Computing Similarity between Cultural Heritage Items using Multimodal Features

Nikolaos Aletras and Mark Stevenson
Department of Computer Science, University of Sheffield
{ n.aletras, m.stevenson } @ dcs.shef.ac.uk

In recent years a vast amount of Cultural Heritage (CH) artefacts have been digitised and made available on-line. For example, the Louvre and the British Museum provide information about exhibits on their web pages\(^1\). In addition, information is also available via sites that aggregate CH information from multiple resources. A typical example is Europeana\(^2\), a web-portal to collections from several European institutions that provides access to over 20 million items including paintings, films, books, archives and museum exhibits.

However, online information about CH artefacts is often unstructured and varies by collection. This makes it difficult to identify information of interest in sites that aggregate information from multiple sources, such as Europeana, or to compare information across multiple collections (such as the Louvre and British Museum). These problems form a significant barrier to accessing the information available in these online collections. A first step towards improving access would be to identify similar items in collections. This could assist with several applications that are of interest to those working in CH including recommendation of interesting items (Wang et al., 2008), generation of virtual tours (Wang et al., 2009), visualisation of collections (Hornbaek and Hertzum, 2011) and exploratory search (Amin et al., 2008).

Information in digital CH collections often includes multiple types of media such as text, images and audio. It seems likely that information from all of these types would help humans to identify similar items and that it could help to identify them automatically.

In this work we make use of information from two media (text and images) to compute the similarity between items in a large collection of CH items. A data set was created by selecting 295 pairs of items from Europeana. Each item corresponds to a metadata record consisting of textual information together with a URI and a link to its thumbnail. Figure 1 shows an item taken from the Europeana website. These are extracted and concatenated to form the textual information associated with each item. In addition, the accompanying thumbnail image (or “preview”) was also extracted to be used as the visual information. The size of these images varies from 7,000 to 10,000 pixels.

A range of similarity measures for text and images are compared and combined. Text similarity is measured by applying a range of corpus-based methods, i.e. tf.idf and LDA (Blei et al., 2003). Image similarity is computed by making use of colour histograms and the Haar wavelet decomposition (Beylkin et al., 1991) to create signatures of images that contain colour and basic shape information. Evaluation is carried out using a set of items from Europeana with similarity judgements that were obtained in a crowdsourcing experiment. We find that combining information from both media produces better results than when either is used alone.

\(^1\)http://www.louvre.fr/, http://www.britishmuseum.org/
\(^2\)http://www.europeana.eu
The main contribution of this work is to demonstrate the usefulness of applying information from more than one medium when comparing CH items. In addition, it explores the effectiveness of different similarity measures when applied to this domain and introduces a data set of similarity judgements that can be used as a benchmark.

Acknowledgments The research leading to these results was carried out as part of the PATHS project (http://paths-project.eu) funded by the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 270082.

References

Alia Amin, Jacco van Ossenbruggen, Lynda Hardman, and Annelies van Nispen. Understanding Cultural Heritage Experts’ Information Seeking Needs. In Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries, pages 39–47, Pittsburgh, PA., 2008.

Gregory Beylkin, Ronald Coifman, and Vladimir Rokhlin. Fast Wavelet Transforms and Numerical Algorithms I. Communications on Pure and Applied Mathematics, 44:141–183, 1991.

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent Dirichlet Allocation. Journal of Machine Learning Research, 3:993–1022, 2003. ISSN 1532-4435.

Kasper Hornbaek and Morten Hertzum. The notion of overview in information visualization. International Journal of Human-Computer Studies, 69:509–525, 2011.

Yiwen Wang, Natalia Stash, Lora Aroyo, Peter Gorgels, Lloyd Rutledge, and Guus Schreiber. Recommendations based on semantically-enriched museum collections. Journal of Web Semantics: Science, Services and Agents on the World Wide Web, 6(4):43–50, 2008.

Yiwen Wang, Lora Aroyo, Natalia Stash, Rody Sambeek, Schuurmans Yuri, Guus Schreiber, and Peter Gorgels. Cultivating personalized museum tours online and on-site. Journal of Interdisciplinary Science Reviews, 34(2):141–156, 2009.