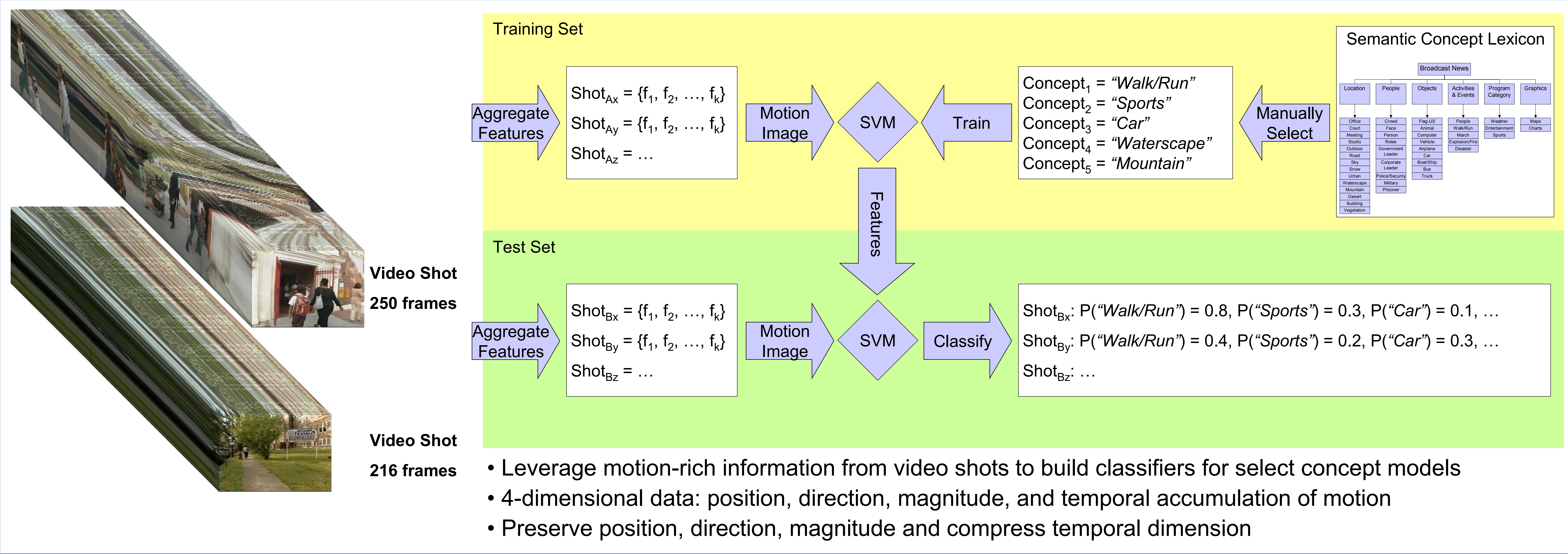


Classification of Video Events using 4-dimensional time-compressed Motion Features

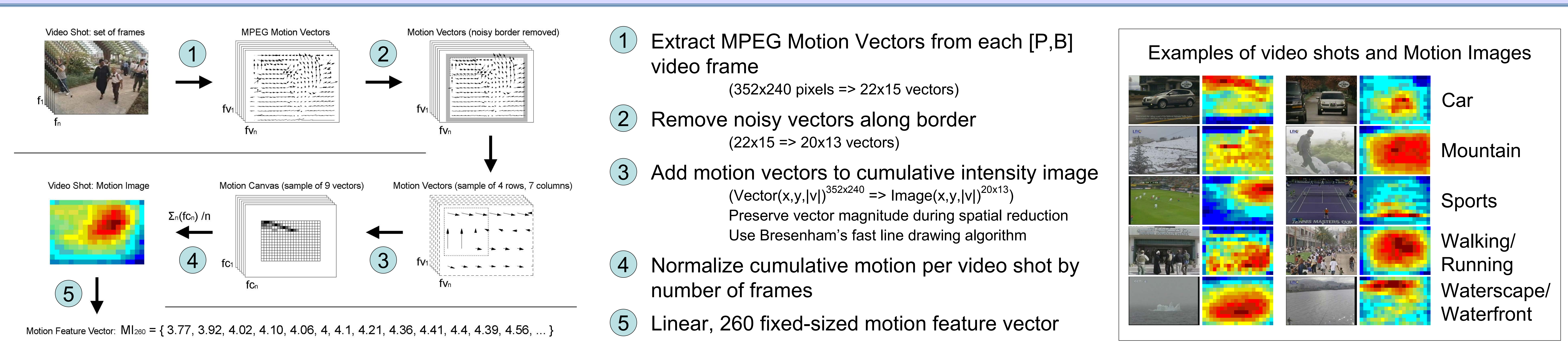
Alexander Haubold, Columbia University

Milind Naphade, IBM TJ Watson Research Center

Problem



Approach



Evaluation

Performance among Similar Features

Comparison to 3 low-level features over 5 selected concepts

- Equivalent conditions: remove noisy border

Motion Direction Histogram

- Compresses time, position, and magnitude
- Histogram bins centered at 10 degree intervals, 36 features
- Average Precision: 11.5%

