TRECVID-2007: Shot Boundary Detection Task Summary

Alan Smeaton
Dublin City University
&
Paul Over
NIST

SB Task Definition

- Shot boundary detection is a fundamental task in any kind of video content manipulation
- Task provides a good entry for groups who wish to “break into” video retrieval and TRECVID gradually
- Task is to identify the shot boundaries with their location and type (cut or gradual) in the given video clip(s)

SB Task Details

- Groups may submit up to 10 runs
- Comparison to human-annotated reference (thanks to Jonathan Lasko, again, but last time)
- Groups were again asked to provide some standard information on the processing complexity of each run:
  - Total runtime in seconds
    - Total decode time in seconds
    - Total segmentation time in seconds
  - Processor description

Participating groups (26 in 2006, 7 re-shows)

1. Asahi Kasei Corporation
2. AT&T Labs
3. Beijing University of Posts and Telecommunications
4. Brno University of Technology
5. COST292 Team
6. Florida International University, FIU-UM
7. Fraunhofer Institute IAIS and University of Bradford
8. HuaZhong University of Science and Technology
9. LIG (Laboratoire d'Informatique de Grenoble)
10. NHK Science and Technical Research Laboratories
11. Philipps University Marburg
12. THU-ICRC (Tsinghua University, Intel China Research Center)
13. University of Karlsruhe (TH)
14. University of Modena and Reggio Emilia (Italy)
15. University of Sheffield

Shot boundary data

- 17 representative news videos
- Total frames: 637,805 (TV2006: 597,043)
- Total transitions: 2,463 (TV2006: 3,785)
- Transition types:
  - **2,236 (90.8%) Cuts (TV2006: 48.7%)**
  - **134 (5.4%) Dissolves (TV2006: 39.9%)**
  - **2 (<0.1%) Fade-out/-in (TV2006: 1.3%)**
  - **91 (3.7%) other (TV2006: 10.1%)**
- Almost all cuts (twice the fraction in past years) thus graduals less important

Evaluation Measures

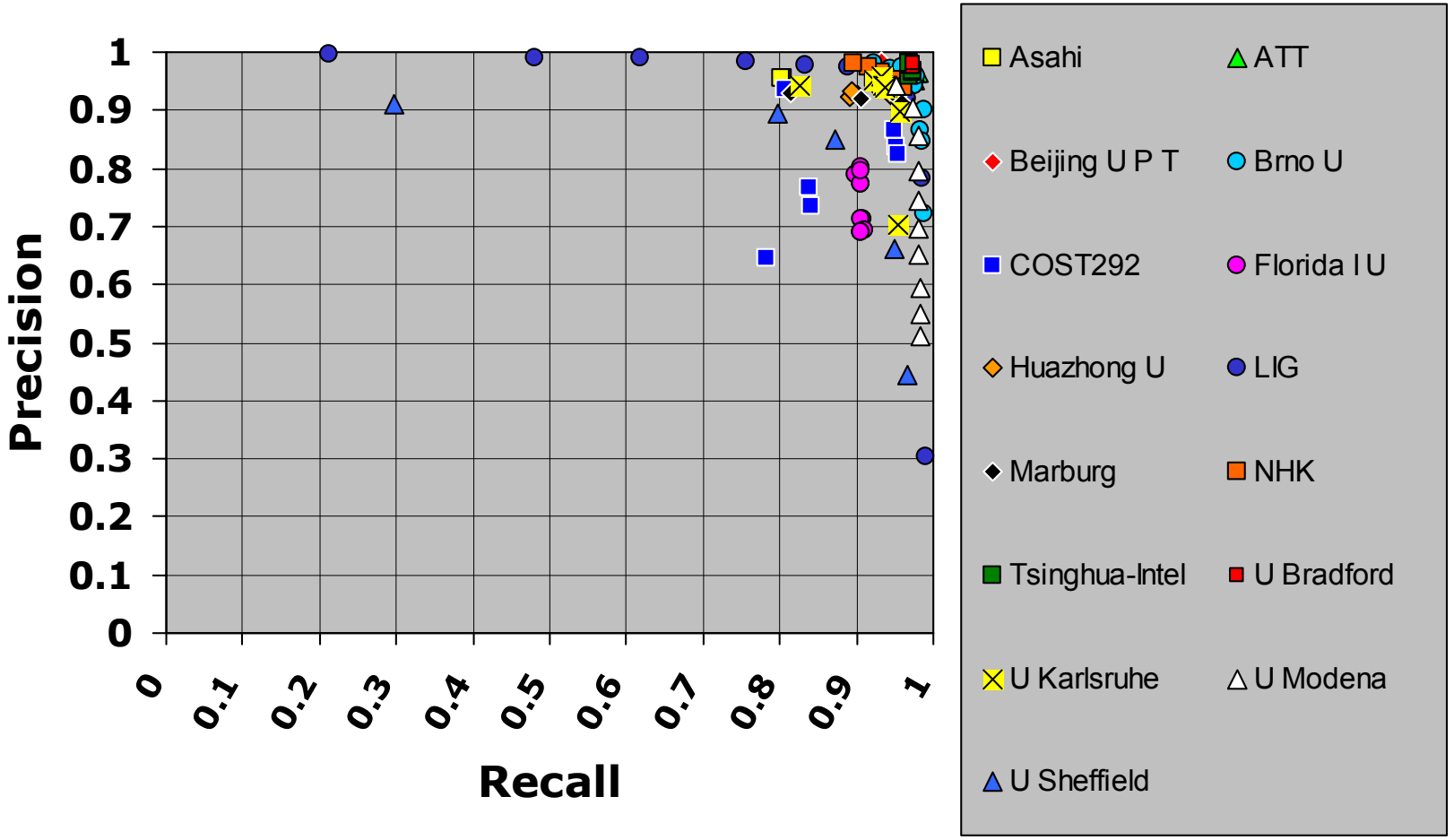
$$\text{Precision} = \frac{\#\ \text{Transitions Correctly Reported}}{\#\ \text{Transitions Reported}}$$

$$\text{Recall} = \frac{\#\ \text{Transitions Correctly Reported}}{\#\ \text{Transitions in Reference}}$$

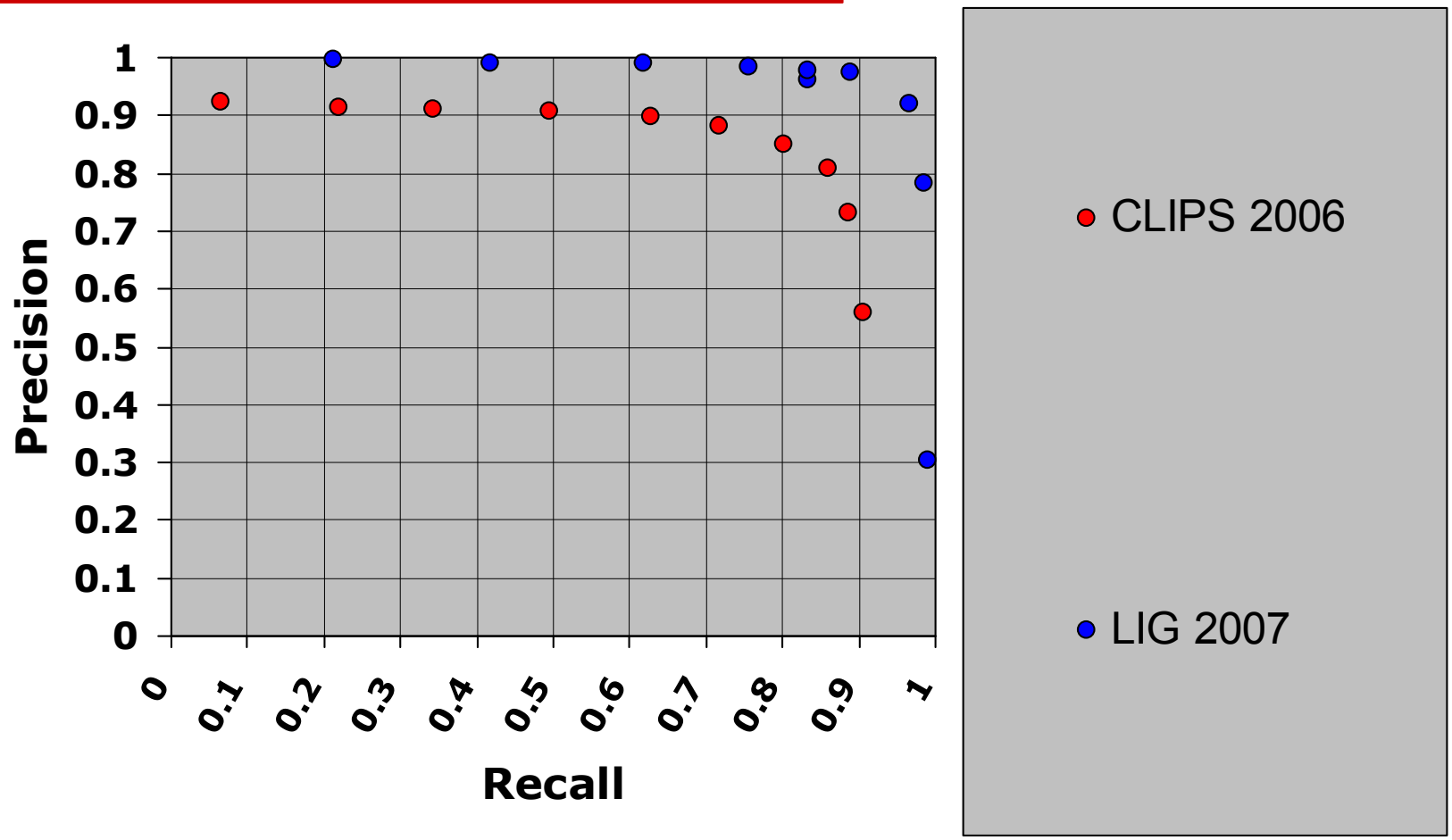
$$\text{Frame Precision} = \frac{\#\ \text{Frames Correctly Reported in Detected Transitions}}{\#\ \text{Frames reported in Detected Transitions}}$$

$$\text{Frame Recall} = \frac{\#\ \text{Frames Correctly Reported in Detected Transitions}}{\#\ \text{Frames in Reference Data for Detected Transitions}}$$

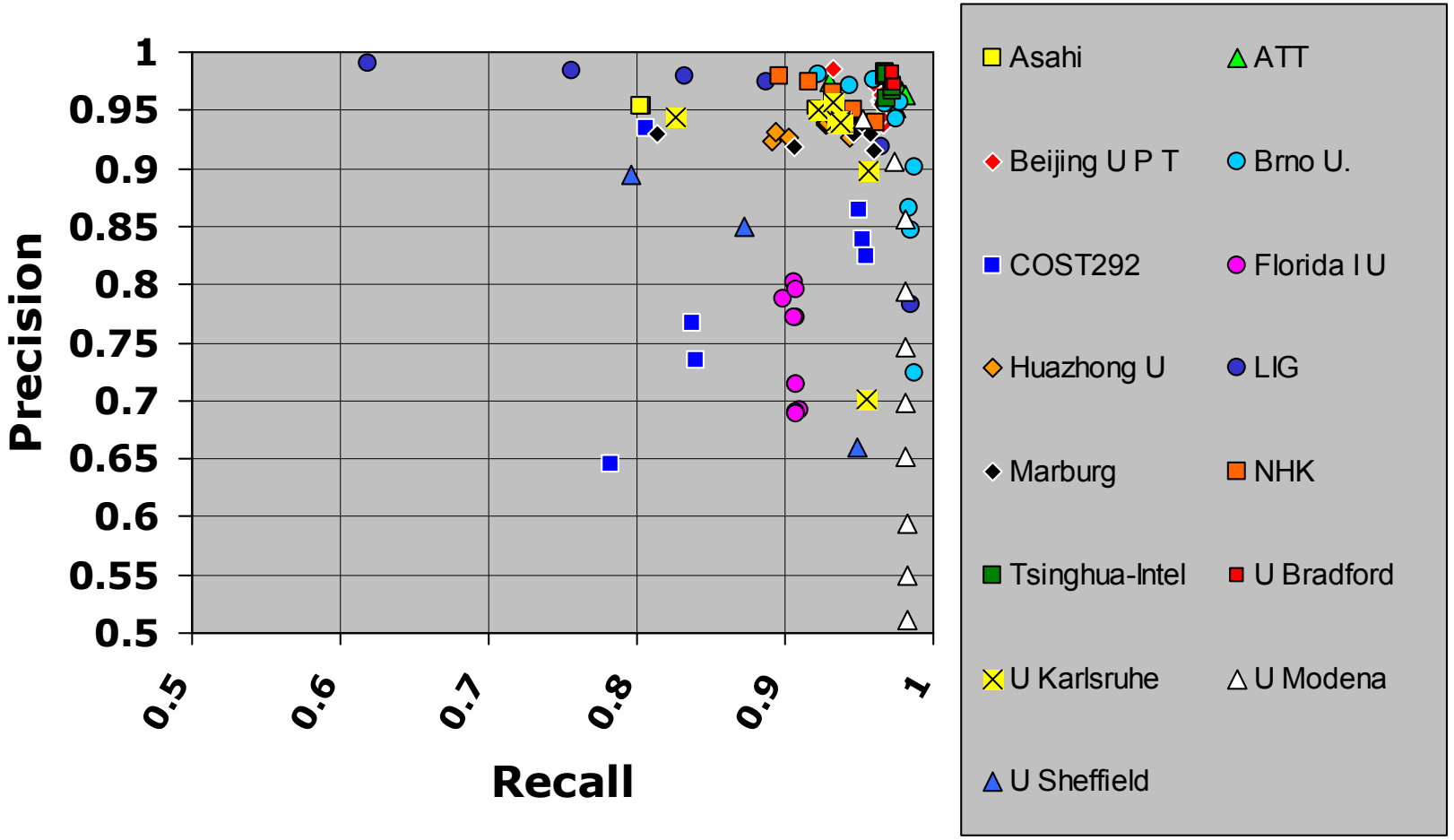
Cuts


Precision
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Recall
Asahi
ATT
Beijing U P T
Brno U
COST292
Florida I U
Huazhong U
LIG
Marburg
NHK
Tsinghua-Intel
U Bradford
U Karlsruhe
U Modena
U Sheffield

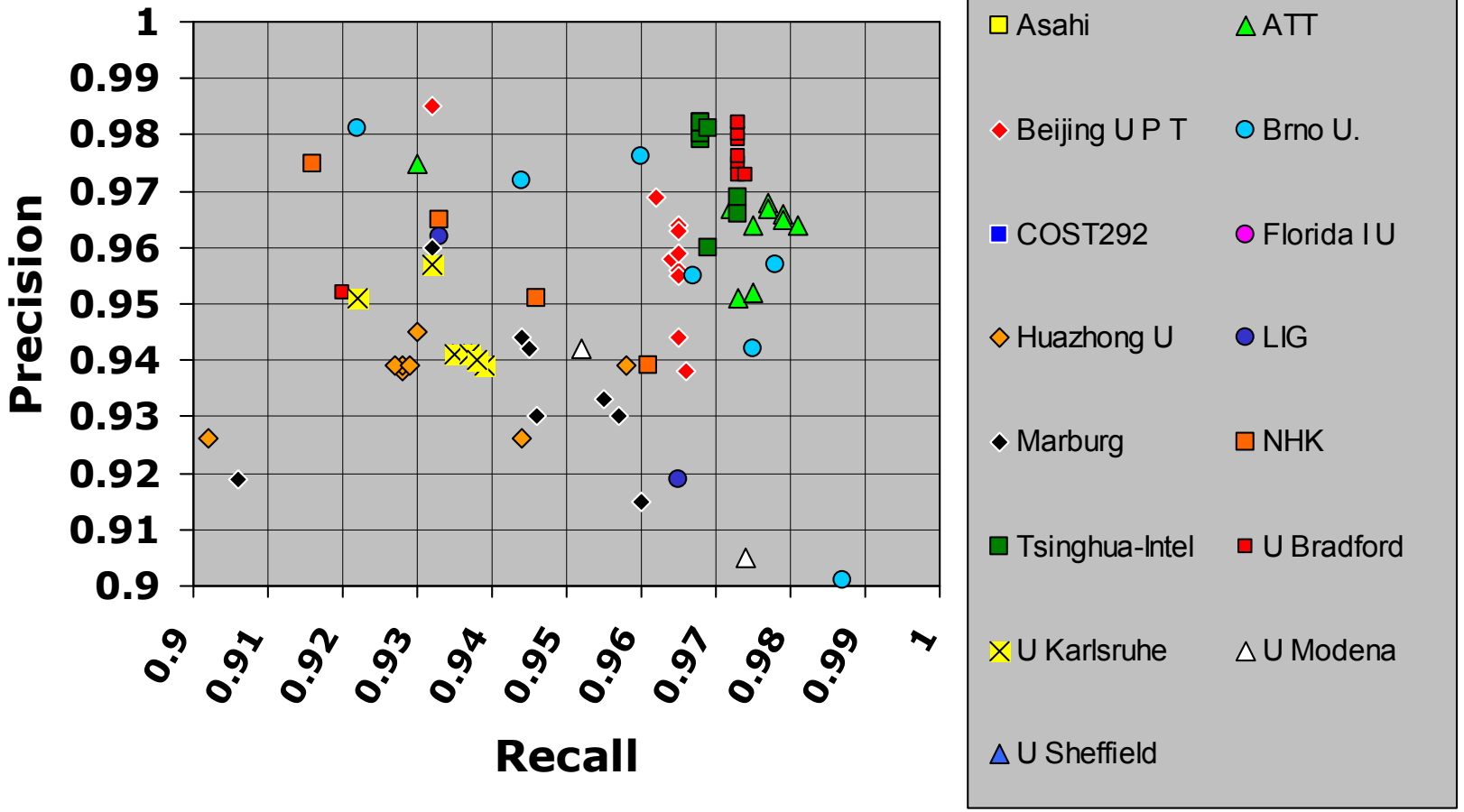
Cuts: same system, different (easier?) data



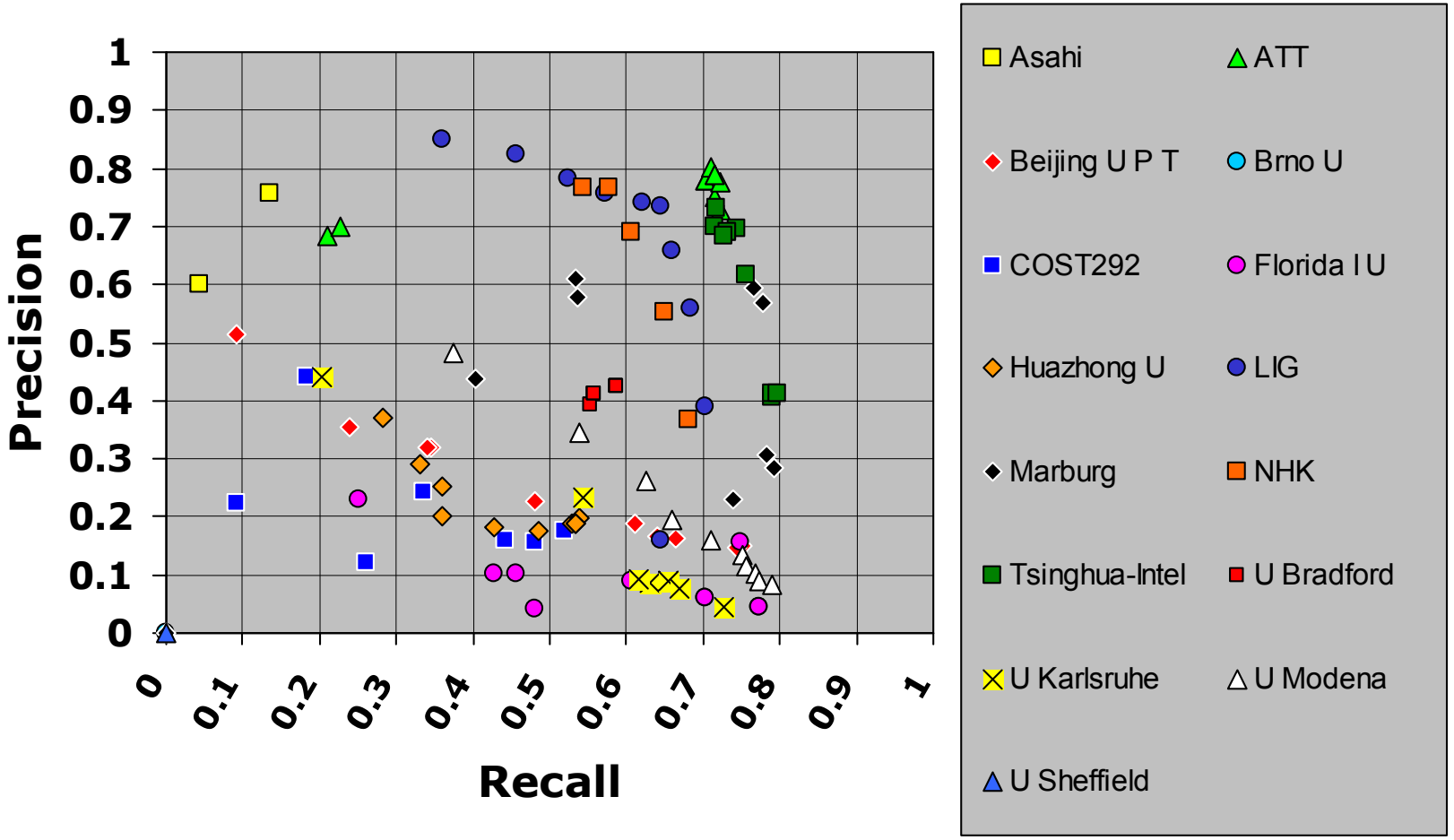
Cuts (zoomed)


