Multi-view network-based social-tagged landmark image clustering

So Yeon Kim, Kyung-Ah Sohn
Department of Computer Engineering, Ajou University, S. Korea
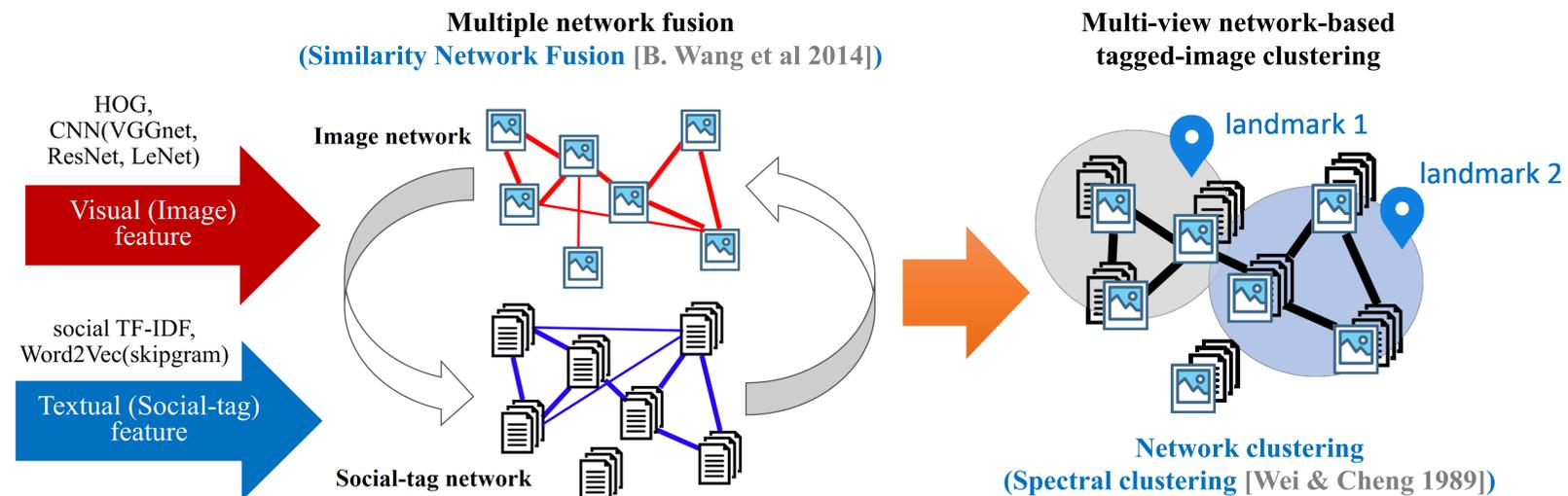e-mail : {jebi1771, kasohn}@ajou.ac.kr

Abstract

- Motivated by the fact that learning multi-modal social data is challenging due to data heterogeneity and noise in user-generated data
- Proposed a multi-view network-based clustering approach that is robust to noise and fully reflects the underlying structure of the comprehensive network
- Experimented with clustering challenging tagged images of landmarks
- Outperforms other previously reported multi-view clustering algorithms and better utilizes the advantages of the network for each view
- Do not need to know the exact network structure of each view and can effectively combine complementary information from different types of data
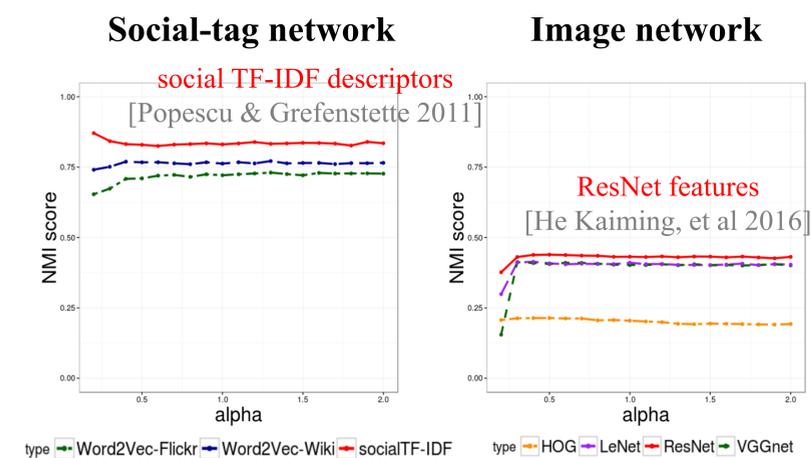- Can be applied to any number of data types

