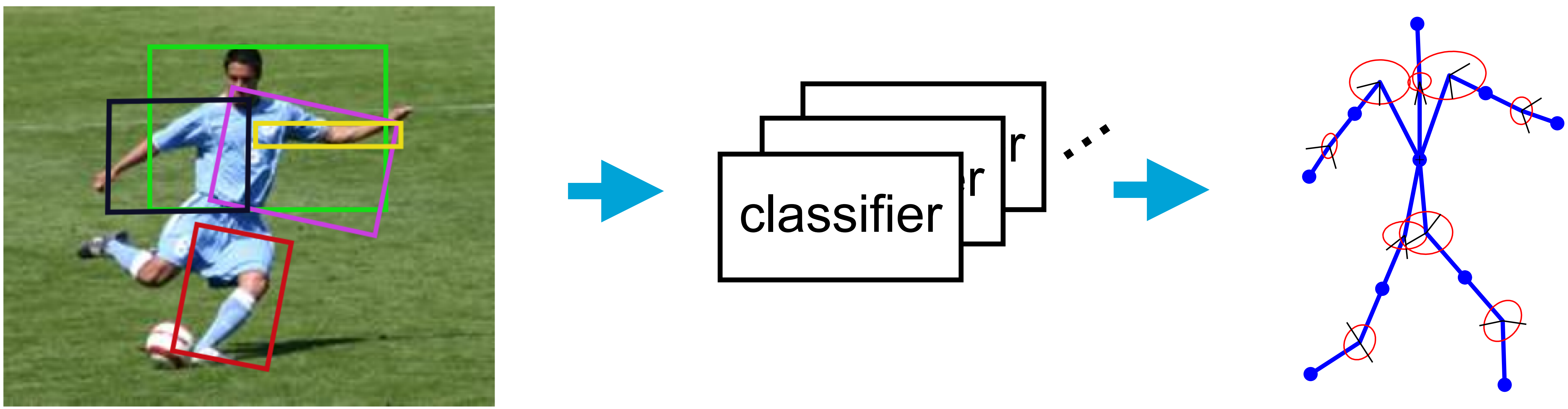
Mixture components

- Allows to model exponentially many trees



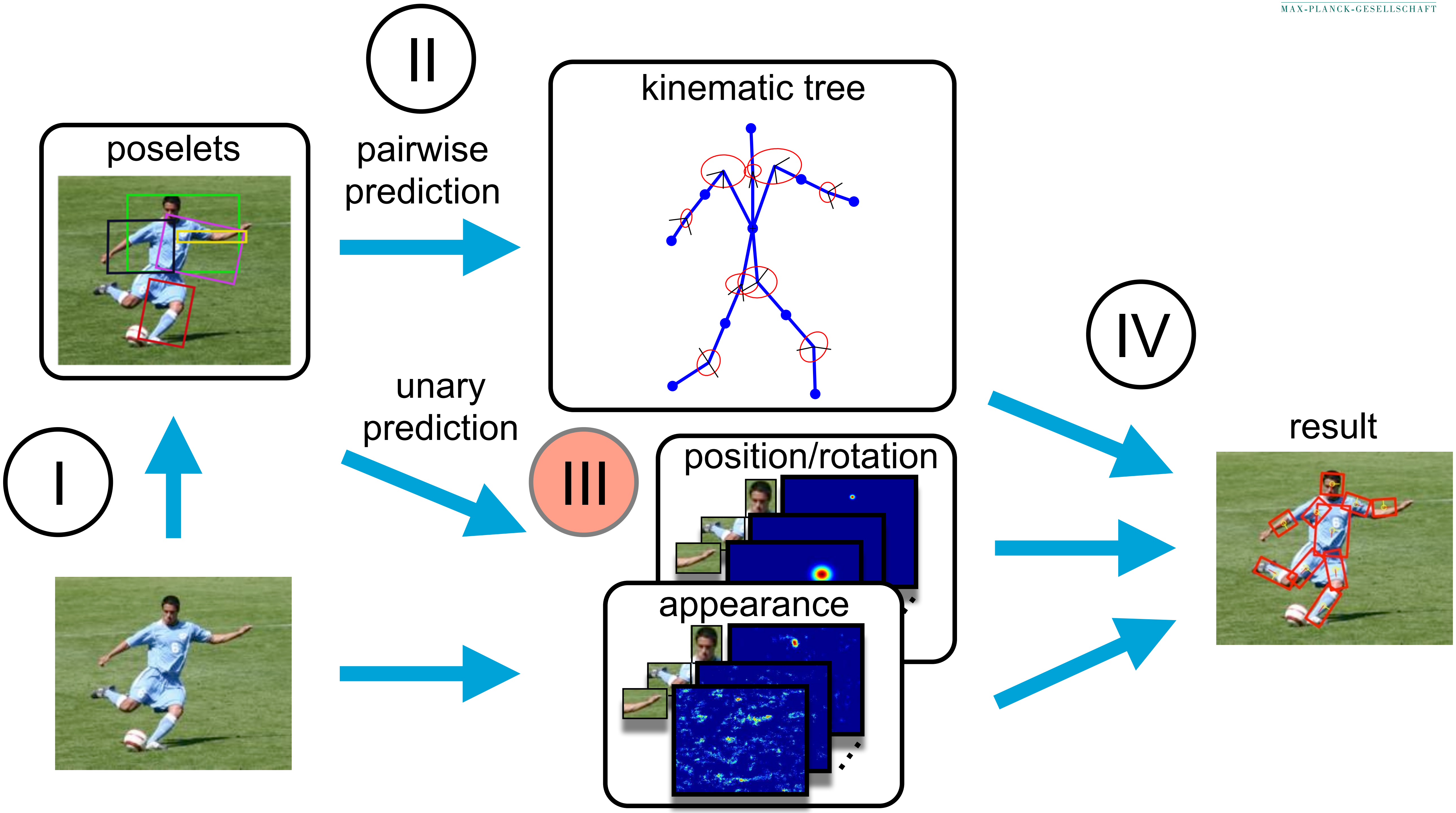
Predicting pairwise parameters

- Test time: how to predict a tree?
  - use multi-class classifier on poselet responses for each pairwise



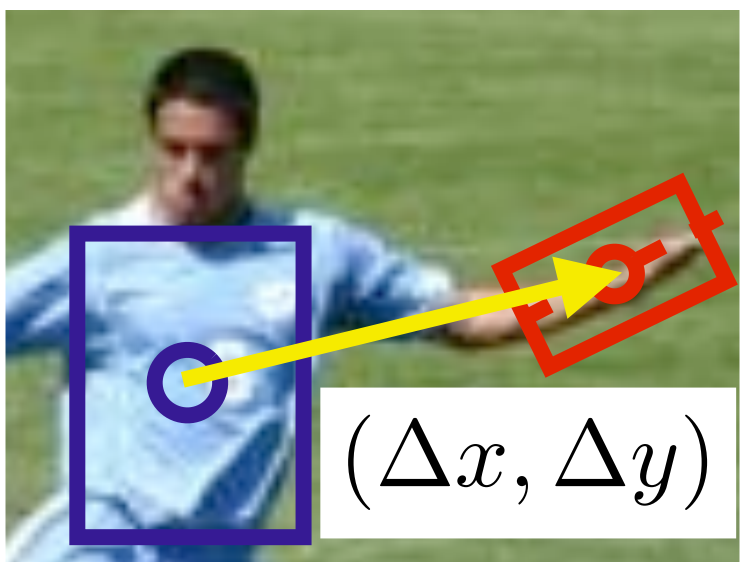
- Predict tree ***before*** pose inference
  - ✓ ***exact and efficient inference***

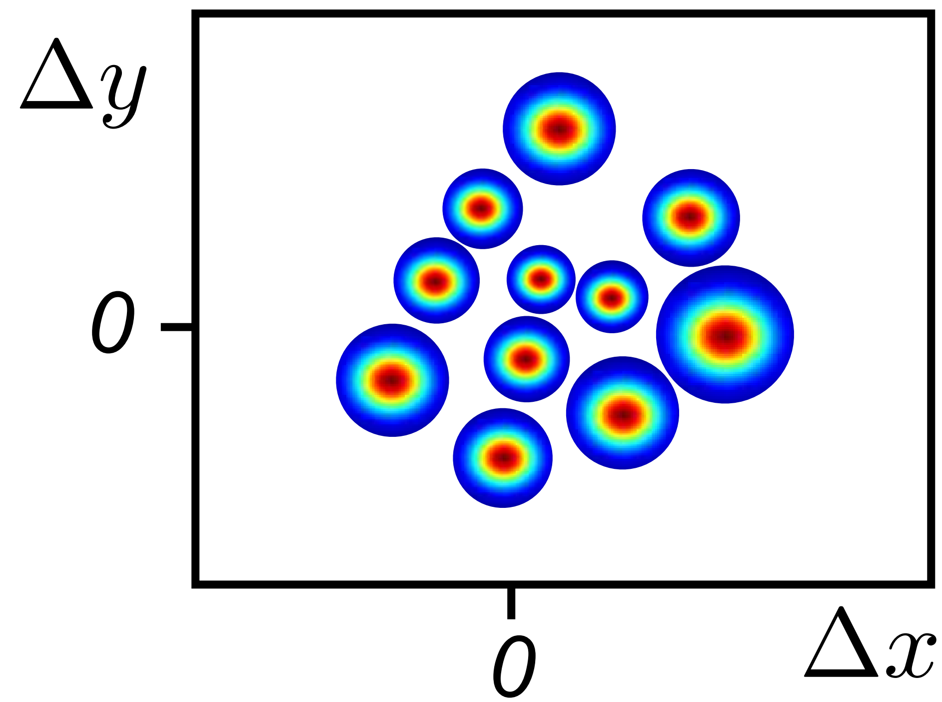
Poselet Conditioned Pictorial Structures