Precision
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.5
0.6
0.7
0.8
0.9
1
Recall
Asahi
ATT
Beijing U P T
Brno U.
COST292
Florida I U
Huazhong U
LIG
Marburg
NHK
Tsinghua-Intel
U Bradford
U Karlsruhe
U Modena
U Sheffield

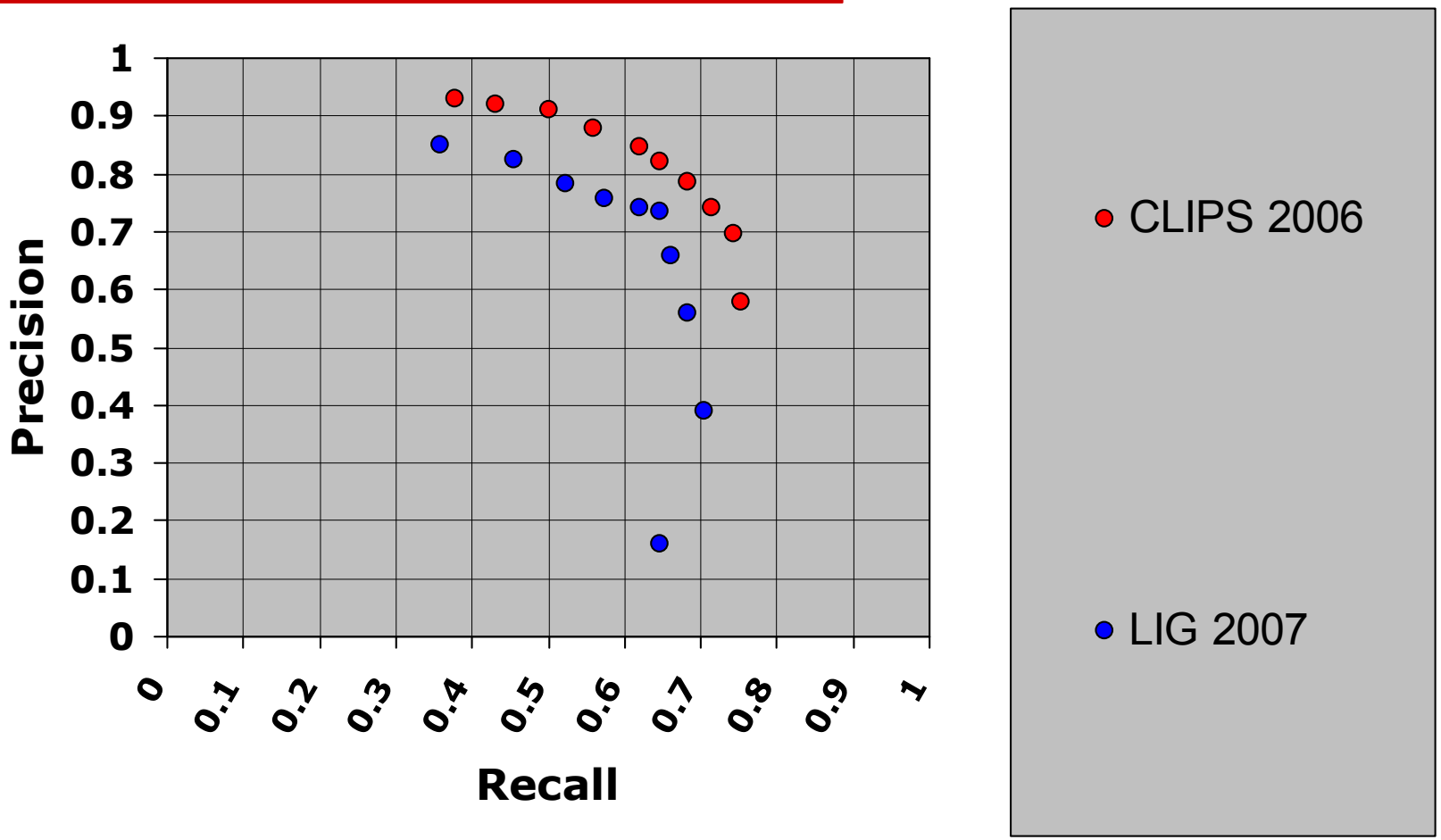
Cuts (zoomed again)



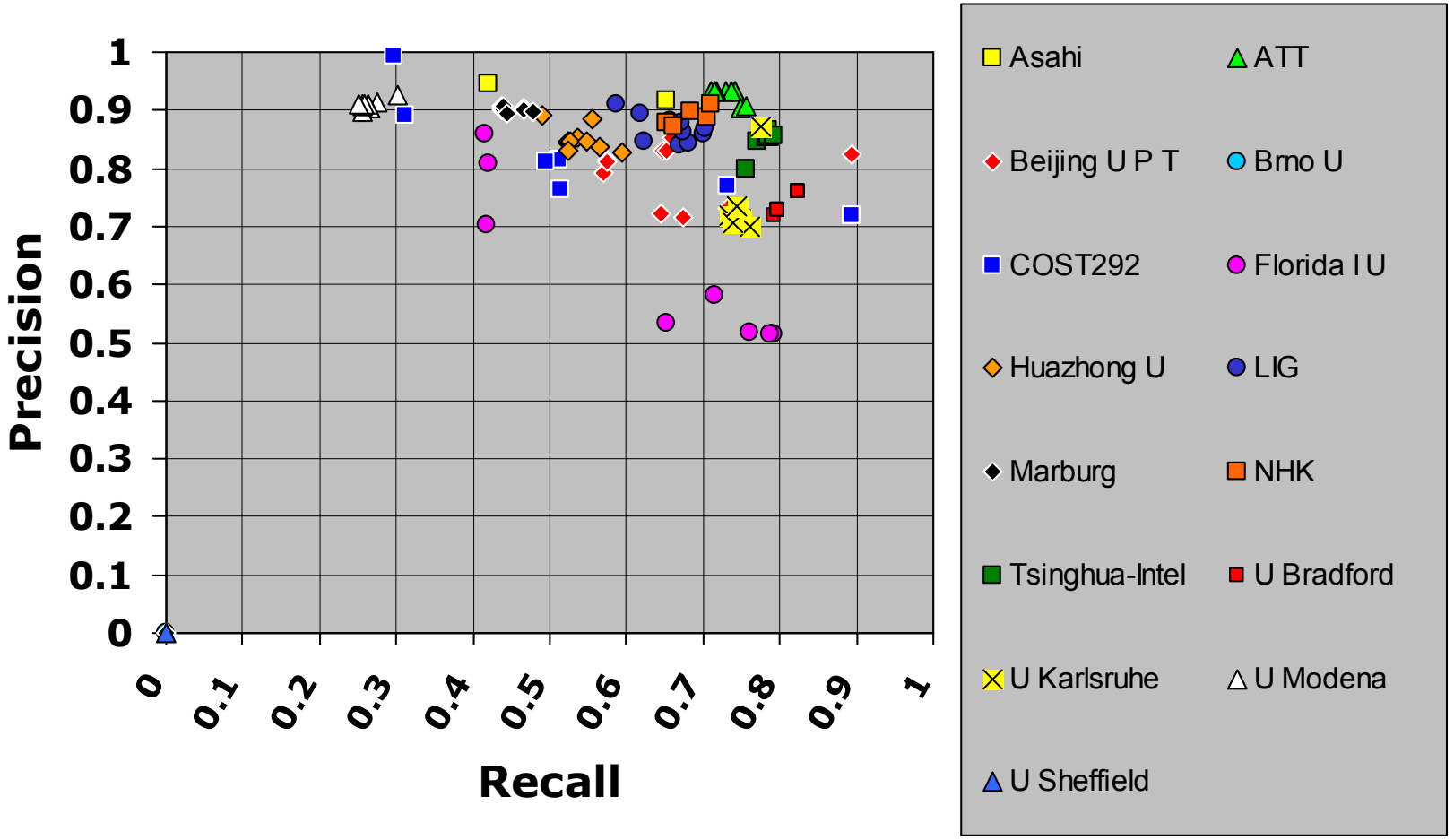
Gradual transitions



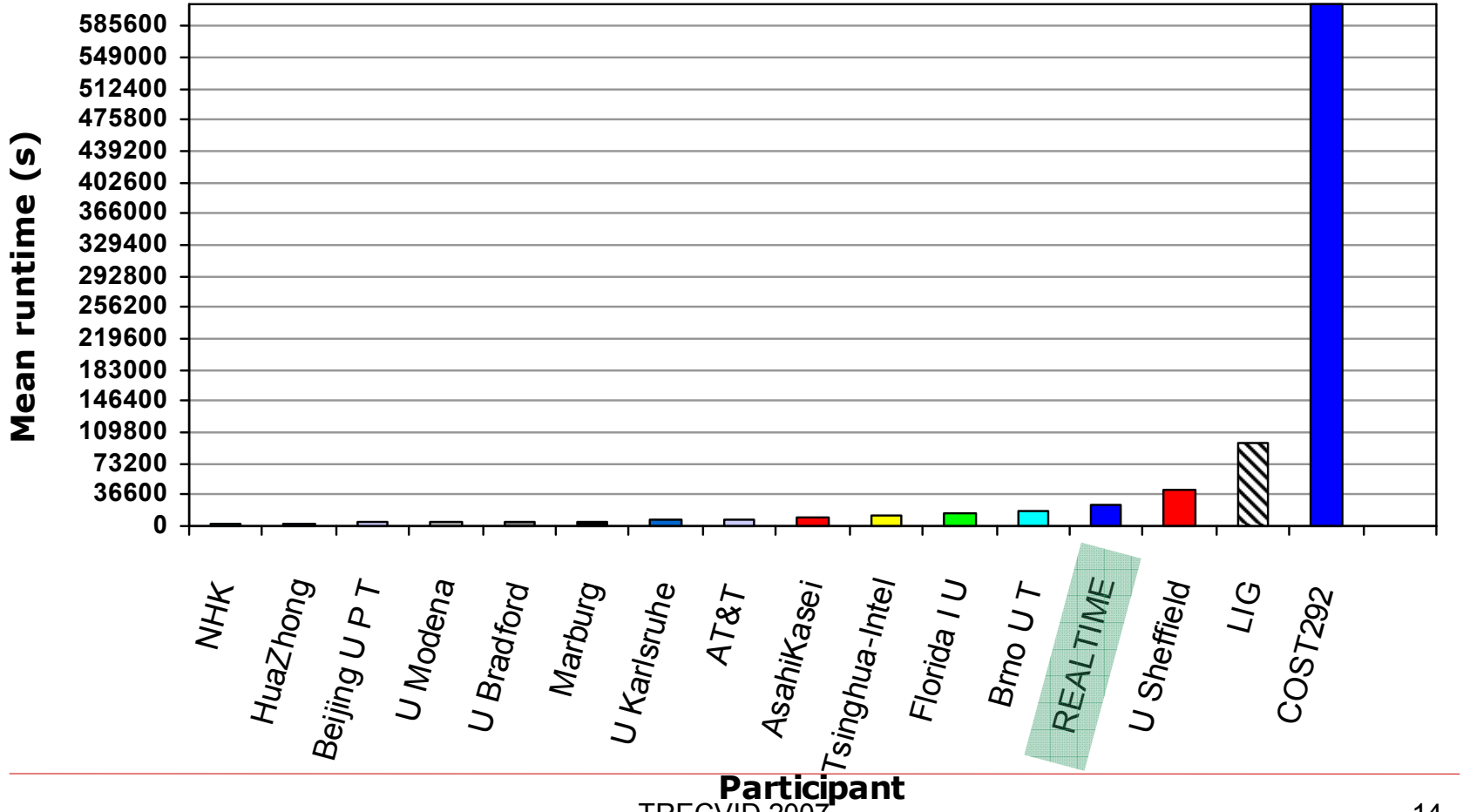
(harder data and/or less agreement with annotation)



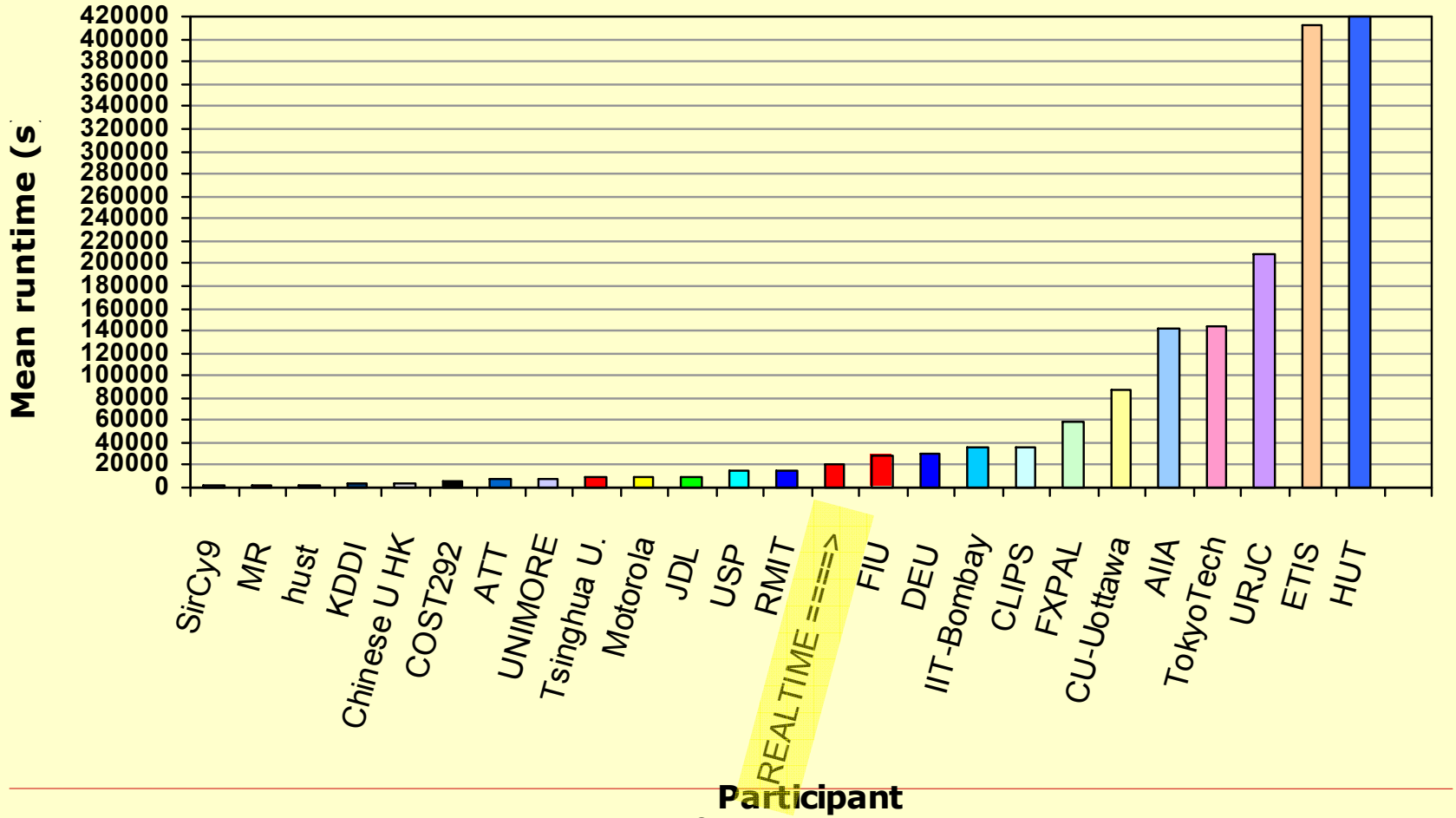
Gradual transitions (Frame-P & -R)



