

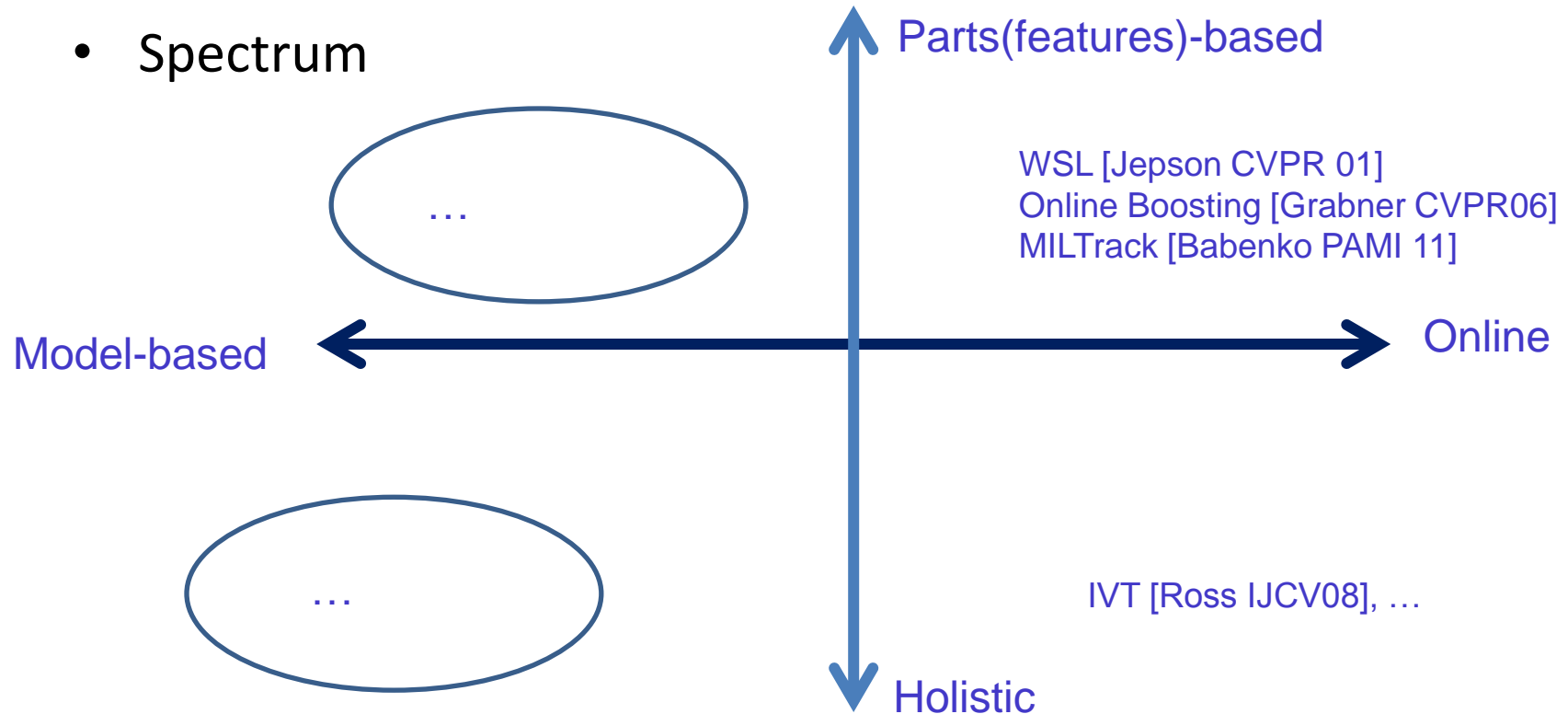
Learning to Track Objects

Qing Wang Ming-Hsuan Yang

University of California at Merced

Online and Model-based

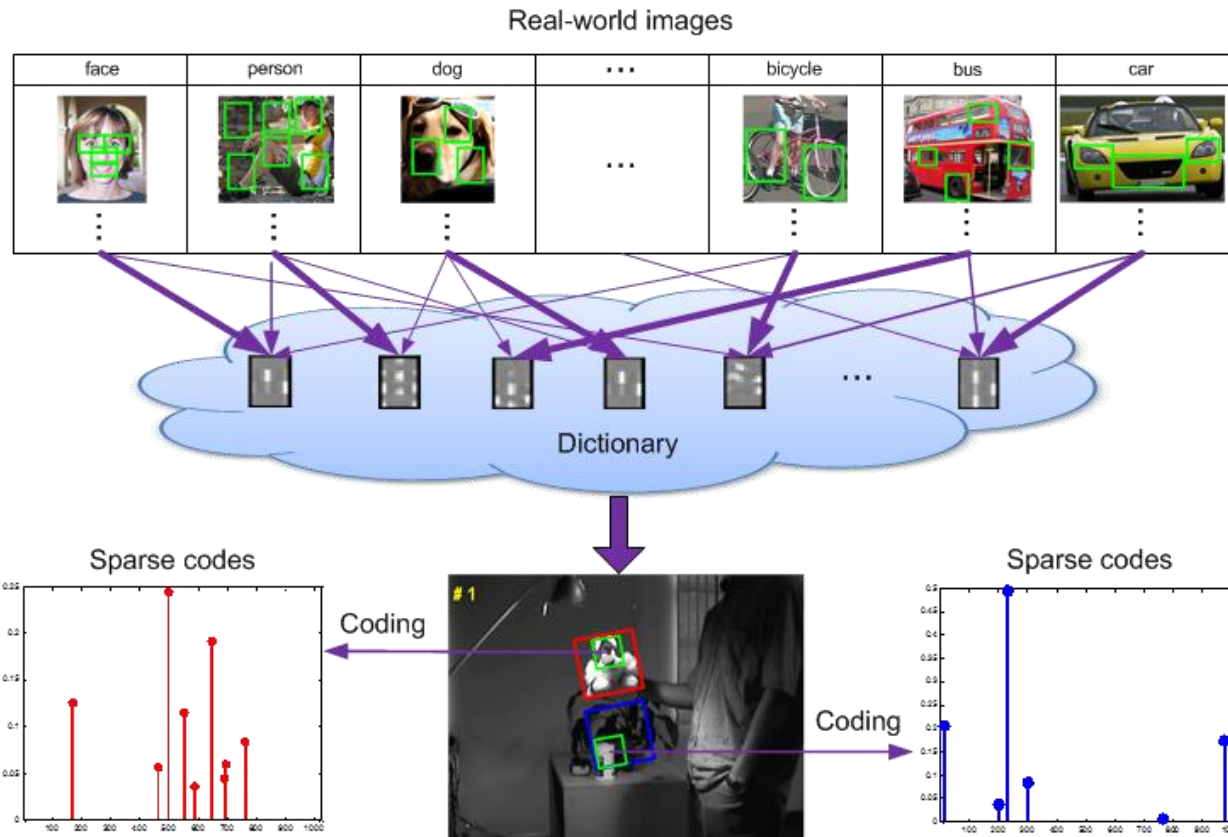
- Spectrum



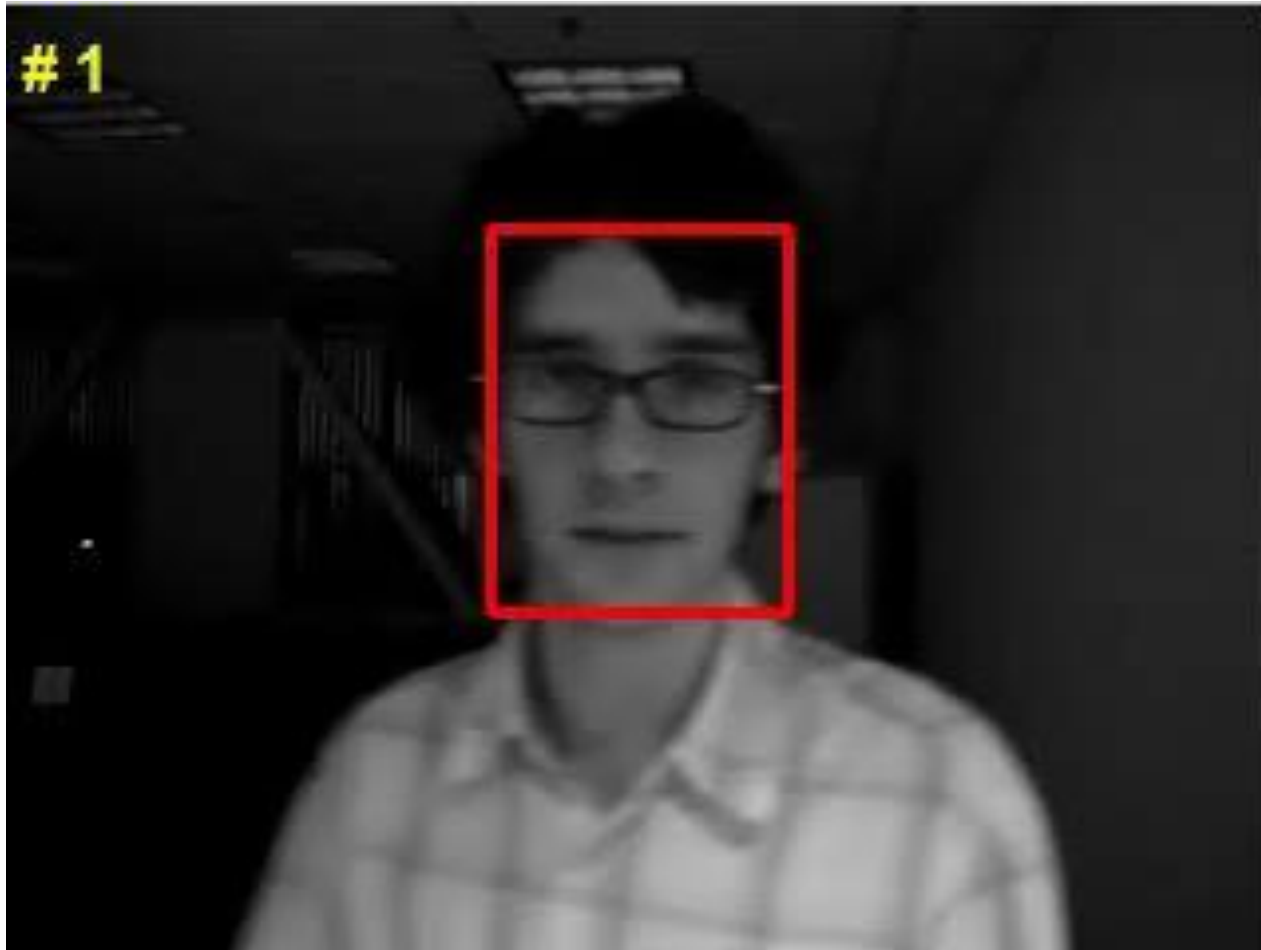
- Object detection, recognition and tracking
- Bridge the gaps

Transferring Visual Prior

- Image patches from real-world image data often share great similarity
- Visual prior
 - Learned offline to exploit structural similarity
 - Used for object representation in online object tracking with update



Preliminary Results



David_indoor
2004

Preliminary Results

**Transferring Visual Prior for
Online Object Tracking**

TrackLab

- A library of tracking code
 - Ongoing work, now with 5 algorithms
- Performance evaluation
 - Benchmark data sets
 - Empirical
 - Theoretical(w/assumption?):performance guarantee
- Ensemble of heterogeneous trackers