Learning unary parameters

- Unaries: part relative position and absolute rotation
- Position: cluster relative position into mixture components



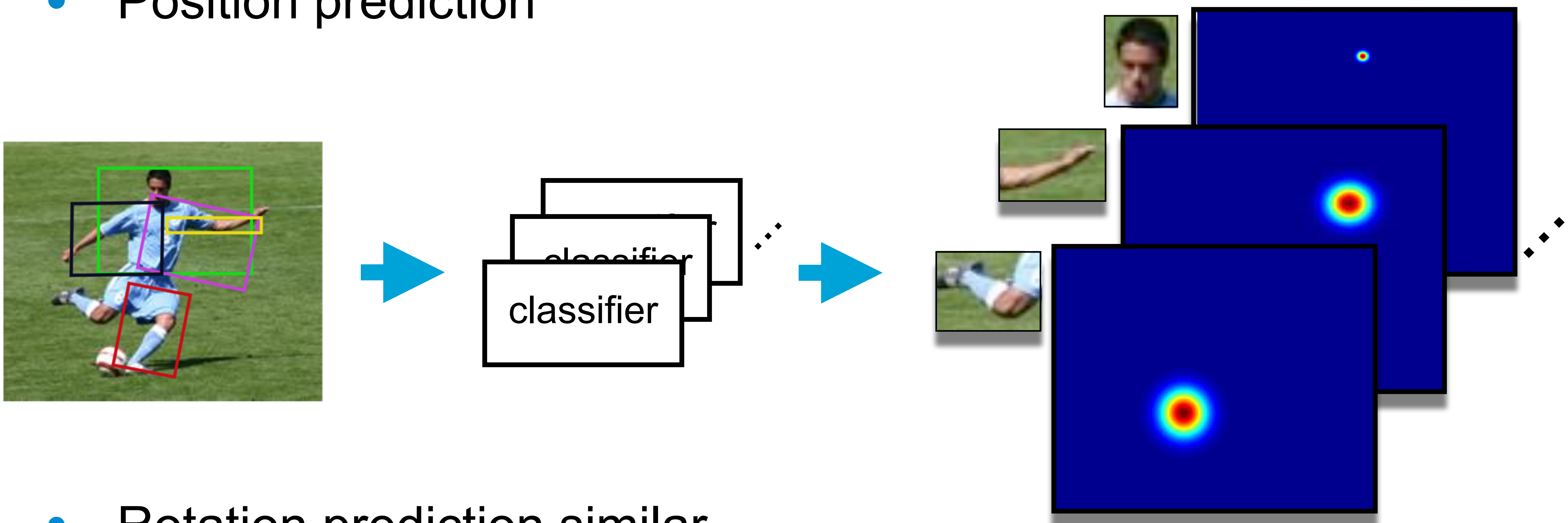


Mixture components

- Rotation: binning of absolute rotation angles

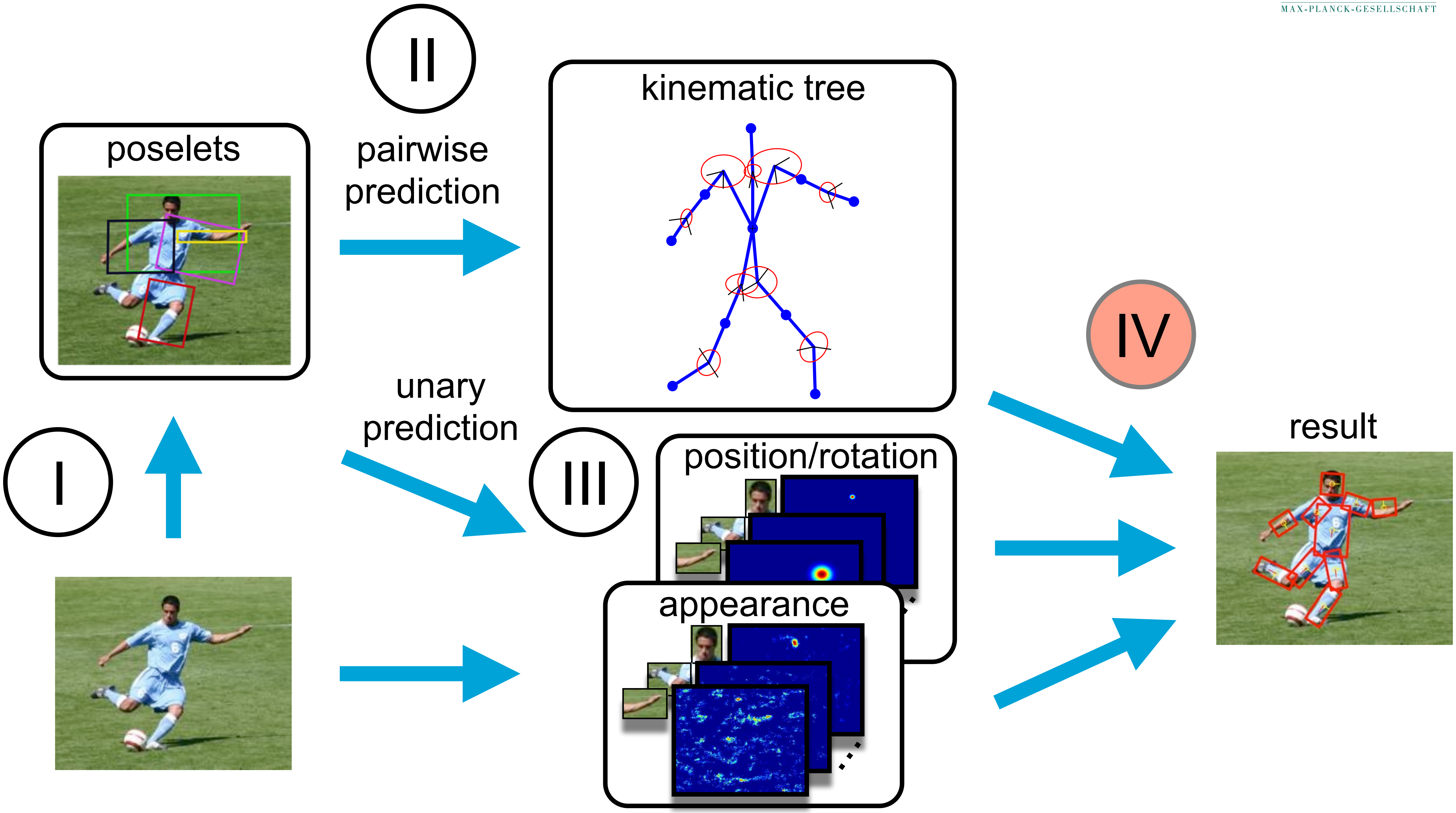
Predicting unary parameters

- Test time: use multi-class classifier for each unary
- Position prediction



- Rotation prediction similar

Poselet Conditioned Pictorial Structures



Experimental evaluation

- Leeds Sports Poses (LSP) [Johnson&Everingham, BMVC'10]
  - 1,000 train, 1,000 test images
  - set parameters using half of training set
  - ***observer-centric*** annotations for testing [Eichner&Ferrari, ACCV'12]