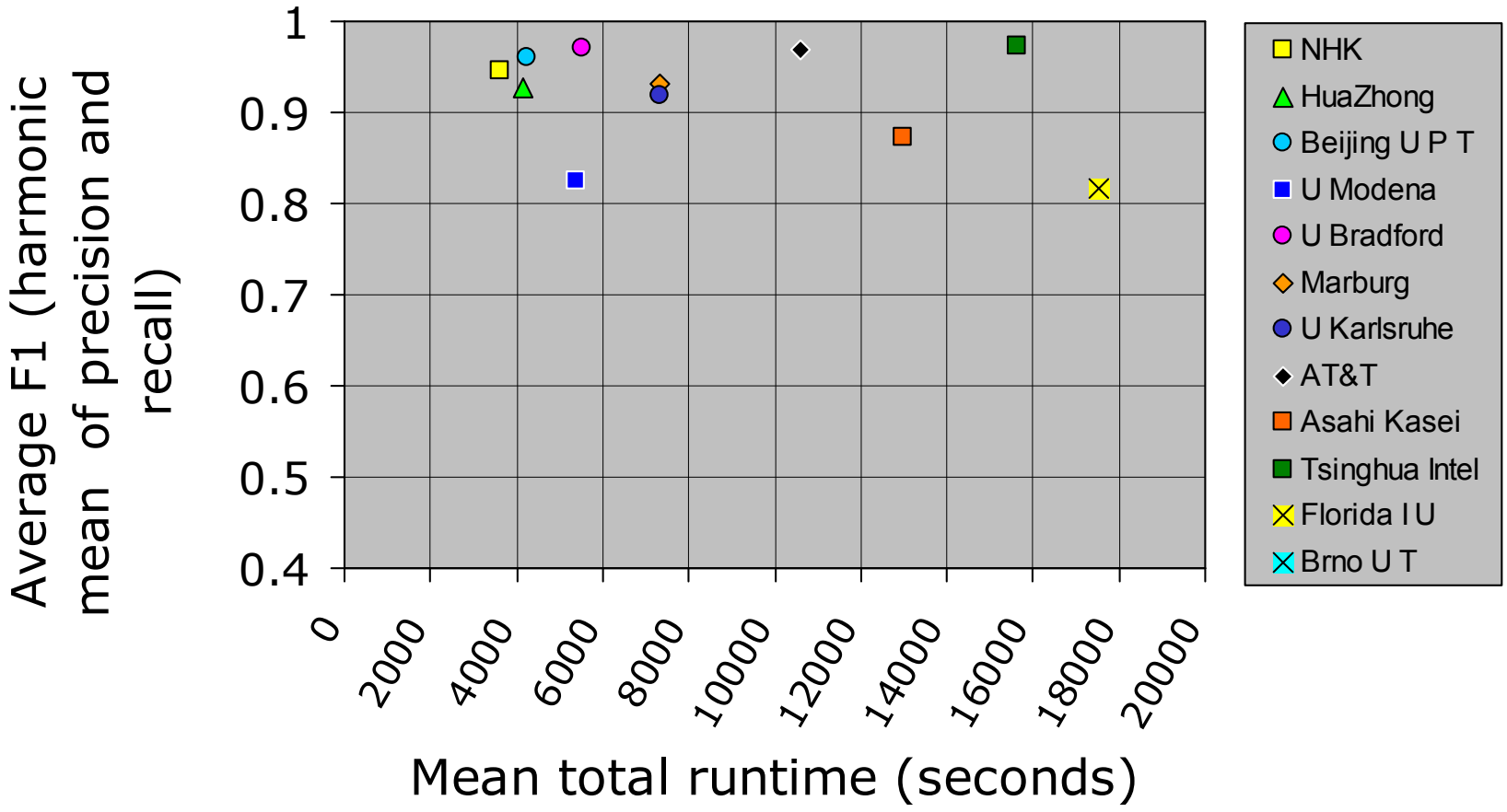
Mean runtime in seconds



Mean runtime in seconds



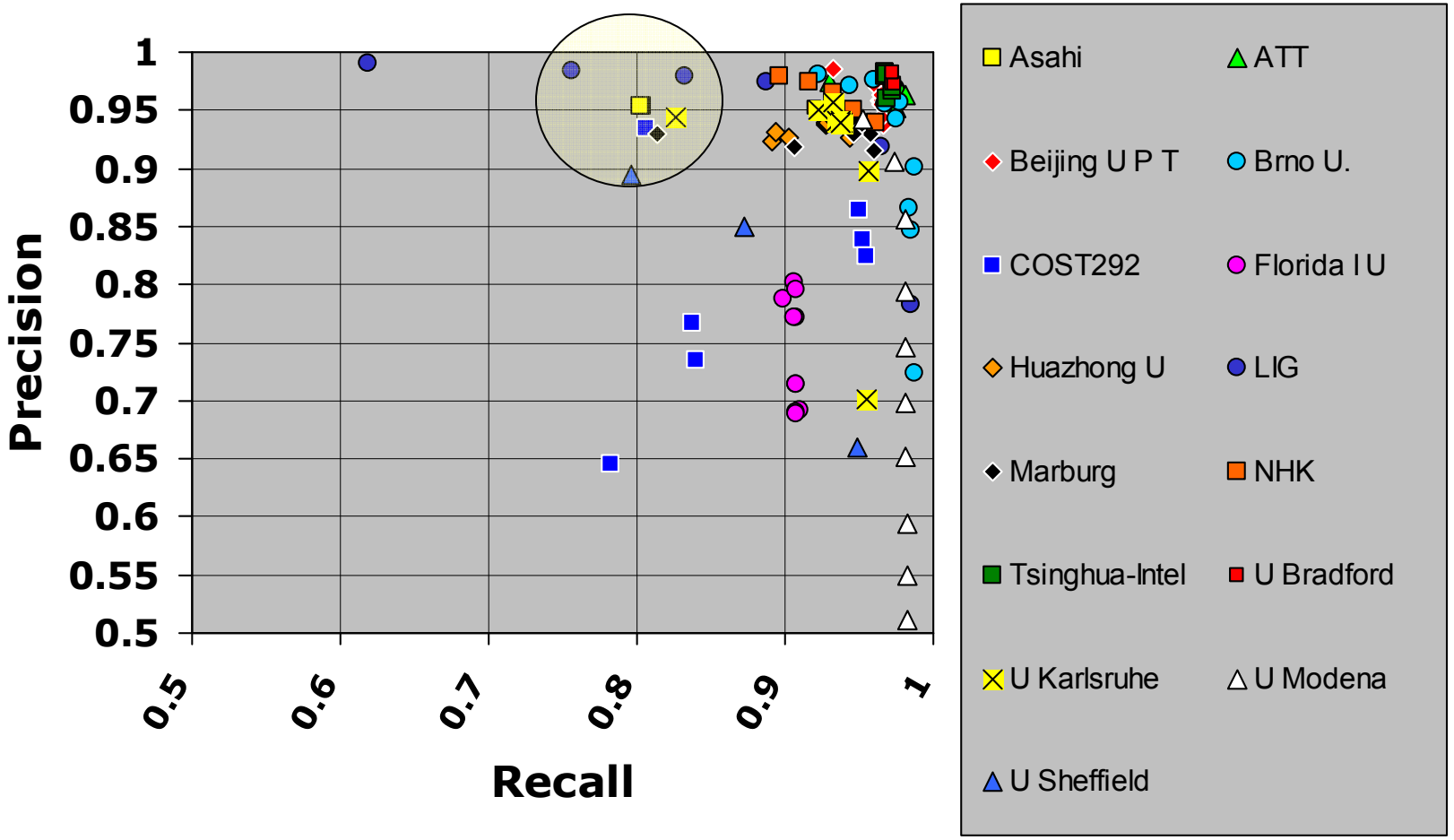
effectiveness on cuts
(for systems faster than realtime)



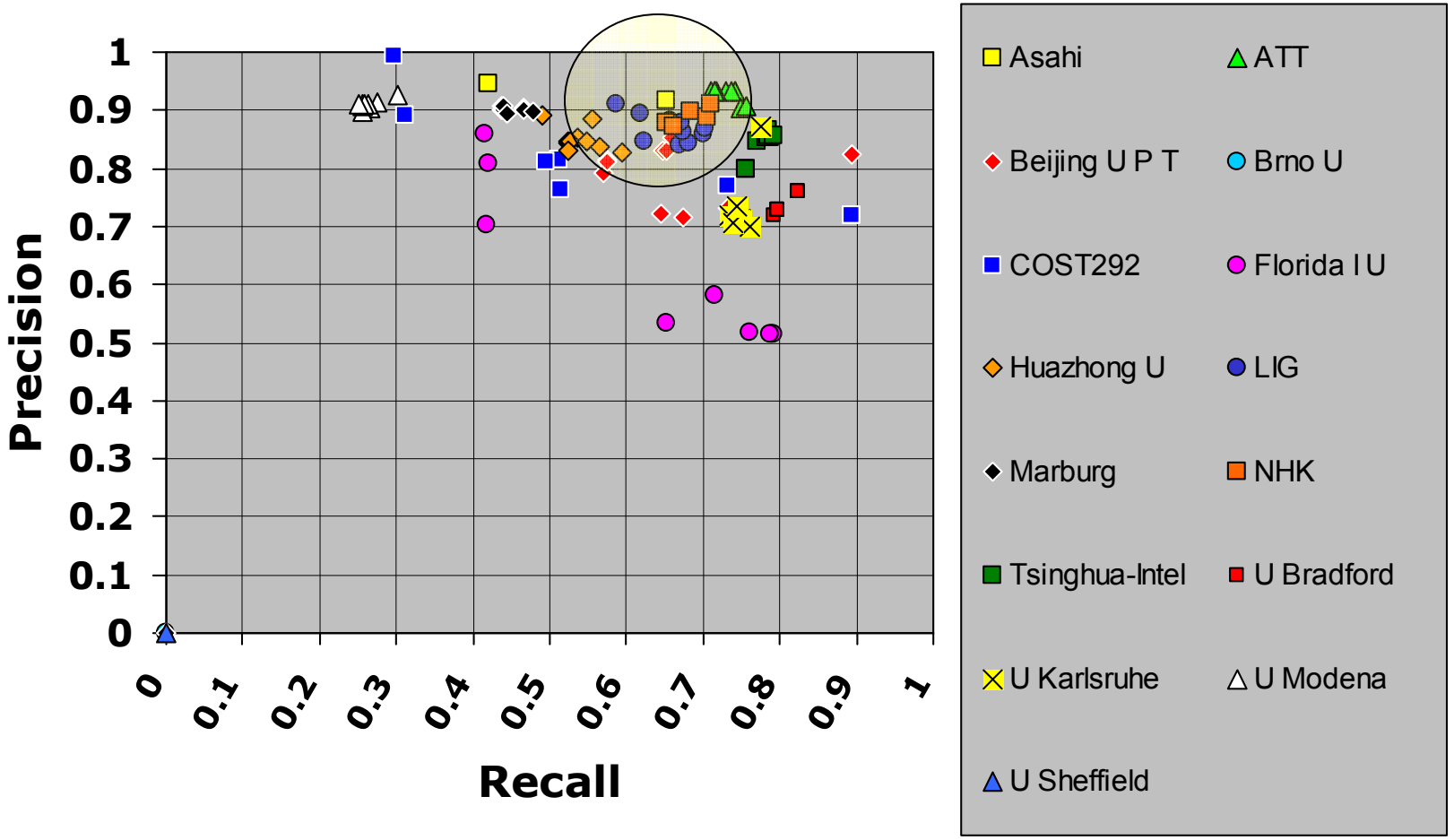
1. Asahi Kasei Corporation

- Use correlations of successive low-resolution gray scale images to detect cuts.
- Longer time correlations (± 50 frames) to detect gradual fades.
- Found normalization necessary for correlation. Colour correlations ineffectual at shot boundary detection. Gradual fade detection requires correlation of >2 images which is difficult, though gradual detector improved performance.

Cuts (zoomed)


Precision
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.5
0.6
0.7
0.8
0.9
1
Recall
Asahi
ATT
Beijing U P T
Brno U.
COST292
Florida I U
Huazhong U
LIG
Marburg
NHK
Tsinghua-Intel
U Bradford
U Karlsruhe
U Modena
U Sheffield

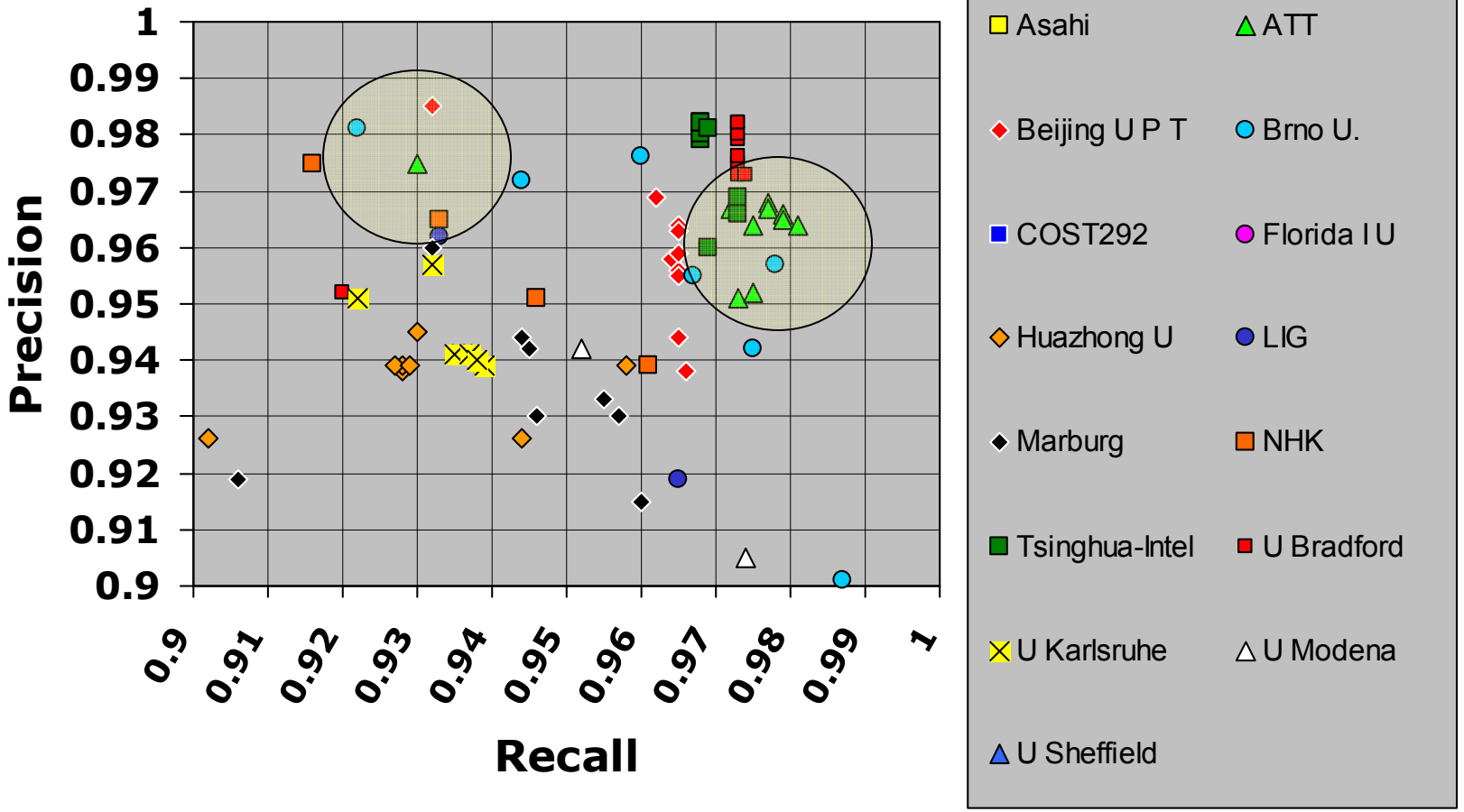
Gradual transitions (Frame-P & -R)



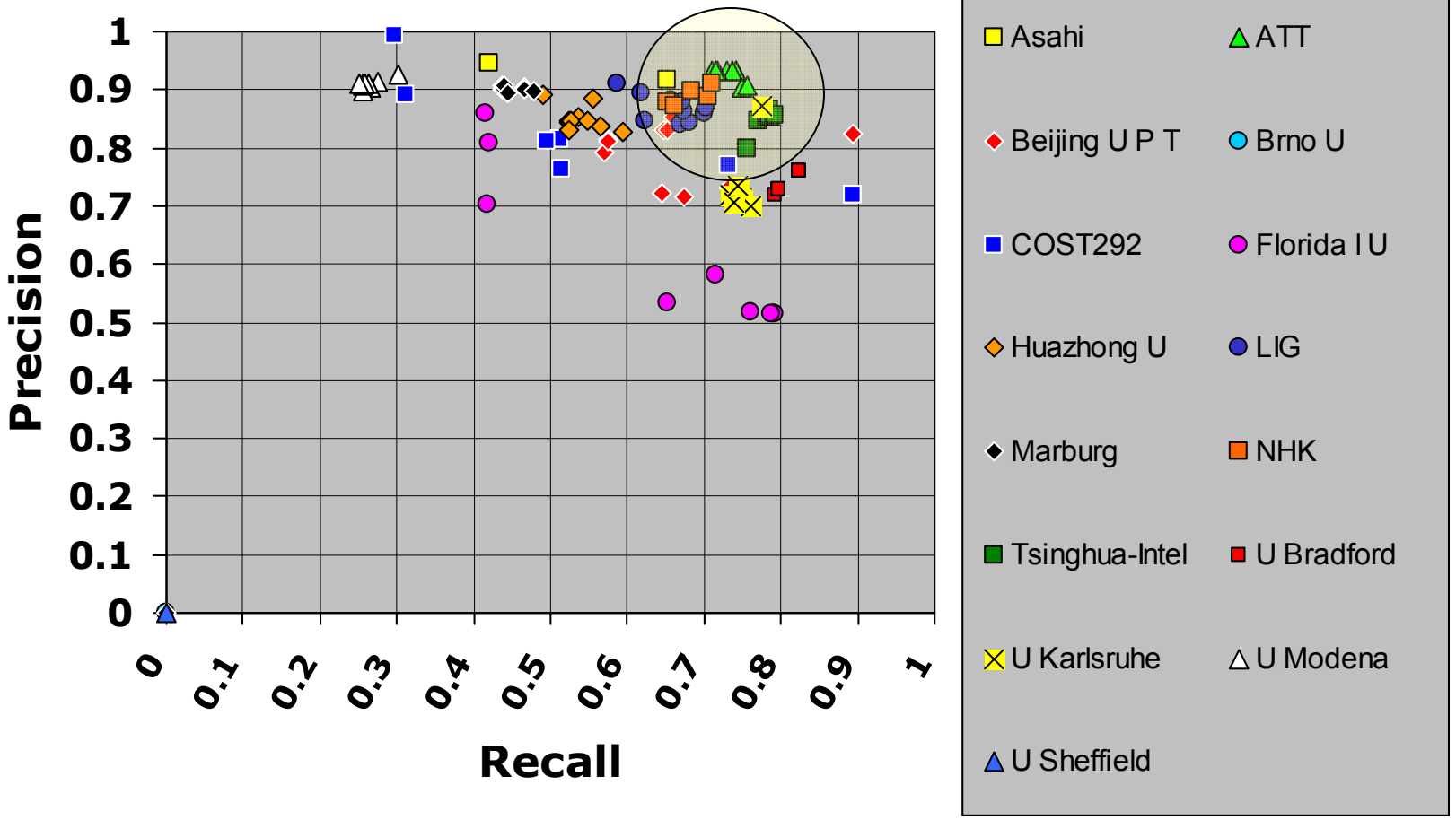
2. AT&T Labs

- Enhanced the 2006 system, new visual features (intra- and inter-frame) from central RoI in frame, and SVM-based verification to boost accuracy
- Speed improved through streamlined processing
- Fusion of independent detectors for cuts, fast dissolves, fade-in, fade-out, dissolves and wipes

Cuts (zoomed again)



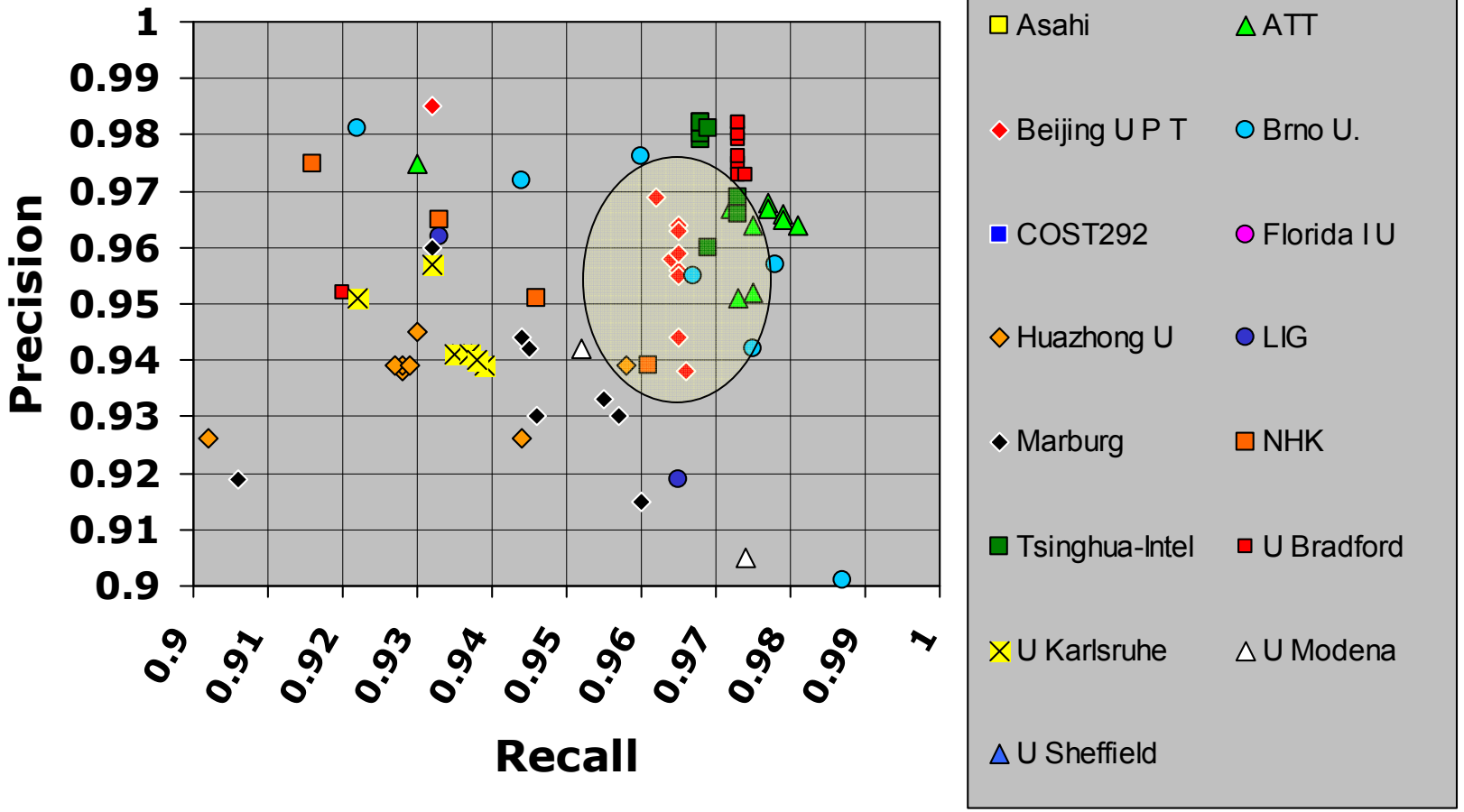
Gradual transitions (Frame-P & -R)



