

Multimodal Data Fusion

AERFAI Summer School on
Pattern Recognition and Learning in Multimedia Systems
June 7-11, 2010, Benicàsim, Spain

April 20, 2010

Dr. Stéphane Marchand-Maillet
Viper group – Computer Science Department – University of Geneva (Switzerland)
<http://viper.unige.ch>

Multimedia data management has significantly evolved, in parallel with the changing usage of the WWW as a multimedia consumption platform. The ease of creation, storage, and consumption has however pinpointed and highlighted issues on the quality of access of multimedia data. Whereas generating content is a low-cost operation and is thus widely generalized, providing accessible handles to that content via annotation, tagging or other description operations is a costly process and therefore generally skipped, ignored or, at best minimized. The effect is scaled by the growing number of users, as content generators or content consumers.

As a result, huge masses of poorly accessible multimedia data lie within the Web, with the risk of mapping a large portion of it onto a large distributed storage of hardly accessible data. Automated recognition of the semantics underlying the content is thus a necessary and crucial operation.

The lecture will present multimodal data fusion strategies under the umbrella of semantic inclusion in multimedia management applications. Each of these topics will be differentiated by the level of implication of the user or multimedia content owner.

After reviewing the general challenge of multimodal fusion as a theoretical problem, we will first review solutions related to handling properly the interpretation of content feature. A second part of the lecture will be dedicated to the construction of practical setups ranging from navigation-based content discovery to learning-based content retrieval. Associated with each of these setups, we will detail how the implication of the user or group of users may be exploited as a rich source of semantic information to bootstrap or conduct the discovery of semantics within the multimedia content. Case studies will be presented to illustrate these models.