Experimental evaluation

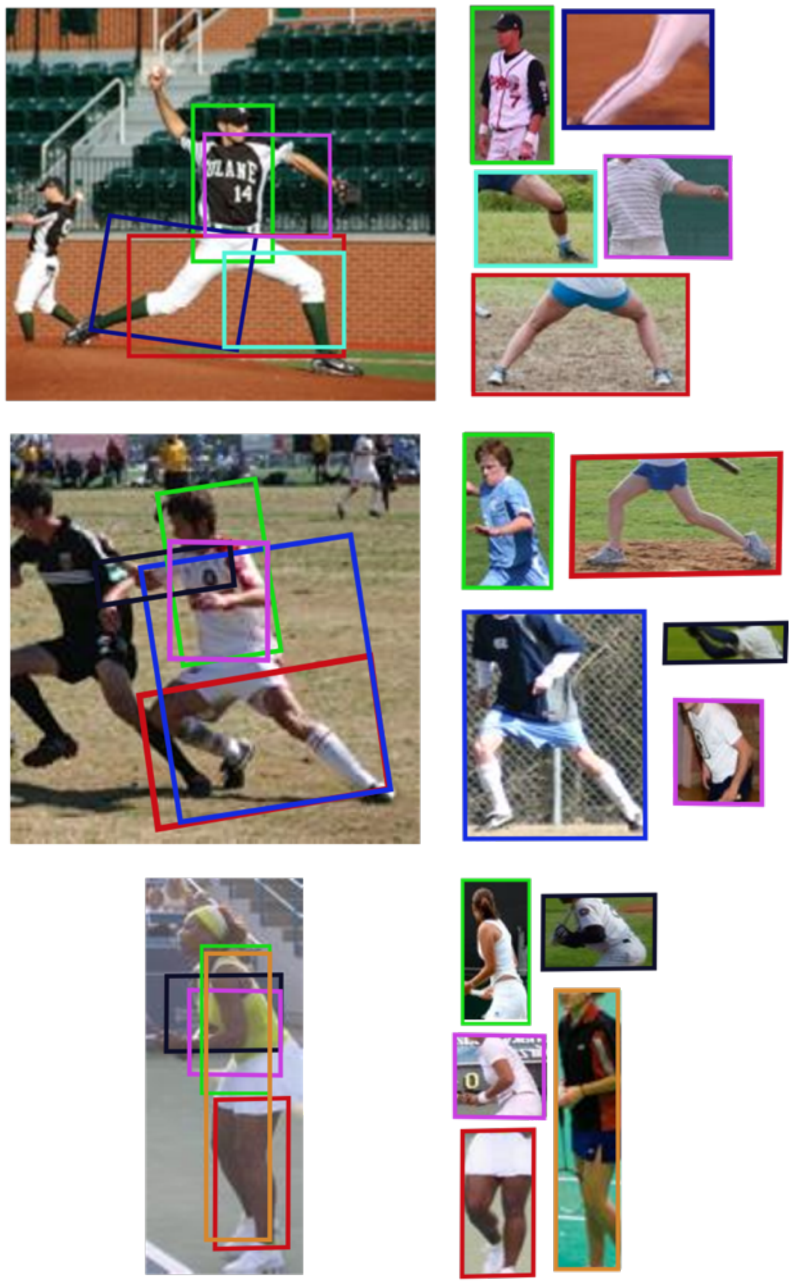
- Leeds Sports Poses (LSP) [Johnson&Everingham, BMVC'10]
  - 1,000 train, 1,000 test images
  - set parameters using half of training set
  - ***observer-centric*** annotations for testing [Eichner&Ferrari, ACCV'12]

- Image Parse (IP) [Ramanan, NIPS'06]
- UIUC People [Tran&Forsyth, ECCV'10]