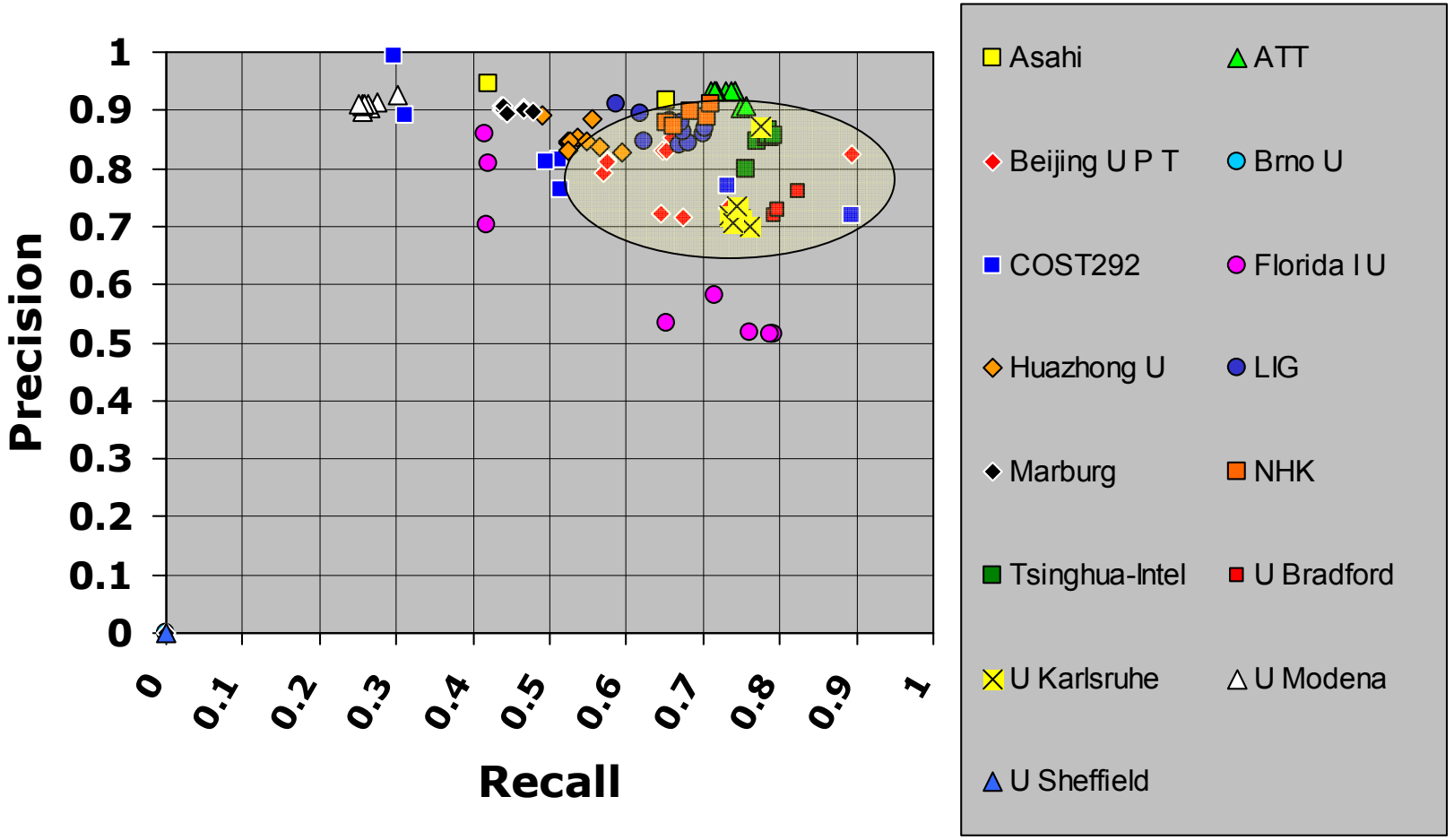
3. Beijing University of Posts and Telecommunications

- 4 independent detectors for cuts (trained SVM), FOI, gradual transitions and motion.moving object
- Gradual transition detector uses motion detector to isolate motion
- CUT detector performed well (P/R typically >95%).
- Motion detector identified many false positives and thus hampered the gradual transition detector's performance

Cuts (zoomed again)



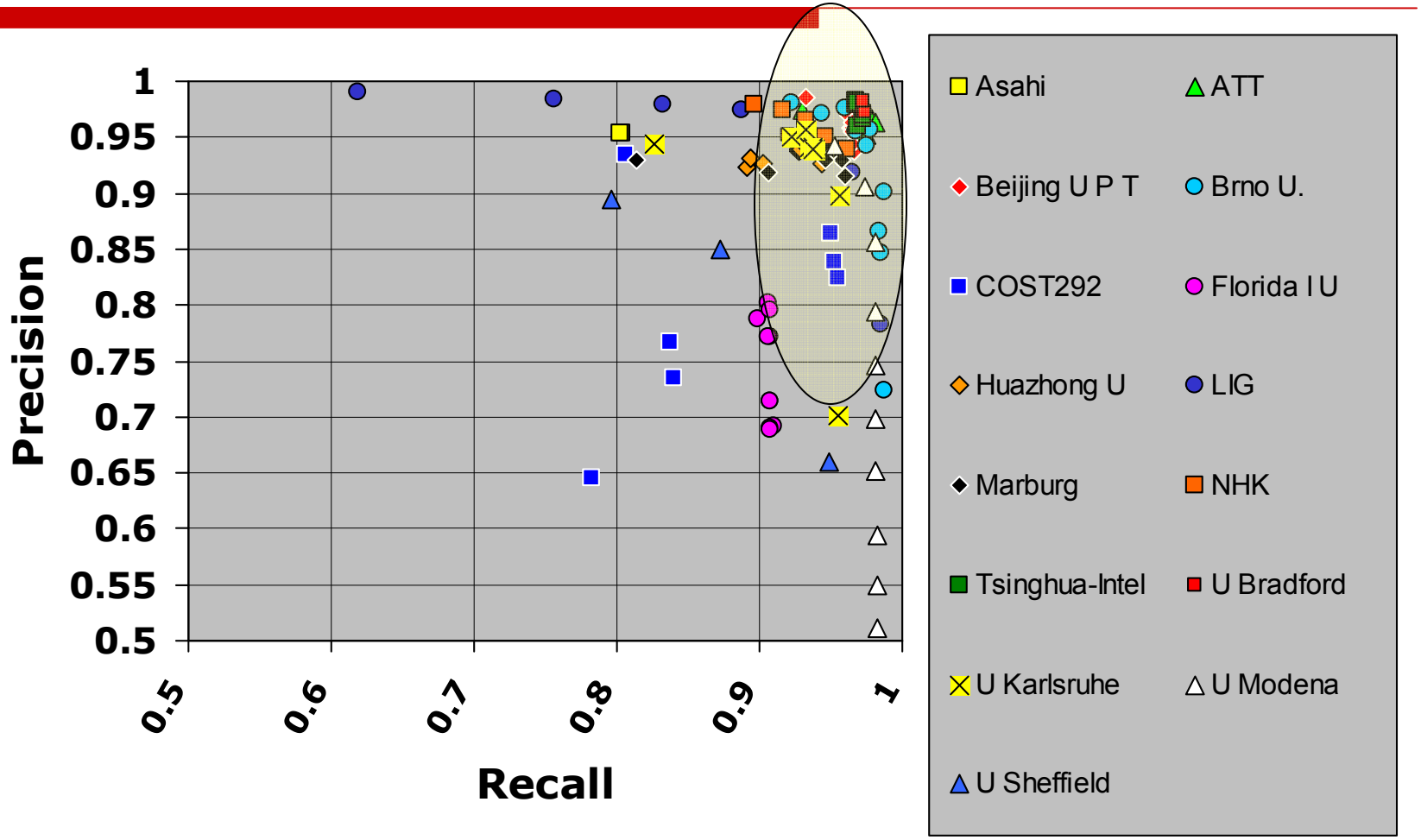
Gradual transitions (Frame-P & -R)



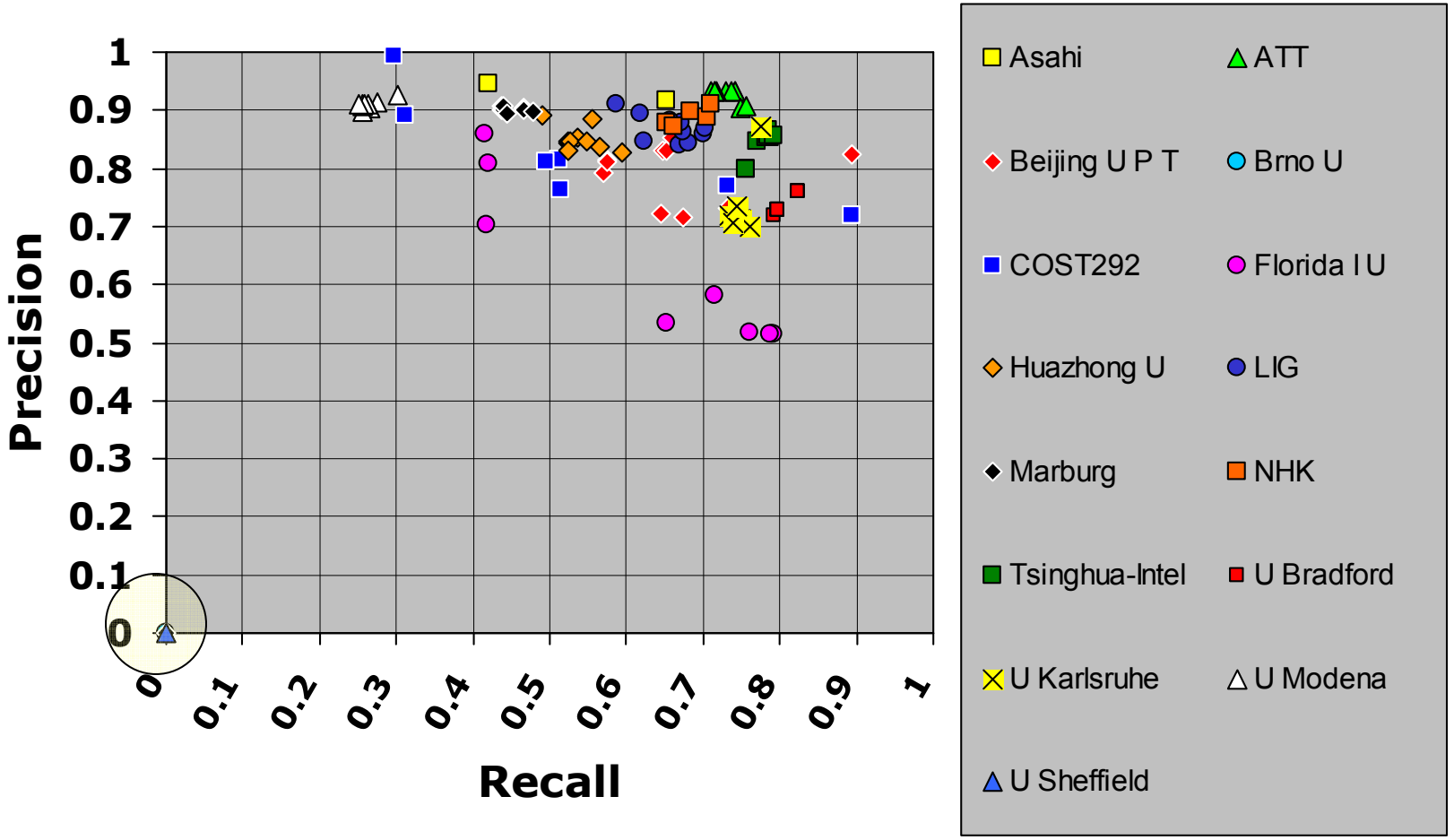
4. Brno University of Technology

- Late/missing paper ?

Cuts (zoomed)


Precision
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.5
0.6
0.7
0.8
0.9
1
Recall
Asahi
ATT
Beijing U P T
Brno U.
COST292
Florida I U
Huazhong U
LIG
Marburg
NHK
Tsinghua-Intel
U Bradford
U Karlsruhe
U Modena
U Sheffield

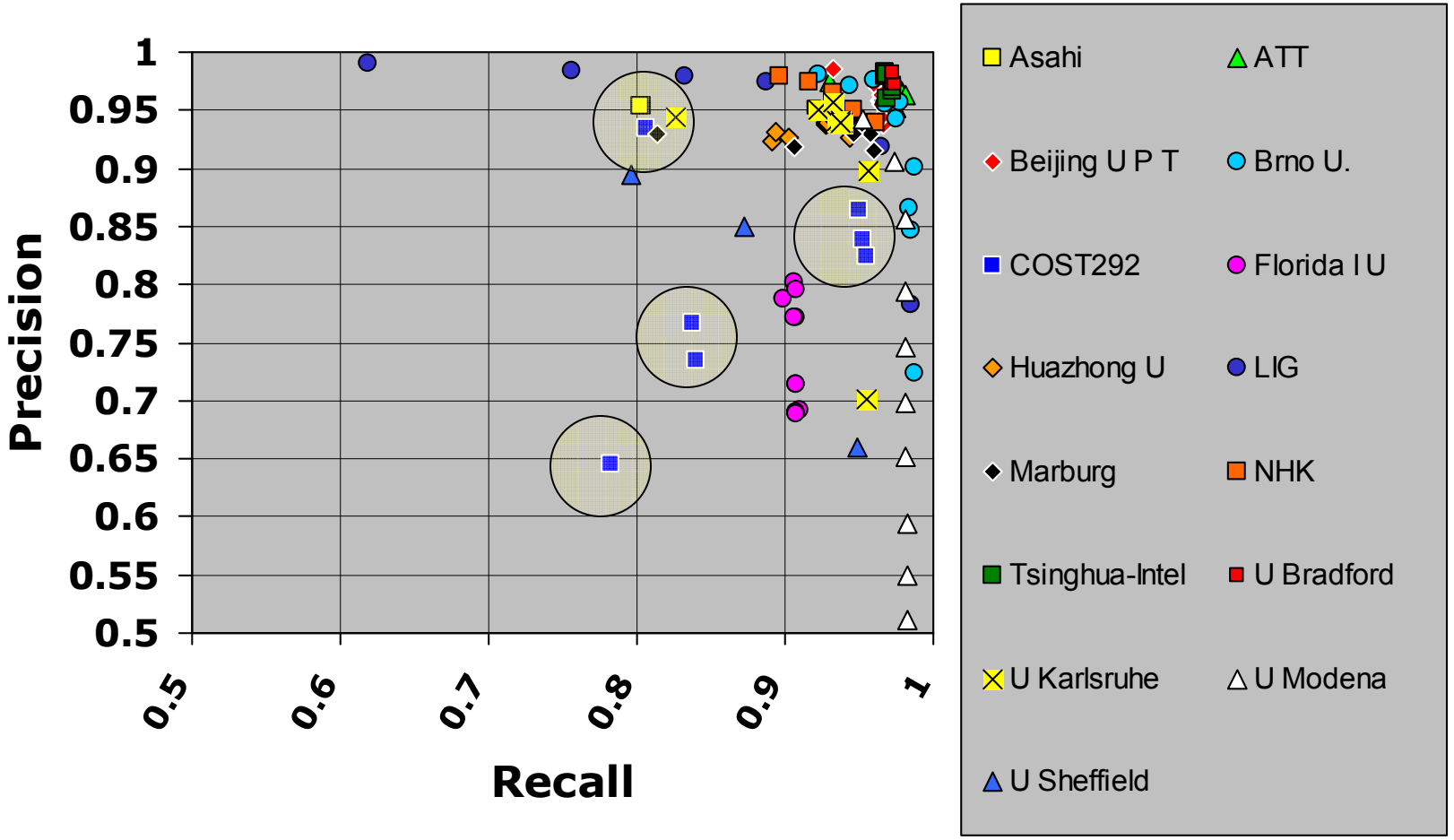
Gradual transitions (Frame-P & -R)



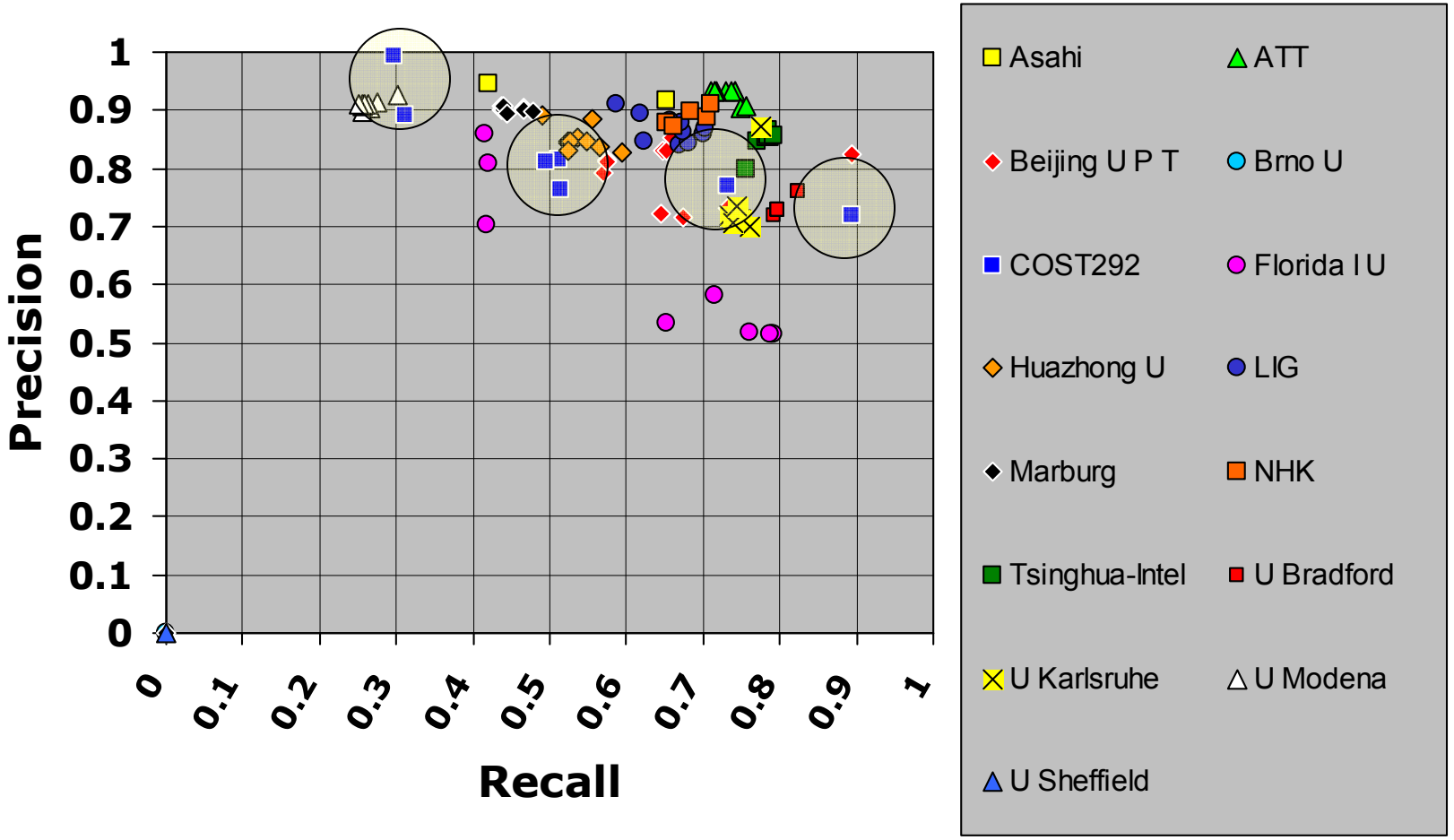
5. COST292 Team

- Large EU-funded collaboration (36 authors)
- Merges four fully independent detectors from University of Brescia, TU Delft, METU, QMUL.
- Uses blind integration – so merge can be done without knowledge of which detectors are used, though uses confidence measures on runs
- Submitted runs a combination of independent runs, and various mergings;
- Merging improves individual performances

Cuts (zoomed)


