A Contribution to
a Multimedia Information Extraction Roadmap

Oliviero Stock

FBK-irst
Trento, Italy
stock@fbk.eu

Critical Technical Challenges

Rather than providing a long list I would like to set our attention to a few specific issues among the many challenges in this area:

1) what type of MM info to extract in relation to its possible use; and how to present it given the available presentation resources (e.g. devices, communication channels, modalities) and the likely attention of the user.

How to attract the user’s attention and interest?

2) as kind of a variation on the above theme I propose in particular MM IE for the benefit of a group of users. There are different cases: for example, a) same info to be presented to a group; b) info segmented so that different people can work on the material collaboratively at the same time; c) info to be intelligently fragmented for individuals in a group, for specific strategic goals, e.g. learning, provoking discussion etc.; d) common info presented on a large common display, complemented by personal info on a personal device.

3) I would also like to emphasize that the affective dimension should have full citizenship in MM IE: for instance sentiment analysis of video, or automatically understanding the bias in a reportage are important challenges for applications based on MM IE.

Remaining Gaps

Affective communication and persuasive communication have only recently become serious object of study. But MM material mostly was produced by humans with a purpose. And when produced automatically, as is atypical now but will be common in the future, it will still mostly be produced with a communication goal. And then in addition to the research issues concerning preexisting material, extracting the right information is part of an overall process that has to take into account the context and the expected use.

I believe that all this is vital for the research field. For building specific applications, these general concepts become more tractable and the road may be shorter.

In general I would say reasonable sentiment and bias analysis for extraction from existing MM material may be possible in 5 years, if adequate research efforts are committed. As for extraction for “intelligent” presentation, there are too many aspects and possible settings for proposing a general timeline

Existing Approaches

Methods and Techniques

For the challenges mentioned above, in addition to the traditional learning techniques for various aspects of the extraction process, and the adaptive combination for the MM situation, I think some deeper understanding of communication is needed. So I believe that understanding rhetorical structure, understanding of cinematography, music and so on will be important if we really want to progress. I believe for instance that MM IE for persuading, requires to build good models of persuasion, and fine tune the models through automatic learning.