Motion Magnitude Histogram

- Compresses time, position, and direction
- Histogram bins centered at magnitudes {1..30}, 30 features
- Average Precision: 7.5%

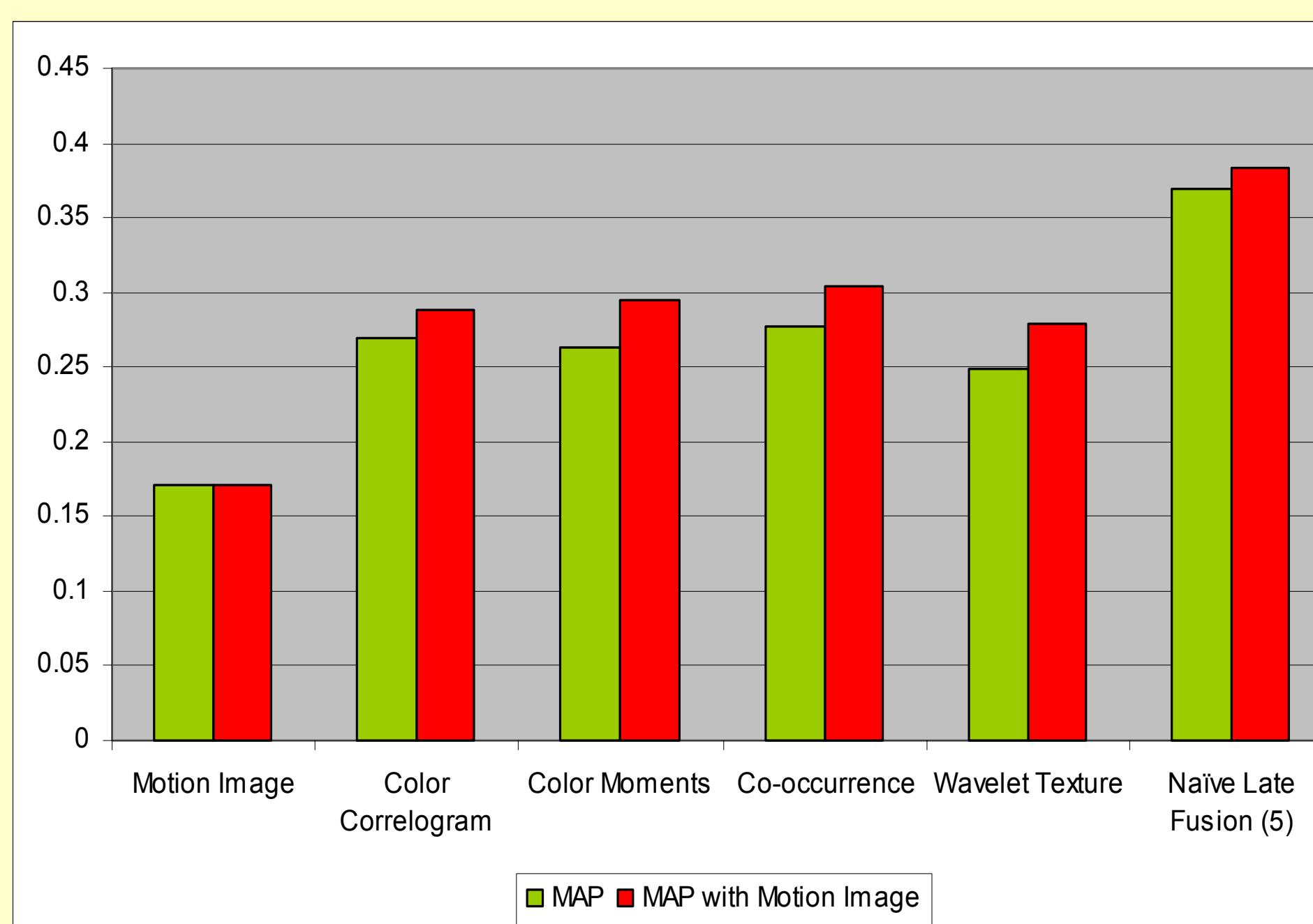
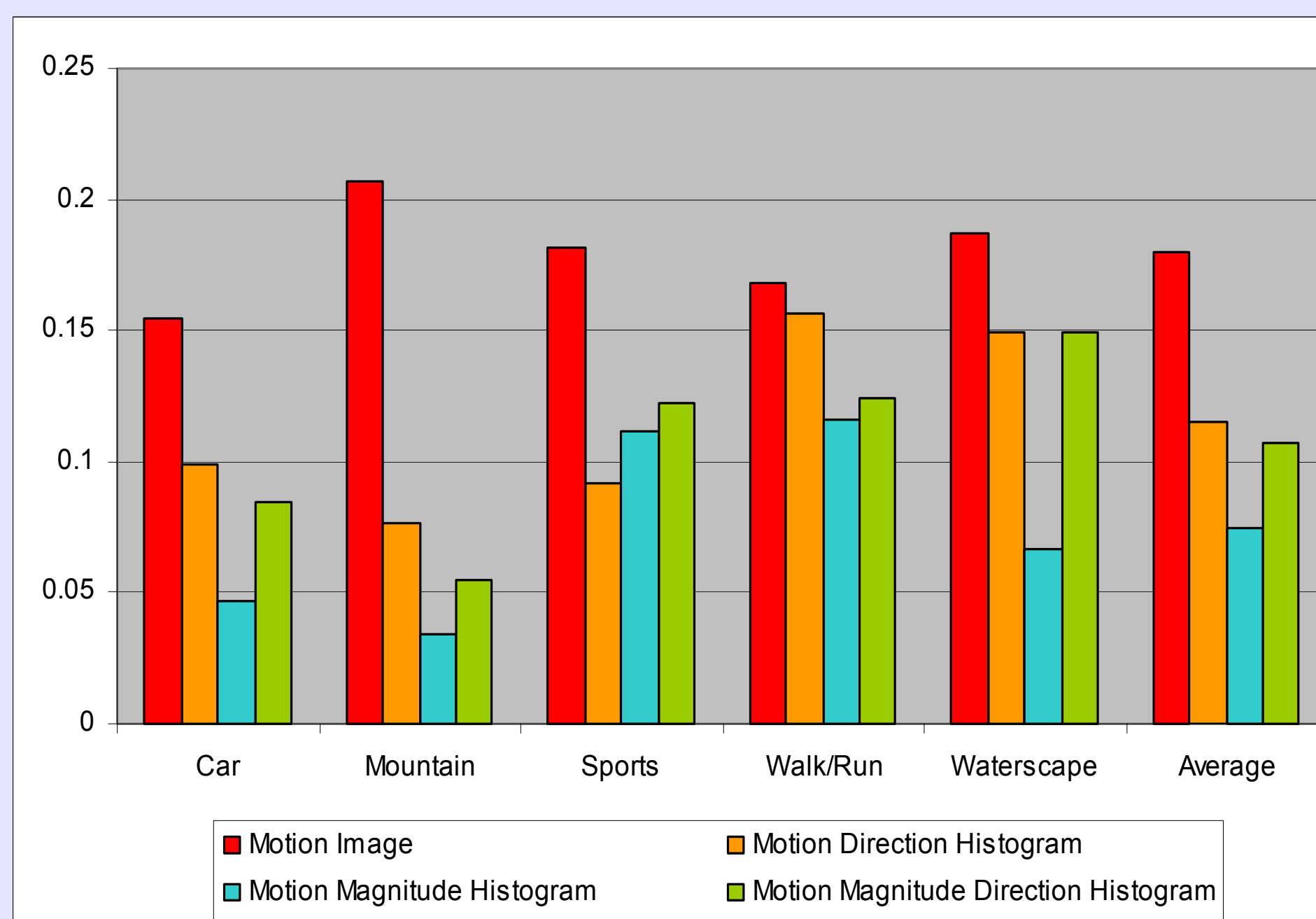
Motion Direction + Magnitude Histogram

