

## Accepted Manuscript

Graph-regularized concept factorization for multi-view document clustering

Kun Zhan, Jinhui Shi, Jing Wang, Feng Tian

PII: S1047-3203(17)30062-7  
DOI: <http://dx.doi.org/10.1016/j.jvcir.2017.02.019>  
Reference: YJVICI 1970

To appear in: *J. Vis. Commun. Image R.*

Received Date: 1 September 2016  
Revised Date: 2 February 2017  
Accepted Date: 24 February 2017



Please cite this article as: K. Zhan, J. Shi, J. Wang, F. Tian, Graph-regularized concept factorization for multi-view document clustering, *J. Vis. Commun. Image R.* (2017), doi: <http://dx.doi.org/10.1016/j.jvcir.2017.02.019>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Graph-regularized concept factorization for multi-view document clustering

Kun Zhan<sup>a</sup>, Jinhui Shi<sup>a</sup>, Jing Wang<sup>b</sup>, Feng Tian<sup>b</sup>

<sup>a</sup>*School of Information Science and Engineering, Lanzhou University, Lanzhou, China*

<sup>b</sup>*Faculty of Science and Technology, Bournemouth University, UK*

---

## Abstract

We propose a novel multi-view document clustering method with the graph-regularized concept factorization (MVCF). MVCF makes full use of multi-view features for more comprehensive understanding of the data and learns weights for each view adaptively. It also preserves the local geometrical structure of the manifolds for multi-view clustering. We have derived an efficient optimization algorithm to solve the objective function of MVCF and proven its convergence by utilizing the auxiliary function method. Experiments carried out on three benchmark datasets have demonstrated the effectiveness of MVCF in comparison to several state-of-the-art approaches in terms of accuracy, normalized mutual information and purity.

*Keywords:* Multi-view learning, concept factorization, document clustering, manifold learning

---

## 1. Introduction

The matrix factorization-based approaches have become popular in document clustering [1, 2]. Nonnegative matrix factorization (NMF) [3] and concept factorization (CF) [1] have produced impressive results. Generally, CF mainly strives to overcome the limitations of NMF while inheriting all its strengths. CF models each concept as a linear combination of the data points, and each data point as a linear combination of the concepts. It aims to interpret the product of the two sets of linear coefficients as an approximation of the original data points. The cluster label of each data point can be easily derived from the obtained linear coefficients. However, CF does not consider the local manifold geometry but the global Euclidean geometry

Download English Version:

<https://daneshyari.com/en/article/4969263>

Download Persian Version:

<https://daneshyari.com/article/4969263>

[Daneshyari.com](https://daneshyari.com)