

Chapter XXII

Improving Multimedia Digital Libraries Usability Applying NLP Sentence Similarity to Multimodal Sentences

Stefano Paolozzi

Consiglio Nazionale delle Ricerche, Italy

Fernando Ferri

Consiglio Nazionale delle Ricerche, Italy

Patrizia Grifoni

Consiglio Nazionale delle Ricerche, Italy

ABSTRACT

This chapter describes multimodality as a means of augmenting information retrieval activities in multimedia digital libraries. Multimodal interaction systems combine visual information with voice, gestures, and other modalities to provide flexible and powerful dialogue approaches. The use of integrated multiple input modes enables users to benefit from the natural approach used in human communication, improving usability of the systems. However, natural interaction approaches may introduce interpretation problems as the systems' usability is directly proportional to users' satisfaction. To improve multimedia digital library usability users can express their queries by means of a multimodal sentence. The authors propose a new approach to match a multimodal sentence with a template stored in a knowledge base to interpret the multimodal sentence and define the multimodal templates similarity.

INTRODUCTION

Multimedia digital libraries are characterized by an integrated management of typologically heterogeneous documents. The management is performed by the use of specific systems to index, search, and automatically extract data that are typical of complex contents of multimedia documents (e.g., video, audio, etc.).

Several researches and studies, over the last few years, address multimedia data analysis, computer vision and object recognition, and multimedia management in large audio and video collections. As the most complex and advanced multimedia information systems, digital libraries are emerging at an increasingly fast rate throughout the world. One of the primary difficulties in building a digital library is to support effective and efficient retrieval of the media objects, including text, image, video, and audio (Baeza-Yates & Ribeiro-Neto, 1998).

Because the information retrieval quality is widely influenced by the user's interaction with the system, it is necessary to modify the multimedia database access approach in order to take into consideration the user's actions. For this reason, the approach has to take into account the complex necessities to define queries not only with terminological data, but also with multimedia data.

Therefore, it is important to change the traditional textual interfaces, which can be used only in keywords or textual metadata searches, with more rich interfaces. These ones must allow the formulation of multidimensional queries, not only by terms but also by images, sounds, and so forth (Flickner, Sawhney, Niblack, & Ashley, 1995). In this way the retrieval can be performed according to indexes composed by text extracted from speech, main images of a sequence, simple images, shapes, colors, sounds, and so forth. Obviously we continue to maintain textual data that are related to particular aspects not specifically related with audiovisual data stored in the multimedia library.

A variety of technologies such as sketch and speech processing, video streaming, multimedia databases, and graphical user interfaces can be integrated to produce a multimodal environment.

The systems' usability is directly proportional to users' satisfaction. To improve multimedia digital library usability the users can express their queries by means of a multimodal sentence. Moreover, the system's ability to give exact or approximate answers can improve users' satisfaction.

In this chapter we analyze different approaches found in literature and propose an original one used to calculate sentence similarity. The proposed approach is based on the evaluation of users' input from the multimodal sentence analysis comparing the given sentence with a knowledge base of short sentences in order to find similarity between them. A multimodal sentence is formed by atomic elements (e.g., glyphs/graphemes, phonemes, etc.).

Because speech represents the most used modality and the prevalent one—prevalent is the modality that permits to express the multimodal sentence's syntax and the semantic while the use of other modalities is devoted to “support” the speech and eventually to resolve ambiguities—multimodal sentences can be described using natural language. In this way we can consider the similarity computation between multimodal sentences as the similarity computation between natural language sentences.

BACKGROUND: A BRIEF STATE-OF-THE-ART OF SENTENCE SIMILARITY ISSUES

There is a wide literature base on measuring the similarity between documents or long texts (e.g., Allen, 1995; Meadow, Boyce, & Kraft, 2000), but there are only few works relating to the measurement of similarity between very short texts (Foltz, Kintsch, & Landauer, 1998) or sentences (Li, McLean, Bandar, O'Shea, & Crockett, 2006).

6 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/improving-multimedia-digital-libraries-usability/19885

Related Content

First Person Singular: A Digital Library Collection that Helps Second Language Learners Express Themselves

Shaoqun Wu and Ian H. Witten (2012). *Multimedia Storage and Retrieval Innovations for Digital Library Systems* (pp. 22-40).

www.irma-international.org/chapter/first-person-singular/64460

An Evaluative Study on the Information Security With Special Reference to University Libraries of West Bengal

Abhijit Chatterjee and Anwesha Chatterjee (2025). *Exploring Digital Metrics in Academic Libraries* (pp. 423-442).

www.irma-international.org/chapter/an-evaluative-study-on-the-information-security-with-special-reference-to-university-libraries-of-west-bengal/367960

Analysis the Typhoon Eyes of Megi from MTSAT Satellite Cloud Images with 3-D Profile Reconstruction

Yueh-Jyun Lee, Ji-Chyun Liu, Shih-Wen Liu, Yuh-Fong Lin and Lung-Chun Chang (2011). *International Journal of Digital Library Systems* (pp. 39-48).

www.irma-international.org/article/analysis-typhoon-eyes-megi-mtsatsat/59887

Building Digital Collections Using Open Source Digital Repository Software: A Comparative Study

George Pyrounakis, Mara Nikolaidou and Michael Hatzopoulos (2014). *International Journal of Digital Library Systems* (pp. 10-24).

www.irma-international.org/article/building-digital-collections-using-open-source-digital-repository-software/105108

Cost Issues in Digitalization and Preservation: A Case Study of a Small-Scale Music Library

Ada Chi Wai Chung (2020). *Digital Libraries and Institutional Repositories: Breakthroughs in Research and Practice* (pp. 45-58).

www.irma-international.org/chapter/cost-issues-in-digitalization-and-preservation/250659