Precision
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.5
0.6
0.7
0.8
0.9
1
Recall
Asahi
ATT
Beijing U P T
Brno U.
COST292
Florida I U
Huazhong U
LIG
Marburg
NHK
Tsinghua-Intel
U Bradford
U Karlsruhe
U Modena
U Sheffield

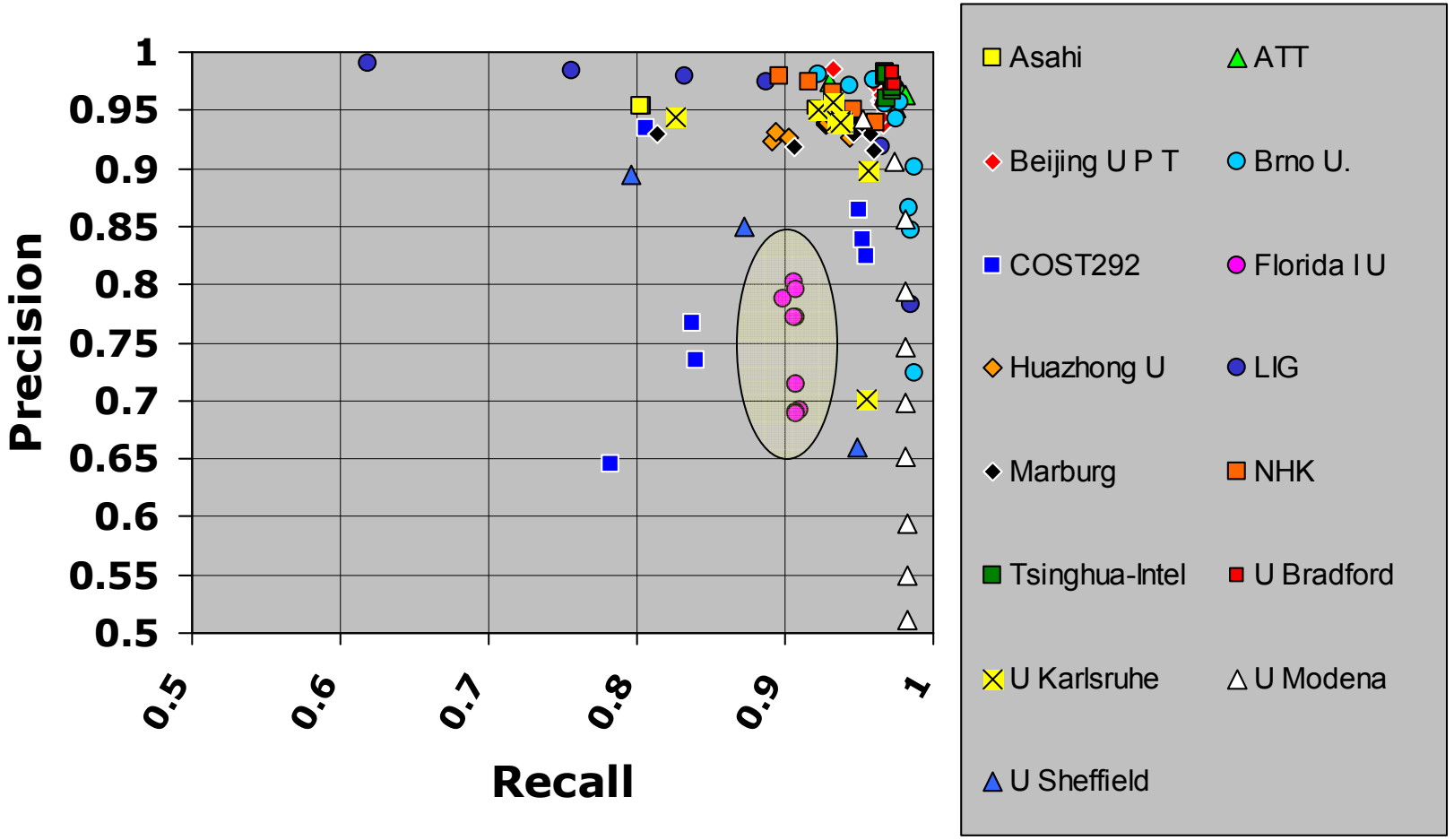
Gradual transitions (Frame-P & -R)



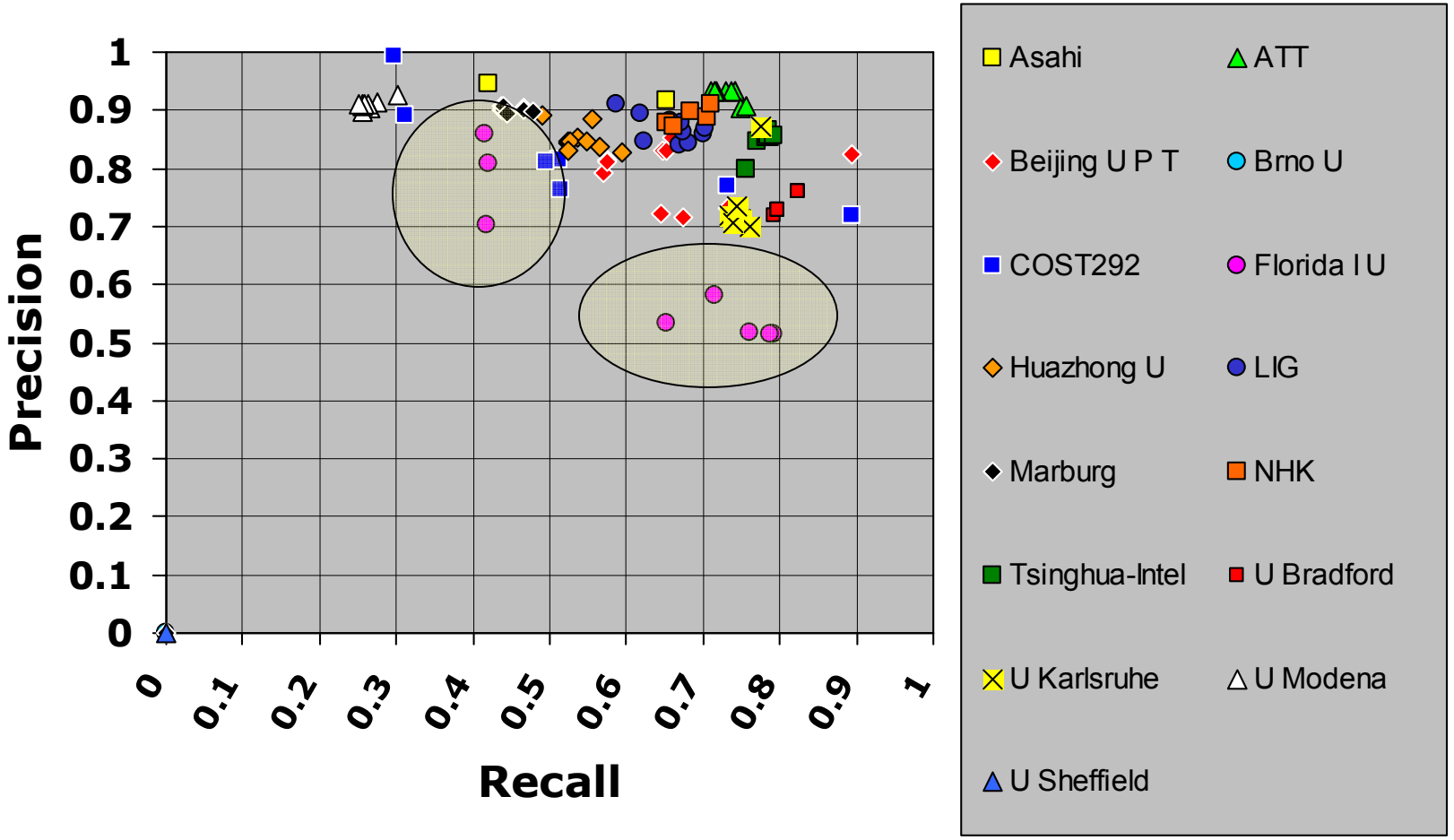
6. Florida International University, FIU-UM

- Late/missing paper ?

Cuts (zoomed)


Precision
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.5
0.6
0.7
0.8
0.9
1
Recall
Asahi
ATT
Beijing U P T
Brno U.
COST292
Florida I U
Huazhong U
LIG
Marburg
NHK
Tsinghua-Intel
U Bradford
U Karlsruhe
U Modena
U Sheffield

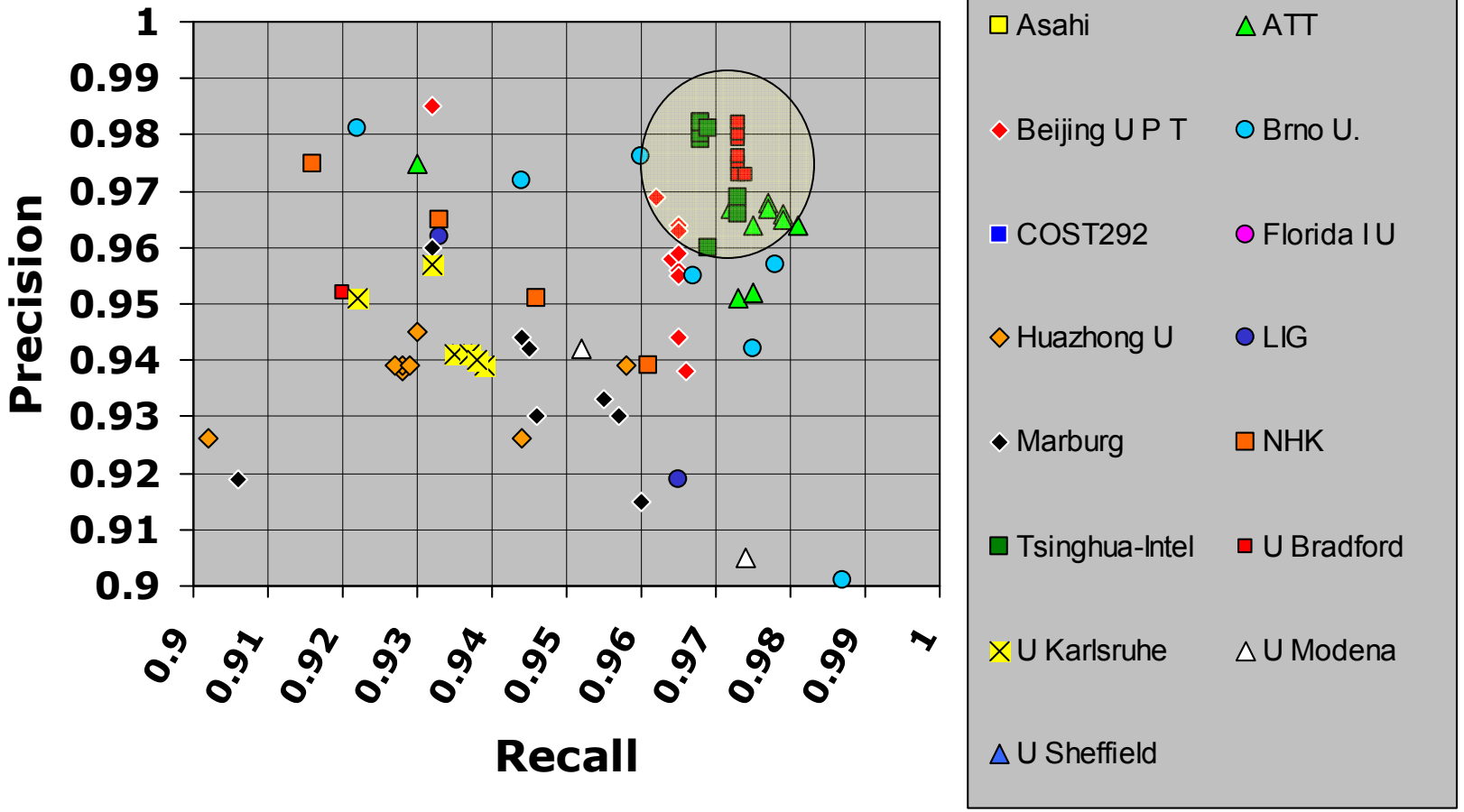
Gradual transitions (Frame-P & -R)



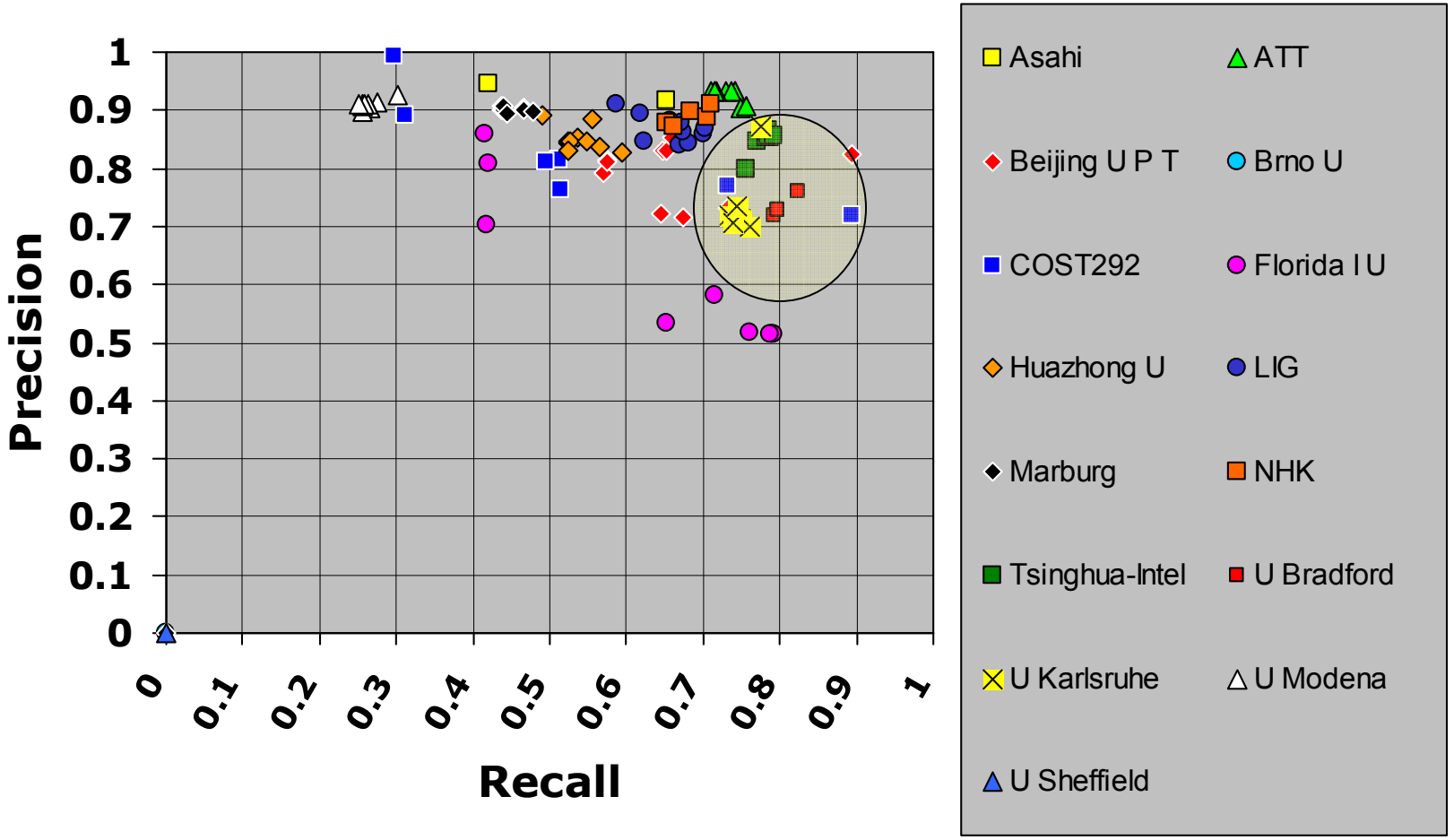
7. U. of Bradford (& Fraunhofer Institute IAIS ?)

- Work in compressed domain, so 5x RT
- Two-phase approach:
  - Extract local indicators from macroblocs to determine candidate cuts via rules
  - Then compute frame-frame similarity using global features

Cuts (zoomed again)



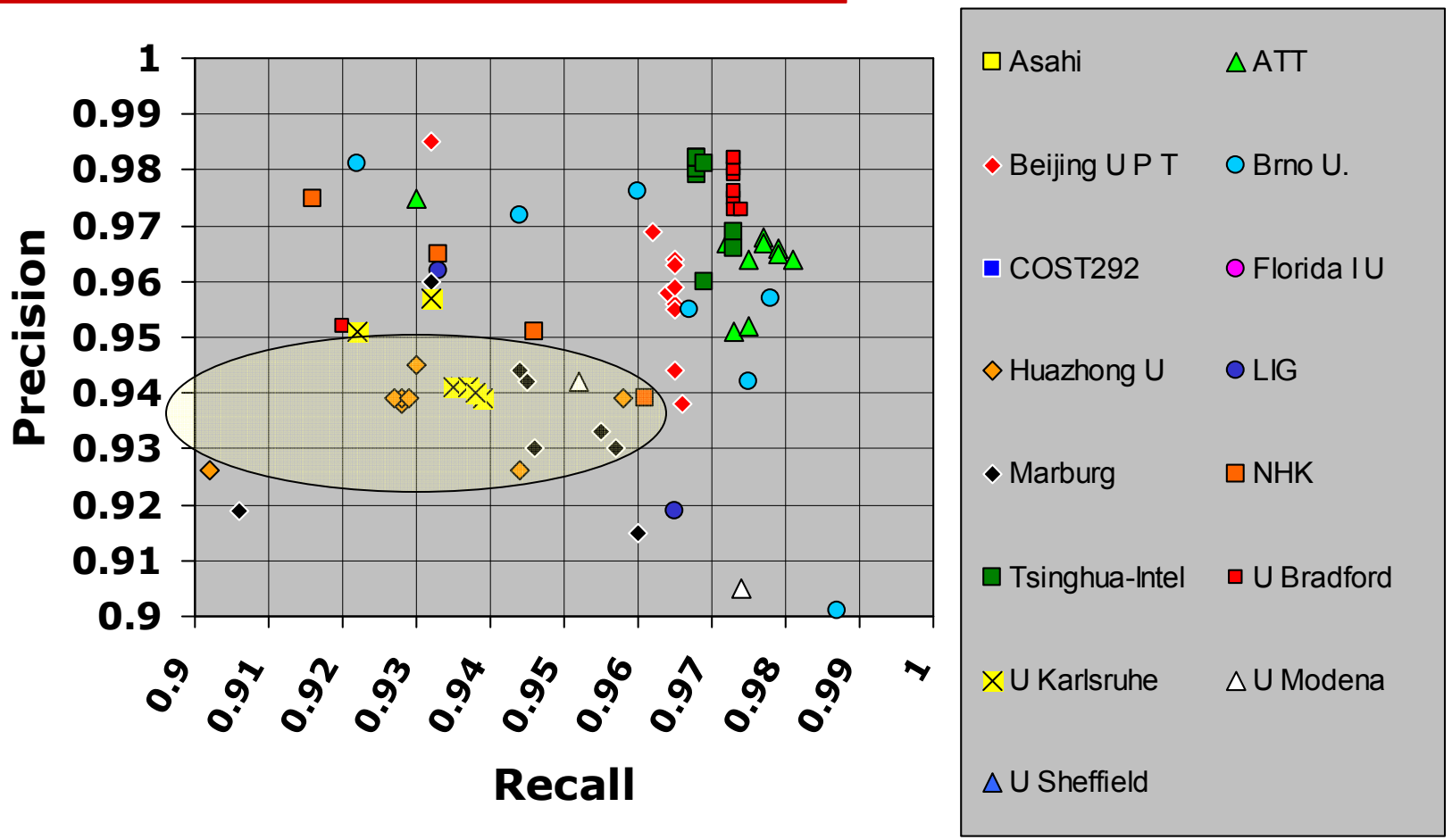
Gradual transitions (Frame-P & -R)