➡ Results in the paper

Results - qualitative evaluation

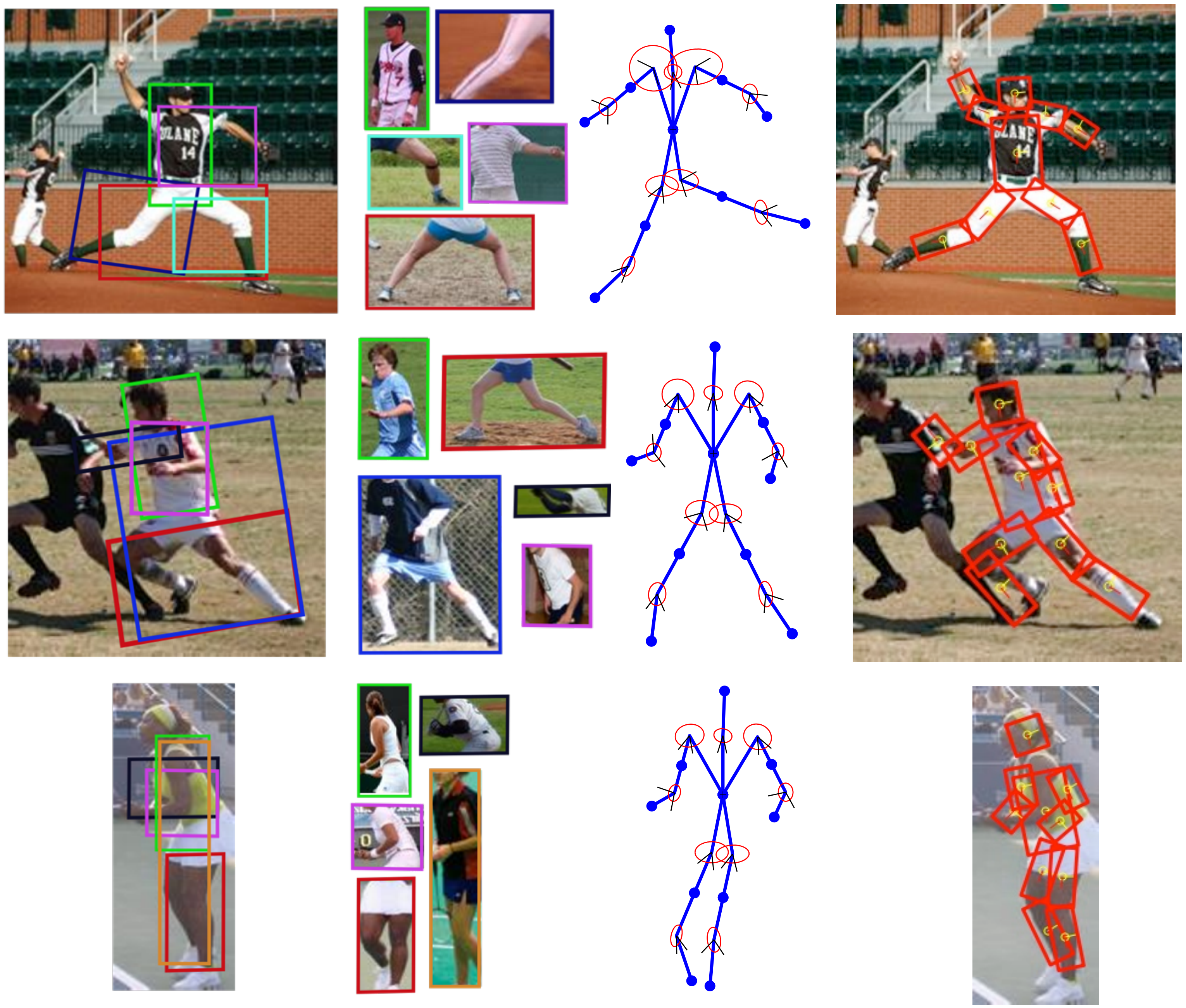
Our method



Top poselet detections **Cluster medoids**

Results - qualitative evaluation

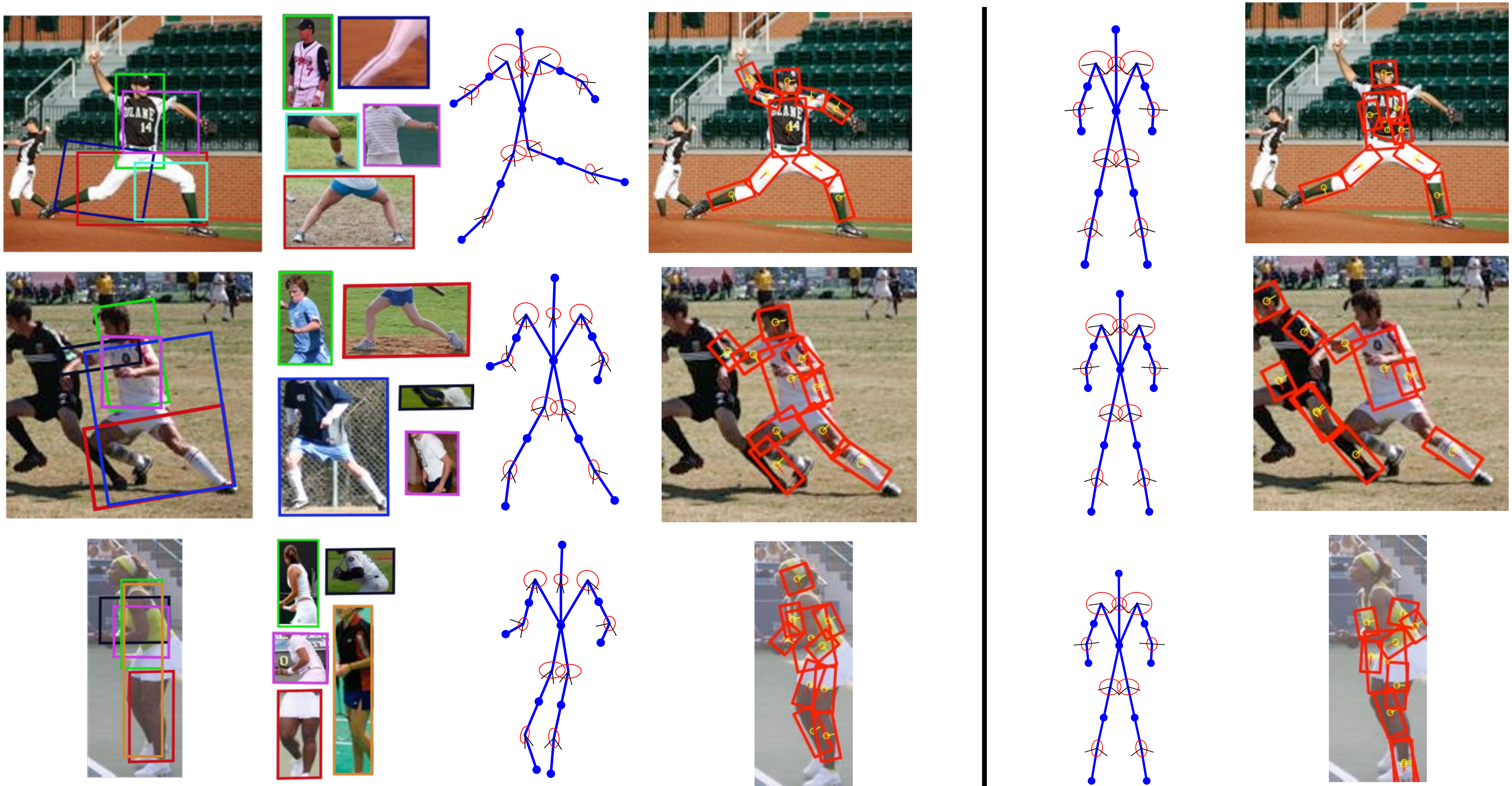
Our method



Top poselet detections | Cluster medoids | Prediction | Result using prediction

Results - qualitative evaluation

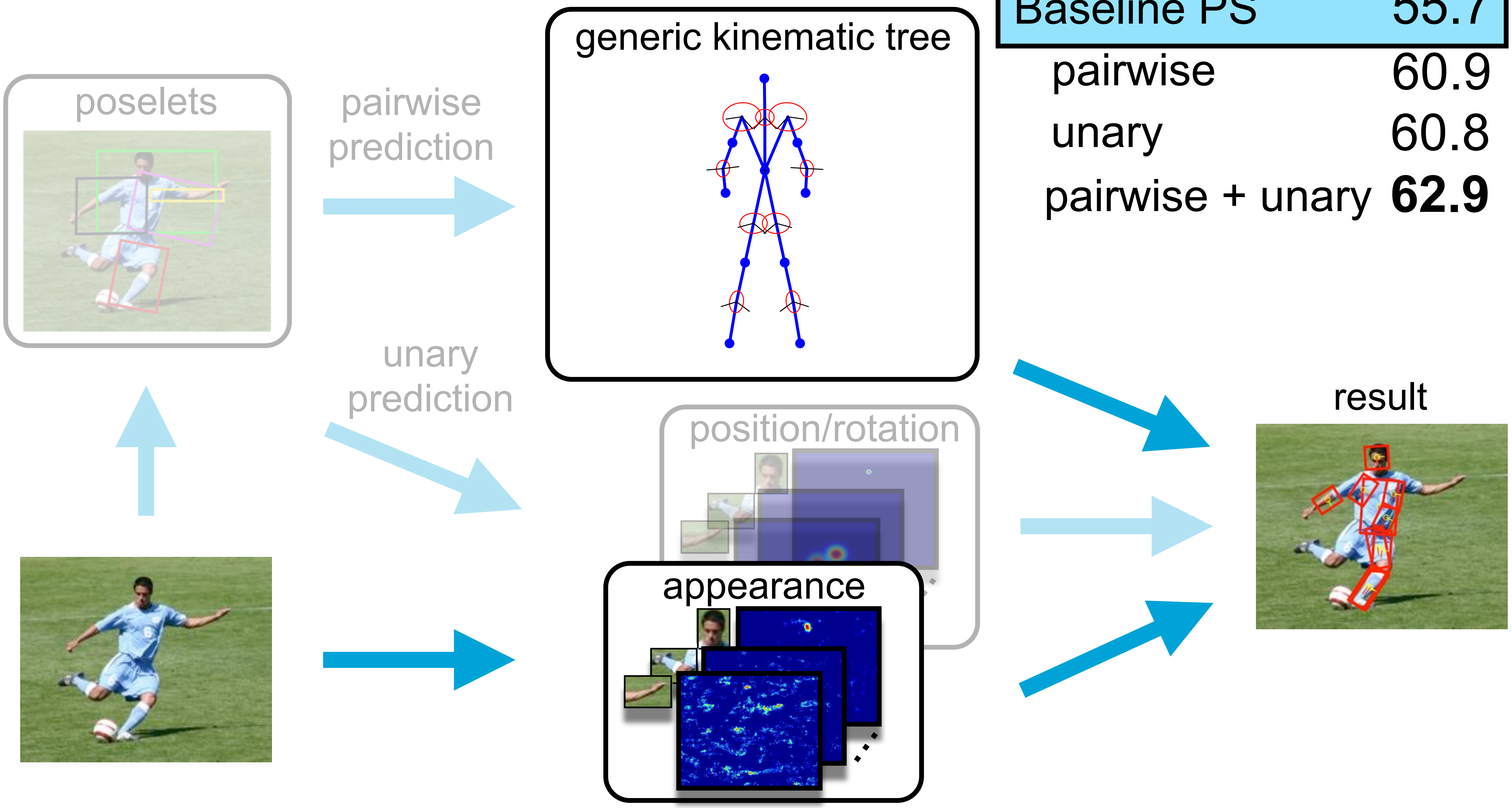
Our method | **Baseline PS**



Top poselet detections | Cluster medoids | Prediction | Result using prediction | Generic tree | Result using generic tree

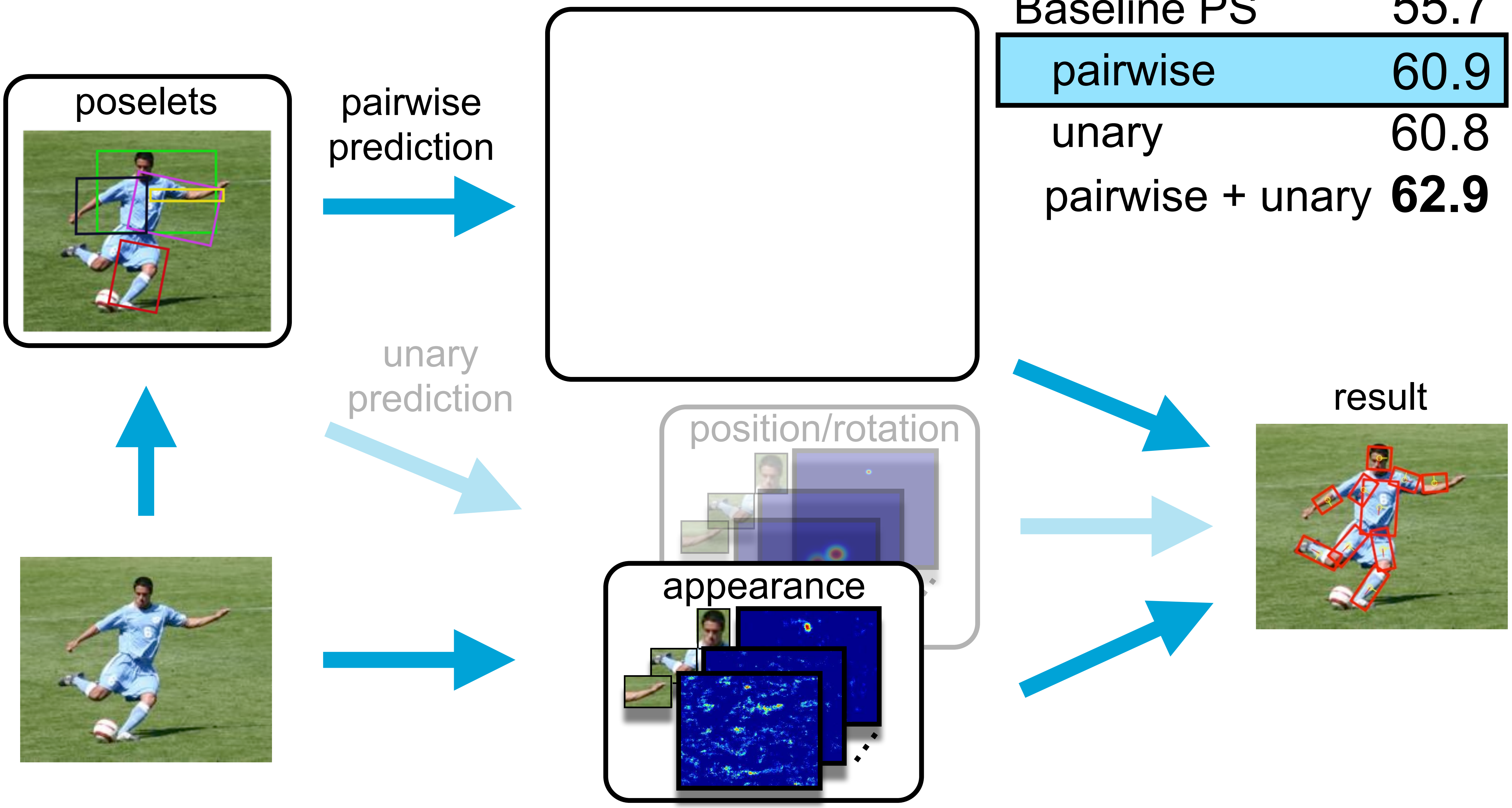
Results (PCP)