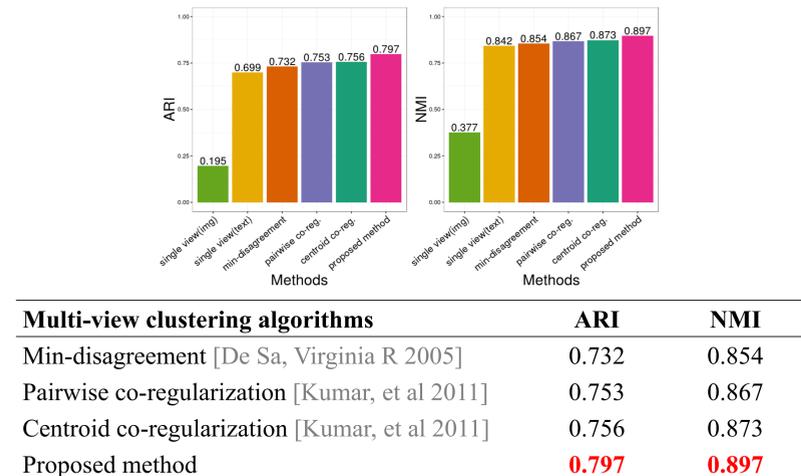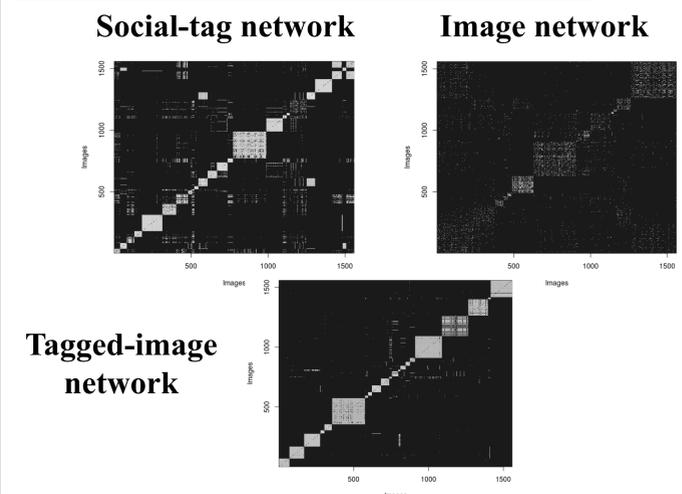## Methods

Flickr social-tagged landmark images

#roma #italy #vatican #church #cathedral

#italy #pisa #tower #campo #attraction

HOG, CNN(VGGnet, ResNet, LeNet)
Visual (Image) feature

social TF-IDF, Word2Vec(skipgram)
Textual (Social-tag) feature

Multiple network fusion
(Similarity Network Fusion [B. Wang et al 2014])

Image network

Social-tag network

Multi-view network-based tagged-image clustering

landmark 1
landmark 2

Network clustering
(Spectral clustering [Wei & Cheng 1989])

Results

Performance comparison of different descriptors

Social-tag network

social TF-IDF descriptors [Popescu & Grefenstette 2011]

type: Word2Vec-Flickr, Word2Vec-Wiki, socialTF-IDF

Image network

ResNet features [He Kaiming, et al 2016]

type: HOG, LeNet, ResNet, VGGnet

Performance comparison between different algorithms

ARI: single view(img) 0.195, single view(txt) 0.699, min-disagreement 0.732, pairwise co-reg 0.753, centroid co-reg 0.756, proposed method 0.797

NMI: single view(img) 0.377, single view(txt) 0.842, min-disagreement 0.854, pairwise co-reg 0.867, centroid co-reg 0.873, proposed method 0.897

Multi-view clustering algorithms	ARI	NMI
Min-disagreement [De Sa, Virginia R 2005]	0.732	0.854
Pairwise co-regularization [Kumar, et al 2011]	0.753	0.867
Centroid co-regularization [Kumar, et al 2011]	0.756	0.873
Proposed method	0.797	0.897

Multi-view network based cluster analysis

Social-tag network

Image network

Tagged-image network

The landmark(cluster) "Wurzburg Residence Germany"

Social-tag network clustering

Image network clustering

Multi-view network-based tagged-image clustering