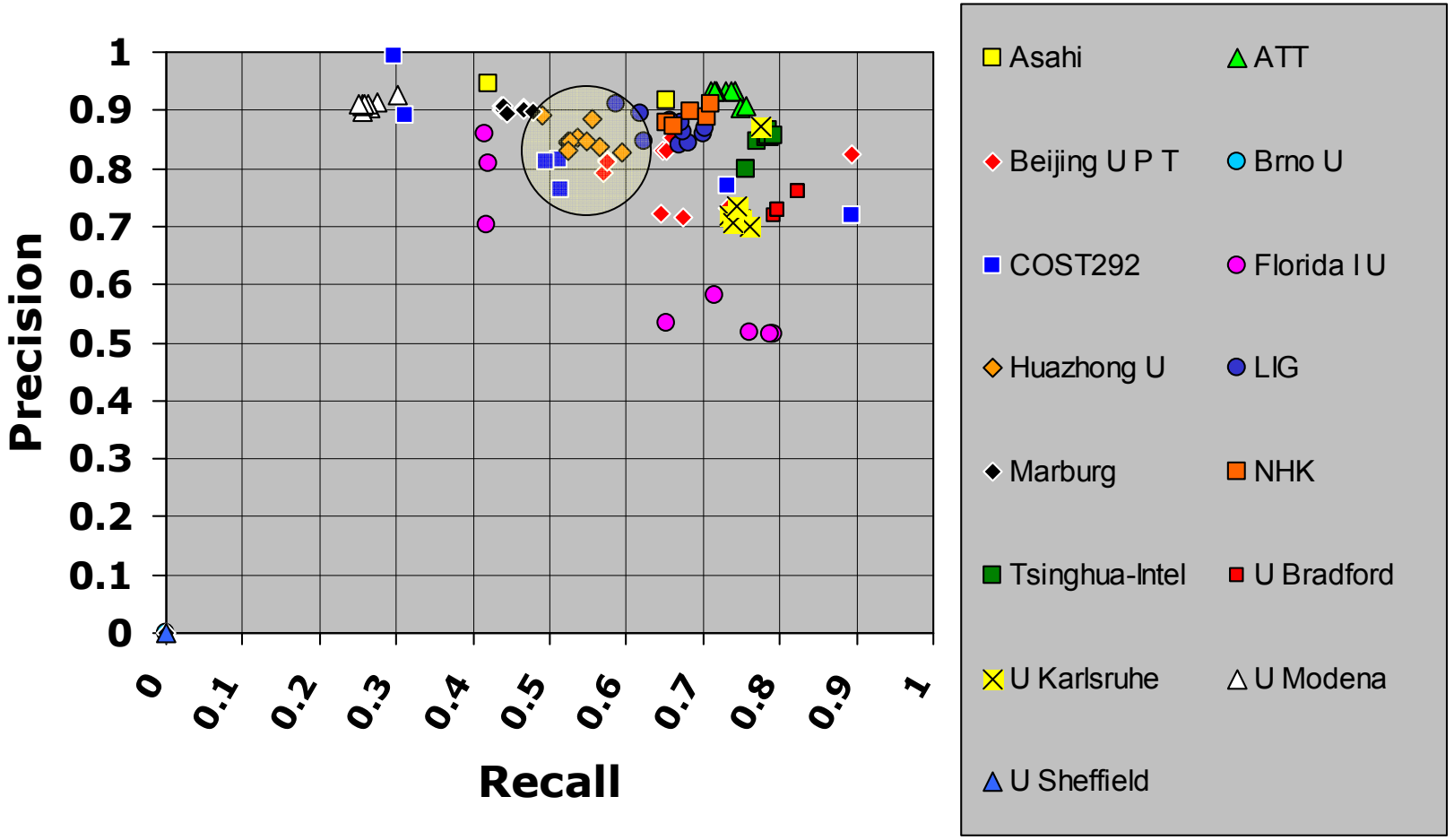
8. HuaZhong Univ. of Science and Tech.

- Late/missing paper - AGAIN !

Cuts (zoomed again)



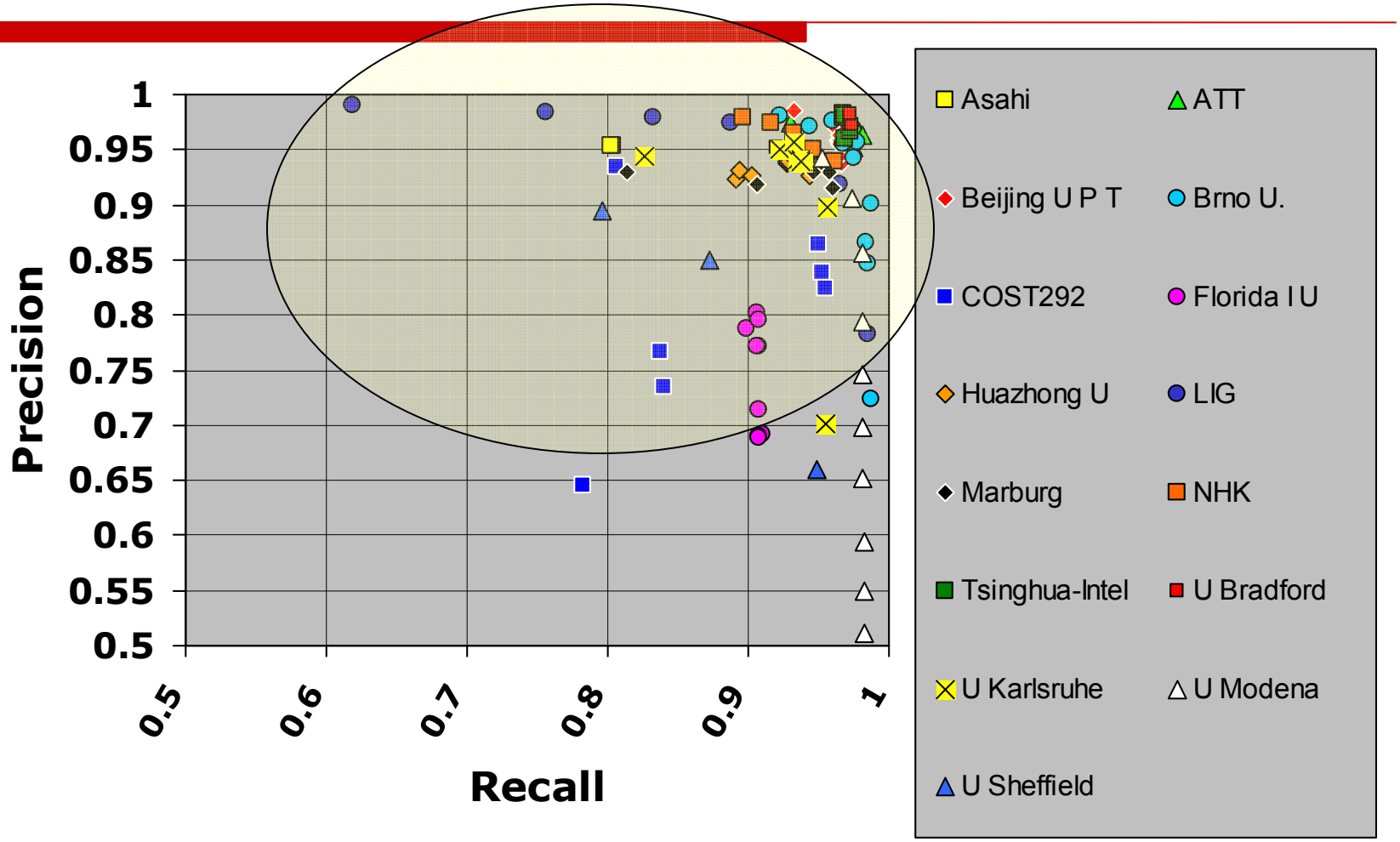
Gradual transitions (Frame-P & -R)



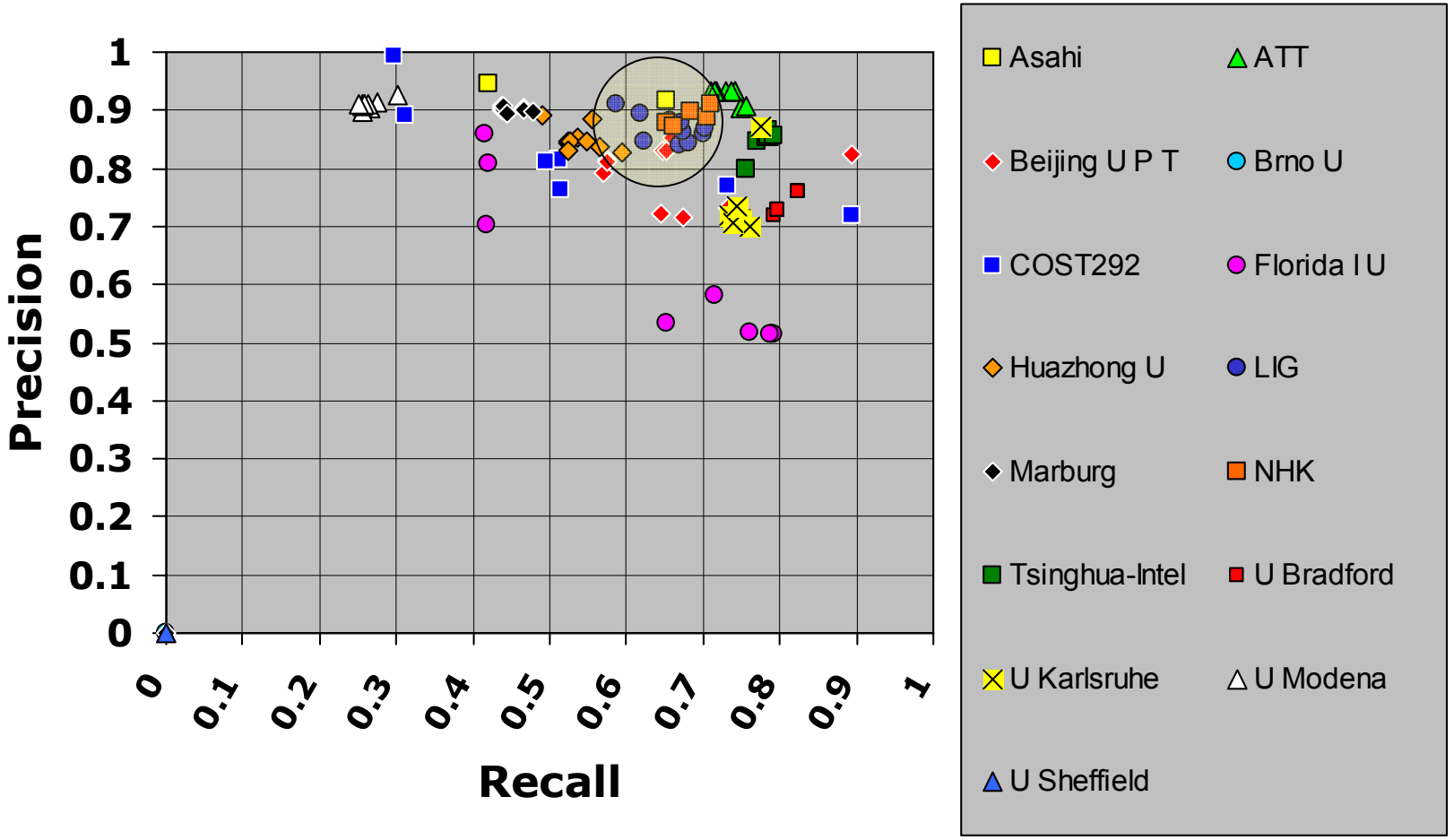
9. LIG (Laboratoire d'Informatique de Grenoble)

- … formerly CLIPS-IMAG, Grenoble
- Repeated 2006 system on 2007, which in turn was the same system as in 2004 and 2005, no new training.
  - Cut detection by image difference with motion compensation and photographic flash detection.
  - GTs by comparing norms of the first and second temporal derivatives of the images.
- Provides a normalising factor (thanks Georges)

Cuts (zoomed)


Precision
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.5
0.6
0.7
0.8
0.9
1
Recall
Asahi
ATT
Beijing U P T
Brno U.
COST292
Florida I U
Huazhong U
LIG
Marburg
NHK
Tsinghua-Intel
U Bradford
U Karlsruhe
U Modena
U Sheffield

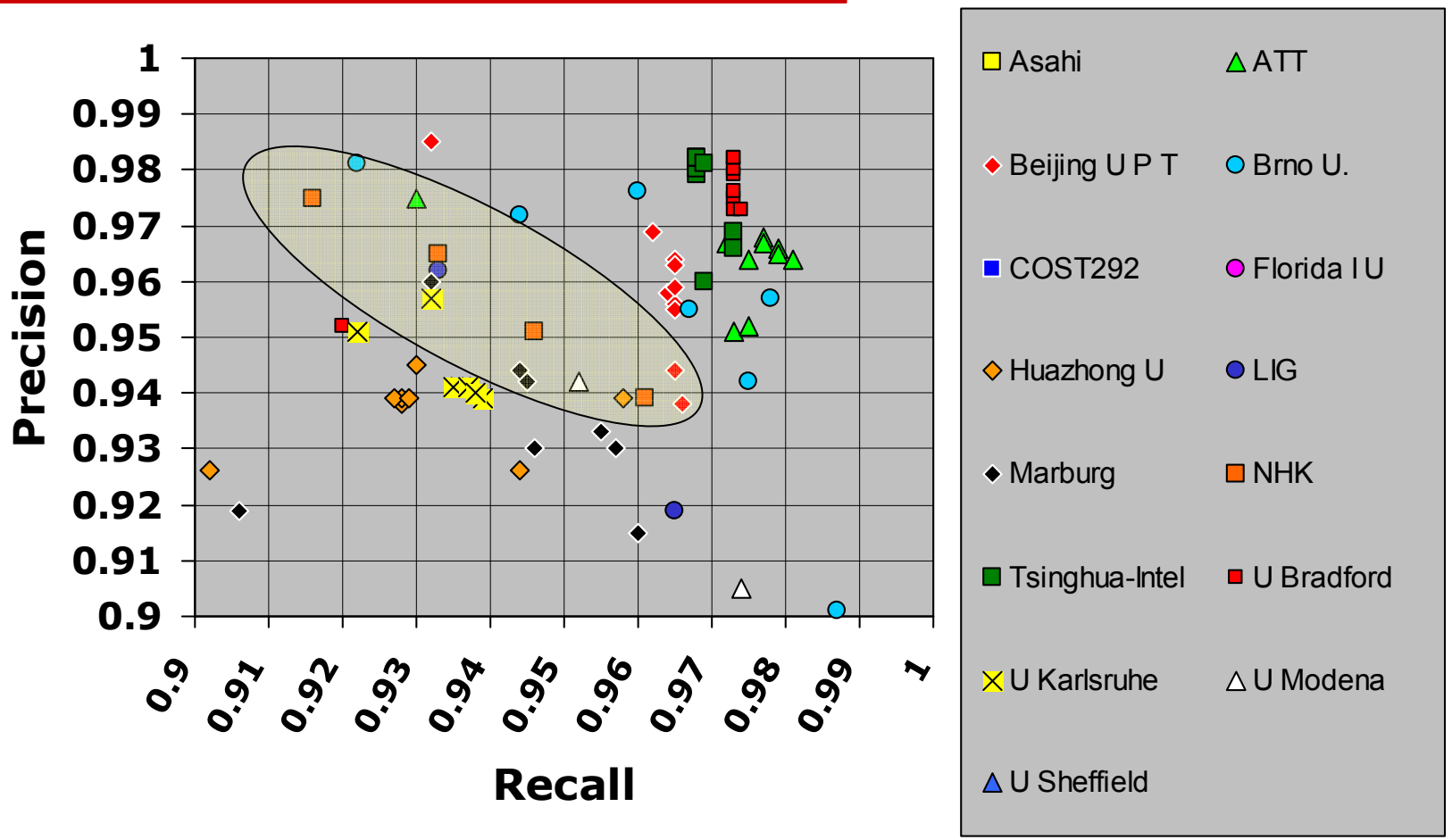
Gradual transitions (Frame-P & -R)



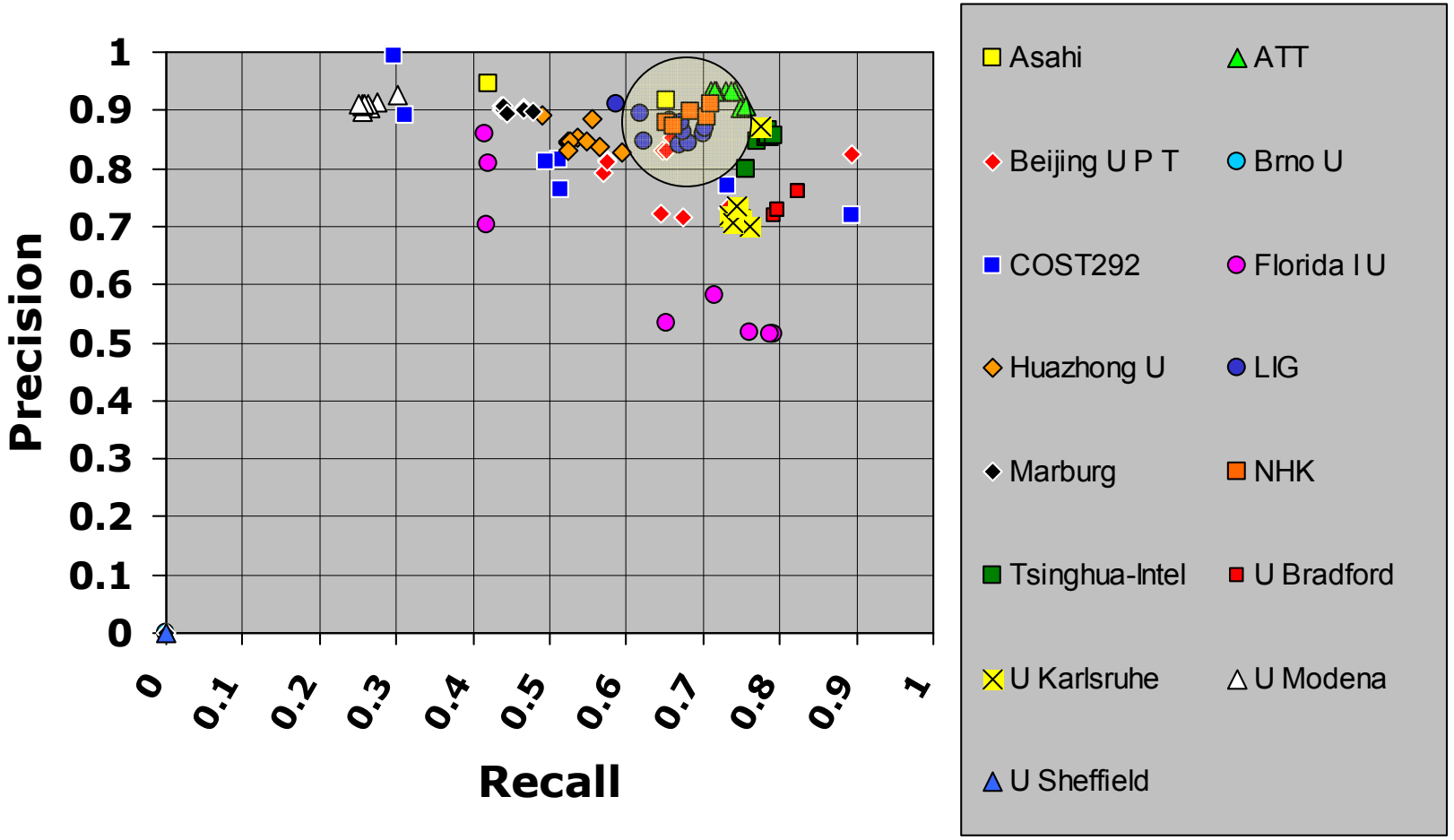
10. NHK Science and Technical Research Laboratories

- Early preprocessing stage filters out obvious non-SBs allowing focus on areas likely to contain SBs
- Very fast, 1/123 of RT
- Independent dissolve, cut, long dissolve and fade in / fade out detectors, fused together
- Gradual and sudden boundaries checked for in parallel. Also to improve performance, there is an early processing stage which filters out obvious non shot boundary candidates.