Baseline PS	55.7
pairwise	60.9
unary	60.8
pairwise + unary	**62.9**



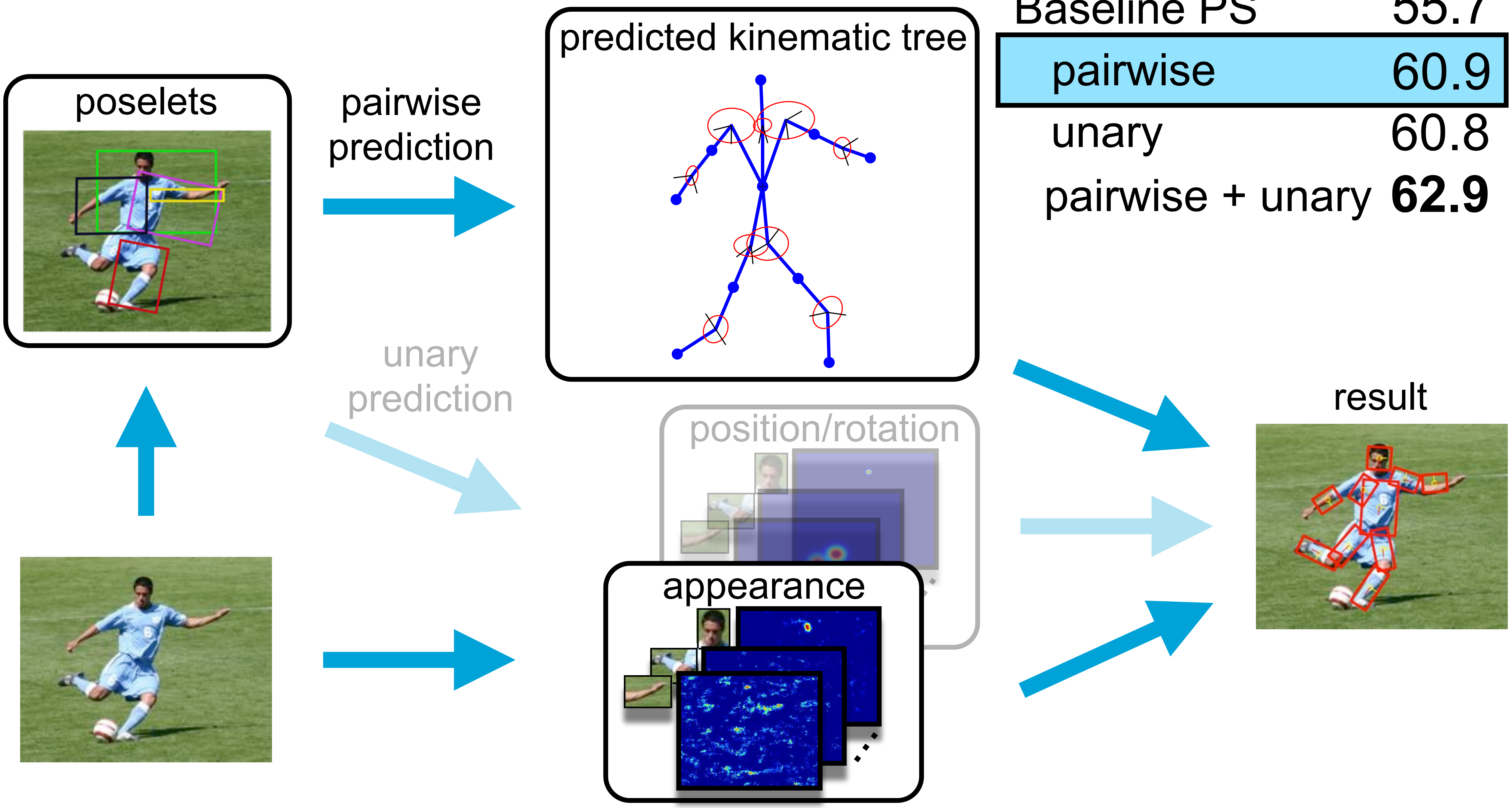
Results (PCP)

Baseline PS	55.7
pairwise	60.9
unary	60.8
pairwise + unary	**62.9**



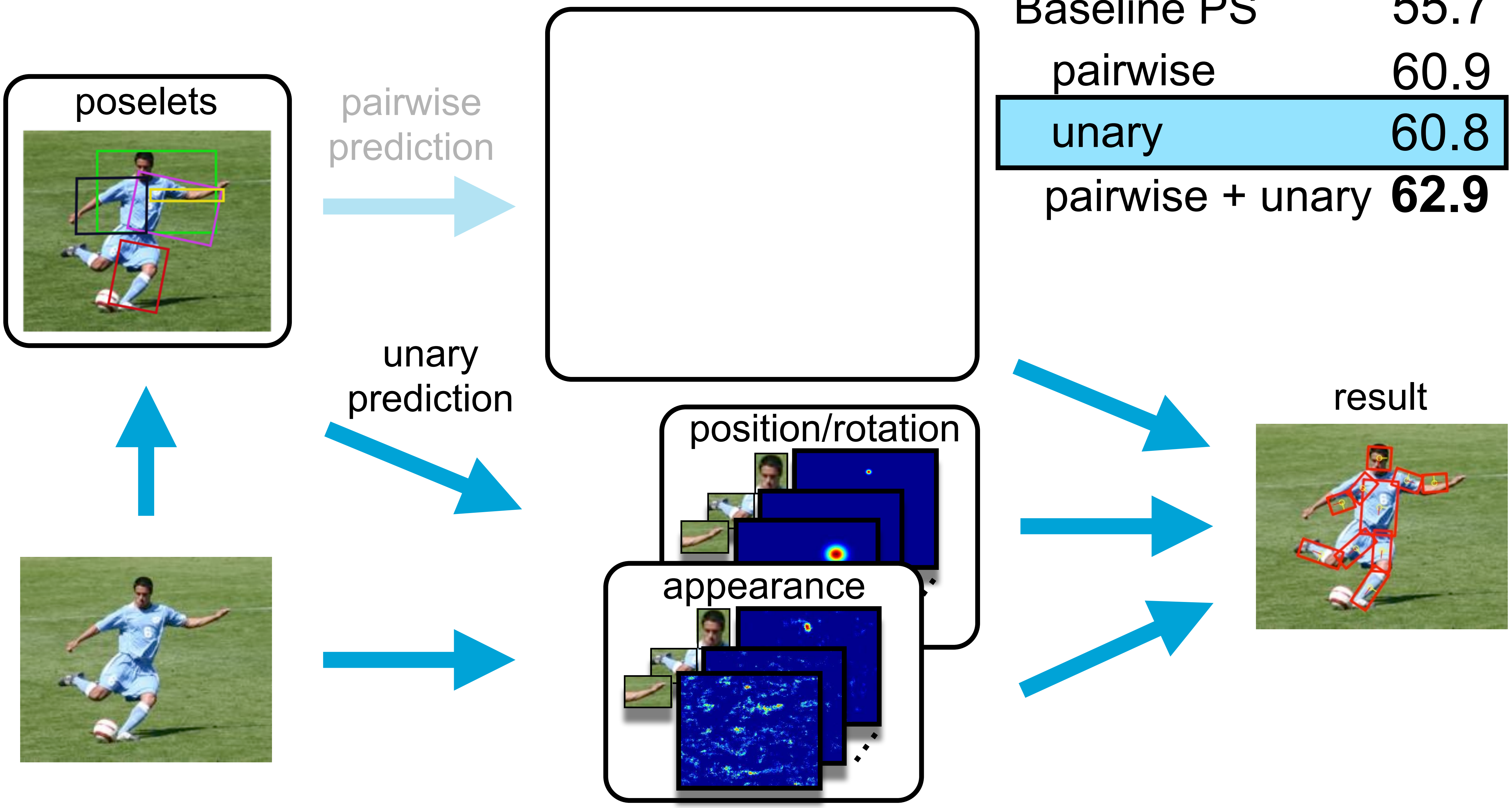
Results (PCP)

