Introduction
Free-hand sketch provides a natural and expressive modality for interaction with computers. This project explores methods to intuitively search video databases using sketches. Although video search is typically performed using keywords that specify content, text is cumbersome for describing scene appearance. Rather, a sketched depiction of a scene represents an orthogonal channel to constrain search. Although sketch based image retrieval (SBIR) has received much attention, the related problem of video retrieval (SVIR) is only sparsely researched – especially the fusion of text and sketch.

Building upon prior SBIR [1] and SVIR [2,3] we describe intermediate results from such a hybrid system. Our sketches describe objects based on their semantics (e.g., horse), a sketch of their motion trajectory, and their colour; collectively representing a natural interface for conveying multiple facets of events.

Semantic Sketch Interface
We provide an intuitive interface for sketch retrieval, allowing users to select the class of object they are searching for a colour palette for retrieval.

The system allows the user to draw a free hand arrow of the trajectory, that is based on a survey conducted in research from [4]. Results are displayed in a standard video search layout.

Video Search
To retrieve relevant videos we use the descriptors of the videos generated previously. We create a descriptor from the query by projecting the query object depiction along the trajectory line. This space is then normalised according to the region of the canvas space used within the query. We then follow the steps for Spatial-Temporal descriptor creation and the colour descriptor expansion.

When ranking the dataset we use a modified version of histogram intersection. This approach is modified to work over the different classes in the video dataset.

At current all items are searched linearly in the dataset, using histogram intersection, within the small dataset evaluated on this style of search is acceptable in larger datasets, clustering approaches would be used to improve retrieval performance.

Evaluation
We demonstrate the results in two ways, first using trajectory results alone then with the combination of trajectory and colour. This was evaluated over a subset of the TSF dataset, this dataset has been used to evaluate two sketch based retrieval systems in recent years. This dataset is composed of horse and snowboarding footage, and is based on the original VideoQ dataset[6]. At current there is no dataset available for semantic sketch based video retrieval, therefore we marked up the dataset for these intermediate results.

Annotated Class + Trajectory

The queries demonstrated above for trajectory and annotated class, these results have a high precision over the first two results, this result is demonstrated in the precision recall graph.

Annotated Class + Trajectory + Colour

We also evaluate our system over the proposed colour expansion these results demonstrate and improved ability of colour, within the horse class, within the person class there was not enough variation in the dataset to successfully differentiate between people.

Multiple Objects Classes
We also demonstrate the ability of our system to be able to identify multiple classes within the top ten results over both class, colour and trajectory we achieve three exact matches.

References