- Compresses time and position
- Histogram is orthogonal combination of above histograms
- Average Precision: 10.7%

Motion Image

- Compresses time
- Aggregated motion vectors, 260 features
- Average Precision: 19%

	Car (%)	Mountain (%)	Sports (%)	Walk/Run (%)	Waterscape / Waterfront (%)	Average (%)	Improvement of Motion Image (%)
Motion Image	0.15463	0.20723	0.18143	0.16791	0.18897	0.179614	---
Motion Direction Histogram	0.09922	0.07659	0.09169	0.15626	0.14956	0.114664	56.6 %
Motion Magnitude Histogram	0.04672	0.0346	0.11187	0.11608	0.06615	0.075084	139.2 %
Motion Magnitude Direction Histogram	0.08495	0.05486	0.12212	0.12379	0.14969	0.107082	67.7 %



Additional Gain over other Visual Features

Comparison to 4 low-level visual features

- Compare Average Precision (AP) of feature detector with AP of Fusion with Motion Image

Color Correlogram

- Global Color and Structure, 166 features
- Fusion with Motion Image: 7.1% improvement in AP

Color Moments

- Local Color from 5x5 grid, 225 features
- Fusion with Motion Image: 12.1% improvement in AP

Co-occurrence Texture

- Global Texture, 96 features
- Fusion with Motion Image: 9.6% improvement in AP

Wavelet Texture Grid

- Local Texture from 3x3 grid, 108 features
- Fusion with Motion Image: 12.2% improvement in AP

Naïve Late Fusion

- 4 Visual Features + Motion Feature, 3.9% improvement in AP

	MAP (avg. prec.)	MAP with Motion Image (avg. prec.)	Improvement
260 dimensional Motion Image	0.170672	---	---
166 dimensional HSV Color Correlogram	0.270124	0.289172	7.1 %
225 dimensional Color Moments Grid	0.262652	0.294532	12.1 %
96 dimensional Cooccurrence Texture	0.27692	0.303466	9.6 %
108 dimensional Wavelet Texture Grid	0.248372	0.27859	12.2 %
Naive Late fusion across detectors of all five features	0.369926	0.384228	3.9 %