Baseline PS	55.7
pairwise	60.9
unary	60.8
pairwise + unary	**62.9**



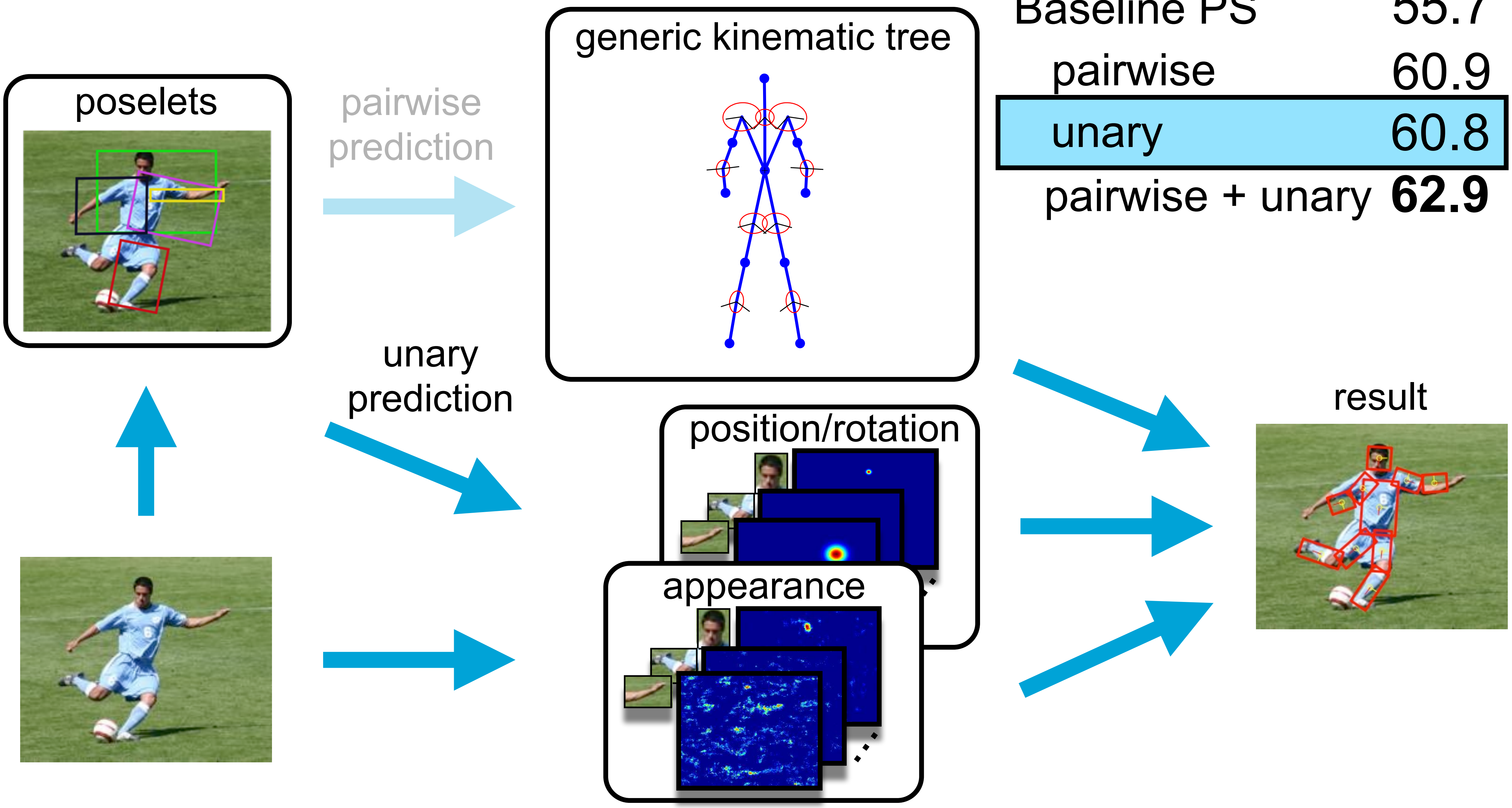
Results (PCP)

Baseline PS	55.7
pairwise	60.9
unary	60.8
pairwise + unary	**62.9**



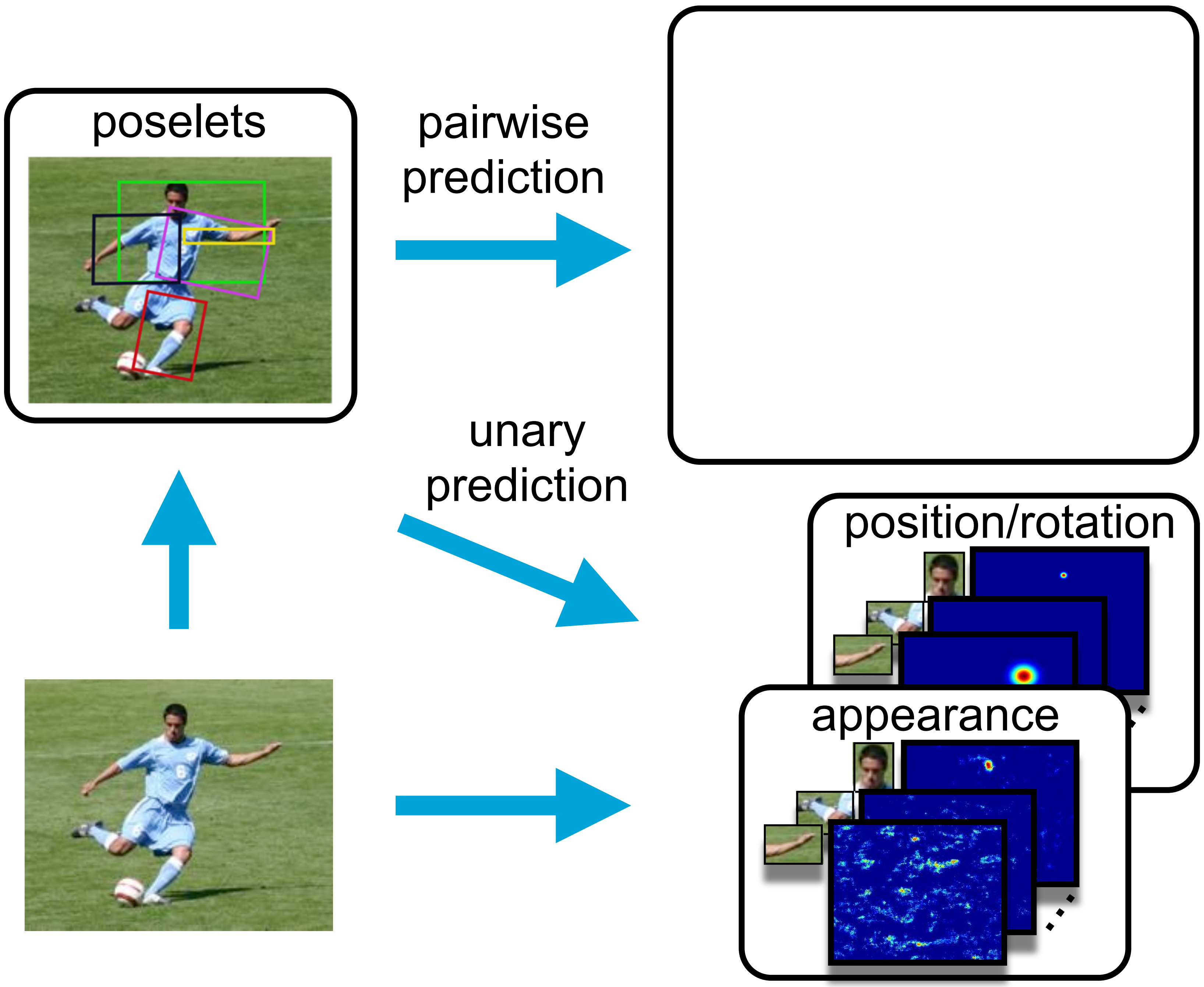
Results (PCP)

