NICTA at TRECVID 2005 Shot Boundary Detection Task

Zhenghua (Jack) Yu, S.V.N Vishwanathan and Alex Smola*

National ICT Australia (NICTA)
{Jack.Yu, SVN.Vishwanathan, Alex.Smola}@nicta.com.au

November 4, 2005

National ICT Australia (NICTA) participated in TRECVID Shot Boundary Detection Task for the first time in 2005. NICTA's system is based on a generic framework that integrates feature extraction from video analysis and statistical machine learning. We are pleased to see that the NICTA system performed very well in TRECVID 2005.

We submitted 10 test runs with results included in Table 1. Major differences between the test runs include turning on/off some modules when making the shot boundary detection decision, and changes to some parameters.

Table 1: NICTA TRECVID 2005 Shot Boundary Detection Results

Sys ID	Plot Char	All		Cut		Grad		Frame	
		Recall	Prec	Recall	Prec	Recall	Prec	Recall	Prec
NICTA04	n	0.878	0.877	0.943	0.901	0.690	0.794	0.766	0.915
NICTA04s	n	0.879	0.878	0.943	0.901	0.690	0.794	0.766	0.915
NICTA06	n	0.881	0.888	0.941	0.915	0.706	0.798	0.774	0.901
NICTA09	n	0.889	0.880	0.928	0.921	0.773	0.761	0.757	0.896
NICTA11	n	0.888	0.881	0.928	0.920	0.774	0.769	0.770	0.895
NICTA132	n	0.901	0.878	0.946	0.914	0.771	0.769	0.771	0.895
NICTA132s	n	0.902	0.878	0.946	0.914	0.771	0.769	0.771	0.895
NICTA13	n	0.904	0.877	0.947	0.914	0.776	0.765	0.769	0.907
NICTA15	n	0.903	0.881	0.947	0.914	0.773	0.781	0.772	0.909
NICTA16	n	0.899	0.880	0.946	0.913	0.760	0.780	0.770	0.902

We also measured the execution time on a PC with a Pentium-4 3GHz CPU. The execution time difference between all submitted runs are almost negligible as the most time-consuming parts remain unchanged in all test runs. Typically it took around 4800 seconds to decode all the 11 TRECVID sequences and around 52000 seconds to segment all the sequences. However we should point out that during the development of our system, the main focus was on the detection accuracy and no effort was devoted to optimize the code for execution speed.

Given this is the first time that we participated in TRECVID, we gained invaluable experience through the exercise. There are several comments that we would like to make:

- TRECVID has provided an invaluable platform for different groups to develop and evaluate their algorithms. We thank TRECVID for the effort.
- The importance of the shot boundary detection task is twofolded: on one hand shot boundary detection is a basic step to analyze edited videos such as movies and news videos. On the other hand, we also believe this task poses an interesting scientific research problem as it can be generalized as a problem of time series segmentation, whilst the existence of various editing techniques calls for effective ways to incorporate prior knowledge in the system. Our initial success at TRECVID reinforced our view and we believe more advances can be made by researchers in this field.

- We have a suggestion to TRECVID Shot Boundary performance evaluation metrics: currently there is an artificial threshold of 5 frames to determine whether a transition is cut or gradual transition. Cuts are only matched with cuts and gradual transitions are only matched with gradual transitions. This means if a transition is correctly detected, but with a wrong length, then it may be double penalized: one false insertion and one false deletion. It is worse than not to detect the transition at all! We suggest to add another method to calculate the "ALL" performance, where matching across cut/grad categories is allowed. This does not require changing any of the current evaluation metrics.