Cuts (zoomed again)



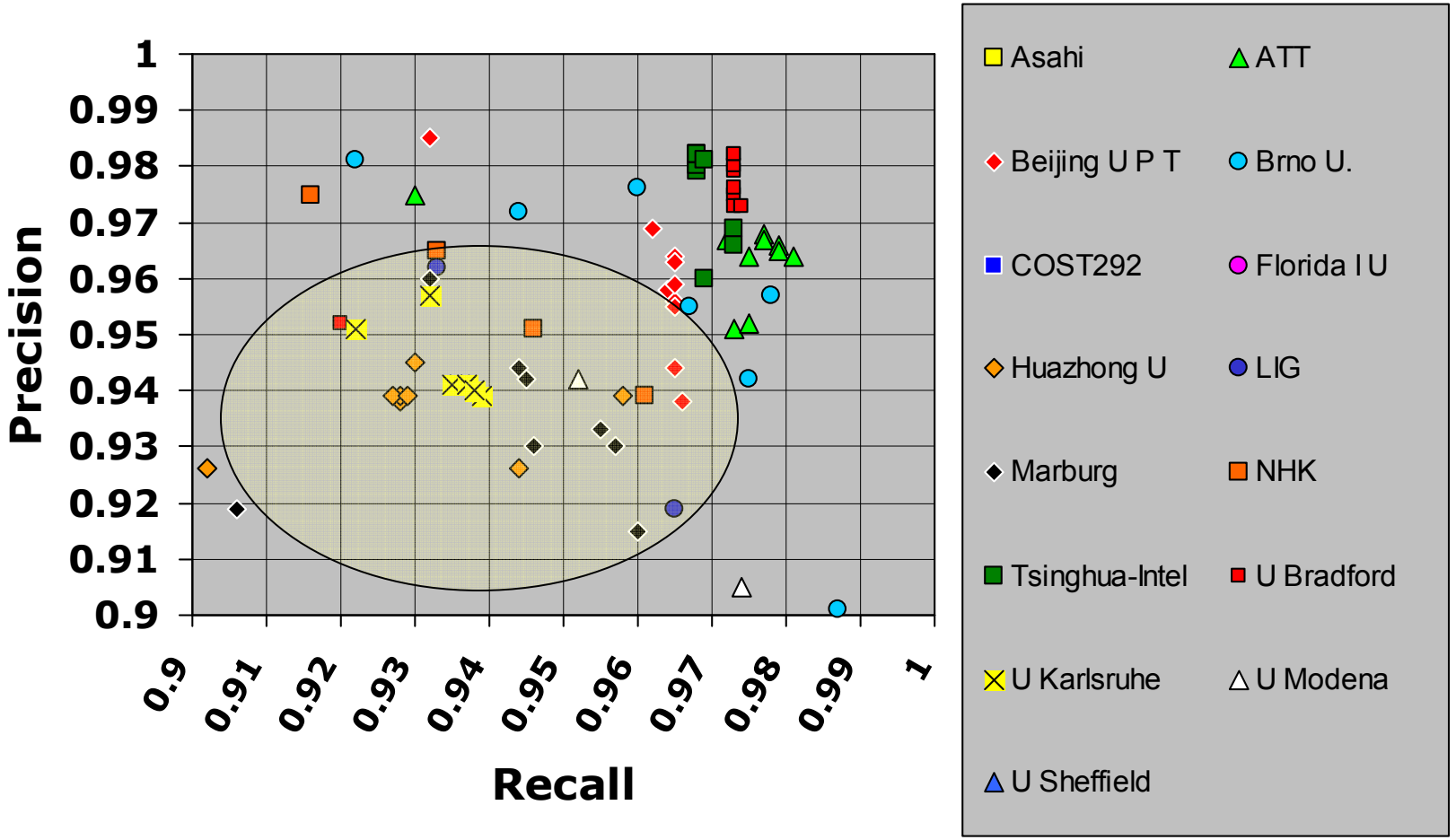
Gradual transitions (Frame-P & -R)



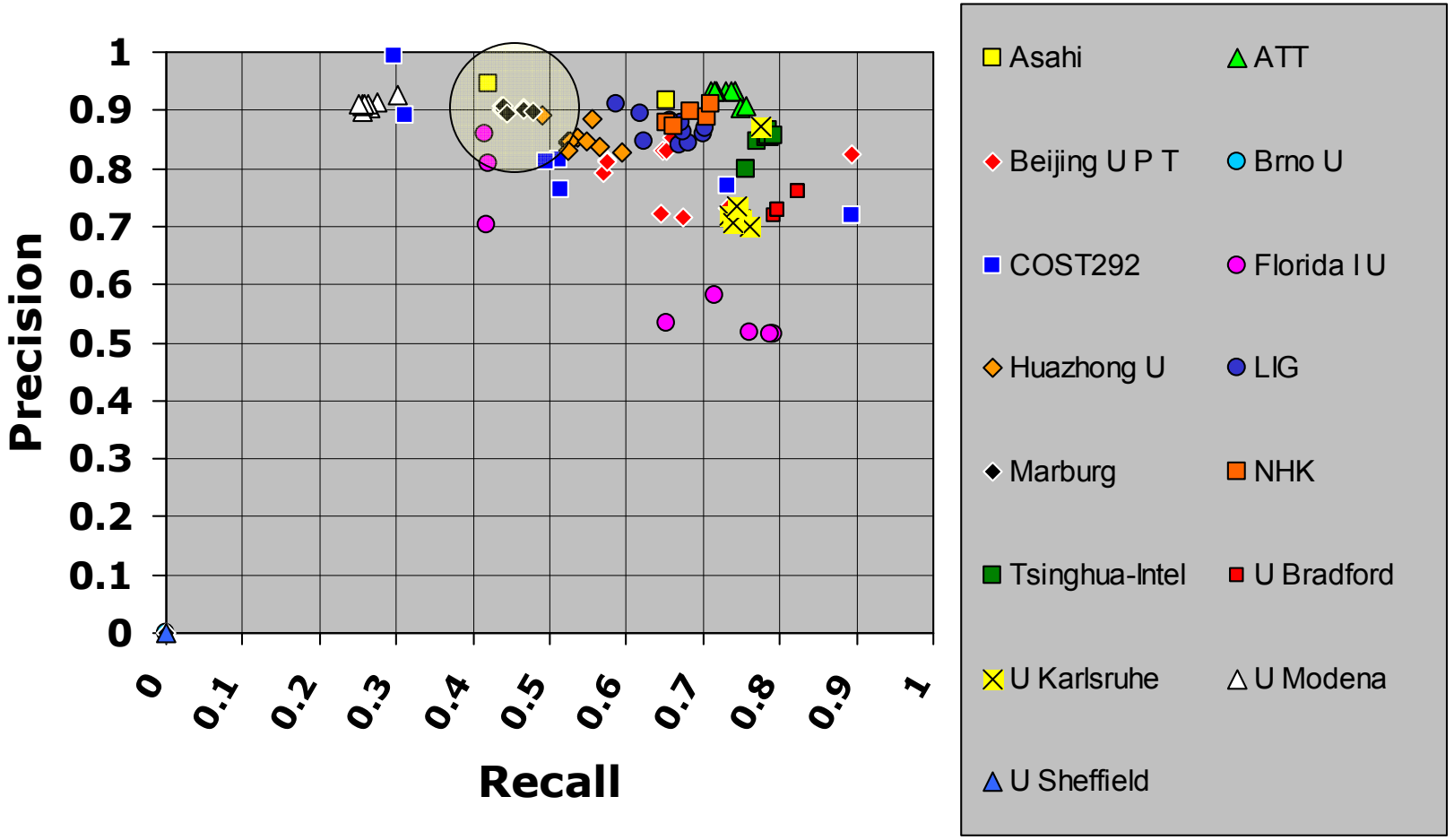
11. Universities of Siegen and of Marburg

- Improved on 2006 by including:
  - metric selection for cut detection based on the evaluation of a clustering result
  - self-supervised learning.
  - false alarm removal method for gradual transitions using camera motion estimation

Cuts (zoomed again)



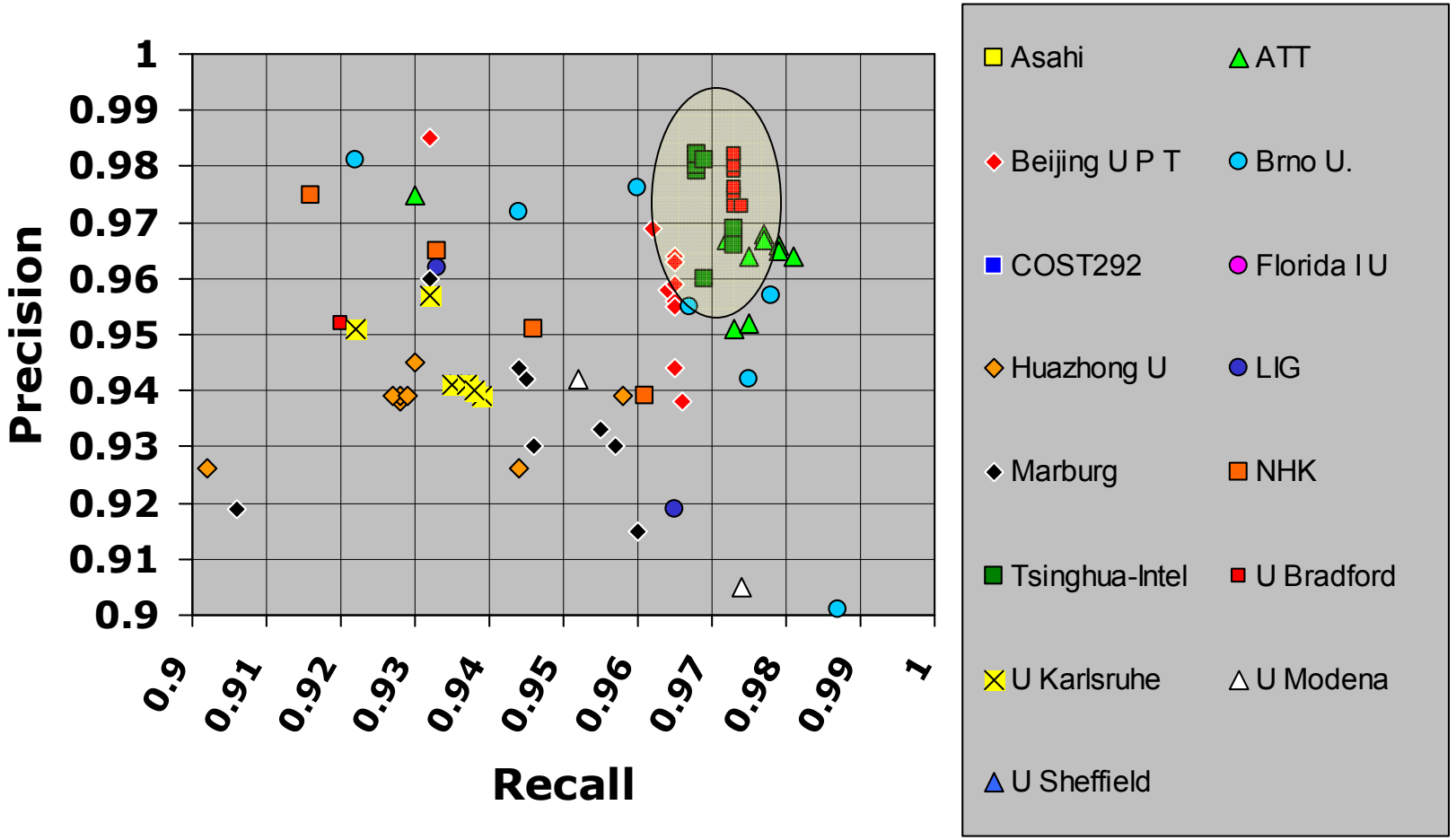
Gradual transitions (Frame-P & -R)



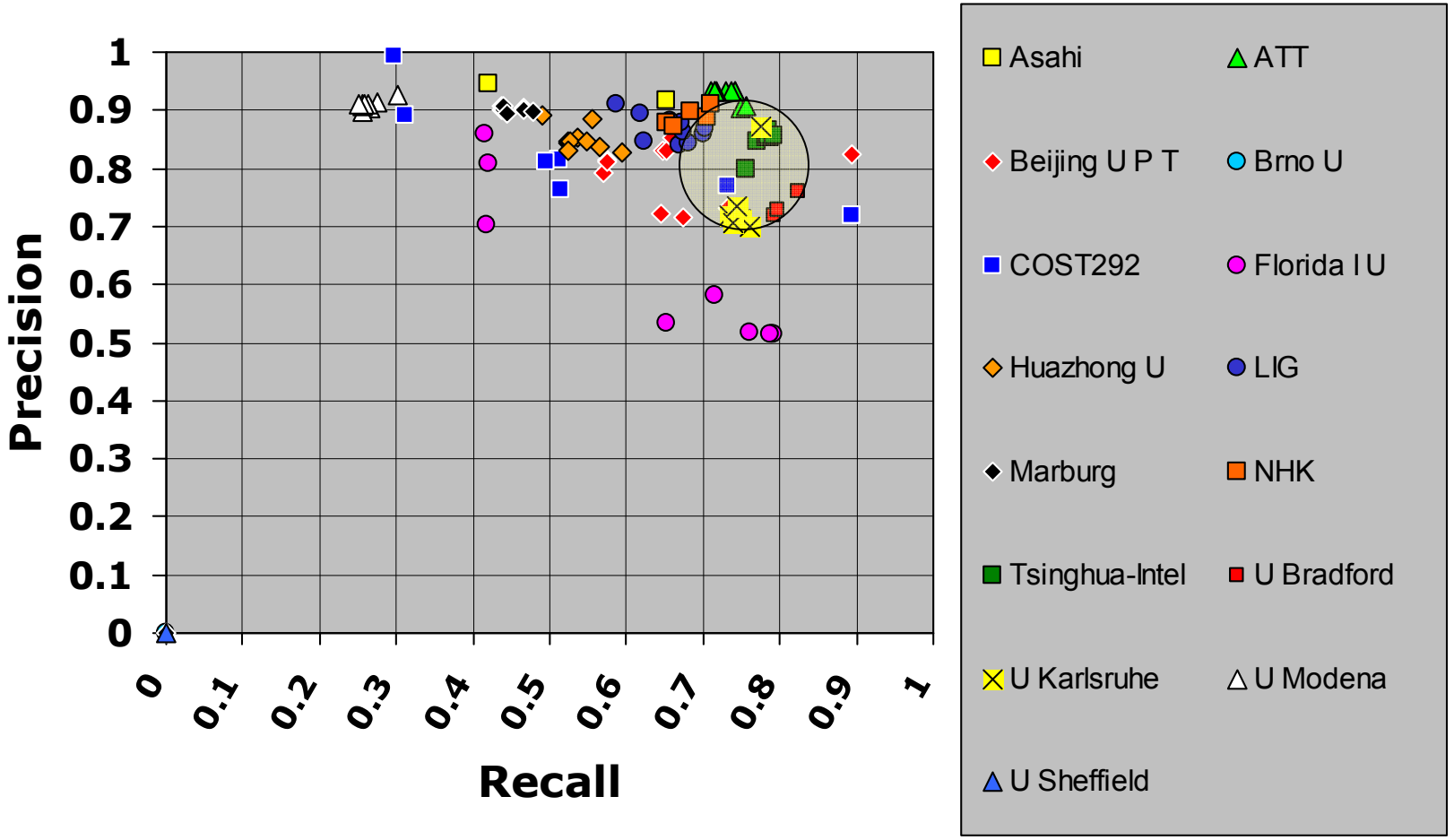
University & Intel China Research Center)

- Improvement on 2006 submission (which was already very good !) and on 2005, so succession
- 2006 …
  - Improvements in the detection of FOIs, flashes and short GTs using an FOI detector, independent CUT and GT detection, and targeted the transitions in video-in-video, which are not SBs;
- 2007 …
  - Different parameters for cut detector and gradual transition detectors
  - Motion detection to remove false alarms in gradual transition detection
  - SIFT-based post-processing for cut and gradual transition detectors.

Cuts (zoomed again)



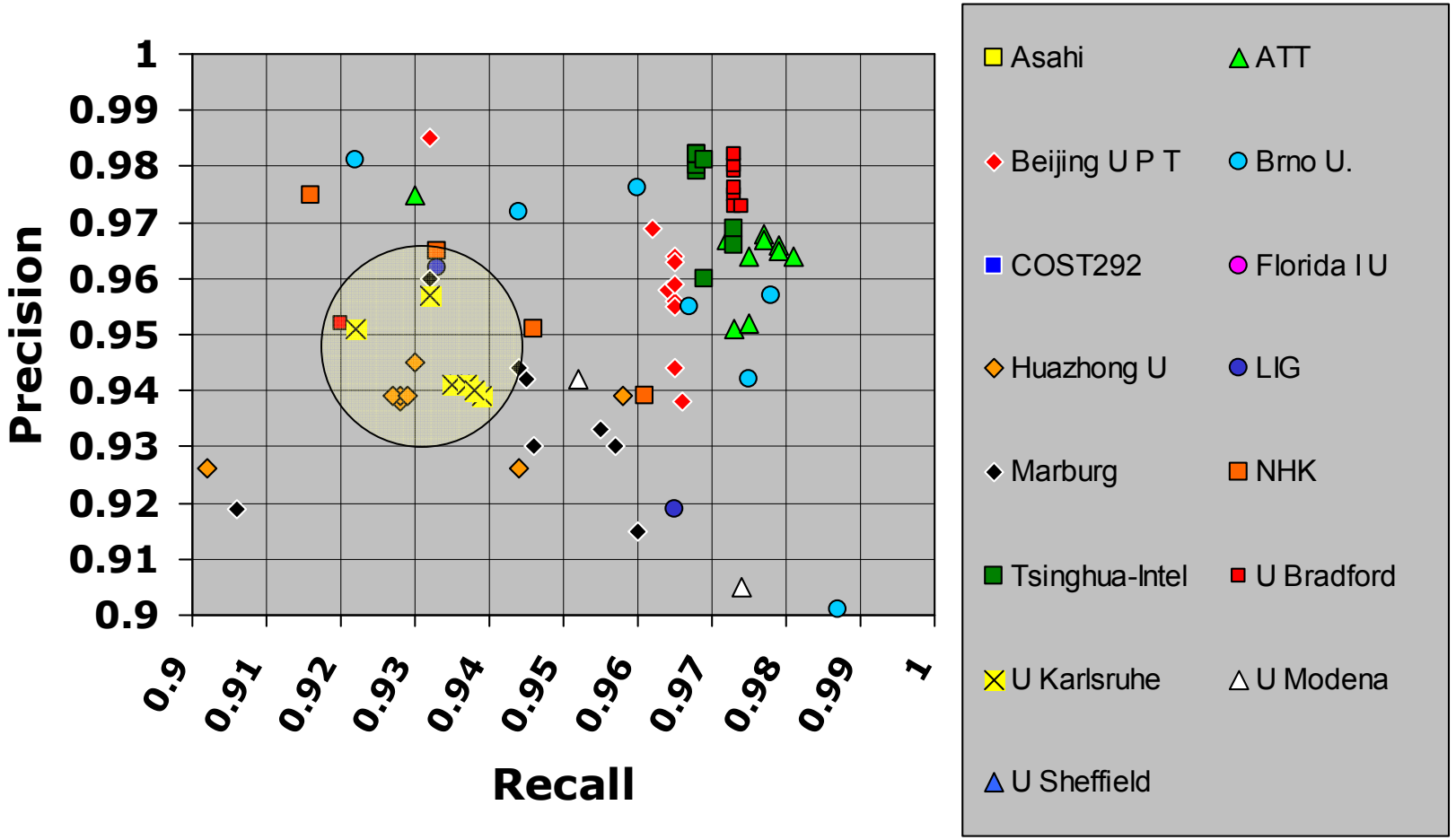
Gradual transitions (Frame-P & -R)



