1 Experiments

1.1 Image recognition evaluation

1.1.1 Data set description

1.1.2 Performance of spatial pyramid matching

SVM	One-level	Two-level	Three-level
Linear	74.17	80	**87.5**
Poly	70	62.5	31.67
RBF	37.5	83.33	74.17
Sigmoid	16.67	16.67	16.67
Histogram Intersection	79.17	82.5	**85**

Table 1: Recognition accuracies (percent) of different spatial pyramids using SVM with different kernels

	One-level	Two-level	Three-level
KNN	50.83	45	48.33

Table 2: Recognition accuracies (percent) of different spatial pyramids using KNN

1.1.3 Performance of earth mover's distance

	RBF	LAP	ID	ISD
Aligned Distance	79.44	76.11	73.89	75.56
Unaligned Distance	78.89	75.56	73.33	75

Table 3: Comparison of recognition accuracy between aligned distance and unaligned distance

1.2 Video recognition

	Gaussian	Laplacian	ISD	ID	Fused kernels
Level 0	44.38 ± 2.13	44.90 ± 2.73	44.01 ± 2.13	45.36 ± 3.13	44.33 ± 2.61
Level 1 (Unaligned)	43.08 ± 3.14	43.85 ± 3.84	43.22 ± 3.11	43.85 ± 3.56	43.56 ± 3.46
Level 1 (Aligned)	43.61 ± 2.97	43.40 ± 3.18	43.46 ± 2.97	43.22 ± 3.11	44.08 ± 3.25

Table 4: Means and standard deviations (percent) of MAPs over six events at different levels using SVM with different kernels.

	Gaussian	Laplacian	ISD	ID	Fused kernels
bxx	40.20 ± 2.57	38.35 ± 2.31	39.93 ± 2.58	38.23 ± 2.08	39.34 ± 2.55
txx	44.28 ± 2.14	44.90 ± 2.73	44.01 ± 2.13	45.36 ± 3.13	44.33 ± 2.61
txc	42.15 ± 4.73	45.01 ± 3.45	43.47 ± 4.56	45.38 ± 3.20	44.11 ± 3.90
tfx	43.76 ± 2.99	44.14 ± 3.36	43.61 ± 3.03	44.05 ± 3.51	44.18 ± 3.22
tfc	43.71 ± 1.37	46.02 ± 1.84	44.93 ± 1.64	46.21 ± 1.83	45.28 ± 1.62
straightforward soft	43.54 ± 2.12	44.77 ± 2.41	43.52 ± 2.08	45.24 ± 2.47	44.79 ± 2.55
Gaussian soft	44.77 ± 2.80	45.23 ± 2.76	44.90 ± 3.01	45.23 ± 2.87	45.20 ± 3.04

Table 5: Means and standard deviations (percent) of MAPs over six events using different mechanisms to build histograms

	Gaussian	Laplacian	ISD	ID	Fused kernels
spherical 128	24.70 ± 1.41	43.04 ± 1.61	26.92 ± 1.00	43.64 ± 0.96	32.91 ± 2.20
spherical 64	23.99 ± 1.40	42.35 ± 1.64	25.62 ± 1.11	43.42 ± 1.18	29.01 ± 1.10
full 128	25.69 ± 7.57	21.39 ± 7.32	26.49 ± 8.38	21.93 ± 7.75	21.79 ± 7.29
full 64	25.23 ± 0.94	29.69 ± 1.81	25.68 ± 1.34	30.74 ± 1.67	26.74 ± 1.63

Table 6: Means and standard deviations (percent) of MAPs over six events using different GMMs

	Recognition accuracy
Kodak $\rightarrow$ Kodak	38.5 ± 12.7
Youtube $\rightarrow$ Kodak	30.0 ± 6.9
Baseline	41.6 ± 11.5

Table 7: Means and standard deviations (percent) of recognition accuracies using concept attributes

Training videos	Testing videos	Original videos	Compressed videos
60	846	38.9 ± 2.9	38.6 ± 2.8
120	786	45.7 ± 2.2	44.5 ± 1.6
180	726	49.5 ± 1.8	48.3 ± 1.9
240	666	52.0 ± 2.1	50.6 ± 2.1

Table 8: Recognition accuracies (percent) using distances calculated on original videos and distances calculated on compressed videos.

References