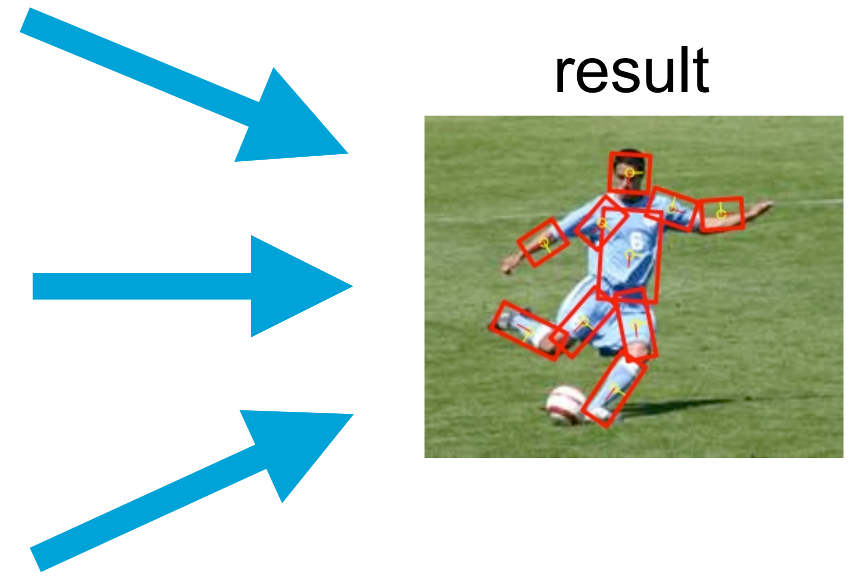
Baseline PS	55.7
pairwise	60.9
unary	60.8
pairwise + unary	**62.9**



Results (PCP)

Baseline PS	55.7
pairwise	60.9
unary	60.8
pairwise + unary	**62.9**

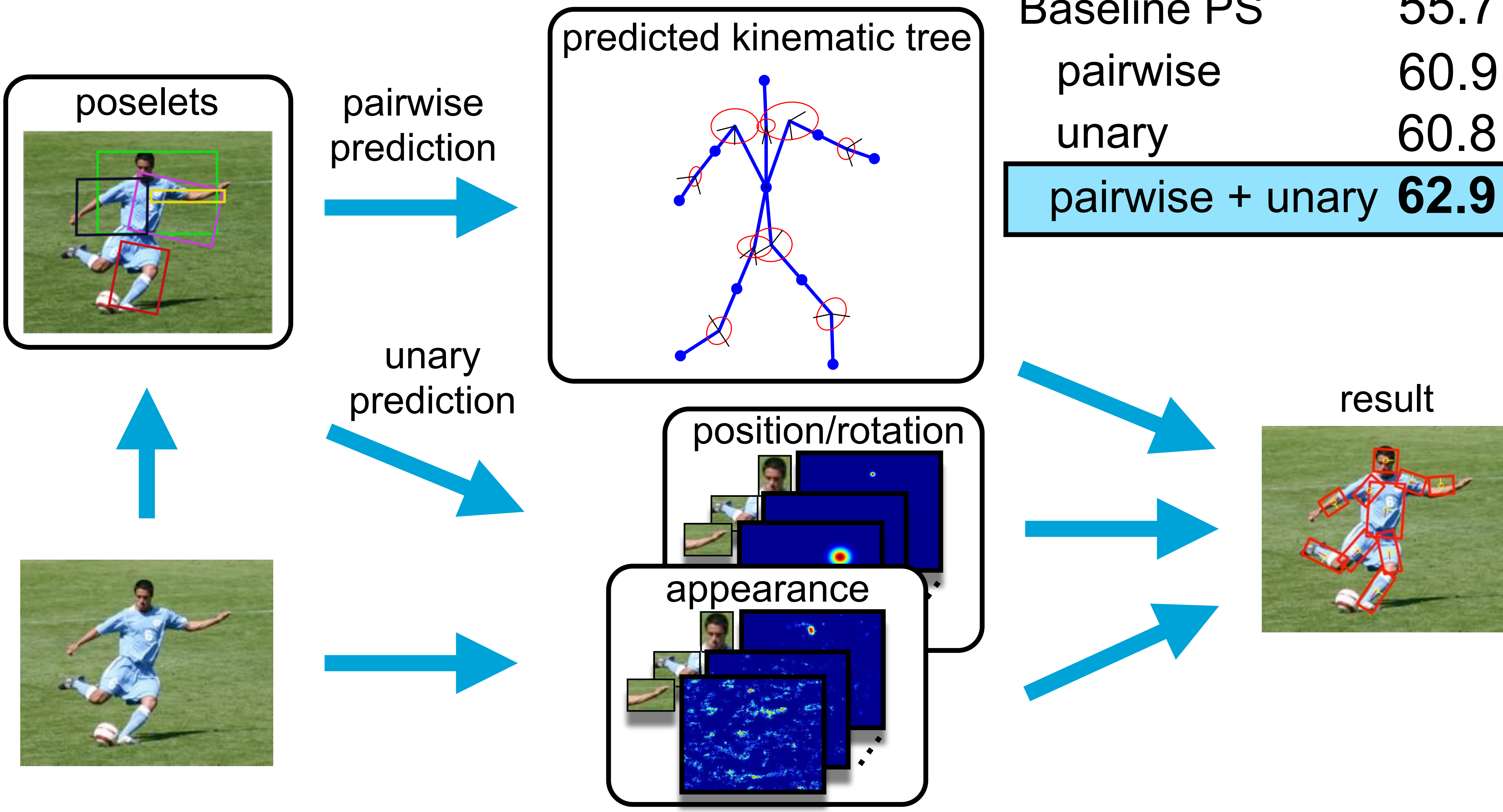




✓ **+7.2% improvement compared to baseline**

Results (PCP)

Baseline PS	55.7
pairwise	60.9
unary	60.8
pairwise + unary	**62.9**



✓ **+7.2% improvement compared to baseline**

Results - comparison to state of the art

Method	Torso	Upper leg	Lower leg	Upper arm	Forearm	Head	*Total*
ours	**87.5**	**75.7**	68.0	54.2	33.9	78.1	*62.9*
Yang&Ramanan, CVPR'11	84.1	69.5	65.6	52.5	35.9	77.1	*60.8*
Eichner&Ferrari, ACCV'12	84.9	73.1	**68.3**	**55.8**	**38.6**	**80.1**	***63.7***

Results - comparison to state of the art

Method	Torso	Upper leg	Lower leg	Upper arm	Forearm	Head	*Total*
ours	**87.5**	**75.7**	68.0	54.2	33.9	78.1	*62.9*
Yang&Ramanan, CVPR'11	84.1	69.5	65.6	52.5	35.9	77.1	*60.8*
Eichner&Ferrari, ACCV'12	84.9	73.1	**68.3**	**55.8**	**38.6**	**80.1**	***63.7***

+2.1%

- ✓ **improve over Yang&Ramanan for all parts but lower arms**

Results - comparison to state of the art

Method	Torso	Upper leg	Lower leg	Upper arm	Forearm	Head	*Total*
ours	**87.5**	**75.7**	68.0	54.2	33.9	78.1	*62.9*
Yang&Ramanan, CVPR'11	84.1	69.5	65.6	52.5	35.9	77.1	*60.8*
Eichner&Ferrari, ACCV'12	84.9	73.1	**68.3**	**55.8**	**38.6**	**80.1**	***63.7***

-0.8%

- ✓ **improve over Yang&Ramanan for all parts but lower arms**
- \- **loose mostly on lower arms compared to Eichner&Ferrari**

Results - comparison to state of the art

Method	Torso	Upper leg	Lower leg	Upper arm	Forearm	Head	*Total*
ours	**87.5**	**75.7**	68.0	54.2	33.9	78.1	*62.9*
Yang&Ramanan, CVPR’11	84.1	69.5	65.6	52.5	35.9	77.1	*60.8*
Eichner&Ferrari, ACCV’12	84.9	73.1	**68.3**	**55.8**	**38.6**	**80.1**	***63.7***

-0.8%

- ✓ **improve over Yang&Ramanan for all parts but lower arms**
- \- **loose mostly on lower arms compared to Eichner&Ferrari**
- ✓ **competitors are complementary as use better local appearance**