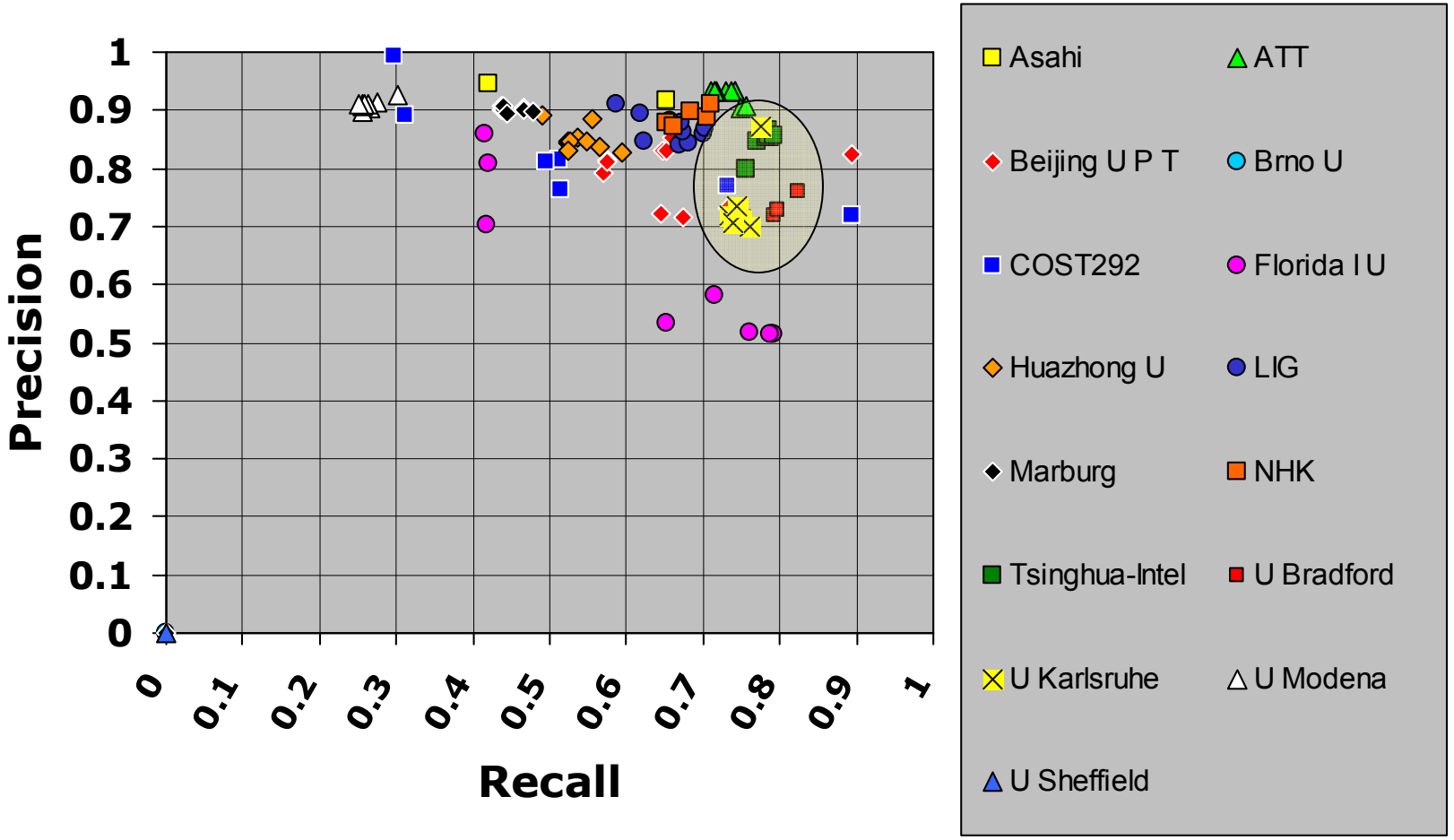
13. University of Karlsruhe (TH)

- 4 independent detectors: cut, fast dissolve, dissolve (using histogram comparisons and attempted dissolve modeling) and fade-out / fade-in (using deviation of pixel intensities).
- Certain heuristic approaches to dissolve proved ineffective.
- Good performance for cuts or fast dissolves but worse for longer transitions. Suggests moving forward using an SVM based approach.

Cuts (zoomed again)



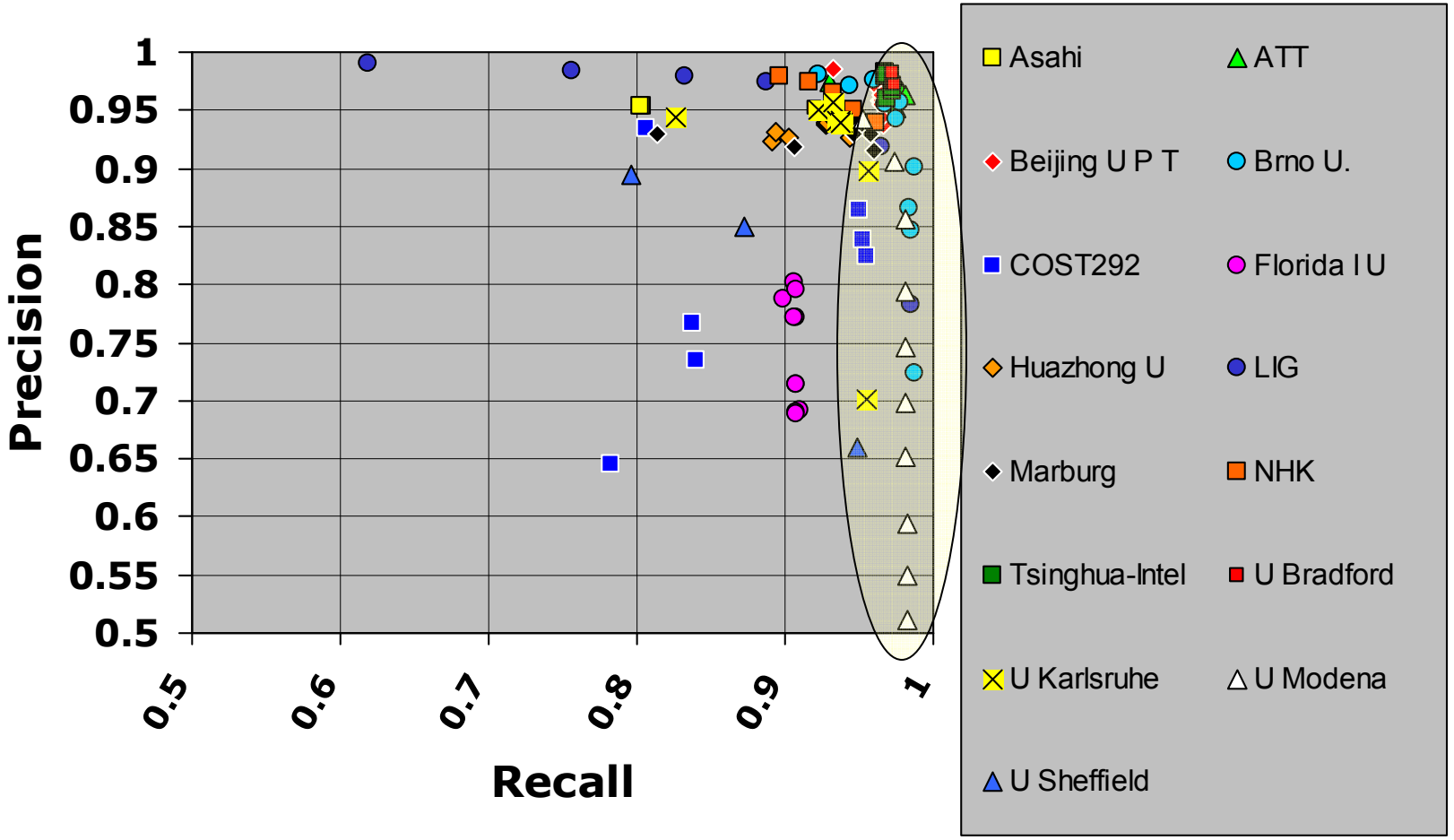
Gradual transitions (Frame-P & -R)



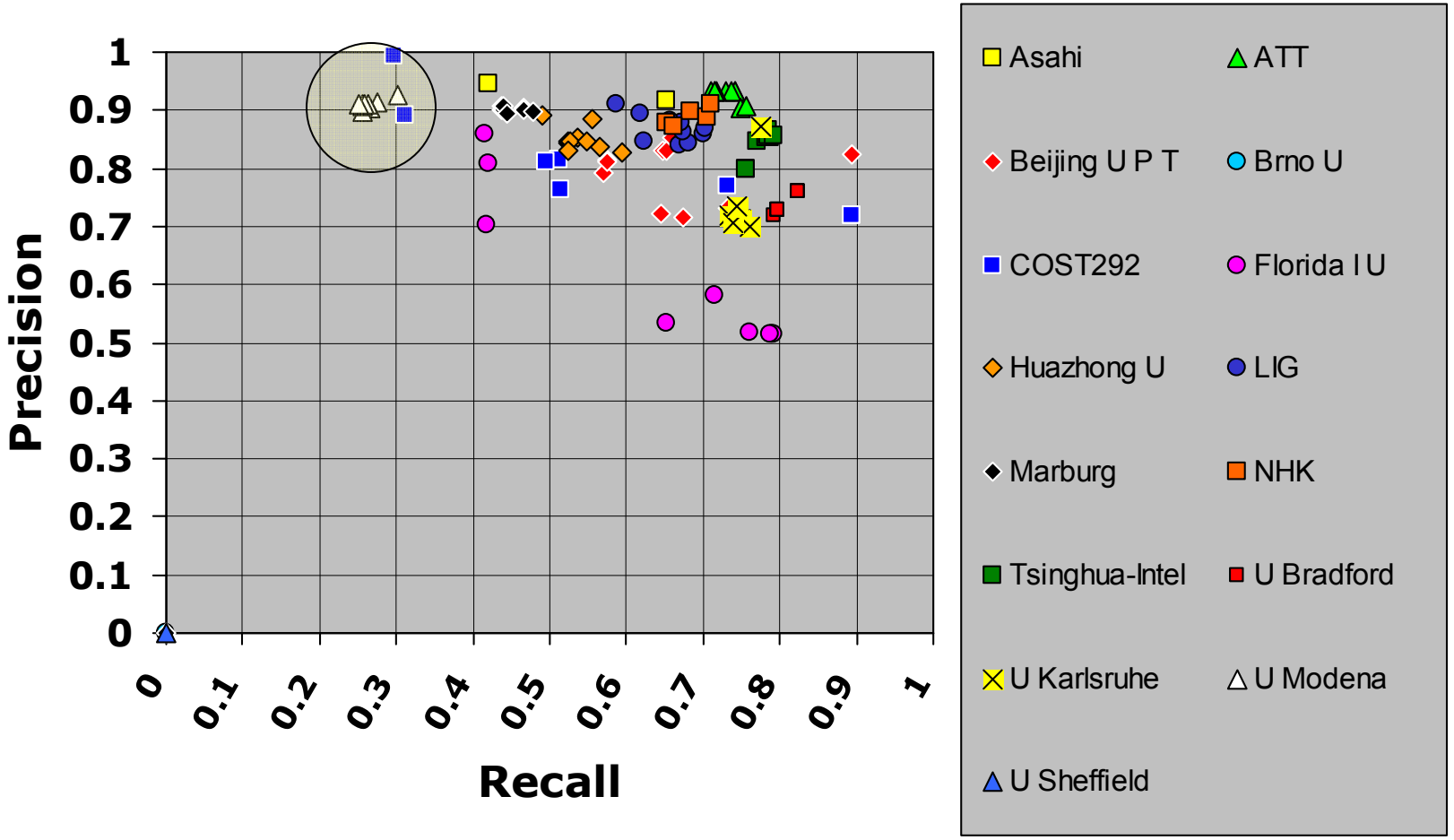
14. University of Modena and Reggio Emilia

- Late/missing paper

Cuts (zoomed)


Precision
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.5
0.6
0.7
0.8
0.9
1
Recall
Asahi
ATT
Beijing U P T
Brno U.
COST292
Florida I U
Huazhong U
LIG
Marburg
NHK
Tsinghua-Intel
U Bradford
U Karlsruhe
U Modena
U Sheffield

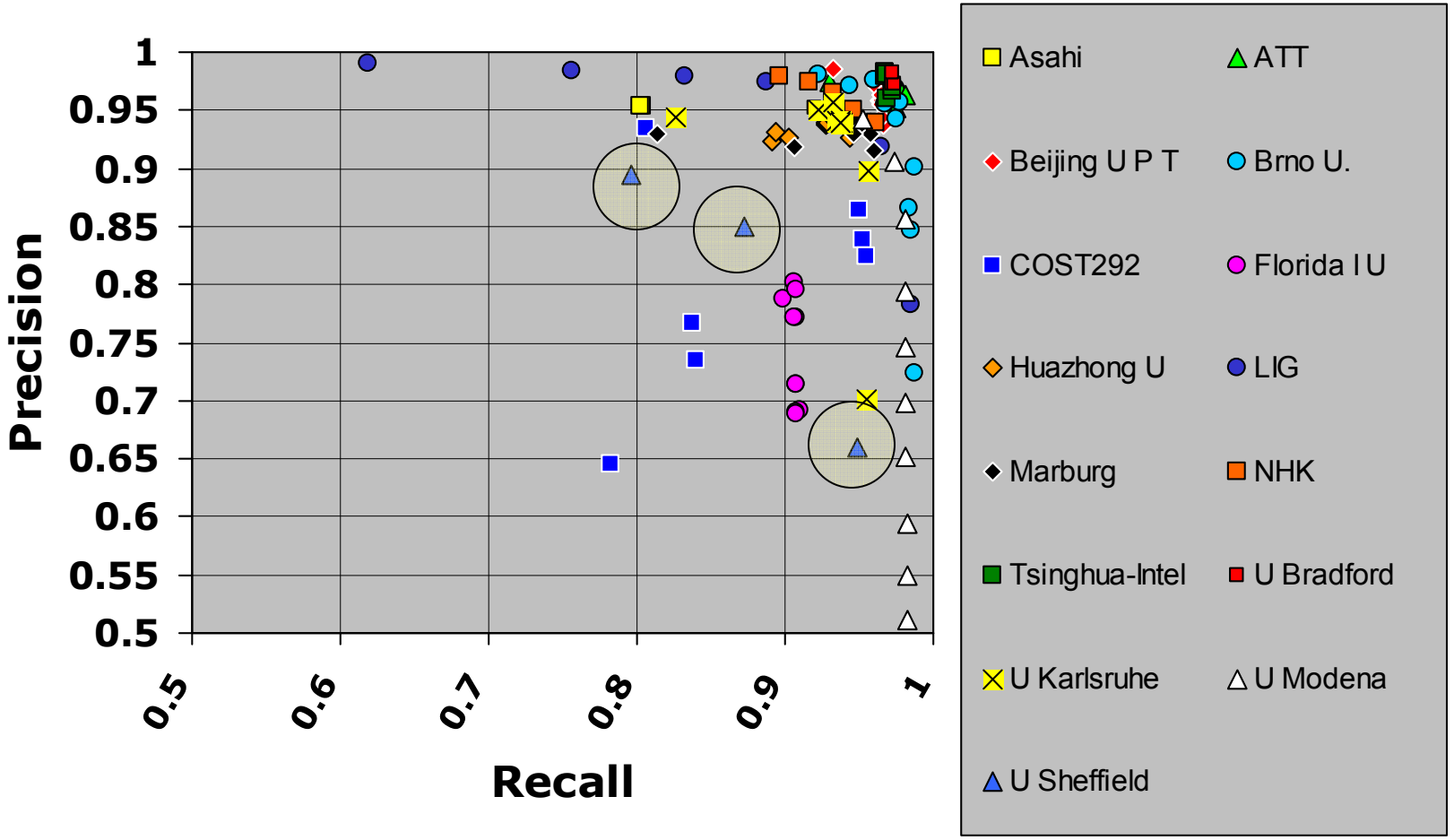
Gradual transitions (Frame-P & -R)



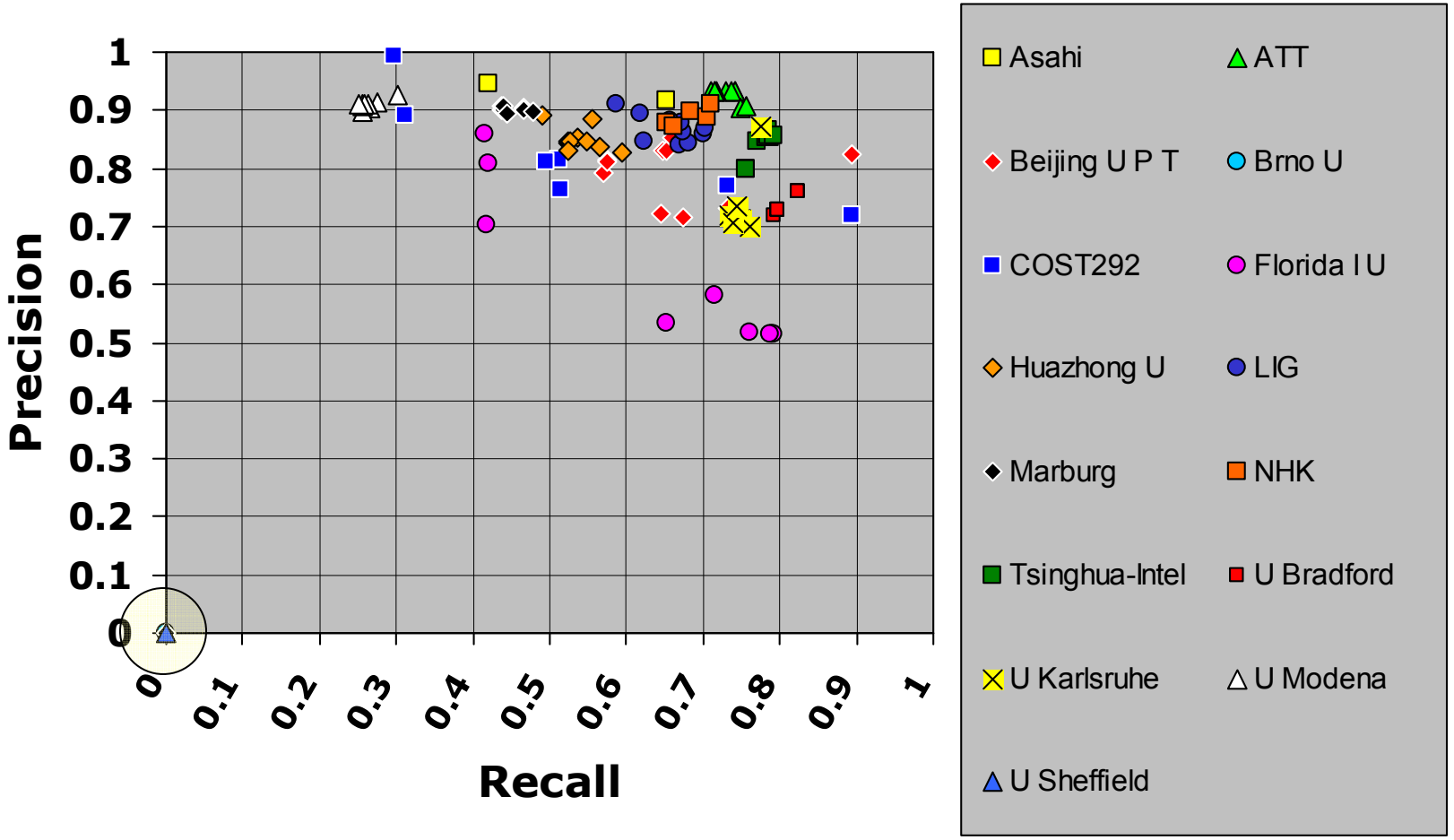
15. University of Sheffield

- Exclusive-OR of successive frames in grayscale to detect cut boundaries.
- Novel, but not so good for graduals !
- Might take approach of others of fusing multiple independent detectors.

Cuts (zoomed)


Precision
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.5
0.6
0.7
0.8
0.9
1
Recall
Asahi
ATT
Beijing U P T
Brno U.
COST292
Florida I U
Huazhong U
LIG
Marburg
NHK
Tsinghua-Intel
U Bradford
U Karlsruhe
U Modena
U Sheffield

Gradual transitions (Frame-P & -R)



Observations

- Last year we said …
  - Excellent performance on cuts and graduals **despite more difficult data**
  - Good effectiveness achievable at significantly less than realtime
  - Despite the continued introduction of novel approaches, novelty =/= improvement with best systems achieving very good performance even on most gradual transitions
- This year … still true
- Even more focus on fusing independent detectors
- In the absence of an application that requires better performance than best systems achieve, it's time to declare victory!
- Experiments can continue outside of TRECVID
  - SB test/training data (2003-2007) distributed to many groups
  - Truth data and evaluation software available from TRECVID website

- Thanks to Jonathan Lasko for each year's consistent annotation

- For citing, there is no TRECVID SBD paper so cite the TRECVID paper, not the URL or notebook (check the website).