

# Chapter IV

## Cross-Modal Correlation Mining Using Graph Algorithms

**Jia-Yu Pan**

*Carnegie Mellon University, USA*

**Hyung-Jeong Yang**

*Chonnam National University, South Korea*

**Christos Faloutsos**

*Carnegie Mellon University, USA*

**Pinar Duygulu**

*Bilkent University, Turkey*

### ABSTRACT

*Multimedia objects like video clips or captioned images contain data of various modalities such as image, audio, and transcript text. Correlations across different modalities provide information about the multimedia content, and are useful in applications ranging from summarization to semantic captioning. We propose a graph-based method, MAGIC, which represents multimedia data as a graph and can find cross-modal correlations using “random walks with restarts.” MAGIC has several desirable properties: (a) it is general and domain-independent; (b) it can detect correlations across any two modalities; (c) it is insensitive to parameter settings; (d) it scales up well for large datasets; (e) it enables novel multimedia applications (e.g., group captioning); and (f) it creates opportunity for applying graph algorithms to multimedia problems. When applied to automatic image captioning, MAGIC finds correlations between text and image and achieves a relative improvement of 58% in captioning accuracy as compared to recent machine learning techniques.*

### Introduction

Advances in digital technologies make possible the generation and storage of large amount of

multimedia objects such as images and video clips. Multimedia content contains rich information in various modalities such as images, audios, video frames, time series, and so forth. However, making

rich multimedia content accessible and useful is not easy. Advanced tools that find characteristic patterns and correlations among multimedia content are required for the effective usage of multimedia databases.

We call a data object whose content is presented in more than one modality a *mixed media* object. For example, a video clip is a mixed media object with image frames, audios, and other information such as transcript text. Another example is a captioned image such as a news picture with an associated description, or a personal photograph annotated with a few keywords (Figure 1). In this chapter, we would use the terms *medium* (plural form *media*) and *modality* interchangeably.

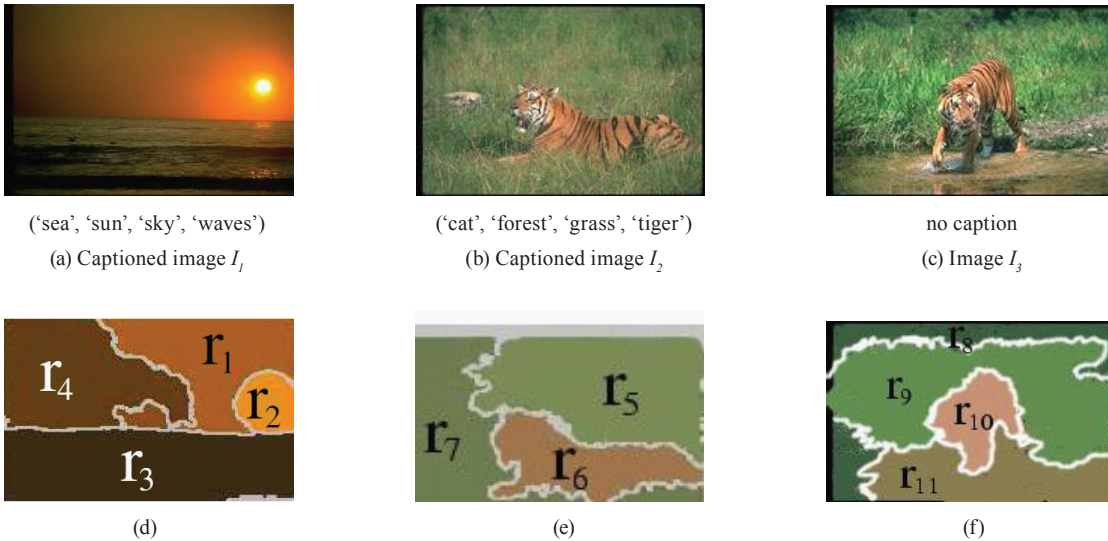
It is common to see correlations among attributes of different modalities on a mixed media object. For instance, a news clip usually contains human speech accompanied with images of static scenes, while a commercial has more dynamic scenes with loud background music (Pan & Faloutsos, 2002). In image archives, caption keywords are chosen such that they describe objects in the

images. Similarly, in digital video libraries and entertainment industry, motion picture directors edit sound effects to match the scenes in video frames.

Cross-modal correlations provide helpful hints on exploiting information from different modalities for tasks such as segmentation (Hsu et al., 2004) and indexing (Chang, Manmatha, & Chua, 2005). Also, establishing associations between low-level features and attributes that have semantic meanings may shed light on multimedia understanding. For example, in a collection of captioned images, discovering the correlations between images and caption words could be useful for image annotation, content-based image retrieval, and multimedia understanding.

The question that we are interested in is “*Given a collection of mixed media objects, how do we find the correlations across data of various modalities?*” A desirable solution should be able to include all kinds of data modalities, overcome noise in the data, and detect correlations between any subset of modalities available. Moreover, in

Figure 1. Three sample images: (a),(b) are captioned with terms describing the content; (c) is an image to be captioned. (d)(e)(f) show the regions of images (a)(b)(c), respectively.



23 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/cross-modal-correlation-mining-using/24901](http://www.igi-global.com/chapter/cross-modal-correlation-mining-using/24901)

## Related Content

---

### OntoWiktionary: Constructing an Ontology from the Collaborative Online Dictionary Wiktionary

Christian M. Meyer and Iryna Gurevych (2012). *Semi-Automatic Ontology Development: Processes and Resources* (pp. 131-161).

[www.irma-international.org/chapter/ontowiktionary-constructing-ontology-collaborative-online/63901](http://www.irma-international.org/chapter/ontowiktionary-constructing-ontology-collaborative-online/63901)

### Distributed "Knowing in Practice" in Workshare Contexts: A Case Study of a Knowledge Management System

Lakshmi Goel (2020). *Current Issues and Trends in Knowledge Management, Discovery, and Transfer* (pp. 222-246).

[www.irma-international.org/chapter/distributed-knowing-in-practice-in-workshare-contexts/244886](http://www.irma-international.org/chapter/distributed-knowing-in-practice-in-workshare-contexts/244886)

### Spatial Navigation Assistance System for Large Virtual Environments: The Data Mining Approach

Mehmed Kantardzic, Pedram Sadeghian and Walaa M. Sheta (2008). *Mathematical Methods for Knowledge Discovery and Data Mining* (pp. 265-283).

[www.irma-international.org/chapter/spatial-navigation-assistance-system-large/26145](http://www.irma-international.org/chapter/spatial-navigation-assistance-system-large/26145)

### Knowledge Discovery in Biomedical Data Facilitated by Domain Ontologies

Amandeep S. Sidhu, Paul J. Kennedy and Simeon Simoff (2007). *Knowledge Discovery and Data Mining: Challenges and Realities* (pp. 189-201).

[www.irma-international.org/chapter/knowledge-discovery-biomedical-data-facilitated/24907](http://www.irma-international.org/chapter/knowledge-discovery-biomedical-data-facilitated/24907)

### Towards Understanding and Implementing Knowledge Management Strategy

Murray Eugene Jennex (2020). *Current Issues and Trends in Knowledge Management, Discovery, and Transfer* (pp. 103-125).

[www.irma-international.org/chapter/towards-understanding-and-implementing-knowledge-management-strategy/244879](http://www.irma-international.org/chapter/towards-understanding-and-implementing-knowledge-management-strategy/244879)