Results - using oracle

Setting	PCP, [%]
prediction	62.9
oracle	**88.1**

+25.2%

- ✓ **using oracle dramatically improves the results**

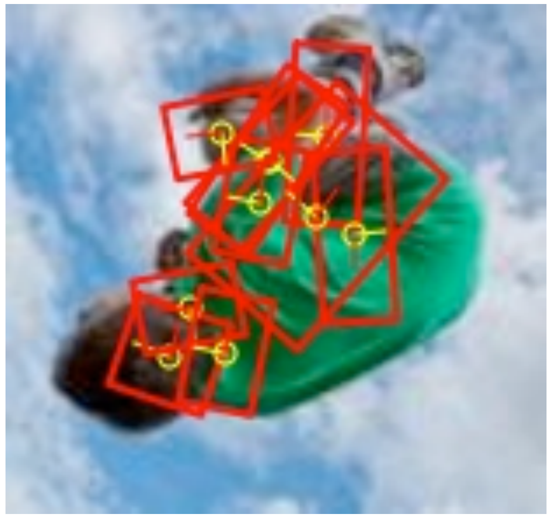
Results - using oracle

Setting	PCP, [%]
prediction	62.9
oracle	**88.1**

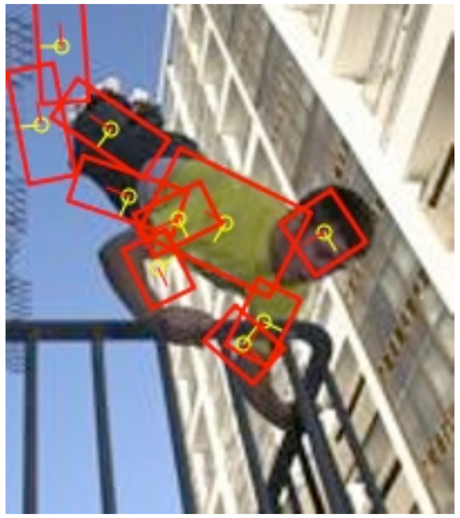
+25.2%

✓ **using oracle dramatically improves the results**

- **What is missing to get 100%?**



Self-occlusion

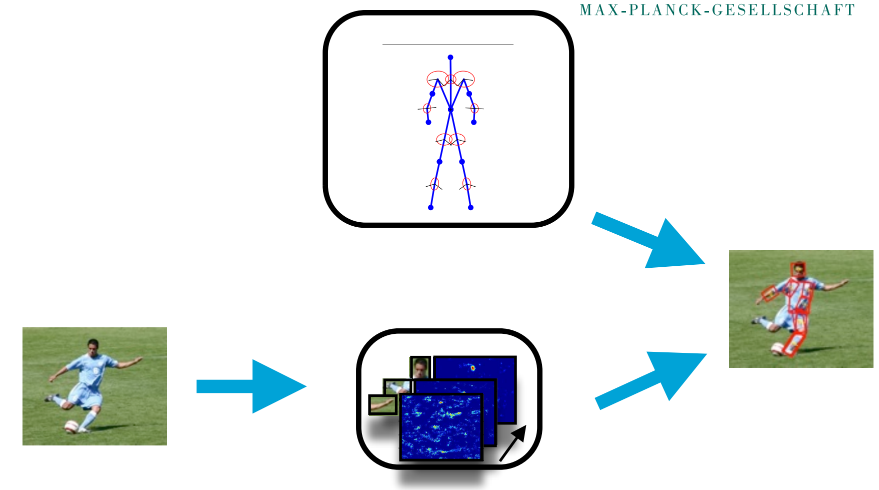


Foreshortening

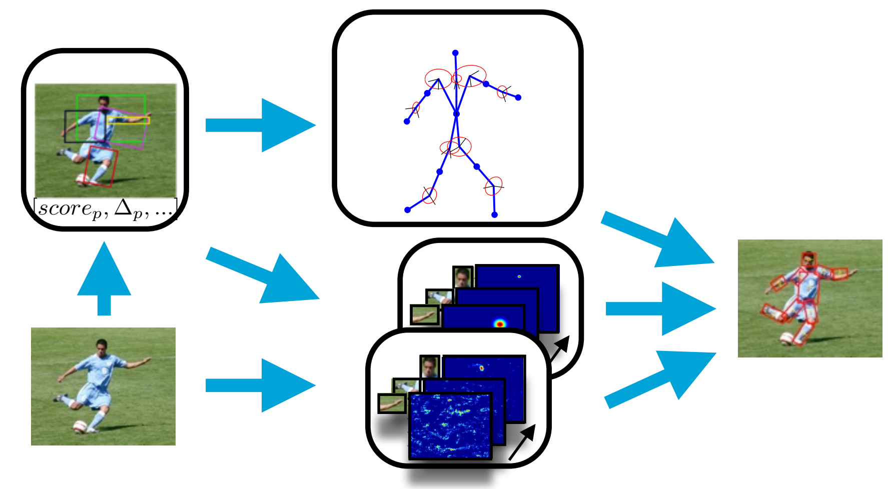
Rare poses

Conclusion

- **Starting point: Pictorial Structures**
  - ✓ exact and efficient
  - \- adjacent part dependencies only



- **Propose image conditioned model**
  - ✓ higher-order part dependencies
  - ✓ exact and efficient

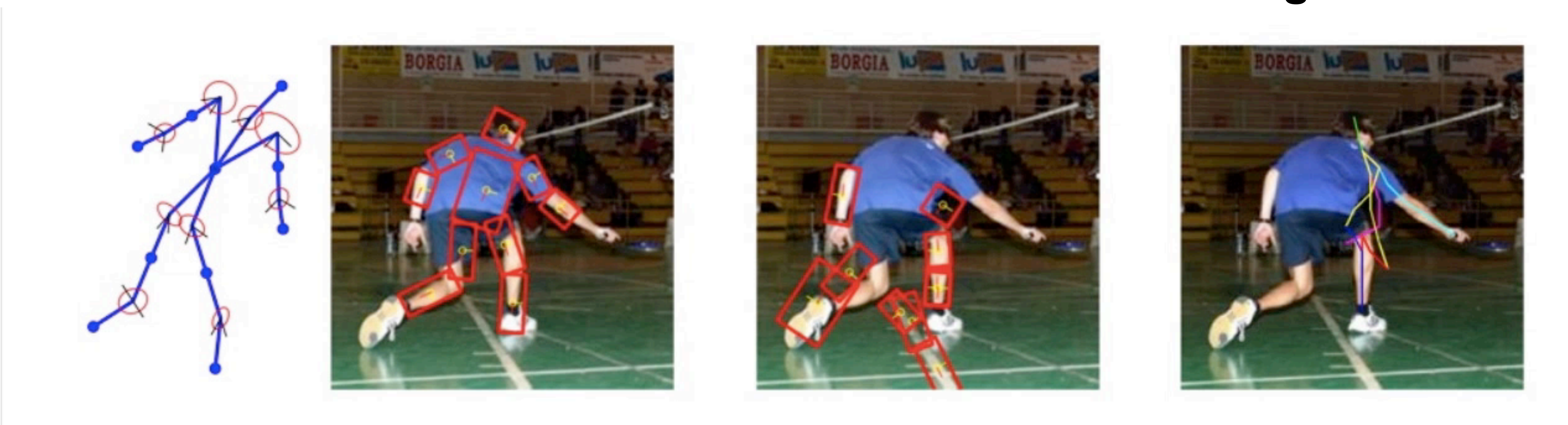


- **Evaluation on public benchmarks**
  - results improve or on par with the state of the art
- **Future work**
  - applications beyond pose estimation: object class detection etc.

Thank you for your attention!

- Poster: **Orals 1C-4**

Our method **Andriluka et al** **Yang&Ramanan**



Results - using true components

Setting	Torso	Upper leg	Lower leg	Upper arm	Forearm	Head	*Total*
Andriluka et al., CVPR'09	80.9	67.1	60.7	46.5	26.4	74.9	*55.7*
+ predict unary rotation (ur)	96.4	91.1	86.1	76.6	60.2	88.5	*81.3*
+ predict unary position (up)	97.1	91.4	80.7	80.2	49.5	90.1	*79.1*
+ predict pairwise (p/wise)	93.2	88.5	81.6	73.6	58.0	87.6	*78.4*
+ ur + up + p/wise	**98.3**	**96.0**	**89.4**	**87.0**	**71.8**	**94.0**	***88.1***

Results - using true components

Setting	Torso	Upper leg	Lower leg	Upper arm	Forearm	Head	*Total*
Andriluka et al., CVPR'09	80.9	67.1	60.7	46.5	26.4	74.9	*55.7*
+ predict unary rotation (ur)	96.4	91.1	86.1	76.6	60.2	88.5	*81.3*
+ predict unary position (up)	97.1	91.4	80.7	80.2	49.5	90.1	*79.1*
+ predict pairwise (p/wise)	93.2	88.5	81.6	73.6	58.0	87.6	*78.4*
+ ur + up + p/wise	**98.3**	**96.0**	**89.4**	**87.0**	**71.8**	**94.0**	***88.1***

- ✓ **Using true components dramatically improves the results**
- ✓ **Not perfect performance due to foreshortening and self-occlusion**

Results - comparison to state of the art

- Image Parse (IP) dataset
- train poselets on LSP, use 100 IP train images for validation

Method	Torso	Upper leg	Lower leg	Upper arm	Forearm	Head	*Total*
ours	**92.2**	74.6	63.7	54.9	39.8	70.7	*62.9*
ours + our CVPR'12	90.7	**80.0**	**70.0**	59.3	37.1	77.6	*66.1*
Andriluka et al., IJCV'12	86.3	66.3	60.0	54.6	35.6	72.7	*59.2*
Yang&Ramanan, CVPR'11	82.9	69.0	63.9	55.1	35.4	77.6	*60.7*
Duan et al., BMVC'12	85.6	71.7	65.6	57.1	36.6	80.4	*62.8*
our CVPR'12	88.8	77.3	67.1	53.7	36.1	73.7	*63.1*
Yang&Ramanan, PAMI'12	85.9	74.9	68.3	63.4	42.7	**86.8**	*67.1*
Johnson&Everingham, CVPR'11	87.6	74.7	67.1	**67.3**	**45.8**	76.8	***67.4***