Project Summary

The rapid adoption and wide availability of on-line instructional video creates great opportunities for efficient access to a broad range of content. However, to reap the benefits, the inner semantic content within the videos must be exposed to browsing and searching systems. To address this need, this project will develop a novel approach to extract, organize, and make accessible the fine-grained semantic content of instructional video. We make the key observation that the common use of instructional slides naturally breaks video into appropriate, digestible, chunks of information. We envision being able to efficiently navigate the semantic space defined by these units of information — *paragraphs of video* — to find vibrant expositions on practically any topic of instruction and to identify videos with similar semantic content. From there, the student will be able view more of the containing video, or jump to alternative video paragraphs that are nearby in the semantic space. This will allow students to view explanations in different styles (by a different instructor), or different contexts (e.g., homogeneous coordinates in math versus homogeneous coordinates in computer graphics). To achieve this vision we will develop scalable methods to divide lectures into sections by presentation slide use, extract multi-modal information from each section (enabling basic search), and build statistical models for the semantics of the sections thereby representing the *semantic space* to enable navigation and improve retrieval of segments conceptually close to queries.

To demonstrate the potential of the approach we will deploy a unified media web-portal for searching and browsing video semantic chunks annotated using semantic web multimedia standards. We will seed the demonstration web portal with hundreds of videos from dozens of classes and talks. Subsequent users will be able to easily submit additional videos. The system itself will be built from modular components whose source code will be disseminated to the research and educational communities.

We will evaluate the hypothesis that this is a particularly effective way to provide access to instructional video in the context of distance education delivery systems. In particular, we will study whether exposing the inner semantics of the videos is helpful and engaging for students in four on-line computer science courses compared with basic video access as is typically provided by on-line course management systems such as Desire2Learn (D2L).

Intellectual Merit. Extracting semantics from video is an important research problem that has attracted much prior effort. The proposed work will make significant progress by focusing on meaningful sections defined by slide use, and extracting semantics at this level of granularity by exploiting the rich multi-modal data within the sections. Specific technical contributions that will result include methods for: 1) robust matching of video frames to slide images; 2) partitioning videos by slide use even when slides themselves are not independently available; 3) extracting semantic information from sources such as slide words, slide structure (e.g., titles, bullets), audio signal, and laser pointer use; 4) statistical modeling of the semantics of the segments; 5) identifying videos that use slides to support gathering such data from the web. Further this project will study how exposing fine-grained video semantics benefits distance learners and students supplementing class attendance with on-line study.

Broader Impacts. The methods developed here will set the stage for building a large repository of video organized by the semantic content of natural sized segments. We are now accustomed to finding information on any topic deep within any on-line text document, but the same does not apply to instructional video. Hence this project has great potential for transformative impact by providing this capability to video searching and browsing systems when slides are used. This will enable significant leveraging of this rich information source to the benefit of distance learners and anyone wanting to find informative video on any topic.

The project will rely heavily on both graduate and undergraduate students and an important project goal is to provide excellent educational opportunities for them. Recruitment of students will include under-represented groups in the computational sciences. The project will provide unique opportunities for students with a broad range of interests (computer vision, statistical analysis of multimedia data, human computer interface, digital document and semantic web technologies, educational technology and systems), and there will be significant value in the interactions across disciplines.