Performance Evaluation

- Taxonomy

Algorithm	Motion Model	Object Representation	Dynamic Model	Search Mechanism	Characteristics
IVT	affine	holistic gray-scale image vector	Gaussian	particle filter	generative
FragT	scaled translation	local gray-scale histograms	-	sampling	generative
VRT	translation	holistic color histograms	-	mean-shift	discriminative
BoostT	translation	holistic representation based on Haar-like, HOG and LBP descriptors	-	sampling	discriminative
SemiT	translation	holistic representation based on Haar-like descriptor	-	sampling	discriminative
BeSemiT	translation	holistic representation based on Haar-like, HOG, and color histograms	-	sampling	discriminative
L1T	affine	holistic gray-level image vector	Gaussian	particle filter	generative
MILT	translation	holistic representation based on Haar-like descriptor	-	sampling	discriminative
VTD	scaled translation	holistic representation based on hue, saturation, intensity, and edge template	Gaussian	particle filter	generative
TLD	scaled translation	holistic representation based on Haar-like descriptor	-	sampling	discriminative

Performance Evaluation

- Benchmark data sets
 - Build on top of existing data sets (PETS, AVSS, TRECVID, LabelMe Video, etc.)
- Characteristics, attributes:

Sequences	Main challenging factors	Resolution	Frames
<i>Sylvester</i>	in-plane/out-of-plane pose change, fast motion, illumination change	320 × 240	1343
<i>Wall-E</i>	scale change, out-of-plane pose change	608 × 256	178
<i>David-indoor</i>	illumination variation, out-of-plane pose change, partial occlusion	320 × 240	461
<i>surfing</i>	fast motion, large scale change, small object, moving camera	320 × 240	870
<i>singer</i>	scale change, significant illumination change	624 × 352	350
<i>shaking</i>	in-plane pose change, significant illumination change	624 × 352	365
<i>Gymnastic</i>	deformation, out-of-plane pose change	426 × 234	765
<i>jumping</i>	image blur, fast motion	352 × 288	312
<i>car</i>	image blur, partial occlusion	320 × 240	280
<i>faceocc</i>	in-plane pose change, partial occlusion	320 × 240	812
<i>PETS2009</i>	heavy occlusion, out-of-plane pose change, distraction from similar objects	768 × 576	146
<i>CAVIAR</i>	heavy occlusion, distraction from similar objects	320 × 240	608
<i>board</i>	background clutter, out-of-plane pose change	640 × 480	698
<i>Avatar</i>	occlusion, out-of-plane pose change, illumination change	704 × 384	192
<i>David-outdoor</i>	low-contrast images, occlusion, out-of-plane pose change	320 × 240	251

Performance Evaluation

- Criteria

	IVT	VRT	FragT	BoostT	SemiT	BeSemiT	L1T	MILT	VTD	TLD
<i>Sylvester</i>	49	17	11	14	8	8	20	12	12	8
<i>WallE2</i>	58	21	34	53	44	-	14	28	18	30
<i>David-indoor</i>	19	114	58	32	20	32	45	30	18	13
<i>surfing</i>	43	9	35	6	3	5	37	51	36	34
<i>singer</i>	9	107	21	14	14	12	3	16	3	23
<i>shaking</i>	130	206	107	25	2	-	22	10	7	167
<i>Gymnastics</i>	10	10	8	15	7	12	123	16	9	12
<i>jumping</i>	5	70	33	10	10	19	7	6	10	6
<i>car</i>	57	42	70	5	10	2	72	4	5	7
<i>faceocc2</i>	18	66	44	22	23	5	33	16	53	9
<i>PETS2009</i>	72	64	154	737	59	79	43	66	73	85
<i>CAVIAR</i>	62	19	73	55	27	49	36	101	53	33
<i>board</i>	93	74	73	118	30	16	152	90	93	118
<i>Avatar</i>	54	160	104	14	50	122	49	48	53	69
<i>David-outdoor</i>	99	42	68	40	108	52	36	41	40	